Alkaline Phosphatase Reporter Transposon for Identification of Genes Encoding Secreted Proteins in Gram-Positive Microorganisms

Carmela M. Gibson and Michael G. Caparon*

Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110-1093

Received 2 July 2001/Accepted 16 October 2001

We describe the construction of TnFuZ, a genetic tool for the discovery and mutagenesis of proteins exported from gram-positive bacteria. This tool combines a transposable element (Tn4001) of broad host range in gram-positive bacteria and an alkaline phosphatase gene (phoZ) derived from a gram-positive bacterium that has been modified by removal of the region encoding its export signal. Mutagenesis of Streptococcus pyogenes with TnFuZ (“FuZ” stands for fusions to phoZ) identified genes encoding secreted proteins whose expression was enhanced during growth in an aerobic environment. Thus, TnFuZ should be valuable for analysis of protein secretion, gene regulation, and virulence in gram-positive bacteria.

Certain bacterial enzymes have activities that depend on their subcellular localization. For example, the alkaline phosphatase encoded by phoA of Escherichia coli is enzymatically active only when it has been transported across the cellular membrane into the periplasmic space. This property has been exploited to engineer PhoA as a molecular sensor of subcellular location (for a review, see reference 22). In a typical application, the region of phoA encoding its own promoter, translation initiation site, and signal peptide is removed and it is then fused with the gene that encodes the protein of interest. Should the fusion partner contribute an export signal, the bacterial cell expressing the hybrid protein will demonstrate alkaline phosphatase activity that can easily be detected through the use of several different assays (20). This strategy has been widely used in the analysis of the topologies of transmembrane proteins, for the identification and analysis of protein export signals, and for the identification of proteins that are targeted for export (19, 20). For the last application, the development of transposon-based methods for the construction of fusions to phoA (21) has allowed analyses to be conducted on a genome-wide scale with a wide variety of gram-negative bacterial species (for a review, see reference 13). Unfortunately, a similar technology has not been available for gram-positive bacteria.

As a group, gram-positive bacteria include species that are important for many industrial processes, for the production of food, and for the production of antibiotics as well as species that are model organisms for the study of development. Many important human and animal pathogens are gram positive, including several that are resistant to multiple antibiotics. While protein export in gram-positive bacteria shares a basic similarity with that in gram-negative bacteria, there are many notable differences and the pathways of export are not well characterized (35). Unlike most gram-negative bacteria, gram-positive bacteria typically secrete numerous proteins into the external environment surrounding the cell. This characteristic of protein secretion is particularly true for the pathogenic gram-positive species, and it is likely that many of these secreted proteins promote virulence (12, 16, 31). Thus, a transposon-based method for identifying mutations in genes that encode proteins targeted for export would be of broad application in the study of protein secretion by and the pathogenicity of gram-positive bacteria. The work reported here describes the construction of TnFuZ, a transposon-based genetic tool for the discovery and study of proteins exported from gram-positive bacteria.

Construction of TnFuZ. An essential component of TaphoA, the prototype element developed for gram-negative bacteria, is the modified phoA reporter (21). Unfortunately, E. coli-derived PhoA is poorly active when it is expressed in gram-positive hosts (27), which may reflect differences in the export pathway from that in gram-negative bacteria and which may result in inefficient dimerization and/or formation of a disulfide bond required for its activity (27). Other secretion reporters, including DNase (29), have been developed for gram-positive bacteria, but the fact that some gram-positive bacteria naturally secrete multiple DNases (37) has limited use of this enzyme. Recently, an alkaline phosphatase (PhoZ) derived from the gram-positive bacterium Enterococcus faecalis (15) has been developed as a reporter that is highly active in gram-positive bacteria (10, 15). The alkaline phosphatase activity of PhoZ, like that of PhoA, is dependent on its export from the cytoplasm (15), and the activity of a derivative of PhoZ lacking its leader peptide can be restored through fusion with a heterologous exported polypeptide (10, 15). The fusion partner can be large, and the resulting chimera can be stable and demonstrate high specific activity (10, 15).

