JOURNAL OF BACTERIOLOGY, May 1997, p. 3122–3126 Vol. 179, No. 10
0021-9193/97/$04.00+0

comYA, a Gene Similar to *comGA* of *Bacillus subtilis*, Is Essential for Competence-Factor-Dependent DNA Transformation in *Streptococcus gordonii*

R. DWAYNE LUNSFORD* AND ARLENE G. ROBLE†

Oral Infection and Immunity Branch, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland

Received 24 December 1996/Accepted 14 March 1997

Tn*4001* mutagenesis identified a new competence gene in *Streptococcus gordonii* Challis designated *comYA*. A *comYA* mutant was completely deficient in transformation and exhibited decreased levels of DNA binding and hydrolysis. The deduced 319-amino-acid ComYA protein exhibited 57% similarity and 33% identity to the ComGA transporter protein of *Bacillus subtilis* and contained the Walker A-box motif conserved in ATP-binding proteins as well as aspartic acid boxes Asp-1 and Asp-2 present in some components of the general secretory pathway of gram-negative bacteria. *comYA* appeared to be part of a putative operon encompassing a *comGB* homolog, designated *comYB*, together with sequences that could encode ComGC- and ComGD-like peptides designated ComYC and ComYD, respectively, as well as other components. The putative ComYC and ComYD peptides had leader sequences similar to the type IV *N*-methylphenylalanine pilins of gram-negative bacteria, but unlike other examples in this class, including *B. subtilis*, they contained an alanine at position −1 of the leader instead of the usual glycine residue. Northern analysis identified a single 6.0-kb *comYA*-containing transcript strictly dependent on exogenous competence factor for expression in ComA1 cells. An identical pattern of expression was seen in wild-type Challis cells grown under conditions of maximal competence but not in cells that were noncompetent.

Oral streptococci play a critical role in the normal microbial ecology of the human oral cavity by functioning as primary colonizers of hard tissue surfaces (29). Members of this group also are the leading cause of subacute bacterial endocarditis and dental caries. Several species are proficient in natural genetic transformation, a competence-phase-specific ability to incorporate free exogenous DNA from the environment. Gram-positive transformable bacteria generally process DNA in a similar fashion. Competent cells bind native DNA in a sequence-independent manner. Following binding, one strand of the helix is nicked, the other is hydrolyzed via an entry nuclease, and the intact single strand is internalized, whereafter it can integrate into the host chromosome or, in the case of a plasmid molecule, establish itself extrachromosomally. Competence in *Streptococcus gordonii*, as well as other streptococci, is an early- to mid-growth-phase state in which nearly every cell in the population is transformable although exact periods of competence vary widely between strains and species. In contrast, *Bacillus subtilis* becomes competent only in late stationary phase and then with only a small subpopulation of cells being able to transform (21). Competence in *S. gordonii* Challis is modulated by unknown environmental stimuli that trigger the production of a secreted protein competence factor (CF) (13). Amplification of the initial stimulus at a critical cell density by CF, at least under in vitro conditions, assures complete competence induction in all cells of a particular culture via the induction of competence-specific gene expression (26). Whether similar levels of competence expression and kinetics of induction exist in the human oral cavity remains to be determined.

Challis CF was originally described as a trypsin-sensitive heat-resistant cationic peptide of small size (14). More recently, the synthetic operon for CF as well as the receptor sensing system has been reported (11). Challis CF was found to be a 19-amino-acid peptide with an estimated pI of 10.1. The molecule is processed from a 50-residue prepeptide with a 31-amino-acid double-glycine-type leader sequence. Previously, our laboratory has described a CF-independent mechanism for competence induction in *S. gordonii* Wicky as well as cell-surface-associated DNA binding and nicking activities in Challis (19, 20). In order to identify additional competence-related loci in this species, we initiated a Tn*4001* mutagenesis project with initial selection being for loss of transformability. Described here are the isolation and characterization of a gene in Challis similar to *comGA* of *B. subtilis* that is required for competence in *S. gordonii*. Biochemical characterization of the Challis mutant together with comparisons to the *B. subtilis* sequences strongly suggested that the gene, designated *comYA*, functions in some aspect of DNA internalization.

