

TITLE: EXPRESSION OF PNEUMOCOCCAL SURFACE PROTEIN A (PspA)**FIELD OF INVENTION**

The present invention relates to high level expression of truncated Pneumococcal Surface Protein A1 (PspA1) in bacteria.

BACKGROUND OF INVENTION

5 *Streptococcus pneumoniae* is an important cause of otitis media, meningitis, bacteremia and pneumonia, and a leading cause of fatal infections in the elderly and persons with underlying medical conditions. An attractive goal for streptococcal vaccination is to reduce carriage in the vaccinated populations and subsequently reduce the incidence of pneumococcal disease.

10 Pneumococcal polysaccharide vaccines marketed under the brand name Pneumovax23 is not effective in children under 2 years of age. The inefficacy of polysaccharide vaccines in this population has been attributed to the immaturity of the infant immune system in the expression of B cell receptors. Conjugation of polysaccharides (PS) to carrier proteins converts it from a T cell-independent antigen to a T cell-dependent antigen.
15 As a T cell-dependent antigen, polysaccharides can raise a response with isotype switching, generation of memory cells and a boostable immunological response.

Membrane proteins exist within and span the membrane across which they serve to transport molecules or facilitate cell adhesion. The proteins may assist in the movement of substances by facilitated diffusion (i.e., passive transport) or active transport. The
20 pneumococcal surface protein A (PspA) is a membrane protein and is another important virulence factor found attached to the cell wall of all *Streptococcus pneumoniae* strains.

The replacement of the universal carrier proteins, such as tetanus toxoid (TT) or CRM197, by a pneumococcal protein in particular PspA1 or a fragment thereof, besides broadening the vaccine coverage, would also prevent the impairment of immune responses caused by the excessive use of the same carrier proteins in conjugate vaccines. An expression
25 plasmid is engineered to contain regulatory sequences that act as enhancer and promoter regions and lead to efficient transcription of the gene carried on the expression vector. The goal of a well-designed expression vector is efficient production of protein, and this may be achieved by synthesizing significant amount of stable messenger RNA. It is possible to
30 design an expression vector that exert a tight control of the expression, and the protein is only produced in high quantity when necessary through the use of suitable expression conditions.

In absence of the tight control of the gene expression, the protein may also be expressed constitutively.

Corynebacterium glutamicum is a gram positive fermentative bacterium that is widely used in the production of mono-sodium glutamate, in high amounts. Owing to its stable genetic characteristics and lack of any endotoxin, *Corynebacterium glutamicum* is classified as GRAS organism (generally regarded safe). The bacterium is not known to secrete any extracellular proteases, hence it becomes an attractive platform for producing heterologous proteins into the medium. This can be achieved using an expression plasmid construct that can synthesize the recombinant protein in high amounts.

10 EP2310502 B1 discloses the use of *P_{tac}* promoter in a construct wherein an *Escherichia coli* strain containing IPTG inducible *ftsZ* and *minCDE* deletion mutation were grown in LB medium.

15 Nokano et.al., J Bacteriol. 1984 Jan;157(1):79-83 discloses use of kanamycin resistance gene in the plasmid construct in order to make the transformed strain which acquire the property of being resistant to kanamycin.

 Masai et.al., Proc Natl Acad Sci USA. 1987 Jul;84(14):4781-5 discloses RepA protein for initiation of R1 plasmid replication and interact with *oriR* sequence.

 Nayak et.al., *Infect. Immun.*-1998-Nayak-3744-51 discloses a live oral recombinant *Salmonella* vaccine strain expressing pneumococcal surface protein A (PspA).

20 Nabors et.al., *Vaccine* 18 (2000) 1743 -54 discloses the expression of recombinant truncated PspA as cytoplasmic protein in *Escherichia coli*.

 Figueredo et.al., *Appl Microbiol Biotechnol* (2017) 101:2305–2317 discloses Production and purification of an untagged recombinant pneumococcal surface protein A (PspA4Pro).

25 The above-mentioned references either disclose the genetic elements of the expression vector or expression of pneumococcal surface protein A in *Escherichia coli* and *Salmonella*.

 The inventors have identified immunogenic fragments of Pneumococcal Surface Protein. Thereafter, the inventors have made extensive efforts to develop expression constructs capable of stable and constitutive or inducible expression of the truncated pneumococcal surface protein A1 (PspA1) in bacteria at high level.

30 Therefore, the present invention contemplates to overcome the challenges of the prior art by preparing expression vectors and recombinant host cells for expression of truncated pneumococcal surface proteins. Further, inventors have prepared vaccine compositions comprising the truncated proteins as carrier proteins.

OBJECTIVE OF THE INVENTION

The main objective of the present invention is to provide an expression construct for the high-level expression of truncated pneumococcal surface protein A1 (PspA1) in bacteria.

Another objective of the present invention is to provide an expression construct
5 capable of stably and constitutively or inducibly expressing a high level of truncated pneumococcal surface protein A1 (PspA1) in bacteria.

SUMMARY OF INVENTION

The present invention provides an expression construct capable of high level expression of truncated pneumococcal surface protein A1 (PspA1) as set forth in SEQ ID
10 NO: 3 and 4.

The present invention provides an expression construct comprising a gene, which encodes for truncated *pspA1*, as set forth in SEQ ID NO: 5 and 6.

The present invention provides an expression construct for the high-level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 3 or 4, in
15 bacteria comprising:

- a) gene encoding truncated *pspA1* as set forth in SEQ ID NO: 5 or 6,
- b) origin of replication,
- c) Antibiotic resistance gene,
- d) A promoter and
- 20 e) ribosomal binding site.

The present invention is also directed to a method for high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) which comprises culturing of bacteria transformed with an expression construct and thereby purifying the expressed protein.

