TITLE

CASB7439 Constructs

BACKGROUND

The mammalian Achaete-Scute homologs are conserved mammalian cognates of the Drosophila Achaete-Scute complex. They are members of the basic helix-loop-helix (bHLH or HLH) gene family that function as lineage-specific transcription factors essential for development. One murine member of the family, MASH2, is essential for development of a placenta, but not important for the development of the embryo proper. HASH2, the human ortholog of the MASH2 gene, was cloned in 1997 by Alders *et al.* (designated "ASCL2" by the authors). Alders *et al.* (1997) *Hum. Molec. Genet.* 6: 859-867.

As described in WO01/62778, expression studies revealed that the HASH2 transcript (termed CASB7439 therein) is over-expressed in colorectal tumors compared to adjacent normal colon and to other tested normal tissues. This gene is over expressed in patients with stage I to IV adenocarcinoma. Thus, the protein can be considered as a cancer antigen useful in an immunotherapeutic approach for ameliorating a subject's cancer, for instance by using recombinant CASB7439 as an antigen specific cancer immunotherapeutic (ASCI).

SUMMARY OF THE INVENTION

Compounds and methods for increasing the recombinant production of CASB7439 polypeptides are provided herein. In some embodiments are provided modified CASB7439 polypeptides, as well as protein constructs comprising such modified CASB7439 polypeptides. Said modified CASB7439 polypeptides comprise at least one modification for enhanced production of CASB7439. Applicants also disclose nucleic acid molecules comprising a polynucleotide sequence encoding modified CASB7439 polypeptides and protein constructs comprising such modified CASB7439 polypeptides as described herein. In certain aspects, Applicants disclose protein constructs having the amino acid (aa) sequences and the nucleotide sequence (DNA) encoding them, as follows:

- LVL055 [SEQ ID NO:1 (aa); SEQ ID NO:2 (DNA)];
- LVL111 [SEQ ID NO:3 (aa); SEQ ID NO:4 (DNA)];
- LVL137 [SEQ ID NO:5 (aa); SEQ ID NO:6 (DNA)];
- LVL141 [SEQ ID NO:7 (aa); SEQ ID NO:8 (DNA)];
- LVL144 [SEQ ID NO:9 (aa); SEQ ID NO:10 (DNA)]; and
- LVL168 [SEQ ID NO:11 (aa); SEQ ID NO:12 (DNA)].

Methods and processes utilizing the nucleic acid molecules described herein for the production of protein constructs are also disclosed.

Applicants also disclose immunogenic compositions, these compositions comprising one or more of the modified CASB7439 polypeptides or constructs described herein, and a pharmaceutically acceptable carrier or excipient, wherein the carrier or excipient may optionally comprise a buffer.

In some embodiments are disclosed the use of the modified CASB7439 polypeptides or protein constructs described herein—or the nucleic acid molecules encoding them—in the preparation of a medicament for treating colorectal cancer. In some embodiments are disclosed protein constructs as disclosed herein for use in therapy, particularly colorectal cancer therapy. In some embodiments are disclosed methods for eliciting an immune response against CASB7439 in a subject with colorectal cancer, the method comprising: (a) selecting a subject with colorectal cancer; and (b) administering to the subject an effective amount of an immunogenic composition comprising the modified CASB7439 polypeptides or construct as described herein and a pharmaceutically acceptable carrier or excipient, wherein the carrier or excipient may optionally comprise a buffer.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1/35. This figure presents a schematic diagram depicting various regions of the CASB7439 polypeptide, specifically: the DNA specific (DNA binding) domain, the bHLH domain, and the proline-rich region. See Example 2 for details. Schematic representations are also shown of the following regions: Nucleus targeting (gray arrow); alpha helical (gray cylinders); coiled-coil (black cylinders); intrinsic disorder (black arrows); and low complexity (gray box).

Figure 2/35. This figure presents a schematic diagram depicting various truncates of the CASB7439 polypeptide (black arrows) aligned with the complete CASB7439 polypeptide. See Example 4 for details.

Figure 3/35. This figure presents a schematic diagram of CASB7439 showing the approximate location of the proline rich region, and identified and predictive epitopes. Various embodiments comprise a CASB7439 polypeptide modified by removal of some or all of the amino acids within the proline-rich region. Such modified CASB7439 polypeptides retain a high portion of epitopes within the CASB7439 polypeptide.

Figure 4/35. An alignment is shown between the sequence of the unmodified CASB7439 polypeptide (SEQ ID NO:13) and the polypeptide sequence resulting from two possible modifications to the CASB7439 polypeptide: the first in which 21 contiguous amino acid residues (133-153, inclusive) are deleted; the second in which 21 contiguous amino acid residues (131-151, inclusive) are deleted. Further, a third modification to the CASB7439 polypeptide can be envisioned in which 21 contiguous amino acid residues (132-152, inclusive) are deleted. As shown in the Figure the resulting modified CASB7439 polypeptide sequence (inset, circled) in the region of the deletion is the same for all three modifications. This is because of the repeated RG amino acid residues at positions 131-132 and 152-153 of SEQ ID NO:13.

Figure 5/35. This figure presents an amino acid sequence alignment between the protein constructs of Example 6. CASB7439 = the HASH2 amino acid sequence (SEQ ID NO:13); LVL088 (SEQ ID NO:27); LVL111 (SEQ ID NO:3); LVL137 (SEQ ID NO:5); LVL138 (SEQ ID NO:33); LVL168 (SEQ ID NO:11).

Figure 6/35. This figure presents an alignment between the HASH2 (CASB7439) amino acid sequence (SEQ ID NO:13) and the following protein constructs: LVL168 (SEQ ID NO:11); pD1/3 (SEQ ID NO:39); and LVL144 (SEQ ID NO:9).

Figure 7/35. Real-time qPCR analysis of CASB7439 mRNA expression in normal (12) and colorectal cancer (83) tissues. See Example 9.

Figure 8/35. Correspondence between CASB7439 mRNA and protein expression in various human colorectal (CRC) samples. ● (black circles) represent the CASB7439 mRNA expression in colon cancer samples, whereas ○ (open circles) represent the CASB7439 mRNA expression in normal adjacent tissues. The corresponding level of CASB7439 protein detected by immunofluorescence is indicated above each sample (**-** no expression, **+/-** very low expression, **+** low expression, **++** strong expression and **+++** very strong expression). CRC: primary tumors, METS: metastasis, NAT: normal adjacent tissues. All samples are colon adenocarcinomas except the samples 22440 and 24762 which are mucinous carcinomas. See Example 9.

Figure 9/35: The bank of peptides covering the entire CASB7439 protein sequence used for splenocytes restimulation after immunization with various CASB7439 constructs is set forth in this table. See Example 10.

Figure 10/35. The CASB7439 peptide matrix used for T-cell immunogenicity investigations is shown. See Example 10.

Figure 11/35. CD4 response expressed as percent double positive (IFNγ/TNFα) CD4 T-cells in four strains of inbred mice (C57BL6, BALBC, CB6F1, and C3H) immunized with LVL111 + AS01B following re-stimulation with CASB7439 overlapping peptides (pools, matrix approach). See Example 10, Inbred Mouse Multistrain Comparison Experiment.

Figure 12/35. CD4 response expressed as percent double positive (IFNγ/TNFα) CD4 T-cells in the four strains of inbred mice immunized with LVL111 + AS15 following re-stimulation with CASB7439 overlapping peptides (pools, matrix approach). See Example 10, Inbred Mouse Multistrain Comparison Experiment.

Figure 13/35. Identification of CD4 Immunogenic CASB7439 peptides in the four inbred mouse strains discussed in the preceding paragraph. Light gray = 0.2-0.4 % double positive (IFNγ/TNFα) CD4 T-cells; dark gray ≥ 0.5% double positive (IFNγ/TNFα) CD4 T-cells. See Example 10, Inbred Mouse Multistrain Comparison Experiment.

Figure 14/35. Identification of CD4 Immunogenic CASB7439 peptides in outbred CD1 mice (AS01B Top panel; AS15 Bottom panel). Light gray = 0.2-0.4 % double positive (IFNγ/TNFα) CD4 T-cells; dark gray ≥ 0.5% double positive (IFNγ/TNFα) T-CD4 cells. See Example 10, Immunogenicity of CASB7439 in Mice: CASB7439 T-cell Immunogenicity in Outbred Mice.

Figure 15/35. CD4 and CD8 T-cell responses expressed as percent double positive (IFNγ/TNFα) in individual CD1 outbred mice immunized with LVL111, LVL168, or LVL144 formulated with AS15, following re-stimulation with CASB7439 overlapping peptides (pools of immunodominant peptides). See Example 10, Studies of LVL111, LVL168, and LVL144 in out-bred Mice.

Figure 16/35. CD4 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled peripheral blood leukocytes (PBL) isolated from HLA A2.1/DR-1 transgenic mice immunized with either LVL168 (SEQ ID NO:11) or LVL144 (SEQ ID NO:9), formulated with AS15, following re-stimulation with CASB7439 overlapping peptides (7 pools of 7 peptides/pool + peptide 24 (SEQ ID NO:76)).

Figure 17/35. CD8 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled PBL isolated from HLA A2.1/DR-1 transgenic mice immunized with either LVL168 (SEQ ID NO:11) or LVL144 (SEQ ID NO:9), formulated with AS15, following re-stimulation with CASB7439 overlapping peptides (7 pools of 7 peptides + peptide 24 (SEQ ID NO:76)).

Figure 18/35. CD4 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in splenocytes isolated from HLA A2.1/DR-1 transgenic mice immunized

with either LVL168 (SEQ ID NO:11) or LVL144 (SEQ ID NO:9), formulated with AS15, following re-stimulation with CASB7439 overlapping peptides (7 pools of 7 peptides + peptide 24 (SEQ ID NO:76)).

Figure 19/35. CD8 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in splenocytes isolated from HLA A2.1/DR-1 transgenic mice immunized with either LVL168 (SEQ ID NO:11) or LVL144 (SEQ ID NO:9), formulated with AS15, following re-stimulation with CASB7439 overlapping peptides (7 pools of 7 peptides + peptide 24 (SEQ ID NO:76)).

Figure 20/35. Time-course analysis of the CASB7439-specific antibody response (IgG1 and IgG2a) triggered upon LVL111 + AS01B or LVL111 + AS15 immunizations in CB6f1 inbred mice. See Example 10, Immunogenicity of CASB7439 in Mice: CASB7439-mediated Humoral Response.

Figure 21/35. CASB7439-specific humoral response in inbred CB6f1 mice (IgG1 and IgG2a) upon immunizations with LVL111, LVL168, or LVL144 formulated with AS15 adjuvant or with AS15 adjuvant alone. See Example 10, Immunogenicity of CASB7439 in Mice: CASB7439-mediated Humoral Response.

Figure 22/35. CASB7439-specific humoral immune response (IgG titer) in CD1 outbred mice immunized with LVL111, LVL168, or LVL144 formulated with AS15 adjuvant, or with AS15 adjuvant alone; naive mice were used as control. See Example 10, CD1 Mouse Studies.

Figure 23/35. TC1/CASB7439 #14 tumor graft in CB6f1 mice immunized with LVL111 formulated with AS01B (top) or AS15 (bottom). See Example 11, Study #1.1. In the top panel, the order of curves on the graph from top to bottom is 'buffer', LVL111 (10µg), LVL111 (10µg) + AS01B, and LVL111 (1µg) + ASO1B. In the bottom panel, the curve at the bottom is LVL111(10µg) + AS15.

Figure 24/35. Survival curves from Example 11, Study #1.2 are presented in top, middle, and bottom panels. TF: tumor-free. (Top) The survival curve of CB6f1 mice immunized with saline buffer, LVL111 without adjuvant, or adjuvant alone (AS01 or AS15) and challenged with TC1/CASB7439 #14 cells. The last surviving mouse from the group immunized with LVL111 died first, followed by the last surviving mouse from the AS01B group, followed by the last surviving mouse from the AS15 group and the last surviving mouse from the buffer group. (Middle) The survival curve of CB6f1 mice immunized with AS15 alone or with the indicated dosages of LVL111 formulated with AS15 and challenged with TC1/CASB7439 #14 cells. The last surviving mouse from the AS15 group expired first. Of the other groups, the 30µg group had the fewest survivors, followed by the 1µg group, then the 10µg group, which had the most survivors. (Bottom) The survival curve of CB6f1 mice immunized with AS01B alone or with the indicated dosages of LVL111 formulated with AS01B and challenged with TC1/CASB7439 #14 cells. In this panel, the last surviving mice from the 10µg and AS01B groups expired on the same day. The 1µg and 30µg groups had the same percentage of survival at the final day.

Figure 25/35. Survival curves from Example 11, Study #1.3 are presented. TF: tumor-free. (Top) Survival curve for CB6f1 mice immunized with LVL111, LVL144, or LVL168 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439#14 cells. (Bottom) Survival curve for CB6f1 mice immunized with LVL111, LVL144, or LVL168 formulated with AS15, or AS15 alone, and challenged with MC38/CASB7439#35 cells.

Figure 26/35. Survival curves from Example 11, Study #1.4 are presented. TF: tumor-free. (Top) Survival curve of CB6f1 mice immunized with LVL144 or LVL168 formulated with AS15, or with AS15 alone, and challenged with TC1/CASB7439 #14 cells. The group immunized with LVL144 had a marginally greater percent survival. (Middle) Survival curve for CB6f1 mice immunized with LVL144 or LVL168 formulated with AS15, or AS15 alone, and challenged with clonal TC1/CASB7439 #14-2 cells. Mice inoculated with LVL144 had a slightly higher percentage of survival. (Bottom) Survival curve for CB6f1 mice immunized with LVL144 or LVL168 formulated with AS15, or

AS15 alone, and challenged with MC38/CASB7439 #35 cells. Mice inoculated with LVL144 had no survivors.

Figure 27/35. Survival curves from Example 11, Study #2.1 are presented. TF: tumor-free. (Top) Survival curve of CB6f1 mice immunized with 6.25 µg of LVL168 and 10 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. The group of mice inoculated with LVL168 had a slightly higher percent survival. (Top middle) Survival curve of CB6f1 mice immunized with 3.1 µg of LVL168 and 5 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. The group of mice inoculated with LVL144 had a slightly higher percent survival. (Bottom middle) Survival curve of CB6f1 mice immunized with 1.55 µg of LVL168 and 2.5 µg of LVL144 formulated with AS15, or with AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. The group of mice inoculated with LVL144 had the highest survival percentage. (Bottom) Survival curve of CB6f1 mice immunized with 0.77 µg of LVL168 and 1.25 µg of LVL144 formulated with AS15, or with AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. The group of mice inoculated with LVL168 had a slightly higher percent survival.

Figure 28/35. Survival curves from Example 11, Study #2.2 are presented. TF: tumor-free. (Top) Survival curve of C57Bl/6 mice immunized with 6.25 µg of LVL168 or 10 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. (Top middle) Survival curve of C57Bl/6 mice immunized with 3.1 µg of LVL168 or 5 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. (Bottom middle) Survival curve of C57Bl/6 mice immunized with 1.55 µg of LVL168 or 2.5 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells. (Bottom) Survival curve of C57Bl/6 mice immunized with 0.77 µg of LVL168 or 1.25 µg of LVL144 formulated with AS15, or AS15 alone, and challenged with TC1/CASB7439 #14-2 cells.

Figure 29/35. Survival curves from Example 11, Study #2.3 are presented. TF: tumor-free. (Top) Survival curve of CB6f1 mice immunized with 6.25 µg of LVL168 or 10 µg

of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival, followed by the group receiving LVL144. There were no survivors in the group receiving only AS15. (Middle) Survival curve of CB6f1 mice immunized with 3.1 μg of LVL168 and 5 μg of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival, followed by the group receiving LVL144. There were no survivors in the group receiving only AS15. (Bottom) Survival curve of CB6f1 mice immunized with 1.55 μg of LVL168 and 2.5 μg of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival, followed by the group receiving LVL144. There were no survivors in the group receiving only AS15. Saline buffer or AS15 alone were used as controls.

Figure 30/35. Survival curves from Example 11, Study #2.3. TF: tumor-free. (Top) Survival curve of CB6f1 mice immunized with 0.77 μg of LVL168 or 1.25 μg of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival, followed by the group receiving LVL144. There were no survivors in the group receiving only AS15. (Middle) Survival curve of CB6f1 mice immunized with 0.19 μg of LVL168 or 0.31 μg of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival. The last surviving mouse from the group receiving LVL144 outlived the last survivor from both the AS15 group and the buffer group. (Bottom) Survival curve of CB6f1 mice immunized with 0.048 μg of LVL168 or 0.078 μg of LVL144 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. The group of mice receiving LVL168 had the highest percent survival. The last surviving mouse from the group receiving LVL144 outlived the last survivor from both the AS15 group and the buffer group. Saline buffer or AS15 alone were used as controls.

Figure 31/35. TC1/CASB7439 #14-2 tumor graft in male and female CB6f1 mice immunized with 1 μg of LVL168 formulated with AS15; AS15 alone was used as control

(top). TC1/CASB7439 #14-2 tumor graft in C57Bl/6 mice immunized with 1 µg of LVL168 formulated with AS15; AS15 alone was used as control (bottom).

Figure 32/35. Top panel: Survival curve of male or female CB6f1 mice immunized with 1 µg of LVL168 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. AS15 alone was used as control in both male and female CB6f1 mice. Among the CB6f1 mice immunized with LVL168 plus AS15, there were 7 tumor-free (TF) males and 3 TF females. There were no TF among the mice receiving AS15 alone. Bottom panel: Survival curve of C57Bl/6 mice immunized with 1 µg of LVL168 formulated with AS15 and challenged with TC1/CASB7439 #14-2 cells. AS15 alone was used as control in both male and female C57Bl/6 mice. Among the C57Bl/6 mice immunized with LVL168 plus AS15, there were 2 TF males and 1 TF females. Among the mice receiving AS15 alone, there were 5 TF males and 1 TF female.

Figure 33/35. CASB7439-specific humoral immune response (IgG titer) in male (black circles) and female (gray circles) CB6F1 mice immunized with 1 µg of LVL168 formulated with AS15, or with AS15 alone (left). CASB7439-specific humoral immune response (IgG titer) in male (black circles) and female (gray circles) C57/BL6 mice immunized with 1 µg of LVL168 formulated with AS15, or with AS15 alone (right).

