Transcriptional regulation, occurrence and putative role of the Pht family of *Streptococcus pneumoniae*

Stéphane Rioux,1 Cécile Neyt,1 Emmanuel Di Paolo,1 Laurence Turpin,1 Nathalie Charland,2 Steve Labbé,3 Marie-Cécile Mortier,1 Tim J. Mitchell,4 Christiane Feron,1 Denis Martin3 and Jan T. Poolman1

Correspondence
Jan T. Poolman
jan.poolman@gskbio.com

1GSK Biologicals, Rixensart, Belgium
2Medicago Inc., Québec, Québec, Canada
3GSK Biologicals, Laval, Québec, Canada
4University of Glasgow, Glasgow, UK

Restricted to the genus *Streptococcus*, the Pht protein family comprises four members: PhtA, PhtB, PhtD and PhtE. This family has the potential to provide a protein candidate for incorporation in pneumococcal vaccines. Based on sequence analysis and on RT-PCR experiments, we show here that the *pht* genes are organized in tandem but that their expression, except that of *phtD*, is monocistronic. PhtD, PhtE, PhtB and PhtA are present in 100, 97, 81 and 62 % of the strains, respectively, and, by analysing its sequence conservation across 107 pneumococcal strains, we showed that PhtD displays very little variability. To analyse the physiological function of these proteins, several mutants were constructed. The quadruple Pht-deficient mutant was not able to grow in a poor culture medium, but the addition of Zn²⁺ or Mn²⁺ restored its growth capacity. Moreover, the *phtD* mRNA expression level increased when the culture medium was depleted in zinc. Therefore, we suggest that these proteins are zinc and manganese scavengers, and are able to store these metals and to release them when the bacterium faces an ion-restricted environment. The data also showed that this protein family, and more particularly PhtD, is a promising candidate to be incorporated into pneumococcal vaccines.

INTRODUCTION

*Streptococcus pneumoniae* is one of the leading causes of infectious morbidity and mortality, responsible for a large spectrum of infections such as otitis media, pneumonia, bacteraemia and meningitis (Hausdorff et al., 2005; McCullers & Tuomanen, 2001). The emergence of antibiotic-resistant strains of this micro-organism has further underlined the need for providing effective prophylactic vaccination (Bridy-Pappas et al., 2005; Lynch & Zhanel, 2005).

Current vaccines comprise epidemiologically dominant serotype-based selections of pneumococcal capsular polysaccharides, conjugated or not to a carrier protein (Dagan et al., 2004; Fedson & Musher, 2004; Mbelle et al., 1999; Smart et al., 1987). However, the vaccine formulations do not cover all serotypes of *S. pneumoniae*, which might be of particular relevance in certain regions that have different dominant serotypes (Dagan et al., 1992). In addition, one may expect that the use of serotype-specific vaccines could allow the positive selection of non-vaccine serotypes (Nunes et al., 2008; Singleton et al., 2007).

An alternative approach involves the development of vaccines that target common pneumococcal antigens. Among multiple candidates, the Pht protein family, restricted to the genus *Streptococcus*, is promising, as it is well conserved across pneumococcal species (Hamel et al., 2004; Zhang et al., 2001), and its members are immune targets in infected individuals and protective antigens in immunized mice upon challenge (Beghetto et al., 2006). Originally, this protein family was independently reported by three groups, and three separate denominations were used: Pht (for pneumococcal histidine triad) (Adamou et al., 2001), Php (for pneumococcal histidine protein) (Zhang et al., 2001) and BVH (Hamel et al., 2004). Those proteins are characterized by a histidine triad motif, HxHxH, repeated five to six times in their amino acid sequences. Four members of this family have been described: PhtA (BH-11-3), PhtB (PhtA/BVH-11) and PhtD (BVH-11-2), which share up to 81 % sequence identity, and PhtE (BVH-3), which is divergent from the three other proteins, showing...
only up to 35% sequence identity with them. PhTe is a longer protein, the only one with six repeats of the histidine triad motif. In mouse immunization studies, all members of the Pht family have been shown to afford a high level of protection to subsequent pneumococcal challenge with a number of different strains/serotypes (Adamou et al., 2001; Godfroid et al., 2011; Hamel et al., 2004; Ogunniyi et al., 2007; Wizemann et al., 2001; Zhang et al., 2001).

Despite their potential importance in vaccination against S. pneumoniae infection, the biological function of these proteins has yet to be determined. Results from antibody-labelling and flow cytometry experiments demonstrated that the Pht proteins are exposed on the surface of the encapsulated bacterium (Hamel et al., 2004), which is in agreement with their relevance as vaccine targets. Based on signature-tagged mutagenesis, it has been suggested that PhtA, PhtB and PhtD are involved in lung-specific variability. Therefore, we have addressed these different aspects with regard to the Pht proteins.

**METHODS**

**Bacterial strains and culture conditions.** S. pneumoniae strain TIGR4 (serotype 4) (Tettelin et al., 2001) was kindly provided by Andrew Camilli (Tufts University School of Medicine, Boston, MA, USA). The WU2 strain (serotype 3) was kindly provided by David E. Briles (University of Alabama at Birmingham, Birmingham, AL, USA). The D39 strain (serotype 2) was kindly provided by J. C. Paton (University of Adelaide, Australia). The type 4 strain was obtained from the CDC (Centers for Disease Control & Prevention, Atlanta, GA, USA).

Pneumococci were routinely grown in Todd–Hewitt broth (THB; Difco) with either 0.5% (w/v) yeast extract or 17% fetal bovine serum (JRH Biosciences) at 37 °C or 8% CO2. When appropriate, erythromycin and/or spectinomycin (Sigma-Aldrich) were added at 0.2 and 1000 μg ml⁻¹, respectively.

Escherichia coli DH5α and JM109 strains (Gibco-BRL, Life Technology) were grown in Luria–Bertani (LB) broth (Difco) with or without 1.5% (w/v) Bacto-agar (Difco) at 37 °C for 16 h. When appropriate, erythromycin or spectinomycin was added to the growth medium at 500 and 100 μg ml⁻¹, respectively.

