Identification of *Haemophilus influenzae* Rd transformation genes using cassette mutagenesis

Brian A. Dougherty† and Hamilton O. Smith

Author for correspondence: Brian A. Dougherty. Tel: +1 203 677 7604. Fax: +1 203 677 6771. e-mail: dougherb@bms.com

Genes required for natural transformation of *Haemophilus influenzae* Rd were identified by a cassette mutagenesis protocol consisting of the following steps: random insertional mutagenesis, phenotypic screening, sequencing of genome sequence tags from the DNA flanking the insertion in the selected mutants and comparison of genome sequence tags to genomic sequence data. The cassette mutagenesis screen for transformation genes resulted in five distinct mutant classes, two of which have been identified in previous studies. Insertions in the three newly identified loci interrupted genes with predicted protein products homologous to a type IV pilin-like protein biogenesis operon, drug-efflux transporters and a phospholipid-biosynthesis enzyme. The most significant finding of this screen is the requirement for type IV pilin-like proteins in genetic transformation of *H. influenzae*. These surface structures are utilized for DNA uptake in a number of Gram-positive and Gram-negative bacteria, and appear to be the common component among the systems for DNA binding.

Keywords: *Haemophilus influenzae*, genetic competence, DNA transformation, mutagenesis, genome sequence

INTRODUCTION

*Haemophilus influenzae* is a Gram-negative facultative anaerobe which resides in the human upper respiratory tract. Isolates are classified either as serotypes A–F or as non-typable, depending on the presence or absence, respectively, of type-specific capsular polysaccharide. *H. influenzae* is also capable of taking DNA from its environment and integrating it into the bacterial chromosome (Smith et al., 1981; Goodgal, 1982). This natural transformation process has been demonstrated for many strains of *H. influenzae* but has been best characterized using strain Rd (Alexander & Leidy, 1951). The transformation frequency for *H. influenzae* Rd grown in rich medium increases from $10^{-7}$ (1 transformed cell per $10^7$ viable cells) in early exponential phase to about $10^{-4}$ in late exponential phase. The addition of cAMP to early-phase cultures increases the frequency of transformation to $10^{-4}$ (Wise et al., 1973; Redfield, 1991). Furthermore, under conditions of transient anaerobiosis (Goodgal & Herriott, 1961) or transfer to starvation medium (Herriott et al., 1970) transformation frequencies are optimized to approximately $10^{-2}$. Transformation is also observed when *H. influenzae* is grown in diffusion chambers implanted into animals (Dargis et al., 1992), suggesting that transformation is not an artifact of *in vivo* cultivation.

A number of genes required for DNA transformation have been identified in *H. influenzae*. These studies have used a variety of molecular methods to identify transformation genes, including complementation of transformation-deficient mutants derived by chemical mutagenesis (Larson et al., 1991), a 'poison-DNA' selection method (Beattie & Setlow, 1971), mini-Tn10 mutagenesis (Tomb et al., 1989), Tn916 mutagenesis (Gwinn et al., 1998) and other genetic selection techniques (Redfield, 1991). It has been a goal of our laboratory to identify the remaining transformation genes. To overcome potential biases inherent in transposon-based

---

†Present address: Department of Applied Genomics, Bristol-Myers Squibb Pharmaceutical Research Institute, 5 Research Parkway, Wallingford, CT 06492-7660, USA.

Abbreviations: sBHI, supplemented brain heart infusion media; Kan', kanamycin-resistance; GST, genome sequence tag; MIV, medium IV; Tfo', transformation-deficient.

The GenBank/EMBL/DDBJ accession number for the sequence reported in this paper is L42023.
screens (e.g. hot-spotting) and yet retain the advantage of gene disruption with an antibiotic-resistance marker, we decided to investigate the utility of the cassette-mutagenesis protocol (Sharetsky et al., 1991). In this procedure, an antibiotic-resistance cassette is ligated to restriction endonuclease-digested chromosomal DNA that has been treated so that the original gene order is maintained during the process. The resulting cassette-mutagenized DNA is then transformed into competent bacteria, resulting in the insertion of the cassette marker throughout the chromosome and the generation of mutants by gene disruption.

