



(51) International Patent Classification:

C07K 19/00 (2006.01) *C12N 15/62* (2006.01)
C07K 14/315 (2006.01) *A61K 39/09* (2006.01)
C12N 15/09 (2006.01) *A61P 31/04* (2006.01)

(21) International Application Number:

PCT/US2013/052352

(22) International Filing Date:

26 July 2013 (26.07.2013)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/676,276 26 July 2012 (26.07.2012) US

(71) Applicants: **GENOCEA BIOSCIENCES, INC.** [US/US]; Cambridge Discovery Park, 100 Acorn Park Drive, 5th Floor, Cambridge, Massachusetts 02140 (US). **PATH VACCINE SOLUTIONS** [US/US]; 2201 Westlake Avenue, Suite 200, Seattle, Washington 98121 (US). **THE UAB RESEARCH FOUNDATION** [US/US]; 701 20th Street South, 770 Admin. Building, Birmingham, Alabama 35294-0107 (US).

(72) Inventors: **MAISONNEUVE, Jean-Francois Lucien**; 739 S.W. 298th Place, Federal Way, Washington 98023 (US). **BRILES, David E.**; 760 Linwood Road, Birmingham, Alabama 35222 (US). **MUNZER, Jon Scott**; 15 Round Table Road, Shrewsbury, Massachusetts 01545 (US).

(74) Agents: **MEDINA, Rolando** et al.; Choate, Hall & Stewart LLP, Two International Place, Boston, Massachusetts 02110 (US).

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

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

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: FUSED ANTIGEN VACCINES AND COMPOSITIONS AGAINST STREPTOCOCCUS PNEUMONIAE

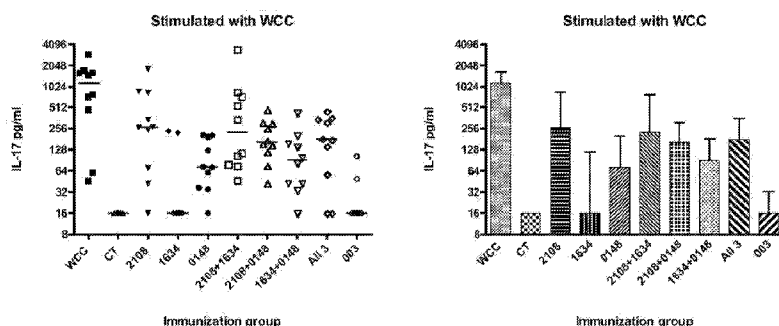


Figure 1

(57) Abstract: *Streptococcus pneumoniae* is a major health concern, especially in very young, elderly, or immunocompromised patients. The present disclosure provides, inter alia, certain highly effective vaccines and pharmaceutical compositions for *Streptococcus pneumoniae* that contain fusion proteins. The vaccines and pharmaceutical compositions may be used therapeutically and/or prophylactically.

FUSED ANTIGEN VACCINES AND COMPOSITIONS AGAINST STREPTOCOCCUS PNEUMONIAE

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. provisional application Serial No. 61/676,276, filed July 26, 2012, the contents of which are herein incorporated by reference in their entirety.

BACKGROUND

[0002] Pneumococcal disease continues to be a leading cause of sickness and death in the United States and throughout the world. Each year, millions of cases of pneumonia, meningitis, bacteremia, and otitis media are attributed to infection with the pathogen *Streptococcus pneumoniae*. *S. pneumoniae* is a Gram-positive encapsulated coccus that colonizes the nasopharynx in about 5-10% of healthy adults and 20-40% of healthy children. Normal colonization becomes pathogenic when *S. pneumoniae* is carried into the Eustachian tubes, nasal sinuses, lungs, bloodstream, meninges, joint spaces, bones, or peritoneal cavity. *S. pneumoniae* has several virulence factors that enable the organism to evade the immune system. Examples include a polysaccharide capsule that prevents phagocytosis by host immune cells, proteases that inhibit complement-mediated opsonization, and proteins that cause lysis of host cells. In the polysaccharide capsule, the presence of complex polysaccharides forms the basis for classifying pneumococci into different serotypes. To date, 93 serotypes of *S. pneumoniae* have been identified.

[0003] Various pharmaceutical compositions have been used to harness an immune response against infection by *S. pneumoniae*. A polyvalent pneumococcal vaccine, PPV-23, was developed for preventing pneumonia and other invasive diseases due to *S. pneumoniae* in the adult and aging populations. The vaccine contains capsular polysaccharides (CPs) from 23 serotypes of *S. pneumoniae*. As T-independent antigens, these CPs induce only short-lived antibody responses, necessitating repeated doses, which increases the risk of immunological tolerance. The antibodies raised against the *S. pneumoniae* capsular polysaccharides, termed anticapsular antibodies, are recognized as protective in adult and immunocompetent individuals. However, children under 2 years of age and immunocompromised individuals, including the elderly, do not respond well to T-independent antigens and, therefore, are not afforded optimal protection by PPV-23. Another *S. pneumoniae* vaccine, Prevnar, includes bacterial polysaccharides from 7 *S. pneumoniae*

serovars conjugated to the diphtheria toxoid protein. This vaccine induces both B and T cell responses. However, because it only protects against 7 pneumococcal serotypes, serotype replacement can render Prevnar ineffective against non-vaccine serovars. Serotype replacement has already been demonstrated in several clinical trials and epidemiologic studies, necessitating development of different formulations of these vaccines. An example is the recently introduced Prevnar 13, directed to 13 pneumococcal serotypes. Furthermore, the two Prevnar formulations are expensive to manufacture, greatly limiting their availability in the developing world. PPV-23, which consists of 23 purified but unconjugated polysaccharides, has broader coverage, but does not provide protection to children under the age of 2 years, a population which is at the highest risk for pneumococcal disease.

[0004] Thus, there remains a need to design more effective pharmaceutical compositions than the current strategies offer. In particular, there remains a need for such compositions incorporating novel or specific antigens that elicit an immune response against *S. pneumoniae*.

SUMMARY OF THE INVENTION

[0005] *Streptococcus pneumoniae* is a major health concern, especially in very young, elderly, or immunocompromised patients. While DNA and protein sequence information for *S. pneumoniae* have been known for some time, and researchers have long attempted to produce vaccines against *S. pneumoniae*, a major challenge was how to elicit an immune response that is both long-lived and effective. Fusion proteins described herein exhibit surprising and unexpected properties, including eliciting immune responses that are long-lived and/or effective. Certain of the *S. pneumoniae* antigens included in fusion proteins described herein were initially identified by screening immune cells from mice either immunized with unencapsulated killed whole *S. pneumoniae* or infected with *S. pneumoniae*, or from healthy human donors, for T-cell specific immune responses. In certain embodiments, such antigens are fused to *S. pneumoniae* capsular antigens, e.g., antigens that stimulate B cell specific immunity. Accordingly, the present disclosure provides, *inter alia*, certain highly effective vaccines against *S. pneumoniae*. The vaccines may be used therapeutically and/or prophylactically. The present disclosure also provides fusion proteins comprising specific antigens and methods for using such fusion proteins to elicit an immune response against *S. pneumoniae*.

[0006] In certain aspects, the present disclosure provides fusion proteins comprising or consisting of (i) one or more immunogenic polypeptides having or comprising the amino acid sequence of one of SEQ ID NOS: 1-8 or an immunogenic fragment or variant thereof; and (ii) one or more *S. pneumoniae* antigens that are predominantly an antibody target. In some embodiments, the immunogenic polypeptide is N-terminal to the *S. pneumoniae* antigen. In certain embodiments, the immunogenic polypeptide is lipidated. In some embodiments, the immunogenic polypeptide has or comprises the amino acid sequence of SEQ ID NO: 2. In some embodiments, the immunogenic polypeptide has or comprises the amino acid sequence of SEQ ID NO: 6.

[0007] In certain embodiments, an immunogenic polypeptide of a fusion protein described herein is directly bonded to a *S. pneumoniae* antigen. In some embodiments, a fusion protein described herein further comprises a linker polypeptide. In some embodiments, a linker polypeptide is positioned between the immunogenic polypeptide and the *S. pneumoniae* antigen. In some embodiments, the linker polypeptide is rigid. In certain embodiments, the linker polypeptide comprises SEQ ID NO: 93. In some embodiments, the linker polypeptide is semi-rigid. In certain embodiments, the linker polypeptide comprises SEQ ID NO: 94.

[0008] In some embodiments, the *S. pneumoniae* antigen of a fusion protein described herein is Pneumococcal surface protein A (PspA) or a derivative thereof. In some embodiments, the PspA derivative is or comprises all or a fragment of the proline-rich region of PspA. In certain embodiments, the PspA derivative is or comprises a non-proline block. In some embodiments, the PspA derivative is or comprises SEQ ID NO: 29. In certain embodiments, the PspA derivative is or comprises SEQ ID NO: 33.

[0009] In some embodiments, the *S. pneumoniae* antigen of a fusion protein described herein is or comprises pneumolysin, a pneumolysoid, or a derivative thereof.

[0010] In additional aspects, the present disclosure provides vaccine formulations comprising one or more fusion protein described herein and a pharmaceutically acceptable carrier.

[0011] In some embodiments, a vaccine formulation comprises a fusion protein described herein and a second immunogenic polypeptide having or comprising an amino acid sequence of one of SEQ ID NOS: 1-11 or an immunogenic fragment thereof. In some embodiments, a vaccine formulation comprises a first fusion protein described herein and a second fusion protein comprising two immunogenic polypeptides, each having or comprising an amino acid sequence of one of SEQ ID NOS: 1-11 or an immunogenic fragment thereof.

[0012] In some embodiments, a vaccine formulation described herein further comprises an adjuvant. In some embodiments, the adjuvant is an agonist of toll-like receptors (TLRs). In some embodiments, the adjuvant is alum.

[0013] In certain embodiments, a vaccine formulation described herein comprises 1-1000 µg of each fusion protein and/ or polypeptide and 25-500 µg of the adjuvant.

[0014] In some aspects, compositions and methods described herein are used for prophylaxis and/or treatment of any pneumococcal disease, disorder, and/or condition due to a pneumococcal infection. In some embodiments, a vaccine composition described herein reduces risk of infection by, and/or treats, alleviates, ameliorates, relieves, delays onset of, inhibits progression of, reduces severity of, and/or reduces incidence of one or more symptoms or features of a pneumococcal disease, disorder, and/or condition. In some embodiments, prophylaxis and/or treatment of pneumococcal infection comprises administering a therapeutically effective amount of a vaccine composition described herein to a subject in need thereof, e.g., in such amounts and/or for such time as is necessary to achieve the desired result. In certain embodiments, a "therapeutically effective amount" of a vaccine composition described herein is an amount effective for treating, alleviating, ameliorating, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of pneumococcal infection.

[0015] In some embodiments, prophylactic, prognostic and/or therapeutic methods involve administering a therapeutically effective amount of one or more vaccine compositions described herein to a subject such that an immune response is stimulated in one or both of T cells and B cells.

[0016] In some embodiments, a vaccine formulation described herein is characterized in that, upon contacting a T_H17 cell, the vaccine formulation induces a T_H17 cell response at least 1.5-fold greater than that induced by a control (e.g., an unrelated antigen) after contacting T_H17 cells. In some embodiments, a vaccine formulation described herein inhibits infection by *S. pneumoniae* in an uninfected subject (e.g., the vaccine formulation is characterized in that, upon administration to an uninfected subject, the vaccine formulation inhibits infection by *S. pneumoniae* in the uninfected subject). In certain embodiments, a vaccine formulation described herein inhibits *S. pneumoniae* colonization in a subject (e.g., the vaccine formulation is characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae* colonization in the subject). In some embodiments, a

vaccine formulation described herein inhibits *S. pneumoniae* symptoms (e.g., the vaccine formulation is characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae* symptoms in the subject). In certain embodiments, a vaccine formulation described herein inhibits *S. pneumoniae*-induced invasive disease (e.g., the vaccine formulation is characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae*-induced invasive disease in the subject). In certain embodiments, the vaccine formulation inhibits *S. pneumoniae*-induced sepsis (e.g., the vaccine formulation is characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae*-induced sepsis in the subject).

[0017] In further aspects, the present disclosure provides methods for treating a subject suffering from or susceptible to *S. pneumoniae* infection, comprising administering to the subject an effective amount of a vaccine formulation described herein.

[0018] In some embodiments, the subject is an uninfected subject, and the method inhibits infection by *S. pneumoniae* in the uninfected subject. In certain embodiments, the method inhibits *S. pneumoniae* colonization in a subject. In some embodiments, the method inhibits *S. pneumoniae* symptoms. In some embodiments, the method inhibits *S. pneumoniae*-induced invasive disease. In certain embodiments, the method inhibits *S. pneumoniae*-induced sepsis.

[0019] In some embodiments, the method treats a subject with one dose. In certain embodiments, the method treats a subject within three doses. In some embodiments, the subject is a human.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The present teachings described herein will be more fully understood from the following description of various illustrative embodiments, when read together with the accompanying drawings. It should be understood that the drawings described below are for illustration purposes only and are not intended to limit the scope of the present teachings in any way.

[0021] FIG. 1 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then stimulated with killed, unencapsulated whole cell *S. pneumoniae*, as described in Example 1. The left panel shows the data in scatter format, and the right panel shows the average and standard deviation for each sample. Immunization group “All 3” represents animals immunized with a combination of SP2108, SP0148, and SP1634.

[0022] FIG. 2 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then stimulated with a combination of three proteins (SP2108, SP0148, and SP1634), as described in Example 1.

[0023] FIG. 3 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then challenged with intranasal administration of *S. pneumoniae*, as described in Example 1. 003 represents a control unrelated antigen.

[0024] FIG. 4 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in BALB/c mice that were immunized with the indicated protein(s) and cholera toxin adjuvant, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 2.

[0025] FIG. 5 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated proteins and cholera toxin adjuvant, then stimulated with the protein of immunization (A) or killed, unencapsulated whole cell *S. pneumoniae* (B), as described in Example 3.

[0026] FIG. 6 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and cholera toxin adjuvant, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 4.

[0027] FIG. 7 shows survival of mice that were immunized with the indicated proteins and the adjuvant alum, then underwent aspiration challenge with *S. pneumoniae* as described in Example 5.

[0028] FIG. 8 shows survival of mice that were immunized with the indicated proteins and the adjuvant alum, then underwent aspiration challenge with *S. pneumoniae* as described in Example 6.

[0029] FIG. 9 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and cholera toxin adjuvant, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 7.

[0030] FIG. 10 shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated proteins and alum, then stimulated with the proteins indicated at upper left, as described in Example 8.

[0031] FIG. 11 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed,

unencapsulated whole cell *S. pneumoniae* plus alum (WCV), then challenged by intranasal administration of *S. pneumoniae*, as described in Example 9.

[0032] FIG. 12 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), then challenged by intranasal administration of *S. pneumoniae*, in two pooled studies as described in Example 10.

[0033] FIG. 13 shows the number of *S. pneumoniae* colonies obtained from a nasal wash in mice that were immunized with the indicated proteins and alum or with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCB), then challenged by intranasal administration of *S. pneumoniae*, as described in Example 11.

[0034] FIG. 14 shows survival of mice that were injected with antibodies or sera specific to the indicated proteins, then underwent aspiration challenge with *S. pneumoniae*, as described in Example 12.

[0035] FIG. 15 shows the percent of animals protected from sepsis in six separate aspiration challenge studies, two of which are described in more detail in Examples 6 and 12.

[0036] FIG. 16A, 16B, and 16C show the mean IL-17A response generated by blood samples from mice that were immunized with the indicated proteins and alum, then stimulated with the proteins indicated at upper right, as described in Example 13.

[0037] FIG. 17 shows the combined number of *S. pneumoniae* colonies, obtained in two separate studies, from nasal washes of mice that were immunized with the indicated fusion proteins adsorbed to alum, then challenged by intranasal administration of *S. pneumoniae*, as described in Example 16.

[0038] FIG. 18 shows hours to moribund for mice that were immunized with the indicated fusion proteins adsorbed to alum, then challenged by intravenous administration of *S. pneumoniae*, as described in Example 17.

[0039] FIG. 19A and 19B shows the concentration of IL-17A generated by blood samples from mice that were immunized with the indicated fusion proteins adsorbed to alum, following stimulation with killed, unencapsulated whole cell *S. pneumoniae*, as described in Example 18.

[0040] FIG. 20 shows the total IgG concentration against CD2 or H70 in blood samples from mice immunized with the indicated fusion proteins adsorbed to alum, then challenged by intravenous administration of *S. pneumoniae*, as described in Example 19.

DETAILED DESCRIPTION

A. Fusion Proteins

[0041] This application describes vaccine formulations that include one or more *S. pneumoniae* fusion proteins, e.g., fusion proteins that include one or more *S. pneumoniae* polypeptides or fragments thereof. The fusion proteins can include a first polypeptide that elicits primarily a T cell response and a second polypeptide that elicits primarily a B cell (e.g., antibody) response. In some embodiments, the first polypeptide elicits an immune response that is primarily a T cell response. For example, when assayed using a suitable immunogenicity assay (e.g., an assay described herein), more than 50%, 60%, 70%, 80%, 90%, or more, of the immune response elicited by the first polypeptide is a T cell response. In some embodiments, the second polypeptide elicits an immune response that is primarily a B cell response. For example, when assayed using a suitable immunogenicity assay (e.g., an assay described herein), more than 50%, 60%, 70%, 80%, 90%, or more, of the immune response elicited by the second polypeptide is a B cell response.

[0042] In certain embodiments, a fusion protein described herein includes one, two or more of the polypeptides listed (or polypeptides encoded by the genes listed) in Table 1 and/or 2. In certain embodiments, a fusion protein includes a polypeptide listed (or encoded by a gene listed) in Table 1 fused to a polypeptide listed (or encoded by a gene listed) in Table 2.

[0043] In some embodiments, a fusion protein comprises an immunogenic peptide having or comprising an amino acid sequence of one of SEQ ID NOS: 1-8, shown in Table 1, or an immunogenic fragment thereof; and a *S. pneumoniae* antigen that is or comprises predominantly an antibody target. In some embodiments, the immunogenic peptide is SP2108 (SEQ ID NO: 6). In some embodiments, the immunogenic peptide is SP0148 (SEQ ID NO: 2). In some embodiments, a fusion protein includes a *S. pneumoniae* antigen that is or comprises predominantly an antibody target, e.g., an antigen shown in Table 2, e.g., PspA (SEQ ID NO: 27) or fragments or variants thereof. In some embodiments, fragments or variants of PspA comprise proline-rich segments with the non-proline block (PR+NPB), for example the CD2 sequence SEQ ID NO: 29. In some embodiments, fragments or variants of PspA comprise proline-rich segments with the non-proline block and 10, 20 30, 40 or more additional amino acids of PspA sequence, for example the H70 sequence (SEQ ID NO: 33). In some embodiments, the *S. pneumoniae* antigen that is predominantly an antibody target comprises pneumolysin or a pneumolysoid.

[0044] Fusion proteins described herein can include polypeptides homologous to the polypeptides of Tables 1 (for example, SP1912, SP1912L, SP0148 with or without a signal sequence, SP2108 with or without a signal sequence) and Table 2 (for example, PspA and fragments and variants thereof). Individual strains of *S. pneumoniae* contain numerous mutations relative to each other, and some of these result in different protein sequences between the different strains. One of skill in the art may readily substitute an amino acid sequence, or a portion thereof, with the homologous amino acid sequence from a different *S. pneumoniae* strain. In certain aspects, a fusion protein described herein includes one or more polypeptides with at least 90%, 95%, 97%, 98%, 99%, or 99.5% identity to the polypeptides of Tables 1 and 2 or an immunogenic fragment thereof. Serotypic variation may be used to design such variants of the polypeptides of Tables 1 and 2.

[0045] In some embodiments, a fusion protein described herein includes a fragment of a polypeptide of Table 1 (for example, fragments of SP1912, SP1912L, SP0148 with or without a signal sequence, SP2108 with or without a signal sequence), and/or a fragment of a polypeptide of Table 2 (for example, PspA and variants thereof). In some embodiments, a fusion protein herein comprises one or more truncation mutants that are close in size to the polypeptides of Tables 1 and 2. For example, they may lack at most one, two three, four, five, ten, or twenty amino acids from one or both termini (referring to component polypeptides in a fusion protein). In certain embodiments, the fragment is a truncated fragment of any of SEQ ID NOS: 1-8 or 27-36, wherein from 1-5, 1-10, or 1-20 amino acid residues are removed from the N-terminus, C-terminus, or both. In certain embodiments, the fragment is a truncated fragment of any of SEQ ID NOS: 1-8 or 27-36, wherein from 1-10 amino acid residues are removed from the N-terminus, C-terminus, or both. For instance, 10 amino acid residues may be removed from each of the N-terminus and C-terminus resulting in a protein with 20 amino acid residues removed. Internal deletions, e.g., of 1-10, 11-20, 21-30, or 31-40 amino acids, are also contemplated.

Table 1. Immunogenic polypeptides for vaccine formulations

Locus tag name and description	Protein SEQ ID No.	DNA SEQ ID No.	DNA GenBank Accession No. (from March 30, 2010)
SP0148 lacking signal sequence	1	37	-
SP0148 including signal sequence	2	38	NC_003028.3 :145,513-

			146,343*
SP0148 consensus lacking signal sequence	3	-	-
SP0148 consensus including signal sequence	4	-	-
SP2108 lacking signal sequence	5	39	-
SP2108 including signal sequence	6	-	NC_003028.3 :2020750-2022021
SP2108 consensus lacking signal sequence	7	-	-
SP2108 consensus including signal sequence	8	-	-
SP1912	9	40	NC_003028.3 :1824672-1824971
SP1912L	10	41	-
SP1912 consensus	11	-	-
SP0641	12		
SP0641N	13	42	
SP0641M	14		-
SP0641N consensus	15	-	-
SP0641M consensus	16	-	-
SP0882	17	-	NC_003028.3 :831804-832628
SP0882N	18	43	-
SP0882 with exogenous signal sequence	19	44	-
SP0882N with exogenous signal sequence	20	45	-
SP0882 consensus	21	-	-
SP0882N consensus	22	-	-
SP0882 consensus with exogenous leader	23	-	-
SP0882N consensus with exogenous	24	-	-

leader			
SP1634	25	-	NC_003028.3 :1534348-1535421
SP0314	26	-	NC_003028.3 :287483-290683

Table 2. Immunogenic polypeptides for vaccine formulations

PspA	27	46	
PR+NRB from PspA with coiled-coil	28		
CD2	29		
PR+NRB from PspA w/o coiled-coil	30		
PR only with coiled-coil	31		
PR only w/o coiled-coil	32		
H70 (PR+NRB from PspA aa 290-410)	33		
Non-proline Block (NPB)	34		
Non-proline Block (NPB)	35		
Non-proline Block (NPB)	36		

*NB: The database sequence incorrectly lists TTG (encoding Leu) at nucleotide positions 541-543. The correct sequence, as shown in SEQ ID NO: 38, has TTC at that codon and encodes Phe. The database sequence further does not include a C-terminal Glu found in certain isolates.

[0046] Particular examples of fusion proteins are provided in Table 3.

Table 3. Immunogenic fusion proteins for vaccine formulations

Locus tag names	Protein SEQ ID No.	DNA SEQ ID No.
SP2108/SP0148	47	67
SP0148/SP2108	48	68
SP2108/SP1912	49	69
SP0148/SP1912	50	70
SP2108/SP1912/SP0148	51	71

SP0148/SP1912/SP2108	52	72
SP2108/SP0148/SP1912	53	73
SP0148/SP2108/SP1912	54	74
SP0148/CD2	55	75
SP0148/H70	56	76
SP2108/CD2	57	77
SP2108/ H70	58	78
SP0148/LC/CD2	59	79
SP0148/ LC/H70	60	80
SP2108/ LC/CD2	61	81
SP2108/LC/H70	62	82
SP0148/LR/CD2	63	83
SP0148/LR/H70	64	84
SP2108/LR/CD2	65	85
SP2108/LR/H70	66	86

[0047] In some embodiments, the fusion protein comprises an N-terminal peptide and a C-terminal peptide. In some embodiments, the N-terminal peptide comprises the immunogenic polypeptide, for example, a polypeptide having an amino acid sequence of one of SEQ ID NOS: 1-8 or immunogenic fragments or variants thereof, and the C-terminal peptide comprises a *S. pneumoniae* antigen that is predominantly an antibody target as described herein.

[0048] In some embodiments, the N-terminal peptide and the C-terminal peptide are directly bound to each other. In some embodiments, the N-terminal peptide and the C-terminal peptide are linked via a linker peptide. The length and/or amino acids of a linker, when present, can be adjusted to obtain a more flexible or rigid linker. Exemplary peptide linkers are shown as SEQ ID NOS: 92-94. A linker can generally be from 1-40, such as 10-30 and specifically 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids in length.

[0049] In some embodiments, the fusion protein comprises an immunogenic fragment of the fusion proteins shown in Table 3. In some embodiments, the fragment is a truncated fusion protein of any of SEQ ID NOS: 47-66. For example, they may lack at most one, two three, four, five, ten, or twenty amino acids from the N-terminus, C-terminus, or both. In some such embodiments, the same number of residues is removed from the N-terminus and the C-terminus, while in other embodiments, a different number of residues is removed from

the N-terminus compared to the C-terminus. Advantageously, in some embodiments the N-terminal polypeptide is not truncated so as not to prevent lipidation.

[0050] In certain aspects, this application provides fusion proteins with at least 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a fusion protein of Table 3. In certain embodiments, a vaccine formulation comprises at least two different polypeptides having an amino acid sequence comprising a sequence at least 90%, 95%, 98%, or 99% identical to any of SEQ ID NOS: 47-66, or an immunogenic fragment thereof.

[0051] In certain embodiments, vaccines or pharmaceutical compositions described herein include one or more fusion proteins containing at least one lipidated polypeptide. In some embodiments, the fusion protein is lipidated. In certain embodiments, the fusion protein is lipidated on the N-terminal peptide. Conjugation to the lipid moiety may be direct or indirect (e.g., via a linker). The lipid moiety may be synthetic or naturally produced. In certain embodiments, a polypeptide from Table 1, 2 or 3 may be chemically conjugated to a lipid moiety. In certain embodiments, a construct may comprise a gene or polypeptide from Table 1, 2 or 3, or an immunogenic fragment or variant thereof, and a lipidation sequence including a lipobox motif. A canonical lipobox motif is shown as SEQ ID NO: 89. A lipidation sequence may be N-terminal or C-terminal to the protein, and may be embedded in a signal or other sequence, or in a fusion protein. Exemplary lipidation sequences include the signal sequence of SP2108 (SEQ ID NO: 90) and the signal sequence of the *E. coli* gene RlpB (SEQ ID NO: 91). A signal sequence may be, for example, an *E. coli* or *S. pneumoniae* signal sequence. Exemplary *E. coli* signal sequences include the mlpA signal sequence (Lin, J.J. *et al.*, “An *Escherichia coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein” Proc Natl Acad Sci U S A. 1978 Oct;75(10):4891-5), the lamB signal sequence (Emr, S.D. *et al.* “Mutations altering the cellular localization of the phage lambda receptor, an *Escherichia coli* outer membrane protein”, Proc Natl Acad Sci U S A. 1978 Dec;75(12):5802-6), or the MBP signal sequence (Bassford, P.J., “Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm” J Bacteriol. 1979 Jul;139(1):19-31). Lpp is an exemplary *E. coli* signal sequence that directs lipidation (Cullen, P.A. *et al.* “Construction and evaluation of a plasmid vector for the expression of recombinant lipoproteins in *Escherichia coli*” Plasmid. 2003 Jan;49(1):18-29.) *E. coli* signal sequences that direct lipidation are also described in Legrain, M. *et al.* (“Production of lipidated meningococcal transferrin binding protein 2 in *Escherichia coli*” Protein Expr Purif. 1995 Oct;6(5):570-8), e.g. the signal sequence of the

gene RlpB (SEQ ID NO: 91) Numerous *S. pneumoniae* signal sequences are known in the art. One such signal sequence is SEQ ID NO: 90.

[0052] In certain embodiments, vaccines or pharmaceutical compositions described herein include a fusion protein comprising a polypeptide encoded by a polynucleotide from Table 1, 2, or 3, a polypeptide from Table 1, 2 or 3, or an immunogenic fragment or variant thereof, and a tag. A tag may be N-terminal or C-terminal. For instance, tags may be added to the nucleic acid or polypeptide to facilitate purification, detection, solubility, or confer other desirable characteristics on the protein or nucleic acid. For instance, a purification tag may be a peptide, oligopeptide, or polypeptide that may be used in affinity purification. Examples include His, GST, TAP, FLAG, myc, HA, MBP, VSV-G, thioredoxin, V5, avidin, streptavidin, BCCP, Calmodulin, Nus, S tags, lipoprotein D, and β galactosidase. Particular exemplary His tags include HHHHHH (SEQ ID NO: 87) and MSYYHHHHHH (SEQ ID NO: 88). In other embodiments, the polypeptide is free of tags such as protein purification tags, and is purified by a method not relying on affinity for a purification tag. In some embodiments, the fused portion is short. This, in some instances, the fusion protein comprises no more than 1, 2, 3, 4, 5, 10, or 20 additional amino acids on one or both termini of the polypeptide of Table 1, 2 or 3.

[0053] In certain embodiments, this application provides nucleic acids encoding one or more of the fusion proteins described herein, such as DNA (e.g., cDNA), RNA, or an analog thereof. The underlying DNA sequences for the polypeptides described above may be modified in ways that do not affect the sequence of the protein product, and such sequences are included in the invention. For instance, the DNA sequence may be codon-optimized to improve expression in a host such as *E. coli*, an insect cell line (e.g., using the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line.

[0054] In certain embodiments, this application provides polynucleotides (such as DNA, cDNA, RNA, or an analog thereof) that are at least 70%, 80%, 90%, 95%, 97%, 98%, 99%, or 100% identical to in the polynucleotide sequences of Table 3, or a variant or portion of said sequence. In certain embodiments, the polynucleotide is 600-2000, 800-1800, 1000-1600, 1200-1400 nucleotides in length. In some embodiments, the polynucleotide is 600-1600, 800-1800, 1000-2000, 2000-3000, or 3000-4000 nucleotides in length. The nucleic acids may be used, for example, for recombinant production of the polypeptides of Table 1 or 2, or the fusion proteins of Table 3, or immunogenic fragments thereof.

[0055] Polynucleotides encoding peptides of Tables 1, 2 or 3 or fragments thereof can be cloned into any of a variety of expression vectors, under the control of a variety of regulatory elements, and fusions can be created with other peptides of Tables 1, 2, or 3 or with other sequences of interest. Methods of cloning nucleic acids are routine and conventional in the art. For general references describing methods of molecular biology which are mentioned in this application, e.g., isolating, cloning, modifying, labeling, manipulating, sequencing and otherwise treating or analyzing nucleic acids and/or proteins, see, e.g., Sambrook, J. et al. (1989). *Molecular Cloning, a Laboratory Manual*. Cold Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Ausubel, F. M. et al. (1995). *Current Protocols in Molecular Biology*, N.Y., John Wiley & Sons; Davis et al. (1986), *Basic Methods in Molecular Biology*, Elsevier Sciences Publishing, Inc., New York; Hames et al. (1985), *Nucleic Acid Hybridization*, IL Press; Dracopoli, N.C. et al. *Current Protocols in Human Genetics*, John Wiley & Sons, Inc.; and Coligan, J. E., et al. *Current Protocols in Protein Science*, John Wiley & Sons, Inc.

