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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
07 March 2019 (07.03.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/043245 A1

(51) International Patent Classification:
A61K 47/64 (2017.01) *C12N 15/63* (2006.01)

(21) International Application Number:
PCT/EP2018/073714

(22) International Filing Date:
04 September 2018 (04.09.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1714128.4 04 September 2017 (04.09.2017) GB
1807669.5 11 May 2018 (11.05.2018) GB

(71) Applicant: **LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE** [GB/GB]; Keppel Street, London WC1E 7HT (GB).

(72) Inventors: **KAY, Emily**; London School of Hygiene and Tropical Medicine Keppel Street, London Greater London WC1E 7HT (GB). **WREN, Brendan**; London School of Hygiene and Tropical Medicine Keppel Street, London Greater London WC1E 7HT (GB). **CUCCUI, Jon**; London School of Hygiene and Tropical Medicine Keppel Street, London Greater London WC1E 7HT (GB).

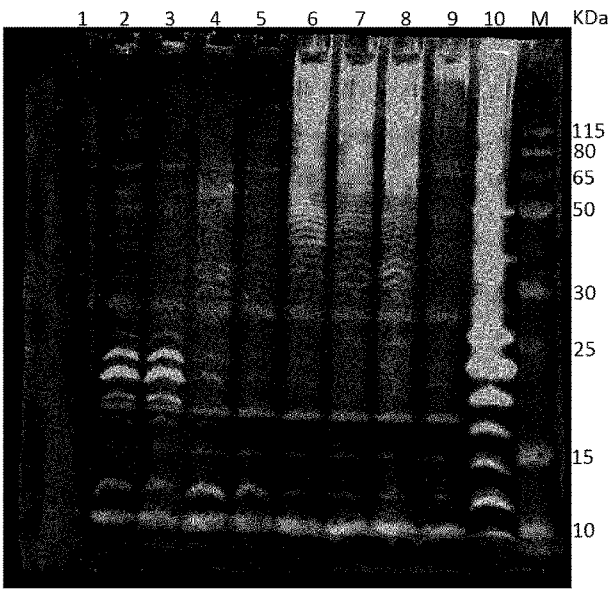
(74) Agent: **SYMBIOSIS IP LIMITED**; NAFIC Sand Hutton, York YO41 1LZ (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

(54) Title: MICROBIAL CELLS EXPRESSING STREPTOCOCCAL SEROTYPES

Figure 4



(57) Abstract: The disclosure relates to a glycoconjugate vaccine comprising one or more polysaccharides covalently linked to one or more carrier polypeptides wherein said glycoconjugate is manufactured by a modified microbial cell; vaccines and immunogenic compositions comprising said glycoconjugate(s); uses and methods of treatment of microbial infections using said vaccines and immunogenic compositions; bioreactors for the manufacture of said glycoconjugates and methods to prepare said glycoconjugate(s) from said modified microbial cells.

WO 2019/043245 A1

[Continued on next page]

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))*
— *with sequence listing part of description (Rule 5.2(a))*

MICROBIAL CELLS EXPRESSING STREPTOCOCCAL SEROTYPES

Field of the Disclosure

The disclosure relates to a glycoconjugate vaccine comprising one or more polysaccharides covalently linked to one or more carrier polypeptides wherein said glyconjugate is manufactured by a modified microbial cell; vaccines and immunogenic compositions comprising said glycoconjugate(s); uses and methods of treatment of microbial infections using said vaccines and immunogenic compositions; bioreactors for the manufacture of said glycoconjugates and methods to prepare said glycoconjugate(s) from said modified microbial cells.

The disclosure provides an efficient means to manufacture vaccines effective in protecting human subjects from microbial infection, particularly in relation to the manufacture of pneumococcal glycoconjugate vaccines, for example, vaccines protective against *Streptococcus pneumoniae*.

Background to the Disclosure

Many modern vaccines are made from protective antigens of the pathogen, isolated by molecular cloning and purified from the materials that give rise to side-effects. These vaccines are known as 'subunit vaccines'. The emergence of new pathogens and the growth of antibiotic resistance have created a need to develop new vaccines and to identify further candidate molecules useful in the development of subunit vaccines. However, although subunit vaccines tend to avoid the side effects of killed or attenuated pathogen vaccines, their 'pure' status means that subunit vaccines do not always have adequate immunogenicity to confer protection. This is typically addressed by the use of carriers and adjuvants to augment immune responses to pathogen antigens.

A "carrier" is an immunogenic molecule which, when bound to a second molecule that augments immune responses to the latter. Some antigens are not intrinsically immunogenic, for example glycans, yet may be capable of generating antibody responses when associated with a foreign protein molecule such as keyhole-limpet haemocyanin or tetanus toxoid. Such antigens contain B-cell epitopes but no T-cell epitopes. The protein moiety of such a conjugate (the "carrier" protein) provides T-cell epitopes which stimulate helper T-cells that in turn stimulate antigen-specific B-cells to differentiate into plasma cells and produce antibody against the antigen. Examples of protein carriers used in vaccines

are Protein D from *Haemophilus influenzae*, CRM197 a non-toxic recombinant variant of the diphtheria toxin of *Corynebacterium diphtheria*, or the tetanus toxoid of *Clostridium tetani.*

Streptococcus pneumoniae, or pneumococcus, is an important human pathogen causing a variety of invasive illnesses such as pneumonia, meningitis and bacteraemia. Serious pneumococcal infections are most common in infants, toddlers, the elderly, and subjects with deficient immune systems such as HIV patients or other underlying conditions. *S. pneumoniae* is the causative agent for pneumonia related deaths in developing countries. *S. pneumonia* causes also less severe infections such as sinusitis and otitis media. These highly prevalent infections have significant impact on healthcare costs, even in developed countries. Presently there are three glycoconjugate pneumococcal vaccines each with a different combination of serotypes. Prevnar® comprises seven serotypes: 4, 6B, 9V, 14, 18C, 19F and 23F, and is suitable for immunisation of infants less than 2 years old and is crosslinked to a diphtheria toxin carrier protein. An alternative thirteen component vaccine also linked to diphtheria toxin comprises serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F and is suitable for human subjects over 65 years old. A further ten serotype pneumococcal vaccine comprises the following serotypes; 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F and 23F. These vaccines are effective in protection against pneumococcal infection and are manufactured by chemical conjugation which is expensive.

There is a need for alternative pneumococcal vaccines which are effective and less expensive to manufacture.

We disclose vaccine and immunogenic compositions comprising carrier polypeptides conjugated with bacterial polysaccharides wherein said bacterial polysaccharides (serotypes) are conjugated to a carrier polypeptide in a microbial based cell expression system adapted to express said carrier polypeptide, one bacterial polysaccharide and an oligosaccharyltransferase. We also disclose bacterial cells that are genetically modified to facilitate the production of said polysaccharide conjugate vaccines that include modifications to the microbial genome which result in the synthesis of full length native or near native bacterial capsular polysaccharides that are immunogenic. Further pneumococcal vaccines comprising serotype 8 are disclosed.

Statements of Invention

According to an aspect of the invention there is provided a microbial cell wherein said cell comprises:

i) a nucleic acid molecule comprising a nucleotide sequence encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides not expressed by said microbial cell;

ii) a nucleic acid molecule comprising a nucleotide sequence encoding at least one heterologous polysaccharide chain length regulator;

iii) a nucleic acid molecule comprising a nucleotide sequence encoding one or more carrier polypeptides that comprise one or more glycosylation motifs; and

iv) a nucleic acid molecule comprising a nucleotide sequence encoding an oligosaccharyltransferase.

In an alternative embodiment of the invention the genome of said microbial cell is modified in one or more endogenous genes encoding a polysaccharide chain length regulator characterised in that said endogenous bacterial capsular polysaccharide chain length regulator is modified or inactivated and substantially lacks expression and/or activity.

Means to mutate and/or inactivate microbial genes are known in the art and include the use of mutagens, homologous recombination and transposon integration to disrupt gene function. More recently CRISPR/Cas 9 has enabled the targeted modification of genomic DNA to introduce permanent genetic modifications with high precision.

In a preferred embodiment of the invention said microbial cell is a bacterial cell, for example an *Escherichia coli* cell.

In a preferred embodiment of the invention said bacterial cell is a lactic acid bacteria of the genus Lactococcus or Lactobacillus.

In an alternative embodiment of the invention said microbial cell is a fungal, algal or yeast cell.

In a preferred embodiment of the invention said nucleic acid encoding polypeptides involved in the synthesis of heterologous bacterial capsular polysaccharide is isolated from a Gram-negative or Gram-positive bacterial species.

In a preferred embodiment of the invention said Gram negative bacterial species is selected from the genus: *Nesseria spp, Klebsiella spp, Haemophilus spp, Pseudomonas spp and Salmonella spp.*

In an alternative preferred embodiment of the invention said Gram positive bacterial species is selected from the genus: *Bacillus spp, Lactococcus spp, Lactobacillus spp, Streptococcus spp and Staphylococcus spp.*

In a preferred embodiment of the invention said *Streptococcus* species is *Streptococcus pneumoniae.*

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides synthesizes a *Streptococcus pneumoniae* polysaccharide serotype selected from one of the 96 serotypes expressed by *Streptococcus pneumoniae.*

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of a heterologous bacterial capsular polysaccharides synthesizes a *Streptococcus pneumoniae* polysaccharide serotype wherein said polysaccharide comprises an acetylated reducing end sugar.

In a further preferred embodiment of the invention said acetylated reducing end sugar comprises a 2-acetamido group.

In an alternative preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of a heterologous bacterial capsular polysaccharides synthesize a *Streptococcus pneumoniae* polysaccharide serotype wherein said polysaccharide comprises glucose as reducing end sugar.

In an alternative preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of a heterologous bacterial capsular polysaccharides synthesize a Streptococcus pneumoniae polysaccharide

serotype wherein said polysaccharide comprises galactose or a galactose derivative as reducing end sugar.

Preferably, said polysaccharide comprising an acetylated reducing end sugar is selected from the group of serotypes consisting of: 1, 4, 5, 12F,12A, 12B, 25F, 25A, 38, 44, 45 or 46.

In a further preferred embodiment said polysaccharide comprising glucose as reducing end sugar is serotype 8.

In a further preferred embodiment said polysaccharide comprising glucose as reducing end sugar either structure confirmed or suggested by the presence of wchA: is 2, 6A, 6B, 7F, 7A, 7B, 7C, 9A, 9L, 9N, 9V, 11F, 11A, 11B, 11C, 11D, 13, 14, 15F, 15A, 15B, 15C, 16F, 16A, 17F, 17A, 18F, 18A, 18B, 18C, 19F, 19A, 19B, 19C, 20, 21, 22F, 22A, 23F, 23A, 23B, 24F, 24A, 24B, 27, 28F, 28A, 32F, 32A, 33F, 33A, 33B, 33D, 34, 35A, 35B, 35C, 36, 40, 41F, 41A, 42, 48.

In an alternative preferred embodiment said polysaccharide comprises galactose or a galactose derivative as the reducing end sugar, either structurally confirmed or suggested by the presence of wcjG or wcjH: 10F, 10A, 10B, 10C, 29, 31, 33C, 35F, 39, 43, 47F, 47A.

Preferably, said serotypes are selected from a group consisting of: 1, 4, 5, 12F, 12A, 12B, 25F, 25A, 38, 44, 45, 46 or 8.

The operons encoding genes involved in the synthesis of capsular polysaccharides are known in the art, for example see Table 1 and Table 5 and accession numbers therein. For example, Bentley et al 2006 sequenced each of 90 *Streptococcus pneumoniae* serotypes.

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides synthesizes a native *Streptococcus pneumoniae* serotype 4 polysaccharide.

In a preferred embodiment of the invention said nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID NO: 1.

In a preferred embodiment of the invention said microbial cell expresses one or more AATGal synthesis genes encoded by SEQ ID NO: 69 and SEQ ID NO: 70

In a preferred embodiment of the invention said nucleic acid molecule comprises a nucleotide sequence as set forth in SEQ ID NO: 57.

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides synthesizes a *Streptococcus pneumoniae* polysaccharide serotype selected from the group consisting of: 1 or 4 or 5.

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of a heterologous bacterial capsular polysaccharides synthesizes a Streptococcus pneumoniae polysaccharide serotype 8 or 1 or 4 or 5.

Preferably, said microbial cell expresses the serotype capsular polysaccharides 1, 4 and 5.

In a further preferred embodiment, said microbial cell expresses the serotype capsular polysaccharide 8, or serotype 1, 4, 5 and 8.

In a preferred embodiment of the invention said microbial cell expresses a serotype capsular polysaccharide selected from the group consisting of: 4, 6B, 9V, 14, 18C, 19F and 23F.

In a preferred embodiment of the invention said microbial cell expresses a serotype capsular polysaccharide selected from the group consisting of: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F.

In a preferred embodiment of the invention said microbial cell expresses serotype 1.

Preferably, said microbial cell expresses serotype 1 encoded by the nucleotide sequence set forth in SEQ ID NO:58.

In a preferred embodiment of the invention said microbial cell expresses serotype 3.

Methods for making immunogenic conjugates comprising Streptococcus pneumoniae serotype 3 are known in the art and disclosed in patent application US2010316666 which is hereby incorporated by reference in its entirety.

In a preferred embodiment of the invention said microbial cell expresses serotype 4.

Preferably, said microbial cell expresses serotype 4 encoded by the nucleotide sequence set forth in SEQ ID NO: 1.

In a preferred embodiment of the invention said microbial cell expresses serotype 5

Preferably, said microbial cell expresses serotype 5 encoded by the nucleotide sequence set forth in SEQ ID NO: 59.

In a preferred embodiment of the invention said microbial cell expresses serotype 6A.

Preferably, said microbial cell expresses serotype 6A encoded by the nucleotide sequence set forth in SEQ ID NO: 74.

In a preferred embodiment of the invention said microbial cell expresses serotype 6B.

Preferably, said microbial cell expresses serotype 6A encoded by the nucleotide sequence set forth in SEQ ID NO: 75.

In a preferred embodiment of the invention said microbial cell expresses serotype 7F.

Preferably, said microbial cell expresses serotype 7F encoded by the nucleotide sequence set forth in SEQ ID NO: 76.

In a preferred embodiment of the invention said microbial cell expresses serotype 9V.

Preferably, said microbial cell expresses serotype 9V encoded by the nucleotide sequence set forth in SEQ ID NO:77.

In a preferred embodiment of the invention said microbial cell expresses serotype 14.

Preferably, said microbial cell expresses serotype 14 encoded by the nucleotide sequence set forth in SEQ ID NO: 78.

In a preferred embodiment of the invention said microbial cell expresses serotype 18C.

Preferably, said microbial cell expresses serotype 18C encoded by the nucleotide sequence set forth in SEQ ID NO: 79.

In a preferred embodiment of the invention said microbial cell expresses serotype 19A.

Preferably, said microbial cell expresses serotype 19A encoded by the nucleotide sequence set forth in SEQ ID NO: 80.

In a preferred embodiment of the invention said microbial cell expresses serotype 19F.

Preferably, said microbial cell expresses serotype 19F encoded by the nucleotide sequence set forth in SEQ ID NO: 81.

In a preferred embodiment of the invention said microbial cell expresses the seroptype 23F.

Preferably, said microbial cell expresses serotype 23F encoded by the nucleotide sequence set forth in SEQ ID NO: 82.

In a preferred embodiment of the invention said microbial cell expresses one or more ribitol genes encoded by the nucleotide sequence set forth in SEQ ID NO: 73.

The nucleic acid molecules comprising nucleotide sequences encoding the serotypes described herein can vary due to e.g. naturally occurring sequence polymorphisms and can comprise a nucleotide sequence which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical over the full-length nucleotide sequence of the sequences set forth in SEQ ID NOs 1, 57-68 and 73-82.

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides is provided on an episomal plasmid.