To study Pht occurrence, in addition to 23 in-house and pneumococcal molecular epidemiology network (PMEN) strains, 34 isolates were provided by T. J. Mitchell (Scotland), six by R. E. Gertz (CDC, Atlanta, USA) (model 3100 or 3130XL). Sequence analyses were performed with Vector NTI 7.1 software (Oxford Molecular), Vector NTI 7.1 software (Informax) or SeqMan software (Lasergene, DNASTAR), and sequences were compared with the available S. pneumoniae TIGR4 genome sequence (http://www.tigr.org) (Peterson et al., 2001).

**DNA treatment and analysis.** E. coli plasmid DNA was obtained by using a plasmid Midi or Mini purification kit (Qiagen). PCR products were purified with the QIAquick PCR purification kit, and DNA digests were purified on 1% (w/v) agarose gel by using the QIAquick gel extraction kit (Qiagen). Restriction and ligation enzymes were obtained from New England BioLabs. The Expand high-fidelity system (Roche) was used for each PCR of these studies. All commercial products were used under conditions recommended by the suppliers.

DNA sequencing was carried out with the Big Dye terminator sequencing kit on an Applied Biosystems automated DNA sequencer (model 3100 or 3130XL). Sequence analyses were performed with MacVector V6.5 software (Oxford Molecular), Vector NTI 7.1 software (Informax) or SeqMan software (Lasergene, DNASTAR), and sequences were compared with the available S. pneumoniae TIGR4 genome sequence (http://www.tigr.org) (Peterson et al., 2001).

**S. pneumoniae genomic DNA extraction.** Chromosomal DNA from each strain was obtained by harvesting confluent overnight growth from one or two heavily inoculated blood agar plates into 1 ml TE (10 mM Tris/HCl, 5 mM EDTA, pH 7.8). The bacterial suspension was centrifuged for 5 min at maximum speed in a microcentrifuge. The pellet was either treated with the QIAamp DNA mini kit protocol (Qiagen) or resuspended in 75 μl TE, lysed by sequential addition of 20 μl lysozyme (100 mg ml⁻¹) and 20 μl proteinase K (20 mg ml⁻¹) and incubated at 37 °C for 45 min. Then, 500 μl of lysis buffer [10 mM Tris/HCl, pH 8.0, 0.14 M NaCl, 0.1 M sodium citrate, 1 mM EDTA, pH 8.0, 0.1% (w/v) sodium deoxycholate] was added and incubated for 10 min at room temperature. At the end of this incubation period, 250 μl ammonium acetate (7.5 mM, pH 7.7) was added to the crude lysate and incubated for 10 min on ice. The viscous DNA was extracted twice with phenol/ chloroform/isooamyl (25:24:1) and precipitated in isopropl alcohol. The resulting DNA was washed with 70% (v/v) ethanol and resuspended in 50 μl TE containing 0.6 μl RNaseA (10 mg ml⁻¹). DNA suspensions were stored at 4 °C.

**RNA isolation.** Total RNA was isolated from pneumococci grown from an OD₆₀₀ of 0.01 in THB to different OD₆₀₀ values to evaluate
gene expression at different growth phases (early exponential, OD\text{600} 0.3; late exponential, OD\text{600} 0.9; stationary, OD\text{600} 1.2). Cells were centrifuged and resuspended in RNA-free Tris-EDTA containing 6 mg lysosyme ml\textsuperscript{-1} and 1 mg sodium deoxycholate ml\textsuperscript{-1}, and incubated at room temperature for 10 min. After incubation, RNA isolation was performed with the Qiagen RNeasy Mini kit following the manufacturer’s instructions. Contaminating genomic DNA was eliminated by incubating RNA samples with 1 U DNase I (μg RNA)\textsuperscript{-1} for 1 h at 37 °C, followed by DNase inactivation with 2.5 mM EDTA for 10 min at 65 °C. Total RNA was quantified by using the Ribogreen RNA quantification kit (Molecular Probes) following the manufacturer’s instructions.

5′-Rapid amplification of cDNA ends (RACE). The method used to identify transcription starts was adapted from that described by Ranasinghe & Hobb’s (1998). Briefly, a primer specific for the 3′ end of the \textit{phtE} gene was used to synthesize the first-strand cDNA from total RNA with the Superscript II reverse transcriptase (Invitrogen), following the manufacturer’s instructions. RNAse A was then added for 1 h at room temperature to generate blunt 3′ ends on the cDNA–RNA hybrid. The hybrid was inserted into EcoRV-digested pKS plasmid (Stratagene) by using T4 DNA ligase (overnight incubation at 16 °C). A PCR was set up to amplify the 5′ end by using another reverse 3′ end-specific \textit{phtE} primer and pKS-specific T7 promoter primer. Sequencing of the pKS–cDNA junction was performed to identify the +1 base.

Transcriptional terminator identification. Terminator identification was performed by using the Wisconsin Sequence Analysis Package version 10.1 (Genetics Computer Group) based on the method described by Brendel & Tritonov (1984).

RT-PCR. RT-PCR studies were performed as follows. RNA (2 μg) was first denatured for 5 min at 65 °C in a mixture containing 10 μM 3′-end gene-specific reverse primer and 20 U RNaseOut in a total volume of 10 μl. The reverse transcription reaction was then carried out by adding 5 mM dithiothreitol, 1 mM dNTP, 15 U ThermoScript reverse transcriptase (Invitrogen), 1× cDNA synthesis buffer and RNase-free sterile water to a volume of 20 μl. The reverse transcription mixture was incubated at 56–58 °C for 1 h, followed by reverse transcriptase denaturation for 5 min at 85 °C. The RNA strand on the RNA–cDNA hybrids was degraded by incubating the reverse transcription solution at 37 °C for 20 min with 1 U RNase H. PCR was carried with 2 μl cDNA by using different 5′ gene-specific forward primers and the 3′ gene-specific reverse primers used for the reverse transcription reaction (0.5 μM final concentrations), 0.2 mM dNTP, Taq DNA polymerase reaction buffer, 2.5 U Taq DNA polymerase (Amersham Biosciences) and sterile water to a volume of 30 μl. The PCR cycle consisted of initial denaturation at 94 °C for 5 min, followed by 25–30 cycles of denaturation at 94 °C for 15–30 s, annealing at 55 °C (phtE, phtD) or 63 °C (phtB, phtD, phtA) for 15–30 s and extension at 72 °C for 1 min, and completed by a final extension step at 72 °C for 5–7 min. A negative control composed of RNA without the reverse transcription reaction was also conducted to exclude DNA contamination in the RNA preparation. PCR products were separated by 1% (w/v) agarose gel electrophoresis and were visualized by ethidium bromide staining.