Recently, molecular analysis of *H. influenzae* has been greatly aided by the determination of the complete 1·83 Mb genome sequence (Fleischmann et al., 1995). Although genome sequence facilitates directed gene knockouts of suspected transformation genes, we reasoned that phenotypic screening of mutagenized cells for transformation deficiencies would provide a more balanced, unbiased answer of what genes are involved in DNA transformation. We therefore used cassette mutagenesis in conjunction with a phenotypic screen for transformation-deficient (Tfo-) mutants and sequencing of PCR-generated genome sequence tags (GSTs) to rapidly map mutations in the transformation genes of *H. influenzae*.

**METHODS**

**Strains and cultivation.** All strains are derived from the original Rd strain (Alexander & Leidy, 1951). The *H. influenzae* strains KW20 (Wilcox & Smith, 1975) and MAP7, a streptomycin-resistant, erythromycin-sensitive derivative of MAP (Catlin et al., 1972) were grown in brain heart infusion broth (Difco) supplemented with 10 μg haemin ml⁻¹ and 2 μg NAD ml⁻¹ (sBHl) as described elsewhere (Barcak et al., 1991). Mutant strains of *H. influenzae* strain KW20 were constructed by transformation with cassette-mutagenized DNA (see below). For antibiotic selection, streptomycin was added to media at 250 μg ml⁻¹ and kanamycin at 7 μg ml⁻¹. All chemicals were purchased from Sigma.

**Plasmids, DNA preparation and DNA manipulation.** Plasmid pUC4K was the source of the kanamycin-resistance (Kan') cassette and was obtained from Pharmacia. Plasmid DNA was prepared by alkaline lysis and purified using Qiagen columns. All polymerases, restriction enzymes and modification enzymes were purchased from New England Biolabs and used according to the manufacturer’s instructions. *H. influenzae* chromosomal DNA was prepared by standard methods (Barcak et al., 1991).

**Mutagenesis technique.** *H. influenzae* Rd was mutagenized by inserting the 1·2 kb Kan’ DNA cassette into various Sau3AI restriction sites throughout the chromosome. The *in vitro* procedure for cassette-mutagenizing chromosomal DNA (Sharetsky et al., 1991) involves five steps: (1) cleavage of the DNA with a suitable restriction enzyme (e.g. EcoRI), (2) intramolecular ligation of the restriction fragments to form circles, (3) partial digestion of the circles with a second enzyme (e.g. Sau3AI) that produces ends compatible with the Kan’ cassette (e.g. BamHI ends), (4) ligation of cut circles and cassette DNA and (5) recutting the ligated DNA with the first enzyme (EcoRI) to linearize the circle. These five steps are outlined in Fig. 1 (a) and have been modified as described below.

Step 1. *H. influenzae* Rd DNA of approximately 50 kb size was cleaved in a 1 ml reaction mixture containing 60 μg DNA and 50 U EcoRI for 2 h at 37 °C. The mixture was phenol-extracted and the DNA recovered by ethanol precipitation, redissolved in 150 μl TE buffer and adjusted to 500 μl in 1× ligation buffer.

Step 2. Intramolecular ligation of DNA fragments is favoured at low DNA concentrations, but in order to avoid dilution of DNA in subsequent steps, the EcoRI fragments were efficiently circularized using a ‘circle generator’ (H. O. Smith, unpublished). The mixing chamber of the generator was a 17×100 mm (Falcon 2059) plastic tube holding the ligation mixture (500 μl) containing 50 U T4 DNA ligase and a magnetic flea turning at about 60 r.p.m. on a magnetic stirring plate. The EcoRI-cut *H. influenzae* Rd DNA (60 μg in 500 μl) was introduced at about 100 μl h⁻¹ using a syringe pump and a 1 ml plastic tuberculin syringe. The reaction was carried out at 23 °C for about 5 h. The circularized DNA was phenol extracted, ethanol precipitated and redissolved in 50 μl TE buffer. Thus, intramolecular ligation was achieved using the circle generator by slowly introducing substrate (unligated EcoRI-generated ends) to the ligation reaction chamber rather than by simply diluting the DNA to low concentrations and adding ligase.

Step 3. DNA circles were digested with Sau3AI in a reaction mixture (100 μl) containing approximately 10 μg circles and 2 U Sau3AI. Incubation was at 23 °C and aliquots were removed every minute from 2 to 8 min into a tube containing 10 μl 0·5 M EDTA, pH 8. The pooled samples were extracted with phenol, precipitated and redissolved in 50 μl TE buffer.