In a preferred embodiment of the invention said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides is stably integrated into the genome of said microbial cell.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is encoded by the wzd gene and wze gene.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 3 and SEQ ID NO: 15;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 3 and SEQ ID NO: 15 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 2 and SEQ ID NO: 14;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 2 and SEQ ID NO: 14 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 4 and SEQ ID NO: 16;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 4 and SEQ ID NO: 16 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 5 and SEQ ID NO: 17;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 5 and SEQ ID NO: 17 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 6 and SEQ ID NO: 18;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 6 and SEQ ID NO: 18 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 7 and SEQ ID NO: 19;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 7 and SEQ ID NO: 19 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 8 and SEQ ID NO: 20;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 8 and SEQ ID NO: 20 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 9 and SEQ ID NO: 21;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 9 and SEQ ID NO: 21 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 10 and SEQ ID NO: 22;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 10 and SEQ ID NO: 22 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 11 and SEQ ID NO: 23;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 11 and SEQ ID NO: 23 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 12 and SEQ ID NO: 24;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 12 and SEQ ID NO: 24 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said nucleic acid encoding a heterologous polysaccharide chain length regulator is selected from the group consisting of:

i) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 13 and SEQ ID NO: 25;

ii) a nucleic acid comprising a nucleotide sequence as set forth in SEQ ID NO: 13 and SEQ ID NO: 25 that has at least 70% nucleotide sequence identity over the full nucleotide sequence and encodes a polypeptide with polysaccharide chain length regulator activity.

In a preferred embodiment of the invention said oligosaccharyltransferase is a *Campylobacter* oligosaccharyltransferase.

In a preferred embodiment of the invention said *Campylobacter* oligosaccharyltransferase is a *Campylobacter jejuni* oligosaccharyltransferase.

In an alternative embodiment of the invention said *Campylobacter* oligosaccharyltransferase is a *Campylobacter sputorum* oligosaccharyltransferase.

In a preferred embodiment of the invention said oligosaccharyltransferase is encoded by a nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID NO: 26, or a nucleotide sequence that has at least 50% nucleotide sequence identity over the full length of the nucleotide sequence set forth in SEQ ID NO: 26.

In a preferred embodiment of the invention said oligosaccharyltransferase is encoded by a nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID NO: 27 or 28, or a nucleotide sequence that has at least 50% nucleotide sequence identity over the full length of the nucleotide sequence set forth in SEQ ID NO: 27 or SEQ ID NO: 28.

In a preferred embodiment of the invention said oligosaccharyltransferase is represented by the amino acid sequence set forth in SEQ ID NO: 29, or an amino acid sequence that is at least 50% identical over the full length of the amino acid sequence set forth in SEQ ID NO: 29.

In a preferred embodiment of the invention said oligosaccharyltransferase is represented by the amino acid sequence set forth in SEQ ID NO: 30, or an amino acid sequence that is at least 50% identical over the full length of the amino acid sequence set forth in SEQ ID NO: 30.

In a preferred embodiment, said oligosaccharyltransferase has at least 55% identity, more preferably at least 60% identity, even more preferably at least 65% identity, still more preferably at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% identity, and most preferably at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity with the full length amino or nucleotide acid sequences as set forth in SEQ ID NOs 26 to 30.

In an alternative embodiment of the invention said oligosaccharyltransferase is encoded by the PglL gene, for example encoded by the PglL gene of Actinobacteria and encoded by the nucleotide sequence set forth in SEQ ID NO: 72.

In a preferred embodiment of the invention said microbial cell is adapted to express a glycosyltransferase encoded by a nucleotide sequence as set forth in SEQ ID NO: 71 (NGT).

In a preferred embodiment, said oligosaccharyltransferase has at least 55% identity, more preferably at least 60% identity, even more preferably at least 65% identity, still more preferably at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% identity, and most preferably at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity with the full length nucleic acid sequences as set forth in SEQ ID NOs 71 or 72.

In a preferred embodiment of the invention said nucleic acid molecule encoding said oligosaccharyltransferase or glycosyltransferase is stably integrated into the genome of said microbial or bacterial cell.

In a preferred embodiment of the invention said one or more carrier polypeptide(s) includes the amino acid motif: Asn-X-Ser or Asn-X-Thr where X is any amino acid except proline.

In an alternative embodiment of the invention said one or more carrier polypeptide(s) includes the amino acid motif: D/E-X-N-X-S/T, wherein X is any amino acid except proline.

In an alternative preferred embodiment of the invention said one or more carrier polypeptide(s) includes the amino acid motif D/E-X-N-X-S/T, wherein X is any amino acid except proline and is selected from the group consisting of: DVNVT, EVNAT, DQNAT,

DNNNT, DNNNS, DQNRT, ENNFT, DSNST, DQNIS, DQNVS, DNNVS, DYNVS, DFNVS, DFNAS, DFNSS, DVNAT, DFNVT or DVNAS.

In a preferred embodiment of the invention the nucleic acid encoding said carrier polypeptide(s) is stably integrated into the microbial genome.

In a preferred embodiment of the invention said endogenous bacterial serotype polysaccharide chain length regulator is modified and encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

i) a nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID NO: 31;

ii) a nucleic acid molecule comprising a nucleotide sequence wherein said sequence is degenerate as a result of the genetic code to the nucleotide sequence defined in (i); and

i) a nucleic acid molecule the complementary strand of which hybridizes under stringent hybridization conditions to the nucleotide sequence in i) and ii) above wherein said nucleic acid molecule encodes a bacterial serotype polysaccharide chain length regulator and is at least 75% identical over the full-length nucleotide sequence set forth in SEQ ID NO: 31,

wherein said nucleic acid molecule is mutated or deleted for all or part of the nucleotide sequence encoding said endogenous bacterial serotype polysaccharide chain length regulator and substantially lacks expression and/or encodes a polypeptide that substantially lacks enzyme activity.

Hybridization of a nucleic acid molecule occurs when two complementary nucleic acid molecules undergo an amount of hydrogen bonding to each other. The stringency of hybridization can vary according to the environmental conditions surrounding the nucleic acids, the nature of the hybridization method, and the composition and length of the nucleic acid molecules used. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed in Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001); and Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes Part I, Chapter 2 (Elsevier, New York, 1993). The T_m is the temperature at which 50% of a given strand of a nucleic acid molecule is hybridized to its complementary strand. The following is an exemplary set of hybridization conditions and is not limiting:

Very High Stringency (allows sequences that share at least 90% identity to hybridize)

Hybridization:	5x SSC at 65°C for 16 hours
Wash twice:	2x SSC at room temperature (RT) for 15 minutes each
Wash twice:	0.5x SSC at 65°C for 20 minutes each

High Stringency (allows sequences that share at least 80% identity to hybridize)

Hybridization:	5x-6x SSC at 65°C-70°C for 16-20 hours
Wash twice:	2x SSC at RT for 5-20 minutes each
Wash twice:	1x SSC at 55°C-70°C for 30 minutes each

Low Stringency (allows sequences that share at least 50% identity to hybridize)

Hybridization:	6x SSC at RT to 55°C for 16-20 hours
Wash at least twice:	2x-3x SSC at RT to 55°C for 20-30 minutes each.

In a preferred embodiment of the invention said endogenous bacterial serotype polysaccharide chain length regulator is modified and comprises a polypeptide comprising an amino acid sequence selected from the group consisting of:

i) a polypeptide comprising or consisting of the amino acid sequence as set forth in SEQ ID NO: 32;

ii) a polypeptide comprising or consisting of an amino acid sequence that is at least 75% identical over the full length of the amino acid sequence set forth in SEQ ID NO: 32 and is a bacterial serotype polysaccharide chain length regulator and

wherein said bacterial serotype polysaccharide chain length regulator according to i) and ii) above is modified and substantially lacks enzyme activity.

In a preferred embodiment, the bacterial serotype polysaccharide chain length regulator has at least 75% 80%, 85% or 90%, or most preferably at least 95%, 96%, 97%, 98% or 99% identity with the full-length nucleotide or amino acid sequences as set forth in SEQ ID NO 31 or 32.

According to a further aspect of the invention there is provided a vaccine or immunogenic composition comprising a glycoconjugate vaccine obtained or obtainable from the microbial cell according to the invention.

In a preferred embodiment of the invention said composition further includes an adjuvant

According to a further aspect of the invention there is provided a vaccine or immunogenic composition according to the invention for use in the treatment of microbial infection or a condition resulting from a microbial infection in a subject.

In preferred embodiment of the invention said subject is human.

In an alternative embodiment of the invention said subject is a non-human animal.

In a preferred embodiment of the invention said microbial infection is a streptococcal infection or a condition resulting from a streptococcal infection.

Preferably the streptococcal infection or condition is caused by a streptococcal species selected from the group consisting of: *Streptococcus pyogenes; S.agalactiae, S.dysgalactiae, S. bovis, S.sanguinis, S.suis, S.mitis, S.mutans and S.pneumoniae.*

In a preferred embodiment of the invention the streptococcal infection or condition is caused by *S.pneumoniae.*

Preferably the subject infected with a streptococcal infection or suffering from a condition associated with a streptococcal infection is an elderly human subject.

Alternatively, the subject infected with a streptococcal infection or suffering from a condition associated with a streptococcal infection is an infant human subject.

According to a further aspect of the invention there is provided a microbial culture comprising a cell according to the invention.

According to a further aspect of the invention there is provided a cell culture vessel comprising a microbial culture according to the invention.

In a preferred embodiment of the invention said cell culture vessel is a fermenter.

According to a further aspect of the invention there is provided a method for the manufacture of a glycoconjugate according to the invention comprising the steps:

i) providing a microbial cell culture according to the invention;

ii) providing cell culture conditions; and

iii) culturing and isolating the glycoconjugate from the microbial cells or microbial cell culture.

According to an aspect of the invention there is provided a glycoconjugate obtained or obtainable by the method according to the invention.

According to a further aspect of the invention there is provided a cell culture vessel comprising a bacterial cell culture according to the invention.

In a preferred embodiment of the invention said cell culture vessel is a fermenter.

According to an aspect of the invention there is provided a microbial cell wherein said cell comprises a modified bacterial serotype polysaccharide chain length regulator selected from the group consisting of:

i) a nucleic acid molecule comprising the nucleotide sequence as set forth in SEQ ID NO: 31;

ii) a nucleic acid molecule comprising a nucleotide sequence wherein said sequence is degenerate as a result of the genetic code to the nucleotide sequence defined in (i); and

iii) a nucleic acid molecule the complementary strand of which hybridizes under stringent hybridization conditions to the nucleotide sequence in i) and ii) above wherein said nucleic acid molecule encodes a bacterial serotype polysaccharide chain length regulator and is at least 75% identical to the nucleotide sequence set forth in SEQ ID NO: 31,

wherein said nucleic acid molecule is mutated or deleted for all or part of the nucleotide sequence encoding said bacterial serotype polysaccharide chain length regulator and substantially lacks expression and/or encodes a polypeptide that substantially lacks enzyme activity.

In a preferred embodiment, the bacterial serotype polysaccharide chain length regulator has at least 75% 80%, 85% or 90%, and most preferably at least, 91%, 93%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity with the full-length nucleotide or amino acid sequences as set forth in SEQ ID NO 31 or 32.

According to a further aspect of the invention there is provided a kit comprising one or more microbial cells according to the invention wherein each microbial cell is engineered

to express at least one serotype.

In a preferred embodiment of the invention said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 4, 6B, 9V, 14, 18C, 19F or 23F.

In an alternative embodiment of the invention said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F or 23F.

In a further alternative embodiment of the invention said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F or 23F.

In a further alternative embodiment of the invention said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 4, 6B, 9V, 14, 18C, 19F or 23F.

In a preferred embodiment of the invention said kit comprises microbial cells that express each of said serotypes.

The kit provides therefore a microbial cell or a plurality of microbial cells which express individually one serotype. In combination said microbial cells express serotypes that can provide a multivalent vaccine composition.

Definitions

Microbial Cell Culture

Bacterial cultures used in the process according to the invention are grown or cultured in the manner with which the skilled worker is familiar, depending on the host organism. As a rule, bacteria are grown in a liquid medium comprising a carbon source, usually in the form of sugars, a nitrogen source, usually in the form of organic nitrogen sources such as yeast extract or salts such as ammonium sulfate, trace elements such as salts of iron, manganese and magnesium and, if appropriate, vitamins, at temperatures of between 0°C and 100°C, preferably between 10°C and 60°C, while gassing in oxygen.

The pH of the liquid medium can either be kept constant, that is to say regulated during the culturing period, or not. The cultures can be grown batchwise, semi-batchwise or continuously. Nutrients can be provided at the beginning of the fermentation or fed in semi-continuously or continuously. The products produced can be isolated from the bacteria as described above by processes known to the skilled worker, for example by extraction, distillation, crystallization, if appropriate precipitation with salt, and/or chromatography. In this process, the pH value is advantageously kept between pH 4 and 12, preferably between pH 6 and 9, especially preferably between pH 7 and 8.

An overview of known cultivation methods can be found in the textbook Bioprocess technology 1. Introduction to Bioprocess technology (Gustav Fischer Verlag, Stuttgart, 1991) or in the textbook by Storhas (Bioreaktoren und periphere Einrichtungen [Bioreactors and peripheral equipment] (Vieweg Verlag, Brunswick/Wiesbaden, 1994)).

The culture medium to be used must suitably meet the requirements of the bacterial strains in question. Descriptions of culture media for various bacteria can be found in the textbook "Manual of Methods for General Bacteriology" of the American Society for Bacteriology (Washington D.C., USA, 1981).

As described above, these media which can be employed in accordance with the invention usually comprise one or more carbon sources, nitrogen sources, inorganic salts, vitamins and/or trace elements.

Preferred carbon sources are sugars, such as mono-, di- or polysaccharides. Examples of carbon sources are glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose. Sugars can also be added to the media via complex compounds such as molasses or other by-products from sugar refining. The addition of mixtures of a variety of carbon sources may also be advantageous. Other possible carbon sources are oils and fats such as, for example, soya oil, sunflower oil, peanut oil and/or coconut fat, fatty acids such as, for example, palmitic acid, stearic acid and/or linoleic acid, alcohols and/or polyalcohols such as, for example, glycerol, methanol and/or ethanol, and/or organic acids such as, for example, acetic acid and/or lactic acid.

Nitrogen sources are usually organic or inorganic nitrogen compounds or materials comprising these compounds. Examples of nitrogen sources comprise ammonia in liquid or gaseous form or ammonium salts such as ammonium sulfate, ammonium chloride, ammonium phosphate, ammonium carbonate or ammonium nitrate, nitrates, urea, amino acids or complex nitrogen sources such as cornsteep liquor, soya meal, soya protein,

yeast extract, meat extract and others. The nitrogen sources can be used individually or as a mixture.

Inorganic salt compounds which may be present in the media comprise the chloride, phosphorus and sulfate salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron.

Inorganic sulfur-containing compounds such as, for example, sulfates, sulfites, dithionites, tetrathionates, thiosulfates, sulfides, or else organic sulfur compounds such as mercaptans and thiols may be used as sources of sulfur for the production of sulfur-containing fine chemicals, in particular of methionine.

Phosphoric acid, potassium dihydrogen phosphate or dipotassium hydrogen phosphate or the corresponding sodium-containing salts may be used as sources of phosphorus.

Chelating agents may be added to the medium in order to keep the metal ions in solution. Particularly suitable chelating agents comprise dihydroxyphenols such as catechol or protocatechuate and organic acids such as citric acid.

The fermentation media used according to the invention for culturing bacteria usually also comprise other growth factors such as vitamins or growth promoters, which include, for example, biotin, riboflavin, thiamine, folic acid, nicotinic acid, panthothenate and pyridoxine. Growth factors and salts are frequently derived from complex media components such as yeast extract, molasses, cornsteep liquor and the like. It is moreover possible to add suitable precursors to the culture medium. The exact composition of the media compounds heavily depends on the particular experiment and is decided upon individually for each specific case. Information on the optimization of media can be found in the textbook "Applied Microbiol. Physiology, A Practical Approach" (Editors P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). Growth media can also be obtained from commercial suppliers, for example Standard 1 (Merck) or BHI (brain heart infusion, DIFCO) and the like.

All media components are sterilized, either by heat (20 min at 1.5 bar and 121°C) or by filter sterilization. The components may be sterilized either together or, if required, separately. All media components may be present at the start of the cultivation or added continuously or batchwise, as desired.

The culture temperature is normally between 15°C and 45°C, preferably at from 25°C to 40°C, and may be kept constant or may be altered during the experiment. The pH of the

medium should be in the range from 5 to 8.5, preferably around 7.0. The pH for cultivation can be controlled during cultivation by adding basic compounds such as sodium hydroxide, potassium hydroxide, ammonia and aqueous ammonia or acidic compounds such as phosphoric acid or sulfuric acid. Foaming can be controlled by employing antifoams such as, for example, fatty acid polyglycol esters. To maintain the stability of plasmids it is possible to add to the medium suitable substances having a selective effect, for example antibiotics. Aerobic conditions are maintained by introducing oxygen or oxygen-containing gas mixtures such as, for example, ambient air into the culture. The temperature of the culture is normally 20°C to 45°C and preferably 25°C to 40°C. The culture is continued until formation of the desired product is at a maximum. This aim is normally achieved within 10 to 160 hours.