BRIEF DESCRIPTION OF THE DRAWINGS

25 **Figure 1:** Panel A shows the diagrammatic representation of pBE31C. Panel B shows sschematic presentation of the domains of *pspA* delineated from the deduced amino acid sequence of Rx1 *pspA1*. Panel C shows the diagrammatic representation of pBE117. Panel D shows the PCR amplicon containing RBS, native signal peptide, truncated *pspA1*, native terminator and trnB.

30 **Figure 2:** Protein sequence coverage of truncated PspA1 in a peptide fingerprint analysis.

Figure 3: Intact mass analysis of the truncated PspA1 expressed in *Corynebacterium glutamicum*.

Figure 4: Truncated PspA1 eluted from ceramic hydroxyapatite (CHT-II) column.

Figure 5: Truncated PspA1 eluted from anion exchange column.

Figure 6: Truncated PspA1 after diafiltration.

Figure 7: SEC-HPLC chromatogram for conjugation reaction kinetics of pneumococcal polysaccharide Serotype 3 (A), 6A (B) and 6B (C).

5 **Figure 8:** Serum antibody titer of immunized rabbits against different conjugates of *Streptococcus pneumoniae* polysaccharide from serotypes 3, 6A, 6B (2.2mcg) with truncated *pspA1* carrier protein.

10 **Figure 9:** Serum antibody titer of immunized rabbits against different conjugates of *Streptococcus pneumoniae* polysaccharide from serotypes 3, 6A, 6B (4.4mcg) with truncated *pspA1* carrier protein.

Figure 10: Panel A shows the diagrammatic representation of pBE114k. Panel B shows the confirmation of pBE114k by restriction digest.

Figure 11: Truncated PspA1 eluted from CHT I chromatography.

Figure 12: Truncated PspA1 eluted from Capto Q Impress chromatography.

15 **Figure 13:** Protein coverage of PspA1 by MALDI.

Figure 14: Intact mass analysis of the truncated PspA1 expressed in *Escherichia coli*.

DETAILED DESCRIPTION OF THE INVENTION

The term PspA1 refers to Pneumococcal Surface Protein A1 from *Streptococcus pneumoniae*.

20 The present invention relates to high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) in bacteria. The bacteria suitable for high level expression of *pspA1* is *Corynebacterium glutamicum* and *Escherichia coli*.

In an embodiment, the present invention relates to high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 3 in *Corynebacterium glutamicum*.

25 In another embodiment, the present invention relates to high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 4 in *Escherichia coli*.

30 The present invention provides an expression construct capable of high level expression of surface protein as set forth in SEQ ID NO: 3.

The present invention also provides an expression construct capable of high level expression of surface protein as set forth in SEQ ID NO: 4.

In another embodiment, the present invention relates to an expression construct for high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 3 in *Corynebacterium glutamicum*.

In yet another embodiment, the present invention relates to an expression construct
5 for high level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 4 in *Escherichia coli*.

The present invention relates to an expression construct which is used for the high-level expression of truncated PspA1 in *Corynebacterium glutamicum* comprising:

- i. ori R origin of replication,
- 10 ii. Kanamycin resistance gene,
- iii. P_{tac} promoter,
- iv. Gene of interest coding for a truncated *pspA1* (SEQ ID NO: 5).

The present invention also relates to an expression construct which is used for the high-level expression of truncated PspA1 in *Escherichia coli* comprising:

- 15 i. pUC origin of replication
- ii. Kanamycin resistance gene
- iii. P_{T7} promoter
- iv. Gene of interest coding for a truncated *pspA1* (SEQ ID NO: 6).

In an embodiment, the expression construct for high level expression of truncated *pspA1*
20 further comprises ribosomal binding site (RBS). RBS is included in the forward primer and a short stretch of DNA containing native terminators are included in the reverse primer used for further amplification of truncated *pspA1*. The RBS (Ribosomal Binding Site) is Triose phosphate isomerase in the expression construct for the expression of truncated PspA1 in *Corynebacterium glutamicum*.

25 In an embodiment of the present invention, the expression construct comprises PspA1 signal peptide coding region (native), truncated *pspA1*, P_{tac} promoter and Ribosomal binding site (RBS) of *triose phosphate isomerase* gene.

In a preferred embodiment, the present invention provides an expression construct for high level expression of truncated PspA1 (SEQ ID NO: 3), in *Corynebacterium glutamicum*
30 which comprises:

- a. gene encoding truncated *pspA1* (SEQ ID NO: 5),
- b. ori R origin of replication (SEQ ID NO: 12),
- c. kanamycin resistance gene (SEQ ID NO: 1),
- d. P_{tac} promoter (SEQ ID NO: 2) and

e. triose phosphate isomerase ribosomal binding site.

In another preferred embodiment, the present invention provides an expression construct for high level expression of truncated PspA1 (SEQ ID NO: 4), in *Escherichia coli* which comprises:

- 5 a) gene encoding truncated *pspA1* (SEQ ID NO: 6),
- b) pUC origin of replication,
- c) kanamycin resistance gene (SEQ ID NO: 1),
- d) *P_{T7}* promoter (SEQ ID NO: 11), and
- e) ribosomal binding site.

10 The pneumococcal surface protein A (PspA) is a membrane protein and is an important virulence factor found attached to the cell wall of all *Streptococcus pneumoniae* strains and is a promising component as has been shown to be highly immunogenic.

In a particular embodiment of the present invention, truncated PspA1 is used as a carrier protein. Carrier proteins, employed in the conjugate vaccines, are preferably proteins that are
15 non-toxic and non-reactogenic and obtainable in large amount and purity. A carrier protein can be conjugated to capsular polysaccharide isolated from pathogenic bacteria to enhance immunogenicity of the polysaccharide. Carrier proteins should be amenable to standard chemical conjugation procedures.

The truncated PspA1 in *Corynebacterium glutamicum* is secreted into the extracellular
20 medium. The secretion into the extracellular medium aids in efficient purification.