Step 4. The 1·2 kb Kan’ cassette was excised from 10 μg pUC4K DNA using BamHI, gel-purified and redissolved in 50 μl TE buffer. Sau3AI-cut circles (5 μl) and cassette DNA (2 μl) were ligated together in a 20 μl reaction containing 2 U T4 DNA ligase at 16 °C for 75 min. The reaction was terminated by treatment at 70 °C for 10 min.

Step 5. One microlitre of 2M NaCl and 1 U EcoRI were added to the ligation mixture from step 4 and incubated at 37 °C for 30 min followed by 10 min at 70 °C. The mutagenized DNA was stored at −20 °C. One microlitre yielded more than 40000 Kan’ clones using frozen competent KW20 cells.

**cAMP/DNA plate screen for transformation mutants.** *H. influenzae* strain KW20 was made competent by the MIV (medium IV) procedure (Herriott et al., 1970; Barcak et al., 1991). Frozen competent cells were transformed with cassette-mutagenized DNA and transformants (~200 colonies per plate) were immediately plated on kanamycin-supplemented sBHI and incubated overnight at 37 °C. The colonies were subsequently replica-plated first onto a kanamycin-supplemented sBHI plate and then onto sBHI plates supplemented with 1 mM cAMP onto which 50 μg MAP7 chromosomal DNA had been freshly spread and allowed to dry (modified from Wise et al., 1973). After overnight growth, colonies from the cAMP/DNA plates were replica-plated onto streptomycin-supplemented sBHI plates. Colonies absent after overnight growth on the streptomycin plates were noted and the corresponding colony from the kanamycin-supplemented replica plate was recovered and frozen for use in subsequent studies.

**DNA-binding assay.** The DNA-binding assay was performed using T4 DNA polymerase to incorporate 32P into MAP7.
Cassette mutagenesis of *H. influenzae*

(a) Cassette mutagenesis

![Diagram of cassette mutagenesis](image)

- Chromosomal DNA
- Step 1: EcoRI digestion
- Step 2: dilute ligation
- Step 3: partial Sau3AI digestion
- Step 4: ligation to cassette
- Step 5: EcoRI linearization
- Kan'
- Transform
- Select Kan'
- Screen

(b) Inverse PCR

![Diagram of inverse PCR](image)

- Ligation product
- PCR product
- SspI digestion
- dilute ligation
- PCR
- Sequence PCR products (genome sequence tags)

**Fig. 1.** (a) Schematic representation of the cassette-mutagenesis procedure. The various steps outlined in the cartoon are explained in greater detail in Methods. (b) Schematic representation of the inverse-PCR strategy used to amplify chromosomal DNA adjacent to the Kan' cassette. The black block represents the cassette DNA and primers 1, 2, 3 and 4 correspond to the PCR primers used in this study (primers 1 and 2 amplify the left SspI junction fragment and primers 3 and 4 amplify the right SspI junction fragment). Note that the orientation of the DNA in the PCR products has been inverted relative to that in the ligation products (the stippled box represents chromosomal DNA proximal to the cassette).

chromosomal DNA. Briefly, a 100 μl reaction containing 12 μg MAP7 DNA and 15 U T4 DNA polymerase in 1 x NEB T4 DNA polymerase buffer was incubated at 37 °C for 25 min to digest the 3' ends of the chromosomal DNA. This step was followed by the addition of dATP, dGTP and dTTP (100 μM final) and 30 μCi [α-32P]dCTP (1110 kBq; Amersham) for a fill-in reaction at 37 °C for 20 min. The chromosomal DNA was purified from unincorporated label by G-50 column chromatography and the volume adjusted to 500 μl with TE buffer (specific activity, 1.4 x 10^6 cpm μg⁻¹). The radiolabelled DNA (10 μl) was added to MIV-competent cells (1 ml) and following incubation at 37 °C for 10 min, the samples were transferred to an ice bath. The sample was divided equally between two tubes in order to measure DNA binding (total cell-associated counts) and DNA uptake (DNase I-resistant counts). The DNA uptake tube was treated with 10 μg DNase I for 5 min followed by the addition of NaCl to a final concentration of 0.5 M. Both samples were then centrifuged and washed once with either 1 ml MIV (DNA binding tube) or MIV containing 0.5 M NaCl (DNA uptake tube). Samples were again centrifuged, the supernatant carefully removed and the radioactivity associated with pellets quantified using a scintillation counter.