The fermentation broth can then be processed further. The biomass may, according to requirement, be removed completely or partially from the fermentation broth by separation methods such as, for example, centrifugation, filtration, decanting or a combination of these methods or be left completely in said broth. It is advantageous to process the biomass after its separation.

However, the fermentation broth can also be thickened or concentrated without separating the cells, using known methods such as, for example, with the aid of a rotary evaporator, thin-film evaporator, falling-film evaporator, by reverse osmosis or by nanofiltration. Finally, this concentrated fermentation broth can be processed to obtain the fatty acids present therein.

Capsular Polysaccharides

Bacterial capsules are formed primarily from long-chain polysaccharides with repeat-unit structures. A given bacterial species can produce a range of capsular polysaccharides with different structures and this help distinguish isolates by serotyping, as is the case with *Streptococcal* antigens. Capsules are virulence factors for many pathogenic bacteria. Capsular polysaccharides are often linked to the cell surface of the bacterium via covalent attachments to either phospholipid or lipid-A molecules although some capsular polysaccharides may be associated with the cell in the absence of a membrane anchor capsular polysaccharides can be either homo- or heteropolymers composed of repeating monosaccharides joined by glycosidic linkages.

The genetic loci necessary for the production of bacterial capsules are known and primarily clustered at a single chromosomal locus, which allows for the coordinate regulation of a large number of genes that may be involved in both the biosynthesis and export of capsular polysaccharides. In most bacterial species, the capsule gene clusters demonstrate conserved sequence and genetic organization.

In the case of human pathogens, a large number of different capsule serotypes have been identified, and certain capsular polysaccharides are associated with specific infections. For example, the *Escherichia coli* K1 antigen, a homopolymer of α2,8-linked N-acetylneuraminic acid (NeuNAc), is the major cause of neonatal meningitis. While bacterial species may demonstrate great structural diversity in synthesizing capsules, chemically identical capsular polysaccharides may also be synthesized by different bacterial species. The *Neisseria meningitidis* group B capsular polysaccharide is identical to the K1 polymer of *E. coli*, and the *E. coli* K18, K22, and K100 antigens have the same constituents and structure as the *Haemophilus influenza* serotype b capsule.

The conservation of capsular polysaccharide serotypes between bacterial species is known.

S. pneumonia is encapsulated presenting specific carbohydrates on its cell wall. The different polysaccharides determine the bacterium's serotype. There are approximately 90 different serotypes of pneumococcus and the protection vaccines confer is typically serotype specific. However, these polysaccharide capsule antigens do not elicit protective levels of antibodies or induce immunological memory especially in children under two years of age and in individuals who are immune suppressed. Currently there are three anti-pneumococcal vaccines based on the formulation of different capsular polysaccharide antigens in clinical use comprising up to 13 of the most commonly disease-causing serotypes. WO2011/151760, which is incorporated by reference in its entirety, discloses a vaccine protective against serotype 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F. Also see WO00/56360 and WO00/62801 each of which is incorporated by reference it their entirety.

Prevnar® comprises seven serotypes: 4, 6B, 9V, 14, 18C, 19F and 23F, and is suitable for immunisation of infants less than 2 years old and is crosslinked to a diphtheria toxin carrier protein. An alternative thirteen component vaccine also linked to diphtheria toxin comprises serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F and is suitable for human subjects over 65 years old. A further ten serotype pneumococcal vaccine

comprises the following serotypes; 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F and 23F. These vaccines are effective in protection against pneumococcal infection and are manufactured by chemical conjugation which is expensive.

Adjuvants and Vaccines

Adjuvants (immune potentiators or immunomodulators) have been used for decades to improve the immune response to vaccine antigens. The incorporation of adjuvants into vaccine formulations is aimed at enhancing, accelerating and prolonging the specific immune response to vaccine antigens. Advantages of adjuvants include the enhancement of the immunogenicity of weaker antigens, the reduction of the antigen amount needed for a successful immunisation, the reduction of the frequency of booster immunisations. Selectively, adjuvants can also be employed to optimise a desired immune response, e.g. with respect to immunoglobulin classes and induction of cytotoxic or helper T lymphocyte responses. In addition, certain adjuvants can be used to promote antibody responses at mucosal surfaces.

Adjuvants can be classified according to their source, mechanism of action and physical or chemical properties. The most commonly described adjuvant classes are gel-type, microbial, oil-emulsion and emulsifier-based, particulate, synthetic and cytokines. More than one adjuvant may be present in the final vaccine product according to the invention. The origin and nature of the adjuvants currently being used or developed is highly diverse. For example, MDP is derived from bacterial cell walls; saponins are of plant origin, squalene is derived from shark liver and recombinant endogenous immunomodulators are derived from recombinant bacterial, yeast or mammalian cells.

There are several adjuvants licensed for veterinary vaccines, such as mineral oil emulsions that are too reactive for human use. Similarly, complete Freund's adjuvant is one of the most powerful adjuvants known.

The vaccine compositions of the invention can be administered by any conventional route, including injection. The administration may be, for example, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or intradermally. The vaccine compositions of the invention are administered in effective amounts. An “effective amount” is that amount of a vaccine composition that alone or together with further doses, produces the desired response. In the case of treating a particular bacterial disease the desired response is providing protection when challenged by an infective agent.

Throughout the description and claims of this specification, the words “comprise” and “contain” and variations of the words, for example “comprising” and “comprises”, means “including but not limited to”, and is not intended to (and does not) exclude other moieties, additives, components, integers or steps. “Consisting essentially” means having the essential integers but including integers which do not materially affect the function of the essential integers.

Throughout the description and claims of this specification, the singular encompasses the plural unless the context otherwise requires. In particular, where the indefinite article is used, the specification is to be understood as contemplating plurality as well as singularity, unless the context requires otherwise.

Features, integers, characteristics, compounds, chemical moieties or groups described in conjunction with a particular aspect, embodiment or example of the invention are to be understood to be applicable to any other aspect, embodiment or example described herein unless incompatible therewith.

An embodiment of the invention will now be described by example only and with reference to the following figures:

Figure 1: Immunoblot of pB4 expressed in various *E. coli* backgrounds. Lysed, whole cell samples were separated by SDS-PAGE on a 4-12 % Bis-Tris gel and detected using anti-serotype 4 and anti-Prevnar primary antisera, and fluorescent secondary antibody. Lane 1: *E. coli* W3110 carrying recombinant pB-4 plasmid; Lane 2: *E. coli* W3110 carrying empty vector pBBR1MCS3; Lane 3: *E. coli* CLM37 carrying recombinant pB-4 plasmid and expressing GalE; Lane 4: *E. coli* W3110ΔCld carrying recombinant pB-4 plasmid; Lane 5: *E. coli* W3110ΔCld carrying recombinant pB-4 plasmid and expressing wzDE; Lane 6: *E. coli* DH5α expressing wzDE; Lane 7: *E. coli* W3110 carrying recombinant pB-4 plasmid and expressing wzDE; Lane 8: *S. pneumoniae* Tigr4; Lane 9: Prevnar 13 loaded at a concentration of 22 ng of polysaccharide. M : Molecular weight marker PageRuler Plus. Polysaccharides don’t always migrate through a gel based on molecular weight; charge, secondary structure and degree of polymerization also play a part, therefore the marker cannot be used to estimate the size of polysaccharide; and

Fig 2: Immunoblot of pB5 expressed in CLM37. Lysed, whole cell samples were separated by SDS-PAGE on a 10 % Bis-Tris gel and detected using anti-serotype 5 and fluorescent

secondary antibody. M: Molecular weight marker PageRuler Plus. Lane 1: *E. coli* W3110 carrying recombinant pB-5 plasmid.

Figure 3: Nucleotide and amino acid sequences of wze and wzd polypeptides;

Figure 4: Immunoblot of pB8 expressed in various *E. coli* backgrounds. Lysed, whole cell samples were separated by SDS-PAGE on a 12 % Bis-Tris gel and detected using anti-serotype 8 primary antisera and fluorescent secondary antibody. Lane 1: *E. coli* W3110 carrying empty vector pBBR1MCS3; Lane 2 and 3: separate single colony isolates of *E. coli* W3110 carrying recombinant pB-8 plasmid; Lane 4 and 5: separate single colony isolates of *E. coli* W3110ΔCld carrying recombinant pB-8 plasmid; Lane 6 and 7: separate single colony isolates of *E. coli* W3110ΔCld carrying recombinant pB-8 plasmid and expressing wzDE; Lane 8 and 9: separate single colony isolates of *E. coli* W3110 carrying recombinant pB-8 plasmid and expressing wzDE; Lane 10: *S. pneumoniae* 573/62 (serotype 8 strain from SSI, Denmark); M : Molecular weight marker PageRuler Plus; and

Figure 5 sequence of capsule loci comprising reducing end sugars with glucose as reducing end sugar.

Materials and Methods

Table 1: Accession numbers for capsule loci with acetylated reducing end sugars.

Serotype	EMBL accession
1	CR931632 (SEQ ID NO 58)
4	CR931635 (SEQ ID NO 1)
5	CR931637 (SEQ ID NO 59)
12F	CR931660 (SEQ ID NO 60)
12A	CR931658 (SEQ ID NO 61)
12B	CR931659 (SEQ ID NO 62)
25F	CR931690 (SEQ ID NO 63)
25A	CR931689 (SEQ ID NO 64)
38	CR931710 (SEQ ID NO 65)
44	CR931717 (SEQ ID NO 66)
45	CR931718 (SEQ ID NO 67)
46	CR931719 (SEQ ID NO 68)

Table 2: Nucleotide Identity tables:

wzd	1	4	5	12F	12A	12B	25F	25A	38	44	45	46
1		89	94	73	72	72	-	-	-	73	89	72
4	89		85	81	80	80	-	-	-	81	98	80
5	94	85		73	72	73	-	-	-	73	86	72
12F	73	81	73		99	99	-	-	-	100	81	99
12A	72	80	72	99		99	-	-	-	99	81	100
12B	72	80	73	99	99		-	-	-	99	81	99
25F	-	-	-	-	-	-		100	99	-	-	-
25A	-	-	-	-	-	-	100		99	-	-	-
38	-	-	-	-	-	-	99	99		-	-	-
44	73	81	73	100	99	99	-	-	-		81	99
45	89	98	86	81	81	81	-	-	-	81		81
46	72	80	72	99	100	99	-	-	-	99	81	

Matches with ≥ 98% coverage
- Match ≤ 13%

wze	1	4	5	12F	12A	12B	25F	25A	38	44	45	46
1		97	93$	75	75	75	65*	65*	65*	75	97	75
4	97		90	75	75	75	-	-	-	75	97	75
5	93$	90		76	76	76	-	-	-	76	90	76
12F	75	75	76		99	99	-	-	-	99	74	99
12A	75	75	76	99		100	-	-	-	99	75	100
12B	75	75	76	99	100		-	-	-	99	75	100
25F	65*	-	-	-	-	-		100	100	-	-	-
25A	65*	-	-	-	-	-	100		100	-	-	-
38	65*	-	-	-	-	-	100	100		-	-	-
44	75	75	76	99	99	99	-	-	-		75	99
45	97	97	90	74	75	75	-	-	-	75		75
46	75	75	76	99	100	100	-	-	-	99	75	

* Match with 74% coverage
$ Match with 93% coverage
Rest match ≥ 95% coverage
- If match < 60%

Table 3: Protein identity tables:

wzd	1	4	5	12F	12A	12B	25F	25A	38	44	45	46
1		90	89	73	72	73	45*	45*	45*	73	88	72
4	90		84	78	78	78	49*	49*	49*	78	97	78
5	89	84		73	72	73	47*	47*	47*	73	85	72
12F	73	78	73		98	100	49*	49*	49*	100	79	98
12A	72	78	72	98		98	49*	49*	49*	98	77	100
12B	73	78	73	100	98		49*	49*	49*	100	79	98
25F	45*	49*	47*	49*	49*	49*		100	100	49*	49*	49*
25A	45*	49*	47*	49*	49*	49*	100		100	49*	49*	49*
38	45*	49*	47*	49*	49*	49*	100	100		49*	49*	49*
44	73	78	73	100	98	100	49*	49*	49*		79	98
45	88	97	85	79	77	79	49*	49*	49*	79		77
46	72	78	72	98	100	98	49*	49*	49*	98	77	

* 94% coverage

Rest are ≥ 99% coverage

wze	1	4	5	12F	12A	12B	25F	25A	38	44	45	46
1		100	89	80	80	80	55	55	55	80	93	80
4	100		89	80	80	80	55	55	55	80	93	80
5	89	89		81	81	81	53	53	53	82	89	81
12F	80	80	81		100	100	51	51	51	99	80	100
12A	80	80	81	100		100	51	51	51	99	80	100
12B	80	80	81	100	100		51	51	51	99	80	100
25F	55	55	53	51	51	51		100	100	51	55	51
25A	55	55	53	51	51	51	100		100	51	55	51
38	55	55	53	51	51	51	100	100		51	55	51
44	80	80	82	99	99	99	51	51	51		80	99
45	93	93	89	80	80	80	55	55	55	80		80
46	80	80	81	100	100	100	51	51	51	99	80	

All matches have ≥ 95% coverage

Table 4: accession number for capsule loci comprising reducing end sugars with glucose as reducing end sugar.

Serotype	Accession number
8 (SEQ ID NO 57)	CR931644

Table 5

Serotype	Capsule sequence accession number
6A	CR931638 (SEQ ID NO: 74)
6B	CR931639 (SEQ ID NO: 75)
7F	CR931643 (SEQ ID NO: 76)
9V	CR931648 (SEQ ID NO: 77)
14	CR931662 (SEQ ID NO: 78)
18C	CR931673 (SEQ ID NO: 79)
19A	CR931675 (SEQ ID NO: 80)
19F	CR931678 (SEQ ID NO: 81)
23F	CR931685 (SEQ ID NO: 82)

Construction of W3110 CldΔKan

First, a construct was assembled consisting of flanks with homology to the ends of Cld and surrounding DNA, attached to a kanamycin resistance determinant. The construct was assembled by PCR amplification of the kanamycin resistance determinant from pEXT22 using Q5 high-fidelity DNA polymerase and primers: KanPromF 5`-agcttgcagtgggcttacat-3` (SEQ ID NO 83) and KanR 5`-gaaatctcgtgatggcaggt-3` (SEQ ID NO 84). Next a left flank was amplified from *E. coli* W3110 DNA using primers: cldLF 5`-cttcaactctcgcctggaac-3` (SEQ ID NO 85) and cldLRKan 5`-atgtaagcccactgcaagcttggcattgttatagccagca-3`(SEQ

ID NO 86), which contains an overlapping region to the kanamycin resistance cassette. A right flank was amplified from *E. coli* W3110 DNA using primers: cldRFKan 5`-acctgccatcacgagatttcgttggtgttctcgccaaact-3`(SEQ ID NO 87), which contains an overlapping region to the kanamycin resistance cassette, and cldRR 5`-ctcctacaccgccaaactgt-3`(SEQ ID NO 88). All PCR reactions were performed using the following cycling conditions: 98 °C for 30s, then 30 rounds of 98 °C 5s, 63 °C for 10s and 72 °C for 20s, followed by a final extension of 72 °C for 2 mins. A soeing PCR was performed using 100ng of the purified PCR of the Kan cassette, cld left flank and cld right flank as templates and the primers cldLF and cldRR. The cycling conditions were as follows: 98 °C for 30s, then 30 rounds of 98 °C 10s, 64 °C for 20s increasing by 0.2 °C every cycle and 72 °C for 1 min, followed by a final extension of 72 °C for 3 mins. The correct size band was extracted from a gel and purified using a Qiagen gel extraction kit according to manufacturer's instructions. Polynucleotide kinase was used to phosphorylate the ends of the PCR product to facilitate blunt end cloning into the SmaI restriction enzyme site of an R6K ori containing plasmid (pBlueX). The construct was verified by Sanger sequencing. Next, competent *E. coli* W3110 cells were transformed with the CldΔKan plasmid and recombinants selected for using kanamycin selection. Recombinant cells were screened to ensure the vector backbone was no longer present and that the wild type cld gene was interrupted with the kan cassette.

