

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 October 2008 (30.10.2008)

PCT

(10) International Publication Number
WO 2008/129256 A2

(51) International Patent Classification:
C12N 15/11 (2006.01) *A61K 31/713* (2006.01)

(21) International Application Number:
PCT/GB2008/001357

(22) International Filing Date: 17 April 2008 (17.04.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0707429.7 18 April 2007 (18.04.2007) GB

(71) Applicant (for all designated States except US): **ISIS INNOVATION LIMITED** [GB/GB]; Ewert House, Ewert Place, Summertown, Oxford OX2 7SG (GB).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **SIMMONS, Alison** [GB/GB]; The Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS (GB).

(74) Agent: **ATKINSON, Jennifer**; Barker Brettell LLP, 138 Hagley Road, Edgbaston, Birmingham B16 9PW (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

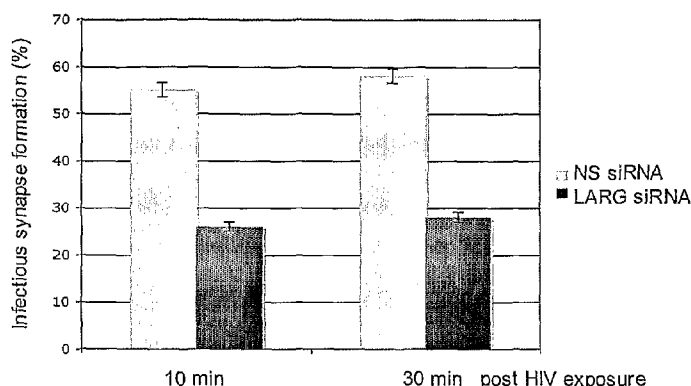
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(54) Title: HIV THERAPY

Figure 4B



(57) Abstract: A composition for inhibiting the expression and/or activity of the leukaemia-associated Rho guanine nucleotide exchange factor (LARG) and/or the Rho protein in a cell and its use in the manufacture of a medicament for the treatment of HIV infection, and a method of treating HIV infection by administering a composition capable of inhibiting the expression and/or activity of LARG and/or Rho protein to a cell or organism.

HIV THERAPY

The present invention relates to compositions for inhibiting the activity or expression of proteins activated by the receptor DC-SIGN, and in particular to the use of such compositions for the prophylactic or therapeutic treatment of HIV infection.

Some 40 million people are infected with HIV worldwide, and new infections occur at the rate of 5 million per year (UNAIDS Update on the Worldwide AIDS Epidemic, December 2001). Immense amounts of time, effort, and money have been invested in pursuit of effective treatments, whether prophylactic or therapeutic, for HIV infection. In the United States, the Food and Drug Administration has approved three different classes of compounds for use in HIV therapy: namely, nucleoside analogs, non-nucleoside reverse transcriptase inhibitors, and protease inhibitors. For some patients combinations of these compounds have proved remarkably effective at reducing viral load. The therapeutic regimens are often very complex, requiring precisely orchestrated administration of multiple pills throughout the day and not tolerating even minor variation in administration. Unfortunately, many patients respond poorly to treatment even when they follow their prescribed therapeutic regimen precisely, and there remains a need for the development of alternative therapies for the treatment and prevention of HIV infection and AIDS.

Dendritic cells (DCs) are important in HIV infection. Dendritic cells (DCs) are pivotal antigen presenting cells (APC) that capture microorganisms in peripheral tissues and migrate to secondary lymphoid organs to initiate naïve T cell priming and adaptive immune responses. DCs express a variety of pattern recognition receptors (PRRs) that recognize conserved molecular motifs on microorganisms. Intracellular

signalling pathways activated by PRRs have been shown to dictate the maturation profile of the DC and thus the nature of the subsequent immune response (Iwasaki A & Medzhitov R *Nat Immunol* 5, 987-95 (2004)). Members of the toll-like receptor (TLR) family are PRRs specializing in sensing pathogens and much information has accumulated about the intracellular signalling pathways employed by TLRs (Oda K & Kitano H *Mol Syst Biol* 2, 2006.0015. (2006)). Other PRRs on DCs include the C-type lectins that mediate self-antigen binding and uptake, in addition to pathogen recognition (Cambi A & Figdor CG *Curr Opin Immunol.* 17, 345-51 (2005); McGreal EP et al *Curr Opin Immunol.* 17, 18-24 (2005)). In contrast to the TLRs, the nature of signalling through C-type lectins is very poorly understood.

DC-specific intracellular adhesion molecule (ICAM) grabbing non-integrin (DC-SIGN) is a C-type lectin expressed by DCs that functions both as an adhesion receptor and as a PRR. As an adhesion receptor, DC-SIGN binds ICAM-3 aiding DC-T cell interaction in the initiation of immune responses (Geijtenbeek TB et al *Cell.* 100, 575-85 (2000)). As a PRR, DC-SIGN binds a variety of viral, bacterial, fungal and parasite pathogens including HIV-1 gp120 (van Kooyk Y & Geijtenbeek TB *Nat Rev Immunol.* 3, 697-709 (2003); Geijtenbeek TB et al *Cell.* 100, 587-97 (2000)). Most DC-SIGN interacting microbes elicit Th2-type responses that result in impaired pathogen clearance and establishment of chronic infection.

25

In the case of HIV-1, DC-SIGN would appear to be the principle surface receptor in monocyte-derived DCs (Turville SG et al *Nat Immunol.* 3, 975-83 (2002)). DC-SIGN captures HIV-1 at low titres through its high affinity interaction with HIV-1 envelope glycoprotein gp120 and internalizes the viral particles to a non-lysosomal compartment characterized by low pH and the presence of the tetraspanins CD81 and

30

CD9, but the absence of MHC II (Geijtenbeek TB et al *Cell*. 100, 587-97 (2000)). Some of the virus is retained in endocytic compartments within the cytoplasm and is either transmitted intact by recycling to permissive CD4+ T cells, allowing efficient infection of by-stander CD4+ T cells, or degraded by the proteasome (Geijtenbeek TB et al *Cell*. 100, 587-97 (2000); Turville SG et al *Blood*. 103, 2170-9 (2004); Moris A et al *Blood*. 103, 2648-54 (2004)). DC-SIGN is required for enhancing T cell infection and is crucial for this process when viral titres are low (Arrighi JF et al *J Virol*. 78, 10848-55 (2004)), at least in part due to its ability to enhance viral synapse formation (Arrighi JF et al *J Exp Med*. 200, 1279-88 (2004)). As with other C-type lectins, very little is understood of the molecular machinery employed to achieve these effects, and how it may be used to HIV's advantage is unknown.

The present invention relates to methods and compositions for preventing or treating HIV infection by disrupting the pathway activated following the recognition of HIV virus particles by DC-SIGN. Accordingly the invention relates to compositions and methods useful for inhibiting the expression and/or activity of the leukaemia-associated Rho guanine nucleotide exchange factor (LARG) and/or a Rho protein. Both LARG and Rho proteins are activated and phosphorylated when HIV-1 gp120 binds to DC-SIGN.

Reference herein to Rho, refers to any of the large family of Rho GTPase proteins, and in particular any Rho protein activated by DC-SIGN when it binds HIV-1 gp120. In particular the Rho protein may be RhoA, RhoA is known to be activated by LARG.

Inhibition of LARG and/or Rho expression and/or activity is intended to refer to the reduction or elimination of the activity of one or both of these proteins in a cell. This may be achieved at (i) the translational level, (ii)

the transcriptional level, and/or (iii) the post-translational level. Inhibition at the translation level may occur by destruction or inactivation of mRNA, or by preventing the translational machinery from accessing the mRNA. Inhibition at the transcriptional level may prevent the
5 production of mRNA or the production of useful mRNA. Inhibition at the post-translational level may prevent the functioning of any protein produced.

According to a first aspect the invention provides a composition for
10 inhibiting the expression and/or activity of the leukaemia-associated Rho guanine nucleotide exchange factor (LARG) and/or the Rho protein in a cell.

The composition may comprise a synthetic siRNA molecule, a synthetic
15 antisense oligonucleotide, a ribozyme or a DNzyme.

A synthetic siRNA molecule, a synthetic antisense oligonucleotide, a ribozyme or a DNzyme may function at the translational level, and inhibit or prevent the translation of mRNA into a functional protein.
20

A synthetic siRNA molecule may be capable of down regulating the expression of LARG or Rho via RNA interference. RNA interference (RNAi) is well known in the art and refers to the process of sequence-specific post-transcriptional gene silencing mediated by small interfering
25 RNAs (siRNA). RNAi works by using siRNAs complementary to the mRNA of a gene whose expression is to be reduced or eliminated. The siRNA binds to the complementary mRNA in the cell to produce double stranded RNA (dsRNA). The presence of dsRNA in cells triggers various defence mechanisms including the activity of ribonuclease III enzyme,
30 Dicer. Small RNA duplexes, usually 19 to 21 nucleotides in length, are generated from dsRNA by the activity of the Dicer enzyme. One strand

of the dsRNA duplex (the anti-sense) is then selectively incorporated into a dsRNA-induced silencing complex (RISC) that contains several proteins and guides the selection of a complementary RNA for cleavage. RNAi works at the translational level and prevents protein production or reduces
5 the amount of a protein produced.

RNAi may be used to silence or reduce gene expression by administration of an siRNA molecule as such, or by transcription of an siRNA from an expression vector and/or retrovirus capable of infecting mammalian cells.

10

Preferably a synthetic siRNA or antisense molecule for use in the composition of the invention has a sequence complementary to part of the LARG or Rho mRNA from a human or any other mammal.

15 Preferably the siRNA or antisense oligonucleotide molecule comprises a sequence of RNA complementary, or substantially complementary, to up to about 50 nucleotides of the LARG or Rho mRNA.