Figure 34/35. CD4 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled PBL isolated from CB6F1 mice immunized with either 1 µg of LVL168 (SEQ ID NO:11) formulated with AS15, or AS15 alone, following re-stimulation with a bank of 46 peptides (see Figure 9) covering the entire CASB7439 protein sequence (top). CD4 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled PBLs isolated from C57/BL6 mice immunized with either 1 µg of LVL168 (SEQ ID NO:11) formulated with AS15, or AS15 alone, following re-stimulation with CASB7439 pooled peptides a bank of 46 peptides (see Figure 9) covering the entire CASB7439 protein sequence (bottom left); CASB7439 peptide 39 (SEQ ID NO:91) (bottom middle); or RPMI (bottom right).

Figure 35/35. CD8 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled PBL isolated from CB6F1 mice immunized with either 1 µg of LVL168 (SEQ ID NO:11) formulated with AS15, or AS15 alone, following re-stimulation with a bank of 46 peptides (see Figure 9) covering the entire CASB7439 protein sequence (top). CD8 T-cell responses (expressed as percent double positive (IFNγ/TNFα)) in pooled PBLs isolated from C57/BL6 mice immunized with either 1 µg of LVL168 (SEQ ID NO:11) formulated with AS15, or AS15 alone, following re-stimulation with a bank of 46 peptides (see Figure 9) covering the entire CASB7439 protein sequence (bottom left); CASB7439 peptide 39 (SEQ ID NO:91) (bottom middle); or RPMI (bottom right).

DETAILED DESCRIPTION

One hurdle to a commercially acceptable immunotherapeutic is the production of commercially acceptable amounts of the recombinant cancer antigen. In general and without limitation, a commercially acceptable production level for a recombinant polypeptide is roughly 5% of total protein produced by the host cell, as estimated on a Coomasie blue stained SDS-PAGE gel.

There are numerous reasons why a native polypeptide sequence may yield only small quantities when produced by recombinant means. Approaches to increasing polypeptide production include varying host expression systems or codon optimization. However, these standard approaches do not always result in acceptable polypeptide production and the production for some polypeptides remains commercially prohibitive. As described in greater detail below, Applicants' investigated whether a modified CASB7439 polypeptide would produce acceptable protein levels by recombinant means.

In some embodiments are provided modified CASB7439 polypeptides, wherein the modification consists of or comprises C-terminal truncation of the CASB7439 polypeptide such that all or a portion of the proline-rich region is removed. In some embodiments the CASB7439 polypeptide modification consists of or comprises deletion of all or a portion of the proline-rich region. In some embodiments are provided

modified CASB7439 polypeptides that retain at least amino acid residue 193 of SEQ ID NO:13, but wherein all or a portion of the proline-rich region is deleted.

In other embodiments are provided protein constructs comprising a modified CASB7439 polypeptide as described herein. In some embodiments, such protein constructs further comprise one or more heterologous polypeptides. In some embodiments, the heterologous polypeptide is a fusion partner, as described in greater detail elsewhere. In particularly suitable embodiments, such protein constructs comprise a heterologous polypeptide located N-terminal of the modified CASB7439 polypeptide. In further particularly suitable embodiments are provided protein constructs wherein the heterologous polypeptide is a fragment of the *H. influenzae* Protein D (as described in greater detail below) located N-terminal of the modified CASB7439 polypeptide.

In some embodiments are provided protein constructs wherein the heterologous polypeptide comprises a polyhistidine region. In further embodiments the polyhistidine sequence is located in the C-terminal portion of the protein construct. In other embodiments the polyhistidine sequence is located at the C-terminus of the protein construct. In further embodiments the polyhistidine sequence is located in the N-terminal portion of the protein construct. In other embodiments the polyhistidine sequence is located at the N-terminus of the protein construct. In suitable embodiments the polyhistidine region comprises ten or more consecutive histidine residues. In particularly suitable embodiments the polyhistidine region comprises ten or more consecutive histidine residues located N-terminal of the modified CASB7439 polypeptide. In some embodiments, more than one polyhistidine region may be included at one or more of the locations described herein.

In some embodiments are provided protein constructs wherein the heterologous polypeptide is a carrier protein chemically conjugated to the modified CASB7439 polypeptide.

In some embodiments are provided protein constructs comprising a heterologous polypeptide that is a polyhistidine region and a heterologous polypeptide that is a carrier protein. In some embodiments are provided protein constructs comprising a heterologous polypeptide that is a polyhistidine region and a heterologous polypeptide

that is a fusion partner. In some embodiments are provided protein constructs comprising a heterologous polypeptide that is a polyhistidine region, a heterologous polypeptide that is a fusion partner, and a heterologous polypeptide that is a carrier protein.

Applicants also disclose embodiments of each of the constructs described in the paragraphs herein in which some or all of any heterologous polypeptide, such as a polyhistidine region, has been removed. Compositions and methods for the removal of heterologous polypeptides are known in the art, including without limitation digestion with endo- or exo- peptidases, chemical modification or breakage of the bond between the heterologous polypeptide and the remainder of the molecule, and the like.

Applicants also disclose herein immunogenic compositions, these compositions comprising any modified CASB7439 polypeptide or protein construct comprising a modified CASB7439 polypeptide as described herein, and a pharmaceutically acceptable carrier or excipient, wherein the carrier or excipient may optionally comprise a buffer. In certain aspects, these compositions further comprise an adjuvant. In certain aspects, these compositions comprise an adjuvant that elicits at least a TH1 immune response. In certain aspects, the adjuvant described herein comprises at least one of: 3D-MPL, QS21, and CpG. In some embodiments the composition comprises 3D-MPL. In other embodiments the composition comprises CpG. In some embodiments the composition comprises QS21. In some embodiments the composition comprises QS21 and cholesterol. In some embodiments the composition comprises cholesterol, 3D-MPL, and QS21 in a liposomal formulation.

Applicants also disclose herein nucleic acid molecules comprising polynucleotide sequences that encode the modified CASB7439 polypeptides and constructs as described herein. In some embodiments are disclosed a vector comprising such nucleic acid molecules as described herein. In some embodiments the vector comprises a prokaryotic expression vector. In some embodiments the vector comprises a eukaryotic expression vector. In some embodiments the polynucleotide sequence that encodes the construct has been codon optimized for expression in a host cell.

Applicants also disclose host cells comprising (i) a nucleic acid molecule as described herein or (ii) a vector comprising a nucleic acid molecule as described herein.

In some embodiments the host cell is selected from the group of: bacterial cells, insect cells and mammalian cells.

Some embodiments provide an immunogenic composition comprising (i) a nucleic acid molecule described herein or (ii) a vector comprising a nucleic acid molecule as described herein, and a pharmaceutically acceptable carrier or excipient.

In some embodiments the use of modified CASB7439 polypeptides or the protein constructs described herein (or the nucleic acid molecules encoding them) in the preparation of a medicament for treating colorectal cancer is provided. In some embodiments are disclosed modified CASB7439 polypeptides or the protein constructs described herein (or the nucleic acid molecules encoding them) for use in therapy, particularly colorectal cancer therapy. In some embodiments are provided methods for eliciting an immune response against CASB7439 in a subject with colorectal cancer, the method comprising: (i) selecting a subject with colorectal cancer; and (ii) administering to the subject an effective amount of an immunogenic composition comprising a modified CASB7439 polypeptide or protein construct as described herein, and a pharmaceutically acceptable carrier or excipient, wherein the carrier or excipient may optionally comprise a buffer. In some embodiments of the methods disclosed herein, the immunogenic composition further comprises an adjuvant. In some embodiments of the method described herein, the adjuvant elicits at least a TH1 immune response. In some embodiments of the method disclosed herein, the adjuvant comprises at least one of: 3D-MPL, QS21, and CpG. In some embodiments of the method described herein, the subject is a human subject.

PROTEIN CONSTRUCTS

CASB7439 Polypeptides

CASB7439, as it has been termed herein, is also known as HASH2 or ASCL2. HASH2 is a 193 amino acid residue polypeptide of human origin (SEQ ID NO:13). See, e.g., Accession No. AAB86993. Herein, reference to CASB7439 features will be made in terms of the polypeptide sequence provided in SEQ ID NO:13.

Proline-rich region

A region of high proline-periodicity of is found in SEQ ID NO:13 in the region of amino acids 127-158. Applicants termed the region of SEQ ID NO:13 from amino acid 133 to amino acid 153, inclusive, the "proline-rich," "polyproline," "polyproline-like," or "polypro" region or domain. Accordingly, as used herein, "proline-rich," "polyproline," "polyproline-like," and "polypro" refers to this region (or domain) of the CASB7439 protein.

Modifications

This disclosure provides protein constructs comprising a CASB7439 polypeptide comprising at least one expression enhancement modification. In some embodiments a portion of the CASB7439 polypeptide is removed, said portion comprising all or a portion of the proline-rich region. In some embodiments the entire C-terminal portion of the CASB7439 polypeptide is removed, resulting in a truncated CASB7439 polypeptide comprising amino acid residues 1-117 of SEQ ID NO:13. In some embodiments a heterologous polypeptide is positioned N-terminal of the CASB7439 polypeptide.

As used herein, "expression enhancement" or "enhanced expression" in the context of a modified CASB7439 polypeptide refers to a modification intended to increase expression levels of a protein construct comprising said modified CASB7439 polypeptide compared to the expression of an otherwise-equivalent protein construct comprising an unmodified CASB7439 polypeptide of SEQ ID NO:13, i.e., same vector, host cell, etc. Any of the methods known in the art for determining protein quantity are acceptable to determine whether a given construct comprising a modified CASB7439 polypeptide has an increased expression level as compared to an otherwise-equivalent protein construct comprising an unmodified CASB7439 polypeptide. In some embodiments production may be assessed using SDS-PAGE and Coomasie Blue staining.

Removal of the Proline-rich Region

In some embodiments the modification of the CASB7439 polypeptide comprises the deletion of some of the proline-rich region. In some embodiments the modification of the CASB7439 polypeptide comprises the deletion of all of the proline-rich region. In some embodiments the modification of the CASB7439 polypeptide comprises the deletion of amino acids outside of the proline-rich region, in addition to deletion of the proline-rich region.

In some embodiments modification is accomplished by the truncation of the C-terminal portion of SEQ ID NO:13. Thus, in some embodiments a protein construct is generated in which the C-terminal amino acid residue of the modified CASB7439 polypeptide corresponds to a residue located N-terminal of the proline-rich region of SEQ ID NO:13. In some embodiments a protein construct is generated in which the C-terminal amino acid residue of the modified CASB7439 polypeptide corresponds to a residue located within the proline-rich region of SEQ ID NO:13. In some embodiments the C-terminal amino acid residue of the modified CASB7439 polypeptide corresponds to amino acid residue 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, or 152 of SEQ ID NO:13.

In some embodiments the modification comprises the deletion of part or all of the proline-rich region. In some embodiments a modified CASB7439 polypeptide sequence corresponds to SEQ ID NO:13 from which 21 or more contiguous amino acid residues have been deleted. These 21 or more contiguous residues correspond to a contiguous amino acid sequence of SEQ ID NO:13 starting at any one of residues 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, or 152, inclusive. In further embodiments, a modified CASB7439 polypeptide sequence corresponds to SEQ ID NO:13 from which 21 contiguous amino acid residues have been deleted, said 21 contiguous residues starting at any one of residues 131, 132, and 133.

In some embodiments the modified CASB7439 polypeptide sequence corresponds to SEQ ID NO:13 from which one or more of residues 133, 134, 135, 136,

137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153 have been deleted. In some embodiments the modified CASB7439 polypeptide sequence corresponds to SEQ ID NO:13 from which two or more contiguous amino acid residues have been deleted in the region of amino acid residues 133-153, inclusive.

In some embodiments the modified CASB7439 polypeptide sequence corresponds to SEQ ID NO:13 from which residues 133 to 153, inclusive, of SEQ ID NO: 13 have been deleted. Exemplary non-limiting embodiments include the following:

- LVL055, SEQ ID NO:1 (aa);
- LVL111, SEQ ID NO:3 (aa);
- LVL137, SEQ ID NO:5 (aa);
- LVL141, SEQ ID NO:7 (aa);
- LVL144, SEQ ID NO:9 (aa); and
- LVL168, SEQ ID NO:11 (aa).

Heterologous Sequences

In some embodiments the modified CASB7439 polypeptide is combined with a heterologous polypeptide. The term "heterologous" with respect to a nucleic acid molecule, a polypeptide or another cellular component, indicates that the component occurs where it is not normally found in nature and/or that it originates from a different source or species than a reference molecule. As used herein, "heterologous" molecules include without limitation fusion proteins, carrier proteins, and purification tags.

In some embodiments the heterologous polypeptide may be chemically conjugated to the modified CASB7439 polypeptide, i.e., a carrier protein. In other embodiments the heterologous polypeptide and the modified CASB7439 polypeptide may be expressed as a single recombinant fusion protein. In other embodiments, a modified CASB7439 polypeptide and heterologous polypeptide are expressed as a single recombinant fusion protein and chemically conjugated to another heterologous polypeptide.

The heterologous polypeptide may assist in providing T helper epitopes, including T helper epitopes recognized by humans. In the case of a protein construct

comprising a CASB7439 polypeptide and a heterologous polypeptide that is a fusion protein, the heterologous polypeptide (fusion partner) may assist in providing such epitopes or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. The fusion partner may be both an immunological fusion partner and an expression enhancing partner.

Carrier Proteins

A carrier protein is chemically conjugated to another polypeptide of interest. The chemical coupling of a polypeptide to a carrier protein can be carried out in any manner known in the art. Thus, for direct covalent coupling it is possible to utilize a carbodiimide, glutaraldehyde or N-[γ-maleimidobutyryloxy]succinimide ester, utilizing common commercially available heterobifunctional linkers such as CDAP and SPDP (using manufacturer's instructions). After the coupling reaction, the polypeptide-carrier protein conjugate can easily be isolated and purified by means of a dialysis method, a gel filtration method, a fractionation method, etc.

The types of carrier proteins used in some embodiments will be readily known to the person skilled in the art. The function of some carrier proteins is to provide cytokine help in order to help induce an immune response against the coupled polypeptide. For example, a non-exhaustive list of carrier proteins which may be used in the present invention include: Keyhole limpet Haemocyanin (KLH), serum albumins such as bovine serum albumin (BSA), inactivated bacterial toxins such as tetanus or diptheria toxins (TT and DT), or recombinant fragments thereof (for example, Domain 1 of Fragment C of TT, or the translocation domain of DT), or the purified protein derivative of tuberculin (PPD).

Fusion Partners

Despite the availability of numerous gene fusion systems, recombinant protein expression in *Escherichia coli* remains difficult. Establishing the fusion partner that best enhances expression of difficult-to-express proteins remains empirical. Common fusion partners include the C termini of maltose-binding protein (MBP), glutathione S-transferase (GST), thioredoxin (TRX), NUS A, ubiquitin (Ub), and Small Ubiquitin-like Modifier (SUMO) tags, each of which is described in the literature. See, e.g., Marblestone (2006) *Prot. Sci.* 15:182-7; Hunt (2005) *Protein Exp. & Purif* 40:1-22; Hammarstom et al. (2002) *Protein Sci.* 11:313-321.

Immunogenic fusion partners include protein D from Haemophilus influenza B and the non-structural protein from influenzae virus, NS1 (hemagglutinin). Another immunological fusion partner is the protein known as LYTA. The C terminal portion of the molecule may be used. LYTA is derived from Streptococcus pneumoniae which synthesize an N-acetyl-L-alanine amidase, amidase LYTA, (coded by the lytA gene García *et al.* (1986) *Gene* 43:265-272). It is possible to use the repeat portion of the Lyta molecule found in the C terminal end starting at residue 178, for example residues 188-305. Another suitable fusion partner is an adenylate cyclase (CyaA) protein or of a fragment thereof, wherein said CyaA fragment retains the property of said adenylate cyclase protein to target Antigen Presenting Cells. See WO2005089792.

In one embodiment of the invention a heterologous polypeptide is Protein D from *Haemophilus influenzae* (EP 0 594 610 B1), or a fragment thereof. Protein D is an IgD-binding protein from *Haemophilus influenzae* (WO 91/18926, granted EP 0 594 610 B1). In some circumstances, for example in recombinant immunogen expression systems it may be desirable to use N-terminal fragments of protein D, for example comprising from about 100 to about 110 N-terminal amino acids of protein D (GB 9717953.5) (pD1/3, SEQ ID NO:39).

H. influenzae Protein D is synthesized as a precursor with an 18 amino acid signal sequence (SEQ ID NO:41, see also Accession Number AAA24998). When the signal sequence is processed during secretion, the cysteine residue at position 19 in the precursor molecule becomes the amino terminal residue.

In one embodiment, a heterologous polypeptide sequence comprises the first approximately one-third of the processed H. influenzae Protein D molecule, or the N-terminal 100-110 amino acids of processed Protein D. In one embodiment, the heterologous protein comprises amino acid residues Met-Asp-Pro joined to the N-terminal 109 amino acids of processed Protein D. In another embodiment, the heterologous protein comprises amino acid residues Met-Asp-Pro joined to amino acid residues 2-109 of processed Protein D (SEQ ID NO:39).

In one embodiment, a modified CASB7439 polypeptide or protein construct is chemically coupled to CyaA protein or a fragment thereof. See WO2005054851. In another embodiment, a modified CASB7439 polypeptide or protein construct is chemically coupled to the B subunit of Shiga Toxin or an immunologically functional equivalent thereof. See US Patent No. 6,613,882; WO02060937; WO2005112991.

Polyhistidine tags, unrelated amino acids.

It is often advantageous to include a heterologous polypeptide which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as a histidine tag, e.g., a polyhistidine sequence (comprised of multiple histidine residues), or an additional sequence for stability during recombinant production.

Histidine tag vectors and kits are commercially available. For instance, the vectors for adding a six histidine tag to a polypeptide are available from Roche. Kits for making and using histidine tagged polypeptides are available from Qiagen, Sigma, Thermo Scientific, GE Healthcare and others. The histidine tag may also be followed by a suitable amino acid sequence that facilitates a removal of the polyhistidine sequence using endopeptidases. An exopeptidase, for instance, the Qiagen TAGZyme, can be used to remove a terminal polyhistidine sequence without the additional sequence.

In some embodiments the polyhistidine sequence is located in the N-terminal region of the protein construct. In further embodiments the polyhistidine sequence is located at the N-terminus of the protein construct. In some embodiments the polyhistidine sequence comprises 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more consecutive histidines. In further embodiments the polyhistidine sequence comprises ten consecutive histidines.