Preparation of Pht mutants. Cloning of mutator vectors was performed in \textit{E. coli} DH5α or JM109 plated on LB agar with the respective antibiotics. Transformation of \textit{E. coli} with plasmid DNA was carried out by using standard methods with CaCl\textsubscript{2}-treated cells (Hanahan, 1985).

Mutator vectors were constructed from the pGEM-T vector (Promega) that replicates in \textit{E. coli} but not in \textit{S. pneumoniae}. They contain recombinant zones that correspond to the upstream and downstream regions of the \textit{pht} genes to be deleted, amplified by PCR, surrounding an antibiotic-resistance gene. To prepare the quadruple Pht-deficient mutant, two different antibiotic-resistance genes had to be used to combine deletion in the two different loci (locus \textit{phtD/phtE}, and locus \textit{phtA/phtB}). An erythromycin resistance gene (ermB), amplified from a derivative of the pIDC9 vector, was selected for the \textit{phtD/phtE} locus. For the \textit{phtA/phtB} locus, a spectinomycin-resistance gene [aad(9) gene], purified from the pR350 plasmid (kindly provided by J. Paton), was used.

Construction of the mutator vectors and the pneumococcal transformation protocol were carried out as fully described by Melin et al. (2010). All deletions were realized from start to stop codon, leaving the promoters and pre- and post-gene sequences untouched.

SDS-PAGE and Western blot analysis. Heat-killd bacterial suspensions were obtained by harvesting the confluent overnight growth from five heavily inoculated blood agar plates into 1 ml sterile PBS (0.14 M NaCl, 2.7 mM KCl, 10 mM Na\textsubscript{2}HPO\textsubscript{4}, 1.8 mM KH\textsubscript{2}PO\textsubscript{4}, pH 7.2), and incubating at 56 °C for 45 min. For harmonization, all bacterial samples were resuspended at 0.02 OD units μl\textsuperscript{-1} (OD\text{600}) and the same volume for each sample was loaded onto the gel. Sample buffer [60 mM Trizma base, 1% (w/v) SDS, 10% (v/v) glycerol, 0.01% (w/v) Bromophenol blue, 2% (w/v) β-mercaptoethanol] was added to the heat-killd suspensions. Preparations were boiled for 5 min, centrifuged at maximum speed in a microcentrifuge for 2 min and separated by SDS-PAGE as described by Laemmli (1970). Proteins were electrophoretically transferred from acrylamide gels onto nitrocellulose membranes (Bio-Rad), according to Towbin et al. (1979). Membranes were probed with a mouse polyclonal antibody raised against PhtD, followed by goat anti-mouse IgG conjugated to alkaline phosphatase (Promega). Enzyme-labelled bands were visualized with an NBT/BCIP substrate system.

Culture growth in ion-deficient medium. The wild-type 4/CDC strain and corresponding Pht mutants were cultured under different conditions of ion depletion or supplementation in a chemically defined synthetic medium (MS) (Sicard, 1964). MS was supplemented with increasing concentrations of Mn\textsuperscript{2+}, Fe\textsuperscript{3+}, Fe\textsuperscript{2+}, Cu\textsuperscript{2+} or Zn\textsuperscript{2+}. OD\text{600} was monitored during exponential phase and at stationary phase. Results were compared with those of the wild-type.

Wild-type WU2 strain was cultured with or without the Zn-specific chelator N,N,N′,N′-tetraakis(2-pyridylmethyl) ethylenediamine (TPEN) to observe the effect of zinc depletion on Pht expression at the RNA (by RT-PCR) and protein (by flow cytometry) levels.

Flow cytometry. WU2 bacteria were grown in THB+0.5% yeast extract at 37 °C, 8% CO\textsubscript{2}, up to exponential phase. Alternatively, 30 μM TPEN was added to the medium. After centrifugation, bacterial pellets were resuspended in a solution containing anti-PhtE, anti-PhtB/D, anti PhtD/E or anti-type 3 polysaccharide monoclonal antibodies as control. After 2 h at 4 °C, the solutions were centrifuged, the bacterial pellets were washed in 2% PBS–BASA and they were then incubated for 1 h at room temperature in AlexaFluor-conjugated (Molecular Probes) goat anti-mouse secondary antibody in 2% PBS–BASA. After washing, cells were fixed in 0.25% PBS–formaldehyde, and fluorescence-activated cell sorting analysis was performed. The median surface fluorescence was recorded.

Quantitative RT-PCR. Total RNA from strain D39 grown to an OD\text{600} ~0.5 (mid exponential) was purified by using the RNeasy Midi kit (Qiagen) and quantified with the Quant-it RiboGreen RNA assay kit (Invitrogen). Samples (1 μg) were processed twice with 1.5 μl RQ1 RNase-free DNase (Promega) for 30 min at 37 °C. The reaction was stopped by the addition of 1 μl DNase STOP followed by incubation.
for 10 min at 65 °C. First-strand cDNA was generated using SuperScript II Reverse Transcriptase (Invitrogen), random primers (Invitrogen) and recombinant RNasin RNase inhibitor (Promega). Real-time PCR was carried out in a 50 µl reaction volume by using the TaqMan PCR core reagents kit (Applied Biosystems), as described by the manufacturer. The following primers and probes were used: gyrB, GGAATTGCGGAAGTGTCAAG (forward), GGAATCGG-AGAAGGCTTCAC (reverse) and TTACATCAGCTTCTTC (probe); phtD, CCACGTGAGCAATATCG (forward), TGACTGCTTGC-TGCTTCTG (reverse) and CGTAAAATCTCTTCTTTGT (probe).

All assays were performed in duplicate (from culture to quantitative PCR) and the relative gene transcription was analysed by the 2−ΔΔCT method (Livak & Schmittgen, 2001) by using gyrB as internal control and growth in THB alone as calibrator.