Preferably the siRNA or antisense oligonucleotide molecule comprises a
20 sequence of RNA complementary, or substantially complementary, to up to about 40 nucleotides of the LARG or Rho mRNA, more preferably to up to about 30 nucleotides of the LARG or Rho mRNA.

Preferably an siRNA for use in the invention comprises a 19 to 21 base
25 pair RNA duplex. Preferably all the nucleotides of one of the strands which form the RNA duplex are complementary to a part of the LARG or Rho mRNA. The RNA duplex preferably has a two-nucleotide overhang at one or both 3' ends. The nucleotides that form the overhang may not be complementary to the LARG or Rho mRNA. Preferably the siRNA
30 comprises between 42 and 46 nucleotides of which between 38 and 42 preferably form an RNA duplex of between 19 and 21 nucleotide pairs,

leaving at least a two nucleotide overhang at each 3' end. Preferably the siRNA is 42 nucleotides in length, 38 of the nucleotides preferably form a 19 base pair RNA duplex, and there is preferably a two-nucleotide overhang at each 3' end. Preferably at least one strand of the 19 base pair
5 RNA duplex is complementary to part of the LARG or Rho mRNA.

Preferably the complementarity between the siRNA and/or the antisense oligonucleotides and the LARG or Rho mRNA is sufficient to prevent translation of the LARG or Rho mRNA.

10

By substantially complementary to the mRNA of LARG or Rho it is intended that the siRNA or antisense oligonucleotide is sufficiently complementary to form a duplex with the mRNA and prevent translation of the mRNA.

15

The siRNA or antisense oligonucleotide may not be complementary along its entire length to the LARG or Rho mRNA. Preferably at least about 35% of the nucleotides in the siRNA or antisense oligonucleotide are complementary with the LARG or Rho mRNA, more preferably at least
20 about 40%, 50%, 60%, 70% 80%, 90% or more of the nucleotides in the siRNA or antisense oligonucleotides are complementary to the LARG or Rho mRNA.

Preferably the siRNA and/or the antisense oligonucleotides do not affect
25 the translation of any genes, other than those encoding LARG or Rho, in the cell to which they are administered. More specifically, preferably the siRNA and/or the antisense oligonucleotides are not complementary to mRNAs other than the LARG or Rho mRNA expressed in the cell to which they are administered.

30

The siRNA molecule may comprise two RNA strands having a region of complementarity of at least about 15 nucleotides, more preferably between about 19 and about 21 nucleotides, in length. The siRNA molecule may optionally further comprise one or two single-stranded overhangs or loops. Alternatively, the siRNA may comprise a single RNA strand having a region of self-complementarity. The single RNA strand may form a hairpin structure with a stem and loop and, optionally, one or more unpaired portions at the 5' and/or 3' portion of the RNA. Preferably the stem comprises an RNA duplex of between 19 and 21 nucleotide pairs.

Preferably the siRNA molecule comprises at least about 8, more preferably between about 8 and about 30, consecutive nucleotides of the LARG or Rho mRNA. Preferably the siRNA molecule comprises between 19 and 21 consecutive nucleotides of the LARG or Rho mRNA.

The double stranded portion of the siRNA molecule may be flanked by single stranded overhangs. Preferably the overhangs are at the 3' end.

siRNA which is a duplex of RNA and DNA may also be used, as may RNA/DNA chimeras. The siRNA may comprise an RNA/DNA chimera wherein the double stranded RNA has 3' overhangs of DNA bases (for example, dTdT), or one or more of the RNA bases in the double stranded section may be DNA bases, or indeed the whole of one strand may be DNA bases.

The siRNA may have the sequence: 5'-GAA ACT CGT CGC ATC TTC C- 3' (SEQ ID NO. 1), preferably the siRNA also has a two nucleotide overhang at the 3' end. Preferably the siRNA also comprises the strand complementary to 5'-GAA ACT CGT CGC ATC TTC C-3' (SEQ ID NO. 1), the two complementary strands preferably form an RNA duplex.

Preferably the complementary strand also has a two-nucleotide overhang at the 3' end. The two-nucleotide overhang at the 3' end is preferably TT.

- 5 The siRNA may have the sequence: 5'-AAG GAA GAG AAG GAT GTT AAT-3' (SEQ ID NO. 2), preferably the siRNA also has a two nucleotide overhang at the 3' end. Preferably the siRNA also comprises the strand complementary to 5'-AAG GAA GAG AAG GAT GTT AAT-3' (SEQ ID NO. 2), the two complementary strands preferably form an RNA duplex.
- 10 Preferably the complementary strand also has a two-nucleotide overhang at the 3' end. The two-nucleotide overhang at the 3' end is preferably TT.

Antisense oligonucleotides may be RNA or DNA or a chimera thereof.

- 15 Preferably antisense oligonucleotides according to the invention are at least about 15 nucleotides in length and hybridise to LARG or Rho mRNA.

- siRNA and/or the antisense oligonucleotides may incorporate non-natural bases, for example, C5 propyne analogs, and/or non-natural backbone linkages such as, phosphorothioates, morpholino oligonucleotides, MEA or DEED phosphoramidates or other modifications, for example 3' capping, to increase stability and enhance resistance to endonucleases.
- 20

- 25 siRNA or antisense oligonucleotides may be synthesised *in vitro* using chemical or enzymatic RNA synthesis techniques well known in the art. If the siRNA molecule is formed from two strands, the strands may be synthesised separately and then annealed.

- 30 Alternatively, siRNA or antisense oligonucleotides may be synthesised intracellularly using a suitable expression vector.

Alternatively to, or in addition to, comprising an siRNA and/or an antisense oligonucleotide, the composition of the invention may comprise an inhibitor of LARG or Rho protein activity. The inhibitor may compete
5 with the LARG and/or the Rho protein or it may interact with the LARG and/or the Rho protein directly to inhibit its activity. The inhibitor may be a small molecule. The small molecule may be a peptide or chemical compound. Suitable small molecules may be identified by screening for a compound that inhibits LARG-Rho interaction, or that inhibits the
10 enzymatic activity of LARG or Rho. Preferably any inhibitor for use in a composition of the invention will have low toxicity to a mammalian cell and will be capable of efficiently entering a mammalian cell. Brefeldin A is an example of an effective inhibitor of LARG and other GEF proteins, but it is highly toxic to cells, and so, without modification, is unlikely to
15 be suitable for use in a composition of the invention.

The composition of the invention may additionally, or alternatively, comprise a ribozyme which inhibits LARG and/or Rho expression and/or activity.

20

Preferably a composition of the invention is able to reduce *in vivo* the level of active LARG and/or Rho protein by at least about 10%, more preferably by at least about 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more, or it may be able to completely prevent LARG and/or Rho
25 protein activity in a cell.

Preferably, by inhibiting the expression and/or activity of the LARG and/or Rho protein in a cell, infection by HIV is reduced or prevented.

According to another aspect the invention provides a vector encoding an siRNA or antisense oligonucleotide specific for the gene or mRNA encoding LARG or Rho.

5 Suitable vectors are well known in the art. The vector may be viral or non-viral (e.g. a plasmid). Generally siRNAs are expressed as stem loops which may be rapidly processed within the cell to form "free" siRNA (Tuschl *Nat Biotech* Vol 20(5), 446-448 (2002)). Vectors for expression of siRNAs are often based on RNA Pol III promoters, since these are
10 particularly suited to accurate expression of very short RNA sequences. Suitable vector systems are described in Brummelkamp TR et al *Science* vol 296, 550-553 (2002); Lee NS et al *Nat Biotech* vol 20, 500-505 (2002); Paul CP et al *Nat Biotech* vol 20, 497-500 (2002).

15 The vector may include sequences that direct autonomous replication, or may include sequences sufficient to allow integration into the host cell DNA.

Useful vectors include, for example, plasmids, cosmids or viral vectors,
20 such as, replication defective retroviruses, adenoviruses and lentiviruses.

Selection of an appropriate vector is within the knowledge of those skilled in the art.

25 Expression of the siRNA and/or the antisense molecule from the vector may be driven by a promoter. The promoter may be activated only in specific cells, thus giving the vector target specificity. For example, the siRNA may only be expressed in cells in which DC-SIGN has been activated.

According to a further aspect the invention provides a host cell transfected with a vector encoding an siRNA or antisense oligonucleotide molecule according to the invention.

- 5 According to another aspect the invention provides a pharmaceutical composition comprising one or more compositions capable of inhibiting the expression and/or activity of LARG and/or Rho and one or more pharmaceutically acceptable excipients, diluents or carriers.
- 10 The pharmaceutical composition may also comprise one or more of fillers, salts, buffers, stabilisers and solubilisers.

The pharmaceutical composition may comprise one or more siRNAs or antisense molecules directed to the genes or mRNA encoding LARG
15 and/or Rho, and/or one or more small molecule inhibitors of the LARG and/or Rho proteins, and/or one or more ribozymes or DNazymes targeted to LARG and/or Rho.

For delivery into cells *in vivo* siRNAs or antisense oligonucleotides may
20 be formulated with a lipid based carrier, including, for example, oil-in-water emulsions, micelles and liposomes. Liposomes are commercially available from Gibco BRL as Lipofectamine™ or Oligofectamine™, which are formed of cationic lipids. Methods for making liposomes are well known in the art. Liposomes may be targeted to a particular tissue by
25 coupling the liposome to a tissue specific ligand, such as, a monoclonal antibody, sugar, glycolipid or protein.

Liposomes may also be used to deliver vectors encoding a siRNA or an antisense oligonucleotide according to the invention.

The skilled man will be able to work out a therapeutically effective amount of a pharmaceutical composition according to the invention to be administered to a subject.

