Protein Expression and Purification 78 (2011) 38–47

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Cloning and optimization of induction conditions for mature PsaA (pneumococcal surface adhesin A) expression in *Escherichia coli* and recombinant protein stability during long-term storage

Ariane Leites Larentis [a,*,1], Ana Paula Corrêa Argondizzo [a,1], Gabriela dos Santos Esteves [a], Ellen Jessouron [b], Ricardo Galler [a], Marco Alberto Medeiros [a,*]

[a] *Fundação Oswaldo Cruz (Fiocruz), Bio-Manguinhos/VDTEC (Vice-Diretoria de Desenvolvimento Tecnológico), Laboratório de Tecnologia Recombinante (LATER), Av. Brasil 4365, 21.040-360, Rio de Janeiro – RJ, Brazil*

[b] *Fundação Oswaldo Cruz (Fiocruz), Bio-Manguinhos/VDTEC (Vice-Diretoria de Desenvolvimento Tecnológico), Laboratório de Tecnologia Bacteriana (LATEB), Av. Brasil 4365, 21.040-360, Rio de Janeiro – RJ, Brazil*

ARTICLE INFO

Article history:
Received 3 December 2010
and in revised form 22 February 2011
Available online 6 March 2011

Keywords:
Recombinant PsaA expression
Experimental design
Central composite design (CCD)
Ion exchange purification
Protein stability
Long-term storage

ABSTRACT

The gene corresponding to mature PsaA from *Streptococcus pneumoniae* serotype 14 was cloned into a plasmid with kanamycin resistance and without a purification tag in *Escherichia coli* to express high levels of the recombinant protein for large-scale production as a potential vaccine candidate or as a carrier for polysaccharide conjugation at Bio-Manguinhos/Fiocruz. The evaluation of induction conditions (IPTG concentration, temperature and time) in *E. coli* was accomplished by experimental design techniques to enhance the expression level of mature recombinant PsaA (rPsaA). The optimization of induction process conditions led us to perform the recombinant protein induction at 25 °C for 16 h, with 0.1 mM IPTG in Terrific Broth medium. At these conditions, the level of mature rPsaA expression obtained in *E. coli* BL21 (DE3) Star by pET28a induction with IPTG was in the range of 0.8 g/L of culture medium, with a 10-fold lower concentration of inducer than usually employed, which contributes to a less expensive process. Mature rPsaA expressed in *E. coli* BL21 (DE3) Star accounted for approximately 30–35% of the total protein. rPsaA purification by ion exchange allowed the production of high-purity recombinant protein without fusion tags. The results presented in this work confirm that the purified recombinant protein maintains its stability and integrity for long periods of time in various storage conditions (temperatures of 4 or −70 °C using different cryoprotectors) and for at least 3 years at 4 or −70 °C in PBS. The conformation of the stored protein was confirmed using circular dichroism. Mature rPsaA antigenicity was proven by anti-rPsaA mouse serum recognition through western blot analysis, and no protein degradation was detected after long periods of storage.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Streptococcus pneumoniae is the causal agent for streptococcal pneumonia and other forms of invasive disease such as meningitis, bacterial sepsis and otitis media (middle ear infection). Approximately 5–10% of adults and 20–40% of children are asymptomatic carriers of this streptococcus species, and only a small percentage of carriers develop invasive disease. Because *S. pneumoniae* is responsible for more than 1 million deaths annually worldwide (mostly occurring in developing countries and in at-risk populations as young children, elderly and immune-compromised individuals) and, recently, is more frequently associated with antibiotic resistance, it is still a considerable health burden for society [1–3].

The Gram-positive serotyping of *S. pneumoniae* is based on surface carbohydrates. More than 90 serotypes are known, and the serotypes most frequently isolated from patients in Brazil were 1, 3, 4, 5 and 14 [4]. Three licensed vaccines against *S. pneumoniae* are commercially available: Pneumovax, Prevnar and Synflorix. Pneumovax (Merck), a 23-valent capsular polysaccharide vaccine licensed in Brazil since 1992, has a strong efficacy in adults but a poor efficacy in infants and young children (<2 years) because they lack mature B cells, which respond better to carbohydrate antigens [1,2,5,6]. The low immunogenicity of polysaccharide vaccines in children and the elderly has led to the development of polyvalent conjugate

* Corresponding authors. Fax: +55 21 38827176.
E-mail addresses: ariane@bio.fiocruz.br (A.L. Larentis), medeiros@bio.fiocruz.br (M.A. Medeiros).
[1] These authors contributed equally to this work.

1046-5928/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2011.02.013

vaccines, which comprise different polysaccharides conjugated to carrier proteins, such as tetanus or diphtheria toxoid [2]. Prevnar or Prevenar (Wyeth), a heptavalent conjugated vaccine licensed in Brazil since 2001, has a strong efficacy in infants and young children (~90%) but provides poor protection against less common serotypes (especially outside USA) [2,3,7]. The conjugated 10-valent vaccine is a commercially named Synflorix (GSK – GlaxoSmithKline Biologicals). Through technology transfer agreement between the Brazilian Ministry of Health/Fiocruz/Bio-Manguinhos and GSK, the 10-valent vaccine was incorporated into the public vaccination calendar of the Brazilian National Immunization Program (*Programa Nacional de Imunizações, PNI*) in March 2010.

The 10-valent vaccine under technology transfer by Bio-Manguinhos/Fiocruz presents a high-cost process production, and while protecting against main serotypes that cause disease in Brazil, the coverage of Brazilian serotypes by this conjugated vaccine is not complete. As an alternative to the expensive conjugate vaccines, diverse virulence proteins have been investigated in the last two decades as potential vaccine antigen candidates [1–3,5–11] or as carriers for polysaccharides, aiming to enhance coverage for serotypes absent from conjugate vaccines. The most important described proteins are: the pneumococcal toxin pneumolysin [12]; neuraminidases [13], such as NanA and NanB; autolysins [14], such as LytA, LytB and LytC; choline-binding surface proteins family, such as pneumococcal surface protein A (PspA) [15], whose choline-binding repeats are responsible for the attachment of PspA to the surface of the pneumococcus, PspC [16] and choline-binding protein A (CbpA) [17]; a lipoprotein called pneumococcal surface adhesin A (PsaA[2]) [18–21]; hyaluronate lyase (Hyl), a hyaluronidase that hydrolyses hyaluronan of the extracellular matrix of tissues and possibly contributes to the invasiveness of *S. pneumoniae* [22]; the putative proteinase maturation protein A (PpmA) [23]; PiuA and PiaA (formerly Pit1A and Pit2A), the lipoprotein components of two *S. pneumoniae* iron ABC transport systems called Piu (pneumococcal iron uptake) and Pia (pneumococcal iron acquisition) [24]; a family of cell surface-exposed pneumococcal homologous proteins named the pneumococcal histidine triad (Pht), which includes PhtA, PhtB, PhtD and PhtE [25], and *S. pneumoniae* histidine proteins (PhP), which includes PhpA, PhpB and PhpC [26] and have also been identified as the BVH-3 and BVH-11 families, respectively [27]; and the PavA protein (pneumococcal adherence and virulence factor A), which has been identified as a pneumococcal fibronectin-binding protein [28]. The majority of pneumococcal serotypes isolated to date express many of these immunogenic proteins [29]. More recently, new targets have been investigated for a pneumococcus subunit vaccine, including the proteins α-enolase (Eno) and the streptococcal lipoprotein rotamase A (SlrA) [30]; the heat shock protein ClpP protease [31]; the pneumococcal protective protein A (PppA), which is a surface-exposed protein [32]; the IgA1 protease, which cleaves the hinge region of the human immunoglobulin A1 heavy chain [33]; the polyamine transport protein D (PotD) [34] and the pilus proteins (the subunits RrgA, RrgB and RrgC), which contribute to adherence and virulence and elicit host inflammatory responses [35]. New potential genes also have been identified by exploiting the genome sequences of *S. pneumoniae* [7,9,36–38] by the screening of libraries from clinical isolates [39], reverse vaccinology [7,38] or proteomics strategies [40,41]. This group of potential targets includes the surface protein Spy0416 [40] and the putative cell wall-associated lipoate protein ligase (Lpl) [41], which were both identified by proteomic analyses; a pneumococcal collagen-like protein (PclA), which was identified by analysis of *S. pneumoniae* sequenced genomes [42]; and two protein antigens conserved among clinical isolates: the protein required for cell wall separation of group B streptococcus (PcsB) and serine/threonine protein kinase (StkP), which were identified by genomic surface libraries [43].

