Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2777198 C 2018/07/24

(11)(21) 2 777 198

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2010/10/06
(87) Date publication PCT/PCT Publication Date: 2011/04/14
(45) Date de délivrance/Issue Date: 2018/07/24
(85) Entrée phase nationale/National Entry: 2012/03/21
(86) N° demande PCT/PCT Application No.: US 2010/051646
(87) N° publication PCT/PCT Publication No.: 2011/044246
(30) Priorités/Priorities: 2009/10/06 (US61/249,253); 2009/10/07 (US61/249,596)

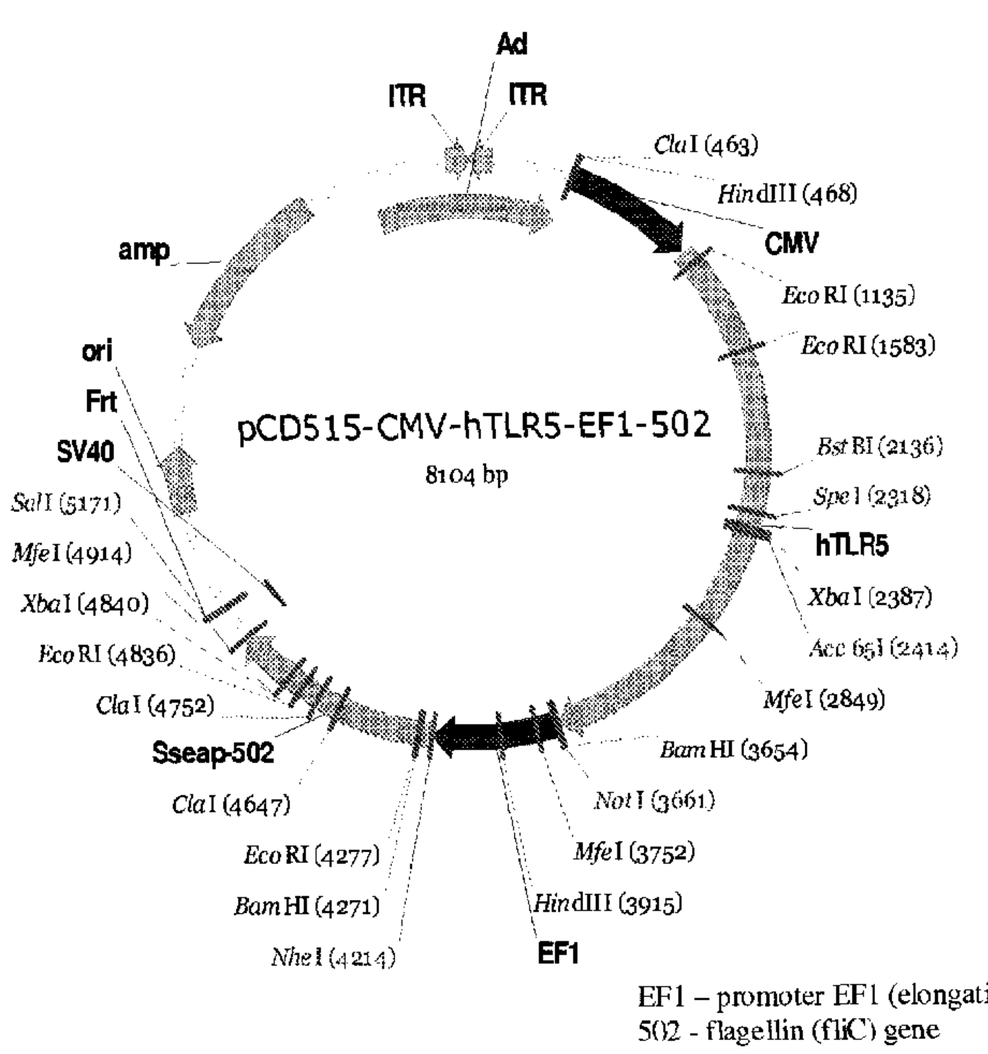
(51) Cl.Int./Int.Cl. *C12N 15/63* (2006.01), *A61K 31/59* (2006.01), *A61K 38/17* (2006.01), *A61K 38/19* (2006.01), *A61K 38/22* (2006.01), *A61K 38/27* (2006.01), *A61P 31/00* (2006.01), *A61P 35/00* (2006.01), *C07K 14/195* (2006.01), *C07K 14/705* (2006.01), *C12N 15/12* (2006.01), *C12N 15/31* (2006.01), *C12N 15/85* (2006.01)

(72) Inventeur/Inventor:
GUDKOV, ANDREI V., US

(73) Propriétaires/Owners:
ROSWELL PARK CANCER INSTITUTE, US;
PANACELA LABS, LLC, RS

(74) Agent: BCF LLP

(54) Titre : UTILISATION DE RECEPTEUR DE TYPE TOLL ET AGONISTE POUR LE TRAITEMENT DU CANCER
(54) Title: USE OF TOLL-LIKE RECEPTOR AND AGONIST FOR TREATING CANCER



(57) Abrégé/Abstract:
The present invention is directed to methods and agents used for treating cancer or infectious diseases by providing toll-like receptors such as toll-like receptor 5 (TLR-5) in combination with providing a toll-like receptor agonists such as flagellin resulting in a cis and in-trans effect that recruits cells involved in both the innate (cis effect) and adaptive (trans effect) immune response to specifically kill cancer cells and cells infected with a pathogen via the NF-κB apoptosis pathway.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 April 2011 (14.04.2011)

PCT

(10) International Publication Number
WO 2011/044246 A1

(51) International Patent Classification:
C12N 15/63 (2006.01)

(21) International Application Number:
PCT/US2010/051646

(22) International Filing Date:
6 October 2010 (06.10.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/249,253	6 October 2009 (06.10.2009)	US
61/249,596	7 October 2009 (07.10.2009)	US

(71) Applicant *(for all designated States except US)*: **CLEVELAND BIOLABS, INC.** [US/US]; 73 High Street, Buffalo, NY 14203 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **GUDKOV, Andrei, V.** [US/US]; 1933 Sweet Road, East Aurora, NY 14052 (US).

(74) Agent: **SCOTT, Teddy, C.**; Polsinelli Shughart PC, 161 North Clark Street, Suite 4200, Chicago, IL 60601 (US).

(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, CL, 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, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

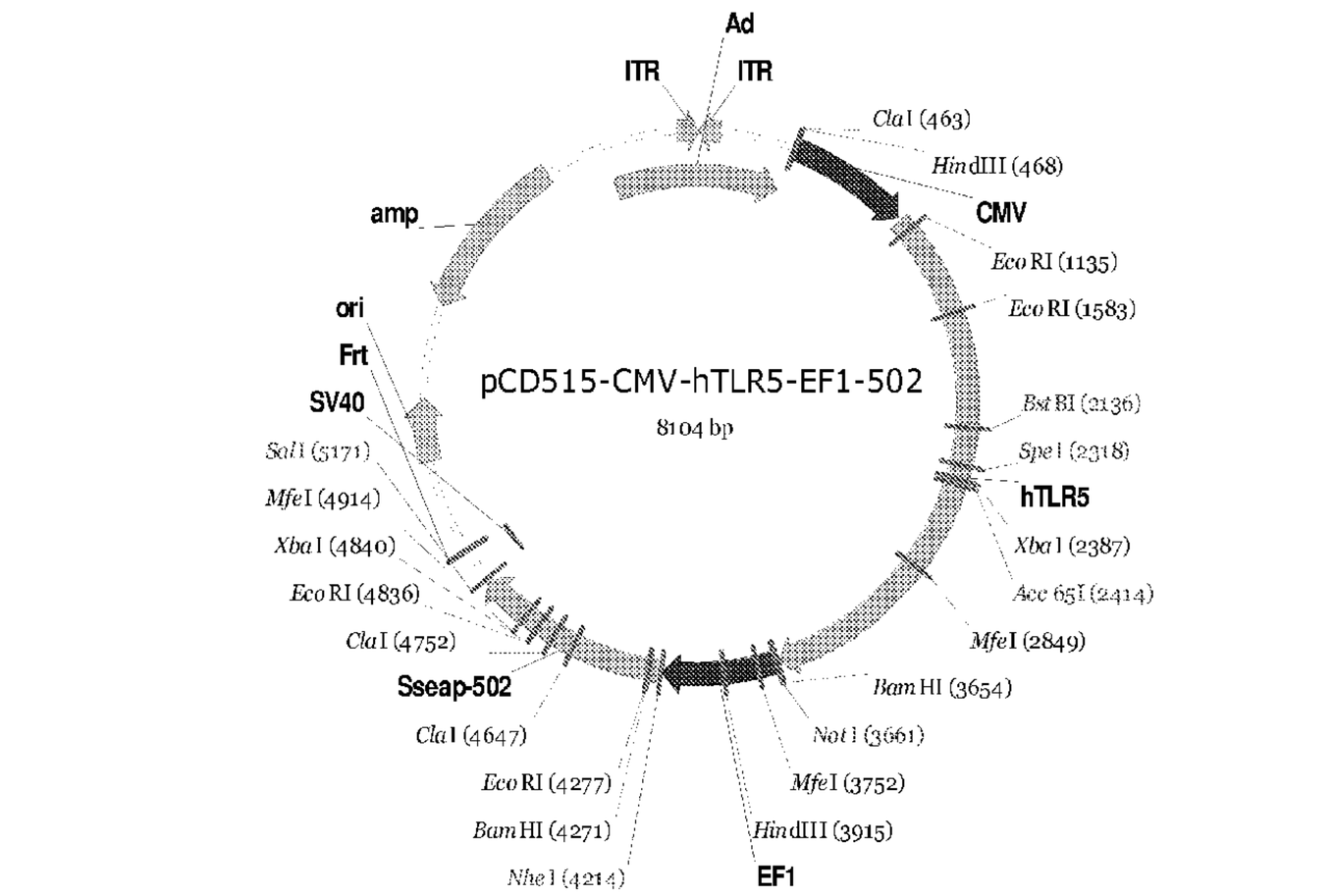
— *of inventorship (Rule 4.17(iv))*

[Continued on next page]

(54) **Title**: USE OF TOLL-LIKE RECEPTOR AND AGONIST FOR TREATING CANCER

(57) **Abstract**: The present invention is directed to methods and agents used for treating cancer or infectious diseases by providing toll-like receptors such as toll-like receptor 5 (TLR-5) in combination with providing a toll-like receptor agonists such as flagellin resulting in a cis and in-trans effect that recruits cells involved in both the innate (cis effect) and adaptive (trans effect) immune response to specifically kill cancer cells and cells infected with a pathogen via the NF-κB apoptosis pathway.

FIGURE 1A



EF1 – promoter EF1 (elongation factor 1α).
502 - flagellin (fliC) gene

WO 2011/044246 A1

WO 2011/044246 A1

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

USE OF TOLL-LIKE RECEPTOR AND AGONIST FOR TREATING CANCER

FIELD OF THE INVENTION

[0001] This invention relates to methods of treating cancer and infectious diseases.

BACKGROUND OF THE INVENTION

[0002] Toll-like receptors are responsible for the recognition of most common patterns of bacterial and viral pathogens. Their activation results in recruitment of innate and subsequently adaptive immune response. Receptor cells of the immune system to the site of presence of antigens is the key step in effective immune response. That is why immunization involves the use of different types of adjuvants. Although the majority of tumors express tumor-specific antigens, they are using a number of mechanisms allowing then to escape immune recognition. It was recently demonstrated in mouse models that activation of TLR5 by its ligand and agonist, bacterial flagellin, results in the induction of antitumor effect against those tumors that express functional TLR5. This opens a general opportunity for considering TLR5 agonists for cancer immunotherapy. There are two major obstacles on the way to reduction of this idea to practice. First, is the rare incidence of tumors expressing functional TLR5 limiting applicability of this approach to only a small subset of tumors. Second, systemic administration of TLR5 agonist leads to activation of TRL5 signaling in all cells that have functional receptor making response unfocused and not tumor-specific. Accordingly, there is a need in the art for a mechanism or method for autocrine activation of TLR receptor signaling in infected or tumor cells with minimal systemic effect thus enabling to attract innate immune response specifically to the infected cell or tumor.

SUMMARY OF THE INVENTION

[0003] The present invention may be directed to a vector comprising a first and second nucleic acid, wherein the first nucleic acid encodes a toll-like receptor and the second nucleic acid encodes a toll-like receptor agonist. The first nucleic acid may encode for a secreted form of a toll-like receptor. The second nucleic acid may be a secreted form of flagellin. The toll-like receptor agonist may be flagellin. The vector may be a mammalian expression vector. The

vector may be expressed from an adenovirus, a lentivirus or a liposome. The secreted form of flagellin may be CBLB502S. The toll-like receptor may be TLR-5.

[0004] The present invention may be directed to a method of treating cancer in a mammal comprising administering to a mammal in need thereof a agent comprising the vector comprising a first and second nucleic acid, wherein the first nucleic acid encodes a toll-like receptor and the second nucleic acid encodes a toll-like receptor agonist. The cancer may be a tumor. The tumor may be derived from the group consisting of prostate, breast, colon, esophagus, stomach, lung, pancreatic, renal, thyroid, ovaries, throat, or the cervix. The tumor may be derived from the group consisting of sarcomas, melenomas, leukemias, and lymphomas. The agent may be administered in trans or outside from the tumor of the mammal. The agent may be administered directly into a tumor of the mammal. The agent may be administered in combination with an immunostimulant. The immunostimulant may be selected from the group consisting of growth hormone, prolactin and vitamin D. The growth hormone may be somatotrophin. The agent may be administered in combination with a cytokine. The cytokine may be a stem cell factor.

[0005] The present invention is also directed to a method for treating an infection in a mammal comprising administering to a mammal in need thereof the agent of agent comprising the vector comprising a first and second nucleic acid, wherein the first nucleic acid encodes a toll-like receptor and the second nucleic acid encodes a toll-like receptor agonist. The cancer may be a tumor. The infection may be derived from the group consisting of viruses, bacteria, protozoan parasites and fungi.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] Figure 1A-1C depicts schematic maps of adenoviral vectors expressing TLR5, CBLB502S and their combination (TLR5 + CBLB502S).

[0007] Figure 2 depicts the results of the ratio of tumor volume in mice over a number of days in tumor cells (A549) transduced with a control vector (without TLR5) or vector expressing TLR5 wherein the mice are treated three days with either CBLB502 or PBS.

[0008] Figure 3 depicts suppression of tumor growth by injection of adenovirus comprising vector coexpressing CBLB502S and Toll-like receptor wherein the adenovirus is injected into syngeneic mice CT26 colon carcinoma tumor and studying the in-cis and in-trans effects of the adenoviral vector constructs.

[0009] Figure 4 shows the domain structure of bacterial flagellin. The Ca backbone trace, hydrophobic core distribution and structural information of F41. Four distinct hydrophobic cores that define domains D1, D2a, D2b and D3. All the hydrophobic side-chain atoms are displayed with the Ca backbone. Side-chain atoms are color coded: Ala, yellow; Leu, Ile or Val, orange; Phe and Tyr, purple (carbon atoms) and red (oxygen atoms). c, Position and region of various structural features in the amino-acid sequence of flagellin. Shown are, from top to bottom: the F41 fragment in blue; three b-folium folds in brown; the secondary structure distribution with a-helix in yellow, b-structure in green, and b-turn in purple; tic mark at every 50th residue in blue; domains D0, D1, D2 and D3; the axial subunit contact region within the proto-element in cyan; the well-conserved amino-acid sequence in red and variable region in violet; point mutations in F41 that produce the elements of different supercoils. Letters at the bottom indicate the morphology of mutant elements: L (D107E, R124A, R124S, G426A), L-type straight; R (A449V), R-type straight; C (D313Y, A414V, A427V, N433D), curly33.