The present invention relates to high level expression of truncated PspA1 in *Corynebacterium glutamicum* wherein the N-terminal region along with proline rich region of truncated *pspA1* gene was amplified from 23F capsular serotype of *Streptococcus pneumoniae*, together with the upstream region.

25 In an embodiment of the present invention, the expression construct comprises truncated *pspA1*, *P_{tac}* promoter and Ribosomal binding site (RBS) of *triose phosphate isomerase* gene.

In an embodiment of the present invention, the expression construct comprises truncated *pspA1*, *P_{T7}* promoter and Ribosomal binding site (RBS).

The expression construct is electroporated into *Corynebacterium glutamicum* and
30 selected on LB plates with kanamycin as selectable marker.

P_{tac} is a strong hybrid promoter composed of the -35 region of the *trp* promoter and the -10 region of the *lacUV5* promoter/operator.

pspA sequences were obtained from GenBank and aligned from the protein database. Primers were designed to specifically amplify native signal peptide coding region, N-terminal

region along with the proline rich region of *pspA* genes belonging to families 1 & 2. The required regions of *pspA1* and *pspA2* were amplified from the available *Streptococcus pneumoniae* clinical isolates. MHC peptide analysis showed that *pspA1* is comparatively more immunogenic than *pspA2*.

- 5 Amino acid sequence encoding truncated PspA1 expressed and produced in *Corynebacterium glutamicum* is set forth in SEQ ID NO: 3 having an intact mass observed to be 35452 Daltons.

Amino acid sequence encoding truncated PspA1 expressed and produced in *Escherichia coli* is set forth in SEQ ID NO: 4 having an intact mass observed to be 41966 Daltons.

- 10 DNA sequence encoding truncated *pspA1* expressed in *Corynebacterium glutamicum* and *Escherichia coli* are set forth in SEQ ID NO: 5 and SEQ ID NO 6 respectively.

- Corynebacterium glutamicum* (previously known as *Micrococcus glutamicus*) used for expressing the truncated *pspA1* is a GRAS organism Gram-positive, rod-shaped bacteria. *Corynebacterium glutamicum* is a GRAS organism. The whole genome sequence of
- 15 *Corynebacterium glutamicum* ATCC 13032 is available which can be grown to higher cell densities and is also genetically stable owing to the lack of a recombination repair system. It has limited restriction-modification system. It shows no autolysis and can maintain metabolic activity under growth-arrested conditions. It has low protease activity favoring recombinant protein production. Its plasticity of metabolism and strong secondary metabolism properties,
- 20 capacity to utilize broad spectrum of carbon sources (pentoses, hexoses, and alternative carbon sources), stress-tolerance to carbon sources make it a promising host for heterologous protein production. These physiological properties make *Corynebacterium glutamicum* accessible to manipulation and cultivation in robust industrial conditions, thus making it a successful industrial workhorse. Heterologous expression of proteins like α - Amylase,
- 25 endoglucanase, endoxylanase, GFP, Xylanase etc., was reported in *Corynebacterium glutamicum*.

In an embodiment of the present invention, the yield of truncated PspA1 is about 500 mg/L, about 400 mg/L, about 300 mg/L, about 250 mg/L, about 220 mg/L, about 200 mg/L, about 180 mg/L, about 160 mg/L, about 150 mg/L, about 120 mg/L, about 100 mg/L.

- 30 In yet another embodiment, the present invention provides a pneumococcal conjugate vaccine comprising at least one polysaccharides from *Streptococcus pneumoniae* serotypes selected from 1, 2, 3, 4, 5, 6A, 6B, 6C, 6D, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15A, 15B, 15C, 16F, 17F, 18C, 19F, 19A, 20A, 20B, 22F, 23A, 23B, 23F, 24B, 24F, 31, 33F, 34, 35B, 35F, 38, 39 and 45, conjugated with truncated *pspA1* of the present invention or combination

of truncated PspA1 and other carrier proteins such as CRM197, tetanus toxoid, pertussis toxoid; PsaA and the like.

In yet another embodiment, the present invention provides multivalent pneumococcal vaccine composition selected from 10 valent, 14 valent, 15 valent, 17 valent, 18 valent, 19 valent, 20 valent, 22 valent, 23 valent, 24 valent or 25 valent comprising polysaccharides from *Streptococcus pneumoniae* serotypes selected from 1, 2, 3, 4, 5, 6A, 6B, 6C, 6D, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15A, 15B, 15C, 16F, 17F, 18C, 19F, 19A, 20A, 20B, 22F, 23A, 23B, 23F, 24B, 24F, 31, 33F, 34, 35B, 35F, 38, 39 and 45, conjugated with truncated *pspA1* of the present invention or combination of truncated PspA1 and other carrier proteins such as CRM197, tetanus toxoid, pertussis toxoid; *psaA* and the like.

In a preferred embodiment, the present invention provides a multivalent conjugate vaccine comprising at least three polysaccharides from *Streptococcus pneumoniae* serotypes 3, 6A and 6B conjugated to truncated PspA1 of the present invention.

In an embodiment of the present invention provides a conjugate vaccine comprising polysaccharides from *Streptococcus pneumoniae* serotypes 3, 6A and 6B conjugated to truncated PspA1 of the present invention and *Streptococcus pneumoniae* serotypes 1, 4, 5, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F and 33F conjugated to CRM₁₉₇.

The present invention provides formulations containing either 2.2µg or 4.4µg of each of the pneumococcal polysaccharides from serotypes 3, 6A and 6B each conjugated to truncated PspA1 of the present invention and about 2.2 µg each of pneumococcal polysaccharides from serotypes 1, 4, 5, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F and 33F each conjugated to CRM₁₉₇.

In another embodiment, the present invention provides a pneumococcal vaccine composition as a single 0.5 mL dose, the single dose comprising about 2.2 to 4.4 µg of one or more pneumococcal polysaccharides; about 1 µg to about 50 µg of truncated PspA1 of the present invention conjugated to each of the one or more pneumococcal polysaccharides; about 0.2 mg to about 1 mg of aluminum phosphate adjuvant; and an excipient.