The pneumococcal proteins described above present a range of biologic activities, indicating that they act at different stages of the pathogenic process [5]. Immunization with each of these pneumococcal proteins alone, in combination with each other or in combination with capsular polysaccharide-protein conjugates has been shown to elicit a significant level of protection in animal models against one or more *S. pneumoniae* serotypes, and these proteins may be useful as pneumococcal vaccine components [44–46]. Many researchers have indicated that a vaccine cannot be developed with a single antigen, and the best approach is incorporating a mixture of pneumococcal proteins (multi-protein vaccine), which are selected by their ability to protect against pneumococcal bacteremia and sepsis, pneumonia, otitis and other forms of invasive disease, as well as against pneumococcal carriage. Different virulence factors could target more than one mechanism for pneumococcal invasion, adhesion, infection and colonization [11,29,44–46]. Combinations of diverse pneumococcal proteins may provide the best non-serotype-dependent protection against *S. pneumoniae*. Such proteins might serve as a vaccine by themselves but, more likely, could be used in conjunction with polysaccharide-protein conjugates or as carriers for polysaccharides [6,15] or other proteins or adjuvants, stimulating immune responses to non-immunogenic antigens [47].

One of these promising targets for protein vaccine development is the pneumococcal surface adhesin A (PsaA), a 37-kDa cell membrane-associated and surface-exposed lipoprotein, which is hydrophobic, immunogenic, genetically conserved and detected on the surface of all known *S. pneumoniae* serotypes (>90 isolates) [29,48–51]. PsaA presents an identity of 97–100% among the *S. pneumoniae* genomes deposited in the NCBI database. PsaA is essential for the virulence of *S. pneumoniae* [5,51,52]. The high degree of conservation of PsaA and its widespread occurrence among several pneumococcal serotypes makes it a suitable potential candidate antigen for a serotype-independent vaccine covering a broad range of pneumococcal infections [2,51]. PsaA remains an important vaccine target because to induce herd immunity against *S. pneumoniae*, it is crucial to elicit protection against carriage, as acquisition of pneumococcal infection is usually from carriers rather than from infected individuals [29,44,53]. Several studies indicate that PsaA presents a potential application as a vaccine target [2,21,29,44–47,50,53–57].

Based on the results that indicate PsaA as a potential target, this work shows the molecular cloning, expression in *E. coli* and its purification, aiming at the large-scale production of the recombinant protein for testing its potential as vaccine candidate or carrier for polysaccharide conjugation at Bio-Manguinhos/Fiocruz. The study of induction conditions to enhance expression levels of recombinant PsaA (rPsaA) in *E. coli* was performed by experimental design techniques using a central composite design. The purified mature protein was monitored in storage over 3 years to analyze the stability and integrity of the recombinant protein, which are as important as the levels of protein production.

[2] *Abbreviations used*: PsaA, pneumococcal surface adhesin A; CCD, central composite design; IPTG, isopropyl β-D-1-thiogalactopyranoside; TB, terrific broth; PBS, phosphate buffered saline; NCBI, National Center for Biotechnology Information; SDS–PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; BSA, bovine serum albumin; DEAE FF, diethylaminoethyl sepharose fast flow; BCA, bicinchoninic acid; PBSTB, PBS–Tween 20 (0.05%)–BSA (0.25%); IgG, immunoglobulin G.

Materials and methods

Plasmid and expression system

The *psaA* gene from *S. pneumoniae* serotype 14 (strain 1871 from SIREVA Project/Canada) was expressed in the *E. coli* BL21 (DE3) Star strain with IPTG induction of the plasmid pET28a (Novagen).

Bioinformatics and cloning

The software Vector NTI 8.0 was used to perform gene sequence analyses and restriction mapping as well as translated protein analysis for the presence of signal peptides (in the SMART interface, http://smart.embl-heidelberg.de/) and transmembrane regions (in the SOSUI interface, http://bp.nuap.nagoya-u.ac.jp/sosui/sosui_submit.html). The PsaA sequence from *S. pneumoniae* serotype 14 was obtained from the accession number **YP_001836340**.

Cloning of *psaA* from *S. pneumoniae* serotype 14 was carried out by high-fidelity PCR amplification of both the 930-bp and 870-bp fragments from full-length and mature *psaA* (for expression of the protein without the first 21 amino acids), respectively. Forward oligonucleotides for PCR amplification included full-length-psaA-F (CACCATGGAAAAATTAGGTACATTACTC), mature-psaA-F (CCCCATGGCTAGCGGAAAAAAAGATACAAC), and reverse oligonucleotides for both full-length/mature-psaA-R included GGCTCGAGTTATTTTGCCAATCCTCCAG, incorporating *Nco*I (CCATGG) and *Xho*I (CTCGAG) restriction sites (underlined), respectively. The restriction sites were incorporated to allow in-frame cloning of the PCR products into the corresponding restriction sites in the polylinkers of pET28a. Both full-length and mature *psaA* genes were cloned with the stop codon, avoiding the incorporation of the histidine tag of pET28a.

The PCR was performed as follows: approximately 35 ng of genomic DNA from *S. pneumoniae* serotype 14 was submitted to a denaturation step at 94 °C for 5 min, 30 cycles of 94 °C for 1 min, 52 °C for 30 s and 68 °C for 2 min and a finalization step at 72 °C for 7 min; 5 pmol of each primer, 1 U Taq Polymerase High Fidelity (Invitrogen), 2 mM $MgSO_4$ and 200 μM dNTP were used in each reaction (25 μL). The vector pET28a and PCR products were digested with *Nco*I and *Xho*I and spliced with T4 ligase (Promega). *E. coli* TOP10 (Invitrogen) was used for host cloning construction, and the recombinants were confirmed by agarose gel size, digestion and sequencing analysis. The recombinant plasmids were then used to transform *E. coli* BL21 (DE3) Star (Invitrogen), and cells were stored at −70 °C in 25% glycerol.

Expression

Recombinant *E. coli* BL21 (DE3) Star was cultivated in a 10-mL pre-inoculum with the inoculation of 1:1000 glycerol-stored cells in TB (Terrific Broth) medium (23.6 g/L yeast extract, 11.8 g/L tryptone, 9.4 g/L potassium phosphate dibasic (K_2HPO_4), 2.2 g/L potassium phosphate monobasic (KH_2PO_4)) supplemented with 50 μg/mL kanamycin, 1% glucose and 0.4% glycerol at 37 °C for 16 h at 200 rpm in 50-mL Erlenmeyer flasks. A total of 2% of this saturated pre-inoculum was diluted in 100 mL of TB medium in 500-mL Erlenmeyer flasks, resulting in an initial absorbance (measured at 600 nm) of approximately 0.1. Cells were grown at 37 °C and 200 rpm until they reached the exponential phase (approximately 0.75/0.85 measured by absorbance at 600 nm) and were subsequently induced with IPTG. Preliminary tests were performed with *E. coli* BL21 (DE3) Star/pET28a expressing full-length and mature *psaA* genes with 1 mM IPTG at 37 °C for 4 h. For the optimization of mature PsaA expression, the induction conditions (IPTG concentration, temperature and time after induction) were varied with the help of experimental design techniques according to the following description. The expression level at each time after induction was compared with the total protein extract from the culture before IPTG induction (uninduced sample). Cell growth at the final expression time was measured by absorbance at 600 nm ($Abs_{600\ nm}$).

Cells were harvested by centrifugation at 20,817×*g* (14,000 rpm in an Eppendorf 5417R centrifuge) for 5 min at 4 °C. The pellets were stored at −20 °C for further analysis by SDS–PAGE. *E. coli* BL21 (DE3) Star/pET28a was used as negative control.

Chemicals

Bacto™ yeast extract and tryptone were purchased from BD (Becton, Dickinson and Company); potassium salts (K_2HPO_4 and KH_2PO_4), glucose and NaCl were from Merck; glycerol was from Invitrogen; kanamycin was from Sigma; and IPTG was from Promega.

Experimental design

The analysis of the effects of three independent variables (IPTG concentration, temperature and time after induction) on the expression yield and productivity of soluble mature PsaA in recombinant *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* was performed using a central composite design, consisting of all of the combinations of the three variables and the central condition performed in triplicate to analyze experimental error and to check the curvature of responses.

The statistical evaluation of the IPTG concentration, temperature and time effects on the expression yield and productivity of PsaA was performed with the help of STATISTICA 9.1 (Statsoft, USA) using the normalized variable at levels of −1 (lower value of the experimental conditions used), +1 (higher value of the experimental conditions) and 0 (central point condition). Values of independent variables and corresponding levels used in the central composite design (normalized in −1, 0 and +1) are described in Table 1. The normalization of each analyzed variable permits the comparison of their effects without the influence of the magnitude of the interval employed for each. The significance of each linear effect and interactions was determined with the Student's *t*-test, and a 0.1 probability level was used.

Preparation of total protein extract, solubility fractions and SDS–PAGE

The pellets from 1-mL samples of expression cells were resuspended in 20 mM Tris/1 mM EDTA (pH 8) according to the ratio of 25 μL of sample buffer for each $Abs_{600\ nm}$ = 0.1. This cell suspension was subjected to lysozyme treatment (1 mg/mL for 30 min on ice) and disrupted by sonication on ice (5 cycles of 10-s pulses with 30-s intervals, 30% amplitude using an ultrasonic cell disruptor Sonics & Material, Inc.) to obtain the total protein extract for solubility analysis. The total protein extract was centrifuged at 20,817×*g* (14,000 rpm) for 10 min at 10 °C using an Eppendorf

Table 1
Experimental design levels. Values of independent induction variables (IPTG concentration, temperature and induction time) and corresponding levels used in a central composite design (normalized in −1, 0 and +1). The central point (0) corresponds to intermediate values between the −1 and +1 interval of each variable.