[0010] Figure 5 shows a schematic of Salmonella flagellin domains, its fragments, and its interaction with TLR5. Dark bars denote regions of the flagellin gene used to construct fragments comprising A, B, C, A' and B'.

[0011] Figure 6 depicts flagellin derivatives. The domain structure and approximate boundaries (amino acid coordinates) of selected flagellin derivatives (listed on the right). FliC flagellin of Salmonella dublin is encoded within 505 amino acids (aa).

[0012] Figure 7 shows the nucleotide and amino acid sequence for the following flagellin variants: AA' (SEQ ID NO: 7-8), AB' (SEQ ID NO: 9-10), BA' (SEQ ID NO: 11-12), BB' (SEQ ID NO: 13-14), CA' (SEQ ID NO: 15-16), CB' (SEQ ID NO: 17-18), A (SEQ ID NO: 19-20), B (SEQ ID NO: 21-22), C (SEQ ID NO: 23-24), GST-A' (SEQ ID NO: 25-26), GST-B' (SEQ ID NO: 27-28), AA'n1-170 (SEQ ID NO: 29-30), AA'n1-163 (SEQ ID NO: 33-34), AA'n54-170 (SEQ ID NO: 31-32), AA'n54-163 (SEQ ID NO: 335-36), AB'n1-170 (SEQ ID NO: 37-38), AB'n1-163 (SEQ ID NO: 39-40), AA'n1-129 (SEQ ID NO: 41-42), AA'n54-129 (SEQ ID NO: 43-44), AB'n1-129 (SEQ ID NO: 45-46), AB'n54-129 (SEQ ID NO: 47-48), AA'n1-100 (SEQ ID NO: 49-50), AB'n1-100 (SEQ ID NO: 51-52), AA'n1-70 (SEQ ID NO: 53-54) and AB'n1-70 (SEQ ID NO: 55-56). The pRSETb leader sequence is shown in Italic (leader includes Met, which is also amino acid 1 of FliC). The N terminal constant domain is underlined.

The amino acid linker sequence is in Bold. The C terminal constant domain is underlined. GST, if present, is highlighted.

[0013] Figure 8 shows a comparison of amino acid sequences of the conserved amino (Fig. 8A) and carboxy (Fig. 8B) terminus from 21 species of bacteria. The 13 conserved amino acids important for TLR5 activity are shown with shading. The amino acid sequences are identified by their accession numbers from TrEMBL (first letter = Q) or Swiss-Prot (first letter = P).

[0014] Figure 9 shows the amino acid sequence for the human Toll-like receptor 5 protein.

DETAILED DESCRIPTION

[0015] The inventors have made the surprising discovery that the provision of a toll-like receptor, such as toll-like receptor 5 (TLR-5), in combination with a toll-like receptor agonist, such as flagellin, results in a cis and in-trans effect that recruits cells involved in both the innate (cis effect) and adaptive (trans effect) immune response to specifically kill cancer cells and cells infected with a pathogen via the NF-κB apoptosis pathway. While not being bound by theory, the idea implemented in this invention was to (i) overcome the dependence of TLR-mediated immunization strategies on pre-existing TLR expression in a tumor by transducing the tumor with a construct driving expression of TLR; and (ii) to direct the immune response to the tumor by creating local pool of TLR agonist. For example, drug formulations comprising TLR simultaneously induce expression and activate TLR, thereby exposing tumor cells to the host immune system imitating the situation of massive bacterial penetration through the intestinal wall.

[0016] By providing a TLR such as TLR5, and a TLR agonist such as flagellin, to interact and activate both the innate and adaptive immune system, the method can be used to treat tumors derived from the prostate, breast, colon, esophagus, stomach, lung, pancreatic, renal, thyroid, ovaries, throat, or the cervix cancer as well as treating sarcomas, melenomas, leukemias, and lymphomas. Applications of this method are not limited to cancer treatments, as this method can also be used to treat infections derived from viruses, bacteria, protozoan parasites and fungi.

[0017] Variations of providing the TLR and TLR agonist may include vectors, co-expressing the TLR receptor and a secretable form of flagellin that activates TLR activity in the same compromised mammalian cell. The method of the present invention may also include vector

constructs that express the TLR receptor in a mammalian cell and the TLR agonist being administered in trans to the cell. For example, an adenoviral vector may require modification of flagellin to reach its effective synthesis and secretion by mammalian cells.

1. Definitions.

[0018] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. As used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

[0019] For recitation of numeric ranges herein, each intervening number there between with the same degree of precision is explicitly contemplated. For example, for the range of 6-9, the numbers 7 and 8 are contemplated in addition to 6 and 9, and for the range 6.0-7.0, the numbers 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6,9, and 7.0 are explicitly contemplated.

[0020] "Administer" may mean a single dose or multiple doses of an agent or agent.

[0021] "Analog" may mean, in the context of a peptide or polypeptide, a peptide or polypeptide comprising one or more non-standard amino acids or other structural variations from the conventional set of amino acids.

[0022] "Antibody" may mean an antibody of classes IgG, IgM, IgA, IgD or IgE, or fragments, or derivatives thereof, including Fab, F(ab')2, Fd, and single chain antibodies, diabodies, bispecific antibodies, bifunctional antibodies and derivatives thereof. The antibody may be a monoclonal antibody, polyclonal antibody, affinity purified antibody, or mixtures thereof which exhibits sufficient binding specificity to a desired epitope or a sequence derived therefrom. The antibody may also be a chimeric antibody. The antibody may be derivatized by the attachment of one or more chemical, peptide, or polypeptide moieties known in the art. The antibody may be conjugated with a chemical moiety.

[0023] A "derivative" may mean a peptide or polypeptide different other than in primary structure (amino acids and amino acid analogs). Derivatives may differ by being glycosylated, one form of post-translational modification. For example, peptides or polypeptides may exhibit glycosylation patterns due to expression in heterologous systems. If at least one biological activity is retained, then these peptides or polypeptides are derivatives according to the invention. Other derivatives may include fusion peptides or fusion polypeptides having a covalently modified N- or C-terminus, PEGylated peptides or polypeptides, peptides or

polypeptides associated with lipid moieties, alkylated peptides or polypeptides, peptides or polypeptides linked via an amino acid side-chain functional group to other peptides, polypeptides or chemicals, and additional modifications as would be understood in the art.

[0024] A "fragment" may mean a portion of a reference peptide or polypeptide.

[0025] A "homolog" may mean a peptide or polypeptide sharing a common evolutionary ancestor.

[0026] A "leader sequence" may be a nucleic acid encoding any peptide sequence that is linked and translated with a peptide or polypeptide of interest to allow the peptide or polypeptide of interest be properly routed through a eukaryotic cell's endoplasmic reticulum and Golgi complexes for the purposed of extracellular secretion from the cell's membrane. The leader peptide sequence may be derived from alkaline phosphatase. The leader sequence may have a DNA sequence comprising atgctgctgctgctgctgctgctgggcctgaggctacagctct ccctgggc.

[0027] A "liposome" may mean a tiny bubble (vesicle) made out of the same material as a cell membrane. A liposome be filled with drugs and used to deliver drugs for cancer and other diseases. A liposome may be filled with a vector. A liposome membrane may be made of phospholipids, which are molecules that have a head group and a tail group. The head of the liposome may be attracted to water, and the tail, which is made of a long hydrocarbon chain, is repelled by water. The tails may be repelled by water, and line up to form a surface away from the water. The lipids in the plasma membrane may be chiefly phospholipids like phosphatidylethanolamine and phosphatidylcholine. Liposomes may be composed of naturally-derived phospholipids with mixed lipid chains (like egg phosphatidylethanolamine), or of pure surfactant components like DOPE (dioleoylphosphatidylethanolamine).

[0028] A "peptide" or "polypeptide" may mean a linked sequence of amino acids and may be natural, synthetic, or a modification or combination of natural and synthetic.

[0029] "Substantially identical" may mean that a first and second amino acid sequence are at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%,or 99% over a region of 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100 amino acids .

[0030] "Treating," "treatment," or "to treat" each may mean to alleviate, suppress, repress, eliminate, prevent or slow the appearance of symptoms, clinical signs, or underlying pathology of a condition or disorder on a temporary or permanent basis. Preventing a condition or disorder

involves administering a agent of the present invention to a subject prior to onset of the disease. Suppressing a condition or disorder involves administering a agent of the present invention to a subject after induction of the condition or disorder but before its clinical appearance. Repressing the condition or disorder involves administering a agent of the present invention to a subject after clinical appearance of the disease.

[0031] A "variant" may mean means a peptide or polypeptide that differs in amino acid sequence by the insertion, deletion, or conservative substitution of amino acids, but retain at least one biological activity. Representative examples of "biological activity" include the ability to bind to a toll-like receptor and to be bound by a specific antibody. Variant may also mean a protein with an amino acid sequence that is substantially identical to a referenced protein with an amino acid sequence that retains at least one biological activity. A conservative substitution of an amino acid, i.e., replacing an amino acid with a different amino acid of similar properties (e.g., hydrophilicity, degree and distribution of charged regions) is recognized in the art as typically involving a minor change. These minor changes can be identified, in part, by considering the hydropathic index of amino acids, as understood in the art. Kyte *et al., J. Mol. Biol.* 157:105-132 (1982). The hydropathic index of an amino acid is based on a consideration of its hydrophobicity and charge. It is known in the art that amino acids of similar hydropathic indexes can be substituted and still retain protein function. In one aspect, amino acids having hydropathic indexes of ±2 are substituted. The hydrophilicity of amino acids can also be used to reveal substitutions that would result in proteins retaining biological function. A consideration of the hydrophilicity of amino acids in the context of a peptide permits calculation of the greatest local average hydrophilicity of that peptide, a useful measure that has been reported to correlate well with antigenicity and immunogenicity. U.S. Patent No. 4,554,101.

Substitution of amino acids having similar hydrophilicity values can result in peptides retaining biological activity, for example immunogenicity, as is understood in the art. Substitutions may be performed with amino acids having hydrophilicity values within ±2 of each other. Both the hyrophobicity index and the hydrophilicity value of amino acids are influenced by the particular side chain of that amino acid. Consistent with that observation, amino acid substitutions that are compatible with biological function are understood to depend on the relative similarity of the amino acids, and particularly the side chains of those amino acids, as revealed by the hydrophobicity, hydrophilicity, charge, size, and other properties.

[0032] A "vector" may mean a nucleic acid sequence containing an origin of replication. A vector may be a plasmid, a yeast or a mammalian artificial chromosome. A vector may be a RNA or DNA vector. A vector may be either a self-replicating extrachromosomal vector or a vector which integrates into a host genome.

2. Toll-Like Receptor

[0033] Provided herein is a toll-like receptor (TLR), which may be a type of pattern recognition receptor (PRR). The TLR may recognize molecules that are conserved molecular products derived from pathogens that include Gram-positive, Gram-negative bacteria, fungi, and viruses, but are distinguishable from host molecules, collectively referred to as pathogen-associated molecular patterns (PAMPs). The TLR may also recognize endogenous molecules released from injured or dying cells, collectively referred to as damage-associated molecular pattern (DAMPs). A PAMP or DAMP may be a TLR agonist as further described below. The TLR may be a fragment, variant, analog, homolog or derivative that recruits adapter molecules within the cytoplasm of cells in order to propagate a signal. The TLR may be from a human or other mammalian species such as rhesus monkey, mouse, or rat. The TLR may be at least 30-99% identical to a TLR that recruits adapter molecules within the cytoplasm of cells in order to propagate a signal.

[0034] The TLR may be one of the between ten and fifteen types of TLR that are estimated to exist in most mammalian species. The TLR may be one of the 13 TLR (named simply TLR1 to TLR13) that have been identified in humans and mice together, or may be an equivalent form that has been found in other mammalian species. The TLR may be one of the 11 members (TLR1-TLR11) that have been identified in humans.

[0035] The TLR may be expressed by different types of immune cells, and may be located on the cell surface or in the cell cytoplasm. The TLR may be expressed on cancer cells. The TLR may be expressed by normal epithelial cells in the digestive system, normal keratinocytes in the skin, alveolar and bronchial epithelial cells, and epithelial cells of the female reproductive tract. These cells lining an organ may be the first line of defense against invasion of micoorganisms, and TLRs expressed in epithelial cells may have a crucial role in the regulation of proliferation and apoptosis.

[0036] The TLR-expressing cancer cell may be selected from the following table:

Table 1 TLR expression in Human Cancer cells

TYPE OF CANCER	TLR
Gastric cancer	TLR2, TLR4, TLR5, TLR9
Colorectal cancer	TLR2, TLR3, TLR4, TLR5, TLR9
Ovarian cancer	TLR2, TLR3, TLR4, TLR5
Cervical cancer	TLR3, TLR4, TLR5, TLR9
Lung cancer	TLR2, TLR3, TLR4, TLR9
Prostate cancer	TLR4, TLR9
Melanoma	TLR2, TLR3, TLR4
Brain cancer	TLR2, TLR4
Breast cancer	TLR2, TLR3, TLR4, TLR9
Hepatocellular carcinoma	TLR2, TLR3, TLR4, TLR6, TLR9
Laryngeal cancer	TLR2, TLR3, TLR4

[0037] The TLR expressed on cancer cells may upregulate the NF-κB cascade and produce anti-apoptotic proteins that contribute to carcinogenesis and cancer cell proliferation.

[0038] Four adapter molecules of TLRs are known to be involved in signaling. These proteins are known as myeloid differentiation factor 88 (MyD88), Tirap (also called Mal), Trif, and Tram. The adapters activate other molecules within the cell, including certain protein kinases (IRAK1, IRAK4, TBK1, and IKKi) that amplify the signal, and ultimately lead to the induction or suppression of genes that orchestrate the inflammatory response. TLR signaling pathways during pathogen recognition may induce immune reactions via extracellular and intracellular pathways mediated by MyD88, nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), and mitogen-associated protein kinase (MAPK). In all, thousands of genes are activated by TLR signaling, and collectively, the TLR constitute one of the most pleiotropic, yet tightly regulated gateways for gene modulation.

[0039] TLRs together with the Interleukin-1 receptors form a receptor superfamily, known as the "Interleukin-1 Receptor/Toll-Like Receptor Superfamily." All members of this family have in common a so-called TIR (Toll-IL-1 receptor) domain. Three subgroups of TIR domains may exist. Proteins with subgroup I TIR domains are receptors for interleukins that are produced by macrophages, monocytes and dendritic cells and all have extracellular Immunoglobulin (Ig) domains. Proteins with subgroup II TIR domains are classical TLRs, and bind directly or indirectly to molecules of microbial origin. A third subgroup of proteins containing TIR domains (III) consists of adaptor proteins that are exclusively cytosolic and mediate signaling from

proteins of subgroups 1 and 2. The TLR may be a fragment, variant, analog, homolog or derivative that retains either a subgroup I TIR domain, subgroup II TIR domain, or subgroup III TIR domain.

[0040] The TLR may function as a dimer. For example, although most TLRs appear to function as homodimers, TLR2 forms heterodimers with TLR1 or TLR6, each dimer having a different ligand specificity. The TLR may also depend on other co-receptors for full ligand sensitivity, such as in the case of TLR4's recognition of LPS, which requires MD-2. CD14 and LPS Binding Protein (LBP) are known to facilitate the presentation of LPS to MD-2.

a. TLR1

[0041] The TLR may be TLR1, which recognizes PAMPs with a specificity for gram-positive bacteria. TLR1 has also been designated as CD281.

b. TLR5

[0042] The TLR may be Toll-like receptor 5. The protein encoded by the TLR-5 may play a fundamental role in pathogen recognition and activation of innate immunity. TLR-5 may recognize PAMPs that are expressed on infectious agents, and mediate the production of cytokines necessary for the development of effective immunity. TLR-5 may recognize bacterial flagellin, a principal component of bacterial flagella and a virulence factor. The activation of the TLR may mobilize the nuclear factor NF-κB and stimulate tumor necrosis factor-alpha production.