**Inverse PCR and DNA sequencing.** Oligonucleotide primers were synthesized using an Applied Biosystems 380B DNA synthesizer at the Protein/Peptide/DNA Facility of Johns Hopkins University. Primers were designed to be complementary to the Kan' cassette (accession no. X06404) and included tailored restriction sites. The primer sequences were the following: primer 1, 5' GAATTTAATCGCGGCTTCCGAG 3'; primer 2, 5' CCCTAGGCACTCATCAAC 3'; primer 3, 5' AGGAATCTCTGCCAGCGCATCAC 3'; primer 4, 5' GATTCCAGATCTTGATGATGTC 3'. A fifth primer used only for sequencing was the BSOut primer (5' CCCTGCAAGTGTCAGGGATC 3'), designed to read out from the BamHI-SalI-PstI polynucleotyper flanking the Kan' cassette.

The inverse-PCR scheme is diagrammed in Fig. 1 (b). Template for the PCR reaction was prepared by SspI digestion of chromosomal DNA followed by dilute ligation (≤1μg ml⁻¹). Similarly treated KW20 chromosomal DNA was used as a negative control. PCR reactions were performed in a reaction mixture (100 μl) containing 100 ng circularized chromosomal DNA, 1 μM primers, 80 μM dNTPs, 2.5 mM MgCl₂ and 1 x Gibco-BRL Taq DNA polymerase buffer lacking MgCl₂. The PCR reactions were performed using a DNA Thermal Cycler (Perkin Elmer). The Taq DNA polymerase (Life Technologies) was added to the PCR reactions upon reaching 93 °C, and the reaction conditions consisted of 30 cycles of denaturing (93 °C, 1 min), annealing (55 °C, 1 min) and extension (72 °C, 3 min), followed by a 5 min final extension (72 °C).

**DNA sequence analysis.** PCR templates were sequenced using the fluorescent dideoxy terminator method of cycle sequencing (McCombie et al., 1992) on a Perkin Elmer/Applied Biosystems 373A automated DNA sequencer (Smith et al., 1986) following the Applied Biosystems protocols at the DNA Analysis Facility of Johns Hopkins University. DNA sequence information from the inverse PCR templates was mapped to the *H. influenzae* genome sequence using search tools at the TIGR World Wide Web site (http://www.tigr.org/tdb/mdb/hidb/hidb.html). In addition to using the *H. influenzae* annotation (Fleischmann et al., 1995), ORFs surrounding the Kan' insertions were identified using DNA Strider, BLAST homology alignments to the NCBI non-redundant protein database and visual inspection. Readers are referred to the L4023 accession number, the genome sequence publication, and TIGR Worldwide Web site URL cited above for the DNA and protein sequences.
RESULTS

Screen for transformation-deficient mutants

To insert kanamycin resistance (Kan') cassettes throughout the H. influenzae Rd genome, the cassette-mutation protocol (Sharetsky et al., 1991) was used to mutagenize strain KW20 (Fig. 1a; see Methods). Approximately 8000 Kan' colonies were then screened for a Tfo- phenotype. This involved a replica-plate based assay (the cAMP/DNA-plate assay) to identify Tfo- mutants by exposing the colonies to chromosomal DNA from a streptomycin-resistant strain, then challenging the colonies on media containing streptomycin. Mutants which failed to grow in the presence of the antibiotic were deemed candidate Tfo- mutants, and were recovered from a replica plate. Following a rescreening of candidate mutant colonies to reduce false negatives, 26 possible Tfo- clones were identified. These strains were then further analysed by the MIV and DNA-binding assays.