B. Specific polypeptides and nucleic acids for use in *S. pneumoniae* vaccines and immunogenic compositions

[0056] This application describes *S. pneumoniae* fusion proteins. The fusion proteins include one, two or more of the polypeptides (or polypeptides encoded by the genes) listed in Table 1 and/or 2, or variants or fragments thereof as described below. In some embodiments, the fusion proteins contain one or more polypeptides from Table 1 and one or more polypeptides from Table 2. The DNA and protein sequence of each gene and polypeptide may be found by searching for the Locus Tag in the publicly available database, Entrez Gene (on the NCBI NIH web site on the World Wide Web, at www.ncbi.nlm.nih.gov/sites/entrez?db=gene), in the *Streptococcus pneumoniae* TIGR4 genome, and the indicated sequences are also included in this application.

[0057] Certain polypeptides of Table 1 and 2, and variants thereof, are described in greater detail below.

1. SP0148 (SEQ ID NO: 2) and variants thereof

[0058] The protein SP0148 is named “ABC transporter, substrate-binding protein”. Proteins of this class are typically extracellular proteins that interact transiently with a transmembrane protein complex. Such complexes use energy generated by ATP hydrolysis to translocate specific substrates across a cell membrane. SP0148 is a 276 or 277 (depending on the isolate) amino acid protein that contains a conserved PBPb (periplasmic binding

protein) domain, spanning amino acids 40-246, which is typical of membrane-bound transport complexes. In addition, SP0148 has a bacterial extracellular solute-binding proteins family 3 domain which is largely co-extensive with the PBPb domain and extends from amino acid 40 to 244. In some embodiments, a vaccine or other composition comprises a truncation mutant of SP0148 comprising or lacking one or more of said domains and motifs.

[0059] In some embodiments, vaccines or pharmaceutical compositions comprising an *S. pneumoniae* polypeptide include a fusion protein containing a polypeptide containing at least 20 consecutive amino acid residues selected from SP0148. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 275, 250, 225, 200, 175, 150, 125, or 100 consecutive amino acids from SP0148.

[0060] Endogenous SP0148 comprises a signal sequence that directs its secretion and potential lipidation. In some embodiments, the signal sequence of the polypeptide of SEQ ID NO: 2 is partially or fully processed by an expression host, e.g. *E. coli*. In some embodiments, a variant of SP0148 that lacks the signal sequence (SEQ ID NO: 1) is used. The polypeptide of SEQ ID NO: 1 is encoded by the nucleic acid of SEQ ID NO: 37, although other nucleic acid sequences (including codon-optimized sequences) may be used. SEQ ID NO: 38 encodes the full length sequence of SP0148 used in the screens herein.

[0061] Variants of the amino acid sequence and nucleotide sequence of SP0148 may be found in U.S. Patent Application Publication No. 2005/0020813, U.S. Patent Nos. 7,378,514 and 7,504,110, and European Patent Application No. EP1572868 and EP1855717.

[0062] Consensus sequences illustrating combinations of SP0148 sequences from different *S. pneumoniae* serotypes are provided as SEQ ID NOS: 3 and 4. Accordingly, in certain embodiments, the vaccine formulation comprises a fusion protein that includes a polypeptide having an amino acid sequence comprising, or consisting of, either of SEQ ID NOS: 3-4, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising one of SEQ ID NOS: 1 or 2).

[0063] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP0148. Exemplary immunogenic fragments of SP0148 include ALGLVAAGV (SEQ ID NO: 97), ELTGYEIEV (SEQ ID NO: 98), AVNNLSYTK (SEQ ID NO: 99), TYLP AEADI (SEQ ID NO: 100), RYNMAVNNL (SEQ ID NO: 101), DFQQIMVRL (SEQ ID NO: 102), EHTDNPTIL (SEQ ID NO: 103), APIAQNPVN (SEQ ID NO: 104), LPSDQQPYV (SEQ ID NO: 105), YVYPLLAQG (SEQ ID NO: 106), QGLDNLKVI (SEQ ID NO: 107), KYLYAAPI (SEQ ID NO: 108), GELTGYEI (SEQ ID

NO: 109), NPNVLVVKK (SEQ ID NO: 110), KLSKQFFGD (SEQ ID NO: 111), GSPRPFIYE (SEQ ID NO: 112), AVNNLSYTK (SEQ ID NO: 113), KIFDKIGVE (SEQ ID NO: 114), MVRLSDGQF (SEQ ID NO: 115), YVYPLLAQG (SEQ ID NO: 116), VVQATTSAK (SEQ ID NO: 117), TLEKLSKQF (SEQ ID NO: 118), VAAGVLAAC (SEQ ID NO: 119), LDNLKVIEL (SEQ ID NO: 120), and NMAVNNLSY (SEQ ID NO: 121).

2. SP2108 (SEQ ID NO: 6) and variants thereof

[0064] The polypeptide SP2108 is 423 amino acids in length and is alternatively known as MalX, maltose/maltodextrin ABC transporter, or maltose/maltodextrin-binding protein. Much of the protein (amino acids 3-423) is classified as a MalE (Maltose-binding periplasmic) domain. In addition, SP2108 contains a signal sequence that directs its secretion and potential lipidation. In some embodiments, the signal sequence of the polypeptide of SEQ ID NO: 6 is partially or fully processed by an expression host, e.g. *E. coli*. In some embodiments, a vaccine or other composition comprises a truncation mutant of SP2108 comprising one or more of said domains and motifs.

[0065] In some embodiments, the compositions and methods herein call for the use of an SP2108 variant that lacks the signal sequence. This variant is represented by polypeptide sequence SEQ ID NO: 5 and may be encoded by, for example, a nucleic acid according to SEQ ID NO: 39, although due to degeneracy in the genetic code, other DNA sequences (including codon-optimized sequences) may be used.

[0066] In some embodiments, vaccines or pharmaceutical compositions comprising an *S. pneumoniae* polypeptide include a fusion protein containing a polypeptide containing at least 20 consecutive amino acid residues selected from SP2108. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 400, 350, 300, 250, 200, 150, or 100 consecutive amino acids from SP2108.

[0067] Consensus sequences illustrating combinations of SP2108 sequences from different serotypes are provided as SEQ ID NOS: 7 and 8. Thus, in certain embodiments, the vaccine formulation comprises a fusion protein containing a polypeptide having an amino acid sequence comprising, or consisting of, either of SEQ ID NOS: 7-8, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising one of SEQ ID NOS: 5 or 6).

[0068] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP2108. Exemplary immunogenic fragments of SP2108 include AIIDGPWKA (SEQ ID NO: 122), VMMAFYDRV (SEQ ID NO: 123), SIAGINYAK (SEQ

ID NO: 124), VWDPKAKNML (SEQ ID NO: 125), QPLPNISQM (SEQ ID NO: 126), APYDRVGS (SEQ ID NO: 127), APAVIESLV (SEQ ID NO: 128), FYYTYGLLA (SEQ ID NO: 129), SKYAFAGE (SEQ ID NO: 130), TEGAGNLI (SEQ ID NO: 131), LADWTNFIYY (SEQ ID NO: 132), SLVMYYNKD (SEQ ID NO: 133), KEAGVKVTL (SEQ ID NO: 134), KSTAVLGTV (SEQ ID NO: 135), GAKTDDTTK (SEQ ID NO: 136), SQKFVDFLV (SEQ ID NO: 137), QAFKDAKVN (SEQ ID NO: 138), AVIESLVMY (SEQ ID NO: 139), DAKTAANDA (SEQ ID NO: 140), YGVATIPTL (SEQ ID NO: 141), KTAAIIDGP (SEQ ID NO: 142), KAYEKEAGV (SEQ ID NO: 143), AGNGAYVFG (SEQ ID NO: 144), and AWWIPQAVK (SEQ ID NO: 145).

3. SP1912 (SEQ ID NO: 9) and variants thereof

[0069] SP1912 is a hypothetical protein of 99 amino acids. While the protein function is not definitively known, sequence analysis suggests it is a putative thioredoxin.

[0070] In some embodiments, vaccines or pharmaceutical compositions comprising an *S. pneumoniae* polypeptide include a polypeptide containing at least 20 consecutive amino acid residues selected from SP1912. The polypeptide may also be a variant of the at least 20 residue fragment. In certain embodiments, the polypeptide includes no more than 90, 75, 60, 45 or 30 consecutive amino acids from SP1912.

[0071] In some embodiments, the compositions and methods herein call for the use of an SP1912 variant that comprises an exogenous lipidation sequence. In some embodiments, a signal sequence directs lipidation. Thus, the lipidation signal may be, e.g., the signal sequence of SP2108 (SEQ ID NO: 90) or SP0148, or an *E. coli* signal sequence. The exemplary variant SP1912L, comprising the signal sequence of the *E. coli* gene RlpB (SEQ ID NO: 91) is represented by polypeptide sequence SEQ ID NO: 10. SP1912 (SEQ ID NO: 9) and SP1912L (SEQ ID NO: 10) may be encoded, respectively, by nucleic acids according to SEQ ID NO: 40 and 41, although due to degeneracy in the genetic code, other DNA sequences (including codon-optimized sequences) may be used. In some embodiments, the lipidation sequence is provided by fusing a full-length polypeptide such as SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) to SP1912. Exemplary fusions are SEQ ID NOS: 49-54.

[0072] Consensus sequences illustrating combinations of SP1912 sequences from different serotypes are provided as SEQ ID NO: 11. Thus, in certain embodiments, the vaccine formulation comprises a polypeptide having an amino acid sequence comprising, or consisting of, SEQ ID NO: 11, or an immunogenic fragment thereof (e.g., in place of a polypeptide having an amino acid sequence comprising SEQ ID NO: 9).

[0073] In some embodiments a fusion protein comprises or consists of an immunogenic fragment of SP1912. Exemplary immunogenic fragments of SP1912 include KMWMAGLALLGIGSL (SEQ ID NO: 146), LLGIGSLALATKKVA (SEQ ID NO: 147), MAGLALLGIGSLALA (SEQ ID NO: 148), WMAGLALLGIGSLAL (SEQ ID NO: 149), GLALLGIGSLALATK (SEQ ID NO: 150), LALLGIGSLALATKK (SEQ ID NO: 151), FSDMGEIATLYVQVY (SEQ ID NO: 152), KAKKMWMAGLALLGI (SEQ ID NO: 153), ALLGIGSLALATKKVAK (SEQ ID NO: 154), SDMGEIATLYVQVYE (SEQ ID NO: 155), DMGEIATLYVQVYES (SEQ ID NO: 156), AGLALLGIGSLALAT (SEQ ID NO: 157), MGEIATLYVQVYESS (SEQ ID NO: 158), KKMWMAGLALLGIGS (SEQ ID NO: 159), GMKAKKMWMAGLALL (SEQ ID NO: 160), MKAKKMWMAGLALLG (SEQ ID NO: 161), HFSDMGEIATLYVQV (SEQ ID NO: 162), MNGMKAKKMWMAGLA (SEQ ID NO: 163), MWMAGLALLGIGSLA (SEQ ID NO: 164), DHFSDMGEIATLYVQ (SEQ ID NO: 165), RDHFSDMGEIATLYV (SEQ ID NO: 166), NGMKAKKMWMAGLAL (SEQ ID NO: 167), and KMWMAGLALLGIG (SEQ ID NO: 168).

4. Additional Antigens

[0074] In certain embodiments, the vaccine formulations provided herein comprise or further comprise one or more, or two or more, known *S. pneumoniae* antigens. In some embodiments, vaccines or pharmaceutical compositions comprising an *S. pneumoniae* polypeptide include a fusion protein containing a known *S. pneumoniae* antigen. In some instances, the known *S. pneumoniae* antigens are predominantly antibody targets. In some instances, the known *S. pneumoniae* antigens protect from *S. pneumoniae* colonization, or from *S. pneumoniae*-induced sepsis. One appropriate art-recognized class of *S. pneumoniae* antigen is Pneumococcal surface protein A (PspA) (SEQ ID NO: 27) and derivatives of PspA. Derivatives of PspA include proline-rich segments with the non-proline block (PR+NPB, also referred to as PRN and further described below as well as in Daniels, C.C. *et al.* (2010) *Infection and Immunity* 78:2163-72) and related constructs comprising all or a fragment of the proline-rich region of PspA (e.g., regions containing one or more of the sequences PAPAP, PKP, PKEPEQ and PEKP and optionally including a non-proline block). In some embodiments, fragments or variants of PspA comprise proline-rich segments with the non-proline block and 10, 20 30, 40 or more additional amino acids of PspA sequence. H70 (SEQ ID NO: 33) is one exemplary sequence which includes the proline-rich region and non-proline-block encompassing amino acids 290-410 of PspA. An example of the non-proline-block has the exemplary sequence EKSADQQAEEYARRSEEEYNRLTQQQ

(SEQ ID NO: 34), which generally has no proline residues in an otherwise proline-rich area of the non-coiled region of PspA. Other embodiments of non-proline block (NPB) sequences include SEQ ID NOS: 35 and 36 and PspA and its derivatives can include genes expressing similar proline-rich structures (i.e. PKP, PKEPEQ and PEKP), with or without the NPB. The amino acids at either end of the NPB mark the boundaries of the proline-rich region. In one example, the amino-terminal boundary to the PR-region is DLKKAVNE (SEQ ID NO: 95), and the carboxy-terminal boundary is (K/G)TGW(K/G)QENGMW (SEQ ID NO: 96).

Peptides containing the NPB are particularly immunogenic, suggesting that the NPB may be an important epitope. Exemplary immunogenic PspA polypeptide derivatives containing the coiled-coil structure include SEQ ID NOS: 28 and 31. Particular embodiments of the immunogenic PspA polypeptide derivatives lacking the coiled-coil structure have the amino acid sequences shown as SEQ ID NOS: 29, 30 and 32. Immunogenic PspA polypeptides SEQ ID NO: 28-30 include both PR and NPB sequences (PR+NPB). Immunogenic PspA polypeptides of SEQ ID NOS: 31 and 32 include only a PR sequence (PR only) and lack the NPB.

[0075] Another appropriate art-recognized class of *S. pneumoniae* antigen is the pneumolysoids. Pneumolysoids have homology to the *S. pneumoniae* protein pneumolysin (PLY), but have reduced toxicity compared to pneumolysin. Pneumolysoids can be naturally occurring or engineered derivatives of pneumolysin. In some embodiments, a pneumolysoid has at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to pneumolysin. In some embodiments, the pneumolysoid demonstrates less than 1/2, 1/5, 1/10, 1/20, 1/50, 1/100, 1/200, 1/500, or 1/1000 the toxicity of pneumolysin in an assay for one or both of hemolytic activity towards erythrocytes and inhibition of polymorphonuclear leukocytes. Both assays are described in Saunders F.K. *et al.* ("Pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae*, does not require a thiol group for in vitro activity" *Infect Immun.* 1989 Aug;57(8):2547-52). Exemplary pneumolysoids include PdT (a triple mutant further described in Berry, A.M. *et al.* (1995) *Infection and Immunity* 63:1969-74); Pd-A and Pd-B (Paton J.C. *et al.* "Purification and immunogenicity of genetically obtained pneumolysin toxoids and their conjugation to *Streptococcus pneumoniae* type 19F polysaccharide" *Infect Immun.* 1991 Jul;59(7):2297-304); rPd2 and rPd3 (Ferreira *et al.* "DNA vaccines based on genetically detoxified derivatives of pneumolysin fail to protect mice against challenge with *Streptococcus pneumoniae*" *FEMS Immunol Med Microbiol* (2006) 46: 291-297); Ply8, Δ6PLY, L460D (see, e.g., US 2009/0285846 and L. Mitchell, *Protective Immune Responses*

to *Streptococcus pneumoniae* Pneumolysoids, ASM2011 conference abstract, 2011), or a variant thereof. In some embodiments, the pneumolysin has a mutation in the catalytic center, such as at amino acid 428 or 433 or the vicinity.

[0076] Other appropriate *S. pneumoniae* antigens include Choline-binding protein A (CbpA) and derivatives thereof (AD Ogunniyi *et al.*, “Protection against *Streptococcus pneumoniae* elicited by immunization with pneumolysin and CbpA,” *Infect Immun.* 2001 Oct;69(10):5997-6003); Pneumococcal surface adhesin A (PsaA); caseinolytic protease; sortase A (SrtA); pilus 1 RrgA adhesin; PpmA; PrtA; PavA; LytA; Stk-PR; PcsB; RrgB and derivatives thereof. CpbA derivatives include fusion proteins described in WO 2012/134975. Such fusion proteins may comprise one or more copies of the R2 domain, R2₁ and/or R2₂ subdomains of CpbA, or active variants and fragments thereof, or any combination thereof. Such fusion proteins may further comprise a pneumolysoid. The construct YLN, for example, comprises CpbA polypeptides YPT and NEEK, and pneumolysoid L460D.

[0077] In some cases, the antigen is at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identical to the corresponding wild-type *S. pneumoniae* protein. Sequences of the above-mentioned polypeptides, and nucleic acids that encode them, are known; see, for example, the *S. pneumoniae* ATCC 700669 complete genome sequence under GenBank accession number FM211187.1 and linked polypeptide sequences therein.

[0078] Typically, fusion proteins described herein include polypeptides, and fragments and variants thereof, that are immunogenic. These polypeptides may be immunogenic in mammals, for example mice, guinea pigs, or humans. An immunogenic polypeptide is typically one capable of raising a significant immune response in an assay or in a subject. The immune response may be innate, adaptive, humoral, cell-mediated, or mucosal (combining elements of innate, adaptive, humoral and cell-mediated immunity). For instance, an immunogenic polypeptide may increase the amount of IL-17 produced by T cells. Alternatively or additionally, an immunogenic polypeptide may (i) induce production of antibodies, e.g., neutralizing antibodies, that bind to the polypeptide and/or the whole bacteria, (ii) induce T_H17 immunity, (iii) activate the CD4⁺ T cell response, for example by increasing CD4⁺ T cells and/or increasing localization of CD4⁺ T cells to the site of infection or reinfection, (iv) activate the CD8⁺ CTL response, for example by increasing CD8⁺ T cells and/or increasing localization of CD8⁺ T cells to the site of infection or reinfection, (v) induce T_H1 immunity, and/or (vi) activate innate immunity. In some embodiments, an immunogenic polypeptide causes the production of a detectable amount of antibody specific to that antigen.

[0079] In certain embodiments, polypeptides have less than 20%, 30%, 40%, 50%, 60% or 70% identity to human autoantigens and/or gut commensal bacteria (e.g., certain *Bacteroides*, *Clostridium*, *Fusobacterium*, *Eubacterium*, *Ruminococcus*, *Peptococcus*, *Peptostreptococcus*, *Bifidobacterium*, *Escherichia* and *Lactobacillus* species). Examples of human autoantigens include insulin, proliferating cell nuclear antigen, cytochrome P450, and myelin basic protein.

[0080] A polypeptide may comprise one or more immunogenic portions and one or more non-immunogenic portions. The immunogenic portions may be identified by various methods, including protein microarrays, ELISPOT/ELISA techniques, and/or specific assays on different deletion mutants (e.g., fragments) of the polypeptide in question. Immunogenic portions may also be identified by computer algorithms. Some such algorithms, like EpiMatrix (produced by EpiVax), use a computational matrix approach. Other computational tools for identifying antigenic epitopes include PEPVAC (Promiscuous EPitope-based VACCine, hosted by Dana Farber Cancer Institute on the world wide web at immunax.dfci.harvard.edu/PEPVAC), MHCPreD (which uses a partial least squares approach and is hosted by The Jenner Institute on the world wide web at www.jenner.ac.uk/MHCPreD), and Immune Epitope Database algorithms on the world wide web at tools.immuneepitope.org. An immunogenic fragment of a polypeptide described herein comprises at least one immunogenic portion, as measured experimentally or identified by algorithm.

[0081] Thus, in some aspects, this application provides a fusion protein that includes one or more immunogenic fragments of an antigen described herein. The fragments, in some instances, are close in size to the full-length polypeptide or the polypeptide of Table 1, 2 or 3. For example, they may lack at most one, two, three, four, five, ten, twenty, or thirty amino acids from one or both termini. In certain embodiments, the polypeptide is 100-500 amino acids in length, or 150-450, or 200-400, or 250-250 amino acids in length. In some embodiments, the polypeptide is 100-200, 150-250, 200-300, 250-350, 300-400, 350-450, or 400-500 amino acids in length. In certain embodiments, the fragments result from processing, or partial processing, of signal sequences by an expression host, e.g. *E. coli*, an insect cell line (e.g., the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line. The fragments described above or sub-fragments thereof (e.g., fragments of 8-50, 8-30, or 8-20 amino acid residues) preferably have one of the biological activities described below, such as increasing the amount of IL-17 released by at least 1.5 fold or 2 fold or more (e.g., either as an absolute measure or relative to an

immunologically inactive protein). A fragment may be used as the polypeptide in the vaccines described herein or may be fused to another protein, protein fragment or a polypeptide.

[0082] In some embodiments, the fragment is a truncated fragment of any of SEQ ID NOS: 1-36 or 47-66, having from 1-5, 1-10, or 1-20 amino acid residues removed from the N-terminus, C-terminus, or both. In some such embodiments, the same number of residues is removed from the N-terminus and the C-terminus, while in other embodiments, a different number of residues is removed from the N-terminus compared to the C-terminus.

[0083] In certain aspects, this application provides fusion proteins that include one or more immunogenic polypeptides with at least 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a polypeptide of Table 1, 2 or 3. In certain embodiments, the fusion protein comprises at least two different polypeptides having an amino acid sequence comprising a sequence at least 90%, 95%, 98%, or 99% identical to any of SEQ ID NOS: 1-36 or 47-66, or an immunogenic fragment thereof.

[0084] In certain embodiments, this application provides nucleic acids encoding one or more of the polypeptides described above, such as cDNA, DNA, RNA, or an analog thereof. The underlying DNA sequences for the polypeptides described above may be modified in ways that do not affect the sequence of the protein product, and such sequences are included in the invention. For instance, the DNA sequence may be codon-optimized to improve expression in a host such as *E. coli*, an insect cell line (e.g., using the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line.

[0085] In certain embodiments, this application provides nucleic acids (such as cDNA, DNA, RNA, or an analog thereof) that are at least 70%, 80%, 90%, 95%, 97%, 98%, 99%, or 100% identical to a gene in Table 1, 2 or 3, or a variant or portion of said gene. In certain embodiments, the nucleic acid is 600-2000, 800-1800, 1000-1600, 1200-1400 nucleotides in length. In some embodiments, the nucleic acid is 600-1600, 800-1800, or 1000-2000 nucleotides in length. The nucleic acids may be used, for example, for recombinant production of the polypeptides of Tables 1, 2 and 3, or immunogenic fragments or fusion proteins thereof.

C. Immunogenic compositions

[0086] The present disclosure also provides pharmaceutical compositions containing immunogenic polypeptides or polynucleotides encoding these immunogenic polypeptides, including fusion proteins, together with a pharmaceutical carrier. The *S. pneumoniae*

antigens used were identified by screening immune cells from mice either immunized with unencapsulated killed whole *S. pneumoniae* or infected with *S. pneumoniae*, or from healthy human donors, that elicit a T cell specific immune response. The human donors had almost certainly been exposed to *S. pneumoniae* at some point during their lifetimes, because *S. pneumoniae* is a very common disease and colonizing pathogen. In some embodiments, these antigens are fused to known *S. pneumoniae* capsular antigens, known to stimulate a B-cell specific immunity. Thus, the present disclosure contemplates compositions of (e.g., fusion proteins including) *S. pneumoniae* antigens that elicit an immune response in all age groups that is strong and long-lived in immunized or infected mice or humans, for counteracting infection by *S. pneumoniae*.

[0087] In certain embodiments, a vaccine composition described herein induces a T_H17 cell response at least 1.5-fold greater than that induced by a control unrelated antigen (such as the HSV-2 protein ICP47 with the gene name US12) after contacting T_H17 cells. In some embodiments, a vaccine formulation described herein inhibits infection by *S. pneumoniae* in an uninfected subject. In certain embodiments, a vaccine formulation reduces occurrence or duration of *S. pneumoniae* nasopharyngeal colonization in an individual infected by *S. pneumoniae*. In some embodiments, a vaccine formulation inhibits development of sepsis in an individual infected by *S. pneumoniae*. In some embodiments, a vaccine formulation inhibits development of invasive diseases such as pneumonia, meningitis, otitis media, sinusitis or infection of other sites or organs with *S. pneumoniae*.

[0088] An immunogenic composition may also comprise portions of said Streptococcus polypeptides, including fusion proteins, for example deletion mutants, truncation mutants, oligonucleotides, and peptide fragments. In some embodiments, the portions of said polypeptides are immunogenic. The immunogenicity of a portion of a protein is readily determined using the same assays that are used to determine the immunogenicity of the full-length protein. In some embodiments, the portion of the polypeptide has substantially the same immunogenicity as the full-length proteins. In some embodiments, the immunogenicity is no more than 10%, 20%, 30%, 40%, or 50% less than that of the full-length protein (e.g., polypeptides of Tables 1, 2 and 3). The peptide fragments may be, for example, linear, circular, or branched.

[0089] In some embodiments, a vaccine formulation contains one or more fusion proteins described herein in combination with one or more, e.g., two, three, four, or more polypeptides from Table 1 or 2 or immunogenic fragments or variants thereof in a mixture. In some embodiments, the mixture contains both full-length polypeptides and fragments

resulting from processing, or partial processing, of signal sequences by an expression host, e.g. *E. coli*, an insect cell line (e.g., the baculovirus expression system), or a mammalian (e.g., human or Chinese Hamster Ovary) cell line. In some embodiments, a vaccine formulation contains one or more fusion proteins of any of SEQ ID NOS: 55-66 in the absence of any other antigens. In some embodiments, a vaccine formulation contains one or more fusion proteins of any of SEQ ID NOS: 55-66 in combination with one or more additional proteins of any of SEQ ID NOS: 1-26 and 47-54 for eliciting a T cell response, in the absence of other antigens.

[0090] In some embodiments, polypeptides described in Tables 1 and 2 may be used without modification in conjunction with the fusion proteins described herein. In certain embodiments, when smaller related polypeptides are used, such as fragments or the like, and their molecular weight is less than about 5000 daltons, e.g., 1500 to 5000 daltons, modification may be useful in eliciting the desired immune response. For example, the smaller polypeptides can be conjugated to an appropriate immunogenic carrier such as tetanus toxoid, keyhole limpet hemocyanin or the like. In some embodiments, polypeptides described herein may be conjugated to *S. pneumoniae* polysaccharides. The conjugated polysaccharides may be, for example, as described in US Patent 5,623,057, US Patent 5,371,197, or PCT/US2011/023526.

[0091] Some embodiments of the vaccine formulations and immunogenic compositions described herein include an immunogenic polypeptide, including a fusion protein (e.g., a polypeptide of Table 1, 2 or 3) that contains a membrane translocating sequence (MTS), to facilitate introduction of the polypeptide into the mammalian cell and subsequent stimulation of the cell-mediated immune response. Exemplary membrane translocating sequences include hydrophobic region in the signal sequence of Kaposi fibroblast growth factor, the MTS of α -synuclein, β -synuclein, or γ -synuclein, the third helix of the Antennapedia homeodomain, SN50, integrin β 3 h-region, HIV Tat, pAntp, PR-39, abaecin, apidaecin, Bac5, Bac7, *P. berghei* CS protein, and those MTSs described in US Patents 6,248,558, 6,432,680 and 6,248,558.

[0092] In certain embodiments, an antigen (e.g., a polypeptide of Table 1, 2 or 3) is covalently bound to another molecule. This may, for example, increase the half-life, solubility, bioavailability, affinity, or immunogenicity of the antigen, or raise an immune reaction against more than one pathogen. Molecules that may be covalently bound to the antigen include a carbohydrate, biotin, avidin or avidin derivatives, poly(ethylene glycol) (PEG), polysialic acid, N-propionylated polysialic acid, nucleic acids, polysaccharides, and

PLGA. There are many different types of PEG, ranging from molecular weights of below 300 g/mol to over 10,000,000 g/mol. PEG chains can be linear, branched, or with comb or star geometries. In some embodiments, a polypeptide described herein is covalently bound to a moiety that stimulates the immune system. An example of such a moiety is a lipid moiety. In some instances, lipid moieties are recognized by a Toll-like receptor (TLR) such as TLR-2 or TLR-4, and activate the innate immune system.

D. Components of a vaccine or immunogenic composition comprising *S. pneumoniae* antigens or antibodies recognizing the same

[0093] In certain embodiments, a vaccine or immunogenic composition described herein comprises one or more fusion proteins described herein and one or more of the following: an adjuvant, vaccine scaffold, stabilizer, buffer, surfactant, controlled release component, salt, preservative, and/or an antibody specific to a polypeptide included in such composition.

1. Adjuvants

[0094] The vaccine formulations and immunogenic compositions described herein may include an adjuvant. Adjuvants can be broadly separated into two classes, based on their principal mechanisms of action: vaccine delivery systems and immunostimulatory adjuvants (see, *e.g.*, Singh *et al.*, *Curr. HIV Res.* 1:309-20, 2003). In most vaccine formulations, the adjuvant provides a signal to the immune system so that it generates a response to the antigen, and the antigen is required for driving the specificity of the response to the pathogen. Vaccine delivery systems are often particulate formulations, *e.g.*, emulsions, microparticles, immune-stimulating complexes (ISCOMs), nanoparticles, which may be, for example, particles and/or matrices, and liposomes. In contrast, immunostimulatory adjuvants are sometimes derived from pathogens and can represent pathogen associated molecular patterns (PAMP), *e.g.*, lipopolysaccharides (LPS), monophosphoryl lipid (MPL), or CpG-containing DNA, which activate cells of the innate immune system.

[0095] Alternatively, adjuvants may be classified as organic and inorganic. Inorganic adjuvants include alum salts such as aluminum phosphate, amorphous aluminum hydroxyphosphate sulfate, and aluminum hydroxide, which are commonly used in human vaccines. Organic adjuvants comprise organic molecules including macromolecules. An example of an organic adjuvant is cholera toxin.

[0096] Adjuvants may also be classified by the response they induce. In some embodiments, the adjuvant induces the activation of T_H1 cells or T_H2 cells. In other embodiments, the adjuvant induces the activation of B cells. In yet other embodiments, the adjuvant induces the activation of antigen-presenting cells. These categories are not mutually exclusive; in some cases, an adjuvant activates more than one type of cell.