Variable	−1	0	+1
Inducer concentration (*IPTG*) (mM)	0.1	0.55	1
Temperature (*T*) (°C)	20	28.5	37
Induction time (*t*) (h)	6	14	22

5417R centrifuge, and the soluble fraction was obtained in the supernatant. The precipitate, including cell and membrane residues, and proteins expressed into inclusion bodies were resuspended in sample buffer for the preparation of the insoluble fraction. Solubility was analyzed by SDS–PAGE.

For SDS–PAGE analysis, the pellets and solubility fractions were resuspended in a sample buffer (60 mM Tris–HCl, pH 6.8, 2% SDS, 0.5% bromophenol blue, 10% glycerol, 5% 2-mercaptoethanol) and heated for 5 min at 95 °C. Then 12.5% SDS–PAGE was performed with 10 μL of total protein extract in a Mini-PROTEAN 3 cell (Bio-Rad) and run for 90 min at 120 V. The gel was stained with Coomassie blue (50% methanol, 10% acetic acid and 0.1% Coomassie brilliant blue R-250) and destained with 30% ethanol.

Quantitation of PsaA expression by densitometry analysis

The areas of the bands corresponding to mature rPsaA expression in SDS–PAGE were analyzed by densitometry in a Bio-Rad GS-800 Calibrated Densitometer, quantified by the software QuantiOne 4.4.1 and normalized to standard markers of 1 μg of each of the following proteins: 60-kDa BSA, 45-kDa egg albumin, 24-kDa trypsinogen, 18.4-kDa β-lactoglobulin and 14.4-kDa lysozyme. The concentration of expressed protein was obtained using the equation (mg PsaA/L) = ($Abs_{600\ nm}$ × Band in densitometry)/4, where 4 is the concentration factor used in the preparation of total protein extract samples.

Experimental design validation and glucose concentration effects

Experimental design results were validated at 25 °C with 0.1 mM IPTG for 16 h of induction in TB medium, with six replicates, to estimate process variability. In these conditions, PsaA expression was tested in TB with 0.1% (1 g/L) and 0.5% (5 g/L) glucose.

Purification

A 100-mL pellet (2.4 g wet weight cells) from the cultivation of *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* cells, grown in validated conditions, were centrifuged (3220×g or 4000 rpm for 20 min at 10 °C using an Eppendorf 5810R centrifuge) and stored at −20 °C. The stored pellet used for the purification procedures was resuspended in 20 mM Tris HCl and 1 mM EDTA, pH 8.0 (in a volume calculated according to the ratio of 5 mL for each 1 g of wet weight cell pellet), and disrupted by sonication on ice with 30-s pulses (5 cycles with 30% amplitude using an ultrasonic cell disruptor from Sonics & Material, Inc.) until the solution turned clear to obtain the total protein extract, as described above.

The total protein extract was harvested by centrifugation (18,514×g or 12,000 rpm for 20 min at 4 °C using an Eppendorf 5417R centrifuge) to obtain the soluble fraction, which was filtered with a 0.22-μm low protein-binding membrane (Millipore) to submit to purification in an ÄKTA purifier 10 system (Amersham Biosciences). NaCl was added to the sample up to 40 mM, which was submitted to a 5-mL weak anion exchange column (Sepharose – Hitrap DEAE FF/GE Healthcare), previously equilibrated with the same buffer used in the sample. Bound protein was eluted by washing with 20 mM Tris HCl and 150 mM NaCl (pH 8.0). Samples of all purification steps were quantified by BCA (Bicinchoninic Acid Kit for Protein Determination, Sigma–Aldrich) using BSA as a standard protein, and analyzed by 12.5% SDS–PAGE stained with Coomassie blue. Densitometry was used to evaluate the amount, yield and purity of PsaA. The purification procedures were performed in triplicate to evaluate experimental error.

Purified mature rPsaA stability and integrity analyses

Mature rPsaA, purified by ion exchange in a Hitrap DEAE FF column, was submitted to PBS (pH 8.0) dialysis (PBS composition: 8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L Na_2HPO_4 and 0.24 g/L KH_2PO_4). After this procedure, the recombinant protein was stored at 4 or −70 °C.

Samples were quantified by BCA and protein integrity was analyzed by 12.5% SDS–PAGE before and after the dialysis and after several months of storage in PBS at 4 or −70 °C. Storage of recombinant mature PsaA was also monitored over five days at −70 °C in PBS with the addition of 10% glycerol and 10% sucrose as cryoprotectors.

The conformation of purified mature rPsaA, dialyzed and stored for 3 years at −70 and 4 °C, was monitored by circular dichroism using a Jasco J-185 CD Spectrometer. The analyses were performed with proteins at concentrations of approximately 0.160 mg/mL in PBS that was diluted 10-fold.

Western blot analysis

Total protein extracts of full-length PsaA (with the peptide signal) and mature PsaA (without the peptide signal), expressed for 4 h at 37 °C for the preliminary tests, were evaluated on an 18% SDS–PAGE. The samples were diluted five times relative to the preparation of the total protein extract, described before (25 μL of sample buffer for each 1 mL of cell pellet at $Abs_{600\ nm}$ = 0.1). These proteins were subjected to western blot analysis with anti-rPsaA mouse serum (kindly provided by Instituto Butantan).

The protein rPsaA, kept at a 4 °C for 125 days at a final concentration of 0.43 mg/mL as quantified by BCA, was used for western blot analysis with anti-rPsaA mouse serum. A total of 500 and 250 ng of protein was boiled in sample buffer, as described above, cooled and loaded onto an SDS–PAGE and run for 90 min at 120 V. The *S. pneumoniae* serotype 14 pellet was resuspended in sample buffer, diluted 1:5 and 1:10 in the buffer and boiled. The same was done with *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* pellets, both either uninduced or induced for 16 h. The total protein extract used was expressed under the validation conditions of *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* induced at 25 °C for 16 h with 0.1 mM IPTG in TB medium, compared to the uninduced sample (around 2 h of growth).

For both western blots, the proteins were transferred to a nitrocellulose membrane (Bio-Rad) at 90 V for 90 min. Membranes were blocked with PBS–skim milk (4%)–Tween 20 (0.05%) overnight at 4 °C. A primary antibody to rPsaA diluted 1:2000 in PBS–Tween 20 (0.05%)–BSA (0.25%) (PBSTB) was applied to the membrane and incubated for 2 h at room temperature. Membranes were washed three times for 20 min each with PBS–0.05% Tween 20. Conjugated anti-mouse IgG alkaline phosphatase (Sigma A3688) was then applied to the membrane at a 1:30,000 dilution in PBSTB and incubated for 1 h at room temperature. Membranes were washed as described above, subjected to a final wash with PBS and revealed with Western Blue Stabilized Substrate for Alkaline Phosphatase (Promega S3841).

Results

Bioinformatics, cloning and preliminary expression of full-length and mature PsaA

Bioinformatic predictions obtained by the software Vector NTI 8.0 confirmed the presence of a signal peptide at the N-terminal of the PsaA sequence. The program also indicated that amino acids from 4 to 21 in the *psaA* gene may be part of a transmembrane region in the PsaA protein sequence (LGTLLVLFLSAIILVACA). These

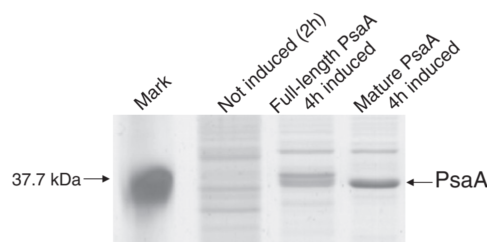

Fig. 1. Electrophoretic analysis of recombinant proteins. SDS–PAGE (18%) of total protein extracts of full-length (with the peptide signal) and mature (without the peptide signal) PsaA expressed in *E. coli* BL21 (DE3) Star/pET28a for 4 h at 37 °C. The samples are diluted five times relative to the standard preparation for the gel (25 μL of sample buffer for each 1 mL of cell pellet at $Abs_{600\ nm} = 0.1$). All bands were recognized by anti-rPsaA mouse serum in western blot analysis.

results indicate that this region (4–21 aa) is likely anchored to the bacterial wall, as predicted by bioinformatic results obtained with Vector NTI 8.0 interfaces. Therefore, two approaches for *psaA* cloning were developed: the design of oligonucleotides for the amplification of the entire gene, from the ATG initiation codon to the TAG termination codon (full-length protein), and a shorter version removing the 21 initial amino acids, including those corresponding to the transmembrane region (mature protein).

The construction pET28a/*psaA-full-length* was expressed in two different bands, one around 37 kDa and another near 32.5 kDa, corresponding to the expected sizes of the full-length and mature proteins, respectively (Fig. 1). An 18% SDS–PAGE was used to separate two bands from the full-length protein because these bands were superposed on a 12.5% and 15% gel concentrations used formerly. All bands were recognized by anti-rPsaA mouse serum in western blot analysis. Both full-length protein bands as well as the mature protein were soluble in 20 mM Tris/1 mM EDTA (pH 8). Densitometry analysis of these bands indicated that the 32.5-kDa band of the mature protein was at least twofold higher than the 37-kDa band of the full-length PsaA.