Characterization of Tfo' mutants by DNA transformation and DNA-binding assays

To confirm and further characterize the Tfo- phenotypes of the mutant strains, each mutant was made competent by the MIV procedure (Herriott et al., 1970; Barcak et al., 1991) and assayed for transformation efficiency. Three of the 26 mutants failed to grow sufficiently to be assayed and three other mutants displayed wild-type levels of transformation by the MIV procedure (data not presented). Of the remaining 20 mutants, the MIV transformation and DNA-binding efficiencies of 15 were further examined, along with wild-type KW20 and one Tfo+ transformant (Table 1). The transformation efficiencies of the Tfo' mutants ranged over six orders of magnitude, from slightly deficient to extremely deficient (i.e. background levels of transformation). This suggested insertions into a variety of sites in the H. influenzae Rd chromosome, which was confirmed by subsequent DNA sequencing of the mutants (see below).

The MIV-competent cells were then assayed for DNA binding using 32P-labelled chromosomal DNA. The DNA-binding and DNA-uptake values listed in Table 1 represent total cell-associated radioactivity and DNase-resistant counts, respectively. All mutants showed at least a 10-fold reduction in DNA binding, with 11/16 mutants displaying background levels of DNA binding. Under DNase I and high-salt conditions, all mutants had background levels of DNA uptake cell-associated counts, with the exception of the BD307 strain. These experiments confirmed that the mutants identified in the cAMP/DNA-plate assay and subsequently by MIV-

<table>
<thead>
<tr>
<th>Strain</th>
<th>Region</th>
<th>MIV*</th>
<th>DNA binding†</th>
<th>DNA uptake†</th>
</tr>
</thead>
<tbody>
<tr>
<td>KW20</td>
<td>Wild-type</td>
<td>3.0 x 10^-8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>BD300</td>
<td>Unknown</td>
<td>3.0 x 10^-8</td>
<td>99.70</td>
<td>100.20</td>
</tr>
<tr>
<td>BD306</td>
<td>Com</td>
<td>5.8 x 10^-9</td>
<td>0.55</td>
<td>0.16</td>
</tr>
<tr>
<td>BD307</td>
<td>Multidrug resistance</td>
<td>1.0 x 10^-8</td>
<td>5.44</td>
<td>4.87</td>
</tr>
<tr>
<td>BD309</td>
<td>Pil</td>
<td>4.0 x 10^-7</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>BD310</td>
<td>Pil</td>
<td>8.5 x 10^-6</td>
<td>0.18</td>
<td>0.36</td>
</tr>
<tr>
<td>BD314</td>
<td>Pbp</td>
<td>7.0 x 10^-6</td>
<td>4.06</td>
<td>0.37</td>
</tr>
<tr>
<td>BD316</td>
<td>Com</td>
<td>4.5 x 10^-6</td>
<td>0.18</td>
<td>0.27</td>
</tr>
<tr>
<td>BD331</td>
<td>Com</td>
<td>1.8 x 10^-7</td>
<td>0.40</td>
<td>0.29</td>
</tr>
<tr>
<td>BD335</td>
<td>Pil</td>
<td>2.2 x 10^-7</td>
<td>0.45</td>
<td>0.28</td>
</tr>
<tr>
<td>BD339</td>
<td>Pil</td>
<td>1.0 x 10^-8</td>
<td>0.53</td>
<td>0.23</td>
</tr>
<tr>
<td>BD341</td>
<td>Pil</td>
<td>2.0 x 10^-8</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td>BD342</td>
<td>Com</td>
<td>1.0 x 10^-8</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>BD343</td>
<td>Multidrug resistance</td>
<td>6.3 x 10^-5</td>
<td>0.21</td>
<td>0.36</td>
</tr>
<tr>
<td>BD344</td>
<td>PgsA</td>
<td>4.6 x 10^-6</td>
<td>0.65</td>
<td>0.46</td>
</tr>
<tr>
<td>BD345</td>
<td>Pbp</td>
<td>7.0 x 10^-5</td>
<td>7.93</td>
<td>0.38</td>
</tr>
<tr>
<td>BD350</td>
<td>Pbp</td>
<td>1.8 x 10^-5</td>
<td>7.52</td>
<td>0.46</td>
</tr>
</tbody>
</table>

* Values presented as transformation frequencies (no. antibiotic resistant c.f.u. divided by total c.f.u.).
† Values calculated as cell-associated c.p.m. for strain divided by cell-associated c.p.m. for KW20, multiplied by 100.
Penicillin-binding protein region