- 5 The pharmaceutical composition may be used for the treatment of HIV infection, or any other pathogen that acts via DC-sign and the activation of LARG and/or Rho,

10 According to a yet further aspect the invention provides the use of a composition capable of inhibiting the expression and/or activity of the LARG and/or Rho protein, or a vector capable of expressing a composition capable of inhibiting the expression and/or activity of the LARG and/or Rho protein, in the manufacture of a medicament for the treatment of HIV or SIV infection.

15

The pharmaceutical composition or medicament may be used to treat HIV-1 and/or HIV-2 and/or SIV infection.

20 The pharmaceutical composition or medicament may be for therapeutic or prophylactic treatment of HIV infection.

25 According to another aspect the invention provides a method of treating HIV infection by administering a composition capable of inhibiting the expression and/or activity of LARG and/or Rho to an infected cell or organism.

The treatment may be therapeutic or prophylactic.

30 According to another aspect the invention provides a composition capable of inhibiting the expression of LARG and/or Rho for use in the treatment of HIV infection.

The skilled man will appreciate that preferred features of any one embodiment and/or aspect of the invention may be applied to all other embodiments and/or aspects of the invention.

5

In particular the skilled man will appreciate that the composition described with reference to the first aspect of the invention can be used with all other aspects of the invention.

10 The present invention will be further described in more detail, by way of example only, with reference to the following figures in which:

Figure 1 – is a table listing at least some of the proteins phosphorylated on DC-SIGN activation;

15

Figure 2A to 2F – shows that DC-SIGN triggering induces Rho activity via LARG. Figure 2A shows Western blot results of DCs left non-stimulated or stimulated with either IB10 or H-200 for 10 min, wherein the protein has been LARG immunoprecipitated and blotted for anti-phosphotyrosine (upper blot) and LARG (lower blot). Figure 2B shows the results of H-200 stimulated or rabbit IgG exposed DCs, which have been lysed and subjected to Rho activation assay over a time period of 30 min. Rho immunoblot following pull-down of GTP-Rho using Rhoketin-agarose is shown in the top panel. Western blot for total levels of Rho in WCL is shown in the lower panel. Figure 2C shows a Rac immunoblot following pull-down of GTP-Rac using PBD-agarose Western blot (top panel) and Rac Western blot of WCL (lower panel) on lysates obtained from H-200 or LPS stimulated DCs. Figure 2D shows that when NS siRNAs and LARG siRNAs were transfected into DCs, LARG siRNAs caused a knockdown in LARG expression.

20

25

30

This was confirmed by Western blot for total LARG protein levels (upper panel) and an actin Western blot (lower panel). Figure 2E shows the results of NS siRNA and LARG siRNA transfected DCs in a Rho activation assay pre and post DC-SIGN activation with H-200. Rho immunoblot following pull-down of GTP-Rho using Rhoketin-agarose is shown in the upper blot and Rho Western blot of whole cell lysates is shown in the lower blot. Figure 2F shows analysis of DCs transfected with NS siRNA and LARG siRNA and 24 hr post transfection either mock infected or exposed to HIV-1 BAL for 10 min. Rhoketin-agarose pull-down of GTP-bound Rho is shown in the upper panel, LARG WCL levels in the middle panel and total Rho levels in the lower panel.

Figure 3 - shows that active Rho and DC-SIGN form a complex. Figure 3A shows anti-DC-SIGN stimulated DCs lysed at different time points, and DC-SIGN immunoprecipitated with anti-DC-SIGN. The precipitates were probed with anti-LARG (WB LARG), anti-Rho (WB Rho) and anti-DC-SIGN (WB DC-SIGN). Figure 3B shows non-stimulated and H-200 stimulated DCs subjected to Rhoketin-Agarose Rho-GTP pull-down and Western blot for Rho (upper panel), DC-SIGN (middle panel) and LARG (lower panel). Total Rho levels are shown in the bottom panel.

Figure 4 - shows HIV utilises the DC-SIGN signalling path to facilitate immune synapse formation. Figure 4A shows DCs transfected with NS siRNA or LARG siRNA and Western blot for LARG performed 24 hr post transfection (top panel). Actin immunoblot is shown in the lower panel. Figure 4B shows NS siRNA or LARG siRNA transfected cells either mock infected or exposed to HIV-1 BAL 24 hr post transfection. The number of viral synapses formed (defined as majority of HIV focused at the

contact zone) were counted and presented in the bar graph. Mean $\pm$ SD of three independent experiments is shown. Figure 4C shows RT (reverse transcriptase) activity of DC-CD4 T+ cell co-cultures post transfection of NS or LARG siRNAs and infection with HIV-1 BAL. Mean $\pm$ SD of five replicate experiments is shown.

Figure 5 – is the mRNA sequence of LARG (SEQ ID NO. 3). The sequence of LARG can also be found in the SwissProt database using the accession number Q9NZN5 and from GenBank using accession number AF180681.

Figure 6 – is the protein sequence of LARG (SEQ ID NO. 4). This sequence can also be found in GenBank using accession number AF180681.

Figure 7 – is the protein sequence of RhoA (SEQ ID NO. 5). This sequence of RhoA can also be found in the SwissProt database using the accession number P61586 or BC001360.

Figure 8 – is the mRNA sequence of RhoA (SEQ ID NO. 6).

Summary

Using an anti-DC-SIGN antibody that induces signalling in DCs in a similar manner to that occurring following contact of DCs with HIV-1 the important role of the Rho guanine nucleotide exchange factor (GEF) LARG and Rho in DC-SIGN signalling was identified. The examples and data presented show that LARG is a therapeutic target for treating HIV. It will be understood that Rho could similarly be used as a therapeutic for HIV. Following DC-SIGN triggering on exposure of DCs to HIV1, LARG activates Rho, and Rho and LARG form a complex with DC-SIGN. The

significance of DC-SIGN mediated Rho activation for HIV-1 pathogenesis is demonstrated using siRNA mediated LARG knockdown. A key function for this activity is identified in DC-SIGN mediated infectious synapse formation and viral replicative capacity, in that the number of infectious
5 synapses observed is markedly reduced in DCs following LARG inhibition, correlating with reduced HIV reverse transcriptase (RT) activity in DC-CD4+ T cell co-cultures.

EXAMPLE 1- Phosphoproteome investigation of DC-SIGN signalling

10

A proteomic study of tyrosine phosphorylated proteins following DC-SIGN activation showed several proteins were phosphorylated, including LARG. The proteins identified as being phosphorylated are shown in Figure 1.

15

To obtain this data immature DCs were stimulated using H-200. H-200 is a rabbit polyclonal anti-DC-SIGN antibody obtained from Santa-Cruz Biotechnology. H-200 recognises an epitope in the stalk domain of DC-SIGN, and causes tetramerisation of DC-SIGN, a feature required for
20 normal ligand internalisation. H-200 stimulates DC-SIGN, mimicking HIV-1 binding. Any proteins tyrosine phosphorylated following DC-SIGN stimulation with H-200 were captured using an immobilised anti-phosphotyrosine antibody. Associating proteins were eluted and subjected to in-solution trypsinolysis LC/MS/MS analysis to identify the
25 proteins (data not shown).

EXAMPLE 2 - DC-SIGN triggering by anti-DC-SIGN or HIV exposure results in LARG-dependent Rho activation

The ability of DC-SIGN to phosphorylate LARG was confirmed by LARG immunoprecipitation and antiphosphotyrosine Western blotting after anti-DC-SIGN activation (Figure 2A).

- 5 The ability of DC-SIGN to activate Rho in DCs was investigated by exposing DCs to plate bound anti-DC-SIGN over 30 min and then subjecting the cells to the Rho activation assay. In the Rho activation assay Rhoketin-agarose was used to pull down GTP bound Rho from DC lysates to compare to the total levels of activated Rho in the lysates.
- 10 DC-SIGN activation resulted in markedly increased Rho activity at 3-5 min (Figure 2B). A comparison was made with the activity of the GTPase Rac, whose activation is associated with DC maturation. Rac-GTP pull-down was performed from DC-SIGN activated cells using the glutathione agarose bound p21-binding domains (GST-PBD) of PAKs with no
- 15 detectable increase in GTP bound Rac (Figure 2C).

The role of LARG in the induction of Rho activity was investigated by LARG knockdown in DCs. Figure 2D shows the effect of LARG siRNA on LARG protein levels 24 hours post transfection. DCs were then left

20 non-stimulated or stimulated using anti-DC-SIGN and were either transfected with NS siRNA (wherein NS siRNA has a sequence that does not match any known human gene) or LARG siRNA. The results show that a reduction in LARG protein levels in DCs leads to a failure to induce Rho activity to levels comparable with NS siRNA transfectants

25 following DC-SIGN triggering (Figure 2E).

To determine whether HIV-1 also induces Rho activity in a LARG dependent manner, DCs were again transfected with either NS siRNA or LARG siRNA for 24 hours. The DCs were then either mock infected or

30 exposed to HIV-1 BAL for 10 min. Rho pull-down revealed increased Rho GTP in HIV-1 exposed DCs where LARG expression was preserved.

In contrast, no such increase in GTP bound Rho was seen in DCs where LARG expression was diminished (Figure 2F). These results demonstrate both DC-SIGN activation by H-200 and by HIV-1 leads to induction of Rho activity in DCs that occurs via the LARG protein.