3. Toll-like Receptor Agonist

[0043] Also provided herein is a TLR agonist. The TLR agonist may be a PAMP, which may be conserved molecular product derived from a pathogen. The pathogen may be a Gram-positive bacterium, Gram-negative bacterium,, fungus, or virus. The TLR agonist may be a damage-associated molecular pattern (DAMP) ligand, which may be an endogenous molecule released from injured or dying cells. A DAMP or PAMP may initiate an immune response through TLR signals and recruit adapter molecules within the cytoplasm of cells in order to propagate a signal. The TLR agonist may be an agonist for the TLR, which may be a ligand from the following in Table 2:

Table 2 TLRs and Ligands

TLR	Ligand DAMP	Ligand PAMP
TLR1		Triacyl lipoproteins
TLR2	Heat Shock proteins	Peptidoglycan
	HMGB1 (high mobility group box 1—amphoterin)	Lipoprotein
		Lipoteichoic acid
		Zymosan
TLR3	Self dsRNA	Viral dsRNA
TLR4	Heat shock proteins	Heat shock proteins
	Fibrinogen	Lipopolysaccharides
	Heparan sulfate	RSV fusion protein
	Fibronectin	MMTV (Mouse mammary tumor virus) envelope proteins
	Hyaluronic acid	Paclitaxel
	HMGB1	
TLR5		flagellin
TLR6		Lipoteichoic acid
		Triacyl lipoproteins
		zymosan
TLR7/TLR8	Self ssRNA	Viral ssRNA
TLR9	Self DNA	Bacterial and viral DNA
TLR10		
TLR11		Profilin

[0044] The TLR agonist may be a fragment, variant, analog, homology or derivative of a PAMP or DAMP that binds a TLR and induces TLR-mediated activity, such as activation of NF-κB activity. The TLR agonsist fragment, variant, analog, homolog, or derivative may be at least 30-99% identical to amino acids of a TLR-agonist and induce TLR-mediated activity.

[0045] The TLR agonist may target a TLR such as TLR-5. The TLR agonist may be an agonist of TLR-5 and stimulate TLR-5 activity. The TLR agonist may be an anti-TLR5 antibody or other small molecule. The TLR agonist may be flagellin.

[0046] The flagellin may also be a flagellin or flagellin-related polypeptide. The flagellin may be from any source, including a variety of Gram-positive and Gram-negative bacterial species. The flagellin may be a flagellin polypeptide from any Gram-positive or Gram-negative bacterial species including, but not limited to, a flagellin polypeptide disclosed in U.S. Pat. Pub. No. 2003/000044429. For example,

the flagellin may have an amino acid sequence from a bacterial species depicted in Figure 7 of U.S. Patent Publication No. 2003/0044429. The nucleotide sequences encoding the flagellin polypeptides listed in Figure 7 of U.S. 2003/0044429 are publicly available at sources including the NCBI Genbank database. The flagellin may also be a flagellin peptide corresponding to an Accession number listed in the BLAST results shown in Fig. 25 of U.S. Patent Pub. 2003/000044429, or a variant thereof. The flagellin may also be a flagellin polypeptide as disclosed in U.S. Patent Appl. Publication No. 2009/0011982. The flagellin may be any one of a flagellin polypeptide as disclosed in Figures 6 and 7 herein.

[0047] The flagellin may be a fragment, variant, analog, homolog or derivative of a flagellin that binds TLR5 and induces TLR5-mediated activity, such as activation of NF-κB activity. A fragment, variant, analog, homolog, or derivative of flagellin may be at least 30-99% identical to amino acids of a flagellin that binds TLR5 and induces TLR5-mediated activity.

[0048] The flagellin may be from a species of Salmonella, a representative example of which is S.dublin (encoded by GenBank Accession Number M84972). The flagellin related-polypeptide may be a fragment, variant, analog, homolog, or derivative of M84972, or combination thereof, that binds to TLR5 and induces TLR5-mediated activity, such as activation of NF-κB activity. A fragment, variant, analog, homolog, or derivative of flagellin may be obtained by rational-based design based on the domain structure of Flagellin and the conserved structure recognized by TLR5.

[0049] The flagellin may comprise at least 10, 11, 12, or 13 of the 13 conserved amino acids shown in Fig. 5 (positions 89, 90, 91, 95, 98, 101, 115, 422, 423, 426, 431, 436 and 452). The flagellin may be at least 30-99% identical to amino acids 1 174 and 418 505 of M84972. Fig. 26 of U.S. Patent Appl Publication No. 2009/0011982,

lists the percentage identity of the amino- and carboxy-terminus of flagellin with known TLR-5 stimulating activity, as compared to M84972.

[0050] The flagellin may be the major component of bacterial flagellum. The flagellin may be composed of three domains (Fig. 4). Domain 1 (D1) and domain 2 (D2) may be discontinuous and may be formed when residues in the amino terminus and carboxy terminus are juxtaposed by the formation of a hairpin structure. The amino and carboxy terminus comprising the D1 and D2 domains may be most conserved, whereas the middle hypervariable domain (D3) may be highly

variable. Studies with a recombinant protein containing the amino D1 and D2 and carboxyl D1 and D2 separated by an Escherichia coli hinge (ND1-2/ECH/CD2) indicate that D1 and D2 may be bioactive when coupled to an ECH element. This chimera, but not the hinge alone, may induce IkBa degradation, NF-kB activation, and NO and IL-8 production in two intestinal epithelial cell lines. The non-conserved D3 domain may be on the surface of the flagellar filament and may contain the major antigenic epitopes. The potent proinflammatory activity of flagellin may reside in the highly conserved N and C D1 and D2 regions (See Figure 4).

[0051] The flagellin may induce NF-kB activity by binding to Toll-like receptor 5 (TLR5). The TLR may recognize a conserved structure that is particular to the flagellin. The conserved structure may be composed of a large group of residues that are somewhat permissive to variation in amino acid content. Smith et al., Nat Immunol. 4:1247-53 (2003).

have identified 13 conserved amino acids in flagellin that are part of the conserved structure recognized by TLR5. The 13 conserved amino acids of flagellin that may be important for TLR5 activity are shown in Fig. 5.

[0052] Numerous deletional mutants of flagellin have been made that retain at least some TLR5 stimulating activity. The flagellin may be such a deletional mutant, and may be a deletional mutant disclosed in the Examples herein. The flagellin may comprise a sequence translated from GenBank Accession number D13689 missing amino acids 185-306 or 444-492, or from GenBank Accession number M84973 missing amino acids 179-415, or a variant thereof.

[0053] The flagellin may comprise transposon insertions and changes to the variable D3 domain. The D3 domain may be substituted in part, or in whole, with a hinge or linker polypeptide that allows the D1 and D2 domains to properly fold such that the variant stimulates TLR5 activity. The variant hinge elements may be found in the *E.coli* MukB protein and may have a sequence as set forth in SEQ ID NOS: 3 and 4, or a variant thereof.

[0054] The flagellin as described above may further comprise a leader sequence. The flagellin further comprising a leader sequence may be CBLB502S.

4. Agent

[0055] This invention also relates to an agent comprising a therapeutically effective amount of a TLR and TLR agonist. The agent may deliver the TLR separately from the TLR agonist. The agent may be a vector. The vector may comprise a first nucleic acid encoding the TLR and a second nucleic acid comprising the TLR agonist. The vector may be capable of transducing

mammalian cells. The vector may be capable of bi-cistronic expression of the TLR and/or TLR agonist using strong promoters. The vector may comprise only a gene encoding the TLR, which may be controlled by a strong promoter. The vector may be delivered into a mammalian cell by a virus or liposome related vector system. The virus vector system may be an adenovirus or a cytomegalovirus.

[0056] The agent may be a liposome harboring the vector. The liposome maybe capable of transducing mammalian cells and delivering the vector for expression.

[0057] The agent may be a drug formulation that simultaneously induces expression and activates the TLR, thereby exposing tumor or infected cells to the host immune system imitating the situation of a massive penetration through the intestinal wall. The agent may be a drug formulation that expresses the TLR in combination with the TLR agonist, and may be delivered systematically in solution for administration such as intramuscularly. The agent may be a drug formulation that expresses the TLR in combination with the TLR agonist, which may be expressed from the same vector, such as an adenoviral or cytomegalovirus vector system. The agent may be a drug formulation that expresses the TLR in combination with the TLR agonist expressed in the form of a nano-particle, which may carry a functional agonist to the cell surface of a mammalian cell.

[0058] The agent may be a pharmaceutical agent comprising the drug formulation described above, which may be produced using methods well known in the art. The agent may also comprise a coagent.

[0059] The vector may comprise a first nucleic acid encoding TLR5 and a second nucleic acid comprising flagellin. The vector may be capable of expressing TLR5 and/or flagellin using a strong promoter. The expression vector may further comprise a leader sequence cloned upstream of the gene encoding the TLR or TLR5 and/or flagellin. The expression vector may be pCD515 based vector system. The expression vector may be pCD515-CMV-hTLR5-EF1-502 as described in Figure 1A. The expression vector may be pCD515-CMV-hTLR5 as described in Figure 1B. The expression vector may be pCD515-CMV-Sseap-502 as described in Figure 1C.

[0060] The agent may be drug formulation that simultaneously induces expression and activates a TLR thereby exposing tumor or infected cells to the host immune system imitating the situation of a massive penetration through the intestinal wall. The drug formulation may be in the form of

a viral expression system harboring the vector. The drug formulation may be an adenovirus expression functional human TLR5 in combination with:

[0061] the TLR agonist, delivered systematically in solution for administration, such as intramuscularly;

[0062] the TLR agonist, expressed from the same adenoviral vector as the TLR; or

[0063] the TLR agonist, expressed in the form of nano-particles carrying functional TLR agonist, such as flagellin, which may be derived from CBLB502, on their surface. The nano-particle may be on the basis of a bacteriophage T7, or fully formed to retain its biological activity. The nano-formulation may provide for dose-dependent, NF-κB-responsive reporter activation, and may result in cell internalization by endocytosis for effective immunization approach (Mobian AP-A).

a. Administration

[0064] Administration of the agents using the method described herein may be orally, parenterally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal administration, or combinations thereof. Parenteral administration includes, but is not limited to, intravenous, intraarterial, intraperitoneal, subcutaneous, intramuscular, intrathecal, and intraarticular. For veterinary use, the agent may be administered as a suitably acceptable formulation in accordance with normal veterinary practice. The veterinarian can readily determine the dosing regimen and route of administration that is most appropriate for a particular animal. The agents may be administered to a human patient, cat, dog, large animal, or an avian.

[0065] The agent may be administered simultaneously or metronomically with other treatments. The term "simultaneous" or "simultaneously" as used herein, means that the agent and other treatment be administered within 48 hours, preferably 24 hours, more preferably 12 hours, yet more preferably 6 hours, and most preferably 3 hours or less, of each other. The term "metronomically" as used herein means the administration of the agent at times different from the other treatment and at a certain frequency relative to repeat administration.

[0066] The agent may be administered at any point prior to another treatment including about 120 hr, 118 hr, 116 hr, 114 hr, 112 hr, 110 hr, 108 hr, 106 hr, 104 hr, 102 hr, 100 hr, 98 hr, 96 hr, 94 hr, 92 hr, 90 hr, 88 hr, 86 hr, 84 hr, 82 hr, 80 hr, 78 hr, 76 hr, 74 hr, 72 hr, 70 hr, 68 hr, 66 hr, 64 hr, 62 hr, 60 hr, 58 hr, 56 hr, 54 hr, 52 hr, 50hr, 48 hr, 46 hr, 44 hr, 42 hr, 40 hr, 38 hr, 36 hr, 34 hr, 32 hr, 30 hr, 28 hr, 26 hr, 24 hr, 22 hr, 20 hr, 18 hr, 16 hr, 14 hr, 12 hr, 10 hr, 8 hr, 6 hr, 4

hr, 3 hr, 2 hr, 1 hr, 55 mins., 50 mins., 45 mins., 40 mins., 35 mins., 30 mins., 25 mins., 20 mins., 15 mins, 10 mins, 9 mins, 8 mins, 7 mins., 6 mins., 5 mins., 4 mins., 3 mins, 2 mins, and 1 mins. The agent may be administered at any point prior to a second treatment of the agent including about 120 hr, 118 hr, 116 hr, 114 hr, 112 hr, 110 hr, 108 hr, 106 hr, 104 hr, 102 hr, 100 hr, 98 hr, 96 hr, 94 hr, 92 hr, 90 hr, 88 hr, 86 hr, 84 hr, 82 hr, 80 hr, 78 hr, 76 hr, 74 hr, 72 hr, 70 hr, 68 hr, 66 hr, 64 hr, 62 hr, 60 hr, 58 hr, 56 hr, 54 hr, 52 hr, 50hr, 48 hr, 46 hr, 44 hr, 42 hr, 40 hr, 38 hr, 36 hr, 34 hr, 32 hr, 30 hr, 28 hr, 26 hr, 24 hr, 22 hr, 20 hr, 18 hr, 16 hr, 14 hr, 12 hr, 10 hr, 8 hr, 6 hr, 4 hr, 3 hr, 2 hr, 1 hr, 55 mins., 50 mins., 45 mins., 40 mins., 35 mins., 30 mins., 25 mins., 20 mins., 15 mins., 10 mins., 9 mins., 8 mins., 7 mins., 6 mins., 5 mins., 4 mins., 3 mins, 2 mins, and 1 mins.

[0067] The agent may be administered at any point after another treatment including about 1min, 2 mins., 3 mins., 4 mins., 5 mins., 6 mins., 7 mins., 8 mins., 9 mins., 10 mins., 15 mins., 20 mins., 25 mins., 30 mins., 35 mins., 40 mins., 45 mins., 50 mins., 55 mins., 1 hr, 2 hr, 3 hr, 4 hr, 6 hr, 8 hr, 10 hr, 12 hr, 14 hr, 16 hr, 18 hr, 20 hr, 22 hr, 24 hr, 26 hr, 28 hr, 30 hr, 32 hr, 34 hr, 36 hr, 38 hr, 40 hr, 42 hr, 44 hr, 46 hr, 48 hr, 50 hr, 52 hr, 54 hr, 56 hr, 58 hr, 60 hr, 62 hr, 64 hr, 66 hr, 68 hr, 70 hr, 72 hr, 74 hr, 76 hr, 78 hr, 80 hr, 82 hr, 84 hr, 86 hr, 88 hr, 90 hr, 92 hr, 94 hr, 96 hr, 98 hr, 100 hr, 102 hr, 104 hr, 106 hr, 108 hr, 110 hr, 112 hr, 114 hr, 116 hr, 118 hr, and 120 hr. The agent may be administered at any point prior after a second treatment of the agent including about 120 hr, 118 hr, 116 hr, 114 hr, 112 hr, 110 hr, 108 hr, 106 hr, 104 hr, 102 hr, 100 hr, 98 hr, 96 hr, 94 hr, 92 hr, 90 hr, 88 hr, 86 hr, 84 hr, 82 hr, 80 hr, 78 hr, 76 hr, 74 hr, 72 hr, 70 hr, 68 hr, 66 hr, 64 hr, 62 hr, 60 hr, 58 hr, 56 hr, 54 hr, 52 hr, 50hr, 48 hr, 46 hr, 44 hr, 42 hr, 40 hr, 38 hr, 36 hr, 34 hr, 32 hr, 30 hr, 28 hr, 26 hr, 24 hr, 22 hr, 20 hr, 18 hr, 16 hr, 14 hr, 12 hr, 10 hr, 8 hr, 6 hr, 4 hr, 3 hr, 2 hr, 1 hr, 55 mins., 50 mins., 45 mins., 40 mins., 35 mins., 30 mins., 25 mins., 20 mins., 15 mins., 10 mins., 9 mins., 8 mins., 7 mins., 6 mins., 5 mins., 4 mins., 3 mins, 2 mins, and 1 mins.

b. Formulation

[0068] The method may comprise administering the agent. Agents provided herein may be in the form of tablets or lozenges formulated in a conventional manner. For example, tablets and capsules for oral administration may contain conventional excipients may be binding agents, fillers, lubricants, disintegrants and wetting agents. Binding agents include, but are not limited to, syrup, accacia, gelatin, sorbitol, tragacanth, mucilage of starch and polyvinylpyrrolidone. Fillers

may be lactose, sugar, microcrystalline cellulose, maizestarch, calcium phosphate, and sorbitol. Lubricants include, but are not limited to, magnesium stearate, stearic acid, talc, polyethylene glycol, and silica. Disintegrants may be potato starch and sodium starch glycollate. Wetting agents may be sodium lauryl sulfate. Tablets may be coated according to methods well known in the art.