**Determination of Pht occurrence.** To select representative strains of S. pneumoniae, the population structure was analysed according to the strain genotype as determined by MLST (multi-locus sequence type; http://www.mlst.net). Based on MLST isolate sequence type (ST), major clonal lineages were determined. For each group, representative strains were selected for occurrence analysis, which was carried out by Western blotting on whole bacterial extracts with anti-PhtD polyclonal antibodies (cross-reactive with PhtA, PhtB and PhtE) and by PCR on pneumococcal genomic DNA with primers specific for PhtA, PhtB, PhtD or PhtE.

**DNA sequencing for PhtD conservation analysis.** DNA of 107 MLST-selected strains was amplified by PCR by using PhtD-specific oligonucleotide primers. The 107 sequences were aligned via the CLUSTAL_X program and the percentage identity between the sequences was computed by using a program written by the authors. Percentage identity was defined as (number of identities/length of shortest sequence) × 100.

**RESULTS**

**Characterization of the pht genes**

**Genomic organization of the pht genes.** In a previous study, DNA sequencing of overlapping clones from an S. pneumoniae strain SP64 genomic library (Hamel et al., 2004) and PCR analyses allowed deduction of the genomic organization of the pht genes and their neighbouring genes in this type 6B strain. The phtA and phtB genes, as well as the phtD and phtE genes, were organized as a pair. BLAST analyses (http://www.tigr.org) (Peterson et al., 2001) indicated that the two gene tandems were located about 161 kbp apart in the S. pneumoniae TIGR4 genome and that the genomic organization was identical to that observed in the SP64 strain (Fig. 1). The same pht gene organization was also found in the 4/CDC strain and in the WU2 strain, with the exception that phtA is missing in the latter (data not shown). Sequencing of the regions surrounding the pht genes on the TIGR4 strain DNA confirmed the latter observation (data not shown). Additional analysis demonstrated that phtA and phtB were separated by 157 bp, whereas phtD and phtE were separated by 209 bp in the TIGR4 strain, which was chosen for further study.

On the phtD–phtE tandem side, a gene showing 72 % sequence similarity to the group A and B streptococci lmb genes, coding for laminin-binding proteins [GenBank accession nos. AAK34689 (Ferretti et al., 2001) and AAD13796 (Spellerberg et al., 1999), respectively], was located 7 bp upstream of the phtD gene (Fig. 1). This gene product was also recently denoted as AdcAII, and described as an ABC transporter-like zinc-binding protein (Loisel et al., 2008). A 1392 bp ORF, located 142 bp upstream of the lmb gene homologue, codes for a protein showing 64 % sequence similarity to the Bacillus subtilis metabolite transporter YfnA protein (accession no. D69814) and 81 % similarity to a putative amino acid permease of Streptococcus pyogenes (accession no. AAK33157) (Ferretti et al., 2001; Kunst et al., 1997). A sequence showing 79 % identity to the first 481 bp of phtE (proposed phtF) was found 226 bp after the phtE stop codon (Fig. 1). This sequence also showed 72 % identity to the phtA, phtB and phtD genes.

On the phtA–phtB tandem side, a 1332 bp ORF, showing 73 % sequence similarity to the Streptococcus salivarius ptsI gene (accession no. P30299) (Gagnon et al., 1992), was located 253 bp upstream of the phtA gene (Fig. 1). No functional ORF was located immediately downstream of the two gene pairs.

**Transcriptional organization of pht genes.** The genomic organization of pht genes suggested that the tandem genes might be coordinately transcribed. Further studies were thus performed to examine this hypothesis. First, putative promoters and ribosome-binding sites of pht genes were identified. 5′-RACE on the phtE gene allowed the identification of its transcription start, from which the promoter region was deduced. The transcription start site (+1) was found to be located 96 bases upstream of the PhTE translation start site, downstream of typical S. pneumoniae −10 and −35 RNA polymerase binding sites (Morrison & Jaurin, 1990) and upstream of a ribosome-binding site (Fig. 2a). Similar sequence organization was found upstream of the phtA, phtB and yfnA genes, indicating the presence of putative promoters (Fig. 2b, c, e). However, due to the close proximity of the lmb gene...
(7 bp), no promoter sequence was identified for the phtD gene. On the other hand, a sequence identical to the –35 sequences of the other pht genes was located upstream of the lmb gene (Fig. 2d). Ribosome-binding sites were observed 5–7 bp upstream of all start codons. Transcription termination sites of pht and adjacent genes were also identified. Computer analysis of predicted mRNA secondary structures suggested the presence of stem–loop terminator-like structures at the 3' ends of genes. Hairpin structures could form with calculated free energies of dissociation (ΔG) of −9.4, −27.0, −16.8 and −21.6 kcal mol⁻¹ for phtB, phtD, phtA and ptsI, respectively, as determined by the method described by Turner et al. (1988) (Fig. 3). In fact, the terminator identified for the phtD gene was identical to that reported by the TIGR website for ORF SP1003, which corresponds to the phtD gene homologue (http://www.tigr.org). No transcription terminators were identified by the TIGR group for the other pht or surrounding genes, probably reflecting differences in the algorithms used by the two studies. Most hairpins ended with a stretch of T residues as typically found in prokaryotic transcription terminators (Rosenberg & Court, 1979) and were located within 70 bp downstream of stop codons (Fig. 3). Interestingly, the phtD terminator sequence (ΔG=−4.7 kcal mol⁻¹) was located 1867 bp downstream of its stop codon and of the phtF gene, the latter ORF containing in-frame stop codons preventing its translation (Fig. 3a). No terminator sequences were identified downstream of yfnA and lmb genes.

The genomic organization suggested that phtE could be part of an operon composed of the yfnA, lmb, phtD, phtE and ptsI genes. Nevertheless, the 5'-RACE (Fig. 2a) and terminator identification (Fig. 3a) indicated that phtE was the first gene transcribed on a bicistronic message, composed of phtE and phtF genes, which was confirmed by RT-PCR. The regions phtE to phtF were amplified by RT-PCR (Fig. 4a, lanes 4 and 6), whereas no amplification product was obtained with the primer pair specific to the region between genes phtD and phtE (Fig. 4a, lane 5), indicating transcriptional termination downstream of phtD (Fig. 3c). As shown in Fig. 4(a) (lanes 1–3), the regions yfnA to phtD were amplified by RT-PCR. Moreover, Loisel et al. (2008) have demonstrated that this phtD transcript also encodes, in addition to yfnA, lmb and phtD, the two genes upstream of yfnA (ccdA, which is involved in the biogenesis of cytochrome c, and spr0904, which shows similarity to...
thioredoxine). Interestingly, the identification of a putative promoter upstream of the lmb gene (Fig. 2d) suggested transcriptional coupling of the phtD and lmb genes.