5

EXAMPLE 3 - Active Rho and DC-SIGN form a signalling complex

As Rho can be an integral part of signalling complexes at cellular membranes, whether Rho or DC-SIGN could interact either constitutively or upon DC-SIGN activation was investigated. DC-SIGN and Rho were not associated in the absence of stimulation. However activation of DC-SIGN either with anti-DC-SIGN (H-200) or HIV-1 BAL led to rapid and transient association of the two proteins as detected by immunoblotting using antibody to Rho or LARG following DC-SIGN immunoprecipitation (Figure 3A). The kinetics of DC-SIGN:LARG:Rho association were tested following stimulation with anti-DC-SIGN over 30 min. Rhoketin agarose was used to pull-down activated Rho and Western blotting for Rho, DC-SIGN and LARG performed. The association of the proteins was maximal between 3-5 min representing a time interval similar to Rho activation (Figure 3B). These observations indicate that DC-SIGN mediated Rho activation occurs at the cell membrane and involves a direct interaction between DC-SIGN, LARG and Rho.

EXAMPLE 4 - HIV-1 relies on DC-SIGN activated Rho and LARG for viral synapse formation

DC-SIGN has previously been shown to promote the formation of infectious synapses between DCs pulsed with HIV-1 and resting CD4+ T cells. The mechanism by which DC-SIGN mediates this effect remains unclear; although DC-SIGN can be detected within the synapse contact site between DCs and CD4+ T cells (Arrighi JF et al *J Exp Med.* 200,

1279-88. (2004)). Rho activity has been shown to be essential for allowing contacts between DCs and T cells in immune synapse formation (Eun SY et al *J Immunol.* 177, 4271-5 (2006); Decker M et al *Microbiol Immunol.* 291, 61-90 (2005)). Thus, whether DC-SIGN mediated Rho
5 activation might be responsible for facilitating viral synapse formation in HIV-1 infection was investigated.

DCs were transfected with either NS siRNA or LARG siRNA and 24 hr post transfection Western blotting was performed to confirm LARG
10 knockdown (Figure 4A). NS or LARG siRNA transfected cells were then either mock infected or pulsed with HIV-1 BAL and incubated with highly purified resting CD4+ T cells for 10 min 30 min before fixation and analysis by confocal microscopy. Infectious synapses defined by viral focusing at the DC-T cell contact site were counted for cells transfected
15 with either NS siRNA or LARG siRNA. The number of conjugates formed between DCs and CD4+ T cells where LARG expression was decreased were reduced by 47% at 10 min and sustained reduction occurred at 30 min (Figure 4B). These results suggest LARG mediated Rho activation induced by DC-SIGN is important for efficient HIV-1 viral
20 synapse formation. Infectious synapse formation facilitates the propagation of HIV-1 in DC-CD4+ T cell co-cultures. To examine whether LARG influences HIV replicative capacity in DC-CD4+ T cell co-cultures, DCs were again transfected with NS or LARG siRNAs and 24 hr post transfection exposed to HIV-1 BAL for 2 hr before
25 introduction of autologous CD4+ T cells. HIV RT activity was then assessed at time points post transfection. A marked reduction in HIV RT activity was observed following LARG knockdown (Figure 4C), supportive of a role for DC-SIGN mediated LARG-Rho activation in facilitating HIV-1 replication.

Methods

DC isolation, stimulation, transfection and proliferative assays.

Buffy coats were obtained from the National Blood Centre. Circulating monocyte derived DCs were isolated from buffy coats by adherence and culture in IL-4 and granulocyte-macrophage colony-stimulating factor (GM-CSF). DC preparations analyzed were more than 98% pure. mDC were isolated from Ficoll-Hypaque-enriched mononuclear cells using the BDCA-1 isolation kit (Miltenyi Biotec). Myeloid DCs were cultured for 24 hr with IL-4 at 10 ng/ml to induce expression of DC-SIGN. DCs were stimulated using plate bound anti-DC-SIGN antibodies at 10 µg/ml or plate bound Fc-ICAM3 (R+D systems). Transfections were performed using the Amaxa transfection reagent according to the manufacturer's instructions. For MLR assays DC's were stimulated for 24 hr before washing and plating at graded doses to 10⁵ allogenic T cells. After 4 days proliferating cell populations were labelled with H3-thymidine for 18 hr.

Antibodies, Rho, Rac and Cdc42 activation assays and siRNAs.

Antibody to LARG was obtained from Abcam; anti-phosphotyrosine and anti-phosphoserine antibodies were obtained from Sigma and Cell Signaling; the H-200 DC-SIGN antibodies were from Santa-Cruz and R&D systems, Rho from Pierce, Cdc42/Rac from Chemicon, β-actin from Sigma, anti-CD36, MHC-II, MHC-I, CD80, CD83, CD86 and CD11c from Pharmingen. Anti-p24 was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr Michael H.Malim. Rho activation assays were performed using the Rho activation kit from Pierce and Cdc42 activation assays were performed using the Cdc42 activation assay kit from Chemicon. Custom high performance purity (HPP) grade siRNAs were obtained from Qiagen. The target DNA sequence used to construct *DC-SIGN* siRNA was 5'-AAG GCT GCA GTG GGT GAG CTT 3' (SEQ ID NO. 7), *MyD88* siRNA 5'-AGG ACC CTA AAT CCA ATA GAA -3' (SEQ ID NO. 8), *LARG*

siRNA 5'-GAA ACT CGT CGC ATC TTC C -3' (SEQ ID NO. 1), and NS siRNA 5'-AAT TCT CCG AAC GTG TCA CGT-3' (SEQ ID NO. 9). The siRNA sequences given represent only one strand of the siRNA duplex.

5

Immunoprecipitation and Western Blot.

For immunoprecipitations, cell lysates were incubated with the corresponding antibody and either Protein A or G-agarose. For Western blot analysis immunoprecipitates or WCL were resolved on SDS-PAGE, transferred to PVDF membrane (Amersham Biosciences) and detected by the indicated antibodies using ECL system (Amersham Biosciences).

10

Transfections

Transient transfection of DCs was carried out using Amaxa primary DC nucleofector kit.

15

Virus stocks, confocal, infectious synapse assay formation assay

Viral stocks were generated by transfection of 293T cells with calcium-phosphate co-precipitated proviral plasmid which encodes for full length HIV-1 BAL strain. Viral titres were determined by RT assay (Amersham). HIV-1 BAL was added to DCs in RPMI supplemented with 10% FCS in the presence or absence of 5 mM EGTA or in the presence or absence of IB10 blocking antibody. Cells were incubated with HIV-1 for 2 hr prior to analysis. CD4+ T cells were prepared by CD8+ dynabead depletion. For infectious synapse assays, 3×10^5 CD4+ T cells were left to adhere on poly-L-lysine-treated glass coverslips for 2 h at 37°C. Mature sorted DCs (10^5) cells were pulsed with HIV-1 BAL for 2 h at 37°C. DCs were washed twice and left to adhere at 37°C on coverslips for 10 min and 30 min to allow contact with previously seeded T cells. Cells were fixed by a 20-min incubation in 3% paraformaldehyde at room temperature and permeabilized with 0.1% Triton X-100 and washed

20

25

30

several times with PBS containing 5% BSA. Cells were stained with anti-Gag antibody. An infectious synapse was measured as a DC-T cell conjugate where the majority of HIV is focused at the contact zone with the CD4+ T cell (>75% HIV). We imaged and quantified in each experiment 60-70 DC-T cell conjugates for the 10 min time point and 170-180 DC-T cell conjugates for the 30 min time point. Confocal microscopy was performed using a Bio-Rad Radiance 2000 laser scanning confocal and analyzed using LaserSharp 2000 software (Bio-Rad). All images were acquired in sequential scanning mode.

10

LC/MS/MS analysis

Material eluted from the anti-phosphotyrosine columns was digested with trypsin and purified using Sepack C18 columns (Waters, Milford, MA, USA), following the manufacturer's instructions. Peptide mixtures were separated on a reversed-phase C18 column (10cm, 75µm ID, Dionex/LC-Packings, Sunnyvale, CA, USA) on a 60 minute gradient 0-55% (solvent A: 98% H₂O, 2% acetonitrile, 0.1% formic acid; solvent B: 80% acetonitrile, 20% H₂O, 0.1% formic acid) using a nano-liquid chromatography system (Ultimate Plus, LC-Packings) coupled to a high capacity iontrap tandem mass spectrometer (Bruker Daltonics, Bremen, Germany). MS/MS spectra were analyzed using Mascot (Matrix Science, UK). Proteins were considered identified when at least two peptides were matched, for which at least four consecutive a or b ions were observed. Tyrosine phosphorylation sites were inspected and verified manually.

25

CLAIMS

1. A composition for inhibiting the expression and/or activity of the leukaemia-associated Rho guanine nucleotide exchange factor (LARG)
5 and/or the Rho protein in a cell.
2. A pharmaceutical composition comprising one or more compositions capable of inhibiting the expression and/or activity of LARG and/or Rho and one or more pharmaceutically acceptable
10 excipients, diluents or carriers.
3. The use of a composition capable of inhibiting the expression and/or activity of the LARG and/or Rho protein, or a vector capable of expressing a composition capable of inhibiting the expression and/or
15 activity of the LARG and/or Rho protein, in the manufacture of a medicament for the treatment of HIV infection.
4. A method of treating HIV infection by administering a composition capable of inhibiting the expression and/or activity of LARG and/or Rho
20 protein to a cell or organism.
5. A composition capable of inhibiting the expression and/or activity of LARG and/or Rho for use in the treatment of HIV infection.
- 25 6. A composition, use or method of any preceding claim wherein the composition comprises a synthetic siRNA molecule or a synthetic antisense oligonucleotide directed to the LARG or Rho mRNA.
7. A composition, use or method of claim 6 wherein the siRNA
30 molecule comprises two RNA strands having a region of complementarity

of between 19 and 21 nucleotides, which allows the strands to form an RNA duplex of 19 to 21 nucleotide pairs.

8. A composition, use or method of claim 6 wherein the siRNA
5 molecule comprises a single RNA strand having a region of self-complementarity, which allows the strands to form an RNA duplex of 19 to 21 nucleotide pairs.