Construction of wzDE expression vector

The *S. pneumoniae* genes wzD and wzE were amplified from a serotype 4 strain. wzdLBamHI 5`-TAAGGATCCggagaaatgatgaaagaacaa-3` (SEQ ID NO 89) and wzeRSalI 5`-TAAGTCGACtccccttccatactatattca-3`(SEQ ID NO 90). The genes were cloned into pEXT21 via the BamHI and SalI restriction sites.

Competent cells and electroporation

Electrocompetent *E. coli* W3110 or W3110ΔCld were prepared by three successive washes in an equal volume of 10% glycerol at 4 °C, followed by centrifugation at 4000 rpm 4°C for 10 mins. Plasmid DNA was introduced into the competent cells by electroporation at 2kV 200Ω 25 μF in a 2mm gap cuvette. Cells were recovered at 37 °C in SOC broth for 1 hr before plating on LB agar containing appropriate antibiotics.

CRISPR mutagenesis

The method of Jiang et al. (2015, Appl Environ Microbiol 81:2506 –2514.) was used for CRISPR mutagenesis. Plasmids pCas (addgene #62225) and pTarget (addgene #62226)

were used as specified in the paper. gRNA design was carried out using ATUM crispr gRNA design tool (https://www.atum.bio/eCommerce/cas9/input)

Expression testing of pB4

E. coli cultures containing various plasmids were grown for 16 hrs with appropriate antibiotics before being diluted into fresh SSOB media to an OD_{600} of 0.03. The media was supplemented with 0.5 mM IPTG and 4 mM $MnCl_2$ and incubated at 28 °C for 24 hrs. All samples were OD_{600} matched, washed with PBS, and then lysed using a Bioruptor ultrasonic processor (Diagenode, Belgium) set on a high pulse rate set for 30s on 30s off for 15 mins. Lysed samples were mixed with SDS-PAGE sample buffer and separated on 4-12% Bis-Tris gel in MOPS buffer (Invitrogen). Samples were electroblotted onto nitrocellulose membrane (GE Healthcare) using a semi-dry transfer unit (AA Hoefer). Membranes were blocked for 1 hr in PBS containing 2% w/v skimmed milk powder. Serotype 4 rabbit anti-capsule antiserum (Statens Serum Institut, Denmark) was used at a dilution of 1:1000 and mouse anti-prevnar antiserum at a dilution of 1:1000 in PBS containing 2% w/v skimmed milk powder and 0.1 % v/v Tween 20. After 1 hr incubation with primary antibody, membranes were washed three times with PBS (0.1% Tween 20) and then incubated for 45 mins with a secondary goat anti-rabbit IgG IRDye800 and a goat anti-mouse IgG IRDye680 conjugate antibody at a dilution of 1:10000. Membranes were washed a further three times in PBS (0.1% Tween 20) and once with PBS before signal detection with the Odyssey LI-COR detection system (LI-COR Biosciences UK Ltd).

Examples

Group 1 capsule loci in *E. coli* have four conserved genes at the beginning of the locus, 3 of which control polymerization and export, independent of capsule structure (wza, wzb, wzc) (Whitfield 2006). In *S. pneumoniae* it is known that polymer length is controlled by a tyrosine kinase phosphoregulatory system. In *S. pneumoniae* the 4 initial genes in the capsule locus are highly conserved. The 2 domains of WzC are contained in 2 separate proteins (CpsC and CpsD) (Yother 2011). In *S. pneumoniae* D39, inactivation of either of these genes resulted in reduced polymer lengths (Bender et al 2003). The phosphatase CpsB is also involved in polymerization but the results of inactivation are variable depending on serotype. Removal of cpsB resulted in higher capsule production in D39 but no change to polymerization (Bender et al 2003), and less capsule production in 19F (Morona et al 2000). CpsB can be compensated for by WzB from the colanic acid locus of *E. coli*, despite the two proteins having different structures and modes of action

(Hagelueken et al 2009). It is therefore unlikely that CpsB would be needed for recombinant expression in *E. coli*. In *S. pneumoniae* the genes cpsA-D are also known as wzg, wzh, wzd and wze.

Each glycan in *E. coli* has its own export and polymerization machinery (Table 6):

Table 6

	O-antigen	ECA	Colanic Acid
Flippase	rfbX	wzxE	WzxC
Polymerase	wbbH	wzyE	WcaD
Chain length	Cld	wzzE	WzC (wzB)

Any of these genes could potentially interfere with recombinant polysaccharide expression in *E. coli*, given enough similarity of glycan structure. Previous results have shown that polymer levels are truncated for serotypes 4, 5 and 12F but not for serotype 8 (Kay et al. 2016). It is unlikely that this is entirely due to incorrect sugar incorporation of GlcNAc as the reducing end sugar, even though the serotypes 4, 5 and 12F were expressed in a wecA positive host strain of *E. coli*. When serotypes 4 and 5 were expressed in CLM37, which lacks wecA, the phenomenom was still seen (Fig 1 lane 3 and fig 2 respectively). Therefore, chain length restriction via cld is likely to be a problem restricted to capsule types that have an acetylated reducing end sugar, which are also the serotypes that are substrates for CjPglB. This list includes serotypes: 1, 4, 5, 12F, 12A, 12B, 25F, 25A, 38, 44, 45 and 46.

Interaction of host and recombinant factors in heterologous production of fully polymerized polysaccharide

E. coli strains were grown in SSOB broth and induced as described above with 0.5 mM IPTG. The expression of recombinant serotype 4 polysaccharide was compared in strains W3110, CLM37 (which lacks wecA) with additional expression of GalE, and W3110ΔCld (where the O-antigen chain length determinant has been interrupted). The additional impact of expressing wzD and wzE was also evaluated. Whole cell, OD-matched lysates, were run on an SDS-PAGE gel alongside positive controls of lysed *S. pneumoniae* Tigr4 cells and Prevnar 13 vaccine (equivalent to 22 ng of serotype 4 capsule. A Western blot

was carried out using anti-SP4 antiserum and anti-Prevnar antiserum. Figure 1 shows that expressing wzDE improved polymerization, but only when the *E. coli* O-antigen chain length regulator (Cld) is removed, otherwise with Cld polymerization is the same as without wzDE (lane 7 vs lane 1). When Cld is removed, no recombinant polysaccharide is produced at all (lane 4), unless wzDE is added (lane 5). When WecA is removed, even with GalE added there is still restriction in chain length, although there is a small amount of more highly polymerized capsule present. Crucially, the level of polymerization achieved by recombinant expression of serotype 4 in an *E. coli* background without Cld and with wzD and wzE, is equivalent to that in Prevnar.

Chain length regulation in additional serotypes of recombinantly expressed *S. pneumoniae* capsule.

E. coli strains were grown in LB broth containing 0.1% glucose and induced after 2 hrs starter culture growth with 0.5 mM IPTG. The expression of recombinant serotype 8 polysaccharide was compared in strains W3110 and W3110ΔCld (where the O-antigen chain length determinant has been interrupted). The additional impact of expressing wzD and wzE was also evaluated. Whole cell, OD-matched lysates, were run on an SDS-PAGE gel alongside positive controls of lysed *S. pneumoniae* 573/62 (serotype 8 strain from SSI, Denmark). A Western blot was carried out using anti-SP8 antiserum. Figure 4 shows that expressing wzDE improved polymerization, with or without the removal of the *E. coli* O-antigen chain length regulator (Cld). This is in contrast to the results for serotype 4, where removal of Cld was necessary to see improved polymerization with wzDE. This suggests that Cld is not compensating for the native chain length regulators when the capsule is expressed in *E. coli*. In contrast to serotype 4, serotype 8 capsule does not have an acetylated reducing end sugar but instead has glucose at the reducing end.

References

Bender, M. H., R. T. Cartee and J. Yother (2003). "Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in Streptococcus pneumoniae." J Bacteriol **185**(20): 6057-6066.

Hagelueken, G., H. Huang, I. L. Mainprize, C. Whitfield and J. H. Naismith (2009). "Crystal structures of Wzb of Escherichia coli and CpsB of Streptococcus pneumoniae, representatives of two families of tyrosine phosphatases that regulate capsule assembly." J Mol Biol **392**(3): 678-688.

Kay, E. J., L. E. Yates, V. S. Terra, J. Cuccui and B. W. Wren (2016). "Recombinant expression of Streptococcus pneumoniae capsular polysaccharides in Escherichia coli." Open Biol **6**(4): 150243.

Linton, D., N. Dorrell, P. G. Hitchen, S. Amber, A. V. Karlyshev, H. R. Morris, A. Dell, M. A. Valvano, M. Aebi and B. W. Wren (2005). "Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway." Molecular Microbiology **55**(6): 1695-1703.

Morona, J. K., J. C. Paton, D. C. Miller and R. Morona (2000). "Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in streptococcus pneumoniae." Mol Microbiol **35**(6): 1431-1442.

Whitfield, C. (2006). "Biosynthesis and assembly of capsular polysaccharides in Escherichia coli." Annu Rev Biochem **75**: 39-68.

Yother, J. (2011). "Capsules of Streptococcus pneumoniae and other bacteria: paradigms for polysaccharide biosynthesis and regulation." Annu Rev Microbiol **65**: 563-581.

Claims

1. A microbial cell wherein said cell comprises:

   i) a nucleic acid molecule comprising a nucleotide sequence encoding one or more polypeptides involved in the synthesis of one or more heterologous streptococcal capsular polysaccharides not expressed by said microbial cell;

   ii) a nucleic acid molecule comprising a nucleotide sequence encoding at least one heterologous polysaccharide chain length regulator;

   iii) a nucleic acid molecule comprising a nucleotide sequence encoding one or more carrier polypeptides that comprise one or more glycosylation motifs; and

   iv) a nucleic acid molecule comprising a nucleotide sequence encoding an oligosaccharyltransferase.

2. The microbial cell according to claim 1 wherein the genome of said microbial cell is modified in one or more endogenous genes encoding a polysaccharide chain length regulator characterised in that said endogenous bacterial capsular polysaccharide chain length regulator is modified or inactivated and substantially lacks expression and/or activity.

3. The microbial cell according to claim 1 or 2 wherein said microbial cell is a bacterial cell.

4. The microbial cell according to any one of claims 1 to 3 wherein said nucleic acid molecule encoding one or more polypeptides involved in the synthesis of one or more heterologous bacterial capsular polysaccharides synthesizes a *Streptococcus pneumoniae* polysaccharide serotype selected from one of the 96 serotypes expressed by *Streptococcus pneumoniae.*

5. The microbial cell according to any one of claims 1 to 4 wherein said polysaccharide comprises an acetylated reducing end sugar and is selected from the group of serotypes consisting of: 1, 4, 5, 12F,12A, 12B, 25F, 25A, 38, 44, 45 or 46.

6. The microbial cell according to any one of claims 1 to 4 wherein said polysaccharide comprises a glucose as an reducing end sugar is selected from the group consisting of serotypes: 2, 6A, 6B, 7F, 7A, 7B, 7C, 8, 9A, 9L, 9N, 9V, 11F, 11A, 11B, 11C, 11D, 13, 14, 15F, 15A, 15B, 15C, 16F, 16A, 17F, 17A, 18F, 18A, 18B, 18C, 19F, 19A, 19B, 19C, 20, 21, 22F, 22A, 23F, 23A, 23B, 24F, 24A, 24B, 27, 28F, 28A, 32F, 32A, 33F, 33A, 33B, 33D, 34, 35A, 35B, 35C, 36, 40, 41F, 41A, 42, 48.

7. The microbial cell according to any one of claims 1 to 4 wherein said polysaccharide comprises a galactose or a galactose derivative as the reducing end sugar is selected from the group of serotypes: 10F, 10A, 10B, 10C, 29, 31, 33C, 35F, 39, 43, 47F, 47A.

8. The microbial cell according to any one of claims 1 to 4 wherein said serotypes are selected from a group consisting of: 1, 4, 5, 8, 12F, 12A, 12B, 25F, 25A, 38, 44, 45 or 46

9. The microbial cell according to any one of claims 1 to 4 wherein said serotypes are selected from the group consisting of: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F.

10. The microbial cell according to any one of claims 1 to 9 wherein said nucleic acid encoding a heterologous polysaccharide chain length regulator is encoded by the wzd gene and wze gene.

11. The microbial cell according to any one of claims 1 to 10 wherein said oligosaccharyltransferase is a *Campylobacter* oligosaccharyltransferase.

12. The microbial cell according to any one of claims 1 to 10 wherein said oligosaccharyltransferase is encoded by the PglL gene encoded by the nucleotide sequence set forth in SEQ ID NO: 72.

13. The microbial cell according to any one of claims 1 to 10 wherein said microbial cell is adapted to express a glycosyltransferase encoded by a nucleotide sequence as set forth in SEQ ID NO: 71.

14. The microbial cell according to any one of claims 1 to 4, 7 and 9 to 10 wherein said microbial cell is adapted to express a ribitol synthesis gene encoded by a nucleotide sequence as set forth in SEQ ID NO: 73.

15. The microbial cell according to any one of claims 1 to 14 wherein said one or more carrier polypeptide(s) includes the amino acid motif: Asn-X-Ser or Asn-X-Thr where X is any amino acid except proline or includes the amino acid motif: D/E-X-N-X-S/T, wherein X is any amino acid except proline.

16. A vaccine or immunogenic composition comprising one or more glycoconjugates obtained or obtainable from the microbial cell according to any one of claims 1 to 15 and optionally including an adjuvant.

17. A vaccine or immunogenic composition according to claim 16 for use in the treatment of microbial infection or a condition resulting from a microbial infection in a subject.

18. The vaccine or immunogenic composition according to the use of claim 17 wherein said subject is human or a non-human animal.

19. The vaccine or immunogenic composition according to the use of claim 17 or 18 wherein said microbial infection is a streptococcal infection or a condition resulting from a streptococcal infection.

20. The vaccine or immunogenic composition according to this use of claim 19 wherein said streptococcal infection is caused by *S.pneumoniae*.

21. A microbial culture comprising a cell according to any one of claims 1 to 15.

22. A cell culture vessel comprising a microbial culture according to claim 21.

23. A method for the manufacture of a glycoconjugate comprising the steps:

i) providing a microbial cell culture according to claim 21;

ii) providing cell culture conditions; and

iii) culturing and isolating the glycoconjugate from the microbial cells or microbial cell culture.

24. A glycoconjugate obtained or obtainable by the method according to claim 23.

25. A vaccine composition comprising a plurality of glyconjugates obtained or obtainable by the method of claim 23.

26. A kit comprising one or more microbial cells according to any one of claims 1 to 15 wherein each microbial cell is engineered to express at least one serotype.

27. The kit according to claim 26 wherein said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 4, 6B, 9V, 14, 18C, 19F or 23F.

28. The kit according to claim 26 wherein said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F or 23F.

29. The kit according to claim 26 wherein said kit comprises a microbial cell genetically engineered to express a serotype selected from the group consisting of: 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F or 23F.

30. The kit according to any one of claims 26 to 29 wherein said kit comprises microbial cells that express each of said serotypes.