Experimental design results for mature PsaA expression in E. coli

The evaluation of the different conditions for the expression of recombinant mature PsaA from *S. pneumoniae* in *E. coli* BL21 (DE3) Star was performed in shaking flasks using a central composite design for three variables: IPTG concentration (0.1 and 1 mM), temperature (20 and 37 °C) and induction time (4 and 22 h), with triplicate experiments at the central point (0.55 mM IPTG, 28.5 °C and 14 h), as described above. Comparison of each induction condition was performed through 12.5% SDS–PAGE analysis of the 32.5-kDa expressed band corresponding to mature PsaA. No band was detected at the same size as PsaA in the total protein extract of *E. coli* BL21 (DE3) Star with pET28a plasmid without the gene (negative control).

In all conditions, recombinant protein was obtained in the soluble form in 20 mM Tris/1 mM EDTA (pH 8). The solubility analysis of the protein expressed in all conditions of IPTG concentration, temperature and time in SDS–PAGE indicated that the insoluble fraction was insignificant. Based on these results, the effect of each variable was obtained by the densitometry analysis of total protein extracts on a 12.5% SDS–PAGE. The areas of the bands on the SDS–PAGE corresponding to the recombinant mature PsaA (32.5 kDa), as calculated in each expression condition by densitometry analysis, are presented in Table 2.

Results presented in Table 2 indicate that cell growth ranged from 1.58 to 6.06 (measured by absorbance at 600 nm), and the production of mature PsaA ranged from 15 to 771 mg/L (resulting in productivities from 0.7 and 35 (mg/L)/h, respectively) in the various tested conditions. At 37 °C, for both tested IPTG concentrations, the expression was reduced over extended periods. Nevertheless, at 20 °C, high expression levels were observed for all induction times and IPTG conditions. At lower temperatures (20 °C), the productivity (concentration of expressed protein per induction time) was kept constant when comparing 6 h with 22 h of induction time. However, the productivity in longer induction times was significantly reduced at higher temperatures (Table 2). Experimental error was evaluated at the central point conditions of the experimental plan (28.5 °C, 0.55 mM IPTG and 14 h of induction). At these conditions, results for cell growth were 4.02 ± 0.2 (measured by absorbance at 600 nm), the concentration of mature PsaA was 201 mg/L ± 18, and productivity was 14.3 (mg/L)/h ± 1.3. The comparison of standard deviations with the triplicate average of mg PsaA/L was less than 10%.

The comparison of the experiments performed with 1 mM and 0.1 mM IPTG showed that similar expression levels were obtained at 20 and 37 °C for all induction times, indicating that the IPTG concentration did not have a significant influence on the expression of the recombinant protein. Inducer concentration, on the other hand, had an influence on the cell growth, which means that a reduction of IPTG to 0.1 mM allowed increased cell growth to be achieved (Table 2).

Statistical analyses of the effects of the variables (IPTG concentration, temperature and time induction) on recombinant mature PsaA expression are presented at Table 3. The adopted experimental design presented no correlation between the variables, and it was possible to evaluate each effect independently.

Table 3 indicates that temperature and time and the interaction between these two variables had statistically significant influence on recombinant PsaA expression yield ($p < 0.1$). The temperature effect was negative (−329, $p = 0.005$), indicating that lower temperatures allowed enhancement of recombinant protein expression. Induction time had a positive effect (171, $p = 0.044$), indicating that longer induction times also increase PsaA expression, mainly because they permit increased cell growth. However, the effect of the interaction between temperature and time was negative (−274, $p = 0.010$); i.e., for the higher temperatures tested, longer times did not favor protein expression, which was confirmed by the experiments performed at 37 °C, as shown above.

For PsaA productivity in (mg/L)/h, both temperature and time had significant and negative effects ((−18.2, $p = 0.032$) and (−14.7, $p = 0.060$), respectively). Statistical analysis indicated that both variables should be reduced to obtain higher levels of PsaA productivities. The concentration of IPTG did not significantly influence PsaA yield expression (−46, $p = 0.480$) or productivity (−3.2, $p = 0.601$), neither its interactions with temperature or time, indicating that the lowest inducer concentration (10-fold lower than the usual concentration) could be used.

Experimental design validation results and glucose concentration effects

Considering that lower temperatures and longer induction times enhanced the production of mature PsaA in recombinant *E. coli* and that IPTG concentration did not influence PsaA production, the experimental plan was validated at 25 °C for 16 h induction with 0.1 mM IPTG. Validation was performed in the same TB culture medium with the addition of 1% glucose, 0.4% glycerol and 50 μg/mL kanamycin. The expression results with 0.1 mM IPTG at 25 °C for 16 h of induction in TB were 718 ± 113 mg/L PsaA and 5.5 ± 0.5 cell growth (measured by absorbance at 600 nm). In comparison to 1 mM IPTG at 37 °C for 6 h of induction, the usual conditions for recombinant protein expression in *E. coli*, PsaA production was 135 mg/L, which is a significant yield increase of

Table 2
Recombinant protein yields. Results for cell growth measured by absorbance at 600 nm ($Abs_{600\ nm}$), recombinant PsaA concentration yield (mg/L) and the productivity of the expressed protein per induction time measured in (mg/L)/h, obtained from densitometry analysis with different expression conditions of IPTG concentration, temperature and induction time defined using a central composite design for three variables, comprising a two-level full-factorial design (runs 1–8) and triplicates at the central point (runs 9–11). Run 12 was made with *E. coli* BL21 (DE3) Star/pET28a (negative control), and no expression was observed at the same size as mature PsaA.

Run	Inducer concentration (*IPTG*)		Temperature (*T*)		Induction time (*t*)		Cell growth ($Abs_{600\ nm}$)	Yield (mg/L)	Productivity (mg/L)/h
1	0.1 mM	−1	20 °C	−1	6 h	−1	3.23	208	34.7
2	0.1 mM	−1	20 °C	−1	22 h	+1	6.06	771	35.0
3	0.1 mM	−1	37 °C	+1	6 h	−1	1.95	181	30.2
4	0.1 mM	−1	37 °C	+1	22 h	+1	4.96	15	0.7
5	1 mM	+1	20 °C	−1	6 h	−1	1.91	217	36.2
6	1 mM	+1	20 °C	−1	22 h	+1	4.20	545	24.8
7	1 mM	+1	37 °C	+1	6 h	−1	1.58	135	22.5
8	1 mM	+1	37 °C	+1	22 h	+1	5.21	95	4.3
9	0.55 mM	0	28.5 °C	0	14 h	0	4.19	219	15.6
10	0.55 mM	0	28.5 °C	0	14 h	0	3.81	183	13.1
11	0.55 mM	0	28.5 °C	0	14 h	0	4.05	200	14.3
12 (C−)	0.55 mM	0	28.5 °C	0	14 h	0	3.18	–	–

Table 3
Statistical evaluation. Effects of *IPTG* concentration, temperature and induction time on mature PsaA expression yield in mg/L (mass of protein obtained by volume of culture medium) and on productivity in (mg/L)/h (concentration of expressed PsaA per induction time). Variables with $p < 0.1$ were considered statistically significant by the Student's *t*-test (in bold). The *p*-value indicates the significance level of the variables and their interactions.

Factor	Effects	Standard errors	*t* (4)	*p*-value
(*a*) *PsaA yield* (*mg/L*)				
Mean	**252**	**25**	**10.0**	**0.001**
Inducer concentration (IPTG)	−46	59	−0.8	0.480
Temperature (*T*)	**−329**	**59**	**−5.6**	**0.005**
Time (*t*)	**171**	**59**	**2.9**	**0.044**
IPTG by *T*	63	59	1.1	0.346
IPTG by *t*	−27	59	−0.5	0.667
T by *t*	**−274**	**59**	**−4.7**	**0.010**
(*b*) *PsaA productivity* (*mg/L*)/*h*				
Mean	**21.0**	**2.4**	**8.7**	**0.001**
Inducer concentration (IPTG)	−3.2	5.6	−0.6	0.601
Temperature (*T*)	**−18.2**	**5.6**	**−3.2**	**0.032**
Time (*t*)	**−14.7**	**5.6**	**−2.6**	**0.060**
IPTG by *T*	1.2	5.6	0.2	0.844
IPTG by *t*	−0.1	5.6	−0.02	0.984
T by *t*	−9.2	5.6	−1.6	0.180

five- to sixfold. In validation condition, the productivity was enhanced to around 45 (mg/L)/h.

At 25 °C and 16 h of induction with 0.1 mM IPTG, tests were performed in TB with the addition of different glucose concentrations (two- and tenfold lower than 1%). The results indicated that there was no expression leak and that the same PsaA level of production was obtained with 0.1%, 0.5% and 1% glucose.