In some embodiments the heterologous polypeptide includes a polyhistidine sequence and one or more unrelated amino acids. By "unrelated amino acids" is intended one or more amino acids that are not a part of the polyhistidine sequence, or fusion protein, and are not naturally a part of the CASB7439 polypeptide sequence. For instance, unrelated amino acids may be those included to provide a peptidase site, they may simply be extraneous sequence such as a cloning artifact, or the like.

NUCLEIC ACID MOLECULES ENCODING CASB7439 CONSTRUCTS

Other embodiments of this disclosure concern recombinant nucleic acids that encode the modified CASB7439 polypeptides and protein constructs as described herein. In certain embodiments, the recombinant nucleic acids are codon optimized for expression in a selected prokaryotic or eukaryotic host cell. To facilitate replication and expression, the nucleic acids can be incorporated into a vector, such as a prokaryotic or a eukaryotic expression vector. Host cells including recombinant modified CASB7439 polypeptides or protein construct-encoding nucleic acids are also a feature of this disclosure. Favorable host cells include prokaryotic (i.e., bacterial) host cells, such as *E. coli*, as well as numerous eukaryotic host cells, including fungal (e.g., yeast) cells, insect cells, and mammalian cells (such as HEK293, CHO and VERO cells).

Expression Vectors

To facilitate replication and expression, the nucleic acids can be incorporated into a vector, such as a prokaryotic or a eukaryotic expression vector. Although the nucleic acids disclosed herein can be included in any one of a variety of vectors (including, for example, bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, pseudorabies, adenovirus, adeno-associated virus, retroviruses and many others), most commonly the vector will be an expression vector suitable for generating polypeptide expression products. In an expression vector, the nucleic acid encoding the protein construct is typically arranged in proximity and orientation to an appropriate transcription control sequence (promoter, and optionally, one or more enhancers) to direct mRNA synthesis. That is, the polynucleotide

sequence of interest is operably linked to an appropriate transcription control sequence. Examples of such promoters include: the immediate early promoter of CMV, LTR or SV40 promoter, polyhedron promoter of baculovirus, *E. coli* lac or trp promoter, phage T7 and lambda PL promoter, and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector typically also contains a ribosome binding site for translation initiation, and a transcription terminator. The vector optionally includes appropriate sequences for amplifying expression. In addition, the expression vectors optionally comprise one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells, such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

Exemplary procedures sufficient to guide one of ordinary skill in the art through the production of recombinant CASB7439 nucleic acids can be found in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, 1989; Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 3d ed., Cold Spring Harbor Press, 2001; Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Associates, 1992 (and Supplements to 2003); and Ausubel *et al.*, *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, 4th ed., Wiley & Sons, 1999.

Exemplary nucleic acid molecules that encode protein constructs are represented by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, and SEQ ID NO:12. Additional sequence variants that share sequence identity with the exemplary nucleic acid molecule can be produced by those of skill in the art. Typically, the nucleic acid variants will encode polypeptides that differ by no more than 1%, or 2%, or 5%, or 10%, or 15%, or 20% of the amino acid residues. That is, the encoded polypeptides share at least 80%, or 85%, more commonly, at least about 90% or more, such as 95%, 98%, 99%, or 99.5% sequence identity. It will be immediately understood by those of skill in the art, that the polynucleotide sequences encoding the protein constructs can themselves share less sequence identity due to the redundancy of the genetic code. In some instances, the modified CASB7439 polypeptide or protein construct has one or more amino acid modification relative to the amino acid sequence

of the exemplary constructs set forth herein (*e.g.*, in addition to the aforementioned modifications). Such differences can be an addition, deletion or substitution of one or more nucleotides or amino acids, respectively. A variant typically differs by no more than about 1%, or 2%, or 5%, or 10%, or 15%, or 20% or 25% or 30% of the nucleotide residues. For example, a variant CASB7439 polypeptide or protein construct encoding nucleic acid can include 1, or 2, or up to 5, or up to about 10, or up to about 15, or up to about 50, or up to about 100 nucleotide differences as compared to the exemplary CASB7439 polypeptide or protein construct encoding nucleic acids of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, and SEQ ID NO:12. Thus, a variant in the context of a CASB7439 polypeptide or protein construct encoding nucleic acid, typically shares at least 70%, or 75%, or 80%, or 85%, or 90%, or 95%, or 98%, or 99%, or 99.5% sequence identity with a reference sequence, *e.g.*, the reference sequences illustrated by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, and SEQ ID NO:12, or any of the other exemplary CASB7439 polypeptide, protein construct encoding nucleic acids disclosed herein. Additional variants can arise through genetic drift, or can be produced artificially using site directed or random mutagenesis, or by recombination of two or more preexisting variants. Such additional variants are also suitable in the context of the CASB7439 polypeptide or protein constructs disclosed herein.

In addition to the variant nucleic acids previously described, nucleic acids that hybridize to one or more of the exemplary nucleic acids represented by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, and SEQ ID NO:12 can also be used to encode CASB7439 polypeptides or protein constructs disclosed herein. One of skill in the art will appreciate that in addition to the percent (%) sequence identity measure discussed herein, another indicia of sequence similarity between two nucleic acids is the ability to hybridize. The more similar are the sequences of the two nucleic acids, the more stringent the conditions at which they will hybridize. The stringency of hybridization conditions are sequence-dependent and are different under different environmental parameters. Thus, hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences.

Generally, the temperature of hybridization and the ionic strength (especially the Na^+ and/or Mg^{++} concentration) of the hybridization buffer will determine the stringency of hybridization, though wash times also influence stringency. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001; Tijssen, *Hybridization With Nucleic Acid Probes*, *Part I: Theory and Nucleic Acid Preparation*, Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Ltd., NY, NY, 1993 and Ausubel *et al. Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc., 1999.

For purposes of the present disclosure, "stringent conditions" encompass conditions under which hybridization will only occur if there is less than 25% mismatch between the hybridization molecule and the target sequence. "Stringent conditions" can be broken down into particular levels of stringency for more precise definition. Thus, as used herein, "moderate stringency" conditions are those under which molecules with more than 25% sequence mismatch will not hybridize; conditions of "medium stringency" are those under which molecules with more than 15% mismatch will not hybridize, and conditions of "high stringency" are those under which sequences with more than 10% mismatch will not hybridize. Conditions of "very high stringency" are those under which sequences with more than 6% mismatch will not hybridize. In contrast, nucleic acids that hybridize under "low stringency conditions" include those with much less sequence identity, or with sequence identity over only short subsequences of the nucleic acid. It will, therefore, be understood that the various variants of nucleic acids that are encompassed by this disclosure are able to hybridize under stringent conditions to at least one of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, and SEQ ID NO:12 over substantially their entire length.

PRODUCTION OF PROTEIN CONSTUCTS

The protein constructs disclosed herein may be produced using well established procedures for the expression and purification of recombinant proteins. Procedures sufficient to guide one of skill in the art can be found in the following references: Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 200; and Ausubel *et al. Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc., 999. Additional and specific details are provided hereinbelow.

Recombinant nucleic acids that encode the protein constructs may be introduced into host cells by any of a variety of well-known procedures, such as electroporation, liposome mediated transfection (*e.g.*, using a commercially available liposomal transfection reagent, such as LIPOFECTAMINE™2000 or TRANSFECTIN™), Calcium phosphate precipitation, infection, transfection and the like, depending on the selection of vectors and host cells.

The host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the inserted polynucleotide sequences. The culture conditions, such as temperature, pH and the like, are typically those previously used with the host cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, *e.g.*, Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, third edition, Wiley- Liss, New York and the references cited therein. In some embodiments, a suitable culture temperature is 16°C; in others 37°C. Expression products corresponding to the nucleic acids of the invention can also be produced in non-animal cells such as plants, yeast, fungi, bacteria and the like. In addition to Sambrook, Berger and Ausubel, details regarding cell culture can be found in Payne *et al.* (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture*; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) and Atlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

Host Cells

Host cells that include recombinant protein construct-encoding nucleic acids are, thus, also a feature of this disclosure. Favorable host cells include prokaryotic (*i.e.*, bacterial) host cells, such as *E. coli*, as well as numerous eukaryotic host cells, including fungal (*e.g.*, yeast, such as *Saccharomyces cerevisiae* and *Picchia pastoris*) cells, insect cells, plant cells, and mammalian cells (such as CHO cells). Recombinant protein construct nucleic acids are introduced (*e.g.*, transduced, transformed or transfected) into host cells, for example, via a vector, such as an expression vector. As described herein, the vector is most typically a plasmid, but such vectors can also be, for example, a viral particle, a phage, etc. Examples of appropriate expression hosts include: bacterial cells, such as *E. coli*, *Streptomyces*, and *Salmonella typhimurium*; fungal cells, such as *Saccharomyces cerevisiae, Pichia pastoris,* and *Neurospora crassa*; insect cells such as *Drosophila* and *Spodoptera frugiperda*; mammalian cells such as 3T3, COS, CHO, BHK, HEK293 or Bowes melanoma; plant cells, including algae cells, etc.

A host cell is optionally chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, glycosylation, (as well as, *e.g.*, acetylation, carboxylation, phosphorylation, lipidation and acylation). Post-translational processing for example, which cleaves a precursor form into a mature form of the protein (for example, by a furin protease) is optionally performed in the context of the host cell. Different host cells such as 3T3, COS, CHO, HeLa, BHK, MDCK, HEK293, WI38, etc. have specific cellular machinery and characteristic mechanisms for such post-translational activities and can be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant protein constructs disclosed herein, stable expression systems are typically used. For example, nucleic acid molecules that stably express a protein construct are introduced into the host cell using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells are allowed to grow for 1-2 days in an enriched media before they are switched to

selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. For example, resistant groups or colonies of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. Host cells transformed with a nucleic acid encoding a protein construct are optionally cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture.

Expressed protein constructs can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, filtration, ultrafiltration, centrifugation, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography (*e.g.*, using any of the tagging systems noted herein), hydroxylapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as desired, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed in the final purification steps. In addition to the references noted herein, a variety of purification methods are well known in the art, including, *e.g.*, those set forth in Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; and Bollag *et al.* (1996) *Protein Methods, 2nd Edition* Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook* Humana Press, NJ, Harris and Angal (1990) *Protein Purification Applications: A Practical Approach* IRL Press at Oxford, Oxford, U.K.; Scopes (1993) *Protein Purification: Principles and Practice 3rd Edition* Springer Verlag, NY; Janson and Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications, Second Edition* Wiley-VCH, NY; and Walker (1998) *Protein Protocols on CD-ROM* Humana Press, NJ.

In certain examples, the nucleic acids are introduced into cells via vectors suitable for introduction and expression in prokaryotic cells, *e.g.*, *E. coli* cells. For example, a nucleic acid including a polynucleotide sequence that encodes a protein construct can be introduced into any of a variety of commercially available or proprietary vectors, such as the pET series of expression vectors (*e.g.*, pET9b and pET2d). Expression of the coding sequence is inducible by isopropyl β-D-1-

thiogalactopyranoside (IPTG), resulting in high levels of protein expression. The polynucleotide sequence encoding the CASB7439 protein construct is transcribed under the phage T7 promoter. Alternate vectors, such as pURV22 that include a heat-inducible lambda pL promoter are also suitable.

The expression vector is introduced (*e.g.*, by electroporation) into a suitable bacterial host. Numerous suitable strains of *E. coli* are available and can be selected by one of skill in the art (for example, BLR DE3, BL21 DE3 and Rosetta DE3 strains have proven favorable for expression of recombinant vectors containing polynucleotide sequences that encode protein constructs).

Optionally, the polynucleotides that encode the protein constructs are incorporated into expression vectors that are suitable for introduction and expression in eukaryotic (*e.g.*, insect or mammalian cells). Favorably, such nucleic acids are codon optimized for expression in the selected vector/host cell. Selected cells can be clonally expanded and characterized for expression of the desired protein construct. Techniques for codon optimization are known in the art. In addition, commercial molecular biology service providers offer codon optimization, among other routine technical services.

Prokaryotic

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the expressed product. For example, when large quantities of a polypeptide or fragments thereof are needed for the production of antibodies, vectors which direct high level expression of fusion proteins that are readily purified are favorably employed. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the coding sequence of interest, *e.g.*, a polynucleotide of the invention as described herein, can be ligated into the vector in-frame with sequences for the amino-terminal translation initiating Methionine and the subsequent 7 residues of beta-galactosidase producing a catalytically active beta galactosidase fusion protein; pIN vectors (Van Heeke & Schuster (1989) *J Biol Chem* 264:5503-5509); pET vectors (Novagen, Madison WI), in

which the amino-terminal methionine is ligated in frame with a histidine tag; and the like. In some embodiments, a suitable vector is pET19; in others, pET24; in others, pET26.

Eukaryotic

Similarly, in yeast, such as *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH can be used for production of the desired expression products. For reviews, *see* Berger, Ausubel, and, *e.g.*, Grant *et al.* (1987; *Methods in Enzymology* 153:516-544). In mammalian host cells, a number of expression systems, including both plasmids and viral-based systems, can be utilized.

IMMUNOGENIC COMPOSITIONS

Pharmaceutically Acceptable Carriers or Excipients

Pharmaceutically acceptable carriers and excipients are well known and can be selected by those of skill in the art. For example, the carrier or excipient can favorably include a buffer. Optionally, the carrier or excipient also contains at least one component that stabilizes solubility and/or stability. Examples of solubilizing/stabilizing agents include detergents, for example, lauryl sarcosine and/or tween. Alternative solubilizing/stabilizing agents include arginine, and glass forming polyols (such as sucrose, trehalose and the like). Numerous pharmaceutically acceptable carriers and/or pharmaceutically acceptable excipients are known in the art and are described, *e.g.*, in *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 5th Edition (975).

Accordingly, suitable excipients and carriers can be selected by those of skill in the art to produce a formulation suitable for delivery to a subject by a selected route of administration.

Suitable excipients include, without limitation: glycerol, Polyethylene glycol (PEG), Sorbitol, Trehalose, N-lauroylsarcosine sodium salt, L –proline, Non detergent sulfobetaine, Guanidine hydrochloride, Urea, Trimethylamine oxide, KCl, Ca^{2+}, Mg^{2+} , Mn^{2+} , Zn^{2+} and other divalent cation related salts, Dithiothreitol, Dithioerytrol, and ß-mercaptoethanol. Other excipients can be detergents (including: Tween80, Tween20,

Triton X-00, NP-40, Empigen BB, Octylglucoside, Lauroyl maltoside, Zwittergent 3-08, Zwittergent 3-0, Zwittergent 3-2, Zwittergent 3-4, Zwittergent 3-6, CHAPS, Sodium deoxycholate, Sodium dodecyl sulphate, Cetyltrimethylammonium bromide).

Adjuvants

Optionally, the immunogenic compositions also include an adjuvant. In the context of an immunogenic composition suitable for administration to a subject, such as a human subject, for the purpose of eliciting an immune response against CASB7439, the adjuvant is selected to elicit a Th1 biased immune response. The adjuvant is typically selected to enhance a Th1 biased immune response in the subject, or population of subjects, to whom the composition is administered.

Th1 Immune Response

A "Th1" type immune response is characterized by the induction of CD4+ T helper cells that produce IL-2 and IFN-γ. In contrast, a "Th2" type immune response is characterized by the induction of CD4+ helper cells that produce IL-4, IL-5, and IL-13.

TLR affectors, including without limitation 3D-MPL

One suitable adjuvant for use in combination with modified CASB7439 polypeptides is a TLR-4-modulator. One example is a non-toxic derivative of lipid A, such as monophosphoryl lipid A or more particularly 3-Deacylated monophoshoryl lipid A (3D–MPL). 3D-MPL is sold under the name MPL by GlaxoSmithKline Biologicals S.A., and is referred throughout the document as MPL or 3D-MPL. See, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094. 3D-MPL primarily promotes CD4+ T cell responses with an IFN-γ (Th1) phenotype. 3D-MPL can be produced according to the methods disclosed in GB2220211 A. Chemically it is a mixture of 3-deacylated monophosphoryl lipid A with 3, 4, 5 or 6 acylated chains. In the compositions of the present invention small particle 3D-MPL can be used. Small particle 3D-MPL has a particle size such that it can be sterile-filtered through a 0.22μm filter. Such preparations are described in WO94/21292.

In other embodiments the lipopolysaccharide can be a β(1-6) glucosamine disaccharide, as described in US Patent No. 6,005,099 and EP Patent No. 0 729 473 B1. One of skill in the art would be readily able to produce various lipopolysaccharides, such as 3D-MPL, based on the teachings of these references. In addition to the aforementioned immunostimulants (that are similar in structure to that of LPS or MPL or 3D-MPL), acylated monosaccharide and disaccharide derivatives that are a sub-portion to the herein structure of MPL are also suitable adjuvants. In other embodiments the adjuvant is a synthetic derivative of lipid A, some of which are described as TLR-4 agonists, and include, but are not limited to: OM174 (2-deoxy-6-o-[2-deoxy-2-[(R)-3-dodecanoyloxytetra-decanoylamino]-4-o-phosphono-β-D-glucopyranosyl]-2-[(R)-3-hydroxytetradecanoylamino]-α-D-glucopyranosyldihydrogenphosphate), (WO 95/14026); OM 294 DP (3S, 9 R) –3--[(R)-dodecanoyloxytetradecanoylamino]-4-oxo-5-aza-9(R)-[(R)-3-hydroxytetradecanoylamino]decan-1,10-diol,1,10-bis(dihydrogenophosphate) (WO 99/64301 and WO 00/0462); and OM 197 MP-Ac DP (3S-, 9R) -3-[(R) -dodecanoyloxytetradecanoylamino]-4-oxo-5-aza-9-[(R)-3-hydroxytetradecanoylamino] decan-1,10-diol,1 -dihydrogenophosphate 10-(6-aminohexanoate) (WO 01/46127).

Other TLR-4 ligands which can be used are alkyl Glucosaminide phosphates (AGPs) such as those disclosed in WO 98/50399 or US Patent No. 6,303,347 (processes for preparation of AGPs are also disclosed), suitably RC527 or RC529 or pharmaceutically acceptable salts of AGPs as disclosed in US Patent No. 6,764,840. Some AGPs are TLR-4 agonists, and some are TLR-4 antagonists. Both are thought to be useful as adjuvants.