Figure 1

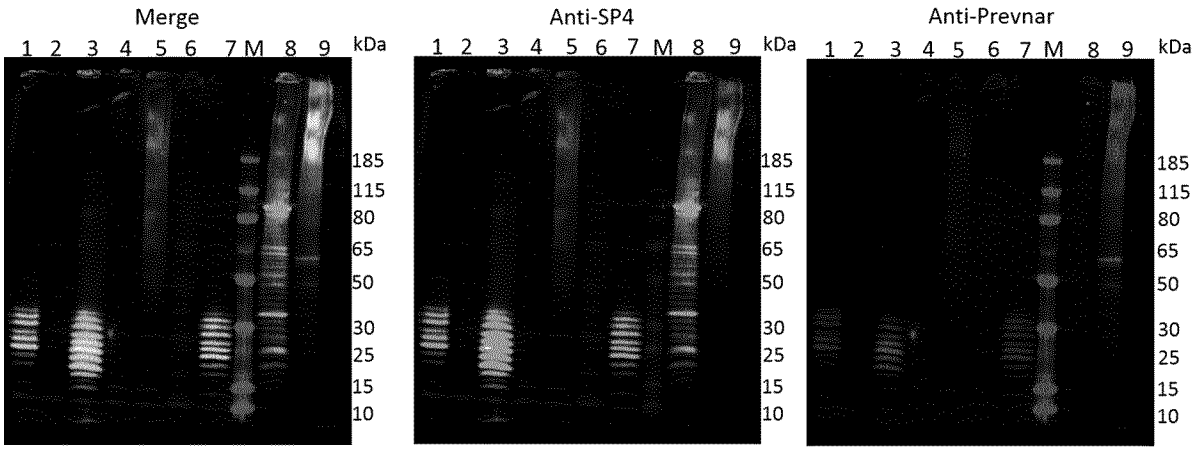


Figure 2

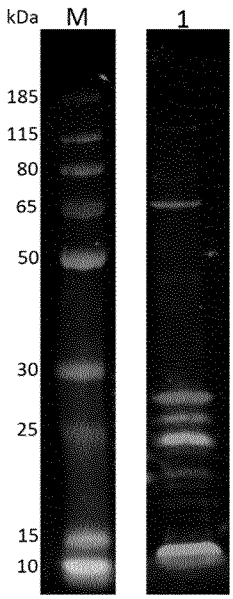


Figure 3: SEQUENCES

>SP1 wzd **(SEQ ID NO: 2)**

ATGAAGGAACAAAACACTTTGGAAATCGATGTATTGCAACTATTCAGAGCTTTATGGAAA
AGAAAGTTGGTCATTTTATTAGTGGCAATTATAACTTCTTCAGTTGCTTTTGCCTACAGT
ACTTTTGTTATCAAACCTGAGTTTACTAGTACGACTCGGATTTATGTAGTTAACCGTAAT
CAGGGAGAGAAGCCTGGTTTAACCAATCAAGACTTGCAGGCAGGAGCATACTTGGTTAAA
GACTATCGTGAAATTGTCCTATCGCAGGATGTTTTGGAGGAAGTTATTTCTGATTTGAAA
CTAGATTTGGTGCCAAAAGGTTTGGCTAATAAAATTAAAGTAACAGTACCAGTTGATACC
CGTATTGTCTCTGTTTCAGTTAGTGATCGAGTTCCTGAAGAGGCAAGCCGTATCGCTAAC
TCTTTAAGAGAAGTAGCTGCTCAAAAAATTATCAGTATTACTCGTGTTTCTGATGTGACA
ACACTGGAGGAGGCAAGACCGGCGATATCACCGTCTTCGCCAAATATTAAACGCAATACA
CTAATTGGTTTTTTGGCAGGGGTGATTGGAACTAGTGTTATAGTTCTTCTTCTTGAACTT
TTGGACACTCGTGTGAAACGTCCGGAAGATATCGAAAATGCACTGCAGATGACACTTTTG
GGAGTTGTACCAAACTTGGGTAAGTTGAAATAG

>SP4 wzd (**SEQ ID NO: 3)**

ATGATGAAAGAACAAAATACGATAGAAATCGATGTATTTCAATTAGTTAAAAGCTTGTGG
AAACGCAAGCTAATGATTTTAATAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCACTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
AATCAAGGAGACAAGCCGGGGTTGACAAATCAGGATTTGCAGGCAGGAACTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGGATGTTTTGGAGGAAGTTGTTTCTGATTTG
AAACTAGATTTGACGCCAAAAGGTTTGGCTAATAAAATTAAAGTGACAGTACCAGTTGAT
ACCCGTATTGTCTCTATTTCAGTTAATGATCGAGTTCCTGAAGAGGCAAGCCGTATCGCT
AACTCTTTGAGAGAAGTAGCTGCTCAAAAAATTATCAGTATTACTCGTGTTTCTGACGTG
ACAACACTGGAGGAGGCAAGGCCGGCGATATCCCCGTCTTCGCCAAATATTAAACGCAAT
ACACTAATTGGTTTTTTGGCAGGGGTGATTGGAACTAGTGTTATAGTTCTTCATCTTGAA
CTTTTGGATACTCGTGTGAAACGTCCGGAAGATATCGAAAATACATTGCAGATGACACTT
TTGGGAGTTGTGCCAAACTTGGGTAAGTTGAAATAG

>SP5 wzd (**SEQ ID NO: 4)**

ATGGAGAAACAAAACACTTTGGAAATCGATGTATTGCAACTATTCAGATCTTTATGGAAA
AGAAAGTTGGTCATTTTATTAGTGGCAATTATAACTTCTTCAGTTGCTTTTGCCTACAGT
ACTTTTGTTATCAAACCTGAGTTTACTAGTACGACTCGGATTTATGTAGTTAGCCGTGAT
CAGGGAGAGAAGTCTGGTTTAACCAATCAAGACTTGCAGGCAGGATCGTACTTGACTAAA
GACTATCGTGAAATTATCCTATCGCAGGATGTTTTGGAGGAAGTTGTTTCTGATTTGAAA
CTAGATTTGACGCCAAAAGGTTTGGCTAATAAAATTAAAGTGGCAGTACCAGTTGATACC
CGTATCGTGTCTATTTCGGTTAATGATCAAGTTCCTGAAGGGGCAAGCCGTATCGCTAAC
TCTTTGAGAGAAGTGGCGGCTCAAAAAATTATCAGTATTACTCGTGTTTCTGATGTGACA
ACACTGGAGGAGGCAAGACCAGCGATATCACCGTCTTCGCCAAATATTAAACTCAATACA
CTAATTGGTTTTTTGGTAGGATTGATTGTCATAAATGTTACAGTTCTTCTTCTTGAACTT
TTGGACACACGAGTGAAACGTCCGGAAGATATCGAAGATGCACTACAGATGACACTTTTG
GGGGTTGTACCAAACTTGGATAAGTTGAAATAG

Figure 3 (continued)

>SP12F wzd (**SEQ ID NO: 5)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTATTTAAAACCTTGTGG
AAACGCAAGCTAATGATTTTATTAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCGCTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
GATCAAGGAGACAAGTCGGGGCTGACAAATCAGGATTTGCAGGCAGGATCTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGAATGTATTGGAAAAAGTAGCGACAAATTTG
AAGTTGGATATACCAGCCAAAACGTTAGCTAGAAAAGTTCAAGTGACTGTACCAGTTGAC
ACTCGTATCGTCTCAATCTCTGTCAAGGATAAACAGCCAGAGGAAGCCAGTCGTATCGCT
AATTCTCTACGAGAAGTTGCTGCAGAAAAGATCATCGCTGTAACGCGAGTATCTGATGTG
ACGACACTTGAAGAAGCGCGACCAGCTACGACTCCCTCTTCTCCAAATGTTGGACGCAAT
TCCTTGTTTGGTTTTTTTGGAGGAGCAGTCGTAACAGTAATTGCTGTTCTTTTGATTGAG
TTGTTCGACATCCGTGTGAAACGTCCTGAAGATGTCGAAGATGTACTGCAAATTCCACTT
TTAGGGGTCGTTCCAGATTTGGACAAAATGAAATAG

>SP12A wzd (**SEQ ID NO: 6)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTATTTAAAACCTTGTGG
AAACGCAAGCTAATGATTTTATTAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCGCTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
GATCAAGGAGACAAGTCGGGGCTGACAAATCAGGATTTGCAGGCAGGAACTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGAATGTATTGGAAAAAGTAGCGACAAATTTG
AAGTTGGATATACCAGCCAAAACGTTAGCTAGAAAAGTTCAAGTGACTGTACCAGTTGAC
ACTCGTATCGTCTCAATCTCTGTCAAGGATAAACAGCCAGAGGAAGCCAGTCGTATCGCT
AATTCTCTACGAGAAGTTGCTGCAGAAAAGATCATCGCTGTAACGCGAGTATCTGATGTA
ACGACACTTGAAGAAGCGCGACCAGCTACGACTCCCTCTTCTCCAAATGTTCGTCTCAAT
TCCTTGTTTGGTTTTTTTGGAGGAGCAGTCGTAACAGTAATTGCTGTTCTTTTGATTGAG
TTGTTCGACATCCGTGTGAAACGTCCTGAAGATGTCGAAGATGTACTGCAAATTCCACTT
TTAGGGGTCGTTCCAGATTTGAACAAAATGAAATAG

>SP12B wzd (**SEQ ID NO: 7)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTATTTAAAACCTTGTGG
AAACGCAAGCTAATGATTTTATTAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCGCTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
GATCAAGGAGACAAGTCGGGGCTGACAAATCAGGATTTGCAGGCAGGATCTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGAATGTATTGGAAAAAGTAGCGACAAATTTG
AAGTTGGATATACCAGCCAAAACGTTAGCTAGAAAAGTTCAAGTGACTGTACCAGTTGAC
ACTCGTATCGTCTCAATCTCTGTCAAGGATAAACAGCCAGAGGAAGCCAGTCGTATCGCT
AATTCTCTACGAGAAGTTGCTGCAGAAAAGATCATCGCTGTAACGCGAGTATCTGATGTA
ACGACACTTGAAGAAGCGCGACCAGCTACGACTCCCTCTTCTCCAAATGTTGGACGCAAT
TCCTTGTTTGGTTTTTTTGGAGGAGCAGTCGTAACAGTAATTGCTGTTCTTTTGATTGAG
TTGTTCGACATCCGTGTGAAACGTCCTGAAGATGTCGAAGATGTACTGCAAATTCCACTT
TTAGGGGTCGTTCCAGATTTGGACAAAATGAAATAG

Figure 3 (continued)
>SP25F wzd (**SEQ ID NO: 8)**

ATGAATAATCGTATAGTAGACATCAGTTTGTTTGATTTATTGAATGTCTTATGGCTTCGT
AAATGGATCATTGTTTTTGTATCATTTGTTTTTGGATGTATTTCTTTTTTATATTATCAA
TTTTTTGTACCCAAAACCTATACAAGTGATACTCGTATATATGTGGTAAGCAAAGGGGAT
AATAAAAATCTGTCTAATGCAGATTTGCAAGCAGGAAGTGCTCTAACAAAGGATTACAAA
GAAATCATTTTATCAGATGAAGTGTTAGAAGAAACAATTTCTGATTTAAAACTAGAATCA
ACTTTGGAATCTCTAGCTTCGAAAATTAAAATTTCTATCCCTGCGGAAACTCGAATTATT
TCCATTAGTGTTACAAATACGAACTCAGATGAAGCTGCTCGAATAGCAAATGGGATAAGG
AAGGTAGCTGCCCTTAAAATCAAGGAAGTAACTCAAGTTACAGATGTCACCACTTTACAG
ACTGCACGCCCTCCCCAAACTCCGTCAGGACCGCATGTGCGAAAATCTACTACAGCAGGG
CTTGTTCTTGGTGCATTCTTAACAGTCTTTCTAGTAGTTGCCAAGGAAATATTAGATGAT
CGCATAAAACGTTTTGAAGAACTTGAAAAATTGGGAATCCCTATTTTAGGCTCAATACCT
TTATCAAAAAATATAAAGTAG

>SP25A wzd (**SEQ ID NO: 9)**

ATGAATAATCGTATAGTAGACATCAGTTTGTTTGATTTATTGAATGTCTTATGGCTTCGT
AAATGGATCATTGTTTTTGTATCATTTGTTTTTGGATGTATTTCTTTTTTATATTATCAA
TTTTTTGTACCCAAAACCTATACAAGTGATACTCGTATATATGTGGTAAGCAAAGGGGAT
AATAAAAATCTGTCTAATGCAGATTTGCAAGCAGGAAGTGCTCTAACAAAGGATTACAAA
GAAATCATTTTATCAGATGAAGTGTTAGAAGAAACAATTTCTGATTTAAAACTAGAATCA
ACTTTGGAATCTCTAGCTTCGAAAATTAAAATTTCTATCCCTGCGGAAACTCGAATTATT
TCCATTAGTGTTACAAATACGAACTCAGATGAAGCTGCTCGAATAGCAAATGGGATAAGG
AAGGTAGCTGCCCTTAAAATCAAGGAAGTAACTCAAGTTACAGATGTCACCACTTTACAG
ACTGCACGCCCTCCCCAAACTCCGTCAGGACCGCATGTGCGAAAATCTACTACAGCAGGG
CTTGTTCTTGGTGCATTCTTAACAGTCTTTCTAGTAGTTGCCAAGGAAATATTAGATGAT
CGCATAAAACGTTTTGAAGAACTTGAAAAATTGGGAATCCCTATTTTAGGCTCAATACCT
TTATCAAAAAATATAAAGTAG

>SP38 wzd (**SEQ ID NO: 10)**

ATGAATAATCGTATAGTAGACATCAGTTTGTTTGATTTATTGAATGTCTTATGGCTTCGT
AAATGGATCATTGTTTTTGTATCATTTGTTTTTGGATGTATTTCTTTTTTATATTATCAA
TTTTTTGTACCCAAAACCTATACAAGTGATACTCGTATATATGTGGTAAGCAAAGGGGAT
AATAAAAATCTGTCTAATGCAGATTTGCAAGCAGGAAGTGCTCTAACAAAGGATTACAAA
GAAATCATTTTATCAGATGAAGTGTTAGAAGAAACAATTTCTGATTTAAAACTAGAATCA
ACTTTGGAATCTCTAGCTTCGAAAATTAAAATTTCTATCCCTGCGGAAACTCGAATTATT
TCCATTAGTGTTACAAATACGAACTCAGATGAAGCTGCTCGAATAGCAAATGGGATAAGG
AAGGTAGCTGCCCTTAAAATCAAGGAAGTAACTCAAGTTACAGATGTCACCACTTTACAG
ACTGCACGCCCTCCCCAAACTCCGTCAGGACCGCATGTGCGAAAATCTACTACAGCAGGG
CTTGTTCTTGGTGCATTCTTAACAGTCTTTCTAGTAGTTGCCAAGGAAATATTAGATGAT
CGCATAAAACGTTTTGAAGAACTTGAAAAATTGGGAATCCCTATTTTAGGCTCAATCCCT
TTATCAAAAAATATAAAGTAG

Figure 3 (continued)

>SP44 wzd (**SEQ ID NO: 11)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTATTTAAAACCTTGTGG
AAACGCAAGCTAATGATTTTATTAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCGCTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
GATCAAGGAGACAAGTCGGGGCTGACAAATCAGGATTTGCAGGCAGGATCTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGAATGTATTGGAAAAAGTAGCGACAAATTTG
AAGTTGGATATACCAGCCAAAACGTTAGCTAGAAAAGTTCAAGTGACTGTACCAGTTGAC
ACTCGTATCGTCTCAATCTCTGTCAAGGATAAACAGCCAGAGGAAGCCAGTCGTATCGCT
AATTCTCTACGAGAAGTTGCTGCAGAAAAGATCATCGCTGTAACGCGAGTATCTGATGTG
ACGACACTTGAAGAAGCGCGACCAGCTACGACTCCCTCTTCTCCAAATGTTGGACGCAAT
TCCTTGTTTGGTTTTTTTGGAGGAGCAGTCGTAACAGTAATTGCTGTTCTTTTGATTGAG
TTGTTCGACATCCGTGTGAAACGTCCTGAAGATGTCGAAGATGTACTGCAAATTCCACTT
TTAGGGGTCGTTCCAGATTTGGACAAAATGAAATAG

>SP45 wzd (**SEQ ID NO: 12)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTAGTTAAAAGCTTGTGG
AAACGCAAGCTAATGATTTTAATAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCACTTTTATTGTTAAGCCAAAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
AATCAAGGAGACAAGTCGGGGTTGACAAATCAGGATTTGCAGGCAGGATCTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGGATGTTTTGGAGGAAGTTGTTTCTGATTTA
AAACTAGATTTGACGCCAAAAGGTTTGGCTAATAAAATTAAAGTGACAGTACCAGTTGAT
ACCCGTATTGTCTCTGTTTCAGTTAATGATCGAGTTCCTGAAGAGGCAAGCCGTATCGCT
AACTCTTTGAGAGAAGTAGCTGCTCAAAAAATTATCAGTATTACTCGTGTTTCTGATGTG
ACAACACTGGAGGAGGCAAGACCGGCGATATCACCGTCTTCGCCAGATATTAAACGCAAT
ACACTAATTGGTTTTTTGGCAGGGGTGATTGGAACTAGTGTTATAGTTCTTCTTCTTGAA
CTTTTGGACACTCGTGTGAAACGTCCGGAAGATATCGAAGATACACTGCAGATGACACTT
TTGGGAGTTGTACCAAACTTGGATAAGTTGAAATAG