[0069] Agents provided herein may also be liquid formulations such as aqueous or oily suspensions, solutions, emulsions, syrups, and elixirs. The agents may also be formulated as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may contain additives such as suspending agents, emulsifying agents, nonaqueous vehicles and preservatives. Suspending agent may be sorbitol syrup, methyl cellulose, glucose/sugar syrup, gelatin, hydroxyethylcellulose, carboxymethyl cellulose, aluminum stearate gel, and hydrogenated edible fats. Emulsifying agents may be lecithin, sorbitan monooleate, and acacia. Nonaqueous vehicles may be edible oils, almond oil, fractionated coconut oil, oily esters, propylene glycol, and ethyl alcohol. Preservatives may be methyl or propyl p-hydroxybenzoate and sorbic acid.

[0070] Agents provided herein may also be formulated as suppositories, which may contain suppository bases such as cocoa butter or glycerides. Agents provided herein may also be formulated for inhalation, which may be in a form such as a solution, suspension, or emulsion that may be administered as a dry powder or in the form of an aerosol using a propellant, such as dichlorodifluoromethane or trichlorofluoromethane. Agents provided herein may also be formulated as transdermal formulations comprising aqueous or nonaqueous vehicles such as creams, ointments, lotions, pastes, medicated plaster, patch, or membrane.

[0071] Agents provided herein may also be formulated for parenteral administration such as by injection, intratumor injection or continuous infusion. Formulations for injection may be in the form of suspensions, solutions, or emulsions in oily or aqueous vehicles, and may contain formulation agents including, but not limited to, suspending, stabilizing, and dispersing agents. The agent may also be provided in a powder form for reconstitution with a suitable vehicle including, but not limited to, sterile, pyrogen-free water.

[0072] Agents provided herein may also be formulated as a depot preparation, which may be administered by implantation or by intramuscular injection. The agents may be formulated with

suitable polymeric or hydrophobic materials (as an emulsion in an acceptable oil, for example), ion exchange resins, or as sparingly soluble derivatives (as a sparingly soluble salt, for example).

c. Dosage

[0073] The method may comprise administering a therapeutically effective amount of the agent to a patient in need thereof. The therapeutically effective amount required for use in therapy varies with the nature of the condition being treated, the length of time desired to activate TLR activity, and the age/condition of the patient. In general, however, doses employed for adult human treatment typically are in the range of 0.001 mg/kg to about 200 mg/kg per day. The dose may be about 1 mg/kg to about 100 mg/kg per day. The desired dose may be conveniently administered in a single dose, or as multiple doses administered at appropriate intervals, for example as two, three, four or more sub-doses per day. Multiple doses may be desired, or required.

[0074] The dosage may be at any dosage such as about 0.1 mg/kg, 0.2 mg/kg, 0.3 mg/kg, 0.4 mg/kg, 0.5 mg/kg, 0.6 mg/kg, 0.7 mg/kg, 0.8 mg/kg, 0.9 mg/kg, 1 mg/kg, 25 mg/kg, 50 mg/kg, 75 mg/kg, 100 mg/kg, 125 mg/kg, 150 mg/kg, 175 mg/kg, 200 mg/kg, 225 mg/kg, 250 mg/kg, 275 mg/kg, 300 mg/kg, 325 mg/kg, 350 mg/kg, 375 mg/kg, 400 mg/kg, 425 mg/kg, 450 mg/kg, 475 mg/kg, 500 mg/kg, 525 mg/kg, 550 mg/kg, 575 mg/kg, 600 mg/kg, 625 mg/kg, 650 mg/kg, 675 mg/kg, 700 mg/kg, 725 mg/kg, 750 mg/kg, 775 mg/kg, 800 mg/kg, 825 mg/kg, 850 mg/kg, 875 mg/kg, 900 mg/kg, 925 mg/kg, 950 mg/kg, 975 mg/kg or 1 mg/kg.

5. Method for Treating Cancer

[0075] Provided herein is a method for treating cancer by administering to a mammal in need thereof the agent. The method provide immunotherapy against cancer by conversion of tumor cells into a TLR agonist-responsive state with targeted intratumor stimulation of TLR, thereby focusing an immune response on the tumor. The method may be be used to treat primary tumors prior to surgical removal in order to reduce the risk of metastasis development, as well as treat of other tumor nodules. The method may comprise intratumor injection. The method may have the step of injecting the agent into a primary tumor prior to surgical removal to reduce the risk of metastasis development, as well as treat other tumor nodules. The method may be used to treat any tumor that is accessible for adenovirus intratumor injection.

[0076] A variety of cancers may be treated according to this invention, including carcinoma, bladder (including accelerated and metastatic bladder cancer), breast, colon (including colorectal

cancer), kidney, liver, lung (including small and non-small cell lung cancer and lung adenocarcinoma), ovary, prostate, testes, genitourinary tract, lymphatic system, rectum, larynx, pancreas (including exocrine pancreatic carcinoma), esophagus, stomach, gall bladder, cervix, thyroid, and skin (including squamous cell carcinoma); hematopoietic tumors of lymphoid lineage including leukemia, acute lymphocytic leukemia, acute lymphoblastic leukemia, B-cell lymphoma, T-cell lymphoma, Hodgkins lymphoma, non-Hodgkins lymphoma, hairy cell lymphoma, histiocytic lymphoma, and Burketts lymphoma; hematopoietic tumors of myeloid lineage including acute and chronic myelogenous leukemias, myelodysplastic syndrome, myeloid leukemia, and promyelocytic leukemia; tumors of the central and peripheral nervous system including astrocytoma, neuroblastoma, glioma, and schwannomas; tumors of mesenchymal origin including fibrosarcoma, rhabdomyoscarcoma, and osteosarcoma; and other tumors including melanoma, xenoderma pigmentosum, keratoactanthoma, seminoma, thyroid follicular cancer, teratocarcinoma, and cancers of the gastrointestinal tract or the abdominopelvic cavity.

[0077] The method may be combined with other methods for treating cancer, including use of an immunostimulant, cytokine, or chemotherapeutic. The immunostimulant may be a growth hormone, prolactin or vitamin D.

6. Treatment of Infected Cells

[0078] Provided herein is a method for treating an infectious disease by the simultaneous delivery of transduced cells by the agent. The method may be used to treat a viral, bacterial, protozoan parasite or fungal infection. The method may be used to treat any infectious disease by using intracellular injection resulting in autocrine activation of TLR signaling of infected cells with minimal systemic effect and thereby enabling to attract innate immune response specific to the infected cells. The method may be combined with other therapies for treating viral, bacterial, protozoan parasite or fungi infections.

[0079] The method may comprise administering the agent. The method may comprise administration of a vaccine comprising the agent, and may be used in combination with any other vaccination, which may comprise a construct expressing an antigen of choice.

Example 1

Synthesis of Bi-cistronic expression TLR5/flagellin vector and Treatment of Tumor Cells

[0080] Vector constructs were created for expressing Toll-like receptor 5 (TLR-5) and flagellin CBLB502. Vector pCD515 was used as a backbone for these constructs. The cDNA sequence of human TLR-5 and the DNA encoding the toll-like receptor agonist's CBLB502 were individually fused with leader peptide derived from alkaline phosphatase enabling routing of the expressed protein through the endoplasmic reticulum (ER) and Golgi towards extracellular secretion.

[0081] The pCD515-CMV-hTLR5-EF1-502s vector construct expressed the secreted form of CBLB502 flagellin (CBLB502S) and the toll-like receptor 5 (TLR5) at the cell surfcace. This adenoviral vector required modification of the CBLB502 to reach its effective synthesis and secretion by mammalian cells. The adenovirus construct comprises the leader nucleic acid sequence (Atgctgctgctgctgctgctgctgggcctgaggctacagctctccctgggc) derived from alkaline phosphatase and was cloned upstream of the truncated Salmonella flagellin (fliC) gene (see Burdelya et al., Science 320:226-230 (2008) to encode a secretable form of flagellin (i.e., CBLB502S). An EF1 (elongation factor 1α) promoter was cloned upstream of this cassette encoding CBLB502S. The TLR5 gene was derived from human and has the amino acid sequence as shown in Figure 9. A CMV promoter was cloned upstream of the TLR5 gene. This construct co-expresses TLR5 and CBLB502S. This construct is shown in Figure 1A.

[0082] The pCD515-CMV-hTLR5 expression vector was constructed to express the form of human TLR-5 (see Figure 9). The adenovirus construct comprises a strong CMV promoter cloned upstream of the hTLR5 cassette. This construct is shown in Figure 1B.

[0083] The pCD515-CMV-Sseap-502 expression vector was constructed to express the secreted flagellin CBLB502 and the toll-like receptor 5 (TLR5). The adenovirus construct comprises a strong CMV promoter cloned upstream of the leader sequence SEAP 502 flagellin (fliC) gene. This construct is shown in FIG. 1C.

Example 2

Synthesis of Bi-cistronic expression TLR5/flagellin vector and Treatment of Tumor Cells

[0084] Two reporter mammalian cell lines, both expressing NF-kB-responsive GFP and differing in their TLR5 status, were transduced with vector constructs pCD515, pCD515-CMV-hTLR5-EF1-502s, pCD515-CMV-hTLR5-502, pCD515-CMV-hTLR5, and pCD515-CMV-Sseap-502 (see Table 3 below).

Table 3 Activity of adenoviral constructs as TLR5 signaling activators

Treatment	Report Line-293-null	Reporter Line-293-TLR
CBLB502	–	+
Ad5 (control) (pCD515)	–	–
Ad5 (TLR5) (pCD515-CMV-hTLR5)	–	–
Ad5(TLR5) + CBLB502 (pCD515-CMV-hTLR5-EF1-502)	+	+
Ad5 (CBLB502S) (pCD515-CMV-Sseap-502)	–	+
Ad5 (TLR5) (pCD515-CMV-hTLR5) + Ad5(CBLB502S) (pCD515-CMV-hTLR5-EF1-502s)	+	+
Ad5 (CBLB502S + TLR5) (pCD515-CMV-hTLR5-EF1-502s)	+	+

[0085] Vector co-expressing TLR5 and TLR5 agonist CBLB502S was sufficient to induce expression of NF-kB reporter in 293-null cells that do not express any of known TLRs and which cannot be activated by TLR5 agonist alone. This experiment demonstrates that TLR5 and flagellin CBLB502S can work in trans or in cis to activate TLR5 signaling.

Example 3

[0086] To test antitumor effects of bi-cistronic adenovirus having (pCD515-CMV-hTLR5-EF1-502s), 10 ml of the adenoviral suspension (1012-1011 IU/ml) were injected into one of two s.c. growing syngeneic tumors in Balb/c mice originating from CT26 mouse colon carcinoma cells when tumors reached 3-5mm in diameter and tumor size was monitored until control non-injected tumors reached size limit requiring termination of the experiment. Control mice were

injected (again, one tumor out of two per mouse) with adenoviral vector expressing red fluorescent protein (RFP). The results of a representative experiment are shown in Figure 3. Almost complete lack of growth of tumors injected with (pCD515-CMV-hTLR5-EF1-502s) was accompanied with reduced growth of the uninjected tumor within the same animal as compared with the tumors in control animals injected with RFP-expressing adenovirus. This result indicates (i) powerful in-cis and (ii) visible in-trans effect of pCD515-CMV-hTLR5-EF1-502s indicative of recruitment of both innate (cis effect) and adaptive (trans effect) immune response. Neither of the other control viruses listed in Table 1 (i.e., AD5 (control) and Ad5 (TLR5)) injected alone had growth suppressive effects on tumors.

[0087] Thus, enforced ectopic expression of TLR5 makes tumor cell types, which originally were TLR5 deficient, highly responsive to TLR5 stimulation resulting in breaking tumor immuno-tolerance, powerful attraction of innate immune response that promotes effective development of adaptive immune response with subsequent general antitumor effect.

CLAIMS

1. A vector comprising a first and second nucleic acid, wherein the first nucleic acid encodes a toll-like receptor and the second nucleic acid encodes an agonist for said toll-like receptor selected from the group consisting of heat shock proteins, high mobility group box 1, fibrinogen, fibronectin, RSV fusion protein, mouse mammary tumor virus envelope proteins, flagellin, and profilin.

2. The vector of claim 1, wherein the toll-like receptor agonist is flagellin.

3. The vector of claim 1, wherein the vector is an expression vector.

4. The vector of claim 3, wherein the vector is a mammalian expression vector.

5. The vector of claim 3, wherein the vector is expressed from an adenovirus, a lentivirus or a liposome.

6. The vector of any one of claims 1 to 5, wherein the first nucleic acid encodes a secreted form of a toll-like receptor.

7. The vector of claim any one of claims 1 to 6, wherein the flagellin is a secreted form of flagellin.

8. The vector of claim 7, wherein the secreted form of flagellin comprises a sequence of thirteen conserved amino acids of flagellin important for Toll-like receptor 5 (TLR5) activity, the sequence of thirteen conserved amino acids being selected from the group consisting of:

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3A(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....VD(X)_2L(X)_3A(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3S(X)_3M(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3N(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2L(X)_3G(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....IK(X)_2I(X)_3S(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....AD(X)_2L(X)_3K(X)_3M(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2L(X)_3K(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2L(X)_3D(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3G(X)_3F(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2I(X)_3G(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2L(X)_3S(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3Q(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2L(X)_3D(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2I(X)_3E(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....IS(X)_2L(X)_3A(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....LD(X)_2M(X)_3E(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....IN(X)_2I(X)_3T(X)_3L(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....VD(X)_2L(X)_3S(X)_3M(X)_5R;$

$LQR(X)_3L(X)_2Q(X)_2N(X)_{12}E.....ID(X)_2L(X)_3S(X)_3L(X)_5R;$ and

$LDT(X)_3K(X)_2Q(X)_2Q\ (X)_{12}E.....AD(X)_2I(X)_3I(X)_3I(X)_5Q;$

wherein $(X)_n$ represents n amino acids.

9. The vector of any one of claims 1 to 8, wherein the toll-like receptor is Toll-like receptor 5 (TLR-5).

10. The vector of any one of claims 1 to 9, wherein the first nucleic acid encodes the amino acid sequence of human Toll-like receptor 5 (TLR-5) and the second nucleic acid encodes the amino acid sequence of flagellin.