[0097] In certain embodiments, the adjuvant induces the activation of T_H17 cells. It may promote the CD4⁺ or CD8⁺ T cells to secrete IL-17. In some embodiments, an adjuvant that induces the activation of T_H17 cells is one that produces at least a 2-fold, and in some cases a 10-fold, experimental sample to control ratio in the following assay. In the assay, an experimenter compares the IL-17 levels secreted by two populations of cells: (1) cells from animals immunized with the adjuvant and a polypeptide known to induce T_H17 activation, and (2) cells from animals treated with the adjuvant and an irrelevant (control) polypeptide. An adjuvant that induces the activation of T_H17 cells may cause the cells of population (1) to produce more than 2-fold, or more than 10-fold more IL-17 than the cells of population (2). IL-17 may be measured, for example, by ELISA or ELISPOT. Certain toxins, such as cholera toxin and labile toxin (produced by enterotoxigenic *E. coli*, or ETEC), activate a T_H17 response. Thus, in some embodiments, the adjuvant is a toxin. Cholera toxin was successfully used in the mouse model to induce protective immunity in conjunction with certain polypeptides from Table 1 (see Examples 1-2). One form of labile toxin is produced by Intercell. Mutant derivatives of labile toxin that are active as adjuvants but significantly less toxic can be used as well. Exemplary detoxified mutant derivatives of labile toxin include mutants lacking ADP-ribosyltransferase activity. Particular detoxified mutant derivatives of labile toxin include LTK7 (Douce *et al.*, “Mutants of *Escherichia coli* heat-labile toxin lacking ADP-ribosyltransferase activity act as nontoxic, mucosal adjuvants” PNAS Vol. 92, pp. 1644-1648, February 1995) and LTK63 (Williams *et al.*, “Innate Imprinting by the Modified Heat-Labile Toxin of *Escherichia coli* (LTK63) Provides Generic Protection against Lung Infectious Disease” The Journal of Immunology, 2004, 173: 7435-7443), LT-G192 (Douce *et al.* “Genetically detoxified mutants of heat-labile toxin from *Escherichia coli* are able to act as oral adjuvants” Infect Immun. 1999 Sep;67(9):4400-6), and LTR72 (“Mucosal adjuvant activity and immunogenicity of LTR72, a novel mutant of *Escherichia coli* heat-labile enterotoxin with partial knockout of ADP-ribosyltransferase activity.” J Exp Med. 1998 Apr 6;187(7):1123-32).

[0098] In some embodiments, the adjuvant comprises a VLP (virus-like particle). One such adjuvant platform, Alphavirus replicons, induces the activation of T_H17 cells using

alphavirus and is produced by Alphavax. In certain embodiments of the Alphavirus replicon system, alphavirus may be engineered to express an antigen of interest, a cytokine of interest (for example, IL-17 or a cytokine that stimulates IL-17 production), or both, and may be produced in a helper cell line. More detailed information may be found in U.S. Patent Nos. 5,643,576 and 6,783,939. In some embodiments, a vaccine formulation is administered to a patient in combination with a nucleic acid encoding a cytokine.

[0099] Certain classes of adjuvants activate toll-like receptors (TLRs) in order to activate a T_H17 response. TLRs are well known proteins that may be found on leukocyte membranes, and recognize foreign antigens (including microbial antigens). Administering a known TLR ligand together with an antigen of interest (for instance, as a fusion protein) can promote the development of an immune response specific to the antigen of interest. One exemplary adjuvant that activates TLRs comprises Monophosphoryl Lipid A (MPL). Traditionally, MPL has been produced as a detoxified lipopolysaccharide (LPS) endotoxin obtained from Gram negative bacteria, such as *S. minnesota*. In particular, sequential acid and base hydrolysis of LPS produces an immunoactive lipid A fraction (which is MPL), and lacks the saccharide groups and all but one of the phosphates present in LPS. A number of synthetic TLR agonists (in particular, TLR-4 agonists) are disclosed in Evans JT *et al.* "Enhancement of antigen-specific immunity via the TLR-4 ligands MPL adjuvant and Ribi.529." Expert Rev Vaccines 2003 Apr;2(2):219-29. Like MPL adjuvants, these synthetic compounds activate the innate immune system via TLR. Another type of TLR agonist is a synthetic phospholipid dimer, for example E6020 (Ishizaka ST *et al.* "E6020: a synthetic Toll-like receptor 4 agonist as a vaccine adjuvant" Expert Rev. Vaccines 2007 Oct; 6(5):773-84). Various TLR agonists (including TLR-4 agonists) have been produced and/or sold by, for example, the Infectious Disease Research Institute (IRDI), Corixa, Esai, Avanti Polar Lipids, Inc., and Sigma Aldrich. Another exemplary adjuvant that activates TLRs comprises a mixture of MPL, Trehalose Dicoynomylate (TDM), and dioctadecyldimethylammonium bromide (DDA). Another TLR-activating adjuvant is R848 (resiquimod).

[0100] In some embodiments, the adjuvant is or comprises a saponin. Typically, the saponin is a triterpene glycoside, such as those isolated from the bark of the *Quillaja saponaria* tree. A saponin extract from a biological source can be further fractionated (e.g., by chromatography) to isolate the portions of the extract with the best adjuvant activity and with acceptable toxicity. Typical fractions of extract from *Quillaja saponaria* tree used as adjuvants are known as fractions A and C.

[0101] A particular form of saponins that may be used in vaccine formulations described herein is immunostimulating complexes (ISCOMs). ISCOMs are an art-recognized class of adjuvants, that generally comprise *Quillaja* saponin fractions and lipids (e.g., cholesterol and phospholipids such as phosphatidyl choline). In certain embodiments, an ISCOM is assembled together with a polypeptide or nucleic acid of interest. However, different saponin fractions may be used in different ratios. In addition, the different saponin fractions may either exist together in the same particles or have substantially only one fraction per particle (such that the indicated ratio of fractions A and C are generated by mixing together particles with the different fractions). In this context, "substantially" refers to less than 20%, 15%, 10%, 5%, 4%, 3%, 2% or even 1%. Such adjuvants may comprise fraction A and fraction C mixed into a ratio of 70-95 A: 30-5 C, such as 70 A : 30 C to 75 A : 5 C, 75 A : 5 C to 80 A : 20 C, 80 A : 20 C to 85 A : 15 C, 85 A : 15 C to 90 A : 10 C, 90 A : 10 C to 95 A : 5 C, or 95 A : 5 C to 99 A : 1 C.

[0102] In certain embodiments, combinations of adjuvants are used. Three exemplary combinations of adjuvants are MPL and alum, E6020 and alum, and MPL and an ISCOM.

[0103] Adjuvants may be covalently bound to antigens (e.g., to one or more polypeptides included in a fusion protein described herein). In some embodiments, the adjuvant may comprise a protein which induces inflammatory responses through activation of antigen-presenting cells (APCs). In some embodiments, one or more of these proteins can be recombinantly fused with an antigen of choice, such that the resultant fusion molecule promotes dendritic cell maturation, activates dendritic cells to produce cytokines and chemokines, and ultimately, enhances presentation of the antigen to T cells and initiation of T cell responses (see Wu et al., Cancer Res 2005; 65(11), pp 4947-4954). In certain embodiments, a polypeptide, including a fusion protein, described herein is presented in the context of the trivalent conjugate system, comprising a fusion protein of *S. pneumoniae* Pneumococcal surface adhesin A (PsaA) with the pneumolysoid PdT and a cell wall polysaccharide (PsaA: PdT-CPs), described in Lu *et al.* ("Protection against Pneumococcal colonization and fatal pneumonia by a trivalent conjugate of a fusion protein with the cell wall polysaccharide." Infect Immun. 2009 May;77(5):2076-83). PdT carries three amino acid substitutions (W433F, D385N, and C428G) which render the molecule nontoxic but do not interfere with its TLR-4-mediated inflammatory properties. Conjugation of a polysaccharide to the fusion of a polypeptide to the TLR-4-agonist PdT enhances immunological response to the polypeptide. In some embodiments, one or more polypeptides

described herein are used in place of PsaA in the trivalent conjugate. The trivalent conjugate system typically includes alum and is usually administered parenterally. Other exemplary adjuvants that may be covalently bound to antigens comprise polysaccharides, pneumolysin, synthetic peptides, lipopeptides, and nucleic acids.

[0104] Typically, the same adjuvant or mixture of adjuvants is present in each dose of a vaccine. Optionally, however, an adjuvant may be administered with the first dose of vaccine and not with subsequent doses (i.e., booster shots). Alternatively, a strong adjuvant may be administered with the first dose of vaccine and a weaker adjuvant or lower dose of the strong adjuvant may be administered with subsequent doses. The adjuvant can be administered before the administration of the antigen, concurrent with the administration of the antigen or after the administration of the antigen to a subject (sometimes within 1, 2, 6, or 12 hours, and sometimes within 1, 2, or 5 days). Certain adjuvants are appropriate for human patients, non-human animals, or both.

2. Additional components of a vaccine or immunogenic composition

[0105] In addition to one or more fusion proteins and adjuvants described above, a vaccine formulation or immunogenic composition may include one or more additional components.

[0106] In certain embodiments, the vaccine formulation or immunogenic composition may include one or more stabilizers such as sugars (such as sucrose, glucose, or fructose), phosphate (such as sodium phosphate dibasic, potassium phosphate monobasic, dibasic potassium phosphate, or monosodium phosphate), glutamate (such as monosodium L-glutamate), gelatin (such as processed gelatin, hydrolyzed gelatin, or porcine gelatin), amino acids (such as arginine, asparagine, histidine, L-histidine, alanine, valine, leucine, isoleucine, serine, threonine, lysine, phenylalanine, tyrosine, and the alkyl esters thereof), inosine, or sodium borate.

[0107] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more buffers such as a mixture of sodium bicarbonate and ascorbic acid. In some embodiments, the vaccine formulation may be administered in saline, such as phosphate buffered saline (PBS), or distilled water.

[0108] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more surfactants such as polysorbate 80 (Tween 80), Triton X-100, Polyethylene glycol tert-octylphenyl ether t-Octylphenoxy polyethoxy ethanol 4-(1,1,3,3-Tetramethylbutyl)phenyl-polyethylene glycol (TRITON X-100); Polyoxyethylenesorbitan

monolaurate Polyethylene glycol sorbitan monolaurate (TWEEN 20); and 4-(1,1,3,3-Tetramethylbutyl)phenol polymer with formaldehyde and oxirane (TYLOXAPOL). A surfactant can be ionic or nonionic.

[0109] In certain embodiments, the vaccine formulation or immunogenic composition includes one or more salts such as sodium chloride, ammonium chloride, calcium chloride, or potassium chloride.

[0110] In certain embodiments, a preservative is included in the vaccine or immunogenic composition. In other embodiments, no preservative is used. A preservative is most often used in multi-dose vaccine vials, and is less often needed in single-dose vaccine vials. In certain embodiments, the preservative is 2-phenoxyethanol, methyl and propyl parabens, benzyl alcohol, and/or sorbic acid.

[0111] In certain embodiments, the vaccine formulation or immunogenic composition is a controlled release formulation.

E. Use of Vaccines

[0112] The *S. pneumoniae* vaccines described herein may be used for prophylactic and/or therapeutic treatment of *S. pneumoniae*. In some embodiments, vaccine compositions may be used to treat, alleviate, ameliorate, relieve, delay onset of, inhibit progression of, reduce risk of infection by, and reduce severity of, and/or reduce incidence of one or more symptoms or features of a pneumococcal disease, disorder, and/or condition. In some embodiments, vaccine compositions may be used to treat, alleviate, ameliorate, relieve, delay onset of, inhibit progression of, reduce severity of, and/or reduce incidence of one or more symptoms or features of pneumococcal infection (e.g., but not limited to sepsis, pneumonia, acute otitis media, bacteremia, and bacterial meningitis). In some embodiments, vaccine compositions may be used to reduce the incidence of pneumococcal infection in subjects administered the composition. In some embodiments, vaccine compositions may be used to reduce the incidence of nasopharyngeal colonization by a pneumococcal organism. In some embodiments, vaccine compositions may be used to reduce the likelihood of chronic infection by a pneumococcal organism. In some embodiments, vaccine compositions may be used to reduce the likelihood of suffering from pelvic inflammatory disease or infertility subsequent to a pneumococcal infection.

[0113] In some embodiments, prophylactic and/or therapeutic protocols involve administering a therapeutically effective amount of one or more vaccine compositions described herein to a healthy subject (i.e., a subject who does not display any

symptoms of, and/or who has not been diagnosed with, pneumococcal infection). For example, healthy individuals may be vaccinated using inventive vaccine compositions prior to pneumococcal colonization, development of pneumococcal infection and/or onset of symptoms of pneumococcal infection; at risk individuals (e.g., patients exposed to individuals suffering from pneumococcal infection, patients with HIV or a compromised immune system, or a subject who is susceptible to pneumococcal infection such as a subject with Job's syndrome (subject lacking T_H17 cell-mediated response) or a subject with agammaglobulinemia (a subject lacking antibody-mediated response) can be treated substantially contemporaneously with (e.g., within 48 hours, within 24 hours, or within 12 hours of) the onset of symptoms of and/or exposure to pneumococcal infection. Of course individuals known to have pneumococcus infection or sepsis may receive treatment at any time.

[0114] In some embodiments, prophylactic and/or therapeutic protocols involve administering a therapeutically effective amount of one or more inventive vaccine compositions to a subject such that an immune response is stimulated in both T cells and B cells.

[0115] In some embodiments, by combining one or more antigens and adjuvants, immune responses (e.g. T cell and/or B cell responses) can be tailored to preferentially elicit the most desirable type of immune response for a given indication, e.g., humoral response, mucosal response, T_H1 -biased T cell response, T_H17 -biased T cell response, IFN- γ secretion by antigen-specific T cells, IL-17A secretion by antigen-specific T cells, cytotoxic T cell response, antibody response, B cell response, adaptive immune response, innate immune response, or a combination of these responses.

[0116] Accordingly, this application provides a method for treating a subject suffering from or susceptible to *S. pneumoniae* infection, comprising administering an effective amount of any of the vaccine formulations described herein. In some aspects, the method inhibits *S. pneumoniae* colonization in an individual. In some aspects, the method inhibits *S. pneumoniae* symptoms, invasive disease or sequelae, such as sepsis, pneumonia, meningitis, otitis media, sinusitis or infection of other sites or organs with *S. pneumoniae*. The subject receiving the vaccination may be a male or a female, and may be a child or adult. In some embodiments, the subject being treated is a human. In other embodiments, the subject is a non-human animal.

1. Prophylactic use

[0117] In prophylactic embodiments, the vaccine is administered to a subject to induce an immune response that can help protect against the establishment of *S. pneumoniae*, for example by protecting against colonization, the first and necessary step in disease. Thus, in some aspects, the method inhibits infection by *S. pneumoniae* in a non-colonized or uninfected subject. In another aspect, the method may reduce the duration of colonization in an individual who is already colonized.

[0118] In some embodiments, the vaccine compositions described herein confer protective immunity, allowing a vaccinated individual to exhibit delayed onset of symptoms or sequelae, or reduced severity of symptoms or sequelae, as the result of his or her exposure to the vaccine. In certain embodiments, the reduction in severity of symptoms or sequelae is at least 25%, 40%, 50%, 60%, 70%, 80% or even 90%. In particular embodiments, vaccinated individuals may display no symptoms or sequelae upon contact with *S. pneumoniae*, do not become colonized by *S. pneumoniae*, or both. Protective immunity is typically achieved by one or more of the following mechanisms: mucosal, humoral, or cellular immunity. Mucosal immunity is primarily the result of secretory IgA (sIgA) antibodies on mucosal surfaces of the respiratory, gastrointestinal, and genitourinary tracts. The sIgA antibodies are generated after a series of events, mediated by antigen-processing cells, B and T lymphocytes, that results in sIgA production by B lymphocytes on mucosa-lined tissues of the body. Humoral immunity is typically the result of IgG antibodies and IgM antibodies in serum. Cellular immunity can be achieved through cytotoxic T lymphocytes or through delayed-type hypersensitivity that involves macrophages and T lymphocytes, as well as other mechanisms involving T cells without a requirement for antibodies. In particular, cellular immunity may be mediated by T_H1 or T_H17 cells.

[0119] Essentially any individual has a certain risk of becoming infected with *S. pneumoniae*. However, certain sub-populations have an increased risk of infection. In some embodiments, a vaccine formulation as described herein is administered to patients that are immunocompromised.

[0120] An immunocompromising condition, e.g., arising from a medical treatment, can expose the individual in question to a higher risk of infection with *S. pneumoniae*. It is possible to treat an infection prophylactically in an individual having an immunocompromised condition before or during treatments known to compromise immune function. By prophylactically treating with an antigenic composition (e.g., a fusion protein described herein), or with antibodies reactive to two or more antigens from Table 1 or 2, or to

an antigen from Table 3, before or during a treatment known to compromise immune function, it is possible to prevent a subsequent *S. pneumoniae* infection or to reduce the risk of the individual contracting an infection due to the immunocompromised condition. Should the individual contract an *S. pneumoniae* infection e.g., following a treatment leading to an immunocompromised condition it is also possible to treat the infection by administering to the individual an antigen composition.

[0121] The following groups are at increased risk of pneumococcal disease or its complications, and therefore it is advantageous for subjects falling into one or more of these groups to receive a vaccine formulation described herein: children, especially those from 1 month to 5 years old or 2 months to 2 years old; children who are at least 2 years of age with asplenia, splenic dysfunction or sickle-cell disease; children who are at least 2 years of age with nephrotic syndrome, chronic cerebrospinal fluid leak, HIV infection or other conditions associated with immunosuppression.

[0122] In another embodiment, at least one dose of a vaccine composition described herein is given to adults in the following groups at increased risk of pneumococcal disease or its complications: all persons 65 years of age; adults with asplenia, splenic dysfunction or sickle-cell disease; adults with the following conditions: chronic cardiorespiratory disease, cirrhosis, alcoholism, chronic renal disease, nephrotic syndrome, diabetes mellitus, chronic cerebrospinal fluid leak, HIV infection, AIDS and other conditions associated with immunosuppression (Hodgkin's disease, lymphoma, multiple myeloma, immunosuppression for organ transplantation), individuals with cochlear implants; individuals with long-term health problems such as heart disease and lung disease, as well as individuals who are taking any drug or treatment that lowers the body's resistance to infection, such as long-term steroids, certain cancer drugs, radiation therapy; Alaskan natives and certain Native American populations.

2. Therapeutic use

[0123] In therapeutic applications, a vaccine formulation described herein may be administered to a patient suffering from *S. pneumoniae* infection, in an amount sufficient to treat the patient. Treating the patient, in this case, refers to reducing *S. pneumoniae* symptoms and/or bacterial load and/or sequelae in an infected individual. In some embodiments, treating the patient refers to reducing the duration of symptoms or sequelae, or reducing the intensity of symptoms or sequelae. In some embodiments, treating the patient refers to eliminating pneumococcal organisms in the nasopharynx or reducing the duration of nasopharyngeal colonization. In some embodiments, the vaccine reduces transmissibility of

S. pneumoniae from the vaccinated patient. In certain embodiments, the reductions described above are at least 25%, 30%, 40%, 50%, 60%, 70%, 80% or even 90%.

[0124] In therapeutic embodiments, a vaccine is administered to an individual post-infection. The vaccine may be administered shortly after infection, e.g. before symptoms or sequelae manifest, or may be administered during or after manifestation of symptoms or sequelae.

[0125] A vaccine formulation described herein can reduce the intensity and/or duration of the various symptoms or sequelae of *S. pneumoniae* infection. Symptoms or sequelae of *S. pneumoniae* infection can take many forms. In some cases, an infected patient develops pneumonia, acute sinusitis, otitis media (ear infection), meningitis, bacteremia, sepsis, osteomyelitis, septic arthritis, endocarditis, peritonitis, pericarditis, cellulitis, or brain abscess.

[0126] Sepsis is a rare but life-threatening complication of *S. pneumoniae* infection, where the bacterium invades the bloodstream and systemic inflammation results. Typically, fever is observed and white blood cell count increases. A further description of sepsis is found in Goldstein, B. *et al.* "International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics." *Pediatr Crit Care Med.* Jan 2005;6 (1):2-8.

3. Assaying vaccination efficacy

[0127] The efficacy of vaccination with vaccines disclosed herein may be determined in a number of ways, in addition to the clinical outcomes described above. First, one may assay IL-17 levels (particularly IL-17A) by stimulating T cells derived from the subject after vaccination with a vaccine composition described herein. The IL-17 levels may be compared to IL-17 levels in the same subject before vaccination. Increased IL-17 (e.g., IL-17A) levels, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, would indicate an increased response to the vaccine. Alternatively (or in combination), one may assay neutrophils in the presence of T cells or antibodies from the patient for pneumococcal killing. Increased pneumococcal killing, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, would indicate an increased response to the vaccine. In addition, one may measure T_H17 cell activation, where increased T_H17 cell activation, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, correlates with an increased response to the vaccine. One may also measure levels of an antibody specific to the vaccine, where increased levels of the specific antibody, such as a 1.5 fold, 2-fold, 5-fold, 10-fold, 20-fold, 50-fold or 100-fold or more increase, are correlated with increased vaccine efficacy. In certain embodiments, two or more of these

assays are used. For example, one may measure IL-17 levels and the levels of vaccine-specific antibody. In certain embodiments, one may assay for additional cytokines and transcription factors in the IL-17 lineage, e.g., cytokines IL-22, IL-23, IL-6 and TGF- β , and transcription factors STAT3, ROR γ t and ROR α . Alternatively, one may follow epidemiological markers such as incidence of, severity of, or duration of pneumococcal infection in vaccinated individuals compared to unvaccinated individuals.

[0128] Vaccine efficacy may also be assayed in various model systems such as the mouse model. For instance, BALB/c or C57BL/6 strains of mice may be used. After administering a vaccine formulation described herein to a subject (as a single dose or multiple doses), the experimenter administers a challenge dose of *S. pneumoniae*. In some cases, a challenge dose administered intranasally is sufficient to cause *S. pneumoniae* colonization (especially nasal colonization) in an unvaccinated animal, and in some cases a challenge dose administered via aspiration is sufficient to cause sepsis and a high rate of lethality in unvaccinated animals. One can then measure the reduction in colonization or the reduction in lethality in vaccinated animals. Examples 1-2 and 4 show the efficacy of polypeptides of Table 1 in inhibiting *S. pneumoniae* nasal colonization following intranasal challenge in the mouse model. Examples 5 and 6 show the efficacy of polypeptides of Table 1 in protecting against sepsis and death following infection with *S. pneumoniae* via aspiration in the mouse model.

F. Use of Immunogenic Compositions

1. Defense against S. pneumoniae infection

[0129] The immunogenic compositions (e.g., vaccine formulations) of the present disclosure are designed to elicit an immune response against *S. pneumoniae*. Compositions described herein (e.g., ones comprising one or more fusion proteins described herein) may stimulate an antibody response or a cell-mediated immune response, or both, in the mammal to which it is administered. In some embodiments, the composition stimulates a T_H1-biased CD4⁺ T cell response, a T_H17-biased CD4⁺ T cell response and/or a CD8⁺ T cell response. In some embodiments, the composition elicits IFN- γ secretion by antigen-specific T cells. In some embodiments, the composition elicits IL-17A secretion by antigen-specific T cells. In some embodiments, the composition stimulates an antibody response (e.g., an IgG, IgM and/or IgA response). In some embodiments, the composition stimulates a B cell response. In some embodiments, the composition stimulates a cytotoxic T cell (CTL) response. In

some embodiments, the composition stimulates an adaptive immune response. In some embodiments, the composition stimulates an innate immune response. In some embodiments, the composition stimulates both a T cell response and a B cell response. In some embodiments, the composition stimulates a T_H1 -biased $CD4^+$ T cell response, T_H17 -biased $CD4^+$ T cell response and/or a $CD8^+$ T cell response, and an antibody and/or B cell response.

[0130] In certain embodiments, the composition (e.g., one comprising one or more fusion proteins described herein) includes a cytokine or nucleotide coding region encoding a cytokine such as IL-17, to provide additional stimulation to the immune system of the mammal. In certain embodiments, the composition comprises a cytokine such as IL-17.

[0131] While not wishing to be bound by theory, in some embodiments a T_H17 cell response is desirable in mounting an immune response to the compositions disclosed herein. In certain embodiments, an active T_H17 response is beneficial in clearing a pneumococcal infection. For instance, mice lacking the IL-17A receptor show decreased whole cell vaccine-based protection from a pneumococcal challenge (Lu *et al.*, "Interleukin-17A mediates acquired immunity to pneumococcal colonization." PLoS Pathog. 2008 Sep 19;4(9)).

[0132] Thus, in some embodiments, the present disclosure includes methods of increasing IL-17 production by administering one or more compositions described herein (e.g., ones comprising one or more fusion proteins described herein) to a subject. Furthermore, the present disclosure includes methods of activating T_H17 cells by administering said compositions to a subject. In certain embodiments, increased IL-17A levels result in increased pneumococcal killing by neutrophils or neutrophil-like cells, for instance by inducing recruitment and activation of neutrophils or neutrophil-like cells. In certain embodiments, this pneumococcal killing is independent of antibodies and complement. However, specific antibody production and complement activation may be useful additional mechanisms that contribute to clearing of a pneumococcal infection.

[0133] Immunogenic compositions containing one or more fusion proteins described herein (or nucleic acids encoding a fusion protein described herein) together with a pharmaceutical carrier are also provided.

[0134] In some instances, the immunogenic composition comprises one or more nucleic acids (e.g., cDNA molecules) encoding one or more polypeptides of SEQ ID NOS: 1-66, such as one or more polypeptides selected from SEQ ID NOS 55-66. In some embodiments these nucleic acids are expressed in the immunized individual, producing the

encoded *S. pneumoniae* antigens, and the *S. pneumoniae* antigens so produced can produce an immunostimulatory effect in the immunized individual.

[0135] Such a nucleic acid-containing immunostimulatory composition may comprise, for example, an origin of replication, and a promoter that drives expression of one or more nucleic acids encoding one or more polypeptides of SEQ ID NOS: 55-66. Such a composition may also comprise a bacterial plasmid vector into which is inserted a promoter (sometimes a strong viral promoter), one or more nucleic acids encoding one or more polypeptides of SEQ ID NOS: 55-66, and a polyadenylation/transcriptional termination sequence. In some instances, the nucleic acid is DNA.

G. Doses and Routes of Administration

1. Dosage forms, amounts, and timing

[0136] The amount of a fusion protein described herein in each vaccine or immunogenic composition dose is selected as an effective amount, which induces a prophylactic or therapeutic response, as described above, in either a single dose or over multiple doses. Preferably, the dose is without significant adverse side effects in typical vaccinees. Such amount will vary depending upon which specific antigen is employed. Generally, it is expected that a dose will comprise 1-1000 µg of each protein, in some instances 2-100 µg, for instance 4-40 µg. In some aspects, the vaccine formulation comprises 1-1000 µg of the polypeptide and 1-350 µg of the adjuvant. In some embodiments, the appropriate amount of antigen to be delivered will depend on the age, weight, and health (e.g. immune-compromised status) of a subject. When present, typically an adjuvant will be present in amounts from 1 µg – 350 µg per dose, for example 50-150 µg, 75-125 µg or 100 µg.

[0137] In some embodiments, only one dose of the vaccine is administered to achieve the results described above. In other embodiments, following an initial vaccination, subjects receive one or more boost vaccinations, for a total of two, three, four or five vaccinations. Advantageously, the number is three or fewer. A boost vaccination may be administered, for example, about 1 month, 2 months, 4 months, 6 months, or 12 months after the initial vaccination, such that one vaccination regimen involves administration at 0, 0.5-2 and 4-8 months. It may be advantageous to administer split doses of vaccines which may be administered by the same or different routes.

[0138] The vaccines and immunogenic compositions described herein may take on a variety of dosage forms. In certain embodiments, the composition is provided in solid or powdered (e.g., lyophilized) form; it also may be provided in solution form. In certain embodiments, a dosage form is provided as a dose of lyophilized composition and at least one separate sterile container of diluent.

[0139] In some embodiments, the composition will be administered in a dose escalation manner, such that successive administrations of the composition contain a higher concentration of composition than previous administrations. In some embodiments, the composition will be administered in a manner such that successive administrations of the composition contain a lower concentration of composition than previous administrations.

[0140] In therapeutic applications, compositions are administered to a patient suffering from a disease in an amount sufficient to treat the patient. Therapeutic applications of a composition described herein include reducing transmissibility, slowing disease progression, reducing bacterial viability or replication, or inhibiting the expression of proteins required for toxicity, such as by 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% or 10% of the levels at which they would occur in individuals who are not treated with the composition.

[0141] In prophylactic embodiments, compositions are administered to a human or other mammal to induce an immune response that can inhibit the establishment of an infectious disease or other condition. In some embodiments, a composition may partially block the bacterium from establishing an infection.

[0142] A therapeutically effective amount may be determined on a population basis, and is not required to be an amount that naturally induces a protective response in a particular subject.

[0143] In some embodiments, the compositions are administered in combination with antibiotics. This co-administration is particularly appropriate for a patient who has recently been exposed (or is suspected of having been recently exposed) to, colonized or infected by *S. pneumoniae*. Many antibiotics are used to treat pneumococcal infections, including penicillin, amoxicillin, amoxicillin/clavulanate, cefuroxime, cefotaxime, ceftriaxone, and vancomycin. The appropriate antibiotic may be selected based on the type and severity of the infection, as well as any known antibiotic resistance of the infection (Jacobs MR “Drug-resistant *Streptococcus pneumoniae*: rational antibiotic choices” Am J Med. 1999 May 3;106(5A):19S-25S).

[0144] In some embodiments, the compositions are administered in combination with one or more additional vaccines, such as Prevnar-13 (Wyeth Pharmaceuticals, Inc.,

Philadelphia, PA) or Pneumovax (Merck & Co., Inc., Whitehouse Station, NJ) . This co-administration can be used, for example, to treat a patient who has been exposed to, colonized or infected by *S. pneumoniae* serotypes that are not represented in said existing vaccine.

2. Routes of administration

[0145] The vaccine formulations and pharmaceutical compositions herein can be delivered by administration to an individual, typically by systemic administration (e.g., by intravenous, intraperitoneal, intramuscular, intradermal, subcutaneous, subdermal, transdermal, intracranial, intranasal, mucosal, anal, vaginal, oral, or buccal routes; by inhalation; or intratracheal or bronchial instillation) or they can be administered by topical application. In some embodiments, the route of administration is intramuscular. In other embodiments, the route of administration is subcutaneous. In yet other embodiments, the route of administration is mucosal. In certain embodiments, the route of administration is transdermal or intradermal.

[0146] Certain routes of administration are particularly appropriate for vaccine formulations and immunogenic compositions comprising specified adjuvants. In particular, transdermal administration is one suitable route of administration for *S. pneumoniae* vaccines comprising toxins (e.g. cholera toxin or labile toxin); in other embodiments, the administration is intranasal. Vaccines formulated with Alphavirus replicons may be administered, for example, by the intramuscular or the subcutaneous route. Vaccines comprising aluminum hydroxide or aluminum phosphate (alum), Monophosphory Lipid A (MPL), Trehalose Dicoynomycolate (TDM), and dioctadecyldimethylammonium bromide (DDA) are suitable (*inter alia*) for intramuscular and subcutaneous administration. A vaccine comprising resiquimod may be administered topically or subcutaneously, for example.