Other suitable TLR-4 ligands, capable of causing a signaling response through TLR-4 (Sabroe et al, JI 2003 p1630-5) are, for example, lipopolysaccharide from gram-negative bacteria and its derivatives, or fragments thereof, in particular a non-toxic derivative of LPS (such as 3D-MPL). Other suitable TLR agonists are: heat shock protein (HSP) 10, 60, 65, 70, 75 or 90; surfactant Protein A, hyaluronan oligosaccharides, heparan sulphate fragments, fibronectin fragments, fibrinogen peptides, b-defensin-2, and muramyl dipeptide (MDP). In one embodiment the TLR agonist is HSP 60, 70 or 90. Other suitable TLR-4 ligands are as described in WO 2003/011223 and in WO 2003/099195, such as compound I, compound II and

compound III disclosed on pages 4-5 of WO2003/011223 or on pages 3-4 of WO2003/099195 and in particular those compounds disclosed in WO2003/011223 as ER803022, ER803058, ER803732, ER804053, ER804057, ER804058, ER804059, ER804442, ER804680, and ER804764. For example, one suitable TLR-4 ligand is ER804057.

Saponin Adjuvants

Other adjuvants that can be used in immunogenic compositions with CASB7439, e.g., on their own or in combination with 3D-MPL, or another adjuvant described herein, are saponins, such as QS21.

Saponins are taught in: Lacaille-Dubois, M and Wagner H. (1996. A review of the biological and pharmacological activities of saponins. Phytomedicine vol 2 pp 363-386). Saponins are steroid or triterpene glycosides widely distributed in the plant and marine animal kingdoms. Saponins are noted for forming colloidal solutions in water which foam on shaking, and for precipitating cholesterol. When saponins are near cell membranes they create pore-like structures in the membrane which cause the membrane to burst. Haemolysis of erythrocytes is an example of this phenomenon, which is a property of certain, but not all, saponins.

Saponins, particularly immunologically active saponin fractions, are known as adjuvants in vaccines for systemic administration. The adjuvant and haemolytic activity of individual saponins has been extensively studied in the art (Lacaille-Dubois and Wagner, supra). For example, Quil A (derived from the bark of the South American tree Quillaja Saponaria Molina), and fractions thereof, are described in US 5,057,540 and "Saponins as vaccine adjuvants", Kensil, C. R., Crit Rev Ther Drug Carrier Syst, 1996, 12 (1-2):1-55; and EP 0 362 279 B1. Particulate structures, termed Immune Stimulating Complexes (ISCOMS), comprising fractions of Quil A are haemolytic and have been used in the manufacture of vaccines (Morein, B., EP 0 109 942 B1; WO 96/11711; WO 96/33739). The haemolytic saponins QS21 and QS17 (HPLC purified fractions of Quil A) have been described as potent systemic adjuvants, and the method of their production is disclosed in US Patent No.5,057,540 and EP 0 362 279 B1. Other saponins which have been used in systemic vaccination studies include those derived

from other plant species such as Gypsophila and Saponaria (Bomford et al., Vaccine, 10(9):572-577, 1992).

Such formulations comprising QS21 and cholesterol have been shown to be successful Th1 stimulating adjuvants when formulated together with an antigen. Thus, for example, CASB7439 can favorably be employed in immunogenic compositions with an adjuvant comprising a combination of QS21 and cholesterol.

Immunostimmulatory Oligonucleic Acids

One suitable adjuvant for use in combination with CASB7439 is a bacterial DNA TLR agonist capable of causing a signaling response through TLR-9, i.e., a TLR-9 agonist, more specifically DNA containing unmethylated CpG nucleotides, in particular sequence contexts known as CpG motifs. CpG-containing oligonucleotides induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Suitably, CpG nucleotides are CpG oligonucleotides. Suitable oligonucleotides for use in the immunogenic compositions of the present invention are CpG containing oligonucleotides, optionally containing two or more dinucleotide CpG motifs separated by at least three, suitably at least six or more nucleotides. A CpG motif is a Cytosine nucleotide followed by a Guanine nucleotide. The CpG oligonucleotides of the present invention are typically deoxynucleotides. In a specific embodiment the internucleotide in the oligonucleotide is phosphorodithioate, or suitably a phosphorothioate bond, although phosphodiester and other internucleotide bonds are within the scope of the invention. In some embodiments are oligonucleotides with mixed internucleotide linkages. Methods for producing phosphorothioate oligonucleotides or phosphorodithioate are described in US Patent Nos. 5,666,153, 5,278,302 and WO 95/26204.

In an alternative embodiment, a TLR agonist is used that is capable of causing a signaling response through TLR-9. In one embodiment, the TLR agonist capable of causing a signaling response through TLR-9 is HSP90. Combinations of different adjuvants, such as those mentioned herein, can also be used in compositions with CASB7439. For example, as already noted, QS21 can be formulated together with 3D-

MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; such as 1:5 to 5 : 1, and often substantially 1 : 1. Typically, the ratio is in the range of 2.5 : 1 to 1 : 1 3D-MPL: QS21. Another combination adjuvant is 3DMPL plus QS21 in a liposomal formulation, also in combination with CpG.

Subjects with Colorectal Cancer

Selection of a subject with colorectal cancer for participation in the methods described herein can be by clinical diagnosis, molecular diagnosis, or a combination thereof. In some embodiments a subject is tested to determine if their cancerous cells express CASB7439 (HASH2). Clinical methods for diagnosing colorectal cancer are known in the art. Molecular techniques and methods for determining whether a cancerous cell or tissue expresses CASB7439 (HASH2) are disclosed in WO01/62778.

EXPERIMENTAL EXAMPLES

Example 1

Two nucleic acid molecules encoding protein constructs comprising a CASB7439 polypeptide were generated in the pET19b vector. A His-tag was juxtaposed to the N-terminal end of the CASB7439 sequence. These clones were termed LVL007 [SEQ ID NO:15(aa); SEQ ID NO:16 (DNA)] and LVL010 [SEQ ID NO:17 (aa); SEQ ID NO:18 (DNA)], respectively). It was observed that an unintended mutation had occurred in the LVL007 clone, resulting in a one amino substitution of proline to leucine (at amino acid residue 17 with reference to SEQ ID NO:13). The nucleic acid molecules were utilized with three different host cell types, namely the *E. coli* strains BLR DE3, BL21 DE3, and Rosetta DE3. Inductions were carried out at two temperatures per induction (16°C or 37°C) and with 1mM of IPTG. See Materials and Methods section, below, for protocol details.

Proteins were harvested for each induction and analyzed by Western blot and by SDS-PAGE with Coomassie-blue staining. In these experiments, the detection of the protein was visible only in western blot.

Example 2

Bioinformatic analyses of the CASB7439 gene were carried out, revealing the following structural features:

- a hairpin RNA structural element located between nucleotide 2 and 27 of SEQ ID NO:14
- 62% GC content of the nucleotide sequence
- nucleotide sequence rich in Arginine codons (26/193 amino acids encoded)
- amino acid sequence with a high level of basic amino acids residues
- amino acid sequence with a high isoelectric point (11.18)
- a low complexity region within the amino acid sequence (residues 7-27 of SEQ ID NO:13
- helix-loop-helix domain (HLH) (IPR000014) within the amino acid sequence (residues 56-108 of SEQ ID NO:13)
- a first region of intrinsic disorder within the amino acid sequence (residues 118-164 of SEQ ID NO:13)
- a second region of intrinsic disorder within the amino acid sequence (residues 118-164 of SEQ ID NO:13)

Applicants also identified a region with a high proline periodicity from amino acid 127 to 158. Applicants designated the region from amino acid 133-153, inclusive, of SEQ ID NO:13 (aa) the proline-rich region. The proline-righ region, along with some of the other features mentioned in this example are depicted in schematic fashion in Figure 1/35, along with regions mentioned elsewhere herein.

Example 3

The unmodified CASB7439 DNA nucleotide sequence (SEQ ID NO:14) was codon-optimized by the commercial DNA synthesis services provider, Geneart (unique identifier 0606597), and cloned. Geneart AG, BioPark, Josef-Engert-Str. 11D-93053 Regensburg, Germany. The CASB7439 nucleotide sequence was obtained from the clone by PCR amplification and used to generate two constructs. In the first construct, LVL016 [SEQ ID NO:29 (aa), SEQ ID NO:30 (DNA)] the CASB7439 sequence was inserted in the pET21b (+) vector in fusion with the pD1/3 protein (from *Haemophilus*

influenzae) at the N-terminal end of the CASB7439 sequence. In the second construct, LVL018 [SEQ ID NO:31 (aa); SEQ ID NO:32 (DNA)], the CASB7439 was inserted in the pET21b (+) vector (without any Protein D component). These constructs were evaluated for production in the BLR DE3 *E. coli* at 16ºC and 37ºC with 1mM of IPTG. Production was detectable by Western blot as an insoluble protein, mainly produced at 16ºC.

Four of the CASB7439 domains were chosen for further investigation to improve protein production: the nucleus targeting domain, the proline-rich region, the DNA binding domain, and the HLH domain.

Example 4

The nucleus targeting domain, the DNA binding domain, the bHLH domain and the proline-rich region of the CASB7439 were chosen to investigate protein production from constructs comprising nucleotide sequences encoding CASB7439 polypeptides. Different sections of the CASB7439 nucleotide sequence were PCR amplified from the codon-optimized Geneart clone described herein and inserted in the pET19b (+) vector, resulting in five constructs having a different CASB7439 sequence (truncated in some fashion or complete); each construct with a polyhistidine tail at the N-terminal of the CASB7439 sequence. See Figure 2/35.

Those constructs shown on Figure 2/35 were evaluated for protein expression in BLR DE3 *E.coli* strain at 37°C and 16°C with 1mM of IPTG. As depicted in Figure 2/35, these constructs comprised full-length CASB7439 [LVL060, SEQ ID NO:19 (aa), SEQ ID NO:20 (DNA)]; CASB7439 truncated to remove the nucleus targeting domain [Const-1, SEQ ID NO:21 (aa), SEQ ID NO:22 (DNA)]; CASB7439 truncated to remove the proline-rich region [LVL055, SEQ ID NO:1 (aa), SEQ ID NO:2 (DNA)]; CASB7439 truncated to remove the DNA binding domain [LVL056, SEQ ID NO:23 (aa), SEQ ID NO:24 (DNA)]; CASB7439 truncated to leave the basic helix-loop-helix (bHLH) domain [LVL057, SEQ ID NO:25 (aa), SEQ ID NO:26 (DNA)]. Due to experimental error, Const-1 lacked a stop codon, resulting in the translation of a portion of the vector nucleotide sequence (resulting with an additional 21 amino acids in the C-terminal region, starting with LED...). The production of LVL055 following incubation at 16°C (C-

terminal truncated) was clearly detectable on a Coomassie-blue stained SDS-PAGE gel.

Example 5

Based on the results described in Example 4, new constructs were designed to (i) remove the proline-rich region while (ii) retaining other regions of the C-terminal portion of the polypeptide. See the description of Figures 3/35 and 4/35, which discuss (i) various identified or predicted epitopes within the polypeptide and (ii) various alternate approaches for removal of the proline-rich region.

A codon optimized synthetic nucleic acid gene sequence coding for a modified CASB7439 polypeptide was then ordered from Geneart (unique identifier 0706840), digested and inserted in parallel in two different vectors to generate the following constructs:

(i) LVL088, comprising the CASB7439 gene with the 21 amino acid proline-rich region deleted, fused with a polyhistidine-tail at the C-terminal in the pET26b (+) vector (SEQ ID NO:27 (aa), SEQ ID NO:28 (DNA)) and

(ii) LVL111, comprising the CASB7439 gene with the 21 amino acid proline-rich region deleted and both an N-terminal- and C-terminal-polyhistidine-tail in the pET19b (+) vector (SEQ ID NO:3 (aa), SEQ ID NO:4 (DNA)).

The LVL111 construct (see SEQ ID NO:3) was cloned following the same strategy used for the LVL055, i.e., using the pET19 (+) vector which includes a polyhistidine-tail composed of 10 histidines followed by the enterokinase site. This resulted in a 23 amino acid fusion partner (SEQ ID NO:42) at the N-terminus.

Other constructs were generated with the codon optimized synthetic gene discussed in Example 3 (unique identifier 0606597), including the following:

- LVL090 (a protein construct comprising (from the N-terminus) Protein D, full-length optimized CASB7439, and a 6-His in pET21 vector) (SEQ ID NO:43 (aa); SEQ ID NO:44 (DNA))
- LVL112 (a protein construct comprising (from the N-terminus) SUMO Protein, full-length optimized CASB7439 in a pSUMO vector) (SEQ ID NO:45 (aa); SEQ ID NO:46 (DNA))

- LVL113 (a protein construct comprising (from the N-terminus) SUMO Protein, pD1/3 protein, full-length optimized CASB7439 in a pSUMO vector) (SEQ ID NO:47 (aa); SEQ ID NO:48 (DNA))
- LVL114 (a protein construct comprising (from the N-terminus) SUMO Protein, CASB7439 truncated at amino acid 117 in a pSUMO vector) (SEQ ID NO:49 (aa); SEQ ID NO:50 (DNA))
- LVL115 (a protein construct comprising (from the N-terminus) SUMO Protein, modified CASB7439 with the proline-rich region removed in a pSUMO vector) (SEQ ID NO:51 (aa); SEQ ID NO:52 (DNA))

No production improvement was observed, although only one repetition was carried out (Data not shown).

Production of LVL088 was analyzed after an overnight induction with 1mM IPTG at 37°C and at 16°C. When analyzed electrophoretically, LVL088 production was visualized only by Western blot. One purification lot from four 800 mL culture flasks resulted in 0.4 mg of protein. On the other hand, LVL111 was easily detectable following Coomassie-blue staining. The recombinant protein LVL111 was produced and purified over 8 lots. Several lots of LVL111 (each from four 800 mL culture flasks) were estimated to produce in the range from 3.3 mg/lot to 10 mg/lot of purified protein. Thus, production of LVL088 was lower than that seen for LVL111. The purified protein was also tested for its reactivity with several anti-CASB7439 monoclonal antibodies and a positive reactivity was obtained in all cases.

Example 6

Removal of the C-terminal polyhistidine-tail from the LVL111 construct was investigated. The nucleic acid molecule encoding the polypro (-) CASB7439 polypeptide without the portion encoding the C-terminal polyhistidine was amplified by PCR from the LVL111 construct and cloned into the pET19b vector. This construct was designated LVL137. The LVL137 construct was essentially LVL111 without the nucleotide sequence encoding a C-terminal polyhistidine tail. SEQ ID NO:5 (aa); SEQ ID NO:6 (DNA). The removal of the C-terminal polyhistidine-tail had no apparent impact

on the production level. As analyzed by coomasie-stained SDS-PAGE gel, LVL137 produced the same level of protein as LVL111.

The nucleic acid molecule encoding the polypro (-) CASB7439 polypeptide was then amplified by PCR from the Geneart (0706840) synthetic CASB7439 sequence and cloned into a pET26 vector with a six polyhistidine tag at the N-terminal (LVL138). SEQ ID NO:33 (aa); SEQ ID NO:34 (DNA). This construct exhibited a significantly reduced level of protein production and was difficult to detect by Western blot.

A nucleic acid molecule encoding a six polyhistidine tag at the N-terminal of the polypro (-) CASB7439 polypeptide was generated in a pET24 kanamycin resistant vector. A new linker design was utilized to remove a hairpin structure created between the polyhistidine-tag sequence and the 5' end of the CASB7439 gene. Specifically, the codons for the first eight amino acid residues of the protein construct were altered to the following nucleotide sequence: 5' ATG CAT CAT CAT CAT CAT CAT GAC… 3' (SEQ ID NO:35). (The original codons were 5' ATG CAC CAT CAC CAT CAC CAT GAT… 3' (SEQ ID NO:36).) This construct was designated LVL160. SEQ ID NO:37 (aa); SEQ ID NO:38 (DNA). By gel and Western blot analysis, protein production appeared as weak as that obtained with LVL138.

A nucleic acid molecule encoding a protein construct with a ten polyhistidine tag at the N-terminal of the polypro (-) CASB7439 polypeptide was generated in a pET24 kanamycin resistant vector. This construct was identified as LVL168. SEQ ID NO:11 (aa); SEQ ID NO:12 (DNA). The production level was comparable to the level observed for LVL111. An alignment of several of the constructs discussed in this Example is found in Figure 5/35.

The results for the constructs discussed in this Example and Example 1 are summarized in Table 1, below.

# LVL	Description	Vector	Production
LVL007	Long His-CASB7439	pET19b(+)	gel-/blot+
LVL010	Long His-CASB7439	pET19b(+)	gel-/blot+
LVL016	pD1/3-CASB7439-OPT-6His	pET21b(+)	gel-/blot +
LVL111	Long His-Del-poly-Pro-domain-His-OPT	pET19b(+)	gel+++/ blot +++

# LVL	Description	Vector	Production
LVL137	Long His-Del-poly-Pro-domain-OPT	pET19b(+)	gel+++/blot +++
LVL138	6His-Del-poly-Pro-domain-OPT	pET26b(+)	gel-/ blot +
LVL160	Hismodif--Del-poly-pro-domain-OPT	pET24b(+)	gel-/blot +
LVL168	10His-del-poly-pro-domain-OPT	pET24b(+)	gel+++/ blot +++

Table 1. Expression levels for constructs discussed in Examples 1, 3, and 6. A construct that results in " gel-/blot+" production generally is visible only by western blot when whole-cell lysate is subjected to SDS-PAGE. A construct that results in " gel+++/blot +++" production is clearly visible by Coomase blue staining when whole-cell lysate is subjected to SDS-PAGE.
6His: Histidine tag = 6 histidine residues
Long His = 23aa (MGHHHHHHHHHHSSGHIDDDDKH) (SEQ ID NO:42)
Del-poly-Pro-domain: CASB7439 modified by deletion of the proline-rich region,
OPT: codon optimized
Hismodif: Histidine tag modified for codon selection (see herein, Example 6),
1/3pD: protein D construct described herein
10His: Histidine tag with 10 histidine residues

Example 7

Another CASB7439 construct was generated in fusion with a fragment of protein D from *Haemophilus influenzae*. As shown in SEQ ID NO:39 (aa) (SEQ ID NO:40 (DNA)), amino acids Met-Asp-Pro were fused to 108 N-terminal amino acids of processed Protein D (amino acid residues 2-109 of processed Protein D). SEQ ID NO:39 is referred to herein as 1/3pD or 1/3 Protein D. The synthetic CASB7439 nucleotide sequence obtained from Geneart (0706840) (proline-rich region deleted) was subcloned in pET26-1/3pD, a kanamycin resistant plasmid which contains the 1/3pD, to obtain LVL141. The sequence of LVL141 is set forth in SEQ ID NO:7 (aa) and SEQ ID NO:8 (DNA). Production was comparable to LVL111.