In another embodiment, the present invention provides a pneumococcal vaccine composition as a single 0.5 mL dose, the single dose comprising about 2.2 to 4.4 µg of one or more pneumococcal polysaccharides; about 1 µg to about 30 µg of truncated PspA1 of the present invention conjugated to each of the one or more pneumococcal polysaccharides; about 1 µg to about 30 µg of CRM₁₉₇ conjugated to each of the one or more pneumococcal polysaccharides; about 0.2 mg to about 1 mg of aluminum phosphate adjuvant; and an excipient.

In another embodiment, the present invention also provides a vaccine for the prevention of invasive disease caused by *Streptococcus pneumoniae*, by administering the conjugated vaccine prepared by conjugating the truncated PspA1 of the present invention with pneumococcal polysaccharides.

5 EXAMPLES

The following examples are provided to illustrate the invention and are merely for illustrative purpose only and should not be construed to limit the scope of the invention.

Example 1: Recombinant Expression of truncated PspA1 in *Corynebacterium glutamicum*.

10 Construction of pBE31C

A small stably replicating broad host range plasmid pBE30 was generated using synthetic plasmid having 2.692 Kb pNG2 *oriR* sequence. Using this plasmid DNA as template, a 1.8 Kb *oriR* region was amplified using the primers pEP-F1 (5'- GCG CGG ACT AGT AGA TCT ATG GTA AAT CTG CGC AGA CAG-3') and pEP-R1 (5'- GCG CGG
15 ACT AGT GAA TTC GGT GAG GTT ATG GCG-3').

Simultaneously, 1.033Kb *kanR* sequence was amplified using pUC4-KIXX template DNA and Kan-F2 (5'- AAG GTC CCG GGA TGG CGA TAG CTA GAC TGG GCG GT-3') and Kan-R2 (5'-AAG GTC CCG GGG GTT GGG CGT CGC TTG GTC GG-3') primers. Both *kanR* gene amplicon and *oriR* amplicon were blunt end ligated to create pBE30 vector.
20 Further, a 0.851 Kb expression cassette containing a tandem of *tac lac UV5* promoter, multiple cloning site, *lacZα* component and TrnB terminator sequence was amplified using the *P_{tac}*-F1 (5'- GG AGC ACT AGT CTG AAA TGA GCT GTT GAC AAT TAA TC-3') and *P_{tac}*-R1 (5'- GG AGC ACT AGT TTT AAA CAT GAG CGG ATA CAT ATT TGA A-3') primers each appending *SpeI* restriction site. The template DNA used for the
25 amplification of expression cassette is pMMB206 (ATCC 37808). Later the expression cassette was cloned into the unique *SpeI* site designed in the pBE30 plasmid. The plasmid thus obtained is designated as pBE31C (Figure 1 A; SEQ ID NO: 14).

Construction of pBE117

Truncated *pspA1* gene (N terminal region along with proline rich region Figure 1B)
30 was amplified from 23F capsular serotype of *Streptococcus pneumoniae*, together with the upstream region. This was cloned in a TA vector (pTZ57R/T procured from Fermentas) using primers PSPAF1_FP (5'ATG AAT AAG AAA AAA ATG ATT TTA ACA AGT CTA GCC 3') and PSPAF1_RP (5'CGA GAG AGA TCT AAA TTA AAA TGT CAA ATG TTC TTA ACA TGC TTT AAT TTT TAT TTT GGT GC 3') and sequence verified. This was

designated as pTZ-*pspA1*. Native terminator sequence was included in the reverse primer PSPAF1_RP. The clade defining regions were mapped in the obtained truncated *pspA1* sequence to confirm that it belongs to family 1 of PspA proteins.

To express truncated *pspA1* in *Corynebacterium glutamicum*, P_{tac} promoter (SEQ ID NO: 2) and Ribosomal binding site (RBS) of *triose phosphate isomerase* gene (SEQ ID NO: 7) belonging to *Corynebacterium*, were chosen. *pspA1* gene along with the entire cassette including RBS, native signal peptide (SEQ ID NO: 8), truncated *pspA1* gene, native terminator (SEQ ID NO: 10) was amplified using primers- SDTICGR0949_FP5 (5'GAG CGA TGG ATC CTA GAA AGG TGT GTT TCA CCC ATG AAT AAG AAA AA 3') and PSPA2_2RP (5' TCA AAT GTT CTT AAC ATG CTT TAA TTT TTA TGG TGC AGG AGC TGG TTG 3') and pTZ-*pspA1* as template. RBS was included in the forward primer SDTICGR0949_FP5. *rrnB* terminator (SEQ ID NO: 13) region was amplified from pBE31C available in-house and ligated to gene encoding for truncated *pspA1* using Splicing by Overlap Extension (SOE) PCR. *rrnB* gene was amplified using TER FP2 (5'ATG TTA AGA ACA TTT GAC ATT TTA ATT TCG GCA CTG GCC GTC GTT 3') and TER_RP3 (5'GCG ATA TGG ATC CCA TGA GCG GAT ACA 3'). PSPA2_2RP and TER FP2 were designed such that there is an overlap of 17 bases. Both the amplicons *pspA1* and *rrnB* were added in the molar ratio of 1:1 and used as template for SOE PCR with the primers SDTICGR0949_FP5 (5'GAG CGA TGG ATC CTA GAA AGG TGT GTT TCA CCC ATG AAT AAG AAA AA 3') and TER_RP3 (5'GCG ATA TGG ATC CCA TGA GCG GAT ACA 3'). Subsequently, the entire amplicon including RBS, native signal peptide (SEQ ID NO: 8), truncated *pspA1* gene, native terminator (SEQ ID NO: 10), *rrnB* terminator was digested with the restriction enzyme appended in the forward (SDTICGR0949_FP5) and reverse primers (TER_RP3) and cloned in an expression vector pBE31C made for *Corynebacterium*. The resulted clone was designated as an expression construct pBE117 (Figure 1C; SEQ ID NO: 15). The expression vector along-with truncated *pspA1* hereinafter is referred to as expression construct. The orientation of the insert (Figure 1D) was confirmed by PCR analysis. The sequence of truncated *pspA1* gene along with its expression cassette was confirmed by DNA sequencing.