3. Formulations

[0147] The vaccine formulation or immunogenic composition may be suitable for administration to a human patient, and vaccine or immunogenic composition preparation may conform to USFDA guidelines. In some embodiments, the vaccine formulation or immunogenic composition is suitable for administration to a non-human animal. In some embodiments, the vaccine or immunogenic composition is substantially free of either endotoxins or exotoxins. Endotoxins may include pyrogens, such as lipopolysaccharide (LPS) molecules. The vaccine or immunogenic composition may also be substantially free of inactive protein fragments which may cause a fever or other side effects. In some embodiments, the composition contains less than 1%, less than 0.1%, less than 0.01%, less than 0.001%, or less than 0.0001% of endotoxins, exotoxins, and/or inactive protein

fragments. In some embodiments, the vaccine or immunogenic composition has lower levels of pyrogens than industrial water, tap water, or distilled water. Other vaccine or immunogenic composition components may be purified using methods known in the art, such as ion-exchange chromatography, ultrafiltration, or distillation. In other embodiments, the pyrogens may be inactivated or destroyed prior to administration to a patient. Raw materials for vaccines, such as water, buffers, salts and other chemicals may also be screened and depyrogenated. All materials in the vaccine may be sterile, and each lot of the vaccine may be tested for sterility. Thus, in certain embodiments the endotoxin levels in the vaccine fall below the levels set by the USFDA, for example 0.2 endotoxin (EU)/kg of product for an intrathecal injectable composition; 5 EU/kg of product for a non-intrathecal injectable composition, and 0.25-0.5 EU/mL for sterile water.

[0148] In certain embodiments, the preparation comprises less than 50%, 20%, 10%, or 5% (by dry weight) contaminating protein. In certain embodiments, a fusion protein described herein is present in the substantial absence of other biological macromolecules, such as other proteins (particularly other proteins which may substantially mask, diminish, confuse or alter the characteristics of the component proteins either as purified preparations or in their function in the subject reconstituted mixture). In certain embodiments, at least 80%, 90%, 95%, 99%, or 99.8% (by dry weight) of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). In some embodiments, the vaccine or immunogenic composition comprising purified subunit proteins contains less than 5%, 2%, 1%, 0.5%, 0.2%, 0.1% of protein from host cells in which the subunit proteins were expressed, relative to the amount of purified subunit. In some embodiments, the desired polypeptides are substantially free of nucleic acids and/or carbohydrates. For instance, in some embodiments, the vaccine or immunogenic composition contains less than 5%, less than 2%, less than 1%, less than 0.5%, less than 0.2%, or less than 0.1% host cell DNA and/or RNA. In certain embodiments, at least 80%, 90%, 95%, 99%, or 99.8% (by dry weight) of biological macromolecules of the same type are present in the preparation (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present).

[0149] It is preferred that the vaccine or immunogenic composition has low or no toxicity, within a reasonable risk-benefit ratio. In certain embodiments, the vaccine or immunogenic composition comprises ingredients at concentrations that are less than LD₅₀ measurements for the animal being vaccinated. LD₅₀ measurements may be obtained in mice

or other experimental model systems, and extrapolated to humans and other animals. Methods for estimating the LD₅₀ of compounds in humans and other animals are well-known in the art. A vaccine formulation or immunogenic composition, and any component within it, might have an LD₅₀ value in rats of greater than 100 g/kg, greater than 50g/kg, greater than 20 g/kg, greater than 10 g/kg, greater than 5 g/kg, greater than 2 g/kg, greater than 1 g/kg, greater than 500 mg/kg, greater than 200 mg/kg, greater than 100 mg/kg, greater than 50 mg/kg, greater than 20 mg/kg, or greater than 10 mg/kg. A vaccine formulation or immunogenic composition that comprises a toxin such as botulinum toxin (which can be used as an adjuvant) should contain significantly less than the LD₅₀ of botulinum toxin.

[0150] The formulations suitable for introduction of the vaccine formulations or pharmaceutical composition vary according to route of administration. Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, intranasal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampoules and vials.

[0151] Injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. In the case of adoptive transfer of therapeutic T cells, the cells can be administered intravenously or parenterally.

[0152] Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of a fusion protein described herein or packaged nucleic acids suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, tragacanth, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the active ingredient in an inert base,

such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art. The pharmaceutical compositions can be encapsulated, e.g., in liposomes, or in a formulation that provides for slow release of the active ingredient.

[0153] A fusion protein described herein, alone or in combination with other suitable components, can be made into aerosol formulations (e.g., they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. Aerosol formulations can be delivered orally or nasally.

[0154] Suitable formulations for vaginal or rectal administration include, for example, suppositories, which consist of the polypeptides or packaged nucleic acids with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a combination of the polypeptides or packaged nucleic acids with a base, including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

H. Preparation and Storage of Vaccine Formulations and Immunogenic Compositions

[0155] The *S. pneumoniae* vaccines and immunogenic compositions described herein may be produced using a variety of techniques. For example, a polypeptide (e.g., a fusion protein) may be produced using recombinant DNA technology in a suitable host cell. A suitable host cell may be bacterial, yeast, mammalian, or other type of cell. The host cell may be modified to express an exogenous copy of one of the relevant polypeptide genes. Typically, the gene is operably linked to appropriate regulatory sequences such as a strong promoter and a polyadenylation sequence. In some embodiments, the promoter is inducible or repressible. Other regulatory sequences may provide for secretion or excretion of the polypeptide of interest or retention of the polypeptide of interest in the cytoplasm or in the membrane, depending on how one wishes to purify the polypeptide. The gene may be present on an extrachromosomal plasmid, or may be integrated into the host genome. One of skill in the art will recognize that it is not necessary to use a nucleic acid 100% identical to the naturally-occurring sequence. Rather, some alterations to these sequences are tolerated and may be desirable. For instance, the nucleic acid may be altered to take advantage of the degeneracy of the genetic code such that the encoded polypeptide remains the same. In some

embodiments, the gene is codon-optimized to improve expression in a particular host. The nucleic acid may be produced, for example, by PCR or by chemical synthesis.

[0156] Once a recombinant cell line has been produced, a polypeptide (e.g., fusion protein) may be isolated from it. The isolation may be accomplished, for example, by affinity purification techniques or by physical separation techniques (e.g., a size column).

[0157] In a further aspect of the present disclosure, there is provided a method of manufacture comprising mixing one or more fusion proteins described herein with a carrier and/or an adjuvant.

[0158] In some embodiments, fusion proteins for inclusion the vaccine formulations and immunogenic compositions may be produced in cell culture. One method comprises providing one or more expression vectors and cloning nucleotides encoding one or more fusion proteins described herein, then expressing and isolating the fusion proteins.

[0159] The immunogenic fusion proteins described herein, and nucleic acid compositions that express the fusion proteins, can be packaged in packs, dispenser devices, and kits for administering nucleic acid compositions to a mammal. For example, packs or dispenser devices that contain one or more unit dosage forms are provided. Typically, instructions for administration of the composition will be provided with the packaging, along with a suitable indication on the label that the composition is suitable for treatment of an indicated condition, such as those disclosed herein.

[0160] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described herein.

[0161] The disclosure is further illustrated by the following examples. The examples are provided for illustrative purposes only. They are not to be construed as limiting the scope or content of the disclosure in any way.

EXAMPLES

Example 1

SP2108, SP0148 and SP1634 polypeptides

[0162] The SP2108 polypeptide (SEQ ID NO: 6), SP0148 polypeptide (SEQ ID NO: 2) and SP1634 polypeptide (SEQ ID NO: 25) were formulated as vaccine compositions using 4 µg of the polypeptide in combination with 1 µg cholera toxin adjuvant (CT). For combinations, 4 µg of each polypeptide was used. The compositions were administered intranasally to C57BL/6 mice three times, one week apart. The mice were then allowed to rest for 3 weeks, and bled at that time for immunogenicity. For this assay, heparinized whole blood was collected from the retrograde orbital sinus. The total PBMC were stimulated with either killed, unencapsulated whole cell *S. pneumoniae* (WCC) or a combination of the three polypeptides in round bottomed tubes for three days. The supernatants were then harvested and evaluated by ELISA for IL-17 levels. Cholera toxin alone (CT) or an unrelated antigen from HSV (003) were used as negative controls. Results of the IL-17 immunogenicity assay are shown in FIGS. 1 and 2, where the left panels show data in scatter format, and the right panels show data as averages with standard deviations. The subjects were allowed to rest an additional 2 weeks, at which time they were challenged with intranasal administration of live, encapsulated *S. pneumoniae*. The subjects were sacrificed a week later, and the number of colony-forming units (CFU) was counted from nasal washes. Results of the colonization assay are shown in FIG. 3.

Example 2

SP0148, SP0314, SP0882, and SP2108 polypeptides tested in the BALB/c mouse

[0163] To determine whether similar immune responses were seen across different mouse genotypes, vaccine compositions were administered to BALB/c mice. Vaccine compositions comprised the polypeptides SP0148 (SEQ ID NO: 2), SP0314 (SEQ ID NO: 26), SP0882 (SEQ ID NO: 17) or SP2108 (SEQ ID NO: 6), and cholera toxin adjuvant (CT). Using a protocol similar to that of Example 1, the mice were immunized, challenged intranasally with *S. pneumoniae*, and the number of CFU was recorded. The results of this colonization experiment are shown in FIG. 4.

Example 3

SP1912, SP2108 and SP0148 polypeptides: IL-17A immunogenicity assay

[0164] The polypeptides SP1912 (SEQ ID NO: 9), SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) were formulated as vaccine compositions with cholera toxin

adjuvant (CT). The vaccine compositions were administered to mice two times, one week apart. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + CT (WCB), and the negative controls were CT alone or recombinant proteins without CT (with the exception of SP1912). Three weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated in a whole blood assay. Briefly, the heparinized whole blood was diluted in media and then cultured in duplicate with A) the protein of immunization, or B) the whole cell vaccine for six days. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 5. Each symbol in the graph represents responses from individual mice, and the line indicates the median response of the group.

Example 4

SP1912, SP2108 and SP0148 polypeptides: colonization assay

[0165] Animals were immunized with vaccine formulations comprising the polypeptides SP1912 (SEQ ID NO:9), SP2108 (SEQ ID NO: 6) or SP0148 (SEQ ID NO: 2) and cholera toxin adjuvant (CT) as described in Example 3, and then challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization (and one week after retroorbital blood collection). Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results are shown in FIG. 6 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. (***) = p-value <0.05).

Example 5

SP1912 polypeptide: aspiration challenge (sepsis assay)

[0166] Polypeptide SP1912 was evaluated for its ability to protect mice from sepsis. Groups of ten mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising either the SP1912 polypeptide (SEQ ID NO: 9) or pneumolysoid (PdT) adsorbed to alum. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + alum (WCB), and the negative control was alum alone. Three weeks after the final immunization, blood was collected for evaluation of IL-17A response and antibody levels, and then one week later, the mice underwent aspiration challenge with 10^7 live strain 0603 (type 6B) *S. pneumoniae*. Animals were monitored for survival for eight days. Results of the aspiration challenge are shown in FIG. 7 as survival curves for each immunized group.

Example 6**Pneumolysoid PdT, SP0148 and SP0641N polypeptides: aspiration challenge (sepsis assay)**

[0167] Polypeptide SP0148 was evaluated for its ability to protect mice from sepsis when immunized singly or in combination with SP0641N and/or pneumolysoid (PdT). Groups of ten mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising polypeptide SP0148 (SEQ ID NO: 2), singly or in combination with polypeptide SP0641N (SEQ ID NO: 13) and/or PdT, adsorbed to alum. The positive control was killed, unencapsulated whole cell *S. pneumoniae* + alum (WCB), and the negative control was alum alone. Three weeks after the final immunization, blood was collected for evaluation of IL-17 and antibody, and then one week later, the mice underwent aspiration challenge with 10^7 live strain 0603 (type 6B) *S. pneumoniae*. Animals were monitored for survival for eight days. The data are shown in FIG. 8 as survival curves for each immunized group.

Example 7**SP1912, SP2108 and SP0148 polypeptides: colonization assay**

[0168] Additional studies were performed essentially as described in Example 4, for a total of four separate studies. Briefly, animals were immunized with vaccine formulations comprising the polypeptides SP1912 (SEQ ID NO: 9), SP2108 (SEQ ID NO: 6), SP0148 (SEQ ID NO: 2), or additionally SP2108 plus SP0148, and cholera toxin adjuvant (CT) as described in Example 3. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* plus CT (WCV), or CT alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Pooled results of four studies are shown in FIG. 9 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. (***) = p-value <0.05). N indicates the total number of animals evaluated. Percentages refer to the number of animals protected from colonization.

Example 8**SP1912 and SP0148 polypeptides: IL-17A immunogenicity assay**

[0169] Groups of ten mice were subcutaneously immunized twice, two weeks apart with vaccine compositions comprising either SP1912 polypeptide (SEQ ID NO: 9), SP0148

polypeptide (SEQ ID NO: 2), or both adsorbed to alum. Control animals were immunized with alum alone. Three weeks after the last immunization, heparinized blood was collected by cardiac puncture and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparinized whole blood was diluted in media and then cultured for six days with the protein(s) of immunization. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 10. Each symbol in the graph represents responses from individual mice, and the line indicates the median response of the group.

Example 9

SP1912 and SP0148 polypeptides: colonization assay

[0170] Animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2) at different doses plus and minus SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results are shown in FIG. 11 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The number of animals protected from colonization out of the number of animals in the group is indicated at the top of the figure.

Example 10

SP1912, SP0148, and SP2108 polypeptides: colonization assay

[0171] In two separate studies, animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2) and SP0148 plus SP1912 (SEQ ID NO: 9), or additionally with SP2108 (SEQ ID NO: 6), SP2108 plus SP0148, and SP2108 plus SP1912, adsorbed to alum. Control animals were immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Pooled results of the two studies are shown in FIG. 12 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The

number of animals protected from colonization out of the number of animals in the group and corresponding percentage of animals protected from colonization are indicated at the top of the figure. (*p<0.05, **p<0.01, ***p<0.001 Dunn's Multiple Comparison Test compared to Alum control)

Example 11

Pneumolysoid L460D, PspA derivative CD2, SP1912, SP0148, and SP2108 polypeptides: colonization assay

[0172] Animals were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the polypeptides SP0148 (SEQ ID NO: 2), SP2108 (SEQ ID NO: 6), SP0148 plus SP2108, and SP0148 plus SP2108 in combination with SP1912 (SEQ ID NO: 9) or known *S. pneumoniae* antigens L460D plus CD2, adsorbed to alum. Two separate studies were conducted. Control animals were immunized with alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results of the second study are shown in FIG. 13 as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The number of animals protected from colonization out of the number of animals in the group is indicated at the top of the figure.

The chart below shows the absolute number and corresponding percentage of animals protected from colonization in the four studies described in Examples 10 and 11.

	1	2	3	4	# not colonized/ total	% not colonized
Alum	0/9	1/10	2/10	4/10	7/39	18%
WCV	9/9	9/10	9/10	9/10	36/39	92%
0148	8/18	4/10	5/9	1/10	18/47	38%
2108		8/10	6/10	4/10	18/30	60%
1912		4/10			4/10	40%
0148 + 2108		6/10	3/10	6/10	15/30	50%
0148 + 1912	13/18	3/10			16/28	57%
2108 + 1912		3/10			3/10	30%
0148 + 2108 + 1912		8/9	8/10	6/10	22/29	76%

0148 + 2108 + L460D + CD2			2/10	6/10	8/20	40%
------------------------------	--	--	------	------	------	-----

Example 12**PspA, SP0148 and SP2108 passive antibody transfer and aspiration challenge (sepsis assay)**

[0173] Groups of ten mice were injected with monoclonal antibodies specific for PspA, heat-inactivated rabbit sera specific for SP0148, SP2108, or combinations of these. Antibody and antisera concentrations and total injection volumes were adjusted with normal rabbit serum (NRS) and PBS. Control animals were injected with NRS, or serum against killed, unencapsulated whole cell *S. pneumoniae* (WCV). One day after injection, the mice underwent aspiration challenge with 10^6 live *S. pneumoniae* type WU-2 (ST-3). Animals were monitored for survival for eight days. The data are shown in FIG. 14 as survival curves for each immunized group.

[0174] FIG. 15 shows the percent of animals protected from sepsis in six separate aspiration challenge studies, two of which are described in more detail in Examples 6 and 12.

Example 13**SP0148, SP2108, and SP1912 fusion proteins: IL-17A immunogenicity assay**

[0175] Groups of four mice were subcutaneously immunized three times, two weeks apart with vaccine compositions comprising the fusion proteins SP2108/0148/1912 (SEQ ID NO: 53), SP0148/2108 (SEQ ID NO: 48), SP0148/1912 (SEQ ID NO: 50) or the comparable combinations of the single antigens SP2108 (SEQ ID NO: 6), SP0148 (SEQ ID NO: 2) and SP1912 (SEQ ID NO: 9), adsorbed to alum. Control animals were immunized with alum alone. Two weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparinized whole blood was diluted in media and then cultured for six days with single (non-fusion) proteins of immunization. The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay are shown in FIG. 16, Panels A-C. Each bar represents the mean $\pm$ SD IL-17A response from each group of four mice in response to the indicated antigen.

Example 14**SP0148, SP2108, and SP1912 fusion proteins: colonization assay**

[0176] Animals are subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the fusion proteins SP2108/0148 (SEQ ID NO: 47),

SP0148/2108 (SEQ ID NO: 48), SP2108/1912 (SEQ ID NO: 49), SP0148/1912 (SEQ ID NO: 50), SP2108/1912/0148 (SEQ ID NO: 51), SP0148/1912/2108 (SEQ ID NO: 52), SP2108/0148/1912 (SEQ ID NO: 53), or SP0148/2108/1912 (SEQ ID NO: 54), adsorbed to alum. Control animals are immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Immunized animals are challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals are euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Results are shown as the colony forming units of bacteria (CFU) per lavage. Each symbol represents a titer from an individual mouse response, and the horizontal line represents the median of the group. The absolute number and corresponding percentage of animals protected from colonization are also indicated.

Example 15

SP0148, SP2108, and SP1912 fusion proteins: aspiration challenge (sepsis assay)

[0177] Groups of ten mice are subcutaneously immunized three times, two weeks apart with vaccine compositions comprising the fusion proteins SP2108/0148 (SEQ ID NO: 47), SP0148/2108 (SEQ ID NO: 48), SP2108/1912 (SEQ ID NO: 49), SP0148/1912 (SEQ ID NO: 50), SP2108/1912/0148 (SEQ ID NO: 51), SP0148/1912/2108 (SEQ ID NO: 52), SP2108/0148/1912 (SEQ ID NO: 53), or SP0148/2108/1912 (SEQ ID NO: 54), adsorbed to alum. Control animals are immunized with killed, unencapsulated whole cell *S. pneumoniae* plus alum (WCV), or alum alone. Three weeks after the final immunization, blood is collected for evaluation of IL-17A and antibody, and then one week later, the mice undergo aspiration challenge with 10^7 live strain 0603 (type 6B) *S. pneumoniae*. Animals are monitored for survival for eight days. Results are reported as survival curves for each immunized group.

Example 16

SP0148, SP2108, SP1912, CD2 and H70 fusion proteins: colonization assay

[0178] In two separate studies, mice were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the fusion proteins SP0148/1912 (SEQ ID NO: 50), SP0148/1912 (SEQ ID NO: 50) plus SP2108 (SEQ ID NO: 6), SP2108/1912 (SEQ ID NO: 49), SP2108/1912 (SEQ ID NO: 49) plus SP0148 (SEQ ID NO: 2), SP2108/CD2 (SEQ ID NO: 57) plus SP0148/1912 (SEQ ID NO: 50), SP0148/CD2 (SEQ ID NO: 57) plus SP2108/1912 (SEQ ID NO: 49), SP2108/H70 (SEQ ID NO: 58) plus SP0148/1912 (SEQ ID NO: 50), and SP0148/H70 (SEQ ID NO: 56) plus SP2108/1912 (SEQ ID NO: 49), adsorbed

to alum. Control animals were immunized with SP2108 (SEQ ID NO: 6) plus SP0148 (SEQ ID NO: 2) plus SP1912 (SEQ ID NO: 9) adsorbed to alum, or alum alone. Immunized animals were challenged intranasally with 10^7 live type 6B *S. pneumoniae* four weeks after the last immunization. Seven days after challenge, animals were euthanized and the nasopharyngeal cavities lavaged and cultured on permissive media to evaluate the *S. pneumoniae* titers. Pooled results of the two studies are shown in FIG. 17 as the colony forming units of bacteria (CFU) per nasal wash. Each symbol represents the number of colonies cultured from an individual mouse, and the horizontal line represents the median of the group. The difference in mean log number of CFUs recovered from immunized mice relative to alum-only control mice is indicated at the top of the figure.

Example 17

SP0148, SP2108, SP1912, CD2 and H70 fusion proteins: intravenous challenge (sepsis assay)

[0179] Mice were subcutaneously immunized three times, two weeks apart with vaccine formulations comprising the fusion proteins SP2108/H70 (SEQ ID NO: 58), SP2108/CD2 (SEQ ID NO: 57), SP0148/H70 (SEQ ID NO: 56), SP0148/CD2 (SEQ ID NO: 55), SP2108/H70 (SEQ ID NO: 58) plus SP0148/1912 (SEQ ID NO: 50), and SP0148/H70 (SEQ ID NO: 56) plus SP2108/1912 (SEQ ID NO: 49), adsorbed to alum. Control animals were immunized with H70 (SEQ ID NO: 33) or CD2 (SEQ ID NO: 29) adsorbed to alum, or alum alone. Immunized animals were challenged intravenously with 368 CFUs of strain A66.1 *S. pneumoniae* three weeks after the last immunization. Mice were observed daily for fourteen days. Results are shown in FIG. 18 as hours to moribund. [NB: In this figure, CD2 is labeled PRN.] Each symbol represents the survival hours of an individual mouse, and the horizontal line represents the median hours to moribund of the group. The table below the figure indicates the mean and median hours to moribund and the standard error.

Example 18

SP0148, SP2108, SP1912, CD2 and H70 fusion proteins: IL-17A immunogenicity assay

[0180] Mice were immunized in two separate studies as described in Example 16. Two weeks after the last immunization, peripheral blood was collected from the retroorbital sinus and evaluated for IL-17A levels in a whole blood assay. Briefly, the heparinized whole blood was diluted in media and then cultured for six days with killed, unencapsulated whole cell *S. pneumoniae* (WCA). The supernatants were harvested and IL-17A levels measured by ELISA. Results of the IL-17A immunogenicity assay for each study are shown in FIG. 19, Panels A and B. Individual symbols represent the cytokine secretion in whole blood from

each mouse. Each horizontal line represents the mean IL-17A response from each group of animals in response to the indicated antigen.

Example 19**SP0148, SP2108, SP1912, CD2 and H70 fusion proteins: antibody response**

[0181] Mice were immunized as described in Example 18 and bled 16-20 days after the last immunization. Total IgG antibody titers against CD2 or H70 were measured by standard ELISA. FIG. 20 shows the mean ($\pm$ SD) IgG titer for the group expressed as the log 10 μ g/ml. [NB: In this figure, CD2 is labeled PRN.]

EQUIVALENTS

[0182] It is to be understood that while the disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

SEQUENCES

SEQ ID NO: 1

SP0148 lacking signal sequence

MCSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDSKDYDKFEKTEWSGVFAGLDADR
 YNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEVVQATTSKQLEAYNAEHTDNPT
 ILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRI
 KELYKDGTTLEKLSKQFFGDTYLPAEADIKE

SEQ ID NO: 2

SP0148 including signal sequence (277 amino acids with N-terminal E)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
 KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
 VQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
 PYVYPLLAQGQDELKSFVDKRIKELYKDGTTLEKLSKQFFGDTYLPAEADIKE

SEQ ID NO: 3

SP0148 consensus lacking signal sequence

MCSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDSKDYDKFE

Q S R N N X

KTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDD

I E

IGGKSTEVVQATTSKQLEAYNAEHTDNPTILNYTKADLQQIMVRLSDGQFDYKIFDKIG

F

VETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTTLEKLSKQ

Y S

FFGDTYLPAEADIK (E)

SEQ ID NO: 4

SP0148 consensus including signal sequence

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELT

G L Q S R N

GYEIEVVRAIFKDS DKYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERA EKYLYAAP

N X

I

IAQNP NVLVVKKDDSSIKSLDDIGGKSTEVVQATTS AKQLEAYNAEHTDNPTILNYTKAD

E

LQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGGQDELK

F

Y

S

SFVDKRIKELYKDGTLEKLSKQFFGDTYLP AEADIK (E)

SEQ ID NO: 5

SP2108 lacking signal sequence

MCGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVM
MAPYDRVGSLSGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLA
KDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPKGMQ
DTEGAGNLIQTQFQEGKTA AIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQK
FVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPLPNISQMSAVWDP AKNMLF DAVSG
QKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 6

SP2108 including signal sequence

>gi|14973620|gb|AAK76167.1| maltose/maltodextrin ABC transporter,
maltose/maltodextrin-binding protein (Streptococcus pneumoniae TIGR4)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVM MAPYDRVGSLSGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMQDTEGAGNLIQTQFQEGKTA AIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
PNISQMSAVWDP AKNMLF DAVSGQKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 7

SP2108 consensus lacking signal sequence

MCGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLD

A

I

KLSLDNQSGNVPDVM MAPYDRVGSLSGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGA

I

X

T

PAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGL

A

LAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYKWPKGMDTEGAGNLIQTQFQE

G P A X H

GKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQK

A

FVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPLPNISQMSA

S A S

VWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 8

SP2108 consensus including signal sequence

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVA

T T V A

KAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMAPYDRVGSLSGSDGQLSEVKLS

I I X

DGAKTDDTTKSLVTAANGKVYGAHAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAF

T

AGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWY

A G P A X

EKWPKGMDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY

H

AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDE

A S A

LTTAVIKQFKNTQPLPNISQMSAVWDPAKNMLFDAVSGQKDAKTAANDAVTLIKETIKQK

S

FGE

SEQ ID NO: 9

SP1912

MNGMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSMDGEIATLYVQVYESSLES LVGGV
IFEDGRHYTFVYENEDLVYEEEV L

SEQ ID NO: 10

SP1912L

MRYLATLLLSLAVLITAGCKKVADDRKLMKTQEELTEIVRDHFSMDGEIATLYVQVYESSLES LVGGVIFEDGRH
YTFVYENEDLVYEEEV L

SEQ ID NO: 11

SP1912 consensus

MNGMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSMDGEIATLY

H A L

VQVYESSLES LVGGVIFEDGRHYTFVYENEDLVYEEEV L

I L S I

SEQ ID NO: 12

SP0641

>gi|14972117|gb|AAK74791.1| serine protease, subtilase family

[Streptococcus pneumoniae TIGR4]

MKKSTVLSLTAAVILAAYAPNEVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKKTAVIKEK
EVVSKNPVIDNNTSNEEAKIKEENSNSKSGDYTDSFVNKNTENPKKEDKVVIYIAEFKDKESGEKAIKELSSLKNT
KVLYTYDRIFNGSAIETTPDNLDKIKQIEGISSVERAQKVQPMNHRKEIGVEEAIDYLSINAPFGKNFDGRG
MVISNIDTGTDRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYNGGKITVEKYDDGRDYFDPHG
MHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFS YKMYSDAGSGFAGDETMFHAIEDSIKHNVDVSVSSGFTGTG
LVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVANNHLKMTDTGNVTRTAAHEDAI AVASAKNQTVFED
KVNIGGESFKYRNIGAFFDKSKITTNEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKK
AMDKGARAIMVVNTVNYNRDNWTELPAMGYEADGTSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDK
LEQYYPIDMESFNSNKPNGVDEKEIDFKFAPDTPDKELYKEDIIVPAGSTSWGPRIDL LKPDVSAPGKNIKSTLN
VINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKE
KSQYFASPRQQAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSA
SAITTDLSLTDRLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAKNKNKF
VESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLGGYDDD GKP KIPGTL
NKGIGGEHGIDKFNPAQVIONRKDKNTTSLDQNP ELF AFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGL
TPSPLVLRSAEEGLISIVNTNKEGENQRDLKVISREHFIRGILNSKSNDAKGKSSKLKVWGLDKWDGLIYNPRG
REENAPESKDNQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLIT
KDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV TYAGEGQGRNRKLDKDGNTI
YEIKGAGDLRGKII EVIALDGSSNFTKIHRIKFANQADEKGMI SYLLVDPDQDSSKYQKLGEIAESKFKNLGNKG
EGSLKKD TTGV EHHHQENEESI KEKSSFTIDRNISTIRDFENKDLKKLIK KKFREVDDFTSETGKRMEEYDYKYD

DKGNIIAYDDGTDLEYETEKLEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNKYKSIEIKATKYDFHS
 KTMFTDLYANINDIVIDGLAFAGDMRLFVKDNDQKKAIEIKIRMPEKIKETKSEYPYVSSYGNVIELGEGDLSKNKP
 DNLTKMESGKIYSDSEKQYLLKDNIILRKGYALKVTTYNPGKTDMLEGNGVYSKEDIAKIQKANPNLRALSETT
 IYADSRNVEDGRSTQSVLMSALDGFNIIRYQVFTFKMNDKGEAIDKDGNLVTDSKLVLFKGDDKEYTGEDKFN
 EAIKEDGSMLFIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIYGLHI
 DNTRDFNIKLVKDGIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNKAVGVHYQFLYDNVKPEVNIDPK
 GNTSIEYADGKSVFNINDKRNNGFDGEIQEQHIYINGKEYTSFNDIKQIIDKTLNIKIVVKDFARNTTVKEFIL
 NKDTGEVSELKPHRVTVTIQNGKEMSSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTF
 IKPVFKKIEEKKEEENKPTFDVSKKKDNPQVNHSQNLNESHKEDLQREEHSQKSDSTKDVDTATVLDKNNISSKST
 TNNPNKLPKTGTASGAQTLAAGIMFIVGIFLGLKKKNQD

SEQ ID NO: 13

SP0641N

MVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAKIKE
 ENSNKSQGDYTDSTFVNKNTENPKKEDKVVIYAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDNL
 DKIKQIEGISSVERAQKVQPMNHARKEIGVEEAIDYLSINAPFGKNFDGRGMVISNIDTGTDRHKAMRIDDD
 AKASMRFKKEDLKGTDKNYWSDKIPHAFNYNGGKITVEKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGI
 DGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIEDSIKHNVDVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVV
 ATGNYATSASSSSWDLVANNHLKMTDTGNVTRTAAHEDAIAVASAKNQTFEFDKVNIGGESFKYRNIGAFFDKSK
 ITTNEGTGKAPSKLKFVYIGKGQDQDLIGLDRGKIAVMDRIYTKDLKNAFKKAMDKGARAIMVNTVNYNRDN
 WTLPAMGYEADGTSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDKLEQYYPIDMESFNSKNPNVGDE
 KEIDFKFAPDTEKELYKEDIIVPAGSTSWGPRIDLLLPDVSAPGKNIKSTLNIVINGKSTYG

SEQ ID NO: 14

SP0641M

MSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQ
 QGAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD
 RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAKNKNKFVESFIHFESV
 EEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEESRSKTLGGYDDDGKPKIPGTLNKGIGGEHGI
 DKFNPAGVIQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSA
 EEGLISIVNTNKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLDGLIYNPRGREENAPESKD
 NQDPATKIRGQFEPIAEGQYFYKFYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNEPEIKLITKDTYHKVKDQ
 YKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEVITYAGEGQGRNRKLDKDGNTIYEIKGAGDLR
 GKIIIEVIALDGSSNFTKIHRIFANQADEKGMISYYLVDPDQDSSKYQ