One set of cassette insertions was found in the 29–42 kb region of the KW20 chromosome containing two clusters of genes apparently involved in cell-wall biosynthesis (Fig. 2). A previous screen for Tfo⁻ mutants resulted in a mutant (JG58; Tomb et al., 1991) with an insertion in the \textit{pbp2} gene in this gene cluster (J.-F. Tomb, unpublished). All three mutants identified here, BD314, BD345 and BD350, had similar DNA binding and uptake values (Table 1; \(\sim 4\%\) DNA binding but background levels of DNA uptake) and had somewhat lowered transformation efficiencies relative to wild-type (50–400-fold reduced). Analysis of the DNA immediately adjacent to the Sau3AI site formed by the cassette-chromosome junction in BD314, BD345 and BD350 indicated that, relative to the wild-type KW20 sequence, there were deletions of 52, 2651 and 2174 bp, respectively. This was apparently due to digestion at a second Sau3AI site in the EcoRI circle and will be addressed in more detail in the Discussion section. Strain BD345 also contained a 505 bp chimeric Sau3AI fragment between the left terminus of the cassette and the chromosomal \textit{pbp2} gene. This chimera was from a region of the KW20 chromosome approximately 530 kb away from the cell-wall biosynthesis region and was apparently ligated along with the Kan' cassette during the cassette-ligation reaction. The deletion of BD314 was within the \textit{rodA} homologue whereas the deletions associated with BD345 and BD350 resulted in the loss of several genes in these strains (Fig. 2).

\textit{com} operon

A second set of insertions (Fig. 3) were found in a genetic locus previously identified by a mini-Tn10 screen for Tfo⁻ mutants (Tomb et al., 1989) and a complementation study (Larson & Goodgal, 1991), and further
characterized as the competence-induced (com) operon (Tomb et al., 1991). Three of the insertions, mutants BD316, BD331 and BD342, were in the same Sau3AI restriction site in the comE gene. The Tfo⁻ phenotypes and lack of DNA binding agree with previously published data for comE mutants (Tomb et al., 1991). BD316 and BD342 were the only other Tfo⁻ mutants listed in Table 1 which harboured chimeric DNA adjacent to the cassette DNA.

The fourth cassette insertion mapping near the com operon was BD306, located approximately 3 kb upstream of the comA-F operon. The Kan’ cassette was associated with a 476 bp deletion of both comORFJ and HI0442 (Fig. 3). Since insertions in comA-F are the only published mutations in the com operon associated with background levels of transformation to date (Tomb et al., 1991), it will be of interest to determine the gene(s) responsible for the loss of transformation in the BD306 mutant.

pil operon

The mutant class with the most insertions was the type IV pilin-like protein class (strains BD309, BD310, BD335, BD339 and BD341). All had background levels of transformation and DNA binding/uptake (Table 1) and had insertions falling within a 4 kb region that includes an operon of genes similar to other type IV pilin genes (Fig. 4). Homology searches indicated that the operon consists of a type IV pilin subunit gene (pilA), two pilin secretion and/or assembly genes (pilB and pilC) and a prepilin peptidase (pilD). A dyad-symmetry element found upstream of multiple H. influenzae competence-induced operons (Karudaparam & Barcak, 1997) is also found upstream of the pil operon; each of the five mutants found in the pil region disrupt this genetic element. The H. influenzae pil operon ORFs are similar to a set of genes required for natural transformation in other bacteria (Chung & Dubnau, 1995; Rudel et al., 1995; Lunsford & Roble, 1997; Porstendorfer et al., 1997). Taken together, the most likely explanation for the Tfo⁻ phenotype of the five pil region mutants was lack of a functional pilin-like protein.

Table 2. Comparison of MICs of dyes, detergents and antibiotics for KW20 (wild-type) and BD307 (deleted for several genes involved in drug efflux)

<table>
<thead>
<tr>
<th>Compound</th>
<th>MIC (µg ml⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KW20</td>
</tr>
<tr>
<td>Acridine orange</td>
<td>50</td>
</tr>
<tr>
<td>Congo red</td>
<td>100</td>
</tr>
<tr>
<td>Crystal violet</td>
<td>3-13</td>
</tr>
<tr>
<td>Ethidium bromide</td>
<td>1-56</td>
</tr>
<tr>
<td>Sodium dodecyl sulfate</td>
<td>50</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>3-13</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>3-13</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>0-39</td>
</tr>
<tr>
<td>Vancomycin</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