11. An agent for use in treating cancer in a mammal in need thereof, wherein the agent comprises the vector of any one of claims 1 to 10.

12. The agent for use of claim 11, wherein the cancer is a tumor.

13. The agent for use of claim 12, wherein the tumor is selected from the group consisting of prostate, breast, colon, esophagus, stomach, lung, pancreatic, renal, thyroid, ovaries, throat, and cervix.

14. The agent for use of claim 12, wherein the tumor is selected from the group consisting of sarcomas, melanomas, leukemias, and lymphomas.

15. The agent for use of any one of claims 12 to 14, wherein the agent is for administration in trans from the tumor of the mammal.

16. The agent for use of claim 11, wherein the agent is for administration in trans from the cancer of the mammal.

17. The agent for use of any one of claims 12 to 14, wherein the agent is for administration directly into the tumor of the mammal.

18. The agent for use of claim 11, wherein the agent is for administration directly into the cancer of the mammal.

19. The agent for use of any one of claims 11 to 18, wherein the agent is for administration in combination with an immunostimulant.

20. The agent for use of claim 19, wherein the immunostimulant is selected from the group consisting of growth hormone, prolactin and vitamin D.

21. The agent for use of claim 20, wherein the growth hormone is somatotrophin.

22. The agent for use of any one of claims 11 to 21, wherein the agent is for administration in combination with a cytokine.

23. The agent for use of claim 22, wherein the cytokine is stem cell factor.

FIGURE 1A

EF1 – promoter EF1 (elongation factor 1α).
502 - flagellin (fliC) gene

FIGURE 1B

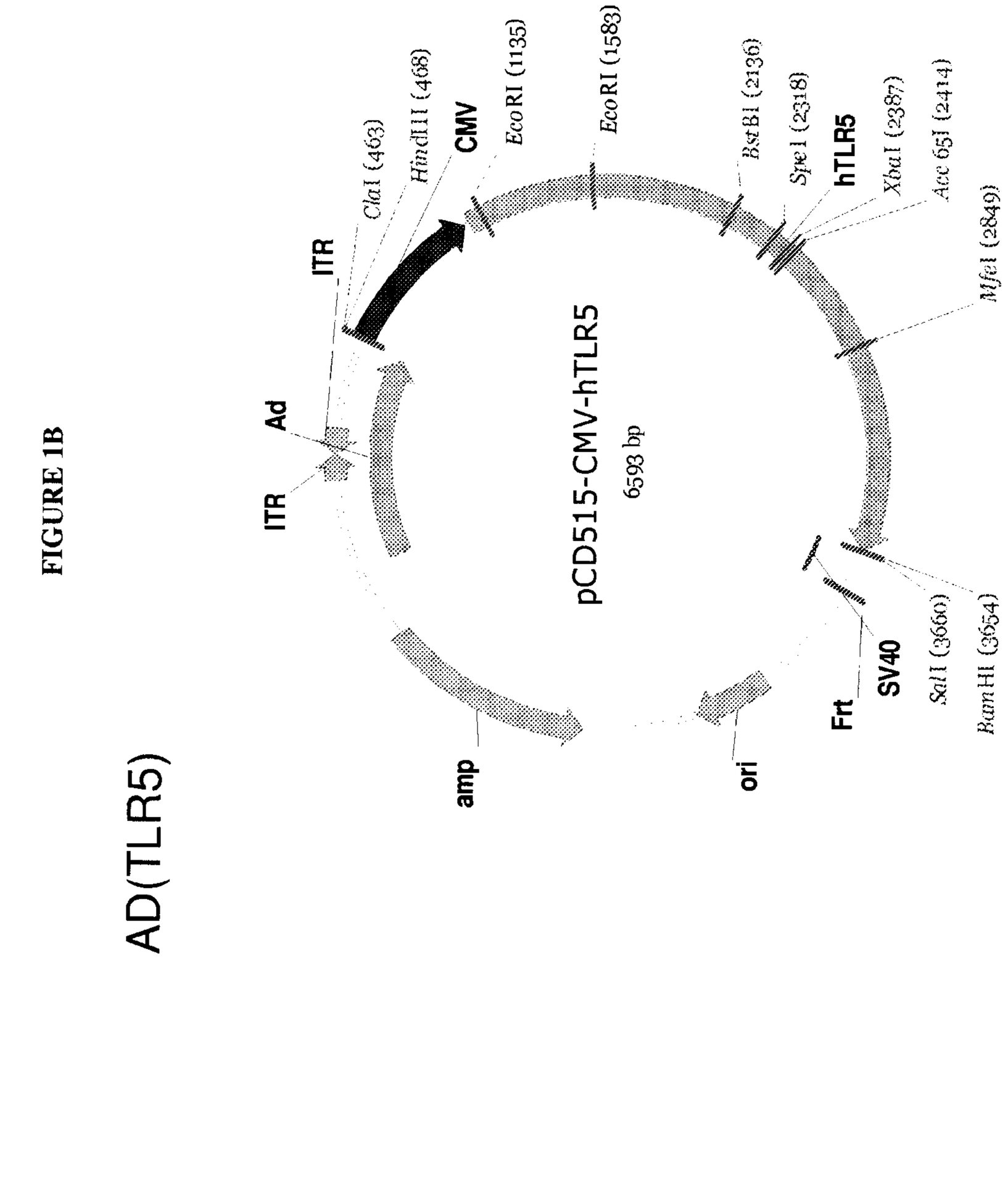
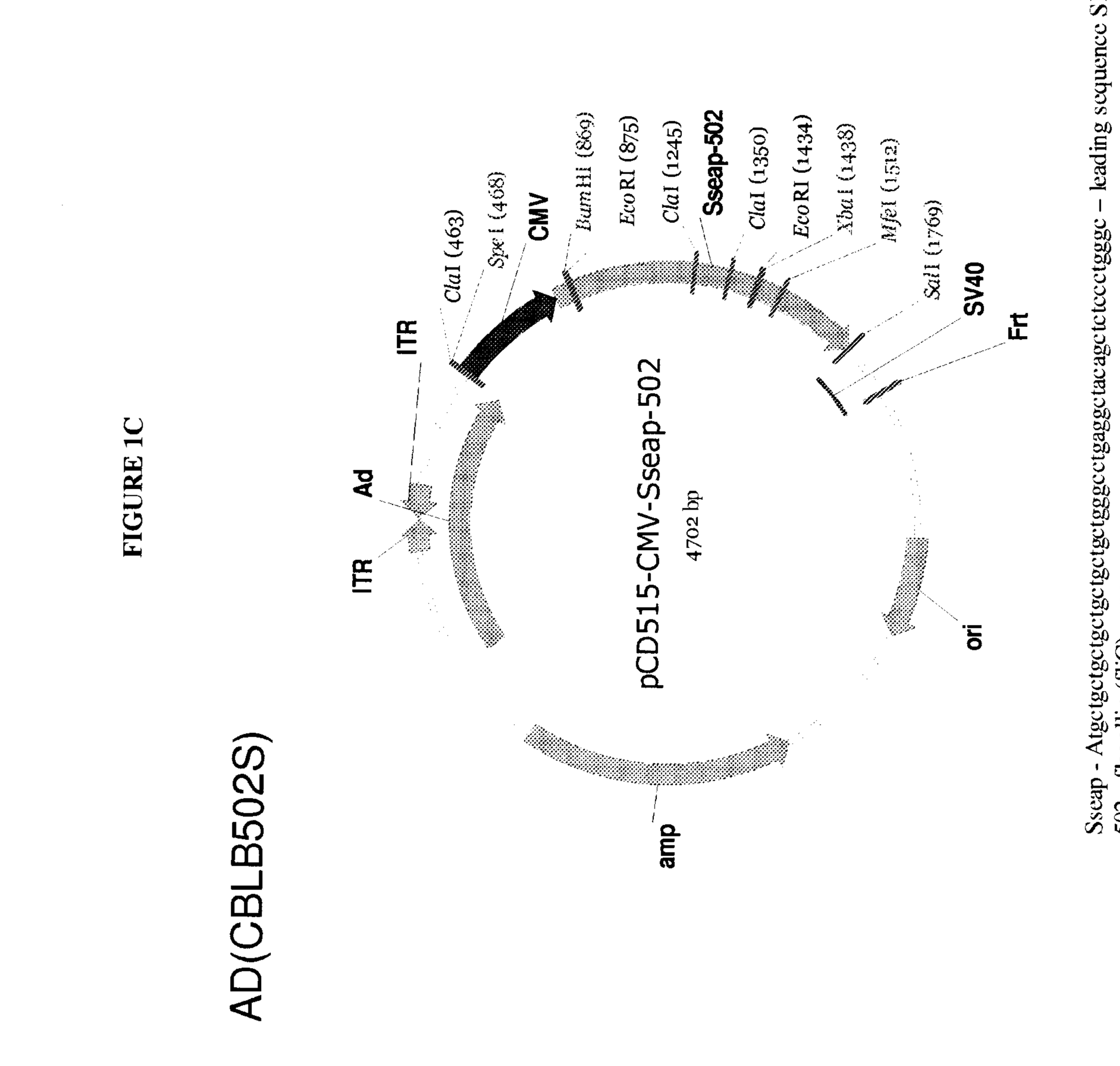

AD(TLR5)
pCD515-CMV-hTLR5
6593 bp
ITR
Ad
ITR
ClaI (463)
HindIII (468)
CMV
EcoRI (1135)
EcoRI (1583)
BstBI (2136)
SpeI (2318)
hTLR5
XbaI (2387)
Acc 65I (2414)
MfeI (2849)
amp
ori
Frt
SV40
SalI (3660)
BamHI (3654)

FIGURE 1C



Sseap - Atgctgctgctgctgctgctgctgggcctgaggctacagctctccctgggc – leading sequence
502 - flagellin (fliC) gene

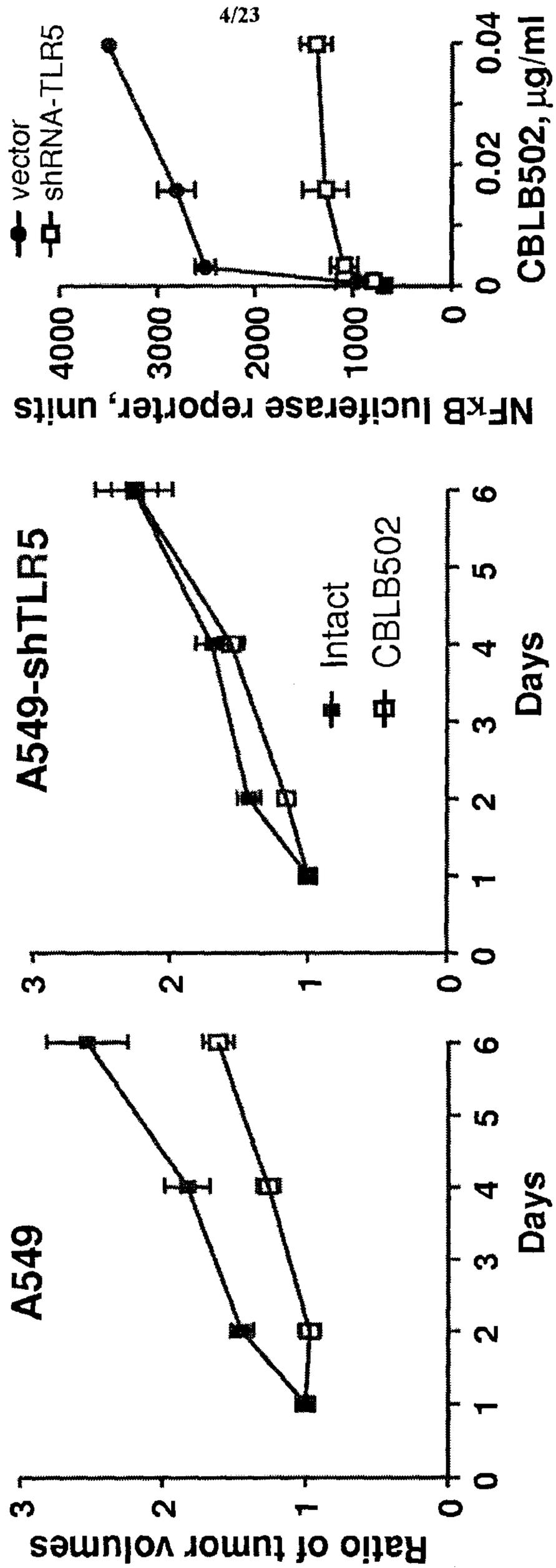
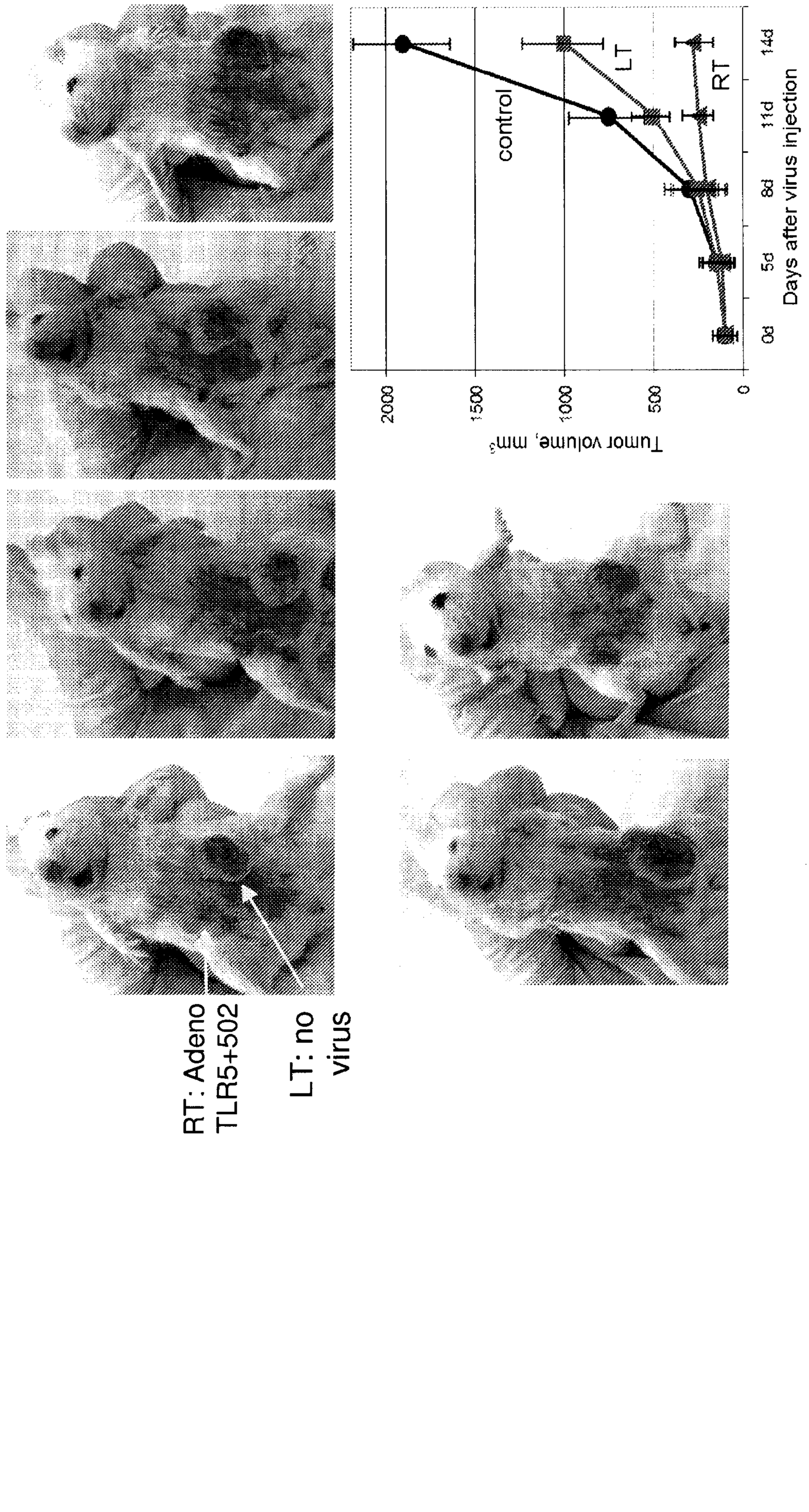

vector
shRNA-TLR5
CBLB502, μg/ml
0
0.02
0.04
NFκB luciferase reporter, units
4000
3000
2000
1000
0
A549-shTLR5
Intact
CBLB502
Days
0
1
2
3
4
5
6
A549
Days
0
1
2
3
Ratio of tumor volumes