SEQ ID NO: 15

SP641N consensus

MVVLADTSSSEDALNISDKEKVA-----ENKEKHENIHSAMETS

- N S VVDKET KD N I

L

QDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAK
K TE TI EG A T TK R

IKEENSNSQGDYTDSEFVNKNTENPKKEDKVYIAEFKDKESG
D- Q H Q
NAH SA G RL

EKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDN
S Q N G Q
G

LDKIKQIEGISSVERAQKVQPMNHARKEIGVEEAIDYL
T I

KSINAPFGKNFDGRGMVISNIDTGTDIRHKAMRIDDDA

KASMRFKKEDLKGTDKNYWLSDKIPHAFNYNGGKITV

EKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGIDGI

APNAQIFSUKMYSDAGSGFAGDETMFHAIEDSIKHNVDDV

SVSSGFTGTGLVGEKYWQAIRALRKAGIPMVATGNY

ATSASSSSWDLVANNHLKMTDTGNVTRTAHEDAIAVA

SAKNQTVFEDKVNIGGESFKYRNIGAFFDKSKITTNEDG
Q N

TKAPSKLKFVYIGKGQDQDLIGLDRGKIAVMDRIYTKDL

KNAFKKAMDKGARAIMVVNTVNYYNRDNWTELPAMGY

EADGEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE

N

DFKDKLEQYYPIDMESFNSNKPNGDEKEIDFKFAPDT

DKELYKEDIIVPAGSTSWGPRIDLLLKPDVSAPGKNIKSTLNVINGKSTYG

SEQ ID NO: 16

SP641M consensus

MSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKI

K T

DLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQG

AGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDK

KYFTIKLHNTSNRPLTFKVSASAITTDSLTDRLKL

V

DETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSS

Y

FDLNAVINVGEAKNKNKFVESFIHFESVEEMEAL

R A

NSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLG

S TD K ME

GYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGV

IQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYP

R D D Q V H E T

LDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNT

NKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSK
K V G

LKVWGLKWDGLIYNPRGREENAPESKDNQDPATKIRGQF

EPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVS

VDFSNPEKIKLITKDTYHKVKDQYKNETLFARDQKEH

PEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQ

GRNRKLDKDGNTIYEIKGAGDLRGKIIIEVIALDGSSNFT
S A

KIHRIKFANQADEKGMISYYLVDPDQDSSKYQ
DH K A E

SEQ ID NO: 17

SP0882

>gi|14972356|gb|AAK75009.1| conserved hypothetical protein (Streptococcus pneumoniae TIGR4)

MNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNP
DISRMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAEFVMEVVKPFIDETYRTKADCQHTAMIGSSL
GGNITQFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFIYVGTEEADDTDKTLMGNIKQAYID
SSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 18

SP0882N

MNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIKRNP
DISRMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAEFVMEVVKPFI

SEQ ID NO: 19

SP0882 with exogenous signal sequence

MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDG
 QNVFNSKESFIGHSWKIIPAICRNPDISMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAEFVMEV
 VKPFIDETYRTKADCQHTAMIGSSLGGNITQFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFI
 YVGTEEADDDTKTMDGNIQAYIDSSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 20

SP0882N with exogenous signal sequence

MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDG
 QNVFNSKESFIGHSWKIIPAICRNPDISMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAEFVMEV
 VKPFI

SEQ ID NO: 21

SP0882 consensus

MNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESF
 I Y

IGHSWKIIPAICRNPDISMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAE

Y H E E

FVMEVVKPFIDETYRTKADCQHTAMIGSSLGGNITQFIGLEYQDQIGCLGVFSSANWLHQ
 EK

EAFNRYFECQKLSPDQRIFIYVGTEEADDDTKTMDGNIQAYIDSSLCYYHDLIAGGVH

I H R

LDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

SEQ ID NO: 22

SP0882N consensus

MNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYEKDTDRSYPVVYFHDGQNVFNSKESF
 I Y

IGHSWKIIPAICRNPDISMIVVAIDNDGMGRMNEYAAWKQESPIPGQQFGGKGVEYAE

Y H E E

FVMEVVKPFI

SEQ ID NO: 23

SP0882 consensus with exogenous signal sequence

MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYE

T T V I

KDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIAKRNPDISMIVVAIDNDGMGRMNE

Y Y H

YAAWKQESPIPGQQFGGKGVEYAEFVMEVVKPFIDETYRTKADCQHTAMIGSSLGNIT

E E

QFIGLEYQDQIGCLGVFSSANWLHQEAFNRYFECQKLSPDQRIFIYVGTEEADDTDKTLM

EK I H

DGNIKQAYIDSSLCYYHDLIAGGVHLDNLVLKVQSGAIHSEIPWSENLPDCLRFFAEKW

R

SEQ ID NO: 24

SP0882N consensus with exogenous signal sequence

MSSKFMKSAAVLGTATLASLLLVACMNQSYFYLMKEHKLKVPYTGKERRVRILLPKDYE

T T V I

KDTDRSYPVVYFHDGQNVFNSKESFIGHSWKIIPAIAKRNPDISMIVVAIDNDGMGRMNE

Y Y H

YAAWKQESPIPGQQFGGKGVEYAEFVMEVVKPFI

E E

SEQ ID NO: 25

SP1634

>gi|14973124|gb|AAK75714.1| hypothetical protein SP_1634 Streptococcus pneumoniae TIGR4

MANIFDYLKDVAYDSYYDLPLNELDILTIEITYLSFDNLVSTLPQRLLDLAPQVPRDPTMLTSKNRLQLLDELA
 QHKRFKNCKLSHFINDIDPELQKQFAAMTYRVSLDLYLIVFRGTDDSIIGWKEDFHLTYMKEIPAQKHALRYLKN
 FFAHHPKQKVILAGHSKGGNLAIYAASQIEQSLQNQITAVYTFDAPGLHQELTQTAGYQRIMDRSKIFIPQGSII
 GMMLEIPAQIIIVQSTALGGIAQHDTFSWQIEDKHVQLDKTNSDSQQVDTTFKEWVATVPDEELQLYFDLFFGT
 ILDAGISSINDLASLKALEYIHHLFVQAQSLTPEERETLGRLTQLLIDTRYQAWKNR

SEQ ID NO: 26

SP0314

>gi|14971788|gb|AAK74491.1| hyaluronidase Streptococcus pneumoniae
 TIGR4MQTKTKKLIVSLSSLVLSGFLLNHYMTIGAEETTTNTIQQSQKEVQYQQRDTKNLVENGDFGQTEDGSSP
 WTGSKAQGWSAWVDQKNSADASTRVIEAKDGAITISSHEKLRALHRMVPIEAKKKYKLRFKIKTDNKGIAKVR
 IIEESGKDKRLWNSATTSGTKDWQTIEADYSPTLDVDKIKLELFYETGTGTVSFKDIELVEVADQLSEDSQTDKQ
 LEEKIDLPIGKKHVFSLADYTYKVENPDVASVKNGLILEPLKEGTTNVIVSKDGKEVKKIPLKILASVKDAYTDRL
 DDWNGIIAGNQYYDSKNEQMAKLNQELEGKVADSLSSISSQADRTYLWEKFSNYKTSANLTATYRKLEEMAKQVT
 NPSSRYQDETIVVRTVRDSMEWMHKKHVYNSEKSIVGNWWDYEIGTPRAINNTLSLMKEYFSDEEIKKYTDVIEKF
 VPDPEHFRKTTDNPFKALGGNLVDMGRVKVIAGLLRKDDQEISSTIRSIEQVFKLVDQGEFGFYQDGSYIDHTNVA
 YTGAYGNVLIDGLSOLLPIVQKTKNPIDKDKMQTMYHWIDKSFAPLLVNGELMDMSRGRSISRANSEGHVAAVEV
 LRGIHRIADMSEGETKQCLQSLVKTIQSDSYVDVFNKLTQYKDISLMQSLSDAGVASVPRPSYLSAFNKMDKT
 AMYNAEKGFGLSLFSSRTLNYEHMNKENKRGWYTS DGMFYLYNGDL SHYSDGYWPTVNPYKMPGTTETDAKRA
 DSDTGKVLPSAFVGT SKLDDANATATMDFTNWNQTLTAHKSWMFLKDKIAFLGSNIQNTSTD TAATTIDQRKLES
 GNPYKVYVNDKEASLTEQEKDYPETQSVFLESFDSKKNIGYFFFKSSISMSKALQKGAWKDINEGQSDKEVENE
 FLTISQAHKQNRDSYGMYLIPNVDRATFNQMIKELESSLIENNETLQSVYDAKQGVGWGIVKYDDSVSTISNQFQV
 LKRGVYTIRKEGDEYKIAYYNPETQESAPDQEVFKKLEQAAQPQVQNSKEKEKSEEEKNHSDQKNLPQTGEGQSI
 LASLGFLLLGAFYLFRRGKNN

SEQ ID NO: 27

PspA

MNKKKMILTSLASVAILGAGFVASSPTFVRAEEAPVANQSKAEKDYDAAVKKSEAAKDYETAKKKAEDA
 QKKYDEDQKKTEAKAEKERKASEKIAEATKEVQQAYLAYLQASNESQRKEADKKIKEATQRKDEAEAAFA
 TIRTTIVVPEPESELAETKKKAEAEATKEAEVAKKKSEEAKEVEVEKNKILEQDAENEKKIDVLQNKVADL
 EKGIAPYQNEVAELNKEIARLQSDLKDAEENNVEDIKEGLEQAITNKKAE LATTQQNIDKTQKDLEDAE
 LELEKVLATLDPEGKTQDELDKEAAEAEELNEKVEALQNVQAELEEEELSKLEDNLKDAETNNVEDYIKEGL
 EEAIATKKAELEKTQKELDAALNELGPDGDEETPAPAPQPEKPAEEPENPAPAPKPEKSADQQAEDDYA
 RRSEEEYNRLTQQQPPKAEKPAPAPQPEQPAPAPKIGWKQENGWYFYNTDGSMATGWLQNNGSWYYLNS
 NGAMATGWLQYNGSWYYLNANGAMATGWLQYNGSWYYLNANGAMATGWLQYNGSWYYLNANGDMATGWLQ
 YNGSWYYLNANGDMATGWAKVHGSWYYLNANGSMATGWVKDGETWYYLEASGSMKANQWFQVSDKWYYVN
 GLGSLSVNTTVDGYKVNANGWV

SEQ ID NO: 28

Immunogenic PspA/PspC polypeptides including the coiled-coil structure (PR
 + NPB)

MSDKIIHLTDDSFDTVDLKDAGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVA
 KLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSQGLKEFLDANLAGSGSG
 HHHHHHHSSGLVPRGSGMKETA AAKFERQHMDSPDLGTDDDDKAMADLKKAVNE
 PEKPAEEPENPAPAPKPAAPQPEKPAAPAPKPEKSADQQAEDDYARRSEEEYNRLTQQ
 QPPKAEKPAPAPVPKPEQPAPAPKTGWGQENGWCRQACGRTRAPPPPLRSGC

SEQ ID NO: 29

Immunogenic PspA/PspC polypeptide CD2

ADLKKAVNEPEKPAEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEEEDYARR
SEEEYNRLTQQQPPKAEKPAPAPVVPKPEQPAPAPKTGWGQENGMW

SEQ ID NO: 30

Immunogenic PspA/PspC polypeptides lacking the coiled-coil structure (PR + NPB)

MAKKAEELEKTPEKPAEEPENPAPAPQPEKSADQQAEEEDYARRSEEEYNRLTQQQPPKA

SEQ ID NO: 31

Immunogenic PspA/PspC polypeptides including the coiled-coil structure (PR only)

MSDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVA
KLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSG
HMHSHHHSSGLVPRGSGMKETAATAKFERQHMDSPDLGTDDDDKAMADLKKAVNE
PETPAPAPAPAPAPTPEAPAPAPAPAPKPAPAPKPAPAPKPAPAPKPAPAPKP
APAPAPAPKPEKPAEKAPAPKPKETPKTGWKQENGMWCRQACGRTRAPPPPLRSG

SEQ ID NO: 32

Immunogenic PspA/PspC polypeptides lacking the coiled-coil structure (PR only)

DLKKAVNEPETPAPAPAPAPAPTPEAPAPAPAPAPKPAPAPKPAPAPKPAPAPKPA
PAPKPAPAPKPAPAPAPAPKPEKPAEKAPAPKPKETPKTGWKQENGMW

SEQ ID NO: 33

Immunogenic PspA/PspC polypeptide H70 (aa 290-410 of PspA)

YFKEGLEKTIAAKKAEELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPK
PEKPAEQPKPEKTDDQQAEEEDYARRSEEEYNRLTQQQPPKAEKPAPAPKTGWKQENGM

SEQ ID NO: 34

Non-proline Block (NPB)

EKSADQQAEEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 35

Non-proline Block (NPB)

DQQAEEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 36

Non-proline Block (NPB)

MEKSADQQAEEEDYARRSEEEYNRLTQQQ

SEQ ID NO: 37

SP0148 lacking signal sequence (nucleotides)

ATGTGCTCAGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCA
CCAAAGCCATTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAA
GATTCTGACAAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGT
TACAATATGGCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCC
CAAAATCCTAATGTCTTGTGCTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAATTCG
ACGGAAGTCGTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAAC
ATCCTTAACCTATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATT
TTTGATAAAATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGAACAATTGAAAGTTATCGAACTTCCAAGC
GACCAACAACCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATC
AAAGAACTTTATAAAGATGGAACCTTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAA
GCTGATATTAAAGAGTAA

SEQ ID NO: 38

SP0148 including signal sequence (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGCTGCTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAATTCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGAACAATTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACCTTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGAGTAA

SEQ ID NO: 39

SP2108 lacking signal sequence (nucleotides)

ATGTGCGGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTA
GACGAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTT
AAAAGTGGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATG
ATGGCTCCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTCAGAAGTGAAATTGAGCGATGGTGCT
AAAACAGACGACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAG
TCACTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAACATTTGCTGACTTGAAAACCTTGCT
AAAGATAGCAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAACTTCTAC
TATACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGT

CTTGCAAACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAA
GATACAGAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCT
TGGAAAGCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAA
GAATATGCTGCATTCCGGTGGTGGTAAAGCTTGGGTCATTCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAA
TTTGTAGACTTCCTTGTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACT
GAGGCTCGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAA
CCACTGCCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGT
CAAAAAGATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAA
TAA

SEQ ID NO: 40

SP1912 (nucleotides)

ATGAATGGTATGAAAGCTAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
TCCGACATGGGGGAAATTGCGACCCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGA

SEQ ID NO: 41

SP1912L (nucleotides)

ATGAGATACCTGGCAACATTGTTGTTATCTCTGGCGGTGTTAATCACCGCCGGGTGCAAAAAAGTTGCAGATGAC
CGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATGGGGGAAATTGCG
ACCCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAGGATGGCCGTCAT
TATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGA

SEQ ID NO: 42

SP0641N (nucleotides)

ATGGTAGTCTTAGCAGACACATCTAGCTCTGAAGATGCTTTAAACATCTCTGATAAAGAAAAAGTAGCAGAAAAAT
AAAGAGAAACATGAAAAATCCATAGTGCTATGGAACTTCACAGGATTTTAAAGAGAAGAAAAACAGCAGTCATT
AAGGAAAAAGAAGTTGTTAGTAAAAATCCTGTGATAGACAATAACACTAGCAATGAAGAAGCAAAAAATCAAAGAA
GAAAAATCCAATAAATCCCAAGGAGATTATACGGACTCATTTGTGAATAAAAAACACAGAAAAATCCCAAAAAAGAA
GATAAAGTTGTCTATATTGCTGAATTTAAAGATAAAGAATCTGGAGAAAAAGCAATCAAGGAACTATCCAGTCTT
AAGAATACAAAAGTTTTATATACTTATGATAGAATTTTAAACGGTAGTGCCATAGAAACAACCTCCAGATAACTTG
GACAAAATTAAACAAAATAGAAGGTATTTTCATCGGTTGAAAGGGCACAAAAAGTCCAACCCATGATGAATCATGCC
AGAAAGGAAATTGGAGTTGAGGAAGCTATTGATTACCTAAAGTCTATCAATGCTCCGTTTGGGAAAAATTTTGAT
GGTAGAGGTATGGTCATTTCAAATATCGATACTGGAACAGATTATAGACATAAGGCTATGAGAATCGATGATGAT
GCCAAAGCCTCAATGAGATTTAAAAAAGAAGACTTAAAGGCACTGATAAAAAATTATTGGTTGAGTGATAAAATC
CCTCATGCGTTCAATTATTATAATGGTGGCAAAATCACTGTAGAAAAATATGATGATGGAAGGGATTATTTTGAC
CCACATGGGATGCATATTGCAGGGATTCTTGCTGGAAATGATACTGAACAAGACATCAAAAACTTTAACGGCATA
GATGGAATTGCACCTAATGCACAAATTTTCTCTTACAAATGTATTCTGACGCAGGATCTGGGTTTGCGGGTGAT
GAAACAATGTTTCATGCTATTGAAGATTCTATCAACACAACGTTGATGTTGTTTCGGTATCATCTGGTTTTACA
GGAACAGGTCTTGTAGGTGAGAAATATTGGCAAGCTATTCGGGCATTAAGAAAAGCAGGCATTCCAATGGTTGTC

GCTACGGGTAACATATGCGACTTCTGCTTCAAGTTCTTCATGGGATTTAGTAGCAAATAATCATCTGAAAATGACC
GACACTGGAAATGTAACACGAACTGCAGCACATGAAGATGCGATAGCGGTCGCTTCTGCTAAAAATCAAACAGTT
GAGTTTGATAAAGTTAACATAGGTGGAGAAAGTTTTAAATACAGAAATATAGGGGCCTTTTTCGATAAGAGTAAA
ATCACAACAAATGAAGATGGAAACAAAAGCTCCTAGTAAATTTAAATTTGTATATATAGGCAAGGGGCAAGACCAA
GATTTGATAGGTTTGGATCTTAGGGGCAAAATTGCAGTAATGGATAGAATTTATACAAAGGATTTAAAAAATGCT
TTTAAAAAAGCTATGGATAAGGGTGCACGCGCCATTATGGTTGTAAATACTGTAAATTACTACAATAGAGATAAT
TGGACAGAGCTTCCAGCTATGGGATATGAAGCGGATGAAGGTACTAAAAGTCAAGTGTTTTCAATTTTCAGGAGAT
GATGGTGTAAAGCTATGGAACATGATTAATCCTGATAAAAAAAGTGAAGTCAAAGAAATAATAAAGAAGATTTT
AAAGATAAATTGGAGCAATACTATCCAATTGATATGGAAAGTTTTAATTCCAACAAACCGAATGTAGGTGACGAA
AAAGAGATTGACTTTAAGTTTGCACCTGACACAGACAAAGAACTCTATAAAGAAGATATCATCGTTCCAGCAGGA
TCTACATCTTGGGGGCCAAGAATAGATTTACTTTTAAACCCGATGTTTCAGCACCTGGTAAAAATATTTAAATCC
ACGCTTAATGTTATTAATGGCAAATCAACTTATGGC

SEQ ID NO: 43

SP0882N DNA

ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT
GTACGTATTCTTCTTCTTAAAGATTATGAGAAAGATACAGACCGTTCCCTATCCTGTTGTATACTTTTCATGACGGG
CAAAATGTTTTTAATAGCAAAGAGTCTTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG
GATATCAGTCGCATGATTGTCTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTCTATGGAGGTG
GTCAAGCCTTTTATC

SEQ ID NO: 44

SP0882 with exogenous signal sequence (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTTGGTAGCTTGC
ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT
GTACGTATTCTTCTTCTTAAAGATTATGAGAAAGATACAGACCGTTCCCTATCCTGTTGTATACTTTTCATGACGGG
CAAAATGTTTTTAATAGCAAAGAGTCTTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG
GATATCAGTCGCATGATTGTCTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTCTATGGAGGTG
GTCAAGCCTTTTATCGATGAGACCTATCGTACAAAAGCAGACTGCCAGCATACGGCTATGATTGGTTCCCTACTA
GGAGGCAATATTACCCAGTTTATCGGTTTGGAAATACCAAGACCAAATTGGTTGCTTGGGCGTTTTTTTCATCTGCA
AACTGGCTCCACCAAGAAGCCTTTAACCGCTATTTTCGAGTGCCAGAACTATCGCCTGACCAGCGCATCTTCATC
TATGTAGGAACAGAAGAAGCAGATGATACAGACAAGACCTTGATGGATGGCAATATCAAACAAGCCTATATCGAC
TCGTCGCTTTGCTATTACCATGATTTGATAGCAGGGGAGTACATCTGGATAATCTTGTGCTAAAAGTTCAGTCT
GGTGCCATCCATAGTGAATCCCTTGGTCAGAAAATCTACCAGATTGTCTGAGATTTTTTGCAGAAAAATGGTAA

SEQ ID NO: 45

SP0882N with exogenous signal sequence (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTTGGTAGCTTGC
ATGAATCAATCCTACTTTTATCTAAAAATGAAAGAACACAAACTCAAGGTTCCTTATACAGGTAAGGAGCGCCGT

GTACGTATTCTTCTTCTTCTTAAAGATTATGAGAAAGATACAGACCGTTCTATCCTGTTGTATACTTTTCATGACGGG
CAAAATGTTTTTAAATAGCAAAGAGTCTTTTCATTGGACATTCATGGAAGATTATCCCAGCTATCAAACGAAATCCG
GATATCAGTCGCATGATTGTCTGTTGCTATTGACAATGATGGTATGGGGCGGATGAATGAGTATGCGGCTTGGAAG
TTCCAAGAATCTCTATCCCAGGGCAGCAGTTTGGTGGTAAGGGTGTGGAGTATGCTGAGTTTGTTCATGGAGGTG
GTCAAGCCTTTTATC

SEQ ID NO: 46

PSPA (nucleotides)

TTGACAAATATTTACGAGGAGGCTTATGCTTAATATAAGTATAGGCTAAAAATGATTATCAGAAAAGAG
GTAAATTTAGATGAATAAGAAAAAATGATTTTAAAGCCTAGCCAGCGTCGCTATCTTAGGGGCTGGT
TTTGTTCGCTCTTCGCCCTACTTTTGTAAAGAGCAGAAGAAGCTCCTGTAGCTAACCAGTCTAAAGCTGAGA
AAGACTATGATGCAGCAGTGAAAAAATCTGAAGCTGCTAAGAAAGATTACGAAACGGCTAAAAAGAAAGC
AGAAGACGCTCAGAAGAAATATGATGAGGATCAGAAGAAAACTGAGGCAAAAGCGGAAAAAGAAAGAAAA
GCTTCTGAAAAGATAGCTGAGGCAACAAAAGAAGTTCAACAAGCGTACCTAGCTTATCTACAAGCTAGCA
ACGAAAGTCAGAGAAAAGAGGCAGATAAGAAGATAAAAGAAGCTACGCAACGCAAAGATGAGGCGGAAGC
TGCATTTGCTACTATTTCGAACAACAATTGTAGTTCCCTGAACCAAGTGAGTTAGCTGAGACTAAGAAAAAA
GCAGAAGAGGCAACAAAAGAAGCAGAAGTAGCTAAGAAAAAATCTGAAGAGGCAGCTAAAGAGGTAGAAG
TAGAGAAAAATAAAATACTTGAACAAGATGCTGAAAACGAAAAGAAAATTGACGTACTTCAAAACAAAGT
CGCTGATTTAGAAAAAGGAATTGCTCCTTATCAAAACGAAGTCGCTGAATTAAATAAAGAAAATTGCTAGA
CTTCAAAGCGATTTAAAAGATGCTGAAGAAAATAATGTAGAAGACTACATTAAAGAAGGTTTAGAGCAAG
CTATCACTAATAAAAAAGCTGAATTAGCTACAACCTCAACAAAACATAGATAAAACTCAAAAAGATTTAGA
GGATGCTGAATTAGAACTTGAAAAAGTATTAGCTACATTAGACCCTGAAGGTAAAACTCAAGATGAATTA
GATAAAGAAGCTGCTGAAGCTGAGTTGAATGAAAAAGTTGAAGCTCTTCAAAACCAAGTTGCTGAATTAG
AAGAAGAAGCTTTCAAACTTGAAGATAATCTTAAAGATGCTGAAACAAACAACGTTGAAGACTACATTAA
AGAAGGTTTAGAAGAAGCTATCGCGACTAAAAAGCTGAATTGGAAAAACTCAAAAAGAATTAGATGCA
GCTCTTAATGAGTTAGGCCCTGATGGAGATGAAGAAGAGACTCCAGCGCCGGCTCCTCAACCAGAAAAAC
CAGCTGAAGAGCCTGAGAAATCCAGCTCCAGCACCAAAACCAGAGAAGTCAGCAGATCAACAAGCTGAAGA
AGACTATGCTCGTAGATCAGAAGAAGAATATAATCGCTTGACCCAACAGCAACCGCCAAAAGCAGAAAAA
CCAGCTCCTGCACCACAACCAGAGCAACCAGCTCCTGCACCAAAAATAGGTTGGAACAAGAAAACGGTA
TGTGGTACTTCTACAATACTGATGGTTCAATGGCGACAGGTTGGCTACAAAACAACGGTTTCATGGTACTA
CCTCAACAGCAATGGCGCTATGGCTACAGGTTGGCTCCAATACAATGGTTTCATGGTATTACCTAAACGCT
AACGGCGCTATGGCGACAGGCTGGCTCCAATACAATGGCTCATGGTACTACCTCAACGCTAACGGCGCTA
TGGCGACAGGCTGGCTCCAATACAATGGCTCATGGTACTACCTCAACGCTAATGGTGATATGGCGACAGG
ATGGCTCCAATACAACGGTTTCATGGTATTACCTCAACGCTAATGGTGATATGGCTACAGGTTGGGCTAAA
GTCCACGGTTTCATGGTACTACCTCAACGCTAACGGTTCAATGGCAACAGGTTGGGTGAAAGATGGAGAAA
CCTGGTACTATCTTGAAGCATCAGGTTCTATGAAAGCAAAACCAATGGTTCCAAGTATCAGATAAATGGTA
CTATGTCAATGGTTTAGGTTCCCTTTTCAGTCAACACAACCTGTAGATGGCTATAAAGTCAATGCCAATGGT
GAATGGGTTTAAAGCCG

SEQ ID NO: 47

SP2108/SP0148 fusion protein (675 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKLT
 GDALGGLDKLSLDNQSGNVPDVMMAFYDRVGSLSGSDGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
 VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
 NDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
 AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKDELTTAVIKQFKNTQPL
 PNISQMSAVWDPKNNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGESGGAKKEGEAASKKEIIVATNGSPKPF
 IYEENGELTGYEIEVVRAIFKDSKDYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNP
 VLVVKKDDSSIKSLDDIGGKSTEVVQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKI
 GVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGGDELKSFVDKRIKELYKDGTLKLSKQFFGDTYLP AEADIK

SEQ ID NO: 48

SP0148/SP2108 fusion protein (674 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPF IYEENGELTGYEIEVVRAIFKDS
 KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPVLVVKKDDSSIKSLDDIGGKSTEV
 VQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
 PYVYPLLAQGGDELKSFVDKRIKELYKDGTLKLSKQFFGDTYLP AEADIKSGSKTADKPADSGSSEVKELTVYVD
 EGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAFYDRVGSLSGSDGQLSEVKLSGDAK
 TDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYY
 TYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAIIIDGPW
 KAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTE
 ARSYAEGKDELTTAVIKQFKNTQPLPNISQMSAVWDPKNNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFG

SEQ ID NO: 49

SP2108/SP1912 fusion protein (521 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKLT
 GDALGGLDKLSLDNQSGNVPDVMMAFYDRVGSLSGSDGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
 VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
 NDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
 AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKDELTTAVIKQFKNTQPL
 PNISQMSAVWDPKNNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGENGMAKMMWAGLALLGIGSLALATKK
 VADDRKLMKTQE ELTEIVRDHFSMDGEIATLYVQVYESSLES LVGGVIFEDGRHYTFVYENEDLVYEEV

SEQ ID NO: 50

SP0148/SP1912 fusion protein (374 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPF IYEENGELTGYEIEVVRAIFKDS
 KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPVLVVKKDDSSIKSLDDIGGKSTEV
 VQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
 PYVYPLLAQGGDELKSFVDKRIKELYKDGTLKLSKQFFGDTYLP AEADIKNGMAKMMWAGLALLGIGSLALA
 TTKVADDRKLMKTQEELTEIVRDHFSMDGEIATLYVQVYESSLES LVGGVIFEDGRHYTFVYENEDLVYEEV

SEQ ID NO: 51

SP2108/SP1912/SP0148 fusion protein (773 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPIYDRVGSLSGSDGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDDELTTAVIKQFKNTQPL
PNISQMSAVWDPKNMLFDVSGQKDAKTAANDAVTLIKETIKQKFGENGMAKKMWMAGLALLGIGSLALATKK
VADDRKLMKTQE
ELTEIVRDHFSDMGEIATLYVQVYESSLESLSVGGVIFEDGRHYTFVYENEDLVYEEVLSGGAKKEGEAASKKEI
IVATNGSPKPFIIYEENGELTGYEIEVVRAIFKDSKDYDKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEY
LYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSD
GQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLKLSKQFFG
DTYLPAEADIK

SEQ ID NO: 52

SP0148/SP1912/SP2108 fusion protein (772 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIIYEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLKLSKQFFGDTYLPAEADIKNGMAKKMWMAGLALLGIGSLALA
TKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLSVGGVIFEDGRHYTFVYENEDLVYEEVLSG
SKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAPI
YDRVGSLSGSDGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDS
KYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPCKMQDTE
GAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFV
FLVATEQQKVLVDKTNEIPANTEARSYAEGKNDDELTTAVIKQFKNTQPLPNISQMSAVWDPKNMLFDVSGQK
AKTAANDAVTLIKETIKQKFGE

SEQ ID NO: 53

SP2108/SP0148/SP1912 fusion protein (773 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPIYDRVGSLSGSDGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDDELTTAVIKQFKNTQPL
PNISQMSAVWDPKNMLFDVSGQKDAKTAANDAVTLIKETIKQKFGESGGAKKEGEAASKKEIIVATNGSPKPF
IYEENGELTGYEIEVVRAIFKDSKDYDKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPN
VLVVKDDSSIKSLDDIGGKSTEVVQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKI
GVETVIKNQGLDNLKVIELPSDQQPYVYPLLAQGQDELKSFVDKRIKELYKDGTLKLSKQFFGDTYLPAEADIK
NGMAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSDMGEIATLYVQVYESSLESLSVGGVI
FEDGRHYTFVYENEDLVYEEVLS

SEQ ID NO: 54

SP0148/SP2108/SP1912 fusion protein (772 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKGSKTADKPADSGSSEVKELTVYVD
EGYKSYIEEVAKAYEKEAGVKVTLKTGDALGGLDKLSLDNQSGNVPDVMMAPIYDRVGSLSGSDGQLSEVKLSGAK
TDDTTKSLVTAANGKVYGAPAVIESLVMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYY
TYGLLAGNGAYVFGQNGKDAKDIGLANDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAIIDGPW
KAQAFKDAKVNYGVATIPTLPNGKEYAAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTE
ARSYAEGKNDELTTAVIKQFKNTQPLPNISQMSAVWDPKNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGEN
GMKAKKMWMAGLALLGIGSLALATKKVADDRKLMKTQEELTEIVRDHFSMDMGEIATLYVQVYESSLESLSVGGVIF
EDGRHYTFVYENEDLVYEEEV