9. A composition, use or method of claim 7 or 8 wherein the siRNA
10 comprises between 19 and 21 nucleotides complementary to 19 to 21 consecutive nucleotides in the LARG or Rho mRNA.

10. A composition, use or method of claim 6, 7, 8 or 9 wherein one
strand of the siRNA comprises the sequence: 5'- GAA ACT CGT CGC
15 ATC TTC C -3' (SEQ ID NO. 1).

11. A composition, use or method of any of claims 1 to 5 comprising an inhibitor of LARG or Rho protein activity.

20 12. A composition, use or method of claim 11 wherein the inhibitor is a small molecule.

13. A composition, use or method of any of claims 1 to 5 comprising a ribozyme or DNAzyme which inhibits the expression and/or activity of
25 LARG and/or Rho.

14. A composition, use or method of any preceding claim wherein the composition is capable of reducing the level of active LARG and/or Rho protein in a cell, or a population of cells, by at least about 10%.

15. A composition, use or method of any preceding claim which is capable of preventing or reducing HIV infection.
16. The pharmaceutical composition according to claim 2, or claim 6
5 to 14, for use for the treatment of HIV-1, HIV-2 and/or SIV infection.
17. The use of claim 3, the method of claim 4 or the composition of claim 5, or any claim dependent thereon, or the pharmaceutical composition of claim 16 wherein the treatment is therapeutic or
10 prophylactic.
18. A vector encoding an siRNA or antisense oligonucleotide specific for a mRNA or a gene encoding the LARG or Rho protein.
- 15 19. A vector according to claim 18 comprising sequences that direct autonomous replication of the vector, or allow integration of at least part of the vector into the host cell DNA.
20. A vector according to claim 18 or 19 selected from the group
20 comprising a plasmid, a cosmid or a viral vector, such as, a replication defective retrovirus, adenovirus or lentivirus.
21. A host cell transfected with a vector encoding an siRNA or antisense oligonucleotide molecule according to any of claims 18 to 20.
25
22. A composition, use or method of claim 6, 7, 8 or 9 wherein one strand of the siRNA comprises the sequence: 5'-AAG GAA GAG AAG GAT GTT AAT-3' (SEQ ID NO. 2).

1/13

Table 2. Summary of Differentially Phosphorylated Proteins		
Functional Category	Phosphorylated in DC-SIGN activated DCs	Phosphorylated in non-stimulated DCs
Signal transduction	(Q9NZN5) Rho guanine nucleotide exchange factor 12 (Leukemia-associated RhoGEF) (P52757) Rho GTPase-activating protein 3 (O14733) MAP kinase kinase 7 (Q9BXU1) Serine/threonine protein kinase 31 (O94804) Serine/threonine protein kinase 10	(Q8IUC4) Rho-GTP binding protein 2 (Q13033) Striatin-3 (Q96P50) centaurin-beta 5
Cytoskeletal and microtubule	(P46821) Microtubule associated protein 1B (P14649) Myosin light chain 1 (P13535) Myosin heavy chain 8 (P06753) Tropomyosin alpha-3 chain (P05976) Myosin light chain 1 (Q9UKX2) Myosin-2 (P68133) Alpha-actin 1 (P09493) Tropomyosin 1 alpha chain (P16157) Ankyrin-1 (P35563) Cyclin-1 (Q9H254) Spectrin beta chain (O60504) Vinexin	(P67936) Tropomyosin alpha-4 chain (P68032) Actin alpha cardiac
Ubiquitination	(Q9UJX3) Anaphase-promoting complex subunit 7 (Q8IX29) F-box only protein 16	
mRNA processing and translation	(O15523) ATP dependent RNA helicase DDX3Y (Q99459) Cdc5-like protein (Q99543) DnaJ homolog subfamily C member 2 (Q96SM3) Probable carboxypeptidase X1 precursor (Q9H7B2) Brix-domain containing protein 1	(P29084) Transcription initiation factor IIE beta subunit (Q04637) eIF-4-gamma 1
Nucleolar proteins, nucleosome assembly and DNA binding	(O94986) Cep152 protein (O43303) Cep 110 protein (P10412) Histone H1.4 (P51449) Nuclear receptor ROR-gamma (O75152) Zinc finger CCCH domain-containing protein 11A (Q14980) NuMA protein	
Transcription	(Q14151) Scaffold attachment factor B2 (P15884) Transcription factor 4	(P19838) NF-kappa-B p105 subunit (Q9Y5R6) Doublesex-related transcription factor 1
Ion channel	(Q9UL51) potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2	(Q92581) Sodium/hydrogen exchanger 6 (O00305) Voltage-dependent L-type calcium channel
Cell surface protein/transmembrane	(P98161) Polycystin 1 precursor (Q13410) Butyrophilin subfamily member 1 (O75949) Transmembrane protein 28 (O95202) Leucine zipper-EF-hand-containing transmembrane protein 1	
Enzyme	(P35520) Cystathionine beta-synthase (Q9H5I1) Histone-lysine N-methyltransferase SUV39H2	(Q01813) Phosphofructokinase 1 (O00750) Phosphatidylinositol-4-phosphate 3-kinase (Q9NYU1) UDP-glucose: glycoprotein glucosyltransferase 2
Scaffold and matrix	(Q9Y6V0) Piccolo protein (Q8WXH2) Junctophilin-3	
Cell survival	(Q8NHY3) GAS2-like protein 2	(Q92614) molecule associated with JAK3 N-terminus (Q9P032) UPF0240 protein C6orf66
Trafficking	(Q10567) beta-adaptin-1 (P53621) Alpha-COP (Q9NV70) Exocyst complex component 1	
Metabolism	(Q00341) Vigilin	
Cell cycle	(O43663) Protein regulator of cytokinesis 1	
Unknown	(Q8IVF6) Ankyrin repeat domain containing protein 18A (Q9NSI6) WD repeat protein 9	(Q15059) RING3-like protein (Q15059) Bromodomain-containing protein 3 (Q8NCN4) RING finger protein 169
Identified proteins are specified by Swiss-Prot accession number. Further information about sequence coverage and peptide sequences can be obtained from the Supplementary Data.		