This clone was then mutated by PCR to remove the four amino acid linker (Ala-Ala-Ala-His) between the 1/3pD and CASB7439 components. The sequence of the resulting construct, designated LVL144, is set forth in SEQ ID NO:9 (aa) and SEQ ID NO:10 (DNA). Production was comparable to LVL111. Figure 6/35 shows an alignment of CASB7439 (SEQ ID NO:13); LVL168 (SEQ ID NO:12), PD1/3 (SEQ ID NO:39), and LVL144 (SEQ ID NO:9).

Example 8

Solubilization tests of LVL168

Generally, purified protein constructs comprising CASB7439 were solublized in 6M guanidine buffer, transferred to 8M urea buffer, and stored in 4M Urea, 20mM HEPES, 1mM TCEP, 150mM NaCl, pH7.0. Preliminary results performed in small volumes suggested that LVL168 could also be solubilized in a phosphate buffer without urea nor any reducing agents, such as EDTA or TCEP, and with or without NaCl. No precipitation was observed after dialysis in those buffers. Solubilization assays were performed to characterize LVL168 protein in the sixteen buffers listed in Table 2. One hundred microliters of purified LVL168 protein samples in 8M Urea buffer, 20mM HEPES, 150mM NaCl, 1mM TCEP, pH7.0 were put into mini dialysis unit 10000 MWCO (Biolynx), previously conditioned with the buffer by two rinses. Dialysis was performed overnight in 100mL of each of the sixteen buffers listed in Table 2 under agitation at 4°C. Samples were then harvested from units, put into eppendorfs and centrifuged for 5 minutes at 20,000g. Soluble proteins stayed in suspension but non-soluble proteins precipitated and lead to pellets after centrifugation. Buffers in which the LVL168 protein was soluble are listed in Table 2.

Supernatants were tested by gel SDS-PAGE and the concentrations of the remaining soluble proteins were measured by the RC DC method using the manufacturer's instructions. A kit for carrying out the RC DC method of the Modifiied Lowry method is available from BioRad Laboratories, 1000 Alfred Nobel Drive, Hercules, CA 94547). Samples were kept at 4°C for a stability test.

Stability tests of LVL168

LVL168 proteins from the solubility assay were kept for 7 days at 4°C. Twenty microliters of each of the samples 2, 3, 5, 6, 9 to 16 were transferred to other tubes and tested for their stability at 4°C, 37°C or room temperature for 2 additional days. Samples were centrifuged for 5 minutes at 20,000g. Supernatants were resolved by SDS-PAGE and profiles were analyzed for degradation, aggregation or loss of protein by comparison to the original purified protein. The protein was stable in the majority of the buffers tested.

Sample #	Buffer	Solubility
1	PBS 1X pH7.3	Insoluble
2	2M Urea, 20mM Tris, 150mM NaCl + 1mM TCEP, pH 8.0	Soluble
3	4M Urea, 20 mM Tris, 150mM NaCl + 1mM TCEP, pH 7.0	Soluble
4	PBS 1X, 1mMTCEP, 1mM EDTA, pH7.3	Insoluble
5	10mM NaAc 1mM EDTA. 1mM, TCEP, 5mM NaCl, pH 5.01	Soluble
6	20mM Imidazol, 150mM, NaCl, 1mM EDTA, 1mM TCEP, pH 6.0	Soluble
7	20mM Bicine, 138 mM NaCl, 1mM EDTA, 1 mM TCEP, pH 8.7	Insoluble
8	50 mM Tris, 250 mM NaCl, 27 mM Sucrose, 0.1% Tween80, 1mMTCEP, pH 8.5	Insoluble
9	50mM Tris, 250mM NaCl, 270mM Sucrose, 0.3ML-A rginine, 1mM TCEP, pH 8.5	Soluble
10	20mM Na_2HPO_4, 150mM NaCl, 1mM EDTA + 1mM TCEP, pH 6.0	Soluble
11	20mM $N3_2HPO_4$, 5Mm NaCl, 1mM EDTA + 1Mm TCEP, pH 6.0	Soluble
12	20mM Na_2HPO_4, 5mM NaCl, 1mM EDTA pH 6.0	Soluble
13.	20mM Na_2HPO_4, 5mM NaCl + 1mM TCEP + pH 6.0	Soluble
14	20mM Na_2HPO_4, 5mM NaCl pH 6.0	Soluble
15	20mM Na_2HPO_4, 5mM NaCl pH 6.5	Soluble
16	20mM Na_2HPO_4, 150mM NaCl pH 6.0	Soluble

Table 2. Solubility and stability tests of LVL168.

For the reader's convenience, a summary of the constructs mentioned herein and their relative production is provided in Table 3.

# LVL	Description	Vector	Production
LVL007	Long His-CASB7439	pET19b(+)	gel-/blot+
LVL010	Long His-CASB7439	pET19b(+)	gel-:blot+
LVL016	pD1/3-CASB7439-OPT-6His	pET21b(+)	gel-/blot +
LVL018	CASB7439-OPT-6His	pET21b(+)	gel-/blot +
Const-1	Long His-CASB7439-OPT-Const-1 (nucleus targeting trunc.)	pET19b(+)	gel-/ blot +
LVL055	Long His-poly-Pro-domain-trunc-OPT	pET19b(+)	gel+++/blot +++
LVL056	Long His-DNA bind trunc-OPT	pET19b(+)	gel-/blot -
LVL057	Long His-bHLH alone-OPT	pET19b(+)	gel-/blot -
LVL060	Long His-C7439-OPT	pET19b(+)	gel-/blot +
LVL088	Del-poly-Pro-domain-OPT-6His	pET26b(+)	gel-/blot +
LVL111	Long His-Del-poly-Pro-domain-His-OPT	pET19b(+)	gel+++/ blot +++
LVL137	Long His-Del-poly-Pro-domain-OPT	pET19b(+)	gel+++/blot +++
LVL138	6His-Del-poly-Pro-domain-OPT	pET26b(+)	gel-/ blot +
LVL141	1/3pD-Del-poly-Pro-domain-OPT-6His	pET26b(+)	gel+++/blot +++
LVL144	1/3pD-Del-poly-Pro-domain-OPTHis mut linker	pET26b(+)	gel+++/ blot +++
LVL160	Hismodif--Del-poly-pro-domain-OPT	pET24b(+)	gel-/blot +
LVL168	10His-del-poly-pro-domain-OPT	pET24b(+)	gel+++/ blot +++

Table 3. See Table 1 abbreviations. The following additional abbreviations are utilized in Table 3:
trunc: truncated
const: construction
mut: sequence mutated to remove unrelated amino acids between the pD1/3 and modified CASB7439 polypeptide
linker: unrelated amino acids between two part of the molecule
DNA bind: DNA binding domain
bHLH alone: Basic helix-loop-helix domain

Materials and Methods

Expression - Small scale inductions:

E. coli BLR (DE3) strains were transformed with constructions mentioned herein, put on Luria Broth (LB) agar plates containing the appropriate antibiotic (40 µg/ml of Kanamycin or 100 µg/ml carbenicilin) and incubated overnight at 37°C. The resulting bacterial lawn was used to inoculate 20mL of LB Alternative Protein Source (APS) medium with the appropriate antibiotic (40 µg/ml of Kanamycin or 100µg per ml of carbenicilin) and put at 37°C to reach an OD600 of 0.5 to 1.0.

Inductions were done by the addition of 1mM IPTG, at either 16°C or 37°C. In order to induce at 16 °C, the growing culture was cooled down on ice 1hr before inducing the expression of the recombinant protein. The culture was then put back to grow at 16 °C and maintained under those conditions for 16hrs. To induce at 37°C, expression of the recombinant proteins was immediately induced by the addition of IPTG to the growing culture. Induction was maintained at 37 °C for 3 hours. Small aliquots of 1mL were taken before and after induction.

Aliquots were then centrifuged (10 min, 4°C, 20000 x g) and bacterial pellets were kept at -20 °C until analysis. The analysis was performed following a BugBuster protein extraction according to the manufacturer's recommendations. BugBuster, manufactured by Novagen, is commercially available from VWR. Expression level was determined by Coomassie-blue staining of SDS-PAGE gels and by Western blots. The solubility was evaluated after detection of the protein in the pellets or in the supernatants from the BugBuster's extraction.

Expression - Large scale inductions:

E. coli BLR (DE3) strains were transformed with constructions as described herein, put on LB agar plates containing the appropriate antibiotic and incubated overnight at 37°C. The resulting bacterial lawn was used to inoculate the desired number of 800 mL flasks of LB APS medium with the appropriate antibiotic (40 µg/ml of Kanamycin or 100µg per ml of carbenicilin) and put at 37°C to reach an OD_{600} of 0.5 to 1.0. The 2.5L flasks containing 800ml of growing culture were then cooled down on ice 1hr before inducing the expression of the recombinant protein by the addition of 1mM

IPTG. To induce at 16ºC, the growing culture was cooled down on ice one hour before induction. The cultures were then put back to grow at 16°C and maintained under those conditions for 16 hours.

Cultures were centrifuged (15 min, 4°C, 6000 x g) and bacterial pellets were kept at -80 °C until purification process. Small aliquots of 1mL were also taken before and after induction for BugBuster and analysed by SDS-PAGE and Western blots before purification.

Purification

Bacterial pellets were resuspended in Lysis buffer (20 mM Tris buffer (pH 8.0) containing 500 mM NaCl, 10mM TCEP) and a mixture of protease inhibitors (Complete EDTA-free). Bacteria were lysed either by sonication using a 13mm probe on a Vibra cell ultrasonic processor, by Emulsiflex C3 (Avestin) or with the Constant Cell disruption system (Constant System). Soluble (supernatant) and insoluble (pellet) components were separated by centrifugation at 20 000g for 20 minutes at 4°C. Because CASB7439 recombinant proteins are expected in the pellets, supernatants were discarded.

Usually, the insoluble components (pellets) were resolubilized in 50mM bicine buffer pH 8.0 containing 6M guanidine HCl, 500mM NaCl, 10mM TCEP, and then centrifuged (20 000 x g for 20 min). The supernatant was loaded on a 5ml His Trap column (GE Healthcare) already preequilibrated in the previous buffer but with TCEP concentration lowered to 1mM. After loading, the column was washed once with the same buffer. The 6M GnHCl chaotrope agent was replaced by 8M Urea (8M Urea, 20mM HEPES, 500mM NaCl, 1mM TCEP, pH 8.0). Elution was performed using a 20mM HEPES buffer (pH 8.0) containing 8M Urea, 500mM NaCl, 1mM TCEP and 250 mM imidazole. Samples from elution fractions were analyzed by SDS-PAGE, and fractions containing the protein of interest were pooled, concentrated on a centrifugal device if needed, and used for the second step of purification.

One or two steps of SEC (size exclusion chromatography) (Superdex 200, GE Healthcare) were done, depending of the purity needed, in 8M Urea, 20mM HEPES, 500 mM NaCl, 1mM TCEP, pH 8,0, before the last desalting step. The desalting was

done on a G25 desalting column preequilibrated with the final 20mM HEPES buffer (pH 7.0) containing 4M Urea, 150mM NaCl and 1mM TCEP.

Protein samples were 1ml-aliquoted in tubes and stored at – 80°C.

Protein analysis—concentration determination

Protein concentration was determined using an RC DC protein assay (Modified Lowry method by BioRad). This assay was used due to the incompatibility of some components of the final buffer with classic Lowry or BCA assays (high contents of urea and reducer agents). Known quantities of final protein preparation were also analyzed by SDS-PAGE, Western blot and for LAL-content to assess aggregation or degradation, *E. coli* protein and LPS contamination level.

Example 9

CASB7439 Expression Validation in Normal and colorectal cancer (CRC) Tissues

Applicants performed real-time qPCR analysis of CASB7439 mRNA expression in 83 CRC samples and 12 normal colon samples obtained from Corixa Corporation. CASB7439 mRNA was at least 10-fold overexpressed in CRC compared to normal colon in 63% of CRC samples tested compared to normal tissues (data reported as average CASB7439/β-actin expression ratio) (Figure 7/35). In addition, Applicants detected CASB7439 protein by immunofluorescence (IF) in human CRC tissues using a rabbit anti-CASB7439 monoclonal antibody (40-12) (antibody discussed elsewhere herein); no CASB7439 protein was detected in normal adjacent tissue sections.

Additional expression data for CASB7439 were generated by GeneLogic (Gene Microarray) analysis using standard techniques. CASB7439 mRNA was found to be highly expressed in CRC samples from all stages of the disease (data not shown). Applicants performed real-time qPCR analysis of CASB7439 mRNA expression in 83 CRC samples and 12 normal colon samples obtained from Corixa Corporation. CASB7439 mRNA was at least 10-fold overexpressed in CRC compared to normal colon in 63% of CRC samples tested compared to normal tissues (data reported as average CASB7439/β-actin expression ratio) (Figure7/35).

Independent literature reports detection of CASB7439, i.e., HASH-2, mRNA by *in situ* hybridization (ISH) at a very low level at the bottom of colonic crypts but not in any other normal tissues except placenta. Jubb et al. (2006) *Oncogene* 25: 3445-3457. In colorectal cancer (CRC) tissues, CASB7439 mRNA has been shown to be 15-fold overexpressed in 70% of human tissues tested.

CASB7439 Protein Expression in CRC and Normal Adjacent Colon Tissues

In order to analyze CASB7439 protein expression by immunofluorescence in normal and CRC tissues, a specific rabbit anti-CASB7439 monoclonal antibody (40-12) was developed. Rabbit hybridomas raised against CASB7439 were obtained from the commercial provider Epitomics Inc. and screened by Applicants. The mAbs produced by the hybridomas were first screened by ELISA using the recombinant protein LVL111 (SEQ ID NO:3). The specificity of the mAbs were verified by Western blot using LVL111 and cell lysate obtained from a recombinant cell line engineered to express CASB7439 (see below, TC1/CASB7439). CASB7439 protein expression in TC1/CASB7439 tumors (grown in mice), SW-620 and SW-480 cells (CASB7439 positive human colorectal cancer cell line), and in HCT-116 cells (CASB7439 negative human colorectal cancer cell line) was assessed by immunofluorescence with the mAbs. The mAb 40-12 was selected for CASB7439 protein expression validation.

Applicants obtained from NDRI (National Disease Research Interchange, 8 Penn Center, 8th Floor, 1628 JFK Boulevard, Philadelphia, PA 19103) OCT-embedded frozen tissue samples of both CRC and normal adjacent tissue from 17 different individuals; for 5 of these individuals, metastatic tissues were also obtained. Applicants screened 17 primary tumors, 17 normal adjacent tissues and 5 metastasis. Out of 17 primary CRC tumors analyzed, 13 were positive for CASB7439 protein expression (76.5%). All these CASB7439-positive samples were adenocarcinomas. None of the normal adjacent tissues expressed detectable levels of CASB7439 protein. Of the metastatic tissues, 3 of 5 were positive (60%) for CASB7439 protein expression.

The correlation between the CASB7439 protein and mRNA expression was studied in 16 CRC tissues by real-time qPCR analysis performed using routine techniques on both cancer and normal adjacent colon tissues. Ki67 staining was used

as a proliferation marker. The correlation between CASB7439 protein and mRNA expression is presented in Figure 8/35. Tissue samples which exhibited CASB7439 protein by immunofluorescence also displayed the highest CASB7439 mRNA levels.

Example 10

Immunogenicity of CASB7439 in Mice: CASB7439 T-cell Immunogenicity in Inbred Mice

To study the T-cell immunogenicity of CASB7439 in mice, a bank of peptides was generated by routine synthesis methods to cover the entire sequence of the CASB7439 protein (Figure 9/35). Each peptide was a 15-mer, overlapping the next peptide by 11 amino acids. To investigate the immunodominant regions, 14 pools of peptides were arranged according to a matrix such that any two pools of peptides had only one common peptide (Figure 10/35).

For T-cell immunogenicity studies in mice, up to 15 inbred mice and up to 10 outbred mice per group have received four intramuscular immunizations every two weeks, with various doses (from 0 to 30μg) of various recombinant CASB7439 proteins formulated with AS01B or AS15, both of which are GSK proprietary adjuvants. Once formulated with antigen, AS01B is composed of MPL (100μg/ml) and QS21 (100μg/ml) in liposomes). Once formulated with antigen, AS15 is composed of MPL (100μg/ml), QS21 (100μg/ml) and CpG7909 (840 μg/ml) in liposomes. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X concentration adjuvant. Therefore, each mouse in the AS01B group received 5 μg MPL and 5 μg QS21. Each mouse in the AS15 group received 5 μg MPL, 5 μg QS21, and 5 μg CpG7909. After the last immunization, spleens or PBLs were obtained from the mice. For CASB7439 immunogenicity analysis in inbred mice, spleens or PBLs were analyzed individually or pooled. For CASB7439 immunogenicity analysis in CD1 outbred mice, each spleen was processed individually.

The PBLs or isolated splenocytes were resuspended at a final concentration of $10x10^6$ cells per ml in RPMI 1640 culture medium with the following additives: Penicillin-Streptomycin (1X), non-essential amino acids (1X), 2-Mercaptoethanol (55mM), sodium pyruvate (100mM), L-Glutamine (200mM). Cells have been stimulated with the

peptides from the CASB7439 sequence (used at a final concentration of 1μg/ml/peptide in RPMI 1640 + additives + 5% heat-inactivated FBS). An intracellular staining (ICS) for CD4, CD8, IL-2, IL-5, IFNγ and TNFα has been performed, according to classical procedures. After stimulation, cells were treated with brefeldin A and fixed. Surface staining of CD4 and CD8 was done with CD4 and CD8 specific monoclonal antibodies. Then, cells were permeabilized and stained with specific monoclonal antibodies raised against IL-2, IL-5, IFNγ and TNFα proteins. Cells have been analyzed by flow cytometry using a BD FACSCanto II. Data are expressed in percentages of double-positive IFNγ+ and TNFα+ CD4 or CD8 T-cells.

Inbred Mouse Multistrain T-cell Immunogenicity Comparison Experiment

Applicants compared the immunogenicity of the CASB7439 protein, formulated with either AS01B or AS15, in C57Bl/6, Balb/C, CB6f1 (first generation progeny of C57Bl/6 crossed with Balb/C mice) and C3H inbred mice purchased from Charles River, Senneville, QC, Canada. Five mice per group were immunized with 10 μg of recombinant CASB7439 protein (LVL111) formulated with either AS01B or AS15. The spleens were collected and pooled (one pool of five spleens per group). The CD4 T-cell response (double-positive IFNγ+ and TNFα+) to LVL111 for each strain is presented in Figure 11/35 for the AS01B formulation and in Figure 12/35 for the AS15 formulation. No CD8 T-cell (double-positive IFNγ+ and TNFα+) response was observed in any of the four inbred mouse strains.