Results obtained for the phtB and phtA genes showed that they were transcribed as monocistronic mRNAs, as was suggested by identification of promoter (Fig. 2b, c) and

---

Fig. 3. Rho-independent transcription terminator sequences of the pht and ptsI genes in the TIGR4 strain. The terminator regions of phtE (a), phtB (b), phtD (c), phtA (d) and ptsI (e) genes are shown. Stop codons are underlined and in bold type, terminator regions are underlined, and sequences underlined with a dotted line indicate the hairpin region of the terminators. ORFs are represented by arrows indicating the direction of transcription over letters in bold type. In (a), the region in italics (phtF gene; putative start codon doubly underlined) shows 78% identity with the first 481 bp of the phtE gene. However, underlined stop codons prevent significant gene translation. Numbers correspond to sequence positions in the sequences with GenBank accession nos AY569979 (a, c) and AY569980 (b, d, e).

Fig. 4. RT-PCR analyses of the pht transcripts in TIGR4. (a) The RT-PCR products with template RNA from cells grown to mid-exponential growth phase were separated in a 1% agarose gel. Lanes 1–8 correspond to regions 1–8 in the schematic representation (b). The RT-PCR products shown in lanes 1–8 were generated by using primer pairs that flanked the corresponding regions depicted in the scheme. The length of each predicted RT-PCR product is indicated under the corresponding interval. The length of each gene is indicated in parentheses.
terminator (Fig. 3b, d) sites. Analysis of the transcriptional organization of phtA and phtB by RT-PCR revealed a phtB-specific amplicon with phtB-specific primers (Fig. 4a, lane 7). No amplification product was obtained by RT-PCR with primers amplifying the region between phtA and phtB (Fig. 4a, lane 8), indicating a monocistronic organization of the phtA and phtB genes. Terminator site identification (Fig. 3e) indicated that ptsI is transcribed as a monocistronic message, which also confirmed that phtA is not part of a polycistronic transcript.

**Construction and use of Pht mutants**

**Characterization of the mutants.** PhtA, PhtB, PhtD and PhtE mutants and quadruple PhtABDE mutant were constructed. To assess the accuracy of the recombination, the genomic DNA of the mutant strains was purified and the recombinant regions were sequenced (data not shown). Furthermore, the mutants were characterized phenotypically by immunoblotting, by using a mouse polyclonal anti-PhtD antibody (Fig. 5). All four Pht isotypes were recognized by this antibody. However, PhtE bands were fainter, confirming the greatest divergence of this Pht from the three other isotypes.

**Influence of various ions on bacterial growth.** Growth of the Pht quadruple mutant was dramatically decreased in MS, compared with that of the wild-type strain and of the different Pht mutants (Fig. 6a). The PhtA, PhtB and PhtE single mutants behaved the same way as the wild-type and the PhtD single mutant (data not shown). When the medium was supplemented with up to 200 μM Fe²⁺, Zn²⁺ or Mn²⁺, growth of the wild-type and of the PhtD-deficient mutant was slightly ameliorated (growth rate versus MS alone: 96–130 %). In contrast, the behaviour of the quadruple mutant was striking. Whereas the addition of 200 μM Fe²⁺ to MS induced only a 25.3 % increase of growth (Fig. 6d), the same concentration of Zn²⁺ or Mn²⁺ restored the growth capacity of the quadruple mutant (Fig. 6b, c). This represents an up to 92.3 % increase in growth rate compared with that obtained in MS alone. However, this recovery of growth rate was delayed, visible only after overnight incubation, as no improvement was visible within the first few hours of culture. The addition of Mg²⁺ at 200 μM did not restore growth completely, as did Zn²⁺ or Mn²⁺, but similar increases in growth rates were obtained when Mg²⁺ was added at 1 mg ml⁻¹ to MS (data not shown). The addition of high concentrations of Cu²⁺ (200 μM) or Zn²⁺ (1 mM) was toxic for wild-type and mutant strains (data not shown).

**Effect of zinc depletion on pht expression.** When the zinc chelator TPEN was added to the culture medium, the expression level of the Pht proteins in the WU2 genetic background was increased, as determined by flow cytometry experiments (Fig. 7a, b, c). As a control, no shift in mean fluorescence was observed with anti-type 3 polysaccharide antibody under the same conditions of zinc depletion (Fig. 7d). At the RNA level, we measured (RT-PCR) an up to 25-fold increase in the phtE transcription level under conditions of zinc depletion (data not shown).

In addition, quantitative RT-PCR experiments were performed with mRNA purified from a pneumococcal isolate (D39 strain) grown in either THB or THB + 25 μM TPEN. The concentration of chelator added to the media was suboptimal, as determined with preliminary experiments, meaning that TPEN did not prevent growth (as observed when TPEN concentration was sufficient to chelate all ions in the media) but delayed it (data not shown). The addition of TPEN to the medium resulted in a 4.84-fold increase in phtD mRNA expression level. Complementation of THB + TPEN with 25 μM ZnSO₄ restored the phtD mRNA expression to a level similar to that observed in medium alone, suggesting that among all ions that can be chelated by TPEN (Zn, Mn, Cd, Co, Ni, Cu, Mg, Ca) Zn has the major impact on phtD expression level.

**Occurrence of Pht in pneumococci**

In total, 74 strains (including 23 PMEN and in-house strains) were investigated. In this set of representative strains, 18 clonal lineages were characterized by 56 different STs, among which the more represented were 81, 90, 124, 156, 162 and 199 (22 strains), and 27 different serotypes, among which the more represented were 19F, 6B, 3 and 23F (47 % of all strains). Forty-six strains (61 %) with 27 different STs belonged to three major clonal groups (founder sequence: ST156, ST81 and ST199).