Figure 1

Figure 2A

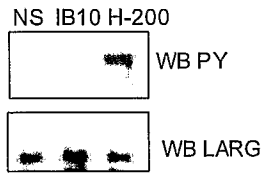


Figure 2B

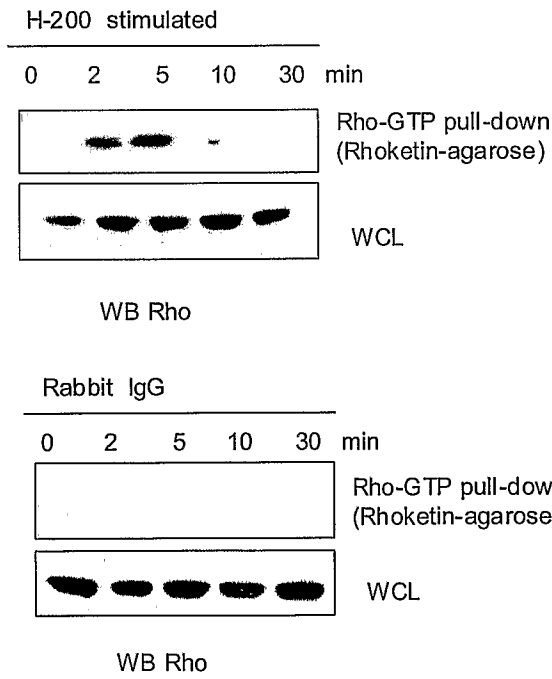


Figure 2C

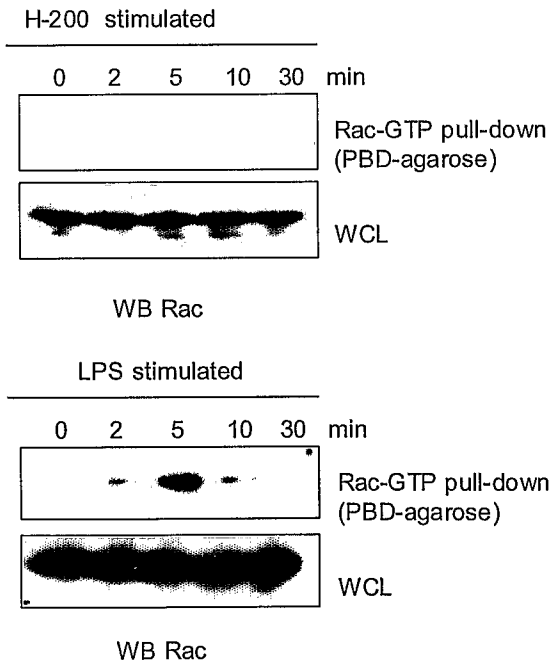


Figure 2D

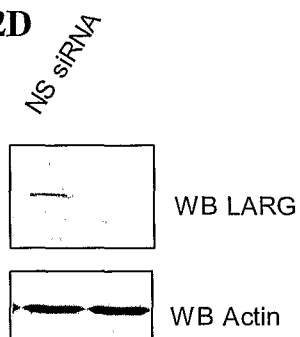


Figure 2F

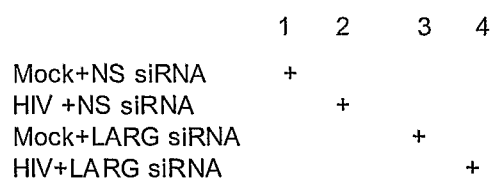


Figure 2E

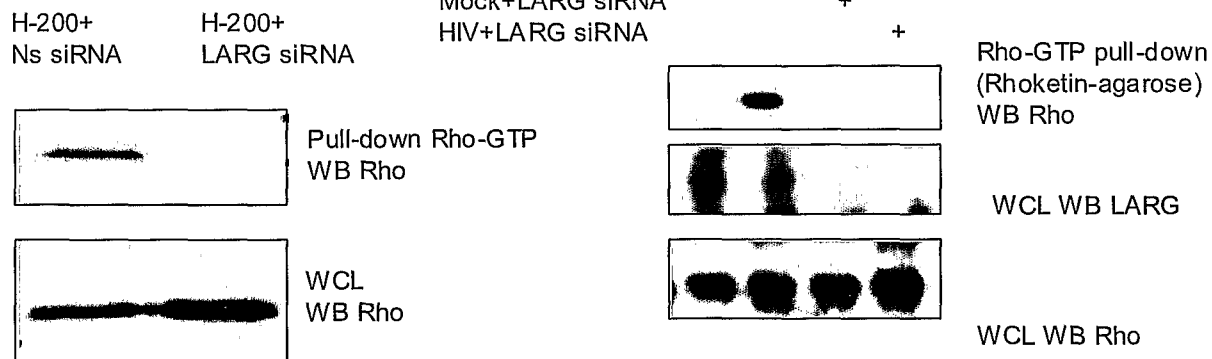


Figure 3A

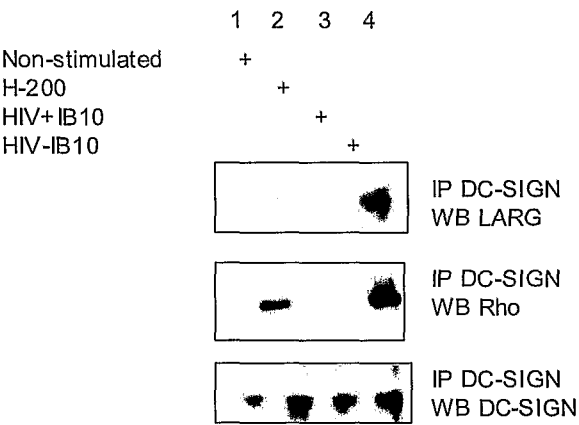
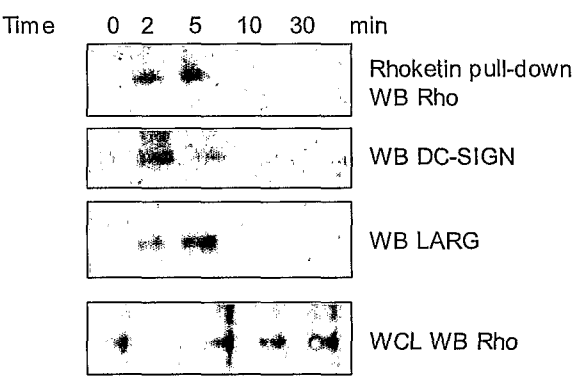


Figure 3B



5/13

Figure 4A

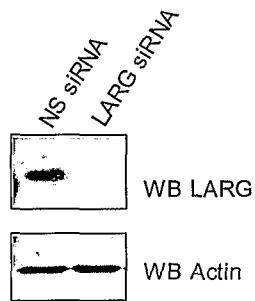


Figure 4B

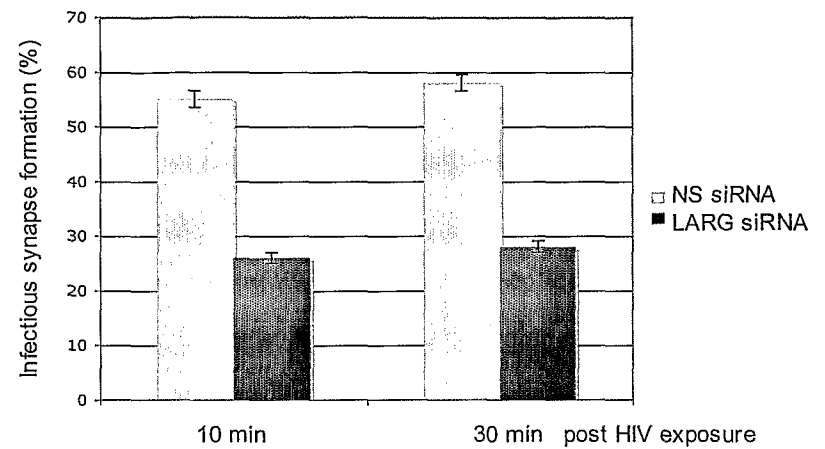
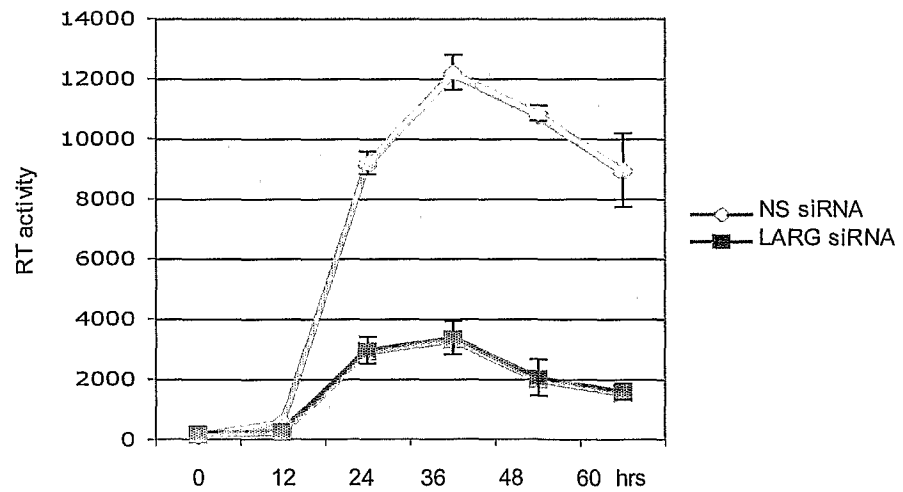


Figure 4C



6/13

1 ggccccaatg agtggcacac agtctactat caccgacagg tttcccctca aaaaacctat
61 aaggcatgga agtatttttg accgagagtc accaacagat aagaagcaga aagttgagcg
121 cattgcatca catgattttg accccacaga tagctcctcc aagaagacaa agtctagttc
181 agaggagagt agatccgaga tatatggtct tgttcagcgt tgcgtaatca tccagaaaga
241 tgacaatgga tttgggctga cggtcagtgg agacaatcca gtcttcgtac agtctgtcaa
301 agaagatgga gcagccatgc gggctggagt acagacaggt gatcgaatca tcaaggtgaa
361 tggaactctg gtgactcatt caaatcatct ggaggtggtg aagctaataa aatctggttc
421 ctatgtagct ctactgttc agggacgccc acctgggtcg cccagattc cacttgccga
481 ctctgaagta gagccgtcag tcattggaca tatgtctccc atcatgacat ctctcattc
541 acctggagca tctgggaata tggagagaat cactagtcct gtgctcatgg gggaggaaaa
601 caatgtggtt cataaccaga aagtagaaat tctgagaaaa atgttacaga aagaacagga
661 acggctacag ttattgcagg aagattacaa ccgaacacct gcccaaagat tgctaaaaga
721 gatccaagag gccaaagaaac acattcctca gctgcaagag cagttatoca aagccacagg
781 ctctgctcag gatggagctg tagttacacc ctccagacct ttaggggaca ccctaacagt
841 cagtgaggca gaaacagatc ctggagatgt actgggcagg actgactgta gcagtggaga
901 tgcttctcgg ccagtagtg acaatgcaga tagtcccaag agtggcccaa aagagagaat
961 ttatctagag gaaaaccag agaaaagtga aacaattcag gacactgaca ctcaatcact
1021 tgtcggaaagt ccctcaaccc gtatagcacc tcatattatt ggagcagaag atgatgattt
1081 tggtagtgaa catgaacaga tcaatggaca gtgcagctgt ttccagagca ttgaattact
1141 aaaatctcgc ccggctcatt tggctgtttt cttacaccat gtagtttcac aatttgacc
1201 tgcgactttg ctctgttatc tctattcaga cctgtataaa cataccaatt ccaaagaaac
1261 tcgtcgcac ttccttgagt ttcattcagtt ctttctagat cgatcagcac acctgaaagt
1321 ttctgttcct gatgaaatgt ctgcagatct agaaaagaga agacctgagc tcattcctga
1381 ggatctgcat cgccactata tccaaactat gcaagaaaga gtccatccag aagttcaaag
1441 gcacttagaa gattttcggc agaaacgtag tatgggactg accttggtg aaagcgagct
1501 gactaaactt gatgcagagc gagacaagga ccgattgact ttggagaagg agcggacatg
1561 tgcagaacag attgttgcca aaattgaaga agtattgatg actgctcagg ctgtagagga
1621 agataagagc tccacatgc agtatgttat tctcatgtat atgaagcatt tgggagtaaa
1681 agtgaaagag cctcgaaatt tggagcacia acgggggtcg attggatttc tccccaaaat
1741 caagcaaagt atgaagaaag ataaagaagg ggaagaaaaa ggaagcgaa gaggattccc
1801 cagcatcctg ggacccccac ggagaccaag ccgtcatgac aacagtgcaa ttggcagagc
1861 catggaacta cagaaggcgc gccaccctaa gcacttatcc acacctcat ctgtgagtcc
1921 tgaacctcag gactctgcca agttgcgcca gagtgggtta gcaaatgaag gaacagacgc
1981 tggataacct cctgccaatt ccatgtcttc tgtagcttca ggggocctt tttcccagga
2041 agggaggaaa gagaatgata caggatcaaa gcaagttgga gaaacatcag cacctggaga
2101 caccttagat ggcacacctc gtactctcaa tactgtcttt gatttcccac cacctccatt
2161 agaccaagtg caggaggagg aatgtgaagt agaaaggtg actgaacatg ggacacaaaa