30 Expression of truncated PspA1

The expression construct was electroporated into *Corynebacterium glutamicum* ATCC 13032 and selected on LB plates with kanamycin as selectable marker. Twenty recombinant *Corynebacterium glutamicum* colonies were picked and analysed by PCR. Five colonies were chosen for constitutive expression of truncated *pspA1*. The recombinant

colonies along with *Corynebacterium glutamicum* ATCC 13032 (as negative control) were inoculated into 10 ml Terrific Broth with 25 µg/ml final concentration of kanamycin and incubated at 35°C with shaking at 200 rpm. After 16 h, secondary inoculation was done in 10 ml of the same medium described above, such that the final OD is 0.1. The cultures were
5 incubated at 35°C with shaking at 200 rpm for 18- 20h. After 18 h, the culture supernatants were checked for the expression of truncated *pspA1*. 30 µl of the supernatant was loaded onto 12 % SDS-PAGE and analysed for truncated *pspA1* expression. A prominent band was seen around 45 kDa.

Western analysis using N terminal epitope specific *pspA* polyclonal antibody
10 (SantaCruz) confirmed the expression of truncated *pspA1*. The expression analysis of recombinant clone 5 was scaled up to 500 ml and the expression of truncated *pspA1* was confirmed at least 3 times. Truncated *pspA1* was initially purified from shake flask experiments using CHT type 1 and Capto Q impress to nearly 99 % purity. Later, after confirming the consistent expression of truncated PspA1 in *C. glutamicum*, the expression
15 was scaled up to 1.5 L.

Purification and validation of truncated PspA1

The 800 ml of culture supernatant containing truncated PspA1 from 1.6 L (0.66 mg/ml), obtained from upstream was dialysed through 10 kDa and concentrated to 260 ml (1.92mg/ml). 70 ml of this dialysed concentrate was subjected to purification using CHT type
20 1 and Capto Q impress. The final recovery of truncated *pspA1* was 162 mg/ L.

MALDI MS/MS analysis of the gel plug containing purified truncated PspA1 from SDS-PAGE gave a clear hit score of 233 with Pneumococcal surface protein A. 31 % (Figure 2) sequence coverage was shown with the PspA protein with NCBI protein id ABY67187.1. The intact molecular weight analysis showed the molecular weight of the truncated *pspA1*
25 expressed is 35.4 kDa. This mass matches with the theoretical molecular weight of truncated *pspA1*. The peak at 17725.4 is a peak of molecule having a charge of 2, hence in the m/z the peak appears at half the intact mass of the truncated *pspA1*. (Figure 3)

Example 2: Production of truncated PspA1

Corynebacterium glutamicum ATCC 13032 harbouring the expression construct
30 comprising truncated *pspA1* (hereinafter referred as *pspA1*) gene was revived from a source bank in LB media. This was used to inoculate a fermenter vessel (5L; CSTR). The parameters that were monitored during the production process were – pH, DO, Temperature (ΔT), Carbon source, Metabolites. Process automation technology (PAT) based feeding strategy was adopted for carrying out Short Fed-Batch of truncated PspA1 fermentation. There was no

induction-based control of the product (truncated PspA1) as this was growth associated product. The harvest OD was approximately 90 (OD_{600nm}) at the time of harvest in a semi-synthetic media. The cells were harvested and washed with PBS, prior to cell disruption in a French Press.

5 Purification of truncated PspA1

16.2 L spent medium, with a total protein concentration of 0.8 mg/ mL, was collected from 20 L fermentation batch. 16.2 L spent media was concentrated to 2.6 L (3.6 mg/mL) by using 10 kDa 0.5 m² cassette followed by Dia filtration against 20mM Potassium phosphate pH-6.8, cond 3.2 ms/cm. This 2.6 L was split into two lots i.e Lot 1 with 1.4 L and Lot 2 with 1.2 L and proceeded to further purification. 500 ml of CHT I resin was packed in HiScale 10 50/40 column. Resin was washed with sterile distilled water followed by equilibration with 8 column volumes (CV) of 20mM Potassium Phosphate pH-6.8 (Buffer A). 1400 mL (3.6 mg /mL) of the spent media concentrate (Lot 1) was loaded onto the column and flow through was collected. Column was washed with 5 column volumes of Buffer A. PspA1 was eluted 15 with Step gradient using 250mM Potassium Phosphate pH-6.8 (Buffer B). The step gradient involved 5 CVs of Step- 40 % B, 5 CVs of Step- 80 % B and 3 CVs of 0.5M Potassium Phosphate buffer. The flow rate was maintained at 80 mL/ min for the entire run. PspA1 was collected in fraction 1 of Step- 80 % B with a fraction volume of 1250 mL. (Figure 4)

Capto Q Impress was used as a second chromatography step in PspA1 purification. 20 250 ml of Capto Q Impress resin was packed in XK 50/20 column. Resin was washed with sterile distilled water followed by equilibration with 5 column volumes (CV) of 20mM Potassium Phosphate and 100mM NaCl, pH-6.8 (Buffer A). 1250 mL of CHT I fraction was diluted to 2300mL with 20mM Potassium Phosphate, pH-6.8, loaded onto the column and flow through was collected. Column was washed with 5 column volumes of Buffer A. PspA1 25 was eluted using 12 CVs of Buffer B (20 mM Potassium Phosphate with 1 M NaCl, pH-6.2) in a linear gradient of 0 to 40 % B and a final step of 3 CVs of 100 % B. Each fraction of 250mL was collected. The flow rate was maintained at 40 mL/ min. PspA1 was collected in linear gradient of 40 % B with pooled fraction volume of 1250 mL. (Figure 5)