FIGURE 2

FIGURE 3


Tumor volume, mm3
2000
1500
1000
500
0
0d
5d
8d
11d
14d
Days after virus injection
control
LT
RT
RT: Adeno
TLR5+502
LT: no
virus

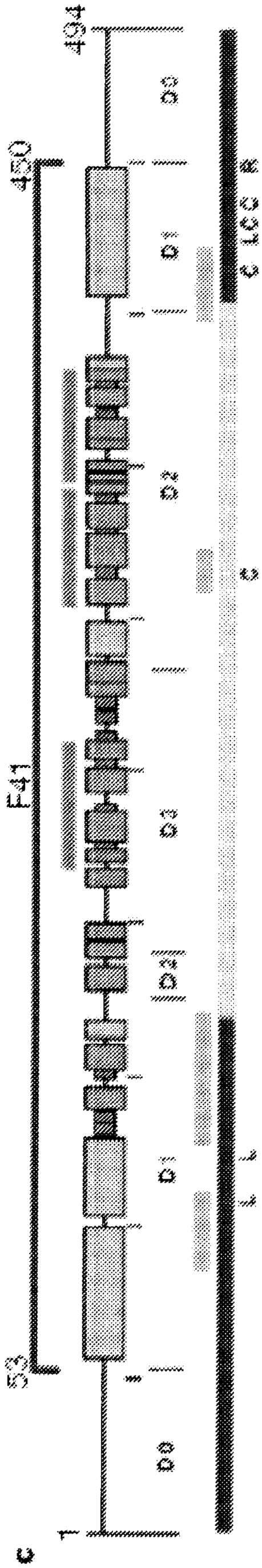


FIGURE 4

FIGURE 5

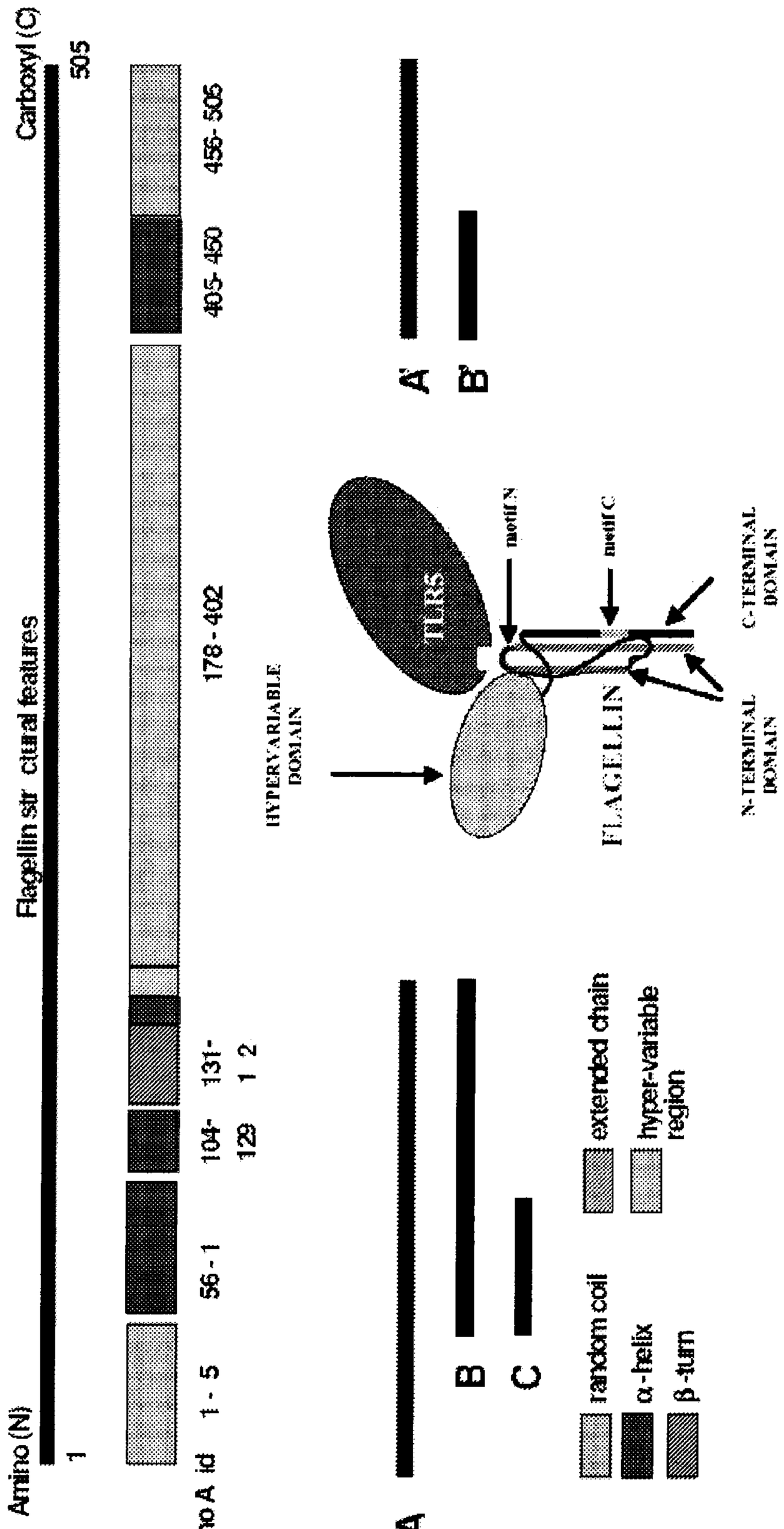

Amino (N)
Carboxyl (C)
505
Flagellin structural features
1
Amino A id
1 - 5
56 - 1
104- 129
131- 1 2
178 - 402
456- 505
405- 450
A
B
C
random coil
α-helix
β-turn
extended chain
hyper-variable region
HYPERVARIABLE DOMAIN
HLR5
motifN
motifC
FLAGELLIN
N-TERMINAL DOMAIN
C-TERMINAL DOMAIN

FIGURE 6

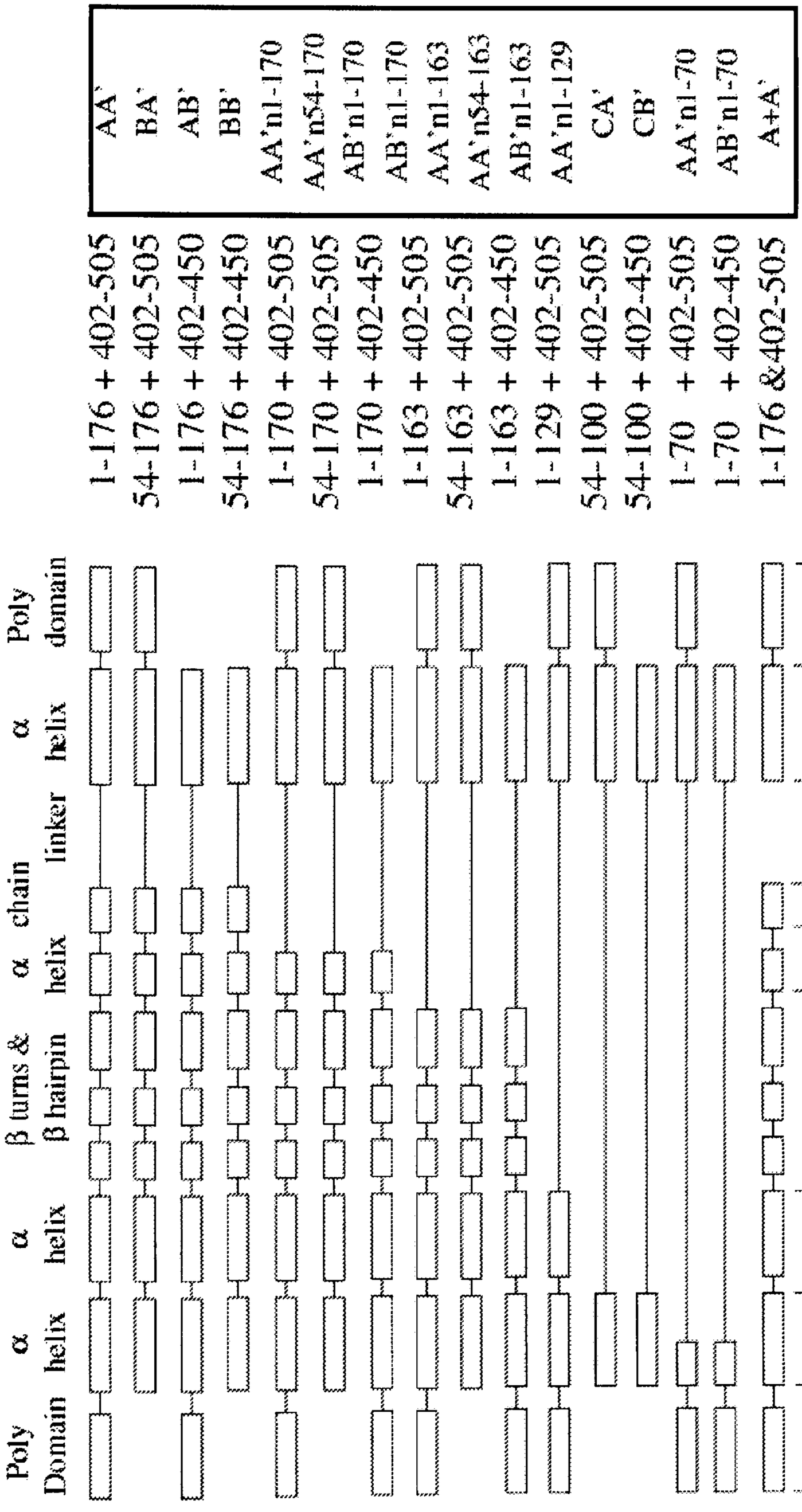

AA'
BA'
AB'
BB'
AA'n1-170
AA'n54-170
AB'n1-170
AB'n1-170
AA'n1-163
AA'n54-163
AB'n1-163
AA'n1-129
CA'
CB'
AA'n1-70
AB'n1-70
A+A'
1-176 + 402-505
54-176 + 402-505
1-176 + 402-450
54-176 + 402-450
1-170 + 402-505
54-170 + 402-505
1-170 + 402-450
1-163 + 402-505
54-163 + 402-505
1-163 + 402-450
1-129 + 402-505
54-100 + 402-505
54-100 + 402-450
1-70 + 402-505
1-70 + 402-450
1-176 &402-505
Poly α domain
α helix linker
α helix
β turns & β hairpin
α chain linker
Poly α Domain helix
402
450
505
163
170
176
129
100
54
1

FIGURE 7

AA'

Nucleotide sequence (990 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATTGATGTGAAAAGC
CTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATT**
CTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCT
AACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTG
GGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAAT
CTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATG
TCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTT
CCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (329 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNV
N**SPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDAVRSSL
GAIQNRFDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTS
VLAQANQVPQNVLSLLR

AB'

Nucleotide sequence (825 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATTGATGTGAAAAGC
CTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATT**
CTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCT
AACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTG
GGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTTAG

Protein sequence (274 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNV
N**SPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDAVRSSL
GAIQNRFDSAITNL

FIGURE 7, CONTINUED

BA'

Nucleotide sequence (831 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAG
GACAACCAGATGAAAATCCAGGTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTG
CAAAAAATTGATGTGAAAAGCCTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGAATT**
TCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCC
GCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTG
GACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAAC
CTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTAT
GCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTT
CTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (276 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNVN
SPGISGGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLG
AIQNRFDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSV
LAQANQVPQNVLSLLR

BB'

Nucleotide sequence (666 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAG
GACAACCAGATGAAAATCCAGGTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTG
CAAAAAATTGATGTGAAAAGCCTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGAATT**
TCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCC
GCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTG
GACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAAC
CTTTAG

Protein sequence (221 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNVN
SPGISGGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLG
AIQNRFDSAITNL

CA'

Nucleotide sequence (603 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAAT
GAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCA
TTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCA

FIGURE 7, CONTINUED

GCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGCCGTATCGAA
GATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTGCAGCAGGCT
GGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCTTTACTGCGT
TAG

Protein sequence (200 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQAT**SPGISGGGGGILDSMG**TLINEDAA
AAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITNLGNTVTNLNSAR
SRIEDADYATEVSNMSKAQILQQAGTSVLAQANQVPQNVLSLLR

CB'

Nucleotide sequence (438 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAAT
GAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCA
TTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCA
GCCATTACCAACCTTTAG

Protein sequence (145 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQAT**SPGISGGGGGILDSMG**TLINEDAA
AAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITNL

A

Nucleotide sequence (639 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
ACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGTCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTGTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATTGATGTGAAAAGC
CTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGA**TGA

Protein sequence (212 AA), last three amino acids are derived from primer and pRSETb polylinker:

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNV
N**SPG**

B

Nucleotide sequence (480 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT

FIGURE 7, CONTINUED

AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAGACTCAATTTAACGGTGTTAAAGTCCTGTCTCAG
GACAACCAGATGAAAATCCAGGTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTG
CAAAAAATTGATGTGAAAAGCCTTGGCCTTGATGGGTTCAATGTTAAT**TCCCCGGGA**TGA

Protein sequence (159 AA), last three amino acids are derived from primer and pRSETb polylinker:

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNVN
SPG

C

Nucleotide sequence (252 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
TCCCCGGGATGA

Protein sequence (83 AA), last three amino acids are derived from primer and pRSETb polylinker:

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQAT**SPG**

GST-A'

Nucleotide sequence (1038 bp), GST highlighted:

ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
CTGGTTCCGCGTGGA**TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATG**
GGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCT
TCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAA
AACCGTTTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCG
CGTAGCCGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAG
ATTCTGCAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTC
CTCTCTTTACTGCGTTAG

Protein sequence (345 AA):

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEF
PNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRY
GVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLY
DALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA
TFGGGDHPPKSDLVPRG**SPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLA
SIDSALSKVDAVRSSLGAIQNRFDSAITNLGNTVTNLNSARSRIEDADYATE
VSNMSKAQILQQAGTSVLAQANQVPQNVLSLLR

FIGURE 7, CONTINUED

GST-B'

Nucleotide sequence (873 bp), GST highlighted:

ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
CTGGTTCCGCGTGGA**TCCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATG**
GGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCT
TCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAA
AACCGTTTTGATTCAGCCATTACCAACCTTTAG

Protein sequence (290 AA):

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEF
PNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRY
GVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLY
DALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA
TFGGGDHPPKSDLVPRG**SPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLA
SIDSALSKVDAVRSSLGAIQNRFDSAITNL

AA'n1-170

Nucleotide sequence (972 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATTGATGTGAAAAGC
CTTGGCCTT**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACA
TTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATT
GATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGT
TTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGC
CGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTG
CAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCT
TTACTGCGTTAG