By PCR on genomic DNA, we found the genes for PhtD, PhtE, PhtB and PhtA in 100, 97, 81 and 62 % of the strains, respectively. Fifty-four per cent of the strains were found to carry the four pht genes in their genome. On immunoblots with polyclonal antibodies raised against PhtD, we detected

---

**Fig. 5.** SDS-PAGE immunoblotting of bacterial extracts. Anti-PhtD antibody was used to probe extracts from the PhtABDE quadruple mutant (a), the PhtE mutant (b), PhtD mutant (c), PhtB mutant (d), PhtA mutant (e) and the wild-type 4/CDC (f) strains. The position of the different Pht bands is indicated on the right, and a molecular mass marker on the left.
**Fig. 6.** *S. pneumoniae* requires Pht proteins to grow in a poor medium. (a) Growth curves of 4/CDC wild-type strain and different Pht-deficient mutants in MS. ○, Wild-type; ▽, PhtD mutant; □, PhtBD mutant; ●, PhtDE mutant; △, PhtBDE mutant; ■, PhtABDE. (b–d) Growth curves of the wild-type (triangles), PhtD-deficient (circles) and Pht quadruple mutant (squares) were also determined in MS with (open symbols) or without (filled symbols) 200 µM Zn²⁺ (b), 200 µM Mn²⁺ (c) or 200 µM Fe²⁺ (d). Each figure depicts the results of one representative experiment of three.

**Fig. 7.** Expression of the Pht proteins is regulated by zinc. WU2 bacterial cells were cultured with or without TPEN 30 µM, a zinc chelator. Next, cells were probed with anti-PhtB/D (a), anti-PhtE (b), anti-PhtD/E (c) or anti-type 3 polysaccharide (d) antibodies followed by AlexaFluor-conjugated goat anti-mouse secondary antibody before they were analysed by flow cytometry. As controls, cells were incubated with the secondary conjugate antibody. Representative FACS plots of the different conditions are shown.

Pht protein family

http://mic.sgmjournals.org

Downloaded from www.microbiologyresearch.org by IP: 54.70.40.11
On: Thu, 03 Jan 2019 08:46:12
PhD in all strains. Likewise, the other Pht isotypes were found by immunoblotting in all strains that carry their respective genes. Notably, due to the highest genetic divergence, PhtE was better detected with a polyclonal antibody specifically raised against it (Fig. 8). Some peculiar Pht isotypes were found, such as a PhtE of a lower size (10 kDa smaller) in six isolates, and of an even smaller size (20 kDa smaller) in eight strains. Likewise, four strains were found to produce a truncated PhtA (Fig. 8), the gene of which was not detected by PCR. Interestingly, these four strains also expressed the 20-kDa-truncated PhtE. Sequencing of the phtA/B locus of phtB-negative strains revealed that the only gene present in this locus was a hybrid between either phtA and phtB or phtA and phtD genes.

Interestingly, sequence analysis demonstrated that the signal sequence encoded by pht genes was specific for each Pht family member. Indeed, the specific signal sequence of a Pht family member differs at least in one position from the signal sequence of another Pht family member (Table 1).

Next, we attempted to determine whether links can be made between the Pht expression profile and the isolate genotype/serotype. In the strains analysed, all serotype 2, 4, 14, 6B and 7F isolates possessed the four Pht isotypes, and all serotype 3, 9, 19F and 22F isolates lacked PhtA or carried a smaller PhtA.

Regarding a potential link between MLST genotype and Pht expression profile, the following features were noted: the 10-kDa-truncated PhtE was found mainly in the genotype ST199 group. The serotypes of these strains are 19F, 19A, 15A, 1 and 6A. The 20-kDa-truncated PhtE was observed in eight isolates that all belonged to the same clonal lineage (founder sequence ST156), but carrying different serotypes (9, 19A, 19F and 14). Finally, strains lacking PhtA were observed in different clonal lineages. Therefore, no major link between lack of PhtA and genotype was identified.

**PhD conservation**

In our study of Pht occurrence, PhD was found to be present among all pneumococcal strains tested, making it the best vaccine candidate among the Pht family. In this respect, it was essential to determine the level of sequence conservation among pneumococcal strains; therefore, DNA sequencing was carried out.

From the analysis of 107 strains (based on MLST classification), it was determined that PhD varies in length between 831 and 853 amino acids and has a molecular mass of around 100 kDa. PhD was found to be highly conserved among the 107 strains tested and only one sequence displayed a stop for a truncated protein (strain 4/75, serotype 4). The proline-rich region contained 13–15 prolines for all strains (in seven strains, only 11–13 prolines). Limited stretches of variability of <4 amino acids were found in the sequence of PhD.

**DISCUSSION**

The Pht proteins are promising candidates to be incorporated in a vaccine against pneumococcal infectious diseases. In that respect, it appeared crucial to investigate how the expression of these proteins is regulated and to better define their role in pneumococcal pathogenesis.

Genome analysis, combining the results obtained in the 4/CDC, TIGR4 and WU2 strains, showed that the four gene homologues are arranged in double tandem. The presence of a fifth, although truncated, member of the pht gene family downstream of the phtE gene was also shown, confirming the finding in a previous study (Adamou et al., 2001). It seems that this truncation is conserved as the same organization was found in the S. pneumoniae strain R6 genome (accession no. AAK99714) (Hoskins et al., 2001).

Our study showed that the tandem organization of the pht genes does not correlate with a pht bicistronic transcription. None of these genes was co-transcribed with its

---

**Fig. 8.** Western blot analysis illustrating the different SDS-PAGE migration patterns of the Pht protein family. Whole-cell extracts were separated by using SDS-PAGE and immunobotted. Strains D39 (serotype 2, ST595), 00-1621 and 00-2292 (lanes 1, 2 and 3, respectively) were probed with a polyclonal anti-PhtD (a), and strains D39, FG404293 and 00-2569 (lanes 4, 5 and 6, respectively) with a polyclonal anti-PhtE (b). Molecular mass markers are shown in lane M.
related pht neighbour under the conditions tested. Promoter and terminator analyses correlated well with traditional RT-PCR studies. We found evidence that the phtB, phtA and phtE genes possess individual putative promoters and that mRNA transcription probably ends soon after the corresponding stop codons. On the other hand, the phtD gene was different in that no promoter was identified in silico for this gene. Instead, promoters, but no transcription terminators, were identified for lmb and yfnA, two genes located upstream of phtD, which tended to indicate that those genes are organized in an operon system. This corroborates the recent finding that phtD may be expressed in a large operon system together with the four genes upstream (Loisel et al., 2008). Nevertheless, the fact that a promoter was identified for yfnA and for lmb indicates that transcription may start at these locations, meaning that phtD-containing transcripts of different length may be produced. In addition, an adCR binding site was identified upstream of the lmb gene (Loisel et al., 2008; Panina et al., 2003), suggesting that a zinc-regulated bicistronic transcript with lmb and phtD may also exist. In line with this, Spellerberg et al. (1999) showed that the group B streptococcal lmb gene is co-transcribed with a gene whose product shows 67% sequence similarity to the first 225 (phtE) and first 228 (phtA, phtD and phtB) amino acids of pht gene products (accession no. AF062533). A comparable genomic arrangement was also observed in the group A streptococcal genome (Ferretti et al., 2001). Furthermore, it was proposed that co-transcription of lmb and phtD indicates a functional link, with the former gene product being involved in pneumococcal adhesion and invasion (Panina et al., 2003).