Capto Q fractions of 4,5,6,7 and 8 were pooled, concentrated/ diafiltered with 20 mM 30 Potassium phosphate pH-6.8. The final recovery obtained was 100 mL of *pspA1* with a concentration of 15.5 mg/mL of protein (Total PspA1 is 1550 mg) from Lot 1. Similar process was followed for Lot 2 and the final recovery obtained was 70 mL of *pspA1* with a concentration of 16 mg/mL (Total PspA1 is 1120 mg) from Lot 2. The purified PspA1 from

both Lot 1 and Lot 2 is 2670 mg from 16.2 L batch (Purified PspA1 yield of 164 mg/L). (Figure 6)

Example 3: Construction of pBE114K

The expression vector pRSET A (commercial vector from Invitrogen) was modified by removing ampicillin resistance gene using *Dra*I restriction digestion and ligating to *Sma*I digested kanamycin coding gene (obtained from pUC4 KIXX). This modified vector was designated as pRSET-km. Truncated PspA1 was expressed in *Escherichia coli* under *P_{T7}* promoter (SEQ ID NO: 11) of pRSET-km. A short stretch of DNA containing native terminators was included in the reverse primer used for further amplification of *pspA1*. Subsequently, the entire truncated *pspA1* gene along with its native terminator was amplified, digested with the restriction enzyme appended in the forward and reverse primers and cloned in pRSET-km vector. The resulted clone was designated as pBE114k (Figure 10A). The expression vector along-with truncated PspA1 hereinafter is referred to as expression construct. The expression construct pBE114k was confirmed by restriction digestion (Figure 10B). The sequence of truncated PspA1 gene along with its expression cassette was confirmed by DNA sequencing.

Expression of truncated PspA1

pBE114k was transformed into *Escherichia coli* DH5 α -T1R chemical competent cells (Procured from Invitrogen) and selected on LB plates with kanamycin as selectable marker. 40 recombinant *Escherichia coli* colonies were picked and analysed by PCR. All the 40 colonies were chosen for inducible expression of truncated PspA1. The recombinant colonies along with *Escherichia coli* (as negative control) were inoculated into 10 ml Terrific Broth with 25 μ g/ml final concentration of kanamycin and incubated at 37°C with shaking at 200 rpm. At mid log phase, 1 mM IPTG was added to induce the expression of PspA1 in *Escherichia coli* (pBE114k). The cultures were incubated at 30°C with shaking at 200 rpm for 16h, after induction. After 16 h, the culture supernatants were checked for the expression of truncated PspA1. The cells were collected by centrifugation, lysed and loaded onto 12 % SDS-PAGE and analysed for truncated PspA1 expression. A prominent band was seen around 65 kDa. Western analysis using N-terminal epitope specific PspA polyclonal antibody (SantaCruz) confirmed the expression of truncated PspA1. The expression analysis of recombinant clone 29 was confirmed for the expression of truncated PspA1 at least 3 times. Truncated PspA1 was initially purified from shake flask experiments using CHT type 1 and Capto Q impress.

Purification and validation of truncated PspA1

40 grams (wet weight) of cell pellet from *Escherichia coli* (pBE114k) was taken and re-suspended in 400ml of 20mM Potassium Phosphate buffer pH-6.8 containing 1 mg/ ml Lysozyme and 1 mM PMSF. The cell suspension was lysed using high pressure homogenizer for 3 passes at 1000 psi. 400 mL of cell lysate with a total protein concentration of 4 mg/ mL was diluted to 750 mL with 20 mM Potassium Phosphate buffer pH-6.8 and purified using CHT I resin followed by Capto Q Impress as a second chromatography.

250 mL of CHT I resin was packed in HiScale 50/20 column. Resin was washed with sterile distilled water followed by equilibration with 5 column volumes (CV) of Buffer A. 750 mL diluted cell lysate was loaded onto the column and flow through was collected. Column was washed with 5 column volumes of Buffer A. PspA1 was eluted with Step gradient using 250mM Potassium Phosphate pH-6.8, cond-29.5 ms/cm (Buffer B). The step gradient involved 5 CVs of Step- 50 % of Buffer B, 5 CVs of Step- 80% of Buffer B and Column was stripped with 3 CVs of 0.5M Potassium Phosphate buffer pH-6.8, Cond- 48ms/cm (Buffer C). The flow rate was maintained at 40 mL/ min for the entire run. The PspA1 protein peak fractions were collected manually, Elution fractions pool of 5, 6 and 7 of Step- 80 % B with a final volume of 350 mL (Figure 11).

Capto Q Impress was used as a second chromatography step in PspA1 purification. 60 ml of Capto Q Impress resin was packed in XK 26/20 column. Resin was washed with sterile distilled water followed by equilibration with 5 column volumes (CV) of 20mM Potassium Phosphate and 100 mM NaCl, pH-6.8, con-13.6ms/cm (Buffer A). 350 mL of CHT I fraction was diluted to 700mL with 1mM Potassium Phosphate, pH-6.8, loaded onto the column and flow through was collected. Column was washed with 5 column volumes of Buffer A. PspA1 was eluted using linear gradient 10 CVs of Buffer B (20 mM Potassium Phosphate with 1 M NaCl, pH-6.2, cond-89ms/cm) in a linear gradient of 0 to 40 % B fractions were collected manually, fraction 4 to 8, 10 and 11 were 60mL and 9 and 12 were 100mL fractions (Figure 12). Final step of 3 CVs of 100% B, 180mL collected as fraction 13. The flow rate was maintained at 10 mL/ min. The PspA1 protein was collected fraction 9 in linear gradient of 40 % B was 100 mL.

Capto Q fraction of 9 was collected, concentrated/ diafiltered with 20mM Potassium phosphate pH-6.8. The final recovery obtained was 100mL of PspA1 with a concentration of 5 mg/mL of protein (Total PspA1 protein is 500 mg) from 400mL cell lysate with a concentration of 4 mg/mL. Recovery % is 31.25.