The second essential component of TaphoA is Tn5, a transposon that both inserts randomly and can accept a reporter gene immediately adjacent to one of the short inverted repeats that defines the end of the element (21). However, Tn5 does not transpose efficiently in gram-positive hosts. A suitable alternative is transposon Tn4001, a Tn5-like element derived from gram-positive bacteria that transposes with a high degree of randomness in gram-positive bacteria and mycoplasmas (11, 18). A modified derivative of Tn4001 which consists of a single

* Corresponding author. Mailing address: Department of Molecular Microbiology, Washington University School of Medicine at Washington University Medical Center, Campus Box 8230, 660 S. Euclid Ave., St. Louis, MO 63110-1093. Phone: (314) 362-1485. Fax: (314) 362-1232. E-mail: caparon@borcin.wustl.edu.
copy of the insertion sequence IS256 modified to accept an antibiotic resistance gene immediately adjacent to its left-end short (28-bp) terminal inverted repeat has been constructed (18). In the present study, we used this site to introduce a modified phoZ as follows. Primers 5′PhoEcoRI (GCGGGTT GTA CGAATTCTAC TGAACAACAAAA AGCGCGCGAAA AAC) and PhoS-SalI (CGTTCTGCTT TGTCGACATT TTG RV site that was introduced into IS256). The gene encoding the Enterococcus faecalis alkaline phosphatase (phoZ) was altered by removal of the region which encoded its signal sequence, and the modified gene (phoZ*) was introduced into the element. As shown, the bottom of the figure shows the DNA sequence and open reading frame which extends across IR_L and joins the open reading frame of phoZ*. Abbreviations: C, ClaI; Ev, EcoRV; Nc, NcoI; P, PstI; S, SalI; Sm, SmaI. A slash indicates a junction of two restriction fragments joined during construction of the element, and restriction sites enclosed by parentheses are inactive.

**FIG. 1.** Construction of TnFuZ. The structure of TnFuZ is shown in the top line. The element contains the left and right inverted repeats (IR_L and IR_R, respectively), the transposase (tnp) of IS256, and the ahpA3 kanamycin resistance determinant contained on ΩKm-2. The gene encoding the Enterococcus faecalis alkaline phosphatase (phoZ) was altered by removal of the region which encoded its signal sequence, and the modified gene (phoZ*) was introduced into the element. As shown, the bottom of the figure shows the DNA sequence and open reading frame which extends across IR_L and joins the open reading frame of phoZ*. Abbreviations: C, ClaI; Ev, EcoRV; Nc, NcoI; P, PstI; S, SalI; Sm, SmaI. A slash indicates a junction of two restriction fragments joined during construction of the element, and restriction sites enclosed by parentheses are inactive.

**Production of insertion libraries.** The group A streptococcus Streptococcus pyogenes is the causative agent of a wide variety of suppurative diseases that affect the skin (impetigo, necrotizing fasciitis, erysipelas) and pharynx (pharyngitis). In addition, S. pyogenes can cause the nonsuppurative sequelae of rheumatic fever and glomerulonephritis. The ability to cause a range of diseases in many different host tissues is likely facilitated by the ability of the organism to secrete upwards of 40 different polypeptides (16). Unfortunately, only a few of these gene products have been directly linked to pathogenesis and the regulation of these putative virulence determinants is still poorly understood, making S. pyogenes an ideal model for evaluation of the utility of TnFuZ. The ColEI-based plasmid used for the construction of TnFuZ does not replicate in gram-positive hosts (18). Thus, transformation of S. pyogenes with the TnFuZ-containing plasmid (pCMG8) by electroporation (6) with selection for the kanamycin resistance determinant of ΩKm-2 yielded between 10^2 and 10^4 transformants μg of DNA^-1, depending on the specific S. pyogenes strain used. The incorporation of an alkaline phosphatase substrate directly into medium did not prove to be an efficient method for detection of alkaline phosphatase activity in colonies, likely because lactic acid bacteria like S. pyogenes acidify their surroundings. However, the use of an alternative method in which the pH of colonies was neutralized after they were lifted from the plates on filters (27) allowed the detection of blue colonies that ranged in color from intensely blue to light blue with the substrate XP (5-bromo-4-chloro-3-indolylphosphate) (Fig. 2). The yield of colonies with detectable alkaline phosphatase activity was between 1 and 4% of the total number of colonies analyzed. Untransformed S. pyogenes strains produced no detectable color (Fig. 2), which was consistent with the observation that there is no gene identified as encoding an alkaline phosphatase in the S. pyogenes strain whose genome sequence has been determined (9).