It is interesting to note that the phtD gene can be transcribed as a polycistronic message with the two other genes, namely yfnA and lmb, which may be involved in transport and specific binding activities, respectively. Indeed, YfnA in S. pneumoniae (Hoskins et al., 2001) and the homologous proteins in Bacillus subtilis (Yamamoto et al., 1997), S. pyogenes (Ferretti et al., 2001) and Streptococcus mutans (Ađić et al., 2002) are thought to be amino acid transporters, members of the superfamily of permeases. The Lmb protein has been described as an ABC transporter-like zinc-binding protein (Loisel et al., 2008) and a putative laminin-binding protein (Spellerberg et al., 1999). Indeed, this protein shows similarities to an adhesin family known as Lra found initially in oral streptococci (Jenkinson, 1994) and since then discovered in other streptococci and other genera (Cockayne et al., 1998). It was suggested that Lra-like proteins are involved in the colonization of human epithelium by streptococci and their subsequent invasion into the bloodstream (Elsner et al., 2002). It is not clear why yfnA, lmb and phtD are associated in an operon system. One plausible hypothesis is that those three proteins are required at the same moment of the bacterial cycle, for invasion or growth, for instance, without necessarily being associated in their functions. Determination of the role of the Pht proteins might give some clue to this genomic association. Regarding their role, we can first speculate that similarities between intra-species Pht proteins are indicative of interexchangeability. These proteins might also share similar functions through their homologous regions and, at the same time, exert distinct activities, even at different development phases of the bacterium, through their variable regions. The results we have obtained from immunoblotting with protein extracts from the various Pht-deficient mutants tend to show that there is no compensation for gene loss by increasing the level of expression of the remaining pht gene products. This feature was also described recently at the RNA level by using RT-PCR (Ogunniyi et al., 2009).

As already mentioned, all Pht proteins share histidine triad motifs (Adamou et al., 2001; Hamel et al., 2004; Zhang et al., 2001), thought to be involved in metal binding. Interestingly, it has been speculated that these motifs might be more involved in zinc binding, especially to generate conformationally functional Pht proteins (Panina et al., 2003). The same authors also hypothesized that a zinc-restricted environment induces the expression of the Pht proteins and favours Streptococcus colonization and invasion. In this context, we carried out experiments in which wild-type and Pht-deficient strains were cultured under different conditions of ion depletion and supplementation. In MS, wild-type and PhD-deficient strains grew more slowly than in rich LB medium, but almost no growth of the quadruple Pht-deficient mutant was observed in the minimal medium. Strikingly, when Zn\(^{2+}\) or Mn\(^{2+}\) was added, growth of the quadruple mutant was restored up to that of the wild-type, and this was particularly visible at concentrations between 20 and 200 μM. However, our results show that growth of the quadruple mutant was

### Table 1. Signal sequence comparison of each Pht family member

<table>
<thead>
<tr>
<th>Signal sequence of:</th>
<th>No. of sequences analysed</th>
<th>Signal sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhtA</td>
<td>7</td>
<td>MKINKKLVGSAAALILSV</td>
</tr>
<tr>
<td>PhtB</td>
<td>8</td>
<td>MKINKKLAGSVASVLALSV</td>
</tr>
<tr>
<td>PhtD</td>
<td>72</td>
<td>MKINKKLAGSVASVLALSV</td>
</tr>
<tr>
<td>PhtE</td>
<td>8</td>
<td>MKFEEKKYAAAGSIVSVSL</td>
</tr>
</tbody>
</table>

Amino acids conserved in at least three Pht family members are shown in bold.
delayed compared with the wild-type. Triple and double mutants, expressing only one of the four PhT proteins, were also constructed and used in this type of experiment. From these investigations it appeared that strains with only one PhT type behave as wild-type strains, suggesting redundancy inside the PhT protein family.

These observations, besides confirming the requirement of Zn\(^{2+}\) and Mn\(^{2+}\) for bacterial growth, argue for a critical role of the PhT family in Zn\(^{2+}\) and Mn\(^{2+}\) uptake. The fact that Zn\(^{2+}\) deprivation induces de novo synthesis of proteins of the PhT family and that Zn\(^{2+}\) addition represses it is a further argument to support a close relationship between PhT and Zn\(^{2+}\). This regulation is likely to occur through AdcR protein regulating zinc uptake in S. pneumoniae. Indeed, putative binding sites for AdcR protein have been found upstream of the phTA, phTB and phTE genes, and of the lmb–phTD operon (Loisel et al., 2008; Panina et al., 2003). Binding of AdcR, induced under conditions of high Zn\(^{2+}\) concentration, inhibits the transcription of the genes under its dependence. Upon direct or indirect zinc starvation conditions, and hence a reduced intracellular concentration of this metal, repression by AdcR is relieved (Brenot et al., 2007; Claverys, 2001). However, and in contrast to what we have observed here, it was recently shown that the addition of zinc in culture medium elicits PhT production (Ogunniyi et al., 2009). Therefore, it is reasonable to estimate that production of PhT follows a bell-shaped curve within a given range of zinc concentrations. Also, the high zinc concentration effects observed by Ogunniyi et al. (2009), leading to increased PhT expression, may have little in vivo relevance as free zinc concentrations available in the human host are very low.