Protein sequence (323 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGL**IPGIS**
GGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNR
FDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSVLAQAN
QVPQNVLSLLR

FIGURE 7, CONTINUED

AA'n1-163

Nucleotide sequence (951 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATT**ATCCCGGGAATT**
TCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCC
GCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTG
GACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAAC
CTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTAT
GCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTT
CTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (316 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKI**IPGISGGGGGIL**
DSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITN
LGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSVLAQANQVPQNVL
SLLR

AA'n54-170

Nucleotide sequence (813 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAG
GACAACCAGATGAAAATCCAGGTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTG
CAAAAAATTGATGTGAAAAGCCTTGGCCTT**ATCCCGGGAATTTCCGGTGGTGGTGGTGGA**
ATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACC
GCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCT
CTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACC
AATCTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAAT
ATGTCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAG
GTTCCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (270 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGL**IPGISG**
GGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRF
DSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSVLAQANQ
VPQNVLSLLR

AA'n54-163

FIGURE 7, CONTINUED

Nucleotide sequence (792 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAG
GACAACCAGATGAAAATCCAGGTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTG
CAAAAAATT**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACA
TTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATT
GATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGT
TTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGC
CGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTG
CAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCT
TTACTGCGTTAG

Protein sequence (263 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKI**IPGISGGGGGILD**
SMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITNL
GNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSVLAQANQVPQNVLS
LLR

AB'n1-170 (or AA'n1-170c402-450)

Nucleotide sequence (807 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATTGATGTGAAAAGC
CTTGGCCTT**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACA
TTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATT
GATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGT
TTTGATTCAGCCATTACCAACCTTTAG

Protein sequence (268 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKIDVKSLGL**IPGIS**
GGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNR
FDSAITNL

AB'n1-163 (or AA'n1-163c402-450)

Nucleotide sequence (786 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC

FIGURE 7, CONTINUED

GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAGACTCAATTTAACGGTGTTAAAGTCCTCTCTCAGGACAACCAGATGAAAATCCAG
GTTGGTGCTAACGATGGTGAAACCATTACCATCGATCTGCAAAAAATT**ATCCCGGGAATT**
TCCGGTGGTGGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCC
GCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTG
GACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAAC
CTTTAG

Protein sequence (261 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQTQFNGVKVLSQDNQMKIQVGANDGETITIDLQKI**IPGISGGGGGIL**
DSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITN
L

AA'n1-129

Nucleotide sequence (849 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAG**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACATTA
ATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGAT
TCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTT
GATTCAGCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGCCGT
ATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTGCAG
CAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCTTTA
CTGCGTTAG

Protein sequence (282 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQ**IPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDA
VRSSLGAIQNRFDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQ
QAGTSVLAQANQVPQNVLSLLR

AA'n54-129

Nucleotide sequence (690 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAG**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATT**
CTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCT
AACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTG
GGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAAT
CTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATG

FIGURE 7, CONTINUED

TCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTT
CCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (229 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQ**IPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDAV
RSSLGAIQNRFDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQ
AGTSVLAQANQVPQNVLSLLR

AB'n1-129

Nucleotide sequence (684 bp):

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACTAACGGGACTAACTCTGATTCC
GATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTGGAAGAAATCGATCGCGTTTCT
AATCAG**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACATTA
ATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATTGAT
TCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGTTTT
GATTCAGCCATTACCAACCTTTAG

Protein sequence (227 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEI
DRVSNQ**IPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDA
VRSSLGAIQNRFDSAITNL

AB'n54-129

Nucleotide sequence (525 bp):

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCG*TTCACTTCTAATATCAAAGGC
CTGACTCAGGCTTCCCGTAACGCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGT
GCGCTGAATGAAATCAACAACAACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT
AACGGGACTAACTCTGATTCCGATCTGAAATCTATCCAGGATGAAATTCAGCAACGTCTG
GAAGAAATCGATCGCGTTTCTAATCAG**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATT**
CTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCT
AACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTG
GGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTTAG

Protein sequence (174 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDP*FTSNIKGLTQASRNANDGI
SIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEID
RVSNQ**IPGISGGGGGILDSMG**TLINEDAAAAKKSTANPLASIDSALSKVDAV
RSSLGAIQNRFDSAITNL

AA'n1-100

Nucleotide sequence (762 bp):

ATGCGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC

FIGURE 7, CONTINUED

GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT**ATCCCGGGAATTTCCGGTGGT**
GGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAG
AAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTT
CGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTGGCAAT
ACGGTAACCAATCTGAACTCCGCGCGTAGCCGTATCGAAGATGCTGACTATGCAACGGAA
GTTTCTAATATGTCTAAAGCGCAGATTCTGCAGCAGGCTGGTACTTCCGTTCTGGCGCAG
GCTAACCAGGTTCCGCAAAACGTCCTCTCTTTACTGCGTTAG

Protein sequence (253 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQAT**IPGISGGGGGILDSMG**TLINEDA
AAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITNLGNTVTNLNSA
RSRIEDADYATEVSNMSKAQILQQAGTSVLAQANQVPQNVLSLLR

AB'n1-100

Nucleotide sequence (597 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGACGGCATTTCTATTGCGCAGACCACTGAAGGTGCGCTGAATGAAATCAACAAC
AACCTGCAGCGTGTGCGTGAGTTGTCTGTTCAGGCCACT**ATCCCGGGAATTTCCGGTGGT**
GGTGGTGGAATTCTAGACTCCATGGGTACATTAATCAATGAAGACGCTGCCGCAGCCAAG
AAAAGTACCGCTAACCCACTGGCTTCAATTGATTCTGCATTGTCAAAAGTGGACGCAGTT
CGTTCTTCTCTGGGGGCAATTCAAAACCGTTTTGATTCAGCCATTACCAACCTTTAG

Protein sequence (198 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDG
ISIAQTTEGALNEINNNLQRVRELSVQAT**IPGISGGGGGILDSMG**TLINEDA
AAAKKSTANPLASIDSALSKVDAVRSSLGAIQNRFDSAITNL

AA'n1-70

Nucleotide sequence (672 bp):

ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGAC**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACA
TTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATT
GATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGT
TTTGATTCAGCCATTACCAACCTTGGCAATACGGTAACCAATCTGAACTCCGCGCGTAGC
CGTATCGAAGATGCTGACTATGCAACGGAAGTTTCTAATATGTCTAAAGCGCAGATTCTG
CAGCAGGCTGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTTCCGCAAAACGTCCTCTCT
TTACTGCGTTAG

Protein sequence (223 AA):

*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNAND**I**
PGISGGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGA

FIGURE 7, CONTINUED

IQNRFDSAITNLGNTVTNLNSARSRIEDADYATEVSNMSKAQILQQAGTSVL
AQANQVPQNVLSLLR

AB'n1-70
Nucleotide sequence (507 bp):
ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
*ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCCGATG*GCACAAGTCATTAATACA
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCCTCACTGAGTTCC
GCTATTGAGCGTCTGTCCTCTGGTCTGCGTATCAACAGCGCGAAAGACGATGCGGCAGGC
CAGGCGATTGCTAACCGCTTCACTTCTAATATCAAAGGCCTGACTCAGGCTTCCCGTAAC
GCTAACGAC**ATCCCGGGAATTTCCGGTGGTGGTGGTGGAATTCTAGACTCCATGG**GTACA
TTAATCAATGAAGACGCTGCCGCAGCCAAGAAAAGTACCGCTAACCCACTGGCTTCAATT
GATTCTGCATTGTCAAAAGTGGACGCAGTTCGTTCTTCTCTGGGGGCAATTCAAAACCGT
TTTGATTCAGCCATTACCAACCTTTAG
Protein sequence (168 ΛΛ):
*MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPM*AQVINTNSLSLLTQNNLN
KSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNANDI
PGISGGGGGILDSMGTLINEDAAAAKKSTANPLASIDSALSKVDAVRSSLGA
IQNRFDSAITNL

FIGURE 8A

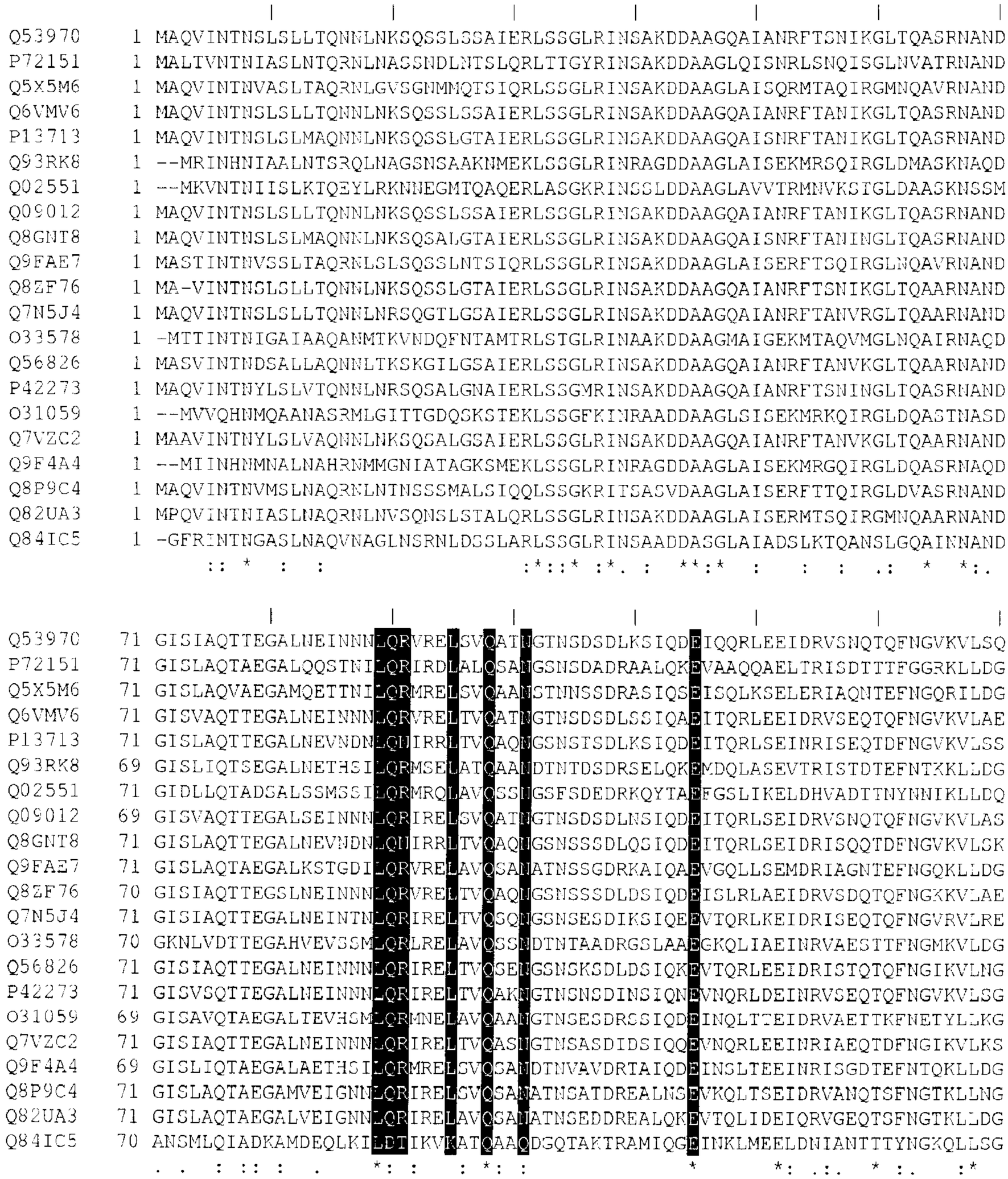

Q53970 1 MAQVINTNSLSLLTQNNLNKSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQASRNAND
P72151 1 MALIVNTNIASLNTQRNLNASSNDLNTSLQRLTTGYRINSAKDDAAGLQISNRLSNQISGLNVATRNAND
Q5X5M6 1 MAQVINTNVASLTAQRNLGVSGNMMQTSIQRLSSGLRINSAKDDAAGLAISQRMTAQIRGMNQAVRNAND
Q6VMV6 1 MAQVINTNSLSLLTQNNLNKSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTANIKGLTQASRNAND
P13713 1 MAQVINTNSLSLMAQNNLNKSQSSLGTAIERLSSGLRINSAKDDAAGQAISNRFTANIKGLTQASRNAND
Q93RK8 1 --MRINHNIAALNTSRQLNAGSNSAAKNMEKLSSGLRINRAGDDAAGLAISEKMRSQIRGLDMASKNAQD
Q02551 1 --MKVNTNIISLKTQEYLRKNNEGMTQAQERLASGKRINSSLDDAAGLAVVTRMNVKSIGLDAASKNSSM
Q09012 1 MAQVINTNSLSLLTQNNLNKSQSSLSSAIERLSSGLRINSAKDDAAGQAIANRFTANIKGLTQASRNAND
Q8GNT8 1 MAQVINTNSLSLMAQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAISNRFTANINGLTQASRNAND
Q9FAE7 1 MASTINTNVSSLIAQRNLSLSQSSLNTSIQRLSSGLRINSAKDDAAGLAISERFTSQIRGLNQAVRNAND
Q8ZF76 1 MA-VINTNSLSLLTQNNLNKSQSSLGTAIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNAND
Q7N5J4 1 MAQVINTNSLSLLTQNNLNRSQGTLGSAIERLSSGLRINSAKDDAAGQAIANRFTANVRGLTQAARNAND
O33578 1 -MTTINTNIGAIAAQANMTKVNDQFNTAMTRLSTGLRINAAKDDAAGMAIGEKMTAQVMGLNQAIRNAQD
Q56826 1 MASVINTNDSALLAQNNLTKSKGILGSAIERLSSGLRINSAKDDAAGQAIANRFTANVKGLTQAARNAND
P42273 1 MAQVINTNYLSLVTQNNLNRSQSALGNAIERLSSGMRINSAKDDAAGQAIANRFTSNINGLTQASRNAND
O31059 1 --MVVQHNMQAANASRMLGITTGDQSKSTEKLSSGFKINRAADDAAGLSISEKMRKQIRGLDQASTNASD
Q7VZC2 1 MAAVINTNYLSLVAQNNLNKSQSALGSAIERLSSGLRINSAKDDAAGQAIANRFTANVKGLTQAARNAND
Q9F4A4 1 --MIINHNMNALNAHRNMMGNIATAGKSMEKLSSGLRINRAGDDAAGLAISEKMRGQIRGLDQASRNAQD
Q8P9C4 1 MAQVINTNVMSLNAQRNLNTNSSSMALSIQQLSSGKRITSASVDAAGLAISERFTTQIRGLDVASRNAND
Q82UA3 1 MPQVINTNIASLNAQRNLNVSQNSLSTALQRLSSGLRINSAKDDAAGLAISERMTSQIRGMNQAARNAND
Q84IC5 1 -GFRINTNGASLNAQVNAGLNSRNLDSSLARLSSGLRINSAADDASGLAIADSLKTQANSLGQAINNAND
:: * : : :*::* :*. : **:* : : : .: * *:.
Q53970 71 GISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDSDLKSIQDEIQQRLEEIDRVSNQTQFNGVKVLSQ
P72151 71 GISLAQTAEGALQQSTNILQRIRDLALQSANGSNSDADRAALQKEVAAQQAELTRISDTTTFGGRKLLDG
Q5X5M6 71 GISLAQVAEGAMQETTNILQRMRELSVQAANSTNNSSDRASIQSEISQLKSELERIAQNTEFNGQRILDG
Q6VMV6 71 GISVAQTTEGALNEINNNLQRVRELTVQATNGTNSDSDLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAE
P13713 71 GISLAQTTEGALNEVNDNLQNIRRLTVQAQNGSNSISDLKSIQDEITQRLSEINRISEQTDFNGVKVLSS
Q93RK8 69 GISLIQTSEGALNETHSILQRMSELATQAANDTNTDSDRSELQKEMDQLASEVTRISTDTEFNTKKLLDG
Q02551 71 GIDLLQTADSALSSMSSILQRMRQLAVQSSNGSFSDEDRKQYTAEFGSLIKELDHVADTTNYNNIKLLDQ
Q09012 69 GISVAQTTEGALSEINNNLQRIRELSVQATNGTNSDSDLNSIQDEITQRLSEIDRVSNQTQFNGVKVLAS
Q8GNT8 71 GISLAQTTEGALNEVNDNLQNIRRLTVQAQNGSNSSSDLQSIQDEITQRLSEIDRISQQTDFNGVKVLSK
Q9FAE7 71 GISLAQTAEGALKSTGDILQRVRELAVQSANATNSSGDRKAIQAEVGQLLSEMDRIAGNTEFNGQKLLDG
Q8ZF76 70 GISIAQTTEGSLNEINNNLQRVRELTVQAQNGSNSSSDLDSIQDEISLRLAEIDRVSDQTQFNGKKVLAE
Q7N5J4 71 GISIAQTTEGALNEINTNLQRIRELTVQSQNGSNSESDIKSIQEEVTQRLKEIDRISEQTQFNGVRVLRE
O33578 70 GKNLVDTTEGAHVEVSSMLQRLRELAVQSSNDTNTAADRGSLAAEGKQLIAEINRVAESTTFNGMKVLDG
Q56826 71 GISIAQTTEGALNEINNNLQRIRELTVQSENGSNSKSDLDSIQKEVTQRLEEIDRISTQTQFNGIKVLNG
P42273 71 GISVSQTTEGALNEINNNLQRIRELTVQAKNGTNSNSDINSIQNEVNQRLDEINRVSEQTQFNGVKVLSG
O31059 69 GISAVQTAEGALTEVHSMLQRMNELAVQAANGTNSESDRSSIQDEINQLTTEIDRVAETTKFNETYLLKG
Q7VZC2 71 GISIAQTTEGALNEINNNLQRIRELTVQASNGTNSASDIDSIQQEVNQRLEEINRIAEQTDFNGIKVLKS
Q9F4A4 69 GISLIQTAEGALAETHSILQRMRELSVQSANDTNVAVDRTAIQDEINSLTEEINRISGDTEFNTQKLLDG
Q8P9C4 71 GISLAQTAEGAMVEIGNNLQRIRELSVQSANATNSATDREALNSEVKQLTSEIDRVANQTSFNGTKLLNG
Q82UA3 71 GISLAQTAEGALVEIGNNLQRIRELAVQSANATNSEDDREALQKEVTQLIDEIQRVGEQTSFNGTKLLDG
Q84IC5 70 ANSMLQIADKAMDEQLKILDTIKVKATQAAQDGQTAKTRAMIQGEINKLMEELDNIANTTTYNGKQLLSG
. . : :: : . *: : : *: : * *: .:. * :. :*