Purified mature rPsaA stability and integrity analysis

Recombinant mature PsaA was purified by ion exchange in a DEAE FF column. Evaluation of total protein amount, PsaA yield and purity associated to each purification step is shown in Table 4. The parameters yield and purity were calculated according to Burgess [58]. The overall PsaA yield indicated that around 55% of the recombinant protein was recovered in an 85% purity of pooled fraction, while a 65% recovery of the protein applied in the DEAE column in the soluble fraction was obtained. Some variations observed in Table 4 were due to the different methodologies used for quantitation. The experimental error, evaluated by means of the purification procedures performed in triplicate, was estimated to be around 15%.

The purified protein was subjected to PBS pH 8.0 dialysis and stored at 4 °C and −70 °C using different cryoprotectors. Protein

Table 4
Purification table. Total protein quantitation performed by BCA method using BSA as a standard protein. PsaA amount, yield and purity of each purification step obtained by densitometry. Results obtained from 2.4 g of wet weight *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* cell pellet (from 100 mL of bacterial culture in validation expression conditions).

Step	Total protein (mg)	PsaA (mg)	Yield (%)[a]	Purity (%)[b]
Total protein extract	277	72	100	30
Soluble fraction	126	61	85[c]	35
DEAE column (pooled peak)	47	40	55	85

[a] The yield at each step was defined as the amount of PsaA at that step divided by the amount in the first step, considered 100% [58].
[b] The purity was determined as the amount of the stain associated with the PsaA band as a fraction of the stain associated with all the bands on the SDS–PAGE [58].
[c] Protein loss in this step was due to centrifugation and filtration procedures. PsaA was not detected on the insoluble fraction.

quantitation and SDS–PAGE (Fig. 2) of each sample were performed to analyze the stability and integrity of the protein during long-term storage. There was no significant loss of recombinant protein in the dialysis. No remarkable protein degradation was verified in samples stored in all conditions described above, as shown by SDS–PAGE analysis (Fig. 2). After several months of storage, a similar intensity was observed for the band corresponding to mature PsaA. The BCA quantitation of purified mature PsaA indicated that the samples maintained a concentration of 0.47 mg/mL at −70 °C and 0.43 mg/mL at 4 °C during storage. These results were confirmed by comparison with a BSA standard curve in the SDS–PAGE (Fig. 2). The errors associated with the BCA protein quantitation method were estimated to be approximately 15%, and no statistical differences were verified in samples stored for long time periods at the different conditions tested.

No considerable loss of protein was observed after cycles of freeze/thaw or after centrifugation (Fig. 2). Protein stability was confirmed for at least 36 months of storage. No remarkable differences were detected at either storage temperatures or with any of the cryoprotectors tested at the temperature of −70 °C (PBS, PBS + 10% glycerol and PBS + 10% sucrose).

Maintenance of protein conformation of purified mature rPsaA was confirmed by circular dichroism after 3 years of storage at −70 or 4 °C in PBS. These results confirm that the protein presents high stability at 4 and −70 °C in PBS for a long period of time.

Western blot analysis

Recombinant PsaA, stored at 4 °C for 125 days, was submitted to western blot analysis using an anti-rPsaA mouse serum. This

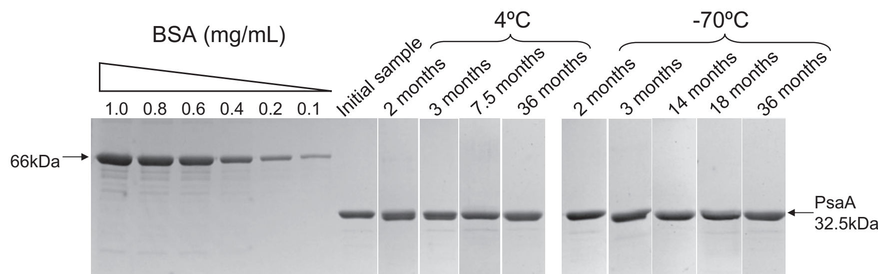

Fig. 2. Purification and stability of recombinant PsaA. Ion exchange-purified mature rPsaA stored in PBS (pH 8.0) at 4 °C or −70 °C at different time periods. BSA standard curve (from 1 to 0.1 mg/mL) was used to estimate yields and integrity of the recombinant protein in SDS–PAGE (12.5%).

serum recognizes purified mature rPsaA, recombinant protein expressed in *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* and the protein in the *S. pneumoniae* serotype 14 extract. These results confirmed there was no protein degradation since only one band was detected in all lanes, which corresponds to the expected size for mature rPsaA, 32.5 kDa (Fig. 3). There was no detectable rPsaA band in the *E. coli* total protein extract not induced with IPTG (negative control). No bands were detected in the western blot using pre-immune serum (negative control).

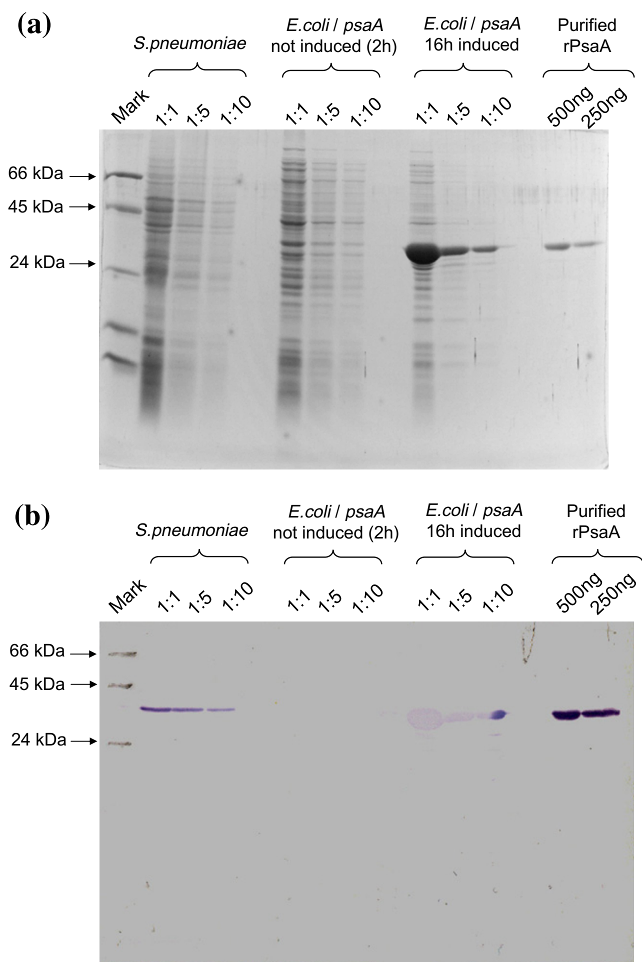

Fig. 3. Antigenicity of recombinant PsaA. SDS–PAGE (12.5%) analysis of total protein extracts from *S. pneumoniae* serotype 14, *E. coli* BL21 (DE3) Star/pET28a/*psaA-mature* (uninduced and induced at validation conditions: 25 °C for 16 h with [IPTG] = 0.1 mM in TB medium) and ion exchange-purified mature rPsaA. (a) Coomassie blue stain and (b) western blot using an anti-rPsaA mouse serum at a dilution of 1:2000.

Discussion

The immunogenic lipoprotein PsaA remains a candidate for a protein vaccine against *S. pneumoniae* [47,51] due to its potential to reduce nasopharyngeal colonization and its conservation among all known species of this pathogen. Several studies employing different vaccine strategies indicate that PsaA is effective at preventing colonization and is protective against carriage [2,6,21,29,44–46,53–56]. Moreover, pneumolysin and PspA have demonstrated the ability to elicit protection against invasive infections [2,29,44,46]. Studies of mouse immunization with PsaA and other antigens, such as detoxified pneumolysin, PspA and PspC, present promising protective results [2,29,44–46,54], supporting PsaA as a suitable target for pneumococcal vaccine development. Because PsaA is also present with a high degree of conservation among the viridans group of streptococcal species [49,57], its use as a vaccine target may raise concerns as to possible alterations in the normal microbiota. Recent work demonstrates that the induction of an immune response against PsaA can be protective against colonization and has a negligible impact on the natural nasopharyngeal and oral microbiota of mice [57].

Studies report that PsaA is expressed in low quantities in its native form in *S. pneumoniae* and that the yield of purified PsaA is 100-fold higher in heterologous expression systems in *E. coli* in comparison to native PsaA [29]. In this work, PsaA from *S. pneumoniae* serotype 14, the most prevalent in Brazil [59] and also one of the most prevalent in the world [60], was cloned for expression at high levels in *E. coli* with the intention of generating a stable protein to develop a recombinant vaccine against pneumococcal diseases. The cloning strategy used a plasmid without ampicillin gene and without tags for purification avoiding possible problems for human use. These characteristics highlight the advantageous and industrial aspects of our plasmid construction.