SEQ ID NO: 55

SP0148/CD2 (389 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKMAMADLKKAVNEPEKPAEEEPENP
APAPKPAPAPQPEKPAPAPAPKPEKSADQQAEDYARRSEEEYNRLTQQQPPKAEKPAPAPVPKPEQPAPAPKTG
WKQENGMWHHHHHH

SEQ ID NO: 56

SP0148/H70 (404 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKYFKEGLEKTTAAKKAELEKTEADL
KKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQAPAPKPEKPAEQPKPEKTDDQQAEDYARRSEEEYNRLTQQ
QPPKAEKPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 57

SP2108/CD2 (536 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMAPIYDRVGSLSGSDGQLSEVKLSGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPCKMQDTEGAGNLIQTQFQEGKTAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDELTTAVIKQFKNTQPL

PNISQMSAVWDPKAKNMLF DAVSGQKDAKTAANDAVTLIKETIKQKFGEMAMADLKKAVNEPEKPAEEEEPENPAPA
PKPAPAPQPEKPAPAPAPKPEKSADQQAEEEDYARRSEEEYNRLTQQQPPKAEKPAPAPVPKPEQPAPAPKTGWKQ
ENGMWHHHHHH

SEQ ID NO: 58

SP2108/H70 (551 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSNGVPDVMMAFYDRVGSLSGDSGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMDTEGAGNLIQTQFQEGKTAIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
PNISQMSAVWDPKAKNMLF DAVSGQKDAKTAANDAVTLIKETIKQKFGYEFKEGLEKTI AAKKAELEKTEADLKK
VNEPEKPAPAPETPAPEAPAEQPKPAPAPQAPAPKPEKPAPAEQPKPEKTDDQQAEEEDYARRSEEEYNRLTQQQPP
KAEKPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 59

SP0148/LC/CD2 (403 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAAASKKEIIVATNGSPKPFIIYEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLP AEADIKGPKPHRIQSTPKGSMAMADLKKAV
NEPEKPAEEEEPENPAPAPKAPAPQPEKPAPAPAPKPEKSADQQAEEEDYARRSEEEYNRLTQQQPPKAEKPAPAP
VPKPEQPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 60

SP0148/LC/H70 (418 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAAASKKEIIVATNGSPKPFIIYEENGELTGYEIEVVRAIFKDS
KYDVKFEKTEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSAKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQ
PYVYPLLAQGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLP AEADIKGPKPHRIQSTPKGSYFKEGLEKTI
AAKKAEELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQAPAPKPEKPAPAEQPKPEKTDDQQAEEEDY
ARRSEEEYNRLTQQQPPKAEKPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 61

SP2108/LC/CD2 (550 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSNGVPDVMMAFYDRVGSLSGDSGQLSEVKLSGDAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPKGMDTEGAGNLIQTQFQEGKTAIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
PNISQMSAVWDPKAKNMLF DAVSGQKDAKTAANDAVTLIKETIKQKFGGPKPHRIQSTPKGSMAMADLKKAVNEP

EKPAEEEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEEEDYARRSEEEYNRLTQQQPPKAEKPAPAPVVK
PEQPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 62

SP2108/LC/H70 (565 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMPYDRVGSLSGQQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPQGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLYDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
PNISQMSAVWDPKMNMLFVAVSGQKDAKTAANDAVTLIKETIKQKFGGPKPHRIQSTPKGSYFKEGLEKTIAAK
KAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPKPEKPAEQPKPEKTDDQQAEEEDYARR
SEEEYNRLTQQQPPKAEKPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 63

SP0148/LR/CD2 (409 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
KYDVKFETEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKLAEATAKEATAKEATAKATAMAMA
DLKKAVNEPEKPAEEEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEEEDYARRSEEEYNRLTQQQPPKAE
KPAPAPVVKPEQPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 64

SP0148/LR/H70 (424 amino acids)

MKKIVKYSSLAALALVAAGVLAACSGGAKKEGEAASKKEIIVATNGSPKPFIEENGELTGYEIEVVRAIFKDS
KYDVKFETEWSGVFAGLDADRYNMAVNNLSYTKERAKEYLYAAPIAQNPNVLVVKDDSSIKSLDDIGGKSTEV
VQATTSKQLEAYNAEHTDNPTILNYTKADFQQIMVRLSDGQFDYKIFDKIGVETVIKNQGLDNLKVIELPSDQQ
PYVYPLLAQGGQDELKSFVDKRIKELYKDGTLEKLSKQFFGDTYLPAEADIKLAEATAKEATAKEATAKATAYFKE
GLEKTIAAKAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPKPEKPAEQPKPEKTDDQ
QAEEEDYARRSEEEYNRLTQQQPPKAEKPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 65

SP2108/LR/CD2 (565 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
GDALGGLDKLSLDNQSGNVPDVMMPYDRVGSLSGQQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFLADWTNFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
NDGSIVGINYAKSWYEKWPQGMQDTEGAGNLIQTQFQEGKTAAIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY

AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
 PNISQMSAVWDPKNNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGELAEATAKEATAKEATAKATAMAMADLK
 KAVNEPEKPAEEEPENPAPAPKPAPAPQPEKPAPAPAPKPEKSADQQAEDYARRSEEEYNRLTQQQPPKAEKPA
 PAPVPKPEQPAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 66

SP2108/LR/H70 (571 amino acids)

MSSKFMKSAAVLGTATLASLLLVACGSKTADKPADSGSSEVKELTVYVDEGYKSYIEEVAKAYEKEAGVKVTLKT
 GDALGGLDKLSLDNQSGNVPDVMMAPIYDRVGSLSGSDGQLSEVKLSDGAKTDDTTKSLVTAANGKVYGAPAVIESL
 VMYYNKDLVKDAPKTFADLENLAKDSKYAFAGEDGKTTAFADWTNFFYYTYGLLAGNGAYVFGQNGKDAKDIGLA
 NDGSIVGINYAKSWYEKWPKGMDTEGAGNLIQTQFQEGKTAIIIDGPWKAQAFKDAKVNYGVATIPTLPNGKEY
 AAFGGGKAWVIPQAVKNLEASQKFVDFLVATEQQKVLVDKTNEIPANTEARSYAEGKNDLTTAVIKQFKNTQPL
 PNISQMSAVWDPKNNMLFDAVSGQKDAKTAANDAVTLIKETIKQKFGELAEATAKEATAKEATAKATAYFKEGLE
 KTIAAKKAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQPAPAPKPEKPAEQPKPEKTDDQQA
 EYARRSEEEYNRLTQQQPPKAEKPAAPAPKTGWKQENGMWHHHHHH

SEQ ID NO: 67

SP2108/SP0148 fusion (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
 GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
 GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
 GGTGATGCTCTAGGAGGCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
 CCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAACA
 GACGACACAATACTCTTGTAAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
 GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAAAACCTTGCTAAAGAT
 AGCAAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
 TATGGACTTCTTGCCGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
 AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
 GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
 GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCACTCTTCCAAATGGAAGAATAT
 GCTGCATTCGGTGGTGGTAAAGCTTGGGTCAATCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
 GACTTCCTTGTGCAACTGAACAACAAAAGTATTATATGATAAGACTAACGAAATCCAGCTAATACTGAGGCT
 CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
 CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
 GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAATCAGGG
 GGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATACCAAAGCCATTT
 ATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGACAAA
 TATGATGTCAAGTTTGAAGACAGAATGGTCAGGTGCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCT
 GTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCAAAATCCTAAT
 GTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAAATCGACGGAAGTCGTT
 CAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAACAT

ACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAAATC
GGTGTTGAAACAGTGATCAAGAACCAAGGTTTGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAACCG
TACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTAT
AAAGATGGAACCTTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA

SEQ ID NO: 68

SP0148/SP2108 fusion (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGACGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAATAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTSCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACCTT
TATAAAGATGGAACCTTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGAC
GAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAA
ACTGGTGTATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATG
GCTCCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTCAGAAGTGAAATTGAGCGATGGTGCTAAA
ACAGACGACACAATAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCA
CTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGCTAAA
GATAGCAAATACGCATTGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAACTTCTACTAT
ACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTT
GCAAACGACGGTTCTATCGTAGGTATCAACTACGTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGAT
ACAGAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGG
AAAGCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAATGGAAGAA
TATGCTGCATTGCGTGGTGGTAAAGCTTGGGTCAATCTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTT
GTAGACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAG
GCTCGTTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCA
CTGCCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAA
AAAGATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAA

SEQ ID NO: 69

SP2108/SP1912 fusion (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAAGCTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT

CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
 GACGACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
 GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAAACCTTGCTAAAGAT
 AGCAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAACTTCTACTATACA
 TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
 AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
 GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
 GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCACTCTTCCAAATGGAAAAGAATAT
 GCTGCATTTCGGTGGTGGTAAAGCTTGGGTCAATTCCTCAAGCCGTTAAGAACCCTTGAAGCTTCTCAAAAATTTGTA
 GACTTCCTTGTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
 CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
 CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
 GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAAATGGT
 ATGAAAGCTAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAAAAA
 GTTGCGAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATG
 GGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAG
 GATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 70

SP0148/SP1912 fusion (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCGAGCCCTTGCTCTTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
 GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
 TTTATCTATGAAGAAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
 AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
 GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCAAAATCCT
 AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGAAAAATCGACGGAAGTC
 GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
 TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
 ATCGGTGTTGAAACAGTGATCAAGAACCAAGTTTGGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
 CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAAGTT
 TATAAAGATGGAACCTTGAAAAATTTGTCTAAACAATCTTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
 AAAAATGGTATGAAAGCTAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
 ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
 TCCGACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
 ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 71

SP2108/SP1912/SP0148 fusion (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAAGTCTACACTTGCTAGCTTGCTTTTGGTAGCTTG
 GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
 GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAGTCACTCTTAAACT

GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAACCTAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAAACCTTGCTAAAGAT
AGCAAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCACTCTTCCAAATGGAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCAATCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAATGGT
ATGAAAGCTAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAAAAA
GTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGACATG
GGGGAATTCGACCCCTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTTGAG
GATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTATGCTCAGGGGGT
GCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCATTTATC
TATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGACAAATAT
GATGTCAAGTTTGAAGAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCTGTC
AACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCTAATGTC
CTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAAATCGACGGAAGTCGTTCAA
GCCACTACATCAGCTAAGCAGTTAGAAGCATAACAATGCTGAACACACGGACAACCCAACCTATCCTTAACCTACT
AAGGCAGACTTCCAACAAATCATGGTACGTTTGGAGCATGGACAATTTGACTATAAGATTTTTGATAAAATCGGT
GTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACCTGAAAGTTATCGAACTTCCAAGCGACCAACAACCGTAC
GTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTATAAA
GATGGAACCTTTGAAAAATGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA

SEQ ID NO: 72

SP0148/SP1912/SP2108 fusion (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGACGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAGAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATAACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGGAGCATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACCTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT

TATAAAGATGGAACCTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAAATGGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCT
ACGAAAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTT
TCCGACATGGGGGAAATTGCGACCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTC
ATTTTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTAGGA
AGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAGGGA
TATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACTGGT
GATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCTCCA
TACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACAGAC
GACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTTGT
ATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGTCTAAAGATAGC
AAATACGCATTTCGCTGGTGAAGATGGTAAAACACTGCTTCCCTAGCTGACTGGACAACTTCTACTATACATAT
GGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCAAAAC
GACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACAGAA
GGTGCTGGAACTTAAATCCAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGAAGCT
CAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATATGCT
GCATTCCGTGGTGGTAAAGCTTGGGTCAATTCCTCAAGCCGTTAAGAACCCTGAAGCTTCTCAAAAATTTGTAGAC
TTCCTTGTTGCACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCTCGT
TCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTGCCA
AACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAAGAT
GCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAACACCACCAC
CACCACCACTGA

SEQ ID NO: 73

SP2108/SP0148/SP1912 fusion (nucleotides)

ATGTCTATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGTCTAAAGAT
AGCAAAATACGCATTTCGCTGGTGAAGATGGTAAAACACTGCTTCCCTAGCTGACTGGACAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAAATCCAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGGA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTCCGGTGGTGGTAAAGCTTGGGTCAATTCCTCAAGCCGTTAAGAACCCTGAAGCTTCTCAAAAATTTGTA
GACTTCTTGTTGCACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA

GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAATCAGGG
GGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCATTT
ATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGACAAA
TATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATGGCT
GTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAATCCTAAT
GTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTCGTT
CAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAACAT
ACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAAATC
GGTGTGTAACAGTGATCAAGAACCAAGGTTTGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAACCG
TACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTTTAT
AAAGATGGAACCTCTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATTAAA
AATGGTATGAAAGCTAAAAAATGTGGATGGCAGGCTTGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACG
AAAAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTCC
GACATGGGGGAAATTGCGACCCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATT
TTTGAGGATGGCCGTCATTATACCTTTGTCTATGAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 74

SP0148/SP2108/SP1912 fusion (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGAC
GAGGGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAGAAGCTGGAGTAAAAGTCACTCTTAAA
ACTGGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATG
GCTCCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTTCAGAAGTGAATTTGAGCGATGGTGCTAAA
ACAGACGACACAATAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCA
CTTGTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGCTAAA
GATAGCAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTAT
ACATATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTT
GCAAACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGAT
ACAGAAGGTGCTGGAACCTTAATCCAACTCAATTCGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGG
AAAGCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAGAA

TATGCTGCATTCCGGTGGTGGTAAAGCTTGGGTCATTCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTT
GTAGACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAG
GCTCGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCA
CTGCCAAACATCTCTCAAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAA
AAAGATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAAACAAAAATTTGGTGAAAAAT
GGTATGAAAGCTAAAAAAATGTGGATGGCAGGCTTGGCTCTGCTAGGTATCGGAAGCCTTGCTCTTGCTACGAAA
AAAGTTGCAGATGACCGTAAGCTCATGAAGACTCAGGAAGAGTTGACAGAGATTGTGCGAGACCATTTTTCCGAC
ATGGGGGAAATTGCGACCCCTTTATGTTCAAGTTTACGAAAGCAGTCTGGAGAGCTTGGTTGGTGGCGTCATTTTT
GAGGATGGCCGTCATTATACCTTTGTCTATGAAAATGAAGACCTAGTCTATGAGGAGGAAGTCTTA

SEQ ID NO: 75

SP0148/CD2 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAAGCTT
TATAAGATGGAACCTTGAAAAATTGTCTAAACAATCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAATGGCCATGGCCGACCTGAAGAAGGCCGTGAACGAGCCGGAGAAACCGGCCGAAGAGGAGCCGGAGAACCCG
GCACCGGCACCTAAACCTGCACCTGCACCGCAGCCTGAGAAGCCTGCCCTGCACCGGCACCGAAACCGGAGAAG
AGCGCCGACCAGCAGGCAGAGGAAGACTACGCCCCGCCGACGGAAGAGGAGTACAATCGCCTGACCCAGCAGCAA
CCGCCGAAGGCAGAAAAGCCGGCCCCGGCCCCGTTCCGAAACCTGAACAGCCGGCACCTGCACCTAAAACAGGC
TGGAAGCAGGAAAACGGCATGTGGCACCACCACCACCACCTGA

SEQ ID NO: 76

SP0148/H70 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAAGCTT

TATAAAGATGGAACCTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAATACTTCAAAGAGGGCTTAGAGAAGACCATCGCCGCCAAGAAGGCCGAGTTAGAGAAAACCGAGGCCGACCTG
AAGAAAGCCGTGAACGAGCCGGAGAAGCCTGCACCTGCACCTGAAACACCTGCCCCCTGAAGCACCGGCCGAACAG
CCTAAGCCGGCACC GGACCTCAACCTGCCCCGGCACCGAAGCCGGAAAAACCGGCCGAGCAGCCGAAACCGGAG
AAAACCGACGACCAGCAGGCAGAAGAGGACTATGCCCGCCGAGCGAGGAAGAGTACAACCGCCTGACCCAGCAG
CAGCCTCCGAAAGCCGAGAAACCGGCCCGGCCCGGAAAAACCGGCTGGAAGCAGGAGAACGGCATGTGGCACCAC
CACCACCACCACTGA

SEQ ID NO: 77

SP2108/CD2 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACCTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAACTTGCTAAAGAT
AGCAAAATACGCATTCGCTGGTGAAGATGGTAAAACCTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGAAACTTAATCCAAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGAAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCAATCCTCAAGCCGTTAAGAACCCTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAATGGCC
ATGGCCGACCTGAAGAAGGCCGTGAACGAGCCGGAGAAACCGGCCGAAGAGGAGCCGGAGAACCCGGCACCGGCA
CCTAAACCTGCACCTGCACCGCAGCCTGAGAAGCCTGCCCCCTGCACCGGCACCGAAACCGGAGAAGAGCGCCGAC
CAGCAGGCAGAGGAAGACTACGCCCGCCGAGCGAAGAGGAGTACAATCGCCTGACCCAGCAGCAACCGCCGAAG
GCAGAAAAGCCGGCCCCGGCCCCGGTTCCGAAACCTGAACAGCCGGCACCTGCACCTAAAACAGGCTGGAAGCAG
GAAAACGGCATGTGGCACCACCACCACCACCACTGA

SEQ ID NO: 78

SP2108/H70 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACCTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAATAAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT

GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAAACCTTGCTAAAGAT
AGCAAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAATCCAAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGAAAAAGAATAT
GCTGCATTCGGTGGTGGTAAAGCTTGGGTCAATTCCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAATAC TTC
AAAGAGGGCTTAGAGAAGACCATCGCCGCCAAGAAGGCCGAGTTAGAGAAAACCGAGGCCGACCTGAAGAAAGCC
GTGAACGAGCCGGAGAAGCCTGCACCTGCACCTGAAACACCTGCCCCCTGAAGCACCGGCCGAACAGCCTAAGCCG
GCACCGGCACCTCAACCTGCCCCGGCACCGAAGCCGAAAAACCGGCCGAGCAGCCGAAACCGGAGAAAACCGAC
GACCAGCAGGCAGAAGAGGACTATGCCCCGCCGAGCGAGGAAGAGTACAACCGCCTGACCCAGCAGCAGCCTCCG
AAAGCCGAGAAACCGGCCCCGGCCCCGAAAACCGGCTGGAAGCAGGAGAACGGCATGTGGCACCACCACCACCAC
CACTGA

SEQ ID NO: 79

SP0148/LC/CD2 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCGAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGAAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGTTTGGACAACCTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAACTT
TATAAAGATGGAACCTTGAAAAATTGTCTAAACAATCTTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGGCCCCGAAGCCGCACCGTATCCAGAGCACCCGAAGGGCAGCATGGCCATGGCAGACCTGAAGAAGGCCGTG
AACGAGCCGGAACCTGCCGAAGAGGAGCCGGAATCCTGCCCCCTGCACCTAAACCTGCACCTGCCCTCAG
CCGGAGAAACCGGCACCTGCACCGGCACCTAAGCCGGAGAAGAGCGCCGATCAGCAGGCCGAGGAGGACTATGCC
CGTCGCAGCGAAGAGGAGTACAACCGCCTGACCCAGCAACAGCCTCCGAAGGCCGAAAAACCGGCCCTGCACCG
GTGCCTAAGCCTGAGCAACCTGCCCCGGCCCCGAAAACCGGCTGGAAGCAGGAAAACGGCATGTGGCACCACCAC
CACCACCACTGA

SEQ ID NO: 80

SP0148/LC/H70 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAAATACCTCTATGCCGCACCAATTGCCCAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAATAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAATATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGTTTGGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGAAGTT
TATAAGATGGAACCTTGAAAAATTGTCTAAACAATCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAAGGCCCCGAAGCCGCACCGYATCCAGAGCACACCGAAAGGCAGCTACTTCAAAGAGGGCTTAGAGAAGACCATC
GCCGCCAAGAAGGCCGAGTTAGAGAAAACCGAGGCCGACCTGAAGAAGGCCGTGAACGAGCCGGAGAAACCTGCA
CCTGCACCGGAGACACCGGCACCTGAAGCCCCCTGCAGAGCAGCCTAAACCTGCCCCCTGCACCTCAACCGGCACCT
GCACCTAAGCCGAAAAACCTGCCGAGCAGCCGAAGCCGGAGAAAACCGACGACCAGCAGGCCGAAGAGGACTAT
GCACGCCGAGCGAAGAGGAGTACAACCGCTTAACCCAACAGCAGCCGCCGAAGGCAGAAAAACCGGCCCTGCC
CCTAAGACCGGTTGGAAGCAGGAGAACGGCATGTGGCACCACCACCACCACCTGA

SEQ ID NO: 81

SP2108/LC/CD2 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAAGTGTACACTTGCTAGCTTGCTTTTGGTAGCTTG
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTTATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGCTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAACA
GACGACACAATAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGCTAAAGAT
AGCAAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAACTTCTACTATACA
TATGGACTTCTTGCCGTAACGGTGCTTACGCTTTTGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGGA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCACTCTTCCAAATGGAAGAATAT
GCTGCATTGCGTGGTGGTAAAGCTTGGGTCATTCTCAAGCCGTTAAGAACCCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAGGCCCCG
AAGCCGCACCGTATCCAGAGCACCCCGAAGGGCAGCATGGCCATGGCAGACCTGAAGAAGGCCGTGAACGAGCCG
GAAAAACCTGCCGAAGAGGAGCCGGAATCCTGCCCTGCACCTAAACCTGCACCTGCCCCCTCAGCCGGAGAAA
CCGGCACCTGCACCGGCACCTAAGCCGGAGAAGAGCGCCGATCAGCAGGCCGAGGAGGACTATGCCCGTCGCAGC
GAAGAGGAGTACAACCGCTGACCCAGCAACAGCTCCGAAGGCCGAAAAACCGGCCCTGCACCGGTGCCCTAAG

CCTGAGCAACCTGCCCCGCCCCGAAAACCGGCTGGAAGCAGGAGAATGGCATGTGGCACCACCACCACCACCAC
TGA

SEQ ID NO: 82

SP2108/LC/H70 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTG
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAATAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAAACCTTGCTAAAGAT
AGCAAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAACTTCTACTATACA
TATGGACTTCTTGCCGGTAACGGTGCTTACGTCCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAAACAGCTGCTATCATCGACGGACCTTGAAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCACTCTTCCAAATGGAAAAGAATAT
GCTGCATTTCGGTGGTAAAGCTTGGGTCAATCCTCAAGCCGTTAAGAACCCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAGGCCCCG
AAGCCGCACCGCATCCAGAGCACACCGAAAGGCAGTACTTCAAAGAGGGCTTAGAGAAGACCATCGCCGCCAAG
AAGGCCGAGTTAGAGAAAACCGAGGCCGACCTGAAGAAGGCCGTGAACGAGCCGGAGAAACCTGCACCTGCACCG
GAGACACCGGCACCTGAAGCCCCTGCAGAGCAGCTAAACCTGCCCCTGCACCTCAACCGGCACCTGCACCTAAG
CCGGA AAAACCTGCCGAGCAGCCGAAGCCGGAGAAAACCGACGACCAGCAGGCCGAAGAGGACTATGCACGCCGC
AGCGAAGAGGAGTACAACCGCTTAACCCAACAGCAGCCGCCGAAGGCAGAAAAACCGGCCCTGCCCCTAAGACC
GGTTGGAAGCAGGAGAACGGCATGTGGCACCACCACCACCACCACTGA

SEQ ID NO: 83

SP0148/LR/CD2 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGACGCCCTTGCTCTTGTTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAGCCA
TTTATCTATGAAGAAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGA AAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAAGA ACTT

TATAAAGATGGAACCTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAACTGGCCGAAGCAACCGCCAAAGAAGCCACCGCCAAAGGAAGCCACAGCCAAGGCCACCGCAATGGCCATGGCC
GACCTGAAGAAGGCCGTGAATGAGCCGGAGAAGCCGGCCGAAGAAGAGCCGGAGAATCCTGCACCGGCCCTAAG
CCTGCACCGGCACCTCAGCCTGAAAAACCGGCACCTGCACCTGCCCCGAAGCCTGAAAAAGAGCGCCGACCAGCAG
GCCGAAGAGGACTATGCCCCCGCAGCGAAGAGGAGTACAACCGTCTGACCCAGCAGCAGCCGCCGAAAGCCGAG
AAACCTGCCCCGGCACCGGTGCCTAAGCCGGAACAGCCTGCCCCCTGCACCGAAGACCGGTTGGAAGCAGGAAAAAC
GGCATGTGGCACCACCACCACCACCCTGA

SEQ ID NO: 84

SP0148/LR/H70 (nucleotides)

ATGAAAAAATCGTTAAATACTCATCTCTTGCAGCCCTTGCTCTTGTGTGCTGCAGGTGTGCTTGCGGCTTGCTCA
GGGGGTGCTAAGAAAGAAGGAGAAGCAGCTAGCAAGAAAGAAATCATCGTTGCAACCAATGGATCACCAAAGCCA
TTTATCTATGAAGAAATGGCGAATTGACTGGTTACGAGATTGAAGTCGTTTCGCGCTATCTTTAAAGATTCTGAC
AAATATGATGTCAAGTTTGAAAAGACAGAATGGTCAGGTGTCTTTGCTGGTCTTGACGCTGATCGTTACAATATG
GCTGTCAACAATCTTAGCTACACTAAAGAACGTGCGGAGAAATACCTCTATGCCGCACCAATTGCCCAAAATCCT
AATGTCCTTGTCGTGAAGAAAGATGACTCTAGTATCAAGTCTCTCGATGATATCGGTGGAAAAATCGACGGAAGTC
GTTCAAGCCACTACATCAGCTAAGCAGTTAGAAGCATACAATGCTGAACACACGGACAACCCAACCTATCCTTAAC
TATACTAAGGCAGACTTCCAACAAATCATGGTACGTTTGAGCGATGGACAATTTGACTATAAGATTTTTGATAAA
ATCGGTGTTGAAACAGTGATCAAGAACCAAGGTTTGGAACAACCTTGAAAGTTATCGAACTTCCAAGCGACCAACAA
CCGTACGTTTACCCACTTCTTGCTCAGGGTCAAGATGAGTTGAAATCGTTTGTAGACAAACGCATCAAGAAGCTT
TATAAAGATGGAACCTCTTGAAAAATTGTCTAAACAATTCTTCGGAGACACTTATCTACCGGCAGAAGCTGATATT
AAACTGGCCGAAGCAACCGCCAAAGAGGCCACCGCCAAAGGAAGCCACCGCCAAAGCCACCGCCTACTTCAAAGAG
GGCTTAGAGAAGACCATCGCCGCCAAGAAGGCCGAGTTAGAGAAAACCGAGGCCGACCTGAAGAAGGCCGTGAAC
GAGCCGGAGAAGCCTGCACCTGCCCCCTGAAACCCCTGCACCGGAAGCACCTGCAGAGCAGCCTAAACCTGCACCT
GCACCGCAACCTGCCCCCTGCACCTAAACCGGAGAAACCGGCAGAGCAGCCGAAGCCGGAGAAAACCGACGACCAG
CAGGCCGAGGAAGACTACGCCGTCGTAGCGAGGAAGAGTACAACCGCCTGACCCAGCAGCAACCGCCGAAAGCC
GAAAAGCCGGCCCTGCACCGAAAACCGGCTGGAAGCAGGAGAACGGCATGTGGCACCACCACCACCACCCTGA

SEQ ID NO: 85

SP2108/LR/CD2 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACCTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAACAA
GACGACACAATAATCTCTTGTAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGGAACCTTGCTAAAGAT
AGCAATACGCATTTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACCTTAATCCAAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGGAAG

GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTTCGGTGGTGGTAAAGCTTGGGTCATTCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAC TGCC
GAGGCAACCGCCAAAGAAGCCACCGCCAAGGAAGCCACAGCCAAGGCCACCGCAATGGCCATGGCCGACCTGAAG
AAGGCCGTGAATGAGCCGAGAAGCCGGCCGAAGAAGAGCCGGAGAATCCTGCACCGGCCCTAAGCCTGCACCG
GCACCTCAGCCTGAAAAACCGGCACCTGCACCTGCCCCGAAGCCTGAAAAGAGCGCCGACCAGCAGGCCGAAGAG
GACTATGCCCCGCCGACGGAAGAGGAGTACAACCGTCTGACCCAGCAGCAGCCGCCGAAAGCCGAGAAACCTGCC
CCGGCACCGGTGCC TAAGCCGGAACAGCCTGCCCCTGACCCGAAGACCGGTTGGAAGCAGGARAACGGCATGTGG
CACCACCACCACCACCCTGA

SEQ ID NO: 86

SP2108/LR/H70 (nucleotides)

ATGTCATCTAAATTTATGAAGAGCGCTGCGGTGCTTGGAACTGCTACACTTGCTAGCTTGCTTTTGGTAGCTTGC
GGAAGCAAACTGCTGATAAGCCTGCTGATTCTGGTTCATCTGAAGTCAAAGAACTCACTGTATATGTAGACGAG
GGATATAAGAGCTATATTGAAGAGGTTGCTAAAGCTTATGAAAAAGAAGCTGGAGTAAAAGTCACTCTTAAACT
GGTGATGCTCTAGGAGGTCTTGATAAACTTTCTCTTGACAACCAATCTGGTAATGTCCCTGATGTTATGATGGCT
CCATACGACCGTGTAGGTAGCCTTGTTCTGACGGACAACCTTTCAGAAGTGAAATTGAGCGATGGTGCTAAAACA
GACGACACAATACTCTTGTAAACAGCTGCTAATGGTAAAGTTTACGGTGCTCCTGCCGTTATCGAGTCACTT
GTTATGTACTACAACAAAGACTTGGTGAAAGATGCTCCAAAAACATTTGCTGACTTGAAAAACCTTGCTAAAGAT
AGCAATACGCATTCGCTGGTGAAGATGGTAAACTACTGCCTTCCTAGCTGACTGGACAAACTTCTACTATACA
TATGGACTTCTTGCCGTAACGGTGCTTACGTCTTTGGCCAAAACGGTAAAGACGCTAAAGACATCGGTCTTGCA
AACGACGGTTCTATCGTAGGTATCAACTACGCTAAATCTTGGTACGAAAAATGGCCTAAAGGTATGCAAGATACA
GAAGGTGCTGGAACTTAATCCAACTCAATTCCAAGAAGGTAAACAGCTGCTATCATCGACGGACCTTGAAAA
GCTCAAGCCTTTAAAGATGCTAAAGTAACTACGGAGTTGCAACTATCCCAACTCTTCCAAATGGAAAAGAATAT
GCTGCATTTCGGTGGTGGTAAAGCTTGGGTCATTCTCAAGCCGTTAAGAACCTTGAAGCTTCTCAAAAATTTGTA
GACTTCCTTGTGCAACTGAACAACAAAAAGTATTATATGATAAGACTAACGAAATCCCAGCTAATACTGAGGCT
CGTTCATACGCTGAAGGTAAAAACGATGAGTTGACAACAGCTGTTATCAAACAGTTCAAGAACACTCAACCACTG
CCAAACATCTCTCAAATGTCTGCAGTTTGGGATCCAGCGAAAAATATGCTCTTTGATGCTGTAAGTGGTCAAAAA
GATGCTAAAACAGCTGCTAACGATGCTGTAACATTGATCAAAGAAACAATCAAACAAAAATTTGGTGAAC TGCC
GAAGCAACCGCCAAAGAGGCCACCGCCAAGGAAGCCACCGCCAAAGCCACCGCCTACTTCAAAGAGGGCTTAGAG
AAGACCATCGCCGCCAAGAAGGCCGAGTTAGAGAAAACCGAGGCCGACCTGAAGAAGGCCGTGAACGAGCCGGAG
AAGCCTGCACCTGCCCTGAAACCCCTGCACCGGAAGCACCTGCAGAGCAGCCTAAACCTGCACCTGCACCGCAA
CCTGCCCCTGACCTAAACCGGAGAAACCGGCAGAGCAGCCGAAGCCGGAGAAAACCGACGACCAGCAGGCCGAG
GAAGACTACGCCCGTCGTAGCGAGGAAGAGTACAACCGCCTGACCCAGCAGCAACCGCCGAAAGCCGAAAAGCCG
GCCCCTGACCGAAAACCGGCTGGAAGCAGGAGAACGGCATGTGGCACCACCACCACCACCCTGA

SEQ ID NO: 87

HHHHHH

SEQ ID NO: 88

MSYYHHHHHH

SEQ ID NO: 89

Canonical lipobox motif

[LIVMFESTAGPC]-[LVIAMFTG]-[IVMSTAGCP]-[AGS]-C

SEQ ID NO: 90

SP2108 signal sequence

MSSKFMKSAAVLGTATLASLLLVAC

SEQ ID NO: 91

E. coli RlpB signal sequence

MRYLATLLLSLAVLITAG[C]

SEQ ID NO: 92

Fusion protein linker (19-mer)

LGGGGSGGGSGGGGSAAA

SEQ ID NO: 93

Fusion protein LR linker (20-mer)

LAEATAKEATAKEATAKATA

SEQ ID NO: 94

Fusion protein LC linker (14-mer)

GPKPHRIQSTPKGS

SEQ ID NO: 95

Amino-terminal boundary to the PR-region

DLKKAVNE

SEQ ID NO: 96

Carboxy-terminal boundary to the PR-region

(K/G)TGW(K/G)QENGMW

CLAIMS

We claim:

1. A fusion protein comprising:
 - (a) a first polypeptide comprising an immunogenic polypeptide having the amino acid sequence of one of SEQ ID NOS: 1-8, or an immunogenic fragment thereof; and
 - (b) a second polypeptide comprising a *Streptococcus pneumoniae* antigen that is predominantly an antibody target.
2. The fusion protein of claim 1, wherein the first polypeptide is N-terminal to the second polypeptide.
3. The fusion protein of claim 2, wherein the first polypeptide is lipidated.
4. The fusion protein of any of the preceding claims, wherein the first polypeptide comprises the amino acid sequence of SEQ ID NO: 2.
5. The fusion protein of any of the preceding claims, wherein the first polypeptide comprises the amino acid sequence of SEQ ID NO: 6.
6. The fusion protein of any of the preceding claims, wherein the first polypeptide is directly bonded to the second polypeptide.
7. The fusion protein of any of the preceding claims, further comprising a linker polypeptide.
8. The fusion protein of claim 7, wherein the linker polypeptide is positioned between the first polypeptide and the second polypeptide.
9. The fusion protein of claim 8, wherein the linker polypeptide is rigid.
10. The fusion protein of claim 9, wherein the linker polypeptide comprises SEQ ID NO: 93.
11. The fusion protein of claim 8, wherein the linker polypeptide is semi-rigid.