**Characterization of fusion strains.** Since virulence factors are frequently subject to coordinate regulation, screening for TaphoA insertions that are regulated by the same environmental signals that control expression of known virulence factors has proved to be a powerful strategy for identification of novel virulence factors in gram-negative bacteria (22). Oxygen is one environmental signal that is known to be involved in the regulation of virulence determinants in S. pyogenes (36). To eval-

**FIG. 1.** Construction of TnFuZ. The structure of TnFuZ is shown in the top line. The element contains the left and right inverted repeats (IR_L and IR_R, respectively), the transposase (tnp) of IS256, and the ahpA3 kanamycin resistance determinant contained on ΩKm-2. The gene encoding the Enterococcus faecalis alkaline phosphatase (phoZ) was altered by removal of the region which encoded its signal sequence, and the modified gene (phoZ*) was introduced into the element. As shown, the bottom of the figure shows the DNA sequence and open reading frame which extends across IR_L and joins the open reading frame of phoZ*. Abbreviations: C, ClaI; Ev, EcoRV; Nc, NcoI; P, PstI; S, SalI; Sm, SmaI. A slash indicates a junction of two restriction fragments joined during construction of the element, and restriction sites enclosed by parentheses are inactive.

**Production of insertion libraries.** The group A streptococcus Streptococcus pyogenes is the causative agent of a wide variety of suppurative diseases that affect the skin (impetigo, necrotizing fasciitis, erysipelas) and pharynx (pharyngitis). In addition, S. pyogenes can cause the nonsuppurative sequelae of rheumatic fever and glomerulonephritis. The ability to cause a range of diseases in many different host tissues is likely facilitated by the ability of the organism to secrete upwards of 40 different polypeptides (16). Unfortunately, only a few of these gene products have been directly linked to pathogenesis and the regulation of these putative virulence determinants is still poorly understood, making S. pyogenes an ideal model for evaluation of the utility of TnFuZ. The ColEI-based plasmid used for the construction of TnFuZ does not replicate in gram-positive hosts (18). Thus, transformation of S. pyogenes with the TnFuZ-containing plasmid (pCMG8) by electroporation (6) with selection for the kanamycin resistance determinant of ΩKm-2 yielded between 10^2 and 10^4 transformants μg of DNA^-1, depending on the specific S. pyogenes strain used. The incorporation of an alkaline phosphatase substrate directly into medium did not prove to be an efficient method for detection of alkaline phosphatase activity in colonies, likely because lactic acid bacteria like S. pyogenes acidify their surroundings. However, the use of an alternative method in which the pH of colonies was neutralized after they were lifted from the plates on filters (27) allowed the detection of blue colonies that ranged in color from intensely blue to light blue with the substrate XP (5-bromo-4-chloro-3-indolylphosphate) (Fig. 2). The yield of colonies with detectable alkaline phosphatase activity was between 1 and 4% of the total number of colonies analyzed. Untransformed S. pyogenes strains produced no detectable color (Fig. 2), which was consistent with the observation that there is no gene identified as encoding an alkaline phosphatase in the S. pyogenes strain whose genome sequence has been determined (9).

**Characterization of fusion strains.** Since virulence factors are frequently subject to coordinate regulation, screening for TaphoA insertions that are regulated by the same environmental signals that control expression of known virulence factors has proved to be a powerful strategy for identification of novel virulence factors in gram-negative bacteria (22). Oxygen is one environmental signal that is known to be involved in the regulation of virulence determinants in S. pyogenes (36). To eval-
Non-oxygen-regulated groups, were analyzed, and in each case
strains, including representatives from both the oxygen- and
dependency Systems Inc.). Chromosomes from a total of 10 fusion
according to the recommendations of the manufacturer (Fli-
dideoxynucleotides (Big Dye terminators; PE Applied Biosys-
DNA sequencing reaction with the primer EnPhoR1 (TGCCT
as described previously (6) and directly used as templates in a
TnPhoZ activity (JT9) (Table 1) appears bright blue, while a TnFuZ
insertion strain (JT0) that did not express activity remains colorless.

FIG. 2. Detection of colony-associated PhoZ activity. Colonies from overnight culture on solid Todd-Hewitt yeast extract medium were transferred to a nitrocellulose filter and exposed to the chromogenic alkaline phosphatase substrate XP as described in the text. A representative S. pyogenes TnFuZ insertion strain that expressed TnPhoZ activity (JT9) (Table 1) appears bright blue, while a TnFuZ insertion strain (JT0) that did not express activity remains colorless.