Cloning was performed by two approaches: producing the full-length protein (encoded by an open reading frame of 930 bp [29,49]), with the signal peptide identified as a transmembrane region, and producing the mature protein without this leader sequence, which was also identified by bioinformatics. Cleavage of this signal sequence results in a mature protein with 290 amino acids (with a theoretic molecular weight of 32.5 kDa and a hydrophilic character), with an isoelectric point of 5 [29,51]. The results obtained for rPsaA solubility indicated that the mature protein, without the transmembrane region (signal peptide), is retained in the soluble fraction. This result corroborates the bioinformatic prediction that overall the mature protein presents a hydrophilic character. Although this leader sequence is described as hydrophobic in the literature [29], which corroborates our *in silico* identification of this signal peptide as a transmembrane region (LGTLLVLFLSAIILVACA), the full-length PsaA was also obtained in soluble form. With the expression of the full-length protein, two bands were detected in total protein extracts, both recognized by anti-rPsaA mouse serum in western blot analysis. One explanation

for this observable fact is that the *E. coli* apparatus cleaves the peptide sequence of the full-length PsaA. Densitometry analysis of these bands indicated that the presence of the signal peptide significantly reduces the expression of the protein (about twofold) when comparing the 37-kDa full-length PsaA with the 32.5-kDa mature protein. The low level of expression of rPsaA with the lipoprotein signal sequence in *E. coli* is also described in the literature. It is discussed that the native PsaA leader peptide interferes with the production of large quantities of functional PsaA in *E. coli*, which could reduce or even hinder the expression of the recombinant protein in this host [29]. Therefore, the mature PsaA clone was used for optimization experiments.

Other studies also report the cloning of PsaA in *E. coli* without its signal sequence [29,50,53,55,61]. To achieve higher levels of expression of the functional protein in *E. coli*, the native leader sequence of rPsaA from *S. pneumoniae* type 6B was replaced by the *Borrelia burgdorferi* OspA (outer surface protein A) signal peptide since it is shorter and less hydrophobic than the PsaA signal peptide [29]. However, the hybrid construct expressed less protein than that produced without the peptide signal. Cullen et al. [61] expressed mature PsaA in a vector (pDUMP) designed specifically for recombinant lipoprotein expression in the *E. coli* system to produce proteins with a near-native conformation and with greater immunogenic potential. The plasmid permits the expression of fusion proteins containing the *E. coli* major outer membrane lipoprotein (Lpp) signal sequence, the lipoprotein signal peptidase (LSP) recognition site and the outer membrane targeting signal. A high level of expression of antigenic PsaA exposed on the *E. coli* cell surface was obtained [61]. Arêas et al. [50] expressed PsaA from *S. pneumoniae* serotype 6B without its signal sequence in highly soluble levels in the salt-induced *E. coli* BL21 (SI) expression system but bearing a histidine tag for one-step purification. Gor et al. [53] expressed PsaA without the first 21 amino acids from *S. pneumoniae* strain A66.1 and purified the protein using metal affinity chromatography from *E. coli* BL21 (DE3) pLysS/pET29b+. The levels of rPsaA obtained in these approaches are similar to those described here without any purification tag, being therefore more appropriate for a product intended for human use. Seo et al. [55] expressed the functional *psaA* gene (i.e., without its signal sequence) from *S. pneumoniae* type 19F in *E. coli* XL1-Blue in a triple fusion protein (rPsaA-intein-chitin binding domain). Arêas et al. [50] and Gor et al. [53] report that this protein without its signal sequence was also expressed in a soluble form, as described in our work. In the construct developed by Seo et al. [55], the protein was obtained in inclusion bodies when induced at 37 °C, partially in the insoluble fraction at 30 and 25 °C and only in the soluble form at 20 °C. Altogether, the higher levels of soluble mature PsaA in comparison to full-length PsaA led our group to use the protein without its signal sequence.

Although PsaA has been expressed in different heterologous systems and in a larger number of immunologic assays, there are, to our knowledge, no reports in the literature that approached the expression optimization enabling the development of a process to produce recombinant PsaA in *E. coli*. The experimental design techniques are able to evaluate the effect of induction variables, identify their interactions and predict conditions that optimize the production of recombinant protein, enhancing expression yields and reducing costs for bioreactor applications. In all conditions in which the expression was performed, including different temperatures, times and IPTG and glucose concentrations, the mature PsaA protein was obtained in the soluble fraction. Even at 37 °C, the protein was not significantly segregated into inclusion bodies, as is usual in *E. coli* systems for overexpressed recombinant proteins [62]. However, at 37 °C, for both tested IPTG concentrations, the expression was reduced over extended induction times, indicating a possible effect of recombinant protein degradation. Degradation of heterologous expressed proteins after long cultivation processes is likely due to the action of proteinases [62]. On the other hand, PsaA is expressed for long periods, maintaining productivity at the lower temperatures tested, without the degradation at long induction times observed at 37 °C, indicating that low temperatures favor the improvement of PsaA expression. In summary, longer induction times can be used to express mature PsaA at lower temperatures, confirming the interaction between these two variables. Therefore, the validation of induction conditions for IPTG concentration, temperature and time led us to perform the recombinant protein induction at 25 °C for 16 h with 0.1 mM IPTG in TB medium, to obtain economically feasible process conditions and optimize cost reductions. At these conditions, levels of mature PsaA expression in *E. coli* BL21 (DE3) Star by pET28a induction with IPTG were obtained in the range of 0.8 g/L of culture medium, which is in the range of the highest yields of soluble recombinant proteins described in the literature [62], with 10-fold lower inducer concentrations than has been previously employed, which is a major contribution to the development of a process with low production costs. The scale-up of protein production for bioreactors will enable us to achieve increased cell growth (by adjustment of aeration/stirring), and inducer concentrations should be adjusted to identify the lowest possible concentration that will avoid high costs and negative effects on cell growth. Mature rPsaA expressed in *E. coli* BL21 (DE3) Star comprised around 30% of the total protein, which is similar to results obtained by Arêas et al. [50] for this protein in the NaCl-induced *E. coli* strain BL21 (SI).

At the validation condition (25 °C and 16 h of induction with 0.1 mM IPTG), tests performed in TB with the addition of different glucose concentrations indicated that there was no expression leak and that the same PsaA level of production was obtained with 0.1%, 0.5% and 1% glucose. These experiments were made to evaluate the repression of the T7 promoter (avoiding expression before IPTG addition) and to insure growth without glucose limitation. Decreased glucose levels did not reduce bacterial growth in this condition of rich medium, i.e., medium with a high yeast extract concentration to insure an adequate carbon and nitrogen source.

rPsaA purification by ion exchange allowed the production of high quantities of recombinant protein without fusion tags, at roughly 55% recovery in an 85% purity fraction. These results are comparable to the homogeneities described in the literature for the expression of PsaA in *E. coli* with the help of purification tags [50,53]. These tags should be avoided in human health products, as described above, allowing the expression of the recombinant target with a conformation closer to the original folded protein. One of the major advantages of our construct is its ability to produce a protein ready for use in biopharmaceutical applications, eliminating the step of removing the expression tag after purification, which increases process costs.

The results presented in this work confirmed that the purified recombinant protein maintains its stability and integrity for at least 3 years at 4 and −70 °C in PBS, which is promising for the use of rPsaA as a product intended for human health. Mature rPsaA antigenicity was demonstrated by anti-rPsaA mouse serum recognition of the protein in *S. pneumoniae* serotype 14 by western blot analysis. Data regarding protein conformation comparing the original sample and after long-term storage were confirmed using circular dichroism. The results obtained by SDS–PAGE, protein quantitation and circular dichroism analyses did not indicate protein degradation during long-term storage. For industrial purposes, storage conditions are as important as expression levels when evaluating the potential of a protein as a cost-effective vaccine because one of the major problems is recombinant protein stability, although these analyses are frequently neglected. Many proteins for biopharmaceutical use have failed in the early stages of

development due to protein instability during purification or storage conditions.

The experimental design tool to optimize protein expression and storage conditions to insure stability and integrity for long time periods can be applied to other targets to optimize process productions in recombinant systems and to reduce costs per target once an effective vaccine might contain multiple proteins. *E. coli*, proposed as a heterologous host to produce these targets, is still the most frequently used expression system because of its low cost and high level of production of recombinant proteins. Among the many producing systems available for heterologous expression, *E. coli* remains one of the most attractive due to its ability to grow rapidly and at high cell densities on inexpensive substrates, its well-characterized genetics and the availability of an increasingly large number of cloning vectors and mutant host strains [62]. In this way, the results presented here suggest the potential to produce *S. pneumoniae* antigens in *E. coli* as an alternative because of its capacity to obtain high levels of soluble proteins. Our results concerning the optimization of PsaA in shaking flasks demonstrate the potential to develop a product and scale-up protein production for bioreactors to achieve higher levels of PsaA expression in the soluble form in *E. coli*.

An efficacious vaccine against pneumococcal diseases is one of the priorities for the public health system in Brazil. Despite the recent technology transfer agreement between Fiocruz/Bio-Manguinhos and GSK for the 10-valent conjugated polysaccharide vaccine, a less expensive alternative protein vaccine is still being studied in the Technological Development Department at Bio-Manguinhos. As discussed above, the requirements for potential vaccine antigens candidates include proteins evolved in the different pneumococcal virulence mechanisms of invasion, adhesion, infection and colonization and proteins present in several *S. pneumoniae* serotypes, mainly those widespread in Brazil, to substitute the expensive polysaccharide conjugate vaccine. These potential targets, including PsaA, can be used in a protein subunit vaccine and can also be employed as carriers for polysaccharide after conjugation, potentiating the vaccine immune response. We are currently studying the applicability of the proposed strategy to other pneumococcal proteins.