Drug-efflux transporter region

One class of Tfo⁻ mutants consisted of two insertions, BD307 and BD343, in a cluster of genes with convergently oriented directions of transcription (Fig. 5). The mtrR, acrA and acrB genes are homologous to components of multidrug-resistance efflux transporters (Lewis, 1994; Nikaido, 1994), representing a transcriptional regulator, a membrane fusion protein and an efflux subunit, respectively. The BD307 mutant, which was severely affected in MIV transformation (Table 1), harboured a 5276 bp deletion that removes DNA from part or all of four ORFs in this region (HI0894–0897, Fig. 5). Consistent with E. coli acrA mutants (Ma et al., 1993) and H. influenzae acrA mutants (Sanchez et al., 1997), the BD307 mutant is hypersensitive to a number of hydrophobic compounds and antibiotics (Table 2); the acrA⁻ BD343 mutant is similar to wild-type KW20 (data not shown). The low DNA-binding and DNA-uptake values for BD307 (Table 1; about 5% for both) were phenotypically distinct from the cell-wall-biosynthesis mutants (low values for DNA-binding and background DNA-uptake values) or the remainder of the Tfo⁻ mutants (background values for both DNA binding and uptake). The BD343 insertion mutated only the fisN gene and showed a 25-fold reduction in MIV transformation. BD307 and BD343, both fisN mutants, showed moderate filamentous growth (4–20 cells in
length) at 37 °C, consistent with the growth characteristics of ftsN mutants (Dai et al., 1993).

**pgsA gene**

An insertion in the pgsA gene of KW20 (Fig. 6) was responsible for a 650-fold reduction in MIV transformation and background levels of DNA binding and uptake (Table 1). In *E. coli*, pgsA is an essential gene catalysing the committed step to the synthesis of acidic phospholipids (Gopalakrishnan et al., 1986; Heacock & Dohan, 1987). It has been reported that the membrane of *H. influenzae* is composed of the phospholipids phosphatidylglycerol (15%) and phosphatidylethanolamine (85%) and that the development of competence did not affect these (Sutrina & Scocca, 1976). A null mutation in *pgsA* would presumably decrease membrane levels of phosphatidylglycerol but not phosphatidylethanolamine.

**DISCUSSION**

In the present study, genes required for transformation in *H. influenzae* were rapidly identified by random insertional mutagenesis followed by phenotypic screening, inverse PCR, sequencing of mutants and comparison to genomic sequence data. We decided to use the cassette mutagenesis technique (Sharetsky et al., 1991) to take advantage of the high transformation efficiency of *H. influenzae* and to investigate the usefulness of the technique to identify transformation genes not mutated in previous screens (Beattie & Setlow, 1971; Tomb et al., 1989; Larson et al., 1991; Gwinn et al., 1998; Redfield, 1991). Cassette mutagenesis techniques based on similar approaches have been described for other transformable organisms, such as *Bacillus subtilis* (Vosman et al., 1986) or yeast (Tettelin et al., 1995). The current study represents a detailed analysis of the cassette mutagenesis procedure in *H. influenzae*.

It is apparent that the cassette-mutagenesis method, combined with the phenotypic screen, was effective, since both previously identified (i.e. *comE* and *phb2*) and newly identified transformation genes were found in this screen. The Kan' cassette inserted into distinct Sau3AI sites in all the mutants studied here, with the exception of the three *comE* mutants (BD316, BD331 and BD342). One complication of the procedure in our hands was that of deletions due to digestion of the EcoRI circle at a second Sau3AI site. A byproduct of this reaction is the generation of small Sau3AI fragments, which occasionally ligate along with the cassette to create chimeras (e.g. BD316, BD342 and BD345). It is apparent that a more limited Sau3AI digestion would remedy the deletion/chimera problem. The publication of complete genome sequence data allows the cassette-mutagenesis strategy to be refined to a high degree; however, the randomness of the insertions is still ultimately limited by restriction enzyme sites. In this study, although the screen was fairly extensive (~ 8000 colonies were screened), other known transformation genes were not isolated, possibly due to factors such as the position of genes relative to the nearest EcoRI sites, the presence of uptake signal sequences (Danner et al., 1980; Smith et al., 1995) on the EcoRI fragment, and the presence of Sau3AI sites in the central portion of the gene coding region. It is unclear why known transformation genes such as the recombination genes *rec-l* and *rec-2* (Beattie & Setlow, 1971) have not been isolated in subsequent screens such as the mini-Tn10 screen (Tomb et al., 1989) or the present study. Also missing from the present study were mutations in regulatory transformation genes such as *sxy* (Williams et al., 1994), also known as *TfoX* (Zulty & Barcak, 1995), which resides on a 3146 bp EcoRI fragment, and *crp* (Chandler, 1992), which was identified in the mini-Tn10 screen (Tomb et al., 1989) and resides on a 1820 bp EcoRI fragment. It is possible that these small EcoRI fragments might have been less efficiently taken up by competent cells and integrated into the genome. In the latter case, an EcoRI fragment containing an insertion into the *sxy* or *crp* genes would have less flanking DNA outside the Kan' cassette for integration. Optimization of mutagenesis procedures for a higher order of randomness is underway, by investigating alternative mutagenesis procedures such as *in vitro* transposition (Gwinn et al., 1997) and integration of plasmids with randomly sheared DNA (B. A. Dougherty, unpublished).