An important facet of the use of TnFuZ is that it generates mutations through insertion into its target genes. Comparison of the sequences obtained for the set of fusion strains described above to data in the streptococcal genome database (9) was done to obtain the sequences of the entire targeted open reading frames, which were then compared to sequences in the Entrez nucleotide sequence database using TBLASTN (1). This analysis revealed that while a relatively modest number of colonies were analyzed to identify fusion strains, the screen developed a rich collection of mutants (Table 2). Strains with mutations in genes encoding surface and cell wall protein antigens were obtained, as well as strains with mutations in genes encoding various enzymatic activities, transporters, and regulators of transcription (Table 2).

The collection identified only three genes that have been previously characterized in S. pyogenes or a closely related streptococcal species (Table 2). Of the surface antigens identified, T protein (encoded by tee6) (33) has been characterized as a cell wall protein anchored by the LPXTG sortase pathway (23). This pathway covalently links proteins via an LPXTG motif located toward the carboxy terminus of the protein to the cell wall peptidoglycan. The function of T protein is unknown, but it is recognized by the immune system during infection. However, there may be as many as 13 proteins in the S. pyo-

### TABLE 1. Alkaline phosphatase activities of various strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Colony-associated activity&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cell-free-supernatant activity&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>JRS4</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.150, dark blue</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT1</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.150, dark blue</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT2</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.150, dark blue</td>
<td>+O&lt;sub&gt;2&lt;/sub&gt; +0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT3</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT4</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT5</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT6</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT7</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT8</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT9</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT10</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT11</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT12</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT13</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT14</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT15</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT16</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
<tr>
<td>JT17</td>
<td>+0.051 to 0.150, light blue</td>
<td>+0.051 to 0.150, light blue</td>
</tr>
</tbody>
</table>

<sup>a</sup> Indicated strains were derived from JRS4 by insertion of TnFuZ.
<sup>b</sup> Cultures on solid medium were incubated under aerobic (+ O<sub>2</sub>) or anaerobic (− O<sub>2</sub>) conditions. Alkaline phosphatase activities were determined by assessment of colony color following staining with the XP substrate and are reported as follows: −, white; +, very light blue; ++, light blue; +++ medium blue; and ++++, dark blue.

<sup>c</sup> Strains were cultured in liquid medium. Alkaline phosphatase activities in cell-free supernatants were determined using the pana-nitrophenylphosphate substrate as described previously (11) and are reported as optical densities at 405 nm as follows: −, 0.000 to 0.024; +, 0.025 to 0.050; ++, 0.051 to 0.150; and +++, >0.150.
genes genome that contain the LPXTG anchor signal, and several of these have been implicated in pathogenesis (9).

In mutant JT8, TnFuZ had inserted into a gene that encodes a protein with homology to the IscA secreted antigen of *Staphylococcus aureus* (17). The function of IscA is unknown. How-ever, as is the case for several other putative *S. pyogenes* or related species.