MALDI MS/MS analysis of the gel plug containing purified truncated PspA1 from SDS-PAGE gave a clear hit score of 223 with Pneumococcal surface protein A. 39 % (Figure 13) sequence coverage was shown with the PspA protein with NCBI protein id WP_050210652.1.

5 The intact molecular weight analysis showed the molecular weight of the truncated PspA1 expressed is 41.96 kDa. Intact mass analysis of the truncated PspA1 expressed in *Escherichia coli* showed 41966 Da (41.9 kDa), matches with the theoretical molecular weight of truncated PspA1. The peak at 20982.7 Da is a peak of molecule having a charge of 2 hence in the m/z the peak appears at half the intact mass of the truncated *pspA1*. (Figure 14).

10 **Example 4: Conjugation of individual pneumococcal polysaccharide to carrier protein to form Polysaccharide- Truncated PspA1 Conjugates**

Activation of Serotype 3 and conjugation with truncated *pspA1*

Approximately, 1:1 ratio of sized serotype 3 (6.0 mL of PS, concentration of 10 mg/mL) and CDAP (100 mg/mL in Acetonitrile (w/v)) was mixed in a glass vial and stirred
15 for 1 min. pH of pneumococcal polysaccharide serotype 3 were adjusted to 9.25 with 0.2M Triethylamine and stirred for 3 min at room temperature (RT). Truncated PspA1 (4.0 mL of conc. 15.0 mg/mL) were added to the activated serotype 3 in the ratio of 1:1 (Truncated PspA1: Serotype 3).

pH of the reaction was adjusted to ~9.05 with 0.2M Triethylamine and the reaction was
20 continued under stirring for 5 hours at room temperature and finally the reaction was quenched by adding excess concentration of glycine.

The reaction mixture was diafiltered using 100 kDa MWCO membrane and purified by size-exclusion chromatography wherein solid line is for polysaccharides and dashed lines for truncated *pspA1* and five-hour reaction is represented by dotted line in chromatogram A
25 (Figure 7). The fractions were analysed by SEC-MALLS anthrone method and fractions containing conjugates were pooled and sterile filtered with 0.2 μ filters. From now this material is called monovalent conjugate bulk (Serotype 3-Truncated PspA1 conjugate).

Activation of Serotype 6A and conjugation with truncated PspA1

Approximately, 1:1 ratio of sized serotype 6A (6.0 mL of PS, concentration of 10
30 mg/mL) and CDAP (100 mg/mL in Acetonitrile (w/v)) was mixed in a glass vial and stirred for 1 min. pH of pneumococcal polysaccharide serotype 6A were adjusted to 9.25 with 0.2M Triethylamine and stirred for 3 min at room temperature (RT). Truncated PspA1 (4.0 mL of conc. 15.0 mg/mL) were added to the activated serotype 6A in the ratio of 1:1 (Truncated PspA1: Serotype 6A).

pH of the reaction was adjusted to ~9.05 with 0.2M Triethylamine and the reaction was continued under stirring for 5 hours at room temperature and finally the reaction was quenched by adding excess concentration of glycine.

The reaction mixture was diafiltered using 100 kDa MWCO membrane and purified by size-exclusion chromatography wherein solid line is for polysaccharides and dashed lines for truncated PspA1 and five-hour reaction is represented by dotted line in chromatogram B (Figure 7). The fractions were analysed by SEC-MALLS, anthrone method and fractions containing conjugates were pooled and sterile filtered with 0.2 μ filters. From now this material is called monovalent conjugate bulk (Serotype 6A-Truncated PspA1 conjugate).

10 Activation of Serotype 6B and conjugation with truncated PspA1

Approximately, 1:1 ratio of sized serotype 6B (6.0 mL of PS, concentration of 10 mg/mL) and CDAP (100 mg/mL in Acetonitrile (w/v)) was mixed in a glass vial and stirred for 1 min. pH of pneumococcal polysaccharide serotype 6B were adjusted to 9.25 with 0.2M Triethylamine and stirred for 3 min at room temperature (RT). Truncated PspA1 (4.0 mL of conc. 15.0 mg/mL) were added to the activated serotype 6B in the ratio of 1:1 (Truncated PspA1: Serotype 6B).

pH of the reaction was adjusted to ~9.05 with 0.2M Triethylamine and the reaction was continued under stirring for 5 hours at room temperature and finally the reaction was quenched by adding excess concentration of glycine.

20 The reaction mixture was diafiltered using 100 kDa MWCO membrane and purified by size-exclusion chromatography wherein solid line is for polysaccharides and dashed lines for truncated PspA1 and five-hour reaction is represented by dotted line in chromatogram C (Figure 7). The fractions were analysed by SEC-MALLS, anthrone method and fractions containing conjugates were pooled and sterile filtered with 0.2 μ filters. From now this material is called monovalent conjugate bulk (Serotype 6B-Truncated PspA1 conjugate).

Example 5: Immunogenicity study of the conjugate vaccine

Two formulations containing either 2.2 μ g or 4.4 μ g of Serotypes 3, 6A and 6B each conjugated to truncated PspA1 was prepared containing 2.2 μ g of serotypes 1, 4, 5, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F and 33F each conjugated to CRM₁₉₇. These conjugates were 30 adsorbed onto Al-hydrogel.

Rabbits having a body weight of 1.5 to 2 kg were grouped in 7 animals each and were immunized with the above-mentioned conjugate formulations. Sera samples were analysed before and after immunization. Serum obtained from the immunized rabbits were analysed for the presence of polysaccharide specific antibody in an indirect ELISA.