FIGURE 8A, CONTINUED

```
                          |                          |                          |
Q53970   141 DNQ-MK--IQVGANDG---------------ETITIDLQ----------KID-VKSLG----LDGFN
P72151   141 SFGTTS--FQVGSNAY---------------ETIDISLQNASASAIGSYQVG-SNGAGTVASVAGTA
Q5X5M6   141 SFSGAS--FQVGANSN---------------QTINFSIG----------SIK-ASSIGGIATATGTE
Q6VMV6   141 NNE-MK--IQVGANDG---------------ETITINLA----------KID-AKTLG----LDGFN
P13713   141 DQK-LT--IQVGANDG---------------ETTDIDLK----------KID-AKQLG----MDTF-
Q93RK8   139 TAQNLT--FQIGANEG---------------QTMSLSIN----------KMD-SE---------SLK
Q02551   139 TATGAATQVSIQASDKAN-------------DLINIDLFNAKGLSAGTITLGSGSTVAGYSALSVAD
Q09012   141 DQT-MK--IQVGANDG---------------ETIEIALD----------KID-AKTLG----LDNFS
Q8GNT8   141 DQK-LT--IQVGANDG---------------ETIDIDLK----------NIN-AQSLG----LDKFN
Q9FAE7   141 SFGSAT--FQVGANAN---------------QTITATTGNFRTNNY-GAQLT-ASASG--AATSGAS
Q8ZF76   140 NTT-MS--IQVGANDG---------------ETIDINLQ----------KID-SKSLG----LGSYS
Q7N5J4   141 DSK-MT--IQVGANDN---------------EVIDIDLK----------KID-KEALN----LGKFT
O33578   140 SFTGKQ--LQIGADSG---------------QTMAINVDSAAATDIGAHKISSASTVVADAALTDTT
Q56826   141 DVTEMK--IQVGANDN---------------ETIGIKLG----------KIN-SEKLN----LKEFS
P42273   141 EKSKMT--IQVGTNDN---------------EVIEFNLD----------KID-NDTLG----VASDK
O31059   139 GNGDRT--VRVYAHDAGLVGSLSQNTTKATFQMRKLEIGDSYTIGGTTYKIG-AETVK--EAMTALK
Q7VZC2   141 NATDMTLSIQVGAKDN---------------ETIDIKID----------RNS-NWNLY----DAVGT
Q9F4A4   139 GFKG-E--FQIGANSN---------------QTVKLDIG----------NMS-AA---------SLG
Q8P9C4   141 DFSGAL--FQVGADAG---------------QTIGINS---------IVDAN-VDSLG--KANFAAS
Q82UA3   141 SFASQI--FQVGANEG---------------ETIDFTD----------------------------
Q84IC5   140 SFSNAQ--FQIGDKAN---------------QTVNATIG----------STN-SAKVGQTRFETGAV
                      . :                      :
```

FIGURE 8B

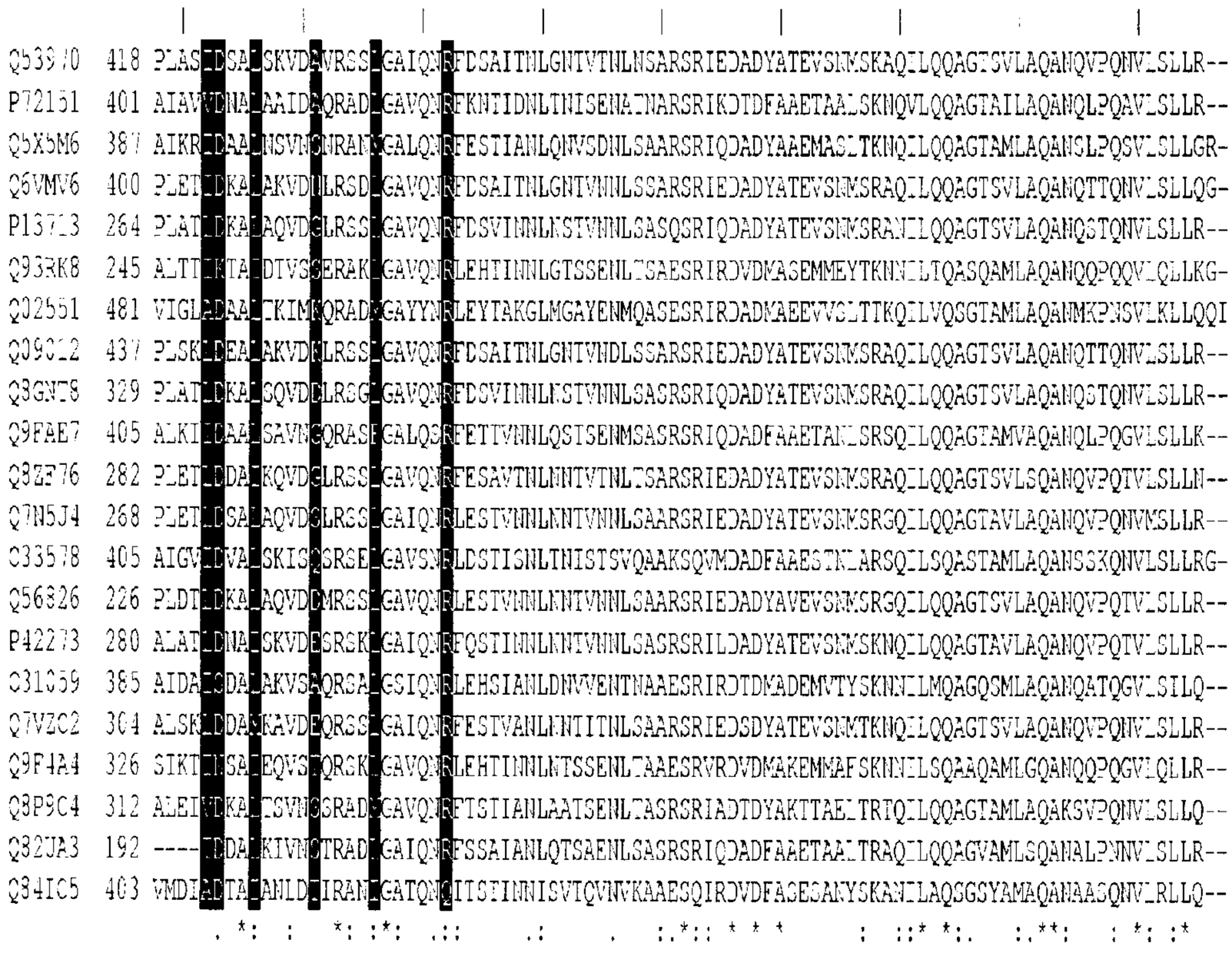
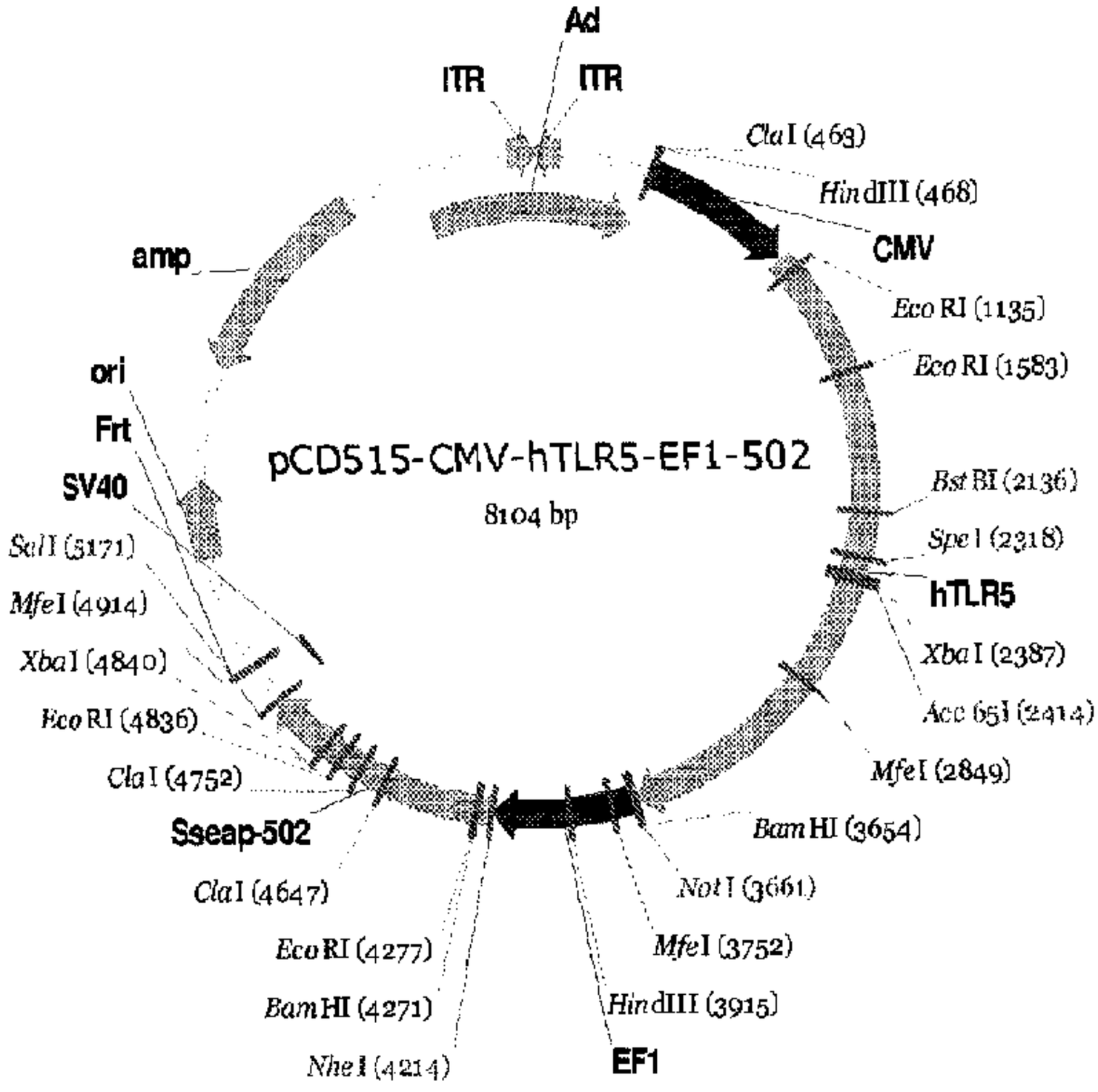


FIGURE 9

Homo Sapiens TLR5

```
   1 mgdhldlllg vvlmagpvfg ipscsfdgri afyrfcnltq vpqvlntter lllsfnyirt
  61 vtassfpfle qlqllelgsq ytpltidkea frnlpnlril dlgsskiyfl hpdafqglfh
 121 lfelrlyfcg lsdavlkdgy frnlkaltrl dlsknqirsl ylhpsfgkln slksidfssn
 181 qiflvcehel eplqgktlsf fslaanslys rvsvdwgkcm npfrnmvlei ldvsgngwtv
 241 ditgnfsnai sksqafslil ahhimgagfg fhnikdpdqn tfaglarssv rhldlshgfv
 301 fslnsrvfet lkdlkvlnla ynkinkiade afygldnlqv lnlsynllge lyssnfyglp
 361 kvayidlqkn hiaiiqdqtf kfleklqtld lrdnalttih fipsipdifl sgnklvtlpk
 421 inltanlihl senrlenldi lyfllrvphl qililnqnrf sscsgdqtps enpsleqlfl
 481 genmlqlawe telcwdvfeg lshlqvlyln hnylnslppg vfshltalrg lslnsnrltv
 541 lshndlpanl eildisrnql lapnpdvfvs lsvldithnk ficecelstf inwlnhtnvt
 601 iagppadiyc vypdsfsgvs lfslstegcd eeevlkslkf slfivctvtl tlflmtiltv
 661 tkfrgfcfic yktaqrlvfk dhpqgtepdm ykydaylcfs skdftwvqna llkhldtqys
 721 dqnrfnlcfe erdfvpgenr ianiqdaiwn srkivclvsr hflrdgwcle afsyaqgrcl
 781 sdlnsalimv vvgslsqyql mkhqsirgfv qkqqylrwpe dfqdvgwflh klsqqilkke
 841 kekkkdnnip lqtvatis
```

Mobilan = AD(TLR5+CBLB502S)



EF1 – promoter EF1 (elongation factor 1α).
502 - flagellin (fliC) gene