The CASB7439 immunodominant peptides identified in inbred mice immunized with CASB7439+AS15 are presented in Figure 13/35. The immunodominant peptides and the with the corresponding regions of SEQ ID NO:13 are:

- C57Bl/6: peptide 39 (SEQ ID NO:91) (aa 153 to 167 of SEQ ID NO:13)
- Balb/c: peptide 24 (aa 93 to 107 of SEQ ID NO:13)
- CB6f1: peptides 7-8 (aa 25 to 43 of SEQ ID NO:13) and 23-24 (aa 89 to 107 of SEQ ID NO:13)
- C3H: peptides 9-11 (aa 33 to 55 of SEQ ID NO:13), 16-19 (aa 61 to 87 of SEQ ID NO:13) and 41-43 (aa 161-183 of SEQ ID NO:13).

Inbred Mouse T-cell Immunogenicity Studies of LVL111, LVL168 and LVL144

Immunogenicity of LVL111, LVL168, and LVL144 were studied in inbred CB6f1 mice. In this experiment, five mice per group (Charles River, Senneville, QC, Canada) have been immunized four times every two weeks with the recombinant CASB7439 proteins (LVL111, LVL168 and LVL144) formulated with AS15. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X concentration adjuvant as described elsewhere herein. The CD4 and CD8 T-cell responses (double positive IFNγ and TNFα) were measured by intracellular staining and flow cytometry analysis in individual inbred CB6f1 mice, after restimulation of splenocytes with the CASB7439 peptides matrix. (Data not shown) The immunizations with the all three constructs triggered a CD4 T-cell immune response against the same immunogenic regions (aa 25 to 43 of SEQ ID NO:13 and aa 89 to 107 of SEQ ID NO:13) in CB6f1 inbred mice.

Immunogenicity of CASB7439 in Mice: CASB7439 T-cell Immunogenicity in Outbred Mice

CD1 outbred mice (Charles River, Senneville, QC, Canada) were immunized four times every two weeks with the recombinant CASB7439 protein (LVL111) formulated with either AS01B or AS15. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X concentration adjuvant as described elsewhere herein. The CD4 and CD8 T-cell responses (double positive IFNγ and TNFα) were measured by intracellular staining and flow cytometry analysis in individual CD1 mice, after restimulation of splenocytes with the CASB7439 peptides matrix.

Tables of the CASB7439 CD4-immunogenic peptides identified in outbred mice immunized with either CASB7439+AS01B or CASB7439+AS15 are presented in Figure 14/35 top and bottom, respectively. The CD4 T-cell immunogenic regions identified in CD1 outbred mice covers the following regions of CASB7439 (expressed in terms of the amino acid positions of SEQ ID NO:13): peptides 8-10 (aa 29 to 51 of SEQ ID NO:13), peptides 16-18 (aa 61 to 83 of SEQ ID NO:13) and peptides 24-26 (aa 93 to 115 of SEQ ID NO:13). The CD8 T-cell immunodominant regions identified in CD1 outbred mice cover the following portions of CASB7439 (SEQ ID NO:13): peptides 16-18 (aa 61

to 83 of SEQ ID NO:13), peptides 24-26 (aa 93 to 115 of SEQ ID NO:13) and peptide 32 (aa 125 to 139 of SEQ ID NO:13). (Data not shown.)

Outbred Mouse T-cell Immunogenicity Studies of LVL111, LVL168 and LVL144

The CASB7439+AS15-mediated CD4 and CD8-T-cell responses induced by LVL111, LVL168 and LVL144 were tested in individual outbred CD1 mice purchased from Charles River, Senneville, QC, Canada. Five mice per group received four intramuscular immunizations every two weeks composed of 10 μg of CASB7439 recombinant protein (LVL111, LVL168 or LVL144) formulated with AS15. (Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X concentration adjuvant as described elsewhere herein. Five mice in the control group were immunized with a saline buffer.

Two weeks after the fourth immunization, the mice were sacrificed. The CD4 and CD8 T-cell responses (double positive IFNγ and TNFα) were measured by intracellular staining and flow cytometry analysis after splenocytes restimulation with four pools of peptides (pool 1: peptides 8-9-10, pool 2: peptides 16-17-18, pool 3: peptides 24-25-26 and pool 4: peptides 30-31-32). The data are presented in Figure 15/35. The same CD4 T-cell immunogenic regions of SEQ ID NO:13, specifically peptides 16-18 (aa 61 to 83 of SEQ ID NO:13) and peptides 24-26 (aa 93 to 115 of SEQ ID NO:13) were activated after immunizations with each of the three CASB7439 constructs formulated with AS15.

HLA A2.1/DR1-transgenic Mouse T-cell Immunogenicity Studies of LVL144 and LVL168

The CASB7439+AS15-mediated CD4 and CD8-T-cell responses induced by two constructs (LVL168 or LVL144) were tested in pooled HLA A2.1/DR1-transgenic mice (3 pools of 3 mice) purchased from the Pasteur Institute, Paris, France. 9 mice per group received four intramuscular immunizations every two weeks composed of 6.25μg of LVL168 or 10μg LVL144, formulated with AS15. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X concentration adjuvant as described elsewhere herein. Two mice in the control group were immunized with the adjuvant alone.

Two weeks after the fourth immunization, the mice were sacrificed. The CD4 and CD8 T-cell responses (double positive IFNγ and TNFα) were measured by intracellular staining and flow cytometry analysis after PBLs and splenocytes restimulation with 7 pools of peptides covering the entire CASB7439 sequence [pool 1: peptides 1-7 (SEQ ID NOS:53-59), pool 2: peptides 8-14 (SEQ ID NOS:60-66), pool 3: peptides 15-21 (SEQ ID NOS:67-73), pool 4: peptides 22-28 (SEQ ID NOS:74-80), pool 5: peptides 29-35 (SEQ ID NOS81-87), pool 6: peptides 36-42 (SEQ ID NOS:88-94) and pool 7: peptides 43-46 (SEQ ID NOS:95-98)]. Moreover, the human HLA DR immunodominant peptide 24 (SEQ ID NO:76) was tested individually. The results presented in Figures 16-19 show that the same CD4 and CD8 T-cell immunodominant region (peptides 22-28 + peptide 24) is activated after immunizations with each of the two CASB7439 constructs formulated with AS15. Thus, peptide 24 is also a human HLA A2.1 immunodominant peptide. These results were confirmed in a second experiment in which the same CASB7439 peptides were used to restimulate splenocytes obtained from 5 individual mice per group immunized with either LVL168 or LVL144 formulated with AS15. (Data not shown.)

Immunogenicity of CASB7439 in Mice: CASB7439-mediated Humoral Response CB6f1 Mouse Studies

CASB7439-mediated humoral response to 10 μg of CASB7439 (LVL111) formulated with either AS01B or AS15, used at 1/10 of the human dose, was assessed in inbred CB6f1 mice (Charles River, Senneville, QC, Canada). Nineteen mice per group received four intramuscular immunizations every two weeks. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of adjuvant (2X concentration). The mice were bled just prior to each immunization and two weeks after the fourth immunization. In another experiment, CASB7439-mediated humoral response to LVL111, LVL168 or LVL144 was assessed in inbred CB6f1 mice (Charles River, Senneville, QC, Canada). Fifty mice per group received four intramuscular immunizations every two weeks with 10 μg of CASB7439 recombinant protein (LVL111, LVL168 or LVL144) formulated with AS15 or with AS15 alone as a control group. Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl of a 2X

concentration adjuvant as described elsewhere herein. The mice were bled two weeks after the fourth immunization.

The anti-CASB7439 IgG1 and IgG2a serum antibody titers have been determined by ELISA, according to standard procedures. Briefly, the CASB7439 recombinant protein was coated in a 96-well plate in phosphate buffer saline (PBS). After blocking with bovine serum albumin 1% in PBS-Tween (0.1%), sera from immunized mice was incubated in the plates for one hour at 37°C. After plate washing, a specific goat anti-IgG1 or IgG2a–HRP labeled mouse antibody (Southern Biotech Associates, Birmingham, AL, USA) was added for one hour at 37°C. After plate washing, a TMB substrate Reagent (BD Pharmingen, Mississauga, ON, Canada) was added for 15 minutes. The colorimetric reaction was stopped by adding 1M sulfuric acid. The OD have been measured at 450nm using a Spectramax spectrophotometer (Molecular Device, Sunnyvale, CA, USA). The antibody titer has been determined with a standard curve, using specific total IgG, IgG1 or IgG2a capture antibodies and an internal CASB7439 positive mouse serum.

Figure 20/35 depicts the time-course analysis of the IgG1 and IgG2a antibody titers in sera of mice immunized with CASB7439 (LVL111) formulated with either AS01B or AS15. The data demonstrate that AS15 triggers a higher Th1 (approximately 10-fold higher IgG2a/IgG1 ratio after 4 immunizations compared to AS01B (IgG2a is a surrogate marker of type 1 immune response). As shown in Figure 21/35, CASB7439 protein formulated with AS15 triggers a very strong IgG1 and IgG2a antibody response in inbred CB6f1 mice, whatever the construct used (LVL111, LVL168 or LVL144).

Immunogenicity of CASB7439 in Mice: CASB7439-mediated Humoral Response CD1 Mouse Studies

CASB7439-mediated humoral response upon CASB7439 immunizations (LVL111, LVL168 or LVL144) was assessed in outbred CD1 mice (Charles River, Senneville, QC, Canada). Ten mice per group received four intramuscular immunizations every two weeks with 10 μg of CASB7439 recombinant protein (LVL111, LVL168 or LVL144) formulated with AS15 (or with AS15 alone as a control group). Each immunization contained 25μl of recombinant CASB7439 protein mixed with 25μl

of a 2X concentration adjuvant as described elsewhere herein.

Two weeks after the fourth immunization, the mice were bled and the anti-CASB7439 total IgG serum antibody titers were determined by ELISA, as already described. CD1 mice immunized with various CASB7439 constructs (LVL111, LVL168 or LVL144) develop a very strong CASB7439-specific humoral immune response (Figure 22/35). The data indicates that a 8 to 10-fold higher anti-CASB7439 total IgG titer was induced by LVL144 as compared to LVL168 and LVL111 respectively in CD1 outbred mice.

Example 11

Anti-Tumor Efficacy of CASB7439 ASCI in Mice: Murine Tumor Model Development

A murine tumor model expressing the CASB7439 protein using TC-1 murine tumor cells or MC38 murine colorectal cancer cells were developed to study CASB7439-mediated *in vivo* tumor protection in mice. TC-1 tumor cells were kindly provided by T. C. Wu (Johns Hopkins University, Baltimore, MD). They were generated from primary lung cells of C57Bl/6 mice by the successive transfer of HPV16 E6 and E7 genes and an activated ras oncogene as described previously (Lin et al. (1996) *Cancer Res.* Jan 1;56(1):21-6. MC38 cells, a murine colorectal adenocarcinoma cell line, were purchased from ATCC (Manassas, VA).

These cells were stably transfected with a construct encoding CASB7439 using the non-liposomal transfectant reagent Fugene-6 (Roche, Mississauga, ON, Canada), according to the manufacturer's recommendations. Briefly, a plasmid (pcDNA3.1/Zeocine) coding for the entire sequence of the CASB7439 protein and containing a zeocine-resistance gene, has been mixed with Fugene-6 and added on top of TC-1 or MC38 tumor cells in serum-free medium overnight. After transfection, cells were allowed to recover for 24 hours in culture medium + 10% heat-inactivated FBS. Cells were then harvested and distributed in 96-well plates (50 µl of cell suspension per well) for limiting dilutions. After cell seeding, the medium was replaced with culture medium + 10% heat-inactivated FBS containing 100µg/ml of zeocine (Invitrogen, Burlington, ON, Canada). After selection, each clone was harvested and allowed to grow in 24-well plates and frozen. The expression of CASB7439 mRNA in

TC1/CASB7439 and MC38/CASB7439 cells was verified by real-time qPCR, using CASB7439-specific TaqMan primers and probe. The real-time qPCR reactions have been performed with a TaqMan reaction kit using a TaqMan 7900 machine (Applied Biosystem, Foster City, CA, USA). The CASB7439 protein expression has been assessed by Western blot and immunofluorescence, using a specific anti-CASB7439 monoclonal antibody (40-12).

The TC-1/CASB7439#14 cell population is not fully clonal but, rather, composed of at least two different subclones. Therefore, another limiting dilution step was done in order to generate a fully clonal TC-1/CASB7439 cell population: the TC-1/CASB7439#14-2 cells.

To determine the optimal dose of TC-1 or MC38 cells for tumor challenge in mice, various doses of cells (from $1x10^5$ to 2×10^6 cells per mouse) were injected in CB6f1 inbred mice. The optimal dose was determined to be 2.5×10^5 cells per mouse for the TC-1/CASB7439#14 and TC1/CASB7439#14-2 challenges; 1×10^5 cells per mouse was found optimal for challenge with the MC38/CASB7439#35 cells.

Anti-tumor Efficacy in Mice: Immunizations and Challenge Experiments

A series of four studies (1.1-1.4) were designed to assess the CASB7439-mediated *in vivo* tumor protection. Generally, in each of these four studies, groups of CB6f1 mice were given four intramuscular immunizations at two week intervals, with various doses (0, 1, 10 or 30 μg) of the specific recombinant CASB7439 protein chosen for the study, formulated with either AS01B or AS15 at a final concentration as described elsewhere herein, with a saline buffer, with the CASB7439 protein alone or with the adjuvants alone. Two weeks after the fourth immunization, CB6f1 mice were challenged by subcutaneous injection with 250,000 TC-1/CASB7439 #14, 100,000 MC38/CASB7439#35 cells or 250,000 TC-1/CASB7439 #14-2 tumor cells. After challenge, the size of the tumors was measured three times a week for each mouse. According to protocol, mice with tumors bigger than 200 mm^2 were sacrificed. Survival curves were then plotted.

An additional three studies were carried out (2.1-2.4) to investigate doses other than 0, 1, 10 or 30 μg of the recombinant CASB7439 protein. Study 2.4 assessed the

correlation between immunogenicity and tumor protection of CASB7439 in male and female C57Bl/6 mice; CB6f1 mice were evaluated in parallel.

Studies 1.1-2.4 are described in greater detail below.

Study #1.1

Applicants investigated the tumor protection conferred by immunization with LVL111 (1 or 10μg) formulated with either AS01B or AS15 against a TC1/CASB7439 #14 challenge in CB6f1 mice. Ten mice per group were intramuscularly immunized four times at two-week intervals with the recombinant CASB7439 protein (LVL111) formulated with either AS01B or AS15 at a final concentration as described elsewhere herein. Two control groups immunized with either a saline buffer or the CASB7439 protein alone were also tested. Two weeks after the fourth immunization, each mouse has been subcutaneously challenged with 250,000 TC-1/CASB7439#14 cells.

Tumor growth was followed for 42 days, but no survival curves were plotted because too few mice were sacrificed at day 42. However, tumor growth curves were plotted for each group. Figure 23/35, top and bottom panels, present the tumor growth curves obtained with the AS01B and the AS15 formulations, respectively. (Note: AS15 is mis-identified in this figure as "AS015".) The curves show that in CB6f1 mice receiving CASB7439 protein formulated with either AS01B or AS15, the TC1/CASB7439#14 tumor growth is slowed.

Study #1.2

Groups of fourteen mice received a different dosage of LVL111 (1, 10 and 30μg) or control. Mice were intramuscularly immunized four times at two week intervals with the different doses of LVL111 formulated with either AS01B or AS15 at a final concentration as described elsewhere herein. Four control groups were also immunized with AS01B-only, AS15-only, saline buffer-only, or 30μg LVL111-only. Two weeks after the fourth immunization, each mouse was subcutaneously challenged with 250,000 TC-1/CASB7439#14 cells. After challenge, the size of the tumors was measured three times a week for each mouse. According to protocol, mice with tumors bigger than 200 mm^2 were sacrificed. A survival curve was drawn.

The results indicate that mice immunized with LVL111 formulated with AS15 were better protected against a TC1/CASB7439#14 tumor challenge than mice immunized with the LVL111 formulated with AS01B. The optimal results in CB6f1 mice were obtained with 10μg of LVL111 formulated with AS15. See Figure 24/35.

Study #1.3

Thirty eight CB6f1 mice per group were intramuscularly immunized four times at two week intervals with 10 μg of LVL111, LVL168 or LVL144 formulated with AS15 at a final concentration as described elsewhere herein. A control group immunized with AS15 alone was also tested.

Two weeks after the fourth immunization, twenty eight mice were subcutaneously challenged with 250,000 TC-1/CASB7439#14 cells and the ten remaining mice have been subcutaneously challenged with 100.000 MC38/CASB7439#35 cells. After challenge, the size of the tumors has been measured three times a week for each mouse. According to protocol, mice with tumors bigger than 200 mm^2 have been sacrificed. The results indicate that LVL111, LVL168 and LVL144, formulated with AS15, are all protective against a TC1/CASB7439#14 or MC38/CASB7439#35 challenge. See Figure 25/35.

Study #1.4

Sixty-six CB6f1 mice per group were intramuscularly immunized four times at two week intervals with 10 μg of the LVL168 or LVL144 formulated with AS15 at a final concentration as described elsewhere herein. A control group immunized with AS15 alone was also tested.

Two weeks after the fourth immunization, twenty two mice per group were subcutaneously challenged with either 250,000 TC-1/CASB7439#14 cells (non-clonal population), 250,000 TC-1/CASB7439#14-2 (clonal population) cells, or 100,000 MC38/CASB7439#35 cells. After challenge, the size of the tumors was measured three times a week for each mouse. According to protocol, mice with tumors bigger than 200 mm^2 were sacrificed. The TC1/CASB7439#14-2 challenge appeared more aggressive than the TC1/CASB7439#14 challenge, as evidenced by the earlier death of the control

mice and the absence of a distinguishable protective effect in this particular study. See Figures 26/35. Subsequent experiments were carried out with the more aggressive TC-1/ CASB7439#14-2, as described elsewhere herein.