>SP46 wzd (**SEQ ID NO: 13)**

ATGATGAAAGAACAAAACACGATAGAAATCGATGTATTTCAATTATTTAAAACCTTGTGG
AAACGCAAGCTAATGATTTTATTAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATAT
AGCGCTTTTATTGTTAAGCCAGAATATACGAGTACCACGCGAATTTACGTAGTGAATCGC
GATCAAGGAGACAAGTCGGGGCTGACAAATCAGGATTTGCAGGCAGGAACTTATCTGGTA
AAAGACTACCGTGAGATTATCCTTTCGCAGAATGTATTGGAAAAAGTAGCGACAAATTTG
AAGTTGGATATACCAGCCAAAACGTTAGCTAGAAAAGTTCAAGTGACTGTACCAGTTGAC
ACTCGTATCGTCTCAATCTCTGTCAAGGATAAACAGCCAGAGGAAGCCAGTCGTATCGCT
AATTCTCTACGAGAAGTTGCTGCAGAAAAGATCATCGCTGTAACGCGAGTATCTGATGTA
ACGACACTTGAAGAAGCGCGACCAGCTACGACTCCCTCTTCTCCAAATGTTCGTCTCAAT
TCCTTGTTTGGTTTTTTTGGAGGAGCAGTCGTAACAGTAATTGCTGTTCTTTTGATTGAG
TTGTTCGACATCCGTGTGAAACGTCCTGAAGATGTCGAAGATGTACTGCAAATTCCACTT
TTAGGGGTCGTTCCAGATTTGAACAAAATGAAATAG

Figure 3 (continued)

>SP1 wze (**SEQ ID NO: 14)**

ATGCCGACATTAGAAATAGCACAAAAAAAACTGGAGTTCATTAAGAAGGCAGAAGAATAT
TACAATACCTTGTGTACAAATATACAGTTGAGCGGAGATAAACTAAAAGTAATTTCCGTT
ACTTCTGTTAATCCGGGGGAAGGAAAAACAACTACTTCCGTAAATATAGCAATGTCATTT
GCGCGTGCAGGCTATAAAACTCTTTTGTTCGATGGCGATATTCGAAATTCAGTTATGTCA
GGATTTTTTAAATCTCGTGAAAAAATTACAGGGCTAACAGAATTTTTATCTGGGACAGCT
GATTTATCTCACGGTTTATGTGATACAAATATTGAAAATTTATTTGTAGTTCAATCGGGA
TCTGTATCACCAAACCCTACAGCCTTGTTACAAAGTAAAAATTTTAATGATATGATTGAA
ACATTGCGTAAATATTTTGATTATATCATTGTTGATACACCACCTATTGGAATTGTTATT
GATGCGGCAATTATCACTCAAAAGTGTGATGCGTCCATCTTGATAACAGCAACAGGTGAG
GCGAATAAACGTGATGTCCAAAAAGCGAAACAACAATTAGAACAAACAGGGGAACTGTTC
CTGGGAGTTGTTTTAAATAAATTGGATATCTCGGTTGATAAGTATGGAGTTTACGGTCGT
AAATAG

>SP4 wze (**SEQ ID NO: 15)**

ATGCCGACATTAGAAATAGCACAAAAAAAACTGGAGTTCATTAAGAAGGCAGAAGAATAT
TACAATGCCTTGTGTACAAATATACAGTTGAGCGGAGATAAACTAAAAGTAATTTCCGTT
ACTTCTGTTAACCCTGGGGAAGGAAAAACAACTACTTCCATAAATATAGCATGGTCGTTT
GCGCGTGCAGGCTATAAAACTCTTTTGATCGATGGCGATACTCGAAATTCAGTTATGTTA
GGAGTTTTTAAATCTCGTGAAAAAATTACAGGGCTAACAGAATTTTTATCTGGGACAGCT
GATTTATCTCACGGTTTATGTGATACAAATATTGAAAATTTATTTGTAGTTCAATCGGGA
TCTGTATCACCAAACCCTACAGCCTTGTTACAAAGTAAAAATTTTAATGATATGATTGAA
ACATTGCGTAAATATTTTGATTATATCATTATTGATACACCGCCTATTGGAATTGTTATT
GATGCGGCAATTATCACTCAAAAGTGTGATGCGTCCATCTTGGTAACAGCAACAGGTGAG
GCGAATAAACGTGATATCCAAAAAGCGAAACAACAATTAAAACAAACAGGGAAACTGTTC
CTAGGAGTTGTTTTAAATAAATTGGATATCTCGGTTAATAAGTATGGAGTTTACGGTTCC
TATGGAAATTATGGTAAAAAATAA

>SP5 wze (**SEQ ID NO: 16)**

ATGCCGACATTAGAAATAGTACAAAAAAAACTGGAGTTCATTAAGAAGACAGAGGAATAT
TACAATGCCTTGTGTACAAACATACAGTTGAGCGGAGATAAACTAAAAGTGATTTCCGTT
ACTTCTGTTAACCCTGGAGAAGGGAAAACAACTACTTCTATAAATATAGCATGGTCTTTT
GCGCGTGCAGGATATAAAACTCTTTTGATAGATGGCGATATCCGAAATTCGGTGATGTCA
GGAGTTTTTAAATCTCGTGAAAAAATTACAGGCTTAACAGAATTTTTATCTGGTACAACT
GATTTATCTCACGGTTTATGTGATACAAATATTGAAAATTTATTTGTAGTTCAATCGGGA
GCTGTATCACCAAACCCTACAGCCTTGTTACAAAGTAAAAATTTTAATGATATGATTGAA
ACCTTGCGTAAATATTTTGATTATATCATTGTTGATACAGCACCTATTGGAGTTGTTATT
GATTCGGCAATTATCACTCAAAAGTGTGATGCGTCTATCTTGGTAACGGCAACGGGTGAG
GTGAATAAACGTGATGTCCCAAAAGCGAAACAACAATTAGAACAAACAGGGAAACTATTT
TTAGGTGTCGTTCTCAATAAATTTAATGTTCAACATGAAAAATATGGTTCTTACGGTAAT
TATGGTAAAAGATAA

Figure 3 (continued)

>SP12F wze (**SEQ ID NO: 17)**

ATGCCAACATTAGAAATCTCACAGGCAAAATTGGATTTTGTAAAAAAGGCAGAGGAAAAT
TATAACGCTTTGTGCACGAACCTACAGTTAAGTGGAGATGATTTGAAAGTATTTTCTATC
ACTTCTGTGAAACAAGGAGAAGGAAAATCAACGACTTCCACCAATATCGCTTGGGCTTTT
GCGCGTGCAGGTTACAAAACGCTGCTGATTGATGGAGATATTCGCAATTCTGTTATGTTA
GGTGTCTTTAAAGCAAGAGATAAGATTACAGGCCTGACAGAATTTTTATCAGGAACTACA
GACCTATCACAAGGGCTTTGTGATACCAATATCGAAAATCTCTTTGTAATTCAGGCTGGC
TCTGTGTCACCGAATCCGACAGCTCTTCTTCAAAGTAAGAATTTCAGTACAATGCTTGAA
ACCTTGCGTAAATATTTTGACTACATCATTGTAGATACTGCTCCTGTCGGTGTCGTGATT
GATGCGGCTATTATTACGCAAAAATGCGATGCTTCTATTTTAGTGACGAAGGCAGGTGAA
ATAAATCGACGGGATATTCAAAAAGCAAAAGAACAGTTGGAACACACAGGGAAGCCGTTT
TTGGGAGTTGTGTTGAATAAATTCGATACTTCAGTAGACAAATACGGTTCTTATGGAAAT
TACGGAAAAAAATAA

>SP12A wze (**SEQ ID NO: 18)**

ATGCCAACATTAGAAATCTCACAGGCAAAATTGGATTTTGTAAAAAAGGCAGAGGAAAAT
TATAACGCTTTGTGCACGAACCTACAGTTAAGTGGAGATGATTTGAAAGTATTTTCTATC
ACTTCTGTGAAACAAGGAGAAGGAAAATCAACGACTTCCACCAATATTGCTTGGGCTTTT
GCGCGTGCAGGTTACAAAACGCTGCTGATTGATGGAGATATTCGCAATTCTGTTATGTTA
GGTGTCTTTAAAGCAAGAGATAAGATTACAGGCCTGACAGAATTTTTATCAGGAACTACA
GACCTATCACAAGGGCTTTGTGATACCAATATCGAAAATCTCTTTGTAATTCAGGCTGGC
TCTGTGTCACCGAATCCGACAGCTCTTCTTCAAAGTAAGAATTTCAGTACAATGCTTGAA
ACCTTGCGTAAATATTTTGACTATATCATTGTAGATACTGCTCCTGTCGGTGTCGTGATT
GATGCGGCTATTATTACGCAAAAATGCGATGCTTCTATTTTAGTGACGAAGGCAGGTGAA
ATAAATCGACGGGATATTCAAAAAGCAAAAGAACAGTTGGAACACACAGGGAAGCCGTTT
TTGGGAGTTGTGTTGAATAAATTCGATACTTCAGTAGACAAATACGGTTCTTATGGAAAT
TACGGAAAAAAATAA

>SP12B wze (**SEQ ID NO: 19)**

ATGCCAACATTAGAAATCTCACAGGCAAAATTGGATTTTGTAAAAAAGGCAGAGGAAAAT
TATAACGCTTTGTGCACGAACCTACAGTTAAGTGGAGATGATTTGAAAGTATTTTCTATC
ACTTCTGTGAAACAAGGAGAAGGAAAATCAACGACTTCCACCAATATTGCTTGGGCTTTT
GCGCGTGCAGGTTACAAAACGCTGCTGATTGATGGAGATATTCGCAATTCTGTTATGTTA
GGTGTCTTTAAAGCAAGAGATAAGATTACAGGCCTGACAGAATTTTTATCAGGAACTACA
GACCTATCACAAGGGCTTTGTGATACCAATATCGAAAATCTCTTTGTAATTCAGGCTGGC
TCTGTGTCACCGAATCCGACAGCTCTTCTTCAAAGTAAGAATTTCAGTACAATGCTTGAA
ACCTTGCGTAAATATTTTGACTATATCATTGTAGATACTGCTCCTGTCGGTGTCGTGATT
GATGCGGCTATTATTACGCAAAAATGCGATGCTTCTATTTTAGTGACGAAGGCAGGTGAA
ATAAATCGACGGGATATTCAAAAAGCAAAAGAACAGTTGGAACACACAGGGAAGCCGTTT
TTGGGAGTTGTGTTGAATAAATTCGATACTTCAGTAGACAAATACGGTTCTTATGGAAAT
TACGGAAAAAAATAA

Figure 3 (continued)

>SP25F wze (**SEQ ID NO: 20)**

ATGGGAAAATTAGAACTAACACGTAAGTACTATCAATTGTACAATGAAACGAAAGAATAT
TTTAATGCCTTACAGACCAATATTCAACTTAGTGGTGATGGACTGAAAGTCATTTCTGTT
ACTTCCACAGTAAAAAACGAAGGCAAATCGACTATATCAGTTGGCTTGTCCATGTCCTTG
GCCTTATCAGGCTATAAAGTCCTGTTACTGGATGTGGATTTACGTAATTCAGTGATGTCA
GGAATGTTTCATTCTAGGGAAAGAATAACAGGTTTGACTGAATATTTATCTGGCAAAGCG
GGAATGGCGGATTCTATGCAGGAAACTACTTTAGATAATCTCTATATATTACAAGCAGGA
GTCGTTTCACCGAATCCCTTATCTCTATTAAGGAGTGTTAAGTTTGAAACCTTGATAGAT
TCTATGAAAAAGTACTTTGACTATATTATAGTAGATACTCCACCAATCGGCCAAGTAATA
GATGCTGCAATTATTGCCCAACAATGTGATGGGATATTTTTGGTCACTGAAATGAGGAAA
ACGACACGCAGACAAATTTTTTCTTCCCTCCAGCAATTGGAACAAACTGGAGTTCCCGTT
TTGGGCCTTGTGGCCAATAAAGCAATTATAAAAAAAGATAGCTATGGCTATTATGGAAAA
AACACCGCGAGCGAGACTGAAAGATATTGA

>SP25A wze (**SEQ ID NO: 21)**

ATGGGAAAATTAGAACTAACACGTAAGTACTATCAATTGTACAATGAAACGAAAGAATAT
TTTAATGCCTTACAGACCAATATTCAACTTAGTGGTGATGGACTGAAAGTCATTTCTGTT
ACTTCCACAGTAAAAAACGAAGGCAAATCGACTATATCAGTTGGCTTGTCCATGTCCTTG
GCCTTATCAGGCTATAAAGTCCTGTTACTGGATGTGGATTTACGTAATTCAGTGATGTCA
GGAATGTTTCATTCTAGGGAAAGAATAACAGGTTTGACTGAATATTTATCTGGCAAAGCG
GGAATGGCGGATTCTATGCAGGAAACTACTTTAGATAATCTCTATATATTACAAGCAGGA
GTCGTTTCACCGAATCCCTTATCTCTATTAAGGAGTGTTAAGTTTGAAACCTTGATAGAT
TCTATGAAAAAGTACTTTGACTATATTATAGTAGATACTCCACCAATCGGCCAAGTAATA
GATGCTGCAATTATTGCCCAACAATGTGATGGGATATTTTTGGTCACTGAAATGAGGAAA
ACGACACGCAGACAAATTTTTTCTTCCCTCCAGCAATTGGAACAAACTGGAGTTCCCGTT
TTGGGCCTTGTGGCCAATAAAGCAATTATAAAAAAAGATAGCTATGGCTATTATGGAAAA
AACACCGCGAGCGAGACTGAAAGATATTGA

>SP38 wze (**SEQ ID NO: 22)**

ATGGGAAAATTAGAACTAACACGTAAGTACTATCAATTGTACAATGAAACGAAAGAATAT
TTTAATGCCTTACAGACCAATATTCAACTTAGTGGTGATGGACTGAAAGTCATTTCTGTT
ACTTCCACAGTAAAAAACGAAGGCAAATCGACTATATCAGTTGGCTTGTCCATGTCCTTG
GCCTTATCAGGCTATAAAGTCCTGTTACTGGATGTGGATTTACGTAATTCAGTGATGTCA
GGAATGTTTCATTCTAGGGAAAGAATAACAGGTTTGACTGAATATTTATCTGGCAAAGCG
GGAATGGCGGATTCTATGCAGGAAACTACTTTAGATAATCTCTATATATTACAAGCAGGA
GTCGTTTCACCGAATCCCTTATCTCTATTAAGGAGTGTTAAGTTTGAAACCTTGATAGAT
TCTATGAAAAAGTACTTTGACTATATTATAGTAGATACTCCACCAATCGGCCAAGTAATA
GATGCTGCAATTATTGCCCAACAATGTGATGGGATATTTTTGGTCACTGAAATGAGGAAA
ACGACACGCAGACAAATTTTTTCTTCCCTCCAGCAATTGGAACAAACTGGAGTTCCCGTT
TTGGGCCTTGTGGCCAATAAAGCAATTATAAAAAAAGATAGCTATGGCTATTATGGAAAA
AACACCGCGAGCGAGACTGAAAGATATTGA

Figure 3 (continued)
>SP44 wze (**SEQ ID NO: 23)**

ATGCCAACATTAGAAATCTCACAGGCAAAATTGGATTTTGTAAAAAAGGCAGAGGAAAAT
TATAACGCTTTGTGCACGAACCTACAGTTAAGTGGAGATGATTTGAAAGTATTTTCTATC
ACTTCTGTGAAACAAGGAGAAGGAAAATCAACGACTTCCATCAATATCGCTTGGGCTTTT
GCGCGTGCAGGTTACAAAACGCTGCTGATTGATGGAGATATTCGCAATTCTGTTATGTTA
GGTGTCTTTAAAGCAAGAGATAAGATTACAGGCCTGACAGAATTTTTATCAGGAACTACA
GACCTATCACAAGGGCTTTGTGATACCAATATCGAAAATCTCTTTGTAATTCAGGCTGGC
TCTGTGTCACCGAATCCGACAGCTCTTCTTCAAAGTAAGAATTTCAGTACAATGCTTGAA
ACCTTGCGTAAATATTTTGACTACATCATTGTAGATACTGCTCCTGTCGGTGTCGTGATT
GATGCGGCTATTATTACGCAAAAATGCGATGCTTCTATTTTAGTGACGAAGGCAGGTGAA
ATAAATCGACGGGATATTCAAAAAGCAAAAGAACAGTTGGAACACACAGGGAAGCCGTTT
TTGGGAGTTGTGTTGAATAAATTCGATACTTCAGTAGACAAATACGGTTCTTATGGAAAT
TACGGAAAAAAATAA