12. The fusion protein of claim 11, wherein the linker polypeptide comprises SEQ ID NO: 94.
13. The fusion protein of any of claims 1-12, wherein the second polypeptide is Pneumococcal surface protein A (PspA) or a derivative thereof.
14. The fusion protein of claim 13, wherein the PspA derivative comprises all or a fragment of the proline-rich region of PspA.
15. The fusion protein of claim 14, wherein the PspA derivative includes a non-proline block.
16. The fusion protein of claim 15, wherein the PspA derivative comprises SEQ ID NO: 29.
17. The fusion protein of claim 15, wherein the PspA derivative comprises SEQ ID NO: 33.
18. The fusion protein of any of claims 1-12, wherein the second polypeptide is a pneumolysoid or a derivative thereof.
19. A vaccine formulation comprising:
 - the fusion protein of any of claims 1-18; and
 - a pharmaceutically acceptable carrier.
20. The vaccine formulation of claim 19, further comprising an immunogenic polypeptide comprising the amino acid sequence of one of SEQ ID NOS: 1-11, or an immunogenic fragment thereof.
21. The vaccine formulation of claim 19, further comprising a second fusion protein comprising two immunogenic polypeptides, each immunogenic polypeptide of the second fusion protein comprising the amino acid sequence of one of SEQ ID NOS: 1-11, or an immunogenic fragment thereof.
22. The vaccine formulation of any of claims 19-21, further comprising an adjuvant.

23. The vaccine formulation of claim 22, wherein the adjuvant is an agonist of toll-like receptors (TLRs).
24. The vaccine formulation of claim 22, wherein the adjuvant is alum.
25. The vaccine formulation of claim 22, wherein the vaccine formulation comprises 1-1000 µg of each fusion protein and 25-500 µg of the adjuvant.
26. The vaccine formulation of any of claims 19-25, characterized in that, upon contacting T_H17 cells, the vaccine formulation induces a T_H17 cell response at least 1.5-fold greater than that induced by a control.
27. The vaccine formulation of any of claims 19-26, characterized in that, upon administration to an uninfected subject, the vaccine formulation inhibits infection by *S. pneumoniae* in the uninfected subject.
28. The vaccine formulation of any of claims 19-27, characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae* colonization in the subject.
29. The vaccine formulation of any of claims 19-28, characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae* symptoms in the subject.
30. The vaccine formulation of any of claims 19-29, characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae*-induced invasive disease in the subject.
31. The vaccine formulation of any of claims 19-29, characterized in that, upon administration to a subject, the vaccine formulation inhibits *S. pneumoniae*-induced sepsis in the subject.

32. A method for treating a subject suffering from or susceptible to *S. pneumoniae* infection, comprising administering to the subject an effective amount of a vaccine formulation according to any of claims 19-31, thereby treating the subject.
33. The method of claim 32, wherein the subject is an uninfected subject, and the vaccine formulation inhibits infection by *S. pneumoniae* in the uninfected subject.
34. The method of claim 32, wherein the vaccine formulation inhibits *S. pneumoniae* colonization in the subject.
35. The method of claim 32, wherein the vaccine formulation inhibits one or more *S. pneumoniae* symptoms in the subject.
36. The method of claim 32, wherein the vaccine formulation inhibits *S. pneumoniae*-induced invasive disease in the subject.
37. The method of claim 32, wherein the vaccine formulation inhibits *S. pneumoniae*-induced sepsis in the subject.
38. The method of claim 32, wherein one dose of the vaccine formulation is administered.
39. The method of claim 32, wherein three doses of the vaccine formulation are administered.
40. The method of any one of claims 32-39, wherein the subject is a human.

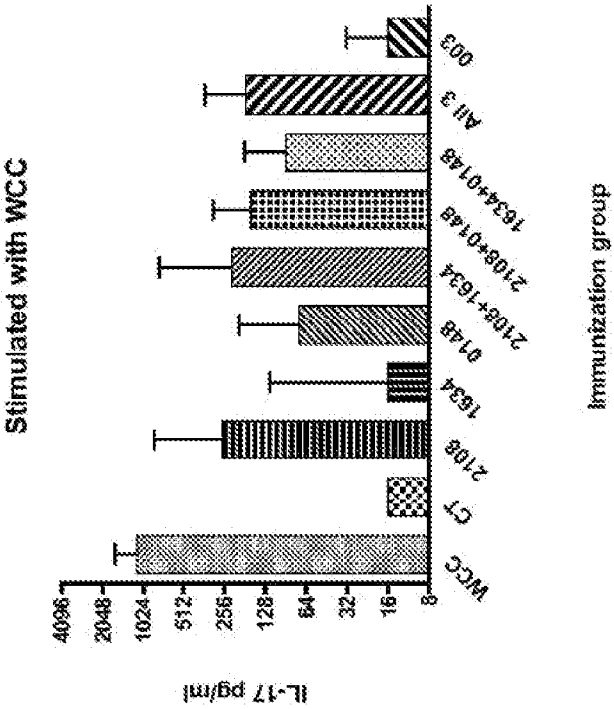
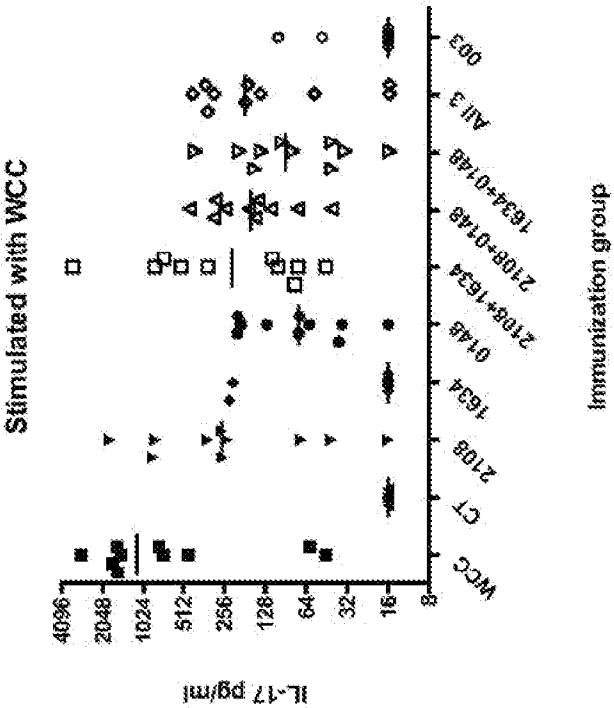


Figure 1



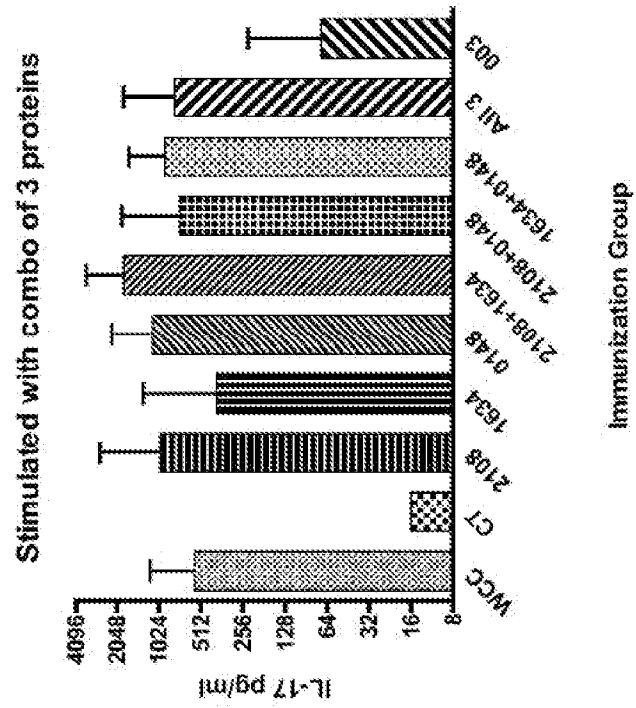
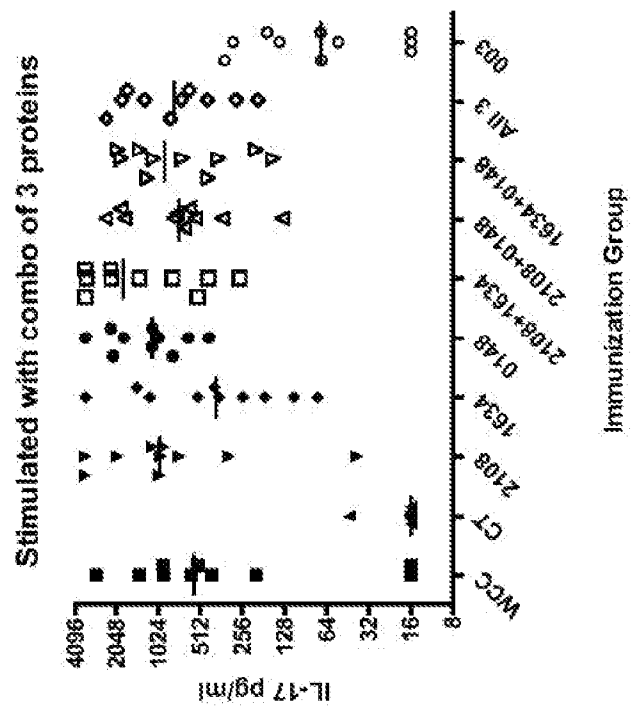