7/13

2221 gccctttcga aagtttgaca gtgtagcttt tggagaaagt caaagtgagg atgaacaatt
2281 tgaaaatgac ttagagacag acccacccaa ctggcagcag cttgttagtc gagaagtgtt
2341 actgggacta aaaccttgtg aaatcaaaag acaggaagtg attaatgaat tgttctacac
2401 tgaaagagct catgttcgaa cactgaaggt tcttgatcaa gtgttctatc agcgagtatc
2461 cagagaagga attctgtcac cctcagagct acggaaaatt ttttcaaact tggaagatat
2521 tcttcaactt catattggat tgaatgaaca aatgaaggct gttcgaaaga gaaatgagac
2581 ctctgttatc gatcagattg ggggaagattt gctgacatgg ttcagcggac caggagagga
2641 gaaattgaaa catgctgctg ctaccttttg cagtaaccaa cctttcgccc tggaaatgat
2701 caaatctcgt cagaaaaagg attctcgatt tcagactttt gtgcaagatg ctgaaagtaa
2761 tccactgtgt cgtcgtcttc aactgaagga tattattccc actcaaatgc aaaggcttac
2821 taagtaccca cttctgttgg ataattatgc caaatacaca gaatggccaa cagaaagggg
2881 gaaggtgaag aaagctgcag atcactgtcg tcagatctta aattatgtaa atcaggctgt
2941 caaggaggca gaaaacaagc agcgcctaga agattatcag cgtcgccttg atacctccag
3001 cctgaagtgt tcagagtacc caaatgttga agagctcagg aatttggatt taacaaaaag
3061 gaagatgatt catgaagggc cattggtttg gaaggtgaat agagataaaa ctattgatth
3121 atacacgttg ctgctggaag acattcttgt attgttataa aagcaggatg atagactggt
3181 tttaaggtgt catagtaaga ttctggcatc tacagctgat agcaaacaca cgtttagccc
3241 tgtcattaag ttgagtacag tgttggttcg acaagtggca acagataaca aagctttatt
3301 cgtcatttcc atgtcagaca atggcgctca gatttatgaa ctggtggcac agacagtttc
3361 tgaaaagact gtctggcagg acctaactctg tcggatggct gcatcagtga aggagcaatc
3421 cacaaagcca attccattac cacagtcaac acctggcgaa ggagataatg atgaagaaga
3481 tccttcaaaa ttaaaagagg agcagcatgg catttcagtc actggtttgc agagtccaga
3541 cagagatttg ggattagaat ctaccttaat atcgtcaaaa cctcagtctc attcactgag
3601 tacctctggg aaatcagagg tacgtgatct gtttgtggct gagagacagt ttgcaaagga
3661 acaacataca gatgggacac taaaggaagt tggagaagat tatcaaactc caatcccaga
3721 ttcacacctg cctgtctcag aagaacggtg ggcattggat gcactaagaa atttgggttt
3781 gttgaagcag ttgctggtgc aacagctagg tttgactgag aagagcgttc aggaagactg
3841 gcaacatttc ccaagataca gaacagcctc tcaggggccg cagacagaca gtgtcatcca
3901 gaactctgaa aatattaagg cctatcattc tgggtgaagga catatgccct ttagaactgg
3961 aactgggtgac attgcaactt gttacagtcc acggacttca actgaatctt ttgctccacg
4021 ggattcagtg ggaactggcac ccaggatag ccaggcaagt aacattttag taatggacca
4081 catgattatg accccagaga tgcctaccat ggagccagaa gggggtcttg atgacagtgg
4141 agagcacttt tttgatgccc gtgaagcaca tagtgatgag aatccatcag aaggtgatgg
4201 agcagttaac aaggaagaga aggatgttaa tttacgcatc tcaggaaact atttgatcct
4261 tgatggctat gaccagtgac aggagagttc cacagatgag gaggttgctt cctcacttac
4321 cctgcagccc atgacaggca tccctgctgt ggaatccacc caccagcagc aacattctcc
4381 tcagaatact cactccgatg gggcaatttc accattcacc cccgaatttc tgggtccagca

8/13

4441 gcgctgggga gctatggagt attcctgttt tgagatccag agtccctcct cttgtgcaga
4501 ttcacagagc cagatcatgg agtacattca taagatagag gctgaccttg aacacttaaa
4561 gaagggtggag gaaagttaca ccattctttg ccaaaggctg gctggatcag ccctcacaga
4621 caagcactca gataaaagtt agagccgcat gtcttgaggg tgactgcagg ttgttggatt
4681 tggagtatcg gccgtgtctc accacatcct ggctccagtg tggatgcaga gagagtgtga
4741 cagaggatct gcctgtgaac cacctgggat tagtcaagtc ccaagggtgc cagagtggga
4801 ctagtctctc acagtgtggc agctgcacta atctgtttgt gagggaaatat ccattccctc
4861 actctactct cctcactatc ggaaattcat tttgattcag aataaaaacc aaatgtatag
4921 agctttgggt gtaggatatg aaattgtact tagatttaag aaaaagagaa aatcagatgt
4981 atttatttga cttcattccg tatttgaaag cacattttaa tttttatttg ccttgttttg
5041 ttttaattga gtagtgagag ttttagacct ttgtctttag tacaccaag gatcaactgc
5101 tcctgaagca aagaggtcag gatggagtat gcagaagttg ggtagagagg ttaagaagag
5161 gaggaagtga gacggggaag ggaggagcac cttgcttcac tatacagtat tccaagctct
5221 ctgctgtcca gtaggctgct tctctgaggg tcacctcaa ggtatgctat gccgtgtggc
5281 tcttatgtgc ccagggtgtg tggtcagaga gtggatgggc ttctcccgcc cctgaggcaa
5341 gcacctcttc tctgtagctg ggatcaacca caattaatag gaatcctcaa cataactaat
5401 agcaggcact tgaaaaatggg tgtgttttct tccattgtcg tcttttctat tgagggtttg
5461 gatttgggtg ggagagggaa gcaaatttta ttttcttgac tataaatttg ttattcttgg
5521 tatcatttca tttttataat tatacattag acattggcac ttgtgtaaaa ctgtccctgc
5581 agattgagca ggaagtaaaa acaaattgaa atgcttcgga tagtggcagg ggtgggggtt
5641 gctaaggaaa ggggtgggatg ggagtgggta aagagtttg gaaggttatc taactgaatc
5701 actactgagt tgaatcatgg ctatcattag ccacaggtag tgctgattat aaggccagta
5761 aaatttttagt acctggaggt ttgactctaa tttttgact gatggtgcca aagggccctg
5821 agtcaaaagg atagccaagg gtggagtgga aaaagatatg gtgaaggcag aactgctcac
5881 accagctcca gaagcacgtc ctctgacttc actgccgcag ctctttccac acggggccgt
5941 gattgaccct aaaaattgca gaccagcaca tacactggac actgcaggca aggtgttgg
6001 aactgcctgc tctaggatga atagtagcgt tagcagcacc ctacagaggg taagtctcag
6061 aattgtaatt tgagaactcc ttcaattata ttgactttct ttggttttcc ttgtgcatta
6121 ggtatatatt ttttgaagta tttttgcaa ctaaaatgaa gtctgtaaat tctgttaata
6181 aacaaggagt tcatccgttg ctcacatctt tcattggtgc cctctgaact ctgtgggttg
6241 ctaggatgta attttaatgc ttccctgcag tccaaagatg attttttcac cacaatggt
6301 aagggatgcc cacctacttt tataaacacc actgcaactt aacaagttta tttatctatg
6361 tccagatttc tgtttctgtc cttaaattgat ctgtgttttt aggtggatca acttggatct
6421 ttagacctca tctataaatt gaaattatat ttttagtcat aagccaagta caatctaact
6481 cagaatggga ttaaaaattt tagaagcaga agctaataa taaatgaagt ttgggatttg
6541 gaactttctg tatctcttag gaggaacaag taaaaccac atggttacat tgtgctgcta
6601 gaaataatgt cattctcaat taaaagggtt acacctgtag ccacttgaca ctaacaccat