Acknowledgments

Bio-Manguinhos, PDTIS (Programa de Desenvolvimento Tecnológico de Insumos para Saúde) and PAPES V (Programa Estratégico de Apoio à Pesquisa em Saúde) from Fundação Oswaldo Cruz (Fiocruz) supported this work. We thank the Bio-Manguinhos technicians for their technical support to this work.

We acknowledge the researchers of the lab of Dr. Filipe Soares Quirino da Silva from Setor de Imunobiológicos/Departamento de Química/INCQS/Fiocruz for their help with the Bio-Rad Densitometer and the software QuantiOne 4.4.1. We also thank Dr. Luciana Leite from Instituto Butantan (SP) for providing the anti-rPsaA mouse serum and Dr. Celso Romero Ramos from Departamento de Helmintologia/IOC/Fiocruz for circular dichroism analyses.

References

[1] M.J. Jedrzejas, Pneumococcal virulence factors: structure and function, Microbiol. Mol. Biol. Rev. 65 (2001) 187–207.
[2] D. Bogaert, P.W.M. Hermans, P.V. Adrian, H.C. Rümke, R. De Groot, Pneumococcal vaccines: an update on current strategies, Vaccine 22 (2004) 2209–2220.
[3] M.M. García-Suárez, F. Vázquez, F.J. Méndez, *Streptococcus pneumoniae* virulence factors, their clinical impact: an update, Enferm. Infec. Microbiol. Clin. 24 (2006) 512–517.
[4] M.C.C. Brandileone, A.L.S.S. Andrade, J.L. DiFabio, M.L.L.S. Guerra, R. Austrian, Appropriateness of a pneumococcal conjugate vaccine in Brazil: potential impact of age and clinical diagnosis, with emphasis on meningitis, J. Infect. Dis. 187 (2003) 1206–1212.
[5] J.C. Paton, Novel pneumococcal surface proteins: role in virulence and vaccine potential, Trends Microbiol. 6 (1998) 85–87.
[6] S.S. Tai, *Streptococcus pneumoniae* protein vaccine candidates: properties, activities and animal studies, Crit. Rev. Microbiol. 32 (2006) 139–153.
[7] M.A. Barocchi, S. Censini, R. Rappuoli, Vaccines in the era of genomics: the pneumococcal challenge, Vaccine 25 (2007) 2963–2973.
[8] E. AlonsoDeVelasco, A.F.M. Verheul, J. Verhoef, H. Snippe, *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines, Microbiol. Rev. 59 (1995) 591–603.
[9] S.K. Hollingshead, D.E. Briles, *Streptococcus pneumoniae*: new tools for an old pathogen, Curr. Opin. Microbiol. 4 (2001) 71–77.
[10] S. Bergmann, S. Hammerschmidt, Versatility of pneumococcal surface proteins, Microbiology 152 (2006) 295–303.
[11] A.D. Ogunniyi, M. Grabowicz, D.E. Briles, J. Cook, J.C. Paton, Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of *Streptococcus pneumoniae*, Infect. Immun. 75 (2007) 350–357.
[12] J.C. Paton, A.M. Berry, R.A. Lock, D. Hansman, P.A. Manning, Cloning and expression in *Escherichia coli* of the *Streptococcus pneumoniae* gene encoding pneumolysin, Infect. Immun. 54 (1986) 50–55.
[13] R.A. Lock, J.C. Paton, D. Hansman, Purification and immunological characterization of neuraminidase produced by *Streptococcus pneumoniae*, Microb. Pathog. 4 (1988) 33–43.
[14] A.M. Berry, R.A. Lock, D. Hansman, J.C. Paton, Contribution of autolysin to virulence of *Streptococcus pneumoniae*, Infect. Immun. 57 (1989) 2324–2330.
[15] L.S. McDaniel, J.S. Sheffield, P. Delucchi, D.E. Briles, PspA, a surface protein of *Streptococcus pneumoniae*, is capable of eliciting protection against pneumococci of more than one capsular type, Infect. Immun. 59 (1991) 222–228.
[16] A. Brooks-Walter, D.E. Briles, S.K. Hollingshead, The pspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia, Infect. Immun. 67 (1999) 6533–6542.
[17] C. Rosenow, P. Ryan, J.N. Weiser, S. Johnson, P. Fontan, A. Ortqvist, H.R. Masure, Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*, Mol. Microbiol. 25 (1997) 819–829.
[18] H. Russell, J.A. Tharpe, D.E. Wells, E.H. White, J.E. Johnson, Monoclonal antibody recognizing a species-specific protein from *Streptococcus pneumoniae*, J. Clin. Microbiol. 28 (1990) 2191–2195.
[19] J.S. Sampson, S.P. O'Connor, A.R. Stinson, J.A. Tharpe, H. Russell, Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus* sp. adhesins, Infect. Immun. 62 (1994) 319–324.
[20] J.A. Tharpe, H. Russell, Purification and seroreactivity of pneumococcal surface adhesin A (PsaA), Clin. Diagn. Lab. Immunol. 3 (1996) 227–229.
[21] D.F. Talkington, B.G. Brown, J.A. Tharpe, A. Koenig, H. Russell, Protection of mice against fatal pneumococcal challenge by immunization with pneumococcal surface adhesion A (PsaA), Microb. Pathog. 21 (1996) 17–22.
[22] M.J. Jedrzejas, R.B. Mewbourne, L. Chantalat, D.T. McPherson, Expression and purification of *Streptococcus pneumoniae* hyaluronate lyase from *Escherichia coli*, Protein Expr. Purif. 13 (1998) 83–89.
[23] K. Overweg, A. Kerr, M. Sluijter, M.H. Jackson, T.J. Mitchell, A.P.J.M. de Jong, R. de Groot, P.W.M. Hermans, The putative proteinase maturation protein A of *Streptococcus pneumoniae* is a conserved surface protein with potential to elicit protective immune responses, Infect. Immun. 68 (2000) 4180–4188.
[24] J.S. Brown, A.D. Ogunniyi, M.C. Woodrow, D.W. Holden, J.C. Paton, Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection, Infect. Immun. 69 (2001) 6702–6706.
[25] J.E. Adamou, J.H. Heinrichs, A.L. Erwin, W. Walsh, T. Gayle, M. Dormitzer, R. Dagan, Y.A. Brewah, P. Barren, R. Lathigra, S. Langermann, S. Koenig, S. Johnson, Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis, Infect. Immun. 69 (2001) 949–958.
[26] Y. Zhang, A.W. Masi, V. Barniak, K. Mountzouros, M.K. Hostetter, B.A. Green, Recombinant PhpA protein, a unique histidine motif-containing protein from *Streptococcus pneumoniae*, protects mice against intranasal pneumococcal challenge, Infect. Immun. 69 (2001) 3827–3836.
[27] J. Hamel, N. Charland, I. Pineau, C. Ouellet, S. Rioux, D. Martin, B.R. Brodeur, Prevention of pneumococcal disease in mice immunized with conserved surface-accessible proteins, Infect. Immun. 72 (2004) 2659–2670.
[28] A.R. Holmes, R. McNab, K.W. Millsap, M. Rohde, S. Hammerschmidt, J.L. Mawdsley, H.F. Jenkinson, The pavA gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence, Mol. Microbiol. 41 (2001) 1395–1408.
[29] B.K. De, J.S. Sampson, E.W. Ades, R.C. Huebner, D.L. Jue, S.E. Johnson, M. Espina, A.R. Stinson, D.E. Briles, G.M. Carlone, Purification and characterization of *Streptococcus pneumoniae* palmitoylated pneumococcal surface adhesin A expressed in *Escherichia coli*, Vaccine 18 (2000) 1811–1821.
[30] P.V. Adrian, D. Bogaert, M. Oprins, S. Rapola, M. Lahdenkari, T. Kilpi, R. de Groot, H. Kayhty, P.W.M. Hermans, Development of antibodies against pneumococcal proteins α-enolase, immunoglobulin A1 protease, streptococcal lipoprotein rotamase A, and putative proteinase maturation protein A in relation to pneumococcal carriage and otitis media, Vaccine 22 (2004) 2737–2742.

[31] H.Y. Kwon, A.D. Ogunniyi, M.H. Choi, S.N. Pyo, D.K. Rhee, J.C. Paton, The ClpP protease of *Streptococcus pneumoniae* modulates virulence gene expression and protects against fatal pneumococcal challenge, Infect. Immun. 72 (2004) 5646–5653.

[32] B.A. Green, Y. Zhang, A.W. Masi, V. Barniak, M. Wetherell, R.P. Smith, M.S. Reddy, D. Zhu, PppA, a surface-exposed protein of *Streptococcus pneumoniae*, elicits cross-reactive antibodies that reduce colonization in a murine intranasal immunization and challenge model, Infect. Immun. 73 (2005) 981–989.