Figure 2



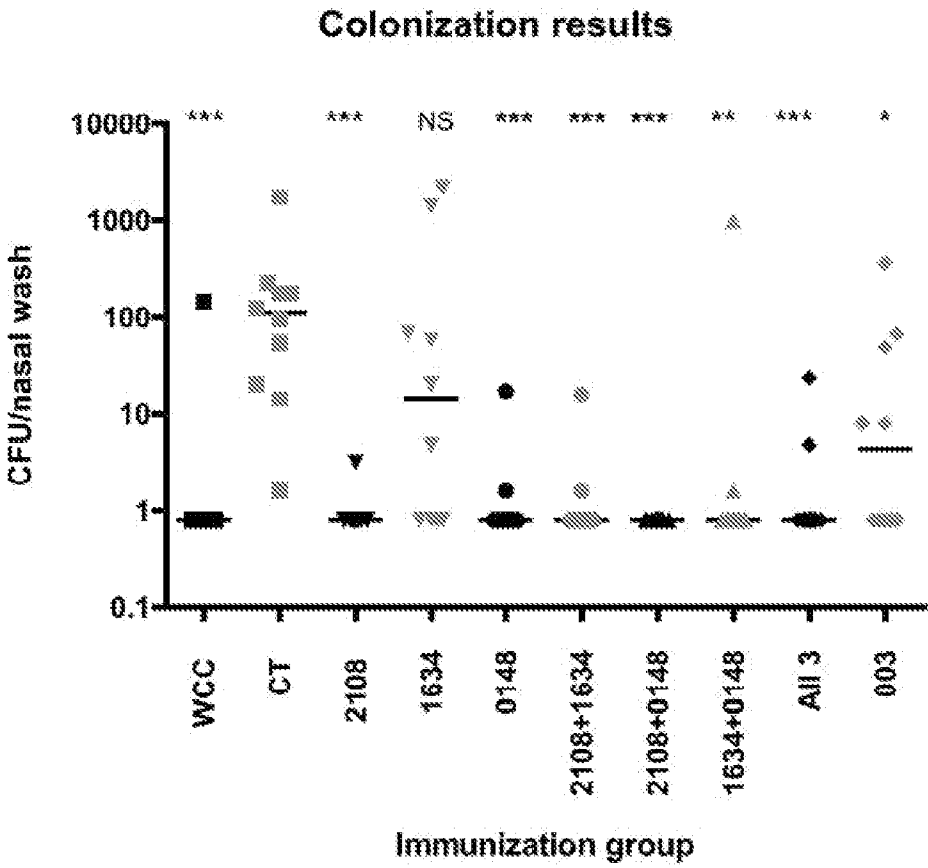


Figure 3

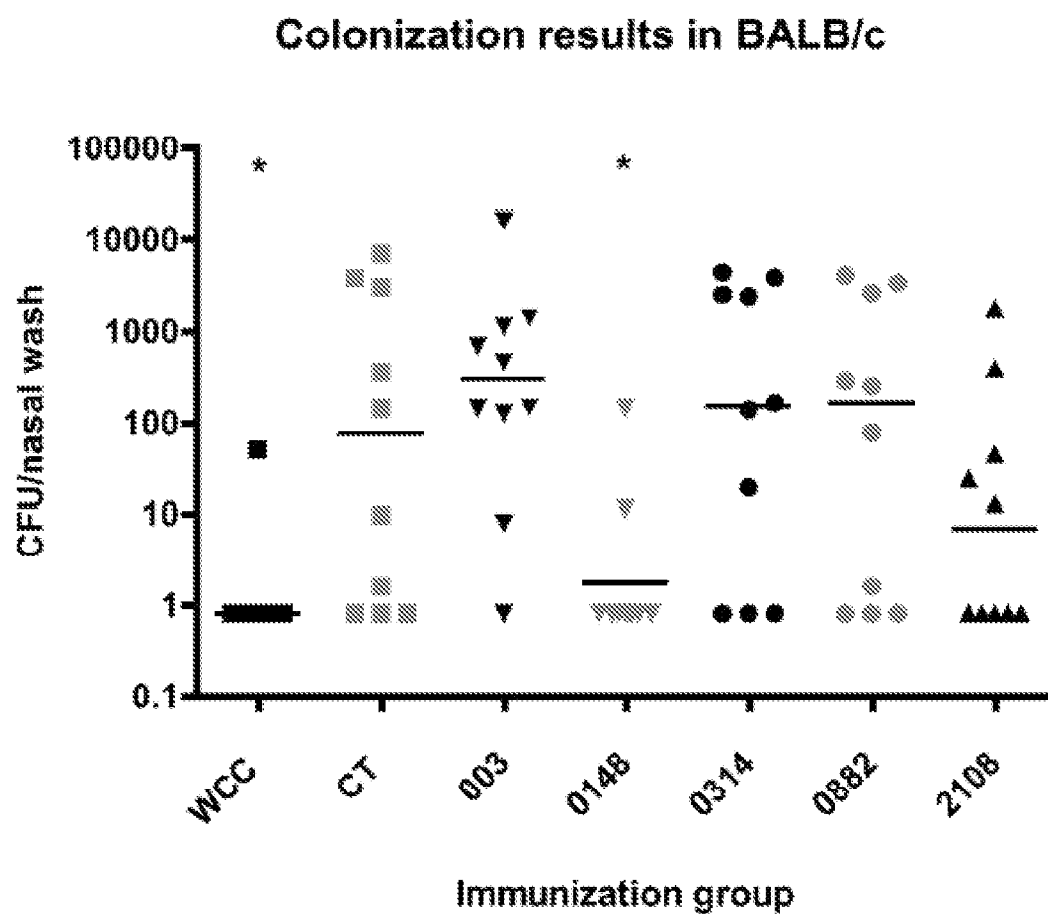


Figure 4

IL-17A Immunogenicity Assay

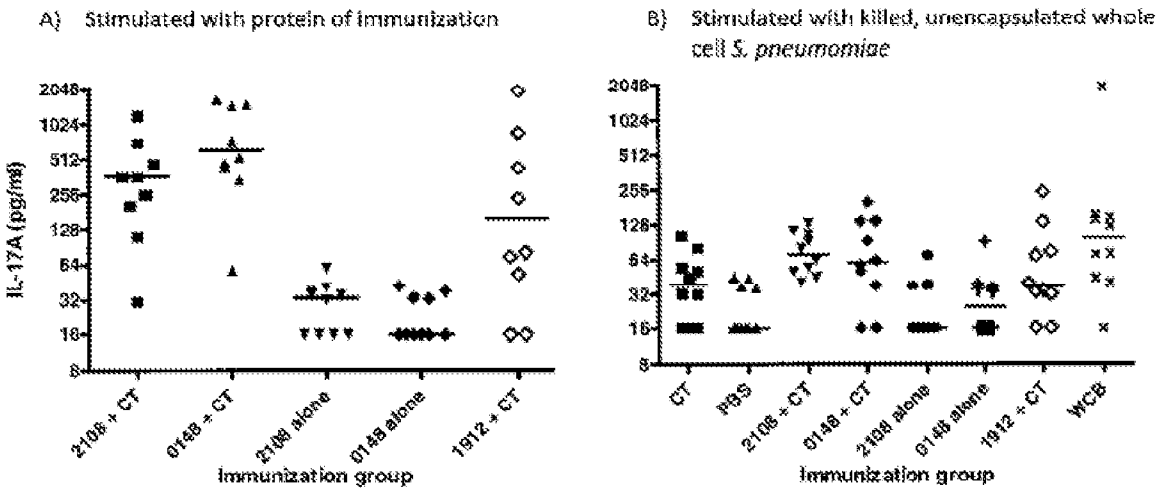


Figure 5

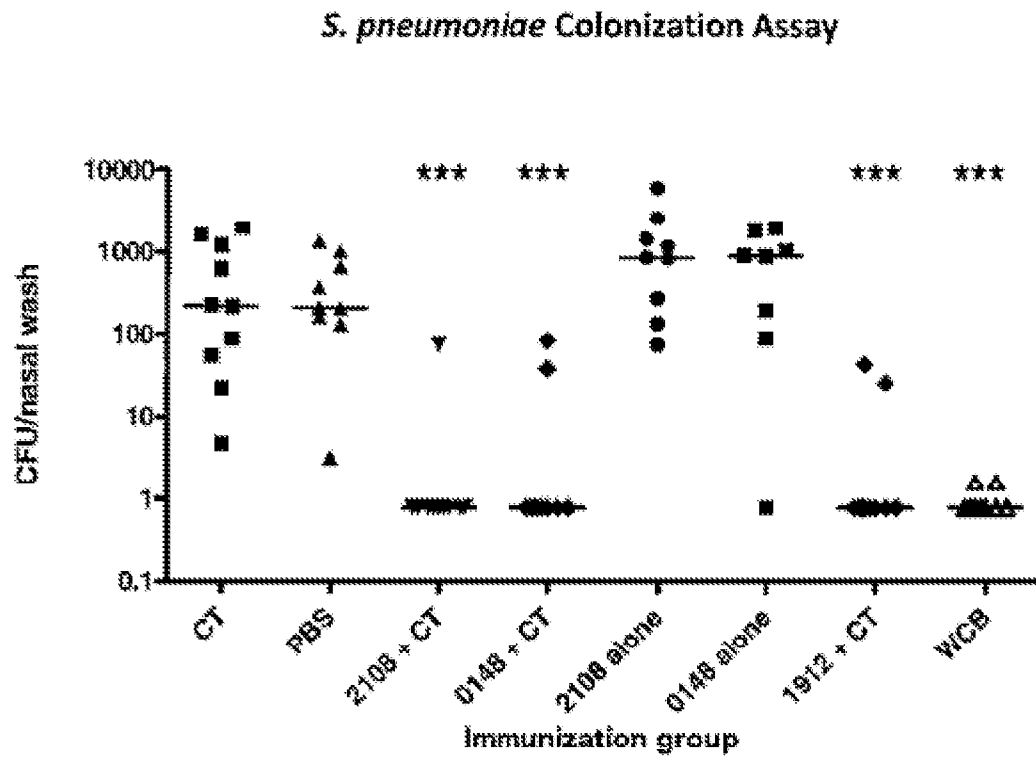


Figure 6

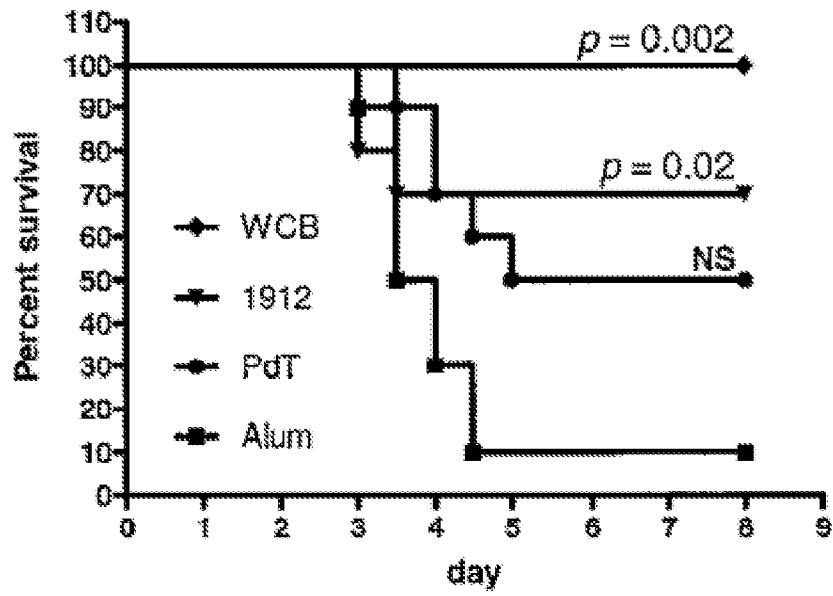
S. pneumoniae Aspiration Challenge

Figure 7

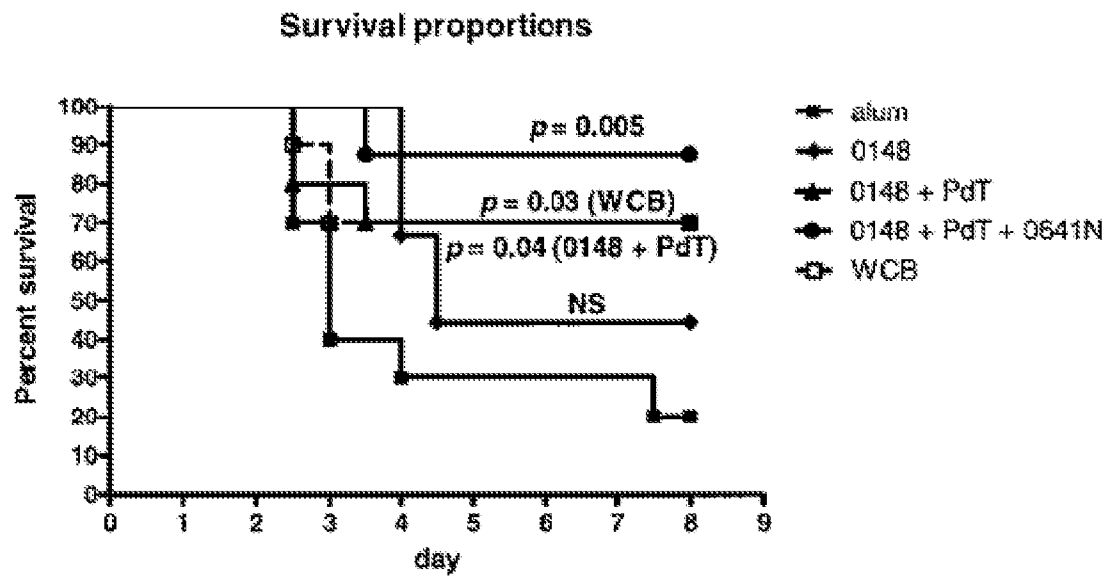
S. pneumoniae Aspiration Challenge

Figure 8

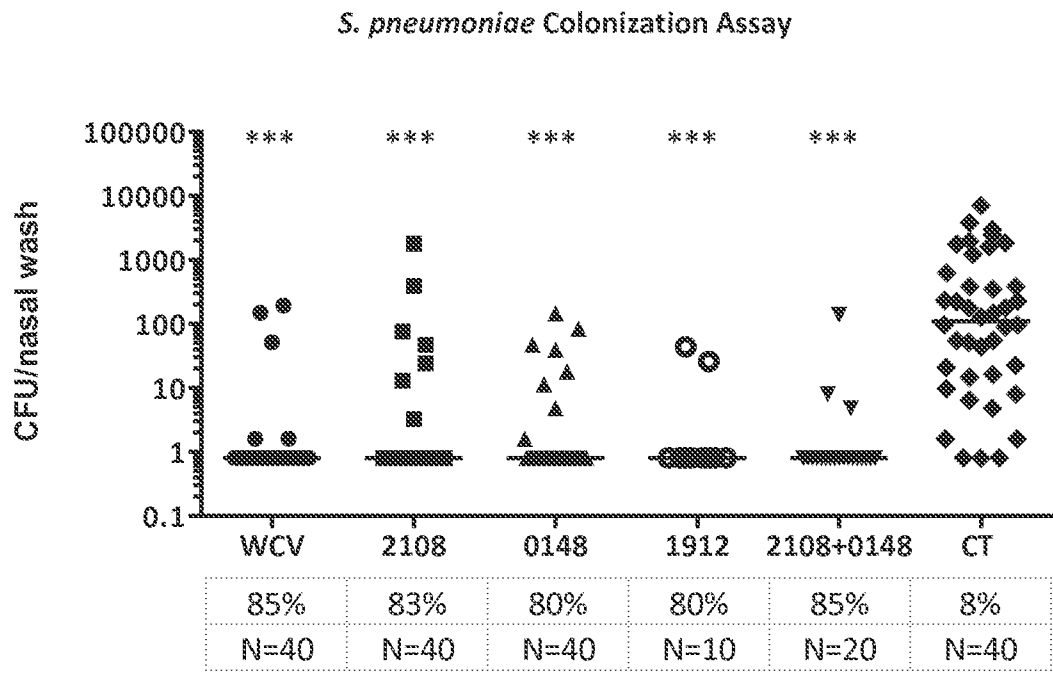


Figure 9

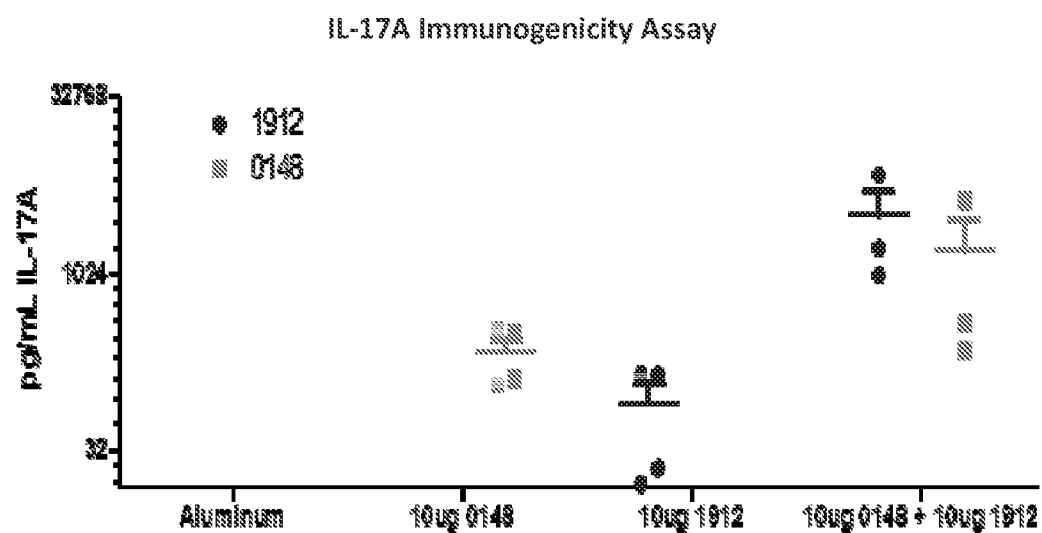


Figure 10

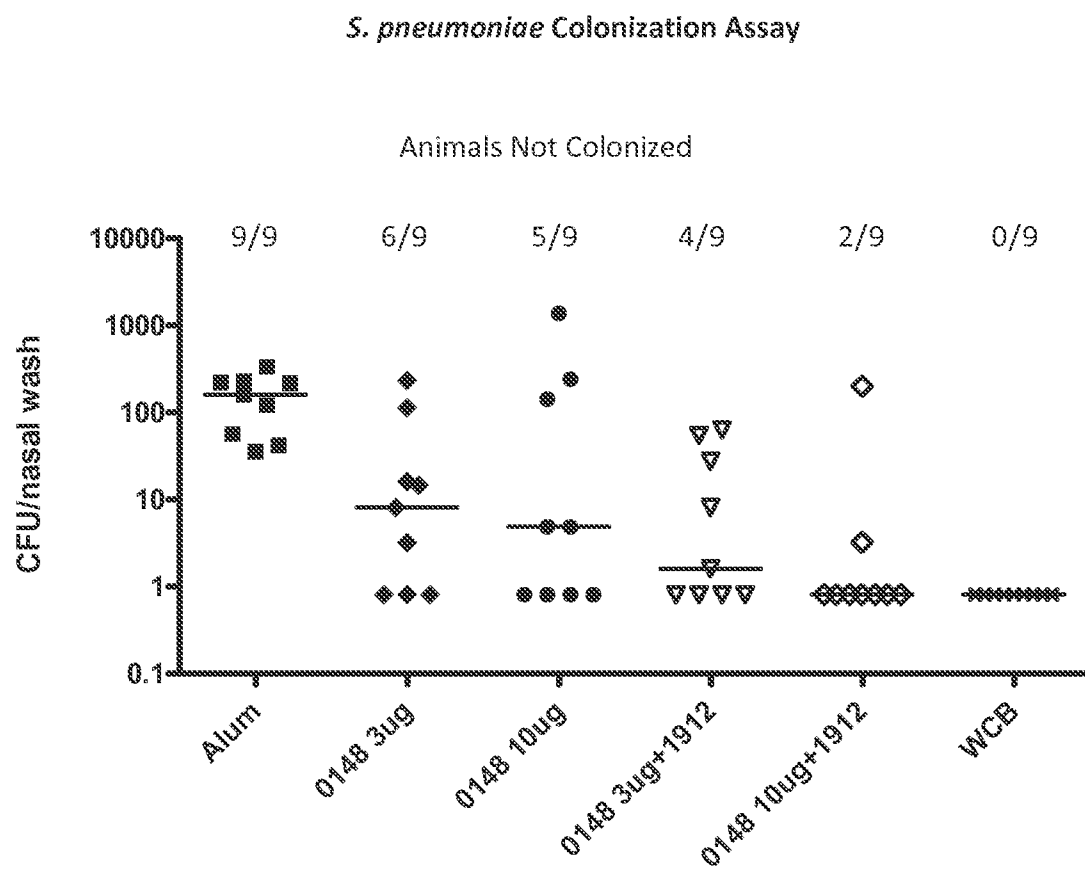


Figure 11

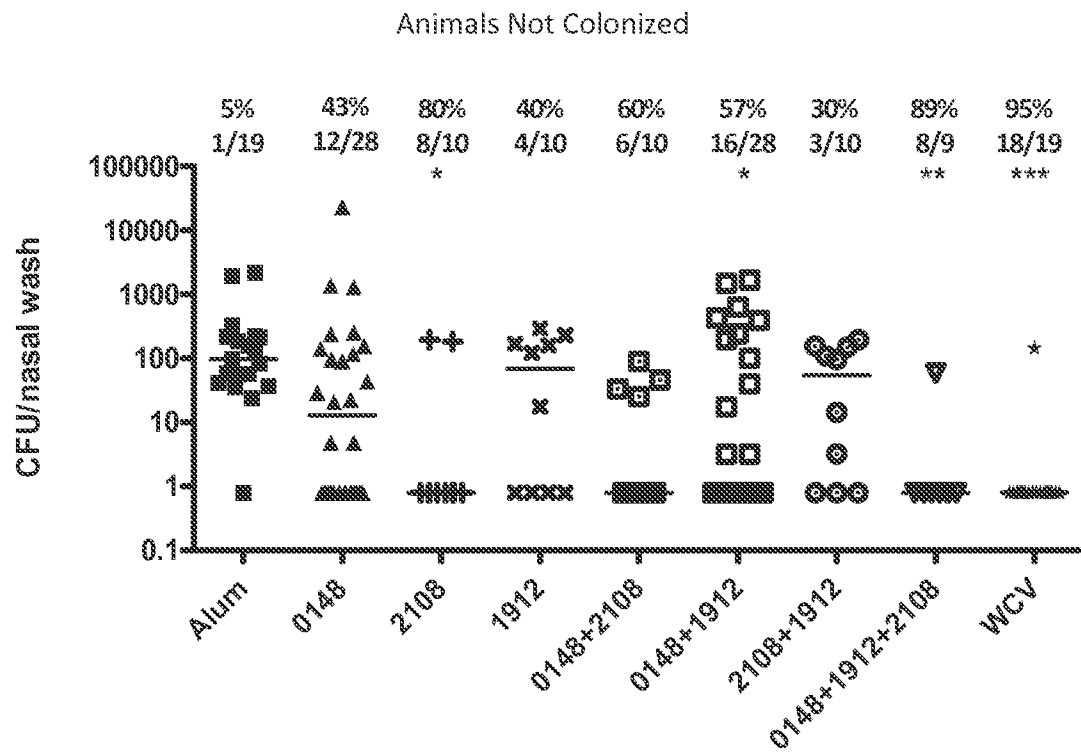
S. pneumoniae Colonization Assay: Pooled Data from 2 Vaccination Studies

Figure 12

S. pneumoniae Colonization Assay

Animals Not Colonized

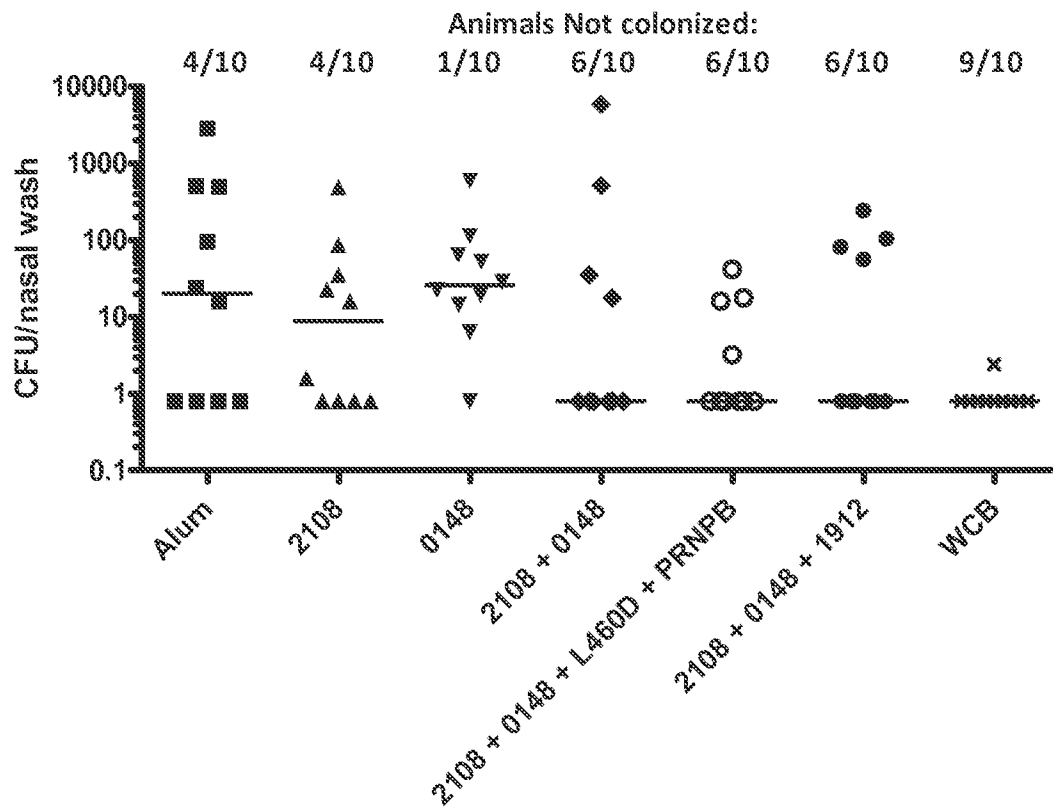


Figure 13

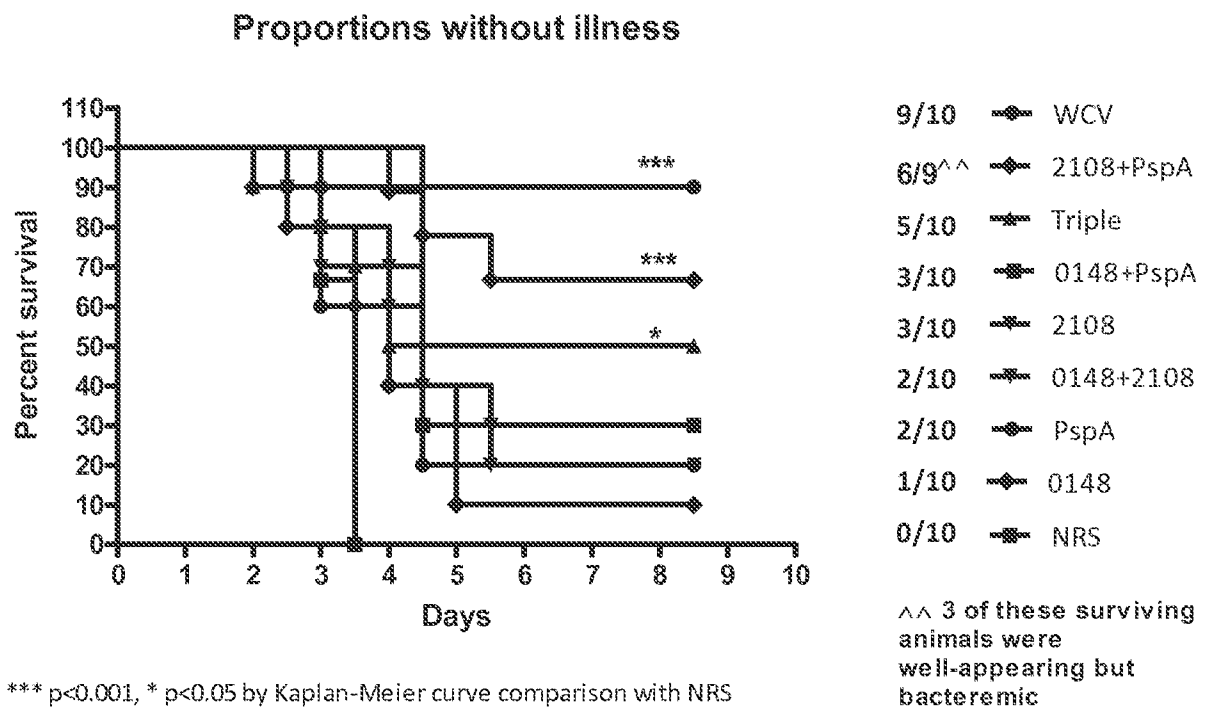
S. pneumoniae Aspiration Challenge: Passive Transfer Study

Figure 14

S. pneumoniae Aspiration Challenge: Summary of Studies

	Sepsis 1	Sepsis 2	Sepsis 3	Sepsis 4	PT1 ^o	PT2*
alum or NRS	20%	10%	30%	20%	0%	20%
0148	44%	10%			10%	
1912		70%		20%		
2108					30%	
PdT		50%		20%		
PspA					10%	70%
0148 + 1912		40%	40%			
0148 + 2108					20%	
0148 + PdT	70%	17%	25%			
0148 + PspA					30%	70%
1912 + PdT				40%		
2108 + PspA					67%	
0148 + PdT + 0641N	90%					
0148 + PdT + 1912		60%	60%			
0148 + 2108 + PspA					50%	70%
WCB	70%	100%	100%	100%	90%	90%

^o 10ug protein dose

*PspA - purified IgG

[^]PspA - ascites

Figure 15

IL-17A Immunogenicity Assay: Fusion Proteins

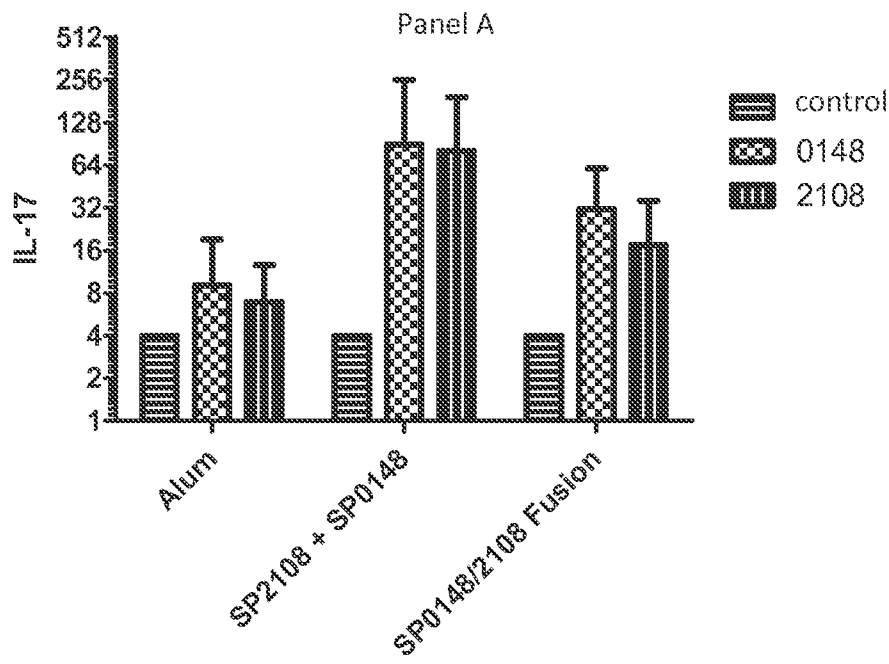


Figure 16A

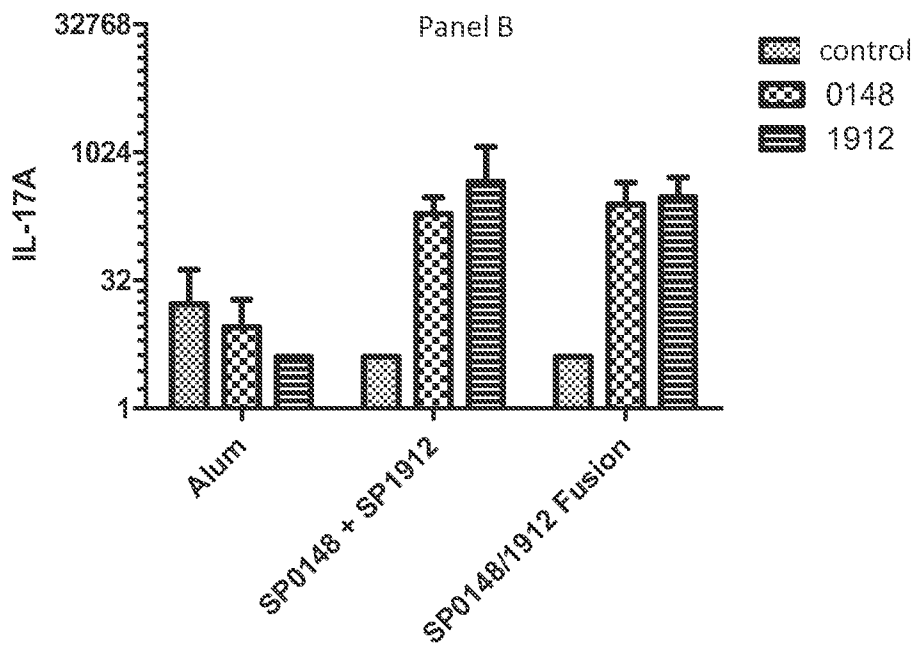


Figure 16B

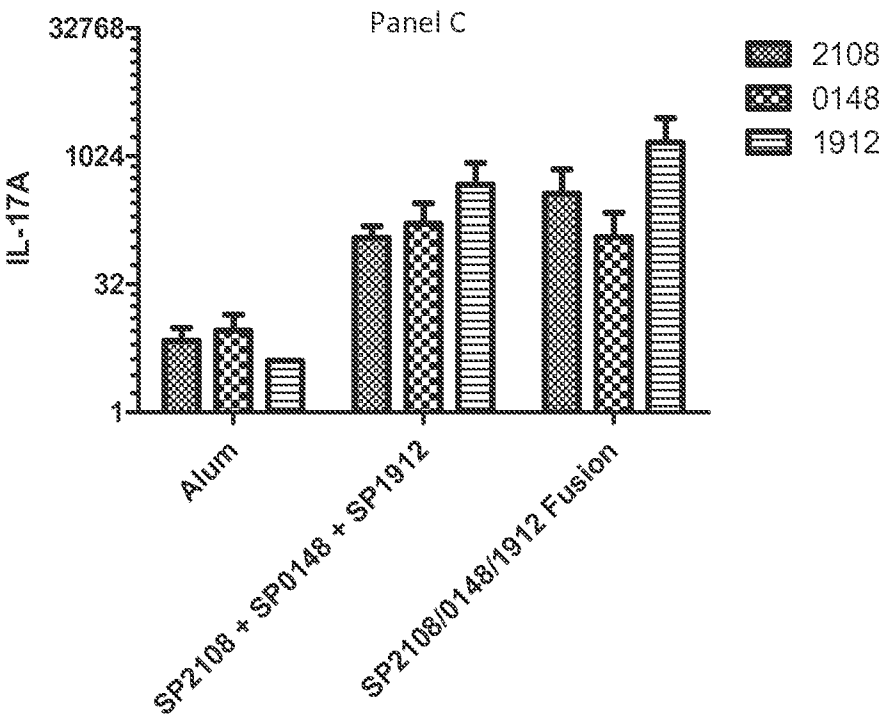


Figure 16C

Combined Intranasal Challenge (Carriage) Studies 1 and 2

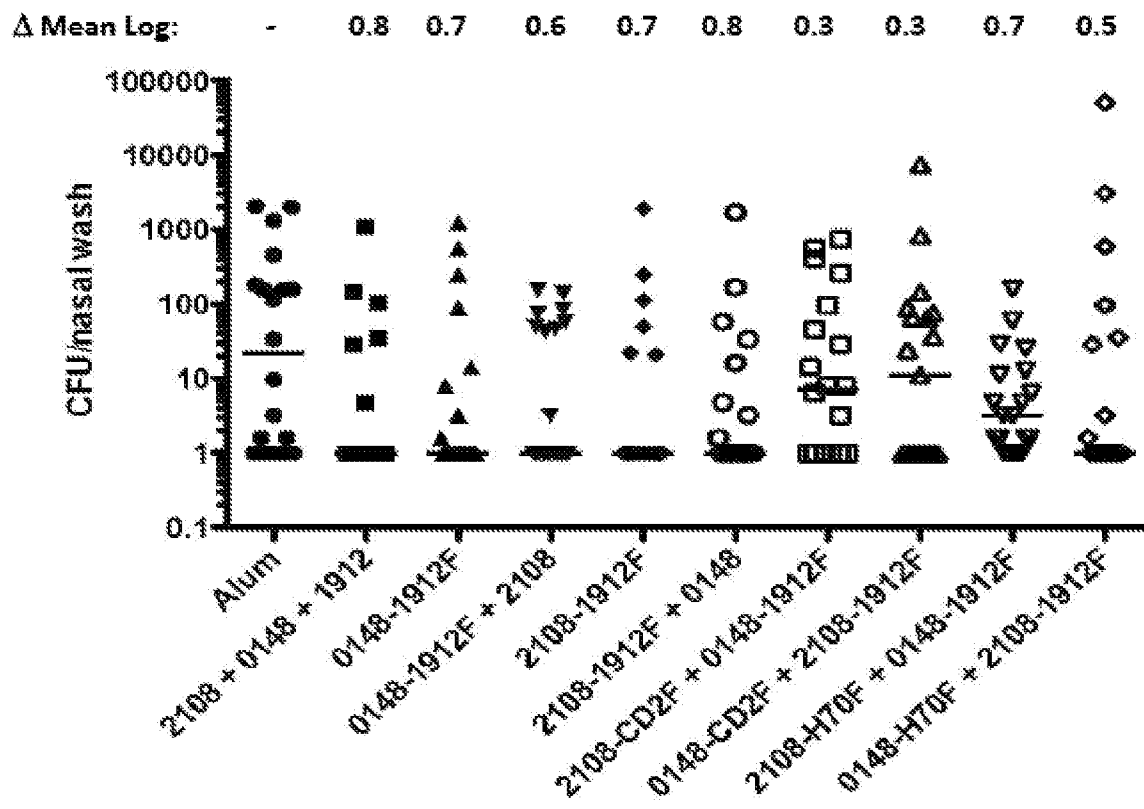
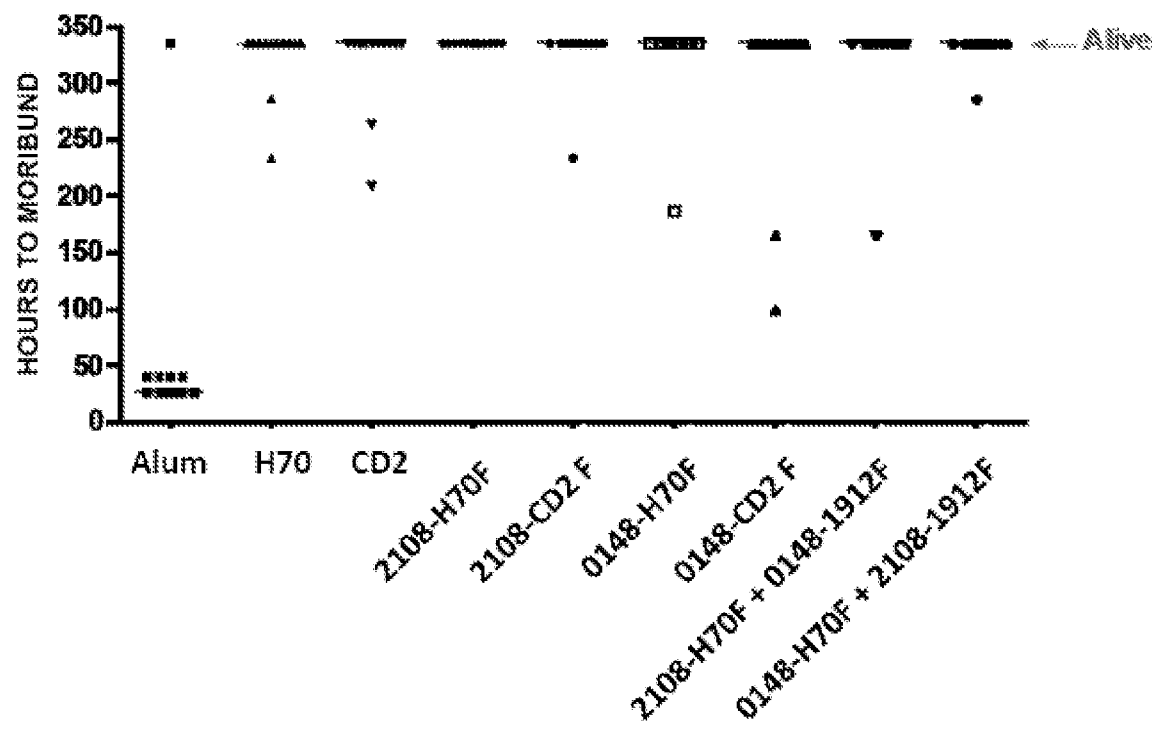


Figure 17

Intravenous Challenge (Sepsis) Study



	Alum	H70	CD2	2108-H70F	2108-CD2F	0148-H70F	0148-CD2F	2108-H70F + 0148-1912F	0148-H70F + 2108-1912F
Mean	56.2	323.4	319.4	336.0	327.4	323.5	302.0	321.7	331.9
Median	25.5	336.0	336.0	336.0	336.0	336.0	336.0	336.0	336.0
Std error	25.5	9.1	11.6	0	8.5	12.4	23.2	14.2	4.0

Figure 18

IL-17A Response to Fusion Proteins

Figure 19A

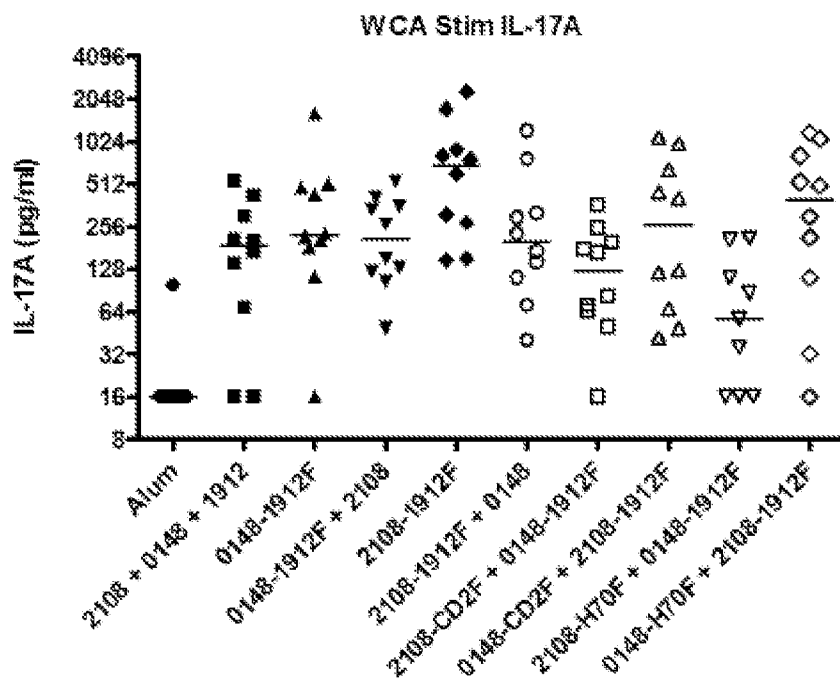
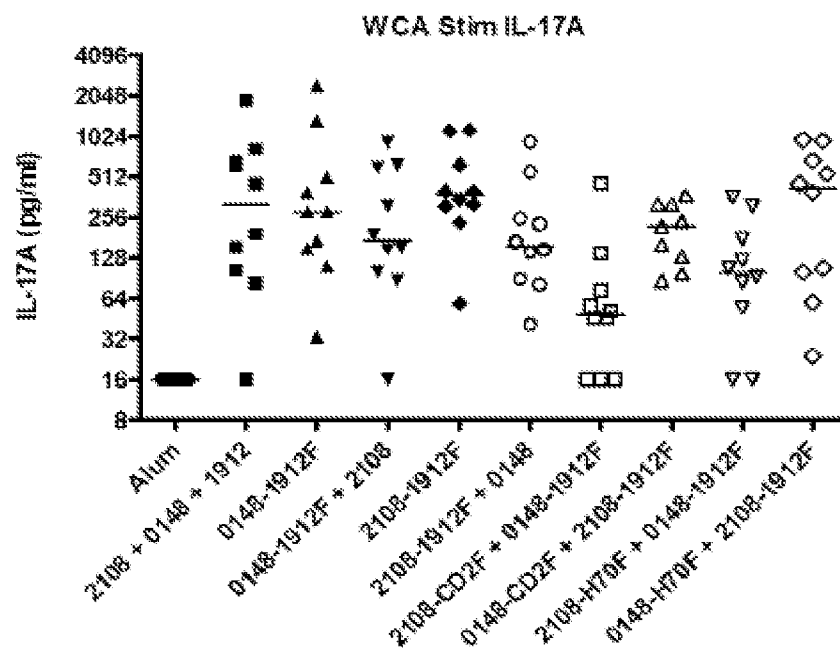


Figure 19B



Antibody Response to Fusion Proteins

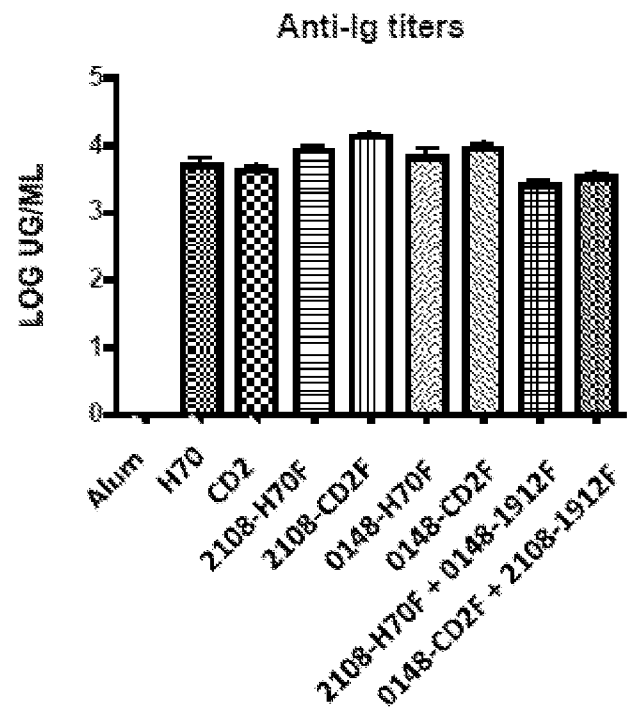


Figure 20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2013/052352

A. CLASSIFICATION OF SUBJECT MATTER

(see extra sheet)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K 19/00, 14/315, C12N 15/09, 15/62, A61K 39/09, A61P 31/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PAJ, Esp@cenet, PCT Online, USPTO DB, WIPO, RUPTO, EAPATIS, PatSearch

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2011/008548 A1 (GENOCEA BIOSCIENCES, INC. et al.) 20.01.2011, abstract, claims, SEQ ID NO: 6, SEQ ID NO: 10, pp. 1, 14, 26, 29, table 1, 2	1-9, 11, 13-16, 18-40
A		10, 12, 17
Y	WO 2008/127179 A1 (LUNDS UNIVERSITETS UTVECKLINGSAKTIEBOLAG et al.) 23.10.2008, abstract, claims, pp. 10, 22	1-9, 11, 13-16, 18-40
Y	EP 2278013 A1 (WYETH HOLDINGS CORPORATION) 26.01.2011, abstract, SEQ ID NO: 218	4
Y	WO 2004/092209 A2 (INTERCELL AG et al.) 28.10.2004, abstract, SEQ ID NO: 233	5

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

30 September 2013 (30.09.2013)

Date of mailing of the international search report

05 December 2013 (05.12.2013)

Name and mailing address of the ISA/ FIPS
Russia, 123995, Moscow, G-59, GSP-5,
Berezhkovskaya nab., 30-1

Authorized officer

E. Tereshkina

Facsimile No. +7 (499) 243-33-37

Telephone No. 8(495)531-65-15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2013/052352

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2010/080424 A1 (E.I. DU PONT DE NEMOURS AND COMPANY et al.) 15.07.2010, abstract, p. 30	9, 11
Y	WO 2007/089866 A2 (UNIVERSITY OF ALABAMA RESEARCH FOUNDATION et al.) 09.08.2007, abstract, SEQ ID NO: 1, p. 14, claim 2	13-16
Y	PATION James C. Purification and Immunogenicity of Genetically Obtained Pneumolysin Toxoids and Their Conjugation to Streptococcus pneumoniae Type 19F Polysaccharide. Infection and Immunity, 1991, Vol. 59, no. 7, p. 2297-2304, abstract	18

INTERNATIONAL SEARCH REPORT
Classification of subject matter

International application No.

PCT/US 2013/052352

C07K 19/00 (2006.01)
C07K 14/315 (2006.01)
C12N 15/09 (2006.01)
C12N 15/62 (2006.01)
A61K 39/09 (2006.01)
A61P 31/04 (2006.01)