>SP45 wze (**SEQ ID NO: 24)**

ATGCCGACATTAGAAATAGCACAAAAAAAACTGGAGTTCGTTAAGAAGGCAGAAGAATAT
TACAATGCCTTGTGTACAAATATACAGTTGAGCGGAGATAAACTAAAAGTAATTTCCGTT
ACTTCTGTTAACCCTGGGGAAGGAAAAACAACTACTTCCGTAAATATAGCAATATCGTTT
GCGTGTGCAGGCTATAAAACTCTTTTGATCGATGGTGATACTCGAAATTCAGTTATGTCA
GGATTTTTTAAATCTCGTGAAAAAATTACAGGGCTAACAGAATTTTTATCTGGGACAGCT
GATTTATCTCACGGTTTATGTGATACAAATATTGAAAATTTATTTGTAGTTCAATCGGGA
ACTGTATCACCAAACCCTACAGCCTTGTTACAAAGTAAAAATTTTAATGATATGATTGAA
ACATTGCGTAAATATTTTGATTATATCATTGTTGATACAGCACCTATTGGAATTGTTATT
GATGCGGCAATTATCACTCAAAAGTGTGATGCGTCCATCTTGGTAACAGCAACAGGTGAG
GTGAATAAACGTGATATCCAAAAAGCGAAACAACAATTAAAACAAACAAAGAAACTGTTC
CTAGGAGTTGTTTTAAATAAATTGGATATCTCGGTTGATAAGTATGGAATTTACGGTTCC
TATGGAAATTATGGTAAAAGATAA

>SP46 wze (**SEQ ID NO: 25)**
ATGCCAACATTAGAAATCTCACAGGCAAAATTGGATTTTGTAAAAAAGGCAGAGGAAAAT
TATAACGCTTTGTGCACGAACCTACAGTTAAGTGGAGATGATTTGAAAGTATTTTCTATC
ACTTCTGTGAAACAAGGAGAAGGAAAATCAACGACTTCCACCAATATTGCTTGGGCTTTT
GCGCGTGCAGGTTACAAAACGCTGCTGATTGATGGAGATATTCGCAATTCTGTTATGTTA
GGTGTCTTTAAAGCAAGAGATAAGATTACAGGCCTGACAGAATTTTTATCAGGAACTACA
GACCTATCACAAGGGCTTTGTGATACCAATATCGAAAATCTCTTTGTAATTCAGGCTGGC
TCTGTGTCACCGAATCCGACAGCTCTTCTTCAAAGTAAGAATTTCAGTACAATGCTTGAA
ACCTTGCGTAAATATTTTGACTATATCATTGTAGATACTGCTCCTGTCGGTGTCGTGATT
GATGCGGCTATTATTACGCAAAAATGCGATGCTTCTATTTTAGTGACGAAGGCAGGTGAA
ATAAATCGACGGGATATTCAAAAAGCAAAAGAACAGTTGGAACACACAGGGAAGCCGTTT
TTGGGAGTTGTGTTGAATAAATTCGATACTTCAGTAGACAAATACGGTTCTTATGGAAAT
TACGGAAAAAAATAA

Figure 3 (continued)
Amino acid sequences: wzd

>SP1 wzd (**SEQ ID NO: 33)**

MKEQNTLEIDVLQLFRALWKRKLVILLVAIITSSVAFAYSTFVIKPEFTSTTRIYVVNRNQGEKPGLT
NQDLQAGAYLVKDYREIVLSQDVLEEVISDLKLDLVPKGLANKIKVTVPVDTRIVSVSVSDRVPEEAS
RIANSLREVAAQKIISITRVSDVTTLEEARPAISPSSPNIKRNTLIGFLAGVIGTSVIVLLLELLDTR
VKRPEDIENALQMTLLGVVPNLGKLK

>SP4 wzd (**SEQ ID NO: 34)**

MMKEQNTIEIDVFQLVKSLWKRKLMILIVALVTGAGAFAYSTFIVKPEYTSTTRIYVVNRNQGDKPGL
TNQDLQAGTYLVKDYREIILSQDVLEEVVSDLKLDLTPKGLANKIKVTVPVDTRIVSISVNDRVPEEA
SRIANSLREVAAQKIISITRVSDVTTLEEARPAISPSSPNIKRNTLIGFLAGVIGTSVIVLHLELLDT
RVKRPEDIENTLQMTLLGVVPNLGKLK

>SP5 wzd (**SEQ ID NO: 35)**

MEKQNTLEIDVLQLFRSLWKRKLVILLVAIITSSVAFAYSTFVIKPEFTSTTRIYVVSRDQGEKSGLT
NQDLQAGSYLTKDYREIILSQDVLEEVVSDLKLDLTPKGLANKIKVAVPVDTRIVSISVNDQVPEGAS
RIANSLREVAAQKIISITRVSDVTTLEEARPAISPSSPNIKLNTLIGFLVGLIVINVTVLLLELLDTR
VKRPEDIEDALQMTLLGVVPNLDKLK

>SP12F wzd (**SEQ ID NO: 36)**

MMKEQNTIEIDVFQLFKTLWKRKLMILLVALVTGAGAFAYSAFIVKPEYTSTTRIYVVNRDQGDKSGL
TNQDLQAGSYLVKDYREIILSQNVLEKVATNLKLDIPAKTLARKVQVTVPVDTRIVSISVKDKQPEEA
SRIANSLREVAAEKIIAVTRVSDVTTLEEARPATTPSSPNVGRNSLFGFFGGAVVTVIAVLLIELFDI
RVKRPEDVEDVLQIPLLGVVPDLDKMK

>SP12A wzd (**SEQ ID NO: 37)**

MMKEQNTIEIDVFQLFKTLWKRKLMILLVALVTGAGAFAYSAFIVKPEYTSTTRIYVVNRDQGDKSGL
TNQDLQAGTYLVKDYREIILSQNVLEKVATNLKLDIPAKTLARKVQVTVPVDTRIVSISVKDKQPEEA
SRIANSLREVAAEKIIAVTRVSDVTTLEEARPATTPSSPNVRLNSLFGFFGGAVVTVIAVLLIELFDI
RVKRPEDVEDVLQIPLLGVVPDLNKMK

>SP12B wzd (**SEQ ID NO: 38)**
MMKEQNTIEIDVFQLFKTLWKRKLMILLVALVTGAGAFAYSAFIVKPEYTSTTRIYVVNRDQGDKSGL
TNQDLQAGSYLVKDYREIILSQNVLEKVATNLKLDIPAKTLARKVQVTVPVDTRIVSISVKDKQPEEA
SRIANSLREVAAEKIIAVTRVSDVTTLEEARPATTPSSPNVGRNSLFGFFGGAVVTVIAVLLIELFDI
RVKRPEDVEDVLQIPLLGVVPDLDKMK

Figure 3 (continued)

>SP25F wzd (**SEQ ID NO: 39)**

MNNRIVDISLFDLLNVLWLRKWIIVFVSFVFGCISFLYYQFFVPKTYTSDTRIYVVSKGDNKNLSNAD
LQAGSALTKDYKEIILSDEVLEETISDLKLESTLESLASKIKISIPAETRIISISVTNTNSDEAARIA
NGIRKVAALKIKEVTQVTDVTTLQTARPPQTPSGPHVRKSTTAGLVLGAFLTVFLVVAKEILDDRIKR
FEELEKLGIPILGSIPLSKNIK

>SP25A wzd (**SEQ ID NO: 40)**

MNNRIVDISLFDLLNVLWLRKWIIVFVSFVFGCISFLYYQFFVPKTYTSDTRIYVVSKGDNKNLSNAD
LQAGSALTKDYKEIILSDEVLEETISDLKLESTLESLASKIKISIPAETRIISISVTNTNSDEAARIA
NGIRKVAALKIKEVTQVTDVTTLQTARPPQTPSGPHVRKSTTAGLVLGAFLTVFLVVAKEILDDRIKR
FEELEKLGIPILGSIPLSKNIK

>SP38 wzd (**SEQ ID NO: 41)**

MNNRIVDISLFDLLNVLWLRKWIIVFVSFVFGCISFLYYQFFVPKTYTSDTRIYVVSKGDNKNLSNAD
LQAGSALTKDYKEIILSDEVLEETISDLKLESTLESLASKIKISIPAETRIISISVTNTNSDEAARIA
NGIRKVAALKIKEVTQVTDVTTLQTARPPQTPSGPHVRKSTTAGLVLGAFLTVFLVVAKEILDDRIKR
FEELEKLGIPILGSIPLSKNIK

>SP44 wzd (**SEQ ID NO: 42)**

MMKEQNTIEIDVFQLFKTLWKRKLMILLVALVTGAGAFAYSAFIVKPEYTSTTRIYVVNRDQGDKSGL
TNQDLQAGSYLVKDYREIILSQNVLEKVATNLKLDIPAKTLARKVQVTVPVDTRIVSISVKDKQPEEA
SRIANSLREVAAEKIIAVTRVSDVTTLEEARPATTPSSPNVGRNSLFGFFGGAVVTVIAVLLIELFDI
RVKRPEDVEDVLQIPLLGVVPDLDKMK

>SP45 wzd (**SEQ ID NO: 43)**

MMKEQNTIEIDVFQLVKSLWKRKLMILIVALVTGAGAFAYSTFIVKPKYTSTTRIYVVNRNQGDKSGL
TNQDLQAGSYLVKDYREIILSQDVLEEVVSDLKLDLTPKGLANKIKVTVPVDTRIVSVSVNDRVPEEA
SRIANSLREVAAQKIISITRVSDVTTLEEARPAISPSSPDIKRNTLIGFLAGVIGTSVIVLLLELLDT
RVKRPEDIEDTLQMTLLGVVPNLDKLK

>SP46 wzd (**SEQ ID NO: 44)**

MMKEQNTIEIDVFQLFKTLWKRKLMILLVALVTGAGAFAYSAFIVKPEYTSTTRIYVVNRDQGDKSGL
TNQDLQAGTYLVKDYREIILSQNVLEKVATNLKLDIPAKTLARKVQVTVPVDTRIVSISVKDKQPEEA
SRIANSLREVAAEKIIAVTRVSDVTTLEEARPATTPSSPNVRLNSLFGFFGGAVVTVIAVLLIELFDI
RVKRPEDVEDVLQIPLLGVVPDLNKMK

Wze

Figure 3 (continued)

>SP1 wze (**SEQ ID NO: 45)**

MPTLEIAQKKLEFIKKAEEYYNTLCTNIQLSGDKLKVISVTSVNPGEGKTTTSVNIAMSFARAGYKTL
LFDGDIRNSVMSGFFKSREKITGLTEFLSGTADLSHGLCDTNIENLFVVQSGSVSPNPTALLQSKNFN
DMIETLRKYFDYIIVDTPPIGIVIDAAIITQKCDASILITATGEANKRDVQKAKQQLEQTGELFLGVV
LNKLDISVDKYGVYGRK

>SP4 wze (**SEQ ID NO: 46)**

MPTLEIAQKKLEFIKKAEEYYNTLCTNIQLSGDKLKVISVTSVNPGEGKTTTSVNIAMSFARAGYKTL
LFDGDIRNSVMSGFFKSREKITGLTEFLSGTADLSHGLCDTNIENLFVVQSGSVSPNPTALLQSKNFN
DMIETLRKYFDYIIVDTPPIGIVIDAAIITQKCDASILITATGEANKRDVQKAKQQLEQTGELFLGVV
LNKLDISVDKYGVYGRK

>SP5 wze (**SEQ ID NO: 47)**

MPTLEIVQKKLEFIKKTEEYYNALCTNIQLSGDKLKVISVTSVNPGEGKTTTSINIAWSFARAGYKTL
LIDGDIRNSVMSGVFKSREKITGLTEFLSGTTDLSHGLCDTNIENLFVVQSGAVSPNPTALLQSKNFN
DMIETLRKYFDYIIVDTAPIGVVIDSAIITQKCDASILVTATGEVNKRDVPKAKQQLEQTGKLFLGVV
LNKFNVQHEKYGSYGNYGKR

>SP12F wze (**SEQ ID NO: 48)**

MPTLEISQAKLDFVKKAEENYNALCTNLQLSGDDLKVFSITSVKQGEGKSTTSTNIAWAFARAGYKTL
LIDGDIRNSVMLGVFKARDKITGLTEFLSGTTDLSQGLCDTNIENLFVIQAGSVSPNPTALLQSKNFS
TMLETLRKYFDYIIVDTAPVGVVIDAAIITQKCDASILVTKAGEINRRDIQKAKEQLEHTGKPFLGVV
LNKFDTSVDKYGSYGNYGKK

>SP12A wze (**SEQ ID NO: 49)**

MPTLEISQAKLDFVKKAEENYNALCTNLQLSGDDLKVFSITSVKQGEGKSTTSTNIAWAFARAGYKTL
LIDGDIRNSVMLGVFKARDKITGLTEFLSGTTDLSQGLCDTNIENLFVIQAGSVSPNPTALLQSKNFS
TMLETLRKYFDYIIVDTAPVGVVIDAAIITQKCDASILVTKAGEINRRDIQKAKEQLEHTGKPFLGVV
LNKFDTSVDKYGSYGNYGKK

>SP12B wze (**SEQ ID NO: 50)**

MPTLEISQAKLDFVKKAEENYNALCTNLQLSGDDLKVFSITSVKQGEGKSTTSTNIAWAFARAGYKTL
LIDGDIRNSVMLGVFKARDKITGLTEFLSGTTDLSQGLCDTNIENLFVIQAGSVSPNPTALLQSKNFS
TMLETLRKYFDYIIVDTAPVGVVIDAAIITQKCDASILVTKAGEINRRDIQKAKEQLEHTGKPFLGVV
LNKFDTSVDKYGSYGNYGKK

Figure 3 (continued)

>SP25F wze (**SEQ ID NO: 51)**

MGKLELTRKYYQLYNETKEYFNALQTNIQLSGDGLKVISVTSTVKNEGKSTISVGLSMSLALSGYKVL
LLDVDLRNSVMSGMFHSRERITGLTEYLSGKAGMADSMQETTLDNLYILQAGVVSPNPLSLLRSVKFE
TLIDSMKKYFDYIIVDTPPIGQVIDAAIIAQQCDGIFLVTEMRKTTRRQIFSSLQQLEQTGVPVLGLV
ANKAIIKKDSYGYYGKNTASETERY

>SP25A wze (**SEQ ID NO: 52)**

MGKLELTRKYYQLYNETKEYFNALQTNIQLSGDGLKVISVTSTVKNEGKSTISVGLSMSLALSGYKVL
LLDVDLRNSVMSGMFHSRERITGLTEYLSGKAGMADSMQETTLDNLYILQAGVVSPNPLSLLRSVKFE
TLIDSMKKYFDYIIVDTPPIGQVIDAAIIAQQCDGIFLVTEMRKTTRRQIFSSLQQLEQTGVPVLGLV
ANKAIIKKDSYGYYGKNTASETERY

>SP38 wze (**SEQ ID NO: 53)**

MGKLELTRKYYQLYNETKEYFNALQTNIQLSGDGLKVISVTSTVKNEGKSTISVGLSMSLALSGYKVL
LLDVDLRNSVMSGMFHSRERITGLTEYLSGKAGMADSMQETTLDNLYILQAGVVSPNPLSLLRSVKFE
TLIDSMKKYFDYIIVDTPPIGQVIDAAIIAQQCDGIFLVTEMRKTTRRQIFSSLQQLEQTGVPVLGLV
ANKAIIKKDSYGYYGKNTASETERY

>SP44 wze (**SEQ ID NO: 54)**

MPTLEISQAKLDFVKKAEENYNALCTNLQLSGDDLKVFSITSVKQGEGKSTTSINIAWAFARAGYKTL
LIDGDIRNSVMLGVFKARDKITGLTEFLSGTTDLSQGLCDTNIENLFVIQAGSVSPNPTALLQSKNFS
TMLETLRKYFDYIIVDTAPVGVVIDAAIITQKCDASILVTKAGEINRRDIQKAKEQLEHTGKPFLGVV
LNKFDTSVDKYGSYGNYGKK