Dintilhac et al. (1997) concluded that, besides Psa, described as an ABC-type Mn\(^{2+}\) permease, and Adc, an ABC-type Zn\(^{2+}\) permease, a third transporter should exist, capable of transporting both Zn\(^{2+}\) and Mn\(^{2+}\). The PhT proteins or the laminin-binding protein would appear as candidates to fulfil this function. Our results are indicative of an additional role for the PhT isotypes. Indeed, the fact that wild-type and PhTD-deficient strains were able to grow in minimal medium in the absence of Zn\(^{2+}\) and Mn\(^{2+}\) is intriguing. In addition, the observation that growth of the quadruple mutant was rescued with a delay when Zn\(^{2+}\) or Mn\(^{2+}\) was added to the minimal medium is also of note. These observations could be explained if we consider that the PhT proteins act as Zn\(^{2+}\) and Mn\(^{2+}\) scavengers, with the function to store and concentrate these divalent cations. When wild-type and PhTD-deficient mutant strains were placed in minimal medium, they were able to start growing immediately as a result of the ions stored previously within the PhT proteins when the bacteria were in a richer medium. In contrast, the quadruple PhT mutants were not able to store these ions when placed in favourable conditions, and therefore could not grow when placed in poor medium. When Zn\(^{2+}\) or Mn\(^{2+}\) were added in excess to minimal medium, time was needed before the ions could be trapped by specific metal permeases or other metal transporters, because they had to ‘find’ them at random in the culture medium, without help from the PhT proteins. This might explain the delay needed for the quadruple PhT mutant to start growing under such conditions. Moreover, a possible scavenging and storage role for the PhT proteins is consistent with the presence of five to six cation-binding domains.

This speculative mechanism of storage, if confirmed, could also be considered as a means for the bacterium to regulate zinc and probably manganese homeostasis. Metal ions such as Zn\(^{2+}\) or Mn\(^{2+}\) are essential trace elements. However, they are potentially harmful to the bacterium when in excess (as observed when 1 mM Zn\(^{2+}\) was added to MS) because they may compete with other elements as cofactors for some critical enzymes. Therefore, it is essential for the bacteria to regulate metal homeostasis, and we suggest that the PhT family is involved in such regulation. Such a system would allow S. pneumoniae to survive when facing ion-restricted environments, for example during the initial stages of the colonization process in the human nasopharynx (Bunker et al., 1984; Harlyk et al., 1997).

The existence of polycistronic transcripts with PhTD might be explained by the requirement of Zn\(^{2+}\) or Mn\(^{2+}\) for Lmb, an Lra family member, and YfnA to exert their function. In partial support of this, it has been suggested that Mn\(^{2+}\) is required for adhesion through the Lra family of proteins, a critical feature for virulence (Dintilhac & Claverys, 1997; Papp-Wallace & Maguire, 2006). In addition, it has been demonstrated in other contexts that laminin binds Zn\(^{2+}\) to promote high-affinity binding between laminin and laminin-binding proteins (Ancsin & Kisilevsky, 1996; Bandyopadhyay et al., 2002). Therefore, we may hypothesize that Lmb needs PhTD to ensure the presence of Zn\(^{2+}\) when Lmb encounters laminin, which enhances binding to the host tissues. The close link between zinc and the PhT proteins may also explain why these proteins have been associated with the inhibition of C3b (Hostetter, 1999; Ogunniyi et al., 2009). Indeed, the cleavage of C3b by factor I in the presence of factor H is regulated by zinc (Blom et al., 2003). By controlling zinc concentration in the bacterial environment, the PhT isotypes might thus contribute to C3b inhibition in some circumstances, which needs to be investigated further.

In a vaccine context, by targeting the PhT protein family, the immune system may impede the bacteria in storing and using critical ions necessary for the invasion process. Consequently, our results confirm the PhT proteins as genuine vaccine candidates against pneumococcal infections. The different members of the PhT family have already been evaluated for their potential as pneumococcal vaccine antigens (Adamou et al., 2001; Godfroid et al., 2011; Hamel et al., 2004; Ogunniyi et al., 2007; Wizemann et al., 2001; Zhang et al., 2001). PhTA, PhTB and PhTD were examined for their ability to protect mice against a subset of pneumococcal isolates (Adamou et al., 2001). PhTD was found to afford the broadest protection, while PhTA
immunization was efficient against a lesser number of the strains tested. This is in line with the results of the present study, where it is shown that PhtA is expressed in 62% of pneumococcal strains, and PhtD in 100%. Although successfully used in two studies, the potential for PhtB to elicit cross-protection is not known as it was evaluated against a single strain only (Adamou et al., 2001; Zhang et al., 2001). However, as we found it in 81% of the strains, one may expect that its inter-strain coverage might not be optimal. PhtE is found in 97% of the strains, indicative of potential broad cross-protection. However, this Pht shares only 32% identity with the three other Pht types, and its C-terminal part, the most immunogenic and conserved, is PhtE-specific. The region of PhtE common with the other Pht types is not accessible to antibodies (Adamou et al., 2001; Hamel et al., 2004). Therefore, PhtD, which is present in all strains tested, has a highly conserved amino acid sequence among pneumococci and also demonstrates potential cross-protection. However, this Pht shares 20% identity with the three other Pht isotypes, and its C-terminal part, the most immunogenic and conserved, is PhtE-specific. The region of PhtE common with the other Pht types is not accessible to antibodies (Adamou et al., 2001; Hamel et al., 2004). Therefore, PhtD, which is present in all strains tested, has a highly conserved amino acid sequence among pneumococci and also demonstrates cross-reactivity with PhtA and PhtB, represents the most suitable candidate amongst the Pht proteins for incorporation into pneumococcal vaccines.

ACKNOWLEDGEMENTS

The study was sponsored by GlaxoSmithKline Biologicals. We wish to thank Sylvie Vanderschrick, Hermenegildo Fernandez-Lopez, Régine Masengo Kalibiona, Isabelle Pinault, Julie Sansfaçon and Nathalie Ramet for their excellent technical assistance. Special thanks to Josée Hamel for her contribution to the initiation and development of the project. We are also grateful to Pascal Cadot and Ulrike Krause for editorial assistance. S. R., C. N., E. D. P., L. T., M.-C. M., C. F., D. M. and J. T. P. are employees of Glaxo SmithKline Biologicals.

REFERENCES


Edited by: S. D. Bentley