Study #2.1

Fifteen CB6f1 mice per group were intramuscularly immunized four times with various equimolar doses of either LVL168 or LVL144 formulated with AS15 at a final concentration as described elsewhere herein. For LVL168, the following antigen doses were used: 0.77, 1.5, 3.1 and 6.25 μg per mouse. For LVL144, the following antigen doses were used: 1.25, 2.5, 5 and 10 μg per mouse. A control group immunized with AS15 alone was also tested. Two weeks after the fourth immunization, all mice were subcutaneously challenged with 250.000 TC-1/ CASB7439 #14-2 (clonal population) cells. After challenge, the size of the tumors was measured three times a week for each mouse. Mice with tumors bigger than 289 mm^2 (17mm X 17mm) were sacrificed. Survival curves for each group were drawn. See Figure 27/35, which depicts the survival curves obtained with equimolar doses of LVL168 and LVL144, formulated with AS15 after a TC1/ CASB7439 #14-2 tumor challenge. TF stands for tumor-free mice. In comparison with the AS15 control group, the four sets of equimolar doses of LVL144 and LVL168 seem to inhibit tumor growth similarly in CB6f1 mice.

Study # 2.2

Fifteen C57Bl/6 mice per group were intramuscularly immunized four times with various equimolar doses of LVL168 or LVL144 formulated with AS15 at a final concentration as described elsewhere herein. For LVL168, the following doses of antigen were used: 0.77, 1.5, 3.1 and 6.25 μg per mouse. For LVL144, the following doses of antigen were used: 1.25, 2.5, 5 and 10 μg per mouse. As control groups, mice were also immunized with AS15 alone or a buffer. Two weeks after the fourth immunization, all mice were subcutaneously challenged with 250.000 TC-1/CASB7439 #14-2 (clonal population) cells. After challenge, the size of the tumors was measured three times a week for each mouse. Mice with tumors bigger than 289 mm^2 (17mm X 17mm) were sacrificed. Survival curves for each group were drawn. See Figure 28/35,

which depicts the survival curves obtained with the different equimolar doses of LVL168 and LVL144, formulated with AS15 after a TC1/CASB7439 #14-2 tumor challenge. TF stands for tumor-free mice. In contrast with the studies conducted with CB6f1 mice, in the present study, no statistically significant difference between the survival of C57Bl/6 mice in the immunized groups, the AS15 group, or the buffer group was observed.

Study # 2.3

Fifteen CB6f1 mice per group were intramuscularly immunized four times with various equimolar doses of LVL168 or LVL144 formulated with AS15 at a final concentration as described elsewhere herein. For LVL168, the following doses were used: 0.04, 0.19, 0.77, 1.5, 3.1 and 6.25 μg per mouse. For LVL144, the following doses were used: 0.078, 0.31, 1.25, 2.5, 5 and 10 μg per mouse. A control group immunized with AS15 alone has also been tested. Naive mice were also used as a second control group. Two weeks after the fourth immunization, all mice were subcutaneously challenged with 250,000 TC-1/CASB7439#14-2 (clonal population) cells. After challenge, the size of the tumors was measured three times a week for each mouse. Mice with tumors bigger than 289 mm^2 (17mm X 17mm) were sacrificed. Survival curves for each group were drawn. See Figures 29/35 and 30/35, which depict the survival curves obtained with the different equimolar doses of LVL168 and LVL144, formulated with AS15 after a TC1/CASB7439 #14-2 tumor challenge. TF stands for tumor-free mice.

The two lowest equimolar doses tested did not appear to significantly inhibit the tumor growth. At the four highest doses, more tumor protection was observed in those mice immunized with LVL168 plus AS15 as compared to LVL144 plus AS15.

Study # 2.4

The lack of protection observed in C57Bl/6 mice immunized with CASB7439 formulated with AS15 against a TC1/CASB7439 #14-2 tumor challenge, as observed in study # 2.2, was further investigated. Fifteen CB6f1 or C57Bl/6 mice per group were intramuscularly immunized four times with 1 μg of LVL168 formulated with AS15, at a final concentration as described elsewhere herein. Male and female groups of mice

immunized with 1 μg of LVL168 + AS15 were studied separately. Separate control male and female groups immunized with AS15 alone were also tested.

Partial bleeds were obtained 7 days after the fourth immunization and the anti-CASB7439 total IgG serum antibody titers were determined by ELISA, as described in Example 10. Similar CASB7439-specific CD4 T-cell responses (approximately 1 x 10^6 ng/ml total IgG) were obtained in both male and female CB6f1 mice immunized with LVL168 + AS15; no antibody response was detected in C57Bl/6 mice. See Figure 33/35.

Partial bleeds were obtained 14 days after the fourth immunization. The CD4 and CD8 T-cell responses (double positive IFNγ and TNFα) were measured by intracellular staining and flow cytometry analysis, as described in Example 10, after PBLs from either CB6f1 or C57Bl/6 mice (5 pools of 3 mice per group) were restiimulated with a pool of peptides covering the entire CASB7439 sequence. In addition, peptide 39 (SEQ ID NO:91) was tested with PBLs from C57Bl/6. Similar CASB7439-specific CD4 T-cell responses (approximately 0.1 % frequency) were obtained in both male and female CB6f1 mice immunized with LVL168 + AS15; a CASB7439-specific CD8 T-cell response (0.05 % frequency) was obtained only in the male mice. See Figures 34/35 and 35/35, top panels. No CASB7439-specific CD4 or CD8 T-cell response was detected in C57Bl/6 mice immunized with 1 μg LVL168 + AS15 after PBL restimulation with a pool of peptides covering the entire CASB7439 sequence or CASB7439 peptide 39 (SEQ ID NO:91) (aa 153 to 167 of SEQ ID NO:13). See Figures 34/35 and 35/35 (bottom panels). Note that the CASB7439 peptide 39 (SEQ ID NO:91) was utilized independently here because it had previously been identified as a low CD4 T cell immunogenic peptide in the splenocytes of C57Bl/6 mice immunized four times with 10 μg LVL168 + AS15. See Figure 13 in Example 10.

Two weeks after the fourth immunization, all mice were subcutaneously challenged with 250,000 TC-1/CASB7439#14-2 cells. After challenge, the size of the tumors was measured three times a week for each mouse. Mice with tumors bigger than 289 mm^2 (17mm X 17mm) were sacrificed. Tumor growth and survival curves for each group were drawn. Figure 31/35 and 32/35 depict, respectively, the tumor growth and survival curves obtained with male or female CB6f1 or C57Bl/6 mice immunized

with 1 μg of LVL168 or LVL144, formulated with AS15, after a TC1/CASB7439 #14-2 tumor challenge. The results indicate that CASB7439 + AS15 tends to yield slower tumor growth and better survival in male CB6f1 mice compared to female CB6f1, but has no effect on tumor protection in C57Bl/6 mice. The lack of CASB7439 tumor protection in C57Bl/6 mice corresponds to the absence of detectable CASB7439-specific CD4 and CD8 T-cells, as well as the absence of a detectable CASB7439-specific antibody response in C57Bl/6 mice. This indicates that, under the experimental conditions used, C57Bl/6 mice are not an ideal mouse model to study immunogenicity and tumor protection by CASB7439. On the other hand, in CB6f1 mice, the observed CASB7439 tumor protection corresponds to the presence of significant CASB7439-specific CD4 T-cells [see immunogenic peptides 7-8 (aa 25 to 43 of SEQ ID NO:13) and 23-24 (aa 89 to 107 of SEQ ID NO:13) Figure 13/35, Example 10] and CASB7439-specific CD8 T-cells, as well as a CASB7439-specific antibody response.

SEQUENCES:

SEQ ID NO:1

LVL055 protein

MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALA

SEQ ID NO:2

LVL055 DNA

atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggatggtggcaccctgccgcgtagcgctccgccggcaccgccggttccggttggttg
tgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggcc
accgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtg
tgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaa
aaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctg
ctggccgaacatgatgcggtgcgtaacgcgctggcctaa

SEQ ID NO:3

LVL111 protein

MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALAGGLRPQAVRPSAPRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWL
GGYHHHHHH

SEQ ID NO:4
LVL111 DNA
atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggatggtggcaccctgccgcgtagcgcaccgccggctccgccggttccggttggttg
tgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggcc
accgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtg
tgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaa
aaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctg
ctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtc
cgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatga
tagcggctgcgaaggtgccctgagcccggcggaacgtgaactgctggattttagcagctggctg
ggcggctatcatcatcatcaccatcattaa

SEQ ID NO:5
LVL137 protein
MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALAGGLRPQAVRPSAPRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWL
GGY

SEQ ID NO:6
LVL137 DNA
atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggatggtggcaccctgccgcgtagcgcaccgccggctccgccggttccggttggttg
tgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggcc
accgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtg
tgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaa
aaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctg
ctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtc
cgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatga
tagcggctgcgaaggtgccctgagcccggcggaacgtgaactgctggattttagcagctggctg
ggcggctattaa

SEQ ID NO:7

LVL141 protein

MDPSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVV
IHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETAAAHMDGGTLPRSAPPA
PPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHV
PHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGGSSEPGSPR
SAYSSDDSGCEGALSPAERELLDFSSWLGGYHHHHHH

SEQ ID NO:8

LVL141 DNA

atggatccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcatta
ttgctcaccgtggtgctagcggttatttaccagagcatacgttagaatctaaagcacttgcgtt
tgcacaacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggtt
attcacgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcatcgta
aagatggccgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacaga
aaactttgaaaccgcggccgcacatatggatggtggcaccctgccgcgtagcgcaccgccggct
ccgccggttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgca
gccgtcgtcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaa
cgaacgtgaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtg
ccgcatggcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatata
ttcgtgcgctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtct
gcgtccgcaggcggttcgtccgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgt
agcgcctatagcagcgatgatagcggctgcgaaggtgccctgagcccggcggaacgtgaactgc
tggattttagcagctggctgggcggctatcatcatcatcaccatcattaa

SEQ ID NO:9

LVL144 protein

MDPSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVV
IHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETDGGTLPRSAPPAPPVPV
GCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGA
SKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGGSSEPGSPRSAYSS
DDSGCEGALSPAERELLDFSSWLGGYHHHHHH

SEQ ID NO:10

LVL144 DNA

atggatccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcatta
ttgctcaccgtggtgctagcggttatttaccagagcatacgttagaatctaaagcacttgcgtt
tgcacaacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggtt
attcacgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcatcgta
aagatggccgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacaga
aaactttgaaaccgatggtggcaccctgccgcgtagcgcaccgccggctccgccggttccggtt
ggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtc
cggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaa
ccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcg
agcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaac
gtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggt
tcgtccgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagc
gatgatagcggctgcgaaggtgccctgagcccggcggaacgtgaactgctggattttagcagct
ggctgggcggctatcatcatcatcaccatcattaa

SEQ ID NO:11

LVL168 protein

MHHHHHHHHHHDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVAR
RNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAG
GLRPQAVRPSAPRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:12

LVL168 DNA

atgcatcatcatcatcatcatcatcatcatcatgacggtggcaccctgccgcgtagcgcaccgc
cggctccgccggttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcg
ttgcagccgtcgtcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgt
cgtaacgaacgtgaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagc
atgtgccgcatggcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtgga
atatattcgtgcgctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggt
ggtctgcgtccgcaggcggttcgtccgagcgcgccgcgtggtggtagcagcgaaccgggtagcc
cgcgtagcgcctatagcagcgatgatagcggctgcgaaggtgccctgagcccggcggaacgtga
actgctggattttagcagctggctgggcggctattaa

SEQ ID NO:13

Accession No. AAB86993, HASH2 protein

MDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVK
LVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPS
APRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGG
Y

SEQ ID NO:14

HASH2 DNA

atggacggcggcacactgcccaggtccgcgccccctgcgccccccgtccctgtcggctgcgctg
cccggcggagacccgcgtccccggaactgttgcgctgcagccggcggcggcgaccggccaccgc
agagaccggaggcggcgcagcggccgtagcgcggcgcaatgagcgcgagcgcaaccgcgtgaag
ctggtgaacttgggcttccaggcgctgcggcagcacgtgccgcacggcggcgccagcaagaagc
tgagcaaggtggagacgctgcgctcagccgtggagtacatccgcgcgctgcagcgcctgctggc
cgagcacgacgccgtgcgcaacgcgctggcgggagggctgaggccgcaggccgtgcggccgtct
gcgccccgcgggccgccagggaccaccccggtcgccgcctcgccctcccgcgcttcttcgtccc
cgggccgcgggggcagctcggagcccggctccccgcgttccgcctactcgtcggacgacagcgg
ctgcgaaggcgcgctgagtcctgcggagcgcgagctactcgacttctccagctggttagggggc
tactga

SEQ ID NO:15
LVL007 protein
MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVLVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAYSSDD
SGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:16
LVL007 DNA
atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggacggcggcacactgcccaggtccgcgccccctgcgccccccgtccttgtcggctg
cgctgcccggcggagacccgcgtccccggaactgttgcgctgcagccggcggcggcgaccggcc
accgcagagaccggaggcggcgcagcggccgtagcgcggcgcaatgagcgcgagcgcaaccgcg
tgaagctggtgaacttgggcttccaggcgctgcggcagcacgtgccgcacggcggcgccagcaa
gaagctgagcaaggtggagacgctgcgctcagccgtggagtacatccgcgcgctgcagcgcctg
ctggccgagcacgacgccgtgcgcaacgcgctggcgggagggctgaggccgcaggccgtgcggc
cgtctgcgccccgcgggccgccagggaccaccccggtcgccgcctcgccctcccgcgcttcttc
gtccccgggccgcgggggcagctcggagcccggctccccgcgttccgcctactcgtcggacgac
agcggctgcgaaggcgcgctgagtcctgcggagcgcgagctactcgacttctccagctggttag
ggggctactga

SEQ ID NO:17
LVL010 protein
MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAYSSDD
SGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:18

LVL010 DNA

atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggacggcggcacactgcccaggtccgcgccccctgcgccccccgtccctgtcggctg
cgctgcccggcggagacccgcgtccccggaactgttgcgctgcagccggcggcggcgaccggcc
accgcagagaccggaggcggcgcagcggccgtagcgcggcgcaatgagcgcgagcgcaaccgcg
tgaagctggtgaacttgggcttccaggcgctgcggcagcacgtgccgcacggcggcgccagcaa
gaagctgagcaaggtggagacgctgcgctcagccgtggagtacatccgcgcgctgcagcgcctg
ctggccgagcacgacgccgtgcgcaacgcgctggcgggagggctgaggccgcaggccgtgcggc
cgtctgcgccccgcgggccgccagggaccaccccggtcgccgcctcgccctcccgcgcttcttc
gtccccgggccgcgggggcagctcggagcccggctccccgcgttccgcctactcgtcggacgac
agcggctgcgaaggcgcgctgagtcctgcggagcgcgagctactcgacttctccagctggttag
ggggctactga

SEQ ID NO:19

LVL060 protein

MGHHHHHHHHHHSSGHIDDDDKHMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPA
TAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRL
LAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAYSSDD
SGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:20

LVL060 DNA

atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggatggtggcaccctgccgcgtagcgctccgccggcaccgccggttccggttggttg
tgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggcc
accgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtg
tgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaa
aaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctg
ctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtc
cgagcgcaccgcgtggtccgccgggtacgacgccggttgcagcgagcccgagccgtgcgagcag
ctctccgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatgat
agcggctgcgaaggtgccctgtctccggcggaacgtgaactgctggattttagcagctggctgg
gcggctattaa

SEQ ID NO:21

Const-1 Protein

MGHHHHHHHHHHSSGHIDDDDKHMATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHG
GASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPS
RASSSPGRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGYLEDPAANKARKEAEL
AAATAEQ

SEQ ID NO:22

Const-1 DNA

atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtga
acgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggc
ggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgc
tgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgca
ggcggttcgtccgagcgcaccgcgtggtccgccgggtacgacgccggttgcagcgagcccgagc
cgtgcgagcagctctccgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctata
gcagcgatgatagcggctgcgaaggtgccctgtctccggcggaacgtgaactgctggattttag
cagctggctgggcggctatctcgaggatccggctgctaacaaagcccgaaaggaagctgagttg
gctgctgccaccgctgagcaataa

SEQ ID NO:23

LVL056 Protein

MGHHHHHHHHHHSSGHIDDDDKHMNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIR
ALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSA
YSSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:24

LVL056 DNA

atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatgaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgca
tggcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgt
gcgctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtc
cgcaggcggttcgtccgagcgcaccgcgtggtccgccgggtacgacgccggttgcagcgagccc
gagccgtgcgagcagctctccgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcgcc
tatagcagcgatgatagcggctgcgaaggtgccctgtctccggcggaacgtgaactgctggatt
ttagcagctggctgggcggctattaa

SEQ ID NO:25
LVL057 Protein
MGHHHHHHHHHHSSGHIDDDDKHMAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSK
VETLRSAVEYIRALQRLLAEHDAVRNALA

SEQ ID NO:26
LVL057 DNA
atgggccatcatcatcatcatcatcatcatcacagcagcggccatatcgacgacgacgaca
agcatatggcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaactggtgaa
cctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaactgagcaaa
gtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggccgaacatg
atgcggtgcgtaacgcgctggcctaa

SEQ ID NO:27
LVL088 Protein
MDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVK
LVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPS
APRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGYHHHHHH

SEQ ID NO:28
LVL088 DNA
atggatggtggcaccctgccgcgtagcgcaccgccggctccgccggttccggttggttgtgcgg
cgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggccaccgc
ggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaa
ctggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaac
tgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggc
cgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtccgagc
gcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatgatagcg
gctgcgaaggtgccctgagcccggcggaacgtgaactgctggattttagcagctggctgggcgg
ctatcatcatcatcaccatcattaa

SEQ ID NO:29

LVL016 Protein

MDPSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVV
IHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETAAAMDGGTLPRSAPPAP
PVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVP
HGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAAS
PSRASSSPGRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGYLEHHHHHH

SEQ ID NO:30

LVL016 DNA

atggatccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcatta
ttgctcaccgtggtgctagcggttatttaccagagcatacgttagaatctaaagcacttgcgtt
tgcacaacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggtt
attcacgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcatcgta
aagatggccgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacaga
aaactttgaaaccgcggccgcaatggatggtggcaccctgccgcgtagcgctccgccggcaccg
ccggttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagcc
gtcgtcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacga
acgtgaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccg
catggcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattc
gtgcgctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcg
tccgcaggcggttcgtccgagcgcaccgcgtggtccgccgggtacgacgccggttgcagcgagc
ccgagccgtgcgagcagctctccgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcg
cctatagcagcgatgatagcggctgcgaaggtgccctgtctccggcggaacgtgaactgctgga
ttttagcagctggctgggcggctatctcgagcaccaccaccaccaccactga

SEQ ID NO:31

LVL018 Protein

MDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVK
LVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPS
APRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGG
YLEHHHHHH