**TABLE 2. Sequence analysis of selected mutants**

<table>
<thead>
<tr>
<th>Strain&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Locus&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Gene(s)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Similar gene(s) (organism)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>% Amino acid identity (% homology)</th>
<th>Putative export signal&lt;sup&gt;e&lt;/sup&gt;</th>
<th>Function</th>
<th>Entrez accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>JT4&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Spy1618</td>
<td>cysK (Bs)</td>
<td>61 (74)</td>
<td>TM (3)</td>
<td>Cysteine biosynthesis</td>
<td>P37887</td>
<td></td>
</tr>
<tr>
<td>JT5&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Spy0092</td>
<td>adeR (Spn)</td>
<td>53 (76)</td>
<td>NI</td>
<td>Transcription repressor</td>
<td>CAA96184</td>
<td></td>
</tr>
<tr>
<td>JT7</td>
<td>Spy2188</td>
<td>trmU (Bs)</td>
<td>63 (76)</td>
<td>NI</td>
<td>tRNA modifier</td>
<td>035020</td>
<td></td>
</tr>
<tr>
<td>JT8&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Spy2192</td>
<td>sa2/ssa4 (Sa)</td>
<td>54 (60)</td>
<td>SS</td>
<td>Secretert antibody</td>
<td>BAA95959/AF144681-1</td>
<td></td>
</tr>
<tr>
<td>JT9</td>
<td>Spy2192</td>
<td>tee-6</td>
<td>54 (60)</td>
<td>SI</td>
<td>Acid phosphatase</td>
<td>CAAT73175</td>
<td></td>
</tr>
<tr>
<td>JT9</td>
<td>Spy1882</td>
<td>lppC</td>
<td>58 (73)</td>
<td>TM (2)</td>
<td>Peptidoglycan synthesis</td>
<td>P70965</td>
<td></td>
</tr>
<tr>
<td>JT11</td>
<td>Spy0763</td>
<td>murA (Bs)</td>
<td>58 (73)</td>
<td>TM (2)</td>
<td>Peptidoglycan synthesis</td>
<td>P70965</td>
<td></td>
</tr>
<tr>
<td>JT13</td>
<td>Spy0242</td>
<td>fasB/gasK1</td>
<td>58 (73)</td>
<td>TM (2)</td>
<td>Peptidoglycan synthesis</td>
<td>P70965</td>
<td></td>
</tr>
<tr>
<td>JT15&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Spy2211</td>
<td>tmp5 (Lc)</td>
<td>44 (60)</td>
<td>TM (2)</td>
<td>Peptidoglycan synthesis</td>
<td>P70965</td>
<td></td>
</tr>
<tr>
<td>JT16</td>
<td>Spy2033</td>
<td>tpt4 (Sc)</td>
<td>59 (72)</td>
<td>SI</td>
<td>ABC transporter</td>
<td>AAB97959</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> An asterisk indicates that alkaline phosphatase activity was decreased in anaerobically grown cultures. See Table 1.

<sup>b</sup> Corresponding open reading frame in the sequenced genome of *S. pyogenes* SF370 (10) (GenBank accession no. AE004992).

<sup>c</sup> Previously characterized for *S. pyogenes* or related species.

<sup>d</sup> Determined by comparison to sequences in EMBL and GenBank databases using TBLASTN (1) and the National Center for Biotechnology Information network services (www.ncbi.nlm.nih.gov/BLAST/). Bs, *B. subtilis*; Spn, *Streptococcus pneumoniae*; Lc, *L. lactis*; Sa, *Staphylococcus aureus*; Sc, *Streptococcus cristatus*.

<sup>e</sup> SS, signal sequence; SI, lipoprotein signal sequence; TM, transmembrane domain; NI, none identified. For prediction of signal sequences by SignalP (25) and transmembrane domains by TopPred 2 (8), we used network services provided by Meta-PredictProtein (dodo.cpmc.columbia.edu/pp/predictprotein.html). Numbers of predicted transmembrane domains are indicated within parentheses.

It was interesting that, of all the fusion proteins produced, none represented any of the known exotoxins of *S. pyogenes*. This may be because at least one toxin appears to have a specialized export mechanism (24). In other cases, it may be that the colony method for screening was not optimal for identification of fusion proteins that do not remain associated with the cell surface or it may be a result of the fact that the number of colonies screened was relatively small. However, the simplicity of various alkaline phosphatase assays and the availability of substrates of diverse chemistries will allow development of screens optimized for detection of proteins that are completely released from the cell. For example, a large-scale screen could be conducted to analyze cell-free culture supernatants harvested from cultures arrayed in a microplate format or the low background obtained in the colony filter assay could be exploited through the use of a highly sensitive substrate.

Most of the characterized fusion proteins had identifiable export signals. As for the two genes which did not encode a putative export signal, it is possible that the location of the fusion junction generated a fortuitous export signal (38) or that the expression of the fusion protein produced cell lysis. Preliminary characterization suggests that neither of these scenarios was the case. However, it should be noted that pathways of protein secretion in gram-positive bacteria are not well understood. In *S. pyogenes* alone, there are several examples of proteins that are found exterior to the cell membrane and that lack a defined export signal (25, 26). The availability of TnFuZ will likely facilitate investigation of alternative pathways of protein secretion in gram-positive bacteria. The broad host range of the transposon used to construct TnFuZ (Tn4001) suggests that this element will find wide application in the analysis of protein secretion by and the virulence of gram-positive bacteria and mycoplasmas.

This work was supported by Public Health Service grant AI38273.

**REFERENCES**