9/13

6661 cgagggcaat taatcaggag aaaagtaaat ataaaaactt tgctottaac gtcacacact
6721 acagggtaat tatgtaattc tctagcactc aattaagaag cttgtaggtt agcaaaatth
6781 cagtttctct agttcagcca gtgtcttgtc cttagtaatt atttgctttc cctgcagttc
6841 aatttctcct tggaactcta acaggactaa ctggaaaaat actttgacca tagctctagt
6901 acttccaaat aaattcatat ctagattttc aaacaaagcc ccaagatgaa gtctaattct
6961 gaaaatatct cacaattttt gaatgttctt tttatcttac cgacctaaact tctcttttcc
7021 ctcccaactt ggctccccc acgttccatt tcattttggc cattatagtc acatccgttt
7081 gactttgggt gtgacatctt cctttcagga tgatggaatt agatgcagtc tgtctttgag
7141 tgagactggg caccatcaga ggagtttttt ctaagtgcg tgactagaag ttgaaatcta
7201 ctccctgct aaagggcaca ggggtgggtg acaagaaaag ctacctccc agagcaagca
7261 ggctgcattt tagctatgga agtgacctgc tgtgacactg tgctctcttc tgtgggctta
7321 atggttcctt tgctatgaag tggcaaatta catgtagagt gtctccttcc ttttcagaga
7381 acagttaatc aaggcaaatc agcaagcccc caaagtgcg taatttaaca tcatgattac
7441 caccttcgaa gctatatatt ttgcatactt taaaatcacc taacttggac tgcttgaatt
7501 acattggctt ttagaaccga aattgtaact atgtattgta tttcatggga ggtatattht
7561 atgagcttht tggctthctt ttttccaca gcaactgcc atgaaggatg aatcccctt
7621 ttaaaaagt gtgtgtgtg ttgtttatth gatthtgagt taggagggat aatagagaag
7681 tccattthaa aattattht agaagctaaa gaaagtaatt atgcttctg tgaattgtct
7741 tttactggca tctttgttht cctctttgat gttagtaaat ttggtgtaat acgtggggct
7801 tccatatht aaagtgaag cthtctctc tgaagtcgat atatggttht gaattactag
7861 agctthtgct agtathtct tccctatatg tcacagagg caccactgag aactgcgtgc
7921 ataggacctc aaaatacaaa attagcagg cctcacagtc agcttctca tggctagtht
7981 tccccctta tattacaatt tttgttht aagtcatht tttcctgata tttccaccac
8041 ttttcagagt catctacaaa atthtcttht cctcaagaaa agagthtctt ttgccttatt
8101 ccttatgcct tccccactg tattgagggt tttgataata aattggtag aaaaaaagt
8161 acctcctaga aggaagcct cccaccatt tccagggtgc aactgctaag cagatatatt
8221 ccaaaaatgg taactgtcat gtgcacactg ttggttatt ttaataagcc tcttctact
8281 agaacatht atthtcttg ttaccatac aatcatgtac tctttaacag aaattgctt
8341 taaaaaatat ctggaactat cthtaaaaa actthtattaa taatcatgta tthttactga
8401 tcacathtg aaatgcctaa aagacttht tgttctaatt atccagatgt accttthgta
8461 aatagctctt ttatgaatta gctgataagg ctgtatgtht ctggaacaaa atattggtca
8521 tctaaaaact tctgttht tggggtctg gaaaatagaa aataagattt caaatattaa
8581 ataagcttaa aggaaccaag catgtgactt ttactthgta tctggttga tttgtctgt
8641 ccaagttata ttagaaaagt tgaggagtcg aggagcctgt aattaatht cctgtcagtg
8701 actgtcagac atcttctgc ttattagagc atccctagca acaaggctaa accttcatga
8761 cagtgtggc aggtgggctt gccctagcat gtggggcagg tttgttht tthtaatagc
8821 tttattgaga tattcacata tcatataatt catccattta agtgtataa tcagatgggt

10/13

8881 tttggtatat tattgat tttt taattgtggt gaaatatata tataacaaaa tttgctattg
8941 taactat tttt taagtataaa gttcagtgag gtaggccagc ttttatgacg cacctgaaga
9001 taacttttgggt tttaatccta ccctaaatga agcatatgcc tctat tttttc ctgtccccag
9061 tgctgggctt cttcacatac caacatttct gtgcttttac caacagcatt ttagctcaga
9121 aggctcgctt actttgggca tttgctgtat tttgg ttttg aaatacttgt aatttgactt
9181 attcttg gga taatatttta gttttcagtc attcaagggc acatttg gct gacggagcta
9241 ataagcgtaa taaaatgctt agtagtttat accatctttt agttaaacat tttaaagatt
9301 aagactattc aaatgtactt tgcggaaatt aacctgtttg tatgcctggc tttgtttaca
9361 aacatatatc tatatctata tatatagata cagatacaga tacatatctg gtggaacaaa
9421 gaaatacctg taccattccc acttgctctt tgatagccac ctctaggaa tggagacaat
9481 aaaactctgt atttgcattg g

Figure 5

11/13

```

1      msgtqstitd rfplkkpirh gsilnrespt dkkqkveria shdfdptdss sktkkssee
61     srseiyclvq rcviiqkddn gfgltvsgdn pvfvqsvked gaamragvqt gdriikvngt
121    lvthsnhlev vkliksgsyv altvqgrppg spqipladse vepsvighms pimtsphspg
181    asgnmerits pvlmgeennv vhnqkveilr kmlqkeqerl qllqedynrt paqrllkeiq
241    eakkkhipqlq eqlskatgsa qdgavvtpsr plgdtltvse aetdpqdvlg rtdcssgdas
301    rpssdnadsp ksgpkeriyl eenpekseti qdtdtqslvg spstriaphi igaedddfgt
361    ehedngqcs cfqsiellks rpahlavflh hvvsqfdpat llcylysdly khtnsketrr
421    iflefhqffl drsahlkvsv pdemsadlek rrpelipedl hrhyiqtmqe rvhpevgrhl
481    edfrqkrsmg ltlaeseltk ldaerdkdrl tlekertcae qivakieevl mtaqaveedk
541    sstmqyvilm ymkhlgvkvk eprnlehkrq rigflpkikq smkkdkegee kgkrrgfpsi
601    lgpprrpsrh dnsaigrame lqkarhpkhl stpssvspep qdsaklrqsg lanegtdagy
661    lpansmssva sgasfsqegg kendtgskqv getsapgdtl dgtprtlnlv fdfppppldq
721    vqeeceever vtehgtpkpf rkfdsvafge sqsedeqfen dletdppnwq qlvsrevllg
781    lkpceikrge vinelfyter ahvrtlkvld qvfyqrvsre gilspselrk ifsnledilq
841    lhiglineqmk avrknetstv idqigedllt wfsgpgeekl khaaatfcsn qpfailemiks
901    rqkkdsrfqt fvdqdaesnpl crrlqlkdi ptqmqrlltky pllldniaky tewpterekv
961    kkaadhcrqi lnyvnqavke aenkqrledy qrrldtsslk lseypnveel rnldltkrkm
1021   ihegplvwkv nrdktidlyt llledilvll qkqddrlvlr chskilasta dskhtfspvi
1081   klstvlvrqv atdnkalfvi smsdngaaiy elvaqtvsek tvwqdlicrm aasvkeqstk
1141   piplpqstpg egdndeedps klkeeqhgis vtglqspdrd lglestliss kpqshslsts
1201   gksevrdlfr aerqfakeqh tdgtlkevge dyqiaipdsh lpvseerwal dalrnlgllk
1261   qllvqqlglt eksvqedwqh fpryrtasqg pqttsvigns enikayhsge ghmpfrtgtg
1321   diatcysprt stesfaprds vglapqdsqa snilvmdhmi mtpemptmep egglddsgeh
1381   ffdareahsd enpsegdgav nkeekdvnlr isgnylildg ydpvqesstd eevassltlq
1441   pmtgipaves thqqqhspqn thsdgaissp tpeflvqgrw gameyscfei qspsscadsq
1501   sqimeyihki eadlehlkkv eesytilcqr lagsaltdkh sdks

```

Figure 6

12/13

MAAIRKKLVIVGDGACGKTCLLIVFSKDQFPEVYVPTVFENYVADIEV
D GKQVELALWDTAGQEDYDRLRPLSYPD TDVILMCFSIDSPDSLENIP
EKWTPEVKHFCPNVPIILVGNKKDLRND EHTRRELAKMKQEPVKPEEG
RDMANRIGAFGYMECSAKTKDGVREVFEMATRAALQARRGKKKSGCLV
L

Figure 7

13/13

```

1   ctcgttagtc cacggtctgg tcttcagcta cccgccttcg tctccgagtt tgcgactcgc
61  ggaccggcgt ccccggcgcg aagaggctgg actcggattc gttgcctgag caatggctgc
121 catccggaag aaactgggtga ttgttggtga tggagcctgt ggaaagacat gcttgctcat
181 agtcttcagc aaggaccagt tcccagaggt gtatgtgccc acagtgtttg agaactatgt
241 ggcagatatc gaggtggatg gaaagcaggt agagttggct ttgtgggaca cagctgggca
301 ggaagattat gatcgctga ggcccctctc ctaccagat accgatgtta tactgatgtg
361 tttttccatc gacagccctg atagtttaga aaacatccca gaaaagtgga cccagaagt
421 caagcatttc tgtcccaacg tgcccatcat cctgggtggg aataagaagg atcttcggaa
481 tgatgagcac acaaggcggg agctagccaa gatgaagcag gagccggtga aacctgaaga
541 aggcagagat atggcaaaca ggattggcgc ttttgggtac atggagtgtt cagcaaagac
601 caaagatgga gtgagagagg tttttgaaat ggctacgaga gctgctctgc aagctagacg
661 tgggaagaaa aaatctgggt gccttgtctt gtgaaacctt gctgcaagca cagcccttat
721 gcggttaatt ttgaagtgtt gtttattaat cttagtgtat gattactggc ctttttcatt
781 tatctataat ttacctaaga ttacaaatca gaagtcattt tgctaccagt atttagaagc
841 caactatgat tattaacgat gtccaacccg tctggcccac cagggtcctt ttgacactgc
901 tctaacagcc ctctctgca ctcccacctg acacaccagg cgctaattca aggaatttct
961 taacttcttg cttctttcta gaaagagaaa cagttggtaa cttttgtgaa ttaggctgta
1021 actactttat aactaacatg tcctgcctat tatctgtcag ctgcaaggta ctctggtgag
1081 tcaccacttc agggctttac tccgtaacag attttgttgg catagctctg ggggtggcag
1141 ttttttgaaa atgggctcaa ccagaaaagc ccaagttcat gcagctgtgg cagagttaca
1201 gttctgtggt ttcatgttag ttaccttata gttactgtgt aattagtgcc acttaatgta
1261 tgttaccaa aataaatata tctaccccag actagatgta gtatTTTTTg tataattgga
1321 tttcctaata ctgtcatcct caaagaaagt gtattggttt tttaaaaaag aaagtgtatt
1381 tggaaataaa gtcagatgga aaattcaaaa aaaaaaaaaa aaaa

```

Figure 8