[33] V. Romanello, M. Marcacci, F. Dal Molin, M. Moschioni, S. Censini, A. Covacci, A.G. Baritussio, C. Montecucco, F. Tonello, Cloning, expression, purification, and characterization of *Streptococcus pneumoniae* IgA1 protease, Protein Expr. Purif. 45 (2006) 142–149.

[34] P. Shah, E. Swiatlo, Immunization with polyamine transport protein PotD protects mice against systemic infection with *Streptococcus pneumoniae*, Infect. Immun. 74 (2006) 5888–5892.

[35] C. Gianfaldoni, S. Censini, M. Hilleringmann, M. Moschioni, C. Facciotti, W. Pansegrau, V. Masignani, A. Covacci, R. Rappuoli, M.A. Barocchi, P. Ruggiero, *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge, Infect. Immun. 75 (2007) 1059–1062.

[36] T.M. Wizemann, J.H. Heinrichs, J.E. Adamou, A.L. Erwin, C. Kunsch, G.H. Choi, S.C. Barash, C.A. Rosen, H.R. Masure, E. Tuomanen, A. Gayle, Y.A. Brewah, W. Walsh, P. Barren, R. Lathigra, M. Hanson, S. Langermann, S. Johnson, S. Koenig, Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection, Infect. Immun. 69 (2001) 1593–1598.

[37] J.C. Paton, P. Giammarinaro, Genome-based analysis of pneumococcal virulence factors: the quest for novel vaccine antigens and drug targets, Trends Microbiol. 9 (2001) 515–518.

[38] S. Bambini, R. Rappuoli, The use of genomics in microbial vaccine development, Drug Discov. Today 14 (2009) 252–260.

[39] K. Shen, J. Gladitz, P. Antalis, B. Dice, B. Janto, R. Keefe, J. Hayes, A. Ahmed, R. Dopico, N. Ehrlich, J. Jocz, L. Kropp, S. Yu, L. Nistico, D.P. Greenberg, K. Barbadora, R.A. Preston, J.C. Post, G.D. Ehrlich, F.Z. Hu, Characterization, distribution, and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*, Infect. Immun. 74 (2006) 321–330.

[40] M.J. Rodríguez-Ortega, N. Norais, G. Bensi, S. Liberatori, S. Capo, M. Mora, M. Scarselli, F. Doro, G. Ferrari, I. Garaguso, T. Maggi, A. Neumann, A. Covre, J.L. Telford, G. Grandi, Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome, Nat. Biotechnol. 24 (2006) 191–197.

[41] C. Morsczeck, T. Prokhorova, J. Sigh, M. Pfeiffer, M. Bille-Nielsen, J. Petersen, A. Boysen, T. Kofoed, N. Frimodt-Møller, P. Nyborg-Nielsen, P. Schrotz-King, *Streptococcus pneumoniae*: proteomics of surface proteins for vaccine development, Clin. Microbiol. Infect. 14 (2008) 74–81.

[42] G.K. Paterson, L. Nieminen, J.M.C. Jefferies, T.J. Mitchell, PclA, a pneumococcal collagen-like protein with selected strain distribution, contributes to adherence and invasion of host cells, FEMS Microbiol. Lett. 285 (2008) 170–176.

[43] C. Giefing, A.L. Meinke, M. Hanner, T. Henics, D.B. Minh, D. Gelbmann, U. Lundberg, B.M. Senn, M. Schunn, A. Habel, B. Henriques-Normark, A. Örtqvist, M. Kalin, A. von Gabain, E. Nagy, Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies, J. Exp. Med. 205 (2008) 117–131.

[44] D.E. Briles, E. Ades, J.C. Paton, J.S. Sampson, G.M. Carlone, R.C. Huebner, A. Virolainen, E. Swiatlo, S. Hollingshead, Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*, Infect. Immun. 68 (2000) 796–800.

[45] D.E. Briles, S. Hollingshead, A. Brooks-Walter, G.S. Nabors, L. Ferguson, M. Schilling, S. Gravenstein, P. Braun, J. King, A. Swift, The potential to use PspA and other pneumococcal proteins to elicit protection against pneumococcal infection, Vaccine 18 (2000) 1707–1711.

[46] D.E. Briles, S.K. Hollingshead, G.S. Nabors, J.C. Paton, A. Brooks-Walter, The potential for using protein vaccines to protect against otitis media caused by *Streptococcus pneumoniae*, Vaccine 19 (2001) S87–S95.

[47] G. Douce, K. Ross, G. Cowan, J. Ma, T.J. Mitchell, Novel mucosal vaccines generated by genetic conjugation of heterologous proteins to pneumolysin (PLY) from *Streptococcus pneumoniae*, Vaccine 28 (2010) 3231–3237.

[48] K.E. Morrison, D. Lake, J. Crook, G.M. Carlone, E. Ades, R. Facklam, J.S. Sampson, Confirmation of *psaA* in all 90 serotypes of *Streptococcus pneumoniae* by PCR and potential of this assay for identification and diagnosis, J. Clin. Microbiol. 38 (2000) 434–437.

[49] I. Jado, A. Fenoll, J. Casal, A. Pérez, Identification of the *psaA* gene, coding for pneumococcal surface adhesin A, in viridans group streptococci other than *Streptococcus pneumoniae*, Clin. Diagn. Lab. Immunol. 8 (2001) 895–898.

[50] A.P.M. Arêas, M.L.S. Oliveira, E.N. Miyaji, L.C.C. Leite, K.A. Aires, W.O. Dias, P.L. Ho, Expression and characterization of cholera toxin B – pneumococcal surface adhesin A fusion protein in *Escherichia coli*: ability of CTB-PsaA to induce humoral immune response in mice, Biochem. Biophys. Res. Commun. 321 (2004) 192–196.

[51] G. Rajam, J.M. Anderton, G.M. Carlone, J.S. Sampson, E.W. Ades, Pneumococcal surface adhesin A (PsaA): a review, Crit. Rev. Microbiol. 34 (2008) 131–142.

[52] A.M. Berry, J.C. Paton, Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*, Infect. Immun. 64 (1996) 5255–5262.

[53] D.O. Gor, X. Ding, Q. Li, J.R. Schreiber, M. Dubinsky, N.S. Greenspan, Enhanced immunogenicity of pneumococcal surface adhesin A by genetic fusion to cytokines and evaluation of protective immunity in mice, Infect. Immun. 70 (2002) 5589–5595.

[54] A.D. Ogunniyi, R.L. Folland, D.E. Briles, S.K. Hollingshead, J.C. Paton, Immunization of mice with combinations of pneumococcal virulence proteins elicits enhanced protection against challenge with *Streptococcus pneumoniae*, Infect. Immun. 68 (2000) 3028–3033.

[55] J.Y. Seo, S.Y. Seong, B.Y. Ahn, I.C. Kwon, H. Chung, S.Y. Jeong, Cross-protective immunity of mice induced by oral immunization with pneumococcal surface adhesin A encapsulated in microspheres, Infect. Immun. 70 (2002) 1143–1149.

[56] L. Baril, D.E. Briles, P. Crozier, J. King, M. Punar, S.K. Hollingshead, J.B. McCormick, Characterization of antibodies to PspA and PsaA in adults over 50 years of age with invasive pneumococcal disease, Vaccine 23 (2004) 789–793.

[57] F.C. Pimenta, E.N. Miyaji, A.P.M. Arêas, M.L.S. Oliveira, A.L.S.S. de Andrade, P.L. Ho, S.K. Hollingshead, L.C.C. Leite, Intranasal immunization with the cholera toxin B subunit-pneumococcal surface antigen A fusion protein induces protection against colonization with *Streptococcus pneumoniae* and has negligible impact on the nasopharyngeal and oral microbiota of mice, Infect. Immun. 74 (2006) 4939–4944.

[58] R.R. Burgess, Preparing a purification summary table, Methods Enzymol. 463 (2009) 29–34.

[59] C.B. Lavall, A.L.S.S. de Andrade, F.C. Pimenta, J.G. de Andrade, R.M. de Oliveira, S.A. Silva, E.C. de Lima, J.L. Di Fabio, S.T. Casagrande, M.C.C. Brandileone, Serotypes of carriage and invasive isolates of *Streptococcus pneumoniae* in Brazilian children in the era of pneumococcal vaccines, Clin. Microbiol. Infect. 12 (2006) 50–55.

[60] W.P. Hausdorff, D.R. Feikin, K.P. Klugman, Epidemiological differences among pneumococcal serotypes, Lancet Infect. Dis. 5 (2005) 83–93.

[61] P.A. Cullen, M. Lo, D.M. Bulach, S.J. Cordwell, B. Adler, Construction and evaluation of a plasmid vector for the expression of recombinant lipoproteins in *Escherichia coli*, Plasmid 49 (2003) 18–29.

[62] J.H. Choi, K.C. Keum, S.Y. Lee, Production of recombinant proteins by high cell density culture of *Escherichia coli*, Chem. Eng. Sci. 61 (2006) 876–885.