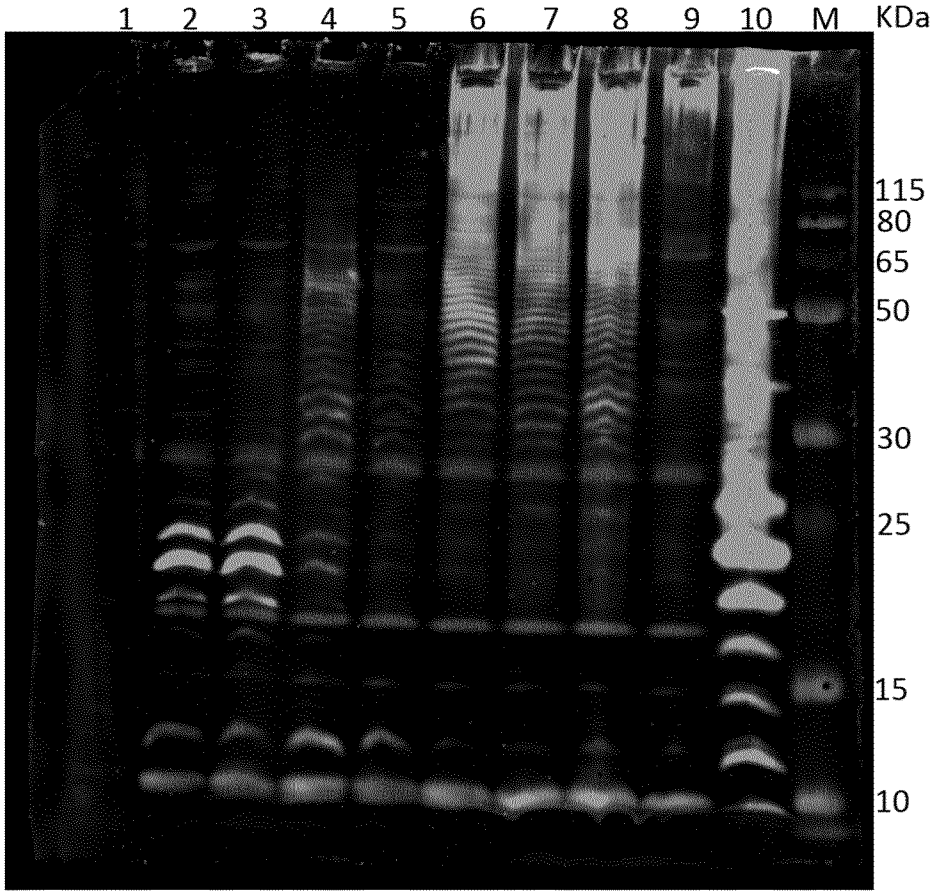
>SP45 wze (**SEQ ID NO: 55)**

MPTLEIAQKKLEFVKKAEEYYNALCTNIQLSGDKLKVISVTSVNPGEGKTTTSVNIAISFACAGYKTL
LIDGDTRNSVMSGFFKSREKITGLTEFLSGTADLSHGLCDTNIENLFVVQSGTVSPNPTALLQSKNFN
DMIETLRKYFDYIIVDTAPIGIVIDAAIITQKCDASILVTATGEVNKRDIQKAKQQLKQTKKLFLGVV
LNKLDISVDKYGIYGSYGNYGKR

>SP46 wze (**SEQ ID NO: 56)**

MPTLEISQAKLDFVKKAEENYNALCTNLQLSGDDLKVFSITSVKQGEGKSTTSTNIAWAFARAGYKTL
LIDGDIRNSVMLGVFKARDKITGLTEFLSGTTDLSQGLCDTNIENLFVIQAGSVSPNPTALLQSKNFS
TMLETLRKYFDYIIVDTAPVGVVIDAAIITQKCDASILVTKAGEINRRDIQKAKEQLEHTGKPFLGVV
LNKFDTSVDKYGSYGNYGKK

Figure 4



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/073714

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K47/64 C12N15/63
ADD.

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)
A61K C12N

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)
EPO-Internal, WPI Data, BIOSIS, COMPENDEX, EMBASE, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/020499 A2 (GLYCOVAXYN AG [CH]) 11 February 2016 (2016-02-11) the whole document, in particular claims and figures and [0078]	1-30
X	WO 2014/072405 A1 (GLYCOVAXYN AG [CH]) 15 May 2014 (2014-05-15) the whole document, in particular the claims	1-30
X	WO 2007/071707 A2 (GLAXOSMITHKLINE BIOLOG SA [BE]; BIEMANS RALPH LEON [BE]; GARCON NATHAL) 28 June 2007 (2007-06-28) the whole document, in particular the claims	16-20, 24,25

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 November 2018	03/12/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Bassias, Ioannis

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/110942 A2 (PFIZER [US]) 30 July 2015 (2015-07-30) the whole document, in particular the claims	16-20, 24,25
X	US 2007/231340 A1 (HAUSDORFF WILLIAM P [BE] ET AL) 4 October 2007 (2007-10-04) the whole document, in particular the claims	16-20, 24,25
X	US 2004/202668 A1 (BOUTRIAU DOMINIQUE [BE] ET AL) 14 October 2004 (2004-10-14) the whole document, in particular the claims	16-20, 24,25
A	EMILY J. KAY ET AL: "Recombinant expression of Streptococcus pneumoniae capsular polysaccharides in Escherichia coli", OPEN BIOLOGY, vol. 6, no. 4, 1 April 2016 (2016-04-01), page 150243, XP055525931, DOI: 10.1098/rsob.150243	1-30
A	CALVIN C. DANIELS ET AL: "A Review of Pneumococcal Vaccines: Current Polysaccharide Vaccine Recommendations and Future Protein Antigens", THE JOURNAL OF PEDIATRIC PHARMACOLOGY AND THERAPEUTICS, vol. 21, no. 1, 1 January 2016 (2016-01-01), pages 27-35, XP055526395, ISSN: 1551-6776, DOI: 10.5863/1551-6776-21.1.27	1-30

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/073714

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016020499 A2	11-02-2016	AR 104469 A1	26-07-2017
		BR 112017002376 A2	05-12-2017
		CA 2956188 A1	11-02-2016
		CN 106795545 A	31-05-2017
		EP 3177325 A2	14-06-2017
		JP 2017523794 A	24-08-2017
		SG 11201700325S A	27-02-2017
		US 2017232093 A1	17-08-2017
		WO 2016020499 A2	11-02-2016
WO 2014072405 A1	15-05-2014	AU 2013343520 A1	04-06-2015
		AU 2017279688 A1	18-01-2018
		BR 112015010283 A2	11-07-2017
		CA 2889767 A1	15-05-2014
		CN 105008539 A	28-10-2015
		EA 201590705 A1	31-05-2016
		EP 2917351 A1	16-09-2015
		JP 6412875 B2	24-10-2018
		JP 2015533511 A	26-11-2015
		KR 20150079964 A	08-07-2015
		SG 11201503308X A	28-05-2015
		US 2015273043 A1	01-10-2015
		WO 2014072405 A1	15-05-2014
WO 2007071707 A2	28-06-2007	AR 058592 A1	13-02-2008
		AR 058706 A1	20-02-2008
		AR 058707 A1	20-02-2008
		AT 520415 T	15-09-2011
		AU 2006327036 A1	28-06-2007
		AU 2006327040 A1	28-06-2007
		AU 2006327041 A1	28-06-2007
		AU 2010241281 A1	25-11-2010
		BR PI0620163 A2	01-11-2011
		BR PI0620193 A2	01-11-2011
		BR PI0620460 A2	16-11-2011
		CA 2633772 A1	28-06-2007
		CA 2634885 A1	28-06-2007
		CA 2634887 A1	28-06-2007
		CA 2808919 A1	28-06-2007
		CA 2816182 A1	28-06-2007
		CL 2013002106 A1	07-03-2014
		CN 101374548 A	25-02-2009
		CN 101378778 A	04-03-2009
		CN 101378779 A	04-03-2009
		CN 103251940 A	21-08-2013
		CN 103585623 A	19-02-2014
		CR 10119 A	29-07-2008
		CR 10122 A	21-08-2008
		CY 1112677 T1	10-02-2016
		CY 1116407 T1	08-02-2017
		CY 1118345 T1	28-06-2017
		CY 1118535 T1	12-07-2017
		DK 1962899 T3	31-10-2011
		DK 1968631 T3	15-06-2015
		DK 1973564 T3	16-01-2017
		DK 2384765 T3	09-01-2017
		EA 200801344 A1	30-12-2008
		EA 200801367 A1	30-12-2008

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/073714

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		EA 200801374 A1	30-12-2008
		EP 1962899 A2	03-09-2008
		EP 1968631 A2	17-09-2008
		EP 1973564 A2	01-10-2008
		EP 2382986 A2	02-11-2011
		EP 2384765 A2	09-11-2011
		EP 2402025 A2	04-01-2012
		EP 3017827 A1	11-05-2016
		EP 3020411 A1	18-05-2016
		EP 3130348 A1	15-02-2017
		ES 2539795 T3	06-07-2015
		ES 2614938 T3	02-06-2017
		ES 2630759 T3	23-08-2017
		HK 1120230 A1	09-12-2011
		HR P20110685 T1	31-10-2011
		HR P20161681 T1	13-01-2017
		HR P20161682 T1	27-01-2017
		HU E032903 T2	28-11-2017
		HU E033081 T2	28-11-2017
		IL 191911 A	24-09-2015
		IL 191913 A	31-07-2012
		IL 192084 A	28-11-2013
		IL 219641 A	31-10-2017
		JO 2813 B1	15-09-2014
		JP 5461838 B2	02-04-2014
		JP 5579387 B2	27-08-2014
		JP 5600337 B2	01-10-2014
		JP 5615323 B2	29-10-2014
		JP 5749689 B2	15-07-2015
		JP 5792920 B2	14-10-2015
		JP 5849125 B2	27-01-2016
		JP 2009520759 A	28-05-2009
		JP 2009520760 A	28-05-2009
		JP 2009520761 A	28-05-2009
		JP 2012211165 A	01-11-2012
		JP 2012229228 A	22-11-2012
		JP 2012232985 A	29-11-2012
		JP 2014205673 A	30-10-2014
		KR 20080081060 A	05-09-2008
		KR 20080081061 A	05-09-2008
		KR 20080081063 A	05-09-2008
		KR 20130103831 A	24-09-2013
		KR 20140036327 A	25-03-2014
		LT 1973564 T	27-12-2016
		LT 2384765 T	10-01-2017
		MA 30062 B1	01-12-2008
		MA 30065 B1	01-12-2008
		MA 30066 B1	01-12-2008
		MX 354843 B	22-03-2018
		MY 147495 A	14-12-2012
		MY 148141 A	15-03-2013
		MY 150719 A	28-02-2014
		MY 160199 A	28-02-2017
		NZ 569076 A	26-08-2011
		NZ 569077 A	26-08-2011
		NZ 569168 A	24-02-2012
		NZ 596500 A	31-05-2013
		PE 10582007 A1	27-11-2007

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		PE 13052007 A1	14-12-2007
		PH 12014502544 A1	14-12-2015
		PL 1973564 T3	28-04-2017
		PL 2384765 T3	31-05-2017
		PT 1962899 E	19-10-2011
		PT 1968631 E	20-07-2015
		PT 1973564 T	24-01-2017
		PT 2384765 T	06-02-2017
		SG 168517 A1	28-02-2011
		SI 1962899 T1	30-11-2011
		SI 1968631 T1	31-07-2015
		SI 1973564 T1	28-02-2017
		TW I415622 B	21-11-2013
		TW 200738257 A	16-10-2007
		TW 201350129 A	16-12-2013
		UA 96583 C2	25-11-2011
		UA 96934 C2	26-12-2011
		US 2009010959 A1	08-01-2009
		US 2009017059 A1	15-01-2009
		US 2009017072 A1	15-01-2009
		US 2010074922 A1	25-03-2010
		US 2015190521 A1	09-07-2015
		US 2015265702 A1	24-09-2015
		US 2016243219 A1	25-08-2016
		US 2018207262 A1	26-07-2018
		WO 2007071707 A2	28-06-2007
		WO 2007071710 A2	28-06-2007
		WO 2007071711 A2	28-06-2007
WO 2015110942 A2	30-07-2015	AU 2015208822 A1	21-07-2016
		CA 2937190 A1	30-07-2015
		CN 106413747 A	15-02-2017
		EP 3096786 A2	30-11-2016
		HK 1232133 A1	05-01-2018
		JP 2017504661 A	09-02-2017
		KR 20160104076 A	02-09-2016
		RU 2016127123 A	28-02-2018
		US 2016324949 A1	10-11-2016
		WO 2015110942 A2	30-07-2015
US 2007231340 A1	04-10-2007	AT 545431 T	15-03-2012
		AU 2007353769 A1	27-11-2008
		BR PI0720624 A2	25-03-2014
		CA 2673477 A1	27-11-2008
		CL 2017001842 A1	04-05-2018
		CN 101610785 A	23-12-2009
		CN 103599529 A	26-02-2014
		DK 2094298 T3	26-03-2012
		DK 2417983 T3	26-08-2013
		EP 2094298 A2	02-09-2009
		EP 2417983 A1	15-02-2012
		ES 2379834 T3	04-05-2012
		ES 2424940 T3	10-10-2013
		HK 1191567 A1	30-12-2016
		IL 199350 A	30-05-2013
		IL 222793 A	31-08-2015
		JP 5523836 B2	18-06-2014
		JP 2010513549 A	30-04-2010

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		KR 20090096731 A	14-09-2009
		KR 20140130567 A	10-11-2014
		KR 20150140411 A	15-12-2015
		PT 2094298 E	03-04-2012
		PT 2417983 E	26-08-2013
		RU 2009125216 A	27-01-2011
		SI 2094298 T1	30-04-2012
		SI 2417983 T1	30-09-2013
		US 2007231340 A1	04-10-2007
		US 2010316666 A1	16-12-2010
		WO 2008143709 A2	27-11-2008
		ZA 200904760 B	28-04-2010
US 2004202668 A1	14-10-2004	AR 040922 A1	27-04-2005
		AT 553775 T	15-05-2012
		AT 553776 T	15-05-2012
		BR 0208595 A	21-02-2006
		CA 2442865 A1	17-10-2002
		CA 2727715 A1	17-10-2002
		CA 2810326 A1	17-10-2002
		CN 1547485 A	17-11-2004
		CN 101112618 A	30-01-2008
		CN 101112619 A	30-01-2008
		CY 1112879 T1	13-04-2016
		CY 1112880 T1	13-04-2016
		CY 1113078 T1	13-04-2016
		CZ 20032698 A3	15-12-2004
		DK 1390066 T3	16-07-2012
		DK 1946771 T3	03-09-2012
		DK 1946772 T3	16-07-2012
		EP 1390066 A2	25-02-2004
		EP 1946771 A1	23-07-2008
		EP 1946772 A1	23-07-2008
		EP 2311488 A2	20-04-2011
		ES 2384041 T3	28-06-2012
		ES 2384065 T3	29-06-2012
		ES 2388690 T3	17-10-2012
		HK 1062891 A1	28-09-2012
		HK 1114779 A1	28-09-2012
		HK 1114781 A1	16-11-2012
		HU 0303996 A2	01-03-2004
		JP 5095595 B2	12-12-2012
		JP 5281224 B2	04-09-2013
		JP 5410498 B2	05-02-2014
		JP 2005508849 A	07-04-2005
		JP 2009102355 A	14-05-2009
		JP 2012051943 A	15-03-2012
		KR 20030087055 A	12-11-2003
		KR 20080048094 A	30-05-2008
		KR 20090026371 A	12-03-2009
		KR 20090120017 A	23-11-2009
		KR 20110036641 A	07-04-2011
		KR 20120005519 A	16-01-2012
		KR 20120005534 A	16-01-2012
		MX PA03008961 A	12-02-2004
		MY 129263 A	30-03-2007
		NO 20034410 A	25-11-2003
		NZ 551621 A	31-07-2008

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/073714

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		NZ 568317 A	29-01-2010
		NZ 581373 A	30-06-2011
		PL 211909 B1	31-07-2012
		PL 211910 B1	31-07-2012
		PL 211911 B1	31-07-2012
		PL 211912 B1	31-07-2012
		PL 211958 B1	31-07-2012
		PT 1390066 E	27-06-2012
		PT 1946771 E	27-08-2012
		PT 1946772 E	09-07-2012
		SI 1390066 T1	31-08-2012
		SI 1946771 T1	30-10-2012
		SI 1946772 T1	31-08-2012
		UA 76952 C2	15-07-2003
		US 2004202668 A1	14-10-2004
		US 2008069835 A1	20-03-2008
		US 2011212124 A1	01-09-2011
		US 2013266609 A1	10-10-2013
		WO 02080965 A2	17-10-2002
		ZA 200307640 B	27-07-2005