SEQ ID NO:32

LVL018 DNA

atggatggtggcaccctgccgcgtagcgctccgccggcaccgccggttccggttggttgtgcgg
cgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggccaccgc
ggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaa
ctggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaac
tgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggc
cgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtccgagc
gcaccgcgtggtccgccgggtacgacgccggttgcagcgagcccgagccgtgcgagcagctctc
cgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatgatagcgg
ctgcgaaggtgccctgtctccggcggaacgtgaactgctggattttagcagctggctgggcggc
tatctcgagcaccaccaccaccaccactga

SEQ ID NO:33

LVL138 Protein

MHHHHHHDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNER
ERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRP
QAVRPSAPRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:34

LVL138 DNA

atgcaccatcaccatcaccatgatggtggcaccctgccgcgtagcgcaccgccggctccgccgg
ttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcg
tcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgt
gaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatg
gcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgc
gctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccg
caggcggttcgtccgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcct
atagcagcgatgatagcggctgcgaaggtgccctgagcccggcggaacgtgaactgctggattt
tagcagctggctgggcggctattaa

SEQ ID NO:35
Modified linker DNA
atgcatcatcatcatcatcatgac

SEQ ID NO:36
Unmodified linker DNA
atgcaccatcaccatcaccatgat

SEQ ID NO:37
LVL160 Protein
MHHHHHHDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRRRRPATAETGGGAAAVARRNER
ERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRP
QAVRPSAPRGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:38
LVL160 DNA
atgcatcatcatcatcatcatgacggtggcaccctgccgcgtagcgcaccgccggctccgccgg
ttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcg
tcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgt
gaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatg
gcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgc
gctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccg
caggcggttcgtccgagcgcgccgcgtggtggtagcagcgaaccgggtagcccgcgtagcgcct
atagcagcgatgatagcggctgcgaaggtgccctgagcccggcggaacgtgaactgctggattt
tagcagctggctgggcggctattaa

SEQ ID NO:39
1/3pD Protein
MDPSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVV
IHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFET

SEQ ID NO:40

1/3pD DNA

atggatccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcatta
ttgctcaccgtggtgctagcggttatttaccagagcatacgttagaatctaaagcacttgcgtt
tgcacaacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggtt
attcacgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcatcgta
aagatggccgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacaga
aaactttgaaacc

SEQ ID NO:41

H. influenzae Protein D

MKLKTLALSLLAAGVLAGCSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQAD
YLEQDLAMTKDGRLVVIHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETK
DGKQAQVYPNRFPLWKSHFRIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAE
TLKVLKKYGYDKKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGY
WVNYNYDWMFKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTV
RKDALPEFFTDVNQMYDALLNKSGATGVFTDFPDTGVEFLKGIK

SEQ ID NO:42

Long polyhistidine tag

MGHHHHHHHHHHSSGHIDDDDKH

SEQ ID NO:43

LVL090 Protein

MDPSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVV
IHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWK
SHFRIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDM
VYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAMA
EVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPAFFTDVNQMY
DVLLNKSGATGVFTDFPDTGVEFLKGIKAAAMDGGTLPRSAPPAPPVPVGCAARRRPASPELLR
CSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVE
YIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSP
RSAYSSDDSGCEGALSPAERELLDFSSWLGGYLEHHHHHH

SEQ ID NO:44

LVL090 DNA

atggatccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcatta
ttgctcaccgtggtgctagcggttatttaccagagcatacgttagagtctaaagcacttgcgtt
tgcacaacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggtt
attcacgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcaccgta
aagatggtcgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacaga
aaactttgaaaccaaagatggcaaacaagcgcaagtttatcctaatcgtttcccactttggaaa
tcacattttagaattcatacctttgaagatgaaattgaatttatccaaggcttagaaaaatcca
ctggcaaaaaagtagggatttatccagaaatcaaagcaccttggttccaccatcaaaatggtaa
agatattgctgctgaaacgctcaaagtgttaaaaaaatatggctatgataagaaaaccgatatg
gtttacttacaaactttcgattttaatgaattaaaacgtatcaaaacggaattacttccacaaa
tgggaatggatttgaaattagttcaattaattgcttatacagattggaaagaaacacaagaaaa
agacccaaagggttattgggtaaactataattacgattggatgtttaaacctggtgcaatggca
gaagtggttaaatatgccgatggtgttggcccaggttggtatatgttagttaataaagaagaat
ccaaacctgataatattgtgtacactccgttggtaaaagaacttgcacaatataatgtggaagt
gcatccttacaccgtgcgtaaagatgcactacccgcgtttttcacagacgtaaatcaaatgtat
gatgtcttattgaataaatcaggggcaacaggtgtatttactgatttcccagatactggcgtgg
aattcttaaaaggaataaaagcggccgcaatggatggtggcaccctgccgcgtagcgctccgcc
ggcaccgccggttccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgt
tgcagccgtcgtcgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtc
gtaacgaacgtgaacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagca
tgtgccgcatggcggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaa
tatattcgtgcgctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtg
gtctgcgtccgcaggcggttcgtccgagcgcaccgcgtggtccgccgggtacgacgccggttgc
agcgagcccgagccgtgcgagcagctctccgggtcgtggtggtagcagcgaaccgggtagcccg
cgtagcgcctatagcagcgatgatagcggctgcgaaggtgccctgtctccggcggaacgtgaac
tgctggattttagcagctggctgggcggctatctcgagcaccaccaccaccaccac

SEQ ID NO:45

LVL112 Protein

MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGGDGGTLPRSAPPAPPVPVGCAAR
RRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLS
KVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPG
RGGSSEPGSPRSAYSSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:46

LVL112 DNA

atgggtcatcaccatcatcatcacgggtcggactcagaagtcaatcaagaagctaagccagagg
tcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagat
cttcttcaagatcaaaaagaccactcctttaagaaggctgatggaagcgttcgctaaaagacag
ggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaagctgatcaggccc
ctgaagatttggacatggaggataacgatattattgaggctcaccgcgaacagattggaggtga
tggtggcaccctgccgcgtagcgctccgccggcaccgccggttccggttggttgtgcggcgcgt
cgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggccaccgcggaaa
ccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaactggt
gaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaactgagc
aaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggccgaac
atgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtccgagcgcacc
gcgtggtccgccgggtacgacgccggttgcagcgagcccgagccgtgcgagcagctctccgggt
cgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatgatagcggctgcg
aaggtgccctgtctccggcggaacgtgaactgctggattttagcagctggctgggcggctat

SEQ ID NO:47

LVL113 Protein

MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGGDPSSHSSNMANTQMKSDKIIIA
HRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHFLDGLTDVAKKFPHRHRKD
GRYYVIDFTLKEIQSLEMTENFETAAAMDGGTLPRSAPPAPPVPVGCAARRRPASPELLRCSRR
RRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLSKVETLRSAVEYIRA
LQRLLAEHDAVRNALAGGLRPQAVRPSAPRGPPGTTPVAASPSRASSSPGRGGSSEPGSPRSAY
SSDDSGCEGALSPAERELLDFSSWLGGY

SEQ ID NO:48

LVL113 DNA

atgggtcatcaccatcatcatcacgggtcggactcagaagtcaatcaagaagctaagccagagg
tcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagat
cttcttcaagatcaaaaagaccactcctttaagaaggctgatggaagcgttcgctaaaagacag
ggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaagctgatcaggccc
ctgaagatttggacatggaggataacgatattattgaggctcaccgcgaacagattggaggtga
tccaagcagccattcatcaaatatggcgaatacccaaatgaaatcagacaaaatcattattgct
caccgtggtgctagcggttatttaccagagcatacgttagaatctaaagcacttgcgtttgcac
aacaggctgattatttagagcaagatttagcaatgactaaggatggtcgtttagtggttattca
cgatcacttttagatggcttgactgatgttgcgaaaaaattcccacatcgtcatcgtaaagat
ggccgttactatgtcatcgactttaccttaaaagaaattcaaagtttagaaatgacagaaaact
ttgaaaccgcggccgcaatggatggtggcaccctgccgcgtagcgctccgccggcaccgccggt
tccggttggttgtgcggcgcgtcgtcgtccggcgagcccggaactgctgcgttgcagccgtcgt
cgccgtccggccaccgcggaaaccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtg
aacgtaaccgtgtgaaactggtgaacctgggctttcaggcgctgcgtcagcatgtgccgcatgg
cggtgcgagcaaaaaactgagcaaagtggaaaccctgcgtagcgcggtggaatatattcgtgcg
ctgcaacgtctgctggccgaacatgatgcggtgcgtaacgcgctggccggtggtctgcgtccgc
aggcggttcgtccgagcgcaccgcgtggtccgccgggtacgacgccggttgcagcgagcccgag
ccgtgcgagcagctctccgggtcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctat
agcagcgatgatagcggctgcgaaggtgccctgtctccggcggaacgtgaactgctggatttta
gcagctggctgggcggctat

SEQ ID NO:49

LVL114 Protein

MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGGDGGTLPRSAPPAPPVPVGCAAR
RRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLS
KVETLRSAVEYIRALQRLLAEHDAVRNALA

SEQ ID NO:50

LVL114 DNA

atgggtcatcaccatcatcatcacgggtcggactcagaagtcaatcaagaagctaagccagagg
tcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagat
cttcttcaagatcaaaaagaccactcctttaagaaggctgatggaagcgttcgctaaaagacag
ggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaagctgatcaggccc
ctgaagatttggacatggaggataacgatattattgaggctcaccgcgaacagattggaggtga
tggtggcaccctgccgcgtagcgctccgccggcaccgccggttccggttggttgtgcggcgcgt
cgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggccaccgcggaaa
ccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaactggt
gaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaactgagc
aaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggccgaac
atgatgcggtgcgtaacgcgctggcc

SEQ ID NO:51

LVL115 Protein

MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGGDGGTLPRSAPPAPPVPVGCAAR
RRPASPELLRCSRRRRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGASKKLS
KVETLRSAVEYIRALQRLLAEHDAVRNALAGGLRPQAVRPSAPRGGSSEPGSPRSAYSSDDSGC
EGALSPAERELLDFSSWLGGY

SEQ ID NO:52

LVL115 DNA

atgggtcatcaccatcatcatcacgggtcggactcagaagtcaatcaagaagctaagccagagg
tcaagccagaagtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagagat
cttcttcaagatcaaaaagaccactcctttaagaaggctgatggaagcgttcgctaaaagacag
ggtaaggaaatggactccttaagattcttgtacgacggtattagaattcaagctgatcaggccc
ctgaagatttggacatggaggataacgatattattgaggctcaccgcgaacagattggaggtga
tggtggcaccctgccgcgtagcgcaccgccggctccgccggttccggttggttgtgcggcgcgt
cgtcgtccggcgagcccggaactgctgcgttgcagccgtcgtcgccgtccggccaccgcggaaa
ccggtggtggtgcggcagcggttgcgcgtcgtaacgaacgtgaacgtaaccgtgtgaaactggt
gaacctgggctttcaggcgctgcgtcagcatgtgccgcatggcggtgcgagcaaaaaactgagc
aaagtggaaaccctgcgtagcgcggtggaatatattcgtgcgctgcaacgtctgctggccgaac
atgatgcggtgcgtaacgcgctggccggtggtctgcgtccgcaggcggttcgtccgagcgcgcc
gcgtggtggtagcagcgaaccgggtagcccgcgtagcgcctatagcagcgatgatagcggctgc
gaaggtgccctgagcccggcggaacgtgaactgctggattttagcagctggctgggcggctat

SEQ ID NO:53
Peptide 1
MDGGTLPRSAPPAPP

SEQ ID NO:54
Peptide 2
TLPRSAPPAPPVPVG

SEQ ID NO:55
Peptide 3
SAPPAPPVPVGCAAR

SEQ ID NO:56
Peptide 4
APPVPVGCAARRRPA

SEQ ID NO:57
Peptide 5
PVGCAARRRPASPEL

SEQ ID NO:58
Peptide 6
AARRRPASPELLRCS

SEQ ID NO:59
Peptide 7
RPASPELLRCSRRRR

SEQ ID NO:60
Peptide 8
PELLRCSRRRRPATA

SEQ ID NO:61
Peptide 9
RCSRRRRPATAETGG

SEQ ID NO:62
Peptide 10
RRRPATAETGGGAAA

SEQ ID NO:63
Peptide 11
ATAETGGGAAAVARR

SEQ ID NO:64
Peptide 12
TGGGAAAVARRNERE

SEQ ID NO:65
Peptide 13
AAAVARRNERERNRV

SEQ ID NO:66
Peptide 14
ARRNERERNRVKLVN

SEQ ID NO:67
Peptide 15
ERERNRVKLVNLGFQ

SEQ ID NO:68
Peptide 16
NRVKLVNLGFQALRQ

SEQ ID NO:69
Peptide 17
LVNLGFQALRQHVPH

SEQ ID NO:70
Peptide 18
GFQALRQHVPHGGAS

SEQ ID NO:71
Peptide 19
LRQHVPHGGASKKLS

SEQ ID NO:72
Peptide 20
VPHGGASKKLSKVET

SEQ ID NO:73
Peptide 21
GASKKLSKVETLRSA

SEQ ID NO:74
Peptide 22
KLSKVETLRSAVEYI

SEQ ID NO:75
Peptide 23
VETLRSAVEYIRALQ

SEQ ID NO:76
Peptide 24
RSAVEYIRALQRLLA

SEQ ID NO:77
Peptide 25
EYIRALQRLLAEHDA

SEQ ID NO:78
Peptide 26
ALQRLLAEHDAVRNA

SEQ ID NO:79
Peptide 27
LLAEHDAVRNALAGG

SEQ ID NO:80
Peptide 28
HDAVRNALAGGLRPQ

SEQ ID NO:81
Peptide 29
RNALAGGLRPQAVRP

SEQ ID NO:82
Peptide 30
AGGLRPQAVRPSAPR

SEQ ID NO:83
Peptide 31
RPQAVRPSAPRGPPG

SEQ ID NO:84
Peptide 32
VRPSAPRGPPGTTPV

SEQ ID NO:85
Peptide 33
APRGPPGTTPVAASP

SEQ ID NO:86
Peptide 34
PPGTTPVAASPSRAS

SEQ ID NO:87
Peptide 35
TPVAASPSRASSSPG

SEQ ID NO:88
Peptide 36
ASPSRASSSPGRGGS

SEQ ID NO:89
Peptide 37
RASSSPGRGGSSEPG

SEQ ID NO:90
Peptide 38
SPGRGGSSEPGSPRS

SEQ ID NO:91
Peptide 39
GGSSEPGSPRSAYSS

SEQ ID NO:92
Peptide 40
EPGSPRSAYSSDDSG

SEQ ID NO:93
Peptide 41
PRSAYSSDDSGCEGA

SEQ ID NO:94
Peptide 42
YSSDDSGCEGALSPA

SEQ ID NO:95
Peptide 43
DSGCEGALSPAEREL

SEQ ID NO:96
Peptide 44
EGALSPAERELLDFS

SEQ ID NO:97
Peptide 45
SPAERELLDFSSWLG

SEQ ID NO:98
Peptide 46
AERELLDFSSWLGGY

We claim:

1. A modified CASB7439 polypeptide comprising at least one expression enhancement modification, wherein the expression enhancement modification comprises at least the deletion of part or all of the proline-rich region of SEQ ID NO:13.

2. A protein construct comprising a modified CASB7439 polypeptide of claim 1 and further comprising one or more heterologous polypeptides.

3. The protein construct of claim 3, wherein the protein construct comprises a construct selected from:

(a) LVL055 (SEQ ID NO:1);

(b) LVL111 (SEQ ID NO:3);

(c) LVL137 (SEQ ID NO:5);

(d) LVL141 (SEQ ID NO:7);

(e) LVL144 (SEQ ID NO:9); and

(f) LVL168 (SEQ ID NO:11).

4. An immunogenic composition comprising the protein construct of any one of claims 1-3, and a pharmaceutically acceptable carrier or excipient, wherein the carrier or excipient may optionally comprise a buffer.

5. The immunogenic composition of claim 4, further comprising an adjuvant.

6. The immunogenic composition of claim 5, wherein the adjuvant elicits at least a Th1 immune response.

7. The immunogenic composition of claim 4, 5, or 6, wherein the adjuvant comprises at least one of: TLR-4 agonist, immunologically active saponin fraction, TLR-9 agonist.

8. The immunogenic composition of claim 7 wherein said TLR-4 agonist is 3D-MPL.

9. The immunogenic composition of claim 7 or 8 wherein said TLR-9 agonist is a CpG oligonucleotide.

10. The immunogenic composition of claim 7, 8 or 9 wherein said immunologically active saponin fraction is QS21.

11. The immunogenic composition of any of claims 7-10, comprising 3D-MPL.

12. The immunogenic composition of any of claims 7-11 comprising CpG.

13. The immunogenic composition of any of claims 7-12 comprising QS21.

14. The immunogenic composition of claim 13, further comprising cholesterol.

15. A nucleic acid molecule comprising a polynucleotide sequence that encodes the protein construct of any of claims 1-3.

16. A method of inducing an irnmuneresponse to CASB7439 in a human or non-human animal comprising administering to the human or non-human animal an effective amount of a composition comprising an adjuvant and a protein comprising a polypeptide sequence selected from the group consisting of:

(a) the polypeptide sequence set forth in SEQ ID NO:9; and

(b) the polypeptide sequence set forth in SEQ ID NO:11.

17. A method of treating a human or non-human animal comprising the steps of:

(a) selecting a human or non-human animal having cancerous cells that express CASB7439; and

(b) administering to the human or non-human animal an effective amount of a composition comprising an adjuvant and a protein comprising a polypeptide selected from the group consisting of:

(i) the construct set forth in SEQ ID NO:9; and

(ii) the construct set forth in SEQ ID NO:11.

18. An immunogenic composition as defined in claims 7-14 for use in a method of inducing an immune response to CASB7439 in a human or non human animal, said method comprising administering an effective amount of said immunogenic composition.

19. Immunogenic composition according to claim 18 which comprises a protein comprising a polypeptide selected from the group consisting of:

(a) the construct set forth in SEQ ID NO:9; and

(b) the construct set forth in SEQ ID NO:11.

20. Use of an immunogenic composition as defined in claims 7-14 in the manufacture of a medicament for use in a method of inducing an immune response to CASB7439 in a human or non human animal, said method comprising administering an effective amount of said immunogenic composition.

21. Use according to claim 20 wherein said immunogenic composition comprises

a protein comprising a polypeptide selected from the group consisting of:

(i) the construct set forth in SEQ ID NO:9; and

(ii) the construct set forth in SEQ ID NO:11.