This study identified five loci involved in *H. influenzae* transformation (Table 1). Mutants in the *com* operon have been isolated in other screens for transformation genes (Tomb et al., 1989; Larson et al., 1991) and, like the *com* mutants in this study (Table 1), show background levels of DNA binding/uptake (Tomb et al., 1991). The *pbp* region may be involved in transformation if the cell wall poses an impediment to DNA transformation, and alteration in expression of *pbp* proteins appears to correlate with the ability to transform DNA (Dargis et al., 1992). The type IV pilin-like protein operon mutants have a loss of DNA-binding phenotype, which has been found in other bacteria where type IV pilin and pilin-like proteins have a demonstrated role in natural transformation (Chung & Dubnau, 1995; Rudel et al., 1995; Lunsford & Roble, 1997; Porstendorfer et al., 1997). This suggests that for all naturally transformable bacteria studied to date, type IV pilin-like proteins are a component of the DNA
receptor/DNA channel that allows for transfer of DNA from the environment to the host cytoplasm. The effect of the phospholipid biosynthesis and multidrug-efflux transport mutations on H. influenzae transformation rates is less clear. It is conceivable that alterations in membrane phospholipid content or the cell membrane structure itself will result in less efficient transformation. The multidrug-resistance efflux transporters are known to pump a broad range of toxic molecules out of the cell (Lewis, 1994; Nikaido, 1994), and mutations of multidrug-resistance genes result in increased sensitivity to these compounds in H. influenzae (Table 2; see also Sanchez et al., 1997). There are several possible mechanisms by which the BD307 deletion/mutation could result in the observed Tfo− phenotype. One possibility is that disruption of the drug-efflux system could affect the proper transport of the DNA-binding complex machinery to the cell surface during the early stages of transformation or the DNA transport process itself (i.e. the fusion of inner and outer membranes by AcrA may be required for proper DNA-channel formation and optimal DNA transport). It is clear that individual, nonpolar disruptions of the genes in each region will be necessary to provide the ultimate answer regarding the relative contribution of each gene to transformation.

The most important finding of this transformation screen is the involvement of type IV pilin-like proteins in H. influenzae transformation. Type IV pilin and pilin-like proteins are processed in a novel manner and are involved in functions ranging from cell adherence to protein export and DNA import (Strom & Lory, 1993; Dubnau, 1997). The pilA–D operon described in this paper should encode the necessary components (pilin subunit, pilin export genes and a prepilin peptidase) to process and export the PilA subunit. It appears that PilA is just one component of the H. influenzae DNA receptor/DNA channel, and further investigation is required to identify all the components and determine their functional interactions in the DNA transformation process.

ACKNOWLEDGEMENTS

B. A. D. is supported by an NIH Training Grant. H. O. S. is an American Cancer Society research professor. We thank Jean-Francois Tomb, Michelle Gwinn and Bill Bishai for valuable discussions, and Jean-Francois Tomb for critical reading of the manuscript.

REFERENCES


Lunsford, R. D & Robie, A. G. (1997). comYA, a gene similar to


Received 25 August 1998; revised 28 September 1998; accepted 8 October 1998.