MATERIALS AND METHODS

Tn*4001* mutagenesis. All strains and plasmids utilized in this study are listed in Table 1. *S. gordonii* Challis-2 (Ch2) wild type was mutagenized as previously described by transformation with the Tn*4001* delivery vector pα (16). A total of 800 Gm^r Em^s clones were subjected to a 96-well plate microtransformation assay as follows. Clones were grown for 18 h in 100 μl of Todd-Hewitt broth (Difco Laboratories, Detroit, Mich.) containing 50 μg of gentamicin per ml. The cultures were diluted 1:100 into drug-free transforming broth containing 5 μg of Sm^r DNA from strain Ch5 per ml (16). Following incubation for 6 h, the cultures were diluted 1:100 into 100 μl of brain heart infusion (Difco) containing 500 μg of streptomycin per ml. Absence of or decreased Sm^r growth in each well was scored after 16 h of incubation. Potential mutants exhibiting transformation defects were rescreened in a standard quantitative tube transformation assay (16).

Mutant characterization. Clones exhibiting defective transformation properties with genomic DNA were tested for the following phenotypes: (i) the ability

* Corresponding author. Present address: Antiinfectives Research UP1345, SmithKline Beecham Pharmaceuticals, 1250 S. Collegeville Rd., P.O. Box 5089, Collegeville, PA 19426-0989. Phone: (610) 917-6512. E-mail: Dwayne_Lunsford-1@sbphrd.com.

† Present address: St. George's University School of Medicine, St. George's, Grenada, West Indies.

TABLE 1. Strains and plasmids

Strain or plasmid	Genotype and/or phenotype of strain or plasmid[a]	Reference or supplier
Spp. and strains		
S. gordonii		
Ch2	Challis Rf^r Xfo^+	20
Ch5	Challis Sm^r Xfo^+	20
CipA11	Challis Rec^-	27
ComA1	Ch2 *comA*::*ermAM* Xfo^- CF^-	20
m3D3	*comD*ΩTn*4001*	
m8G2	*comYA*ΩTn*4001*	
E. coli DH10B	F^- *mcrA* Δ(*mrr hsdRMS mcrBC*) *endA recA*	GIBCO-BRL
Plasmids		
pΩ	Em^r insertion-duplication vector	16
pα	Tn*4001* in pΩ	16
pCC24	IS*256* 1.1-kbp *Bam*HI-*Hin*dIII fragment	16
pCC93	pΩ with the m8G2 *Eco*RI-*Bam*HI junction fragment	
f2H7	*comD*-specific fosmid	
f1D5	*comY*-specific fosmid	
pDL277	Sp^r	D. LeBlanc

[a] Abbreviations: Rf^r, Sm^r, Em^r, Ap^r, and Sp^r, resistant to rifampin, streptomycin, erythromycin, ampicillin, and spectinomycin, respectively; Xfo^+, transformation positive; Xfo^-, transformation negative. Entries lacking reference or source designations were either laboratory stocks or strains or plasmids constructed in this study. Antibiotic concentrations used: for *S. gordonii*, 10 μg of erythromycin, 500 μg of streptomycin, 100 μg of rifampin, and 800 μg of spectinomycin per ml; for *E. coli*, 300 μg of erythromycin, 50 μg of spectinomycin, and 100 μg of ampicillin per ml.

to transform with plasmid DNA, (ii) complementation with exogenous CF, (iii) CF production, (iv) UV light sensitivity, and (v) [^{3}H]DNA uptake and solubilization. Plasmid transformation was tested in the standard tube assay with pDL277 and selection for spectinomycin resistance. CF complementation was tested with filter-sterilized heat-inactivated Ch2-conditioned medium as described elsewhere (20). CF production was assayed on the nonproducing strain ComA1 (20). UV light sensitivity was tested with a model 1800 UV Stratalinker (Stratagene, La Jolla, Calif.). Clones were streaked across brain heart infusion plates together with Ch2 and the recombination-deficient strain CipA11 (27) as controls, and one-half of the plate was exposed to 3,000 μJ of 254-nm light per cm^2. Growth was scored after 24 h of incubation in the dark. [^{3}H]DNA uptake and solubilization activities were assayed essentially as described elsewhere (9, 15) with 4-h cultures in competence medium with the following changes: test DNA was uniformly ^{3}H-labeled *Escherichia coli* genomic DNA (2×10^7 dpm/μg; DuPont-NEN, Boston, Mass.), and the solubilization assay was processed in microcentrifuge tubes. Data were expressed as percentages of wild-type uptake or solubilization.

The characteristics of mutant m8G2 were as follows: defective in transformation for genomic and plasmid DNA, negative for CF stimulation, positive for CF production, resistant to UV light, and having 52% of the DNA uptake and 34% of the solubilization activity of the wild type.

Tn*4001* insertion site characterization. Genomic DNA was isolated from mutants as described elsewhere (18) and subjected to filter hybridization analysis with the IS*256*-specific probe pCC24 (16). Tn*4001* junction fragments were recovered by inverse PCR (IPCR) (22). Briefly, genomic DNA was digested to completion with *Mbo*I and ligated at a low DNA concentration (0.5 μg/ml) to optimize intramolecular reactions. The resulting material served as template in IPCRs with *Pfu* polymerase (Stratagene) performed with the following primer pair: IPCR-L, 5′gcgaagcttgaattCGCCTGATACATACATTTCC3′ (3′ base position 425 of the sequence in reference 5); IPCR-R, 5′gcgtctagaggatccATTGT GGAAGAACTTTGTGG3′ (position 487 of the sequence in reference 5). Lowercase bases for all primers described here represent nonhomologous thermal clamps and restriction sites for subsequent molecular cloning. Reaction conditions were as follows: 1 min, 94°C; 0.5 min, 50°C; and 1.5 min, 72°C, for 30 cycles. Reaction products were digested with *Bam*HI and *Eco*RI and cloned directly into identically cleaved insertion-duplication vector pΩ (16). *E. coli* DH10B (GIBCO-BRL, Gaithersburg, Md.) served exclusively as the cloning host in all manipulations. Plasmids capable of both transforming Ch2 to Em^r and regenerating the original Tn*4001* mutation were retained for further analysis. Concurrently, IPCR inserts were utilized directly as probes to isolate genomic clones containing wild-type loci from a fosmid library (20). m8G2 yielded pCC93 and f1D5. Also isolated was the m3D3-specific fosmid f2H7.

DNA sequence analysis. IPCR products and fosmid clones were utilized directly as sequencing templates (20). The sequence was assembled with AssemblyLIGN and analyzed with MacVector (International Biochemicals, Inc., New Haven, Conn.). Database searches were conducted with the BLAST network system (National Center for Biotechnology Information-National Library of Medicine, Bethesda, Md.). Protein alignments were generated with GAP (Wisconsin Package, version 8.0 for VMS-VAX; Genetics Computer Group, Madison, Wis.).

CF synthesis. Challis CF propeptide was synthesized on a Synergy 432A personal peptide synthesizer (Applied Biosystems, Foster City, Calif.) with fast-moc chemistry and a sequence based on the ComC theoretical peptide assuming that cleavage of the leader sequence occurred just C terminal to the double glycines at positions 30 and 31 (deduced sequence from U80077). The product was purified by C_{18} hydrophobic-affinity high-pressure liquid chromatography, assayed on ComA1, and stored desiccated at 4°C. Aliquots were prepared in phosphate-buffered saline immediately before use. Activity for competence induction was maximal at 1 ng/ml.

Northern hybridization analysis. Ten-microgram samples of total RNA extracted from 4- or 18-h Ch2 cultures and ComA1 cells induced with either pure CF (100 ng/ml) or buffer controls in competence medium were isolated and subjected to Northern analysis as described elsewhere (17). A *comYA* probe was synthesized as a PCR product and corresponded to bp 2268 to 3025 of U81957 synthesized with 5′cgcgaattCAGCTAGCCGCTGTGATTAG3′ and 5′cgcggatc cGGATGGATGGATATGTCCTGCTGC3′. The fragment was isolated by gel electrophoresis in low-melting-point agarose with NACS affinity resin (GIBCO-BRL), labeled with ^{32}P by nick translation (Lofstrand Laboratories Ltd., Gaithersburg, Md.), and utilized at 10^5 dpm/ml in Quick-Hyb solution (Stratagene) at 65°C. Washes were at high stringency (0.1 × SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]–0.1% sodium dodecyl sulfate, 60°C).

Nucleotide sequence accession number. The sequences were submitted to GenBank under accession no. U81957 (*comY* region of f1D5) and U80077 (*comD* region of f2H7).

RESULTS

Generation of transposon mutants and phenotype characterization. Ch2 was mutagenized with Tn*4001* to yield 800 Gm^r Em^s clones that were screened in a microtransformation assay. Of this population, 16 were found to be defective in competence. Southern analysis of genomic DNA with an IS*256*-specific probe indicated that 14 members of this population contained a common insertion site in the *comD* locus represented by m3D3 (reference 11 and data not shown). We previously have shown that Tn*4001* insertion is fairly random after screening transformants solely for Gm^r (16). However, in the context of competence mutants at least, there is clearly a hot spot for insertion of this element into *comD*. The two remaining mutants contained unique insertion sites, with one, m8G2, being the focus of this study. The remaining mutant appeared to involve a regulatory locus and will be described elsewhere. m8G2 was screened for the ability to transform with both genomic and plasmid DNA, production of and stimulation by CF, sensitivity to UV light, and [^{3}H]DNA uptake and solubilization. Cells were found to be completely defective for transformation with either genomic or plasmid DNA and were resistant to UV light, also suggesting that the defect did not involve conventional recombination pathways (27). m8G2 was proficient in CF production but was clearly defective in DNA uptake and processing.

Insertion site characterization. Sequence information flanking transposon insertion sites was generated by IPCR. Genomic DNA from each mutant was digested with *Mbo*I and self-ligated under dilution conditions that enriched the DNA for intramolecular circularization. This material was used as template for PCRs with outward-oriented primers based on the IS*256* sequence (5). Reaction products were cloned directly into pΩ to generate targeting vectors for insertion-duplication mutagenesis with which to confirm that the original mutants were indeed due to Tn*4001* insertions (16). As expected, pCC93 regenerated the original transformation-defective phenotype of m8G2. The insertion site was found to reside within

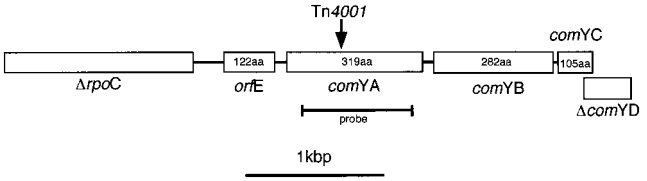

FIG. 1. Genetic structure of the *comY* locus. Rectangles represent ORFs with expression from left to right. Complete ORFs contain the number of amino acid residues in the deduced peptides derived from the DNA sequence. Thick lines, not to scale, represent intervening sequence. Incompletely sequenced ORFs are denoted by Δ. A downward arrow denotes the insertion site of Tn*4001* in the original mutation. The line with bars represents the regions derived by PCR for use as probe in Northern analysis.

an open reading frame (ORF) with similarity to the *comGA* gene of *B. subtilis* (1) (Fig. 1). Insertion was precisely between nucleotides 2620 and 2621 of our sequence. We have designated this putative gene in *S. gordonii comYA*. A total of 4,566 bp of DNA sequence was determined in the region, representing all or portions of six major ORFs (GenBank accession no. U81957). The first, Δ*rpoC*, represented the far C-terminal portion of the β′ subunit of RNA polymerase core enzyme (2). Although the region exhibited significant similarity to several bacterial and eucaryotic DNA-directed RNA polymerases, the highest match was to the *B. subtilis* enzyme, with 82% similarity and 73% identity. Spontaneous rifampin resistance generally maps as point mutations to the β subunit of RNA polymerase. In *