The serum antibody titer in the immunized rabbits were evaluated by indirect ELISA. Micro-titer plates coated with specific polysaccharides were reacted with the serum antibody. Rabbit serum prior to immunization, post 1st and 2nd dose was used for analysis wherein Y axis indicates antibody titer, which is arrived at using inverse of maximum dilution that gave ELISA OD₄₅₀ above the cut-off. Serum antibody titer in pre-immune rabbits were below detection limit. Open bar indicates titer after first dose of vaccine administration while the black solid bar indicates antibody titer after second dose of the vaccine (Figure 8 & 9). There was a dose dependent increase in both truncated *pspA1* conjugated serotype as well as CRM197 conjugated serotype titers. The titers of CRM197 conjugated polysaccharide was not inhibited by the presence of truncated PspA1 conjugates and vice-versa. This indicates that the truncated PspA1 could be used as an alternative carrier protein for polysaccharide protein conjugate vaccine.

The claims:

1. A nucleic acid encoding truncated pneumococcal surface protein A1 (PspA1) comprising the nucleotide sequence of SEQ ID NO: 5 or SEQ ID NO: 6.
2. An expression construct comprising the nucleic acid as claimed in claim 1.
- 5 3. An expression construct capable of high level expression of truncated pneumococcal surface protein A1 (PspA1) as set forth in SEQ ID NO: 3 or SEQ ID NO: 4.
4. An expression construct for the high-level expression of truncated PspA1 (Pneumococcal Surface Protein A1) as set forth in SEQ ID NO: 3 or SEQ ID NO: 4, comprising:
 - 10 a. gene encoding truncated *pspA1* as set forth in SEQ ID NO: 5 or SEQ ID NO: 6;
 - b. origin of replication;
 - c. antibiotic resistance gene;
 - d. a promoter; and
 - 15 e. ribosomal binding site.
5. An expression construct as claimed in claim 4, wherein the antibiotic resistance gene is kanamycin resistance gene.
6. An expression construct as claimed in claim 4, comprising:
 - 20 a. ori R origin of replication;
 - b. kanamycin resistance gene;
 - c. *P_{tac}* promoter; and
 - d. gene of interest coding for a truncated *pspA1* as set forth in SEQ ID NO: 5.
7. An expression construct as claimed in claim 4, comprising:
 - 25 a. pUC origin of replication;
 - b. kanamycin resistance gene;
 - c. *P_{T7}* promoter; and
 - d. gene of interest coding for a truncated *pspA1* as set forth in SEQ ID NO: 6.
8. An expression construct as claimed in claim 7 further comprising triose phosphate isomerase ribosomal binding site as set forth in SEQ ID NO: 7.
- 30 9. An expression construct for high level expression of truncated PspA1 (SEQ ID NO: 3), in *Corynebacterium glutamicum* which comprises:
 - a. gene encoding truncated *pspA1* (SEQ ID NO: 5);
 - b. ori R origin of replication (SEQ ID NO: 12);

- c. kanamycin resistance gene (SEQ ID NO: 1);
 - d. *P_{tac}* promoter (SEQ ID NO: 2); and
 - e. triose phosphate isomerase ribosomal binding site.
- 10. An expression construct for high level expression of truncated PspA1 (SEQ ID NO: 4),
5 in *Escherichia coli* which comprises:
 - a. gene encoding truncated *pspA1* (SEQ ID NO: 6);
 - b. pUC origin of replication;
 - c. kanamycin resistance gene (SEQ ID NO: 1);
 - d. *P_{T7}* promoter (SEQ ID NO: 11); and
 - 10 e. ribosomal binding site.
- 11. A recombinant host cell comprising the expression vector as claimed in any one of claims 2-10.
- 12. The recombinant host cell as claimed in claim 11, wherein the host is *Corynebacterium glutamicum* or *Escherichia coli*.
- 15 13. A method for high level expression of *Streptococcus pneumoniae* truncated PspA1 (Pneumococcal Surface Protein A1) which comprises culturing of *bacteria* transformed with an expression construct as claimed in any one of claims 2-10 and thereby purifying the expressed protein thereof.
- 14. A method of high level of expression and purification of the truncated PspA1 as
20 claimed in claim 13, wherein the yield of truncated *pspA1* is about 500 mg/L, about 400 mg/L, about 300 mg/L, about 250 mg/L, about 220 mg/L, about 200 mg/L, about 180 mg/L, about 160 mg/L, about 150 mg/L, about 120 mg/L, about 100 mg/L.
- 15. A pneumococcal conjugate vaccine comprising at least one polysaccharide from *Streptococcus pneumoniae* serotypes selected from 1, 2, 3, 4, 5, 6A, 6B, 6C, 6D, 7F, 8,
25 9N, 9V, 10A, 11A, 12F, 14, 15A, 15B, 15C, 16F, 17F, 18C, 19F, 19A, 20A, 20B, 22F, 23A, 23B, 23F, 24B, 24F, 31, 33F, 34, 35B, 35F, 38, 39 and 45, conjugated with truncated PspA1 as set forth in SEQ ID NO: 3 or SEQ ID NO: 4, or combination of truncated PspA1 and other carrier proteins selected from group comprising CRM197, tetanus toxoid, pertussis toxoid and PsaA.
- 30 16. The vaccine composition as claimed in claim 15, wherein the vaccine is a multivalent vaccine selected from 10 valent, 14 valent, 15 valent, 17 valent, 18 valent, 19 valent, 20 valent, 22 valent, 23 valent, 24 valent and 25 valent.
- 17. A multivalent conjugate vaccine as claimed in claim 15 comprising at least three

polysaccharides from *Streptococcus pneumoniae* serotypes 3, 6A and 6B conjugated to truncated PspA1 as set forth in SEQ ID NO: 3 or SEQ ID NO: 4, and *Streptococcus pneumoniae* serotypes 1, 4, 5, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F and 33F conjugated to CRM₁₉₇.

- 5 18. Use of truncated pneumococcal surface protein A1 (PspA1) comprising the amino acid sequence of SEQ ID NO: 3 or SEQ ID NO:4 as a carrier protein in a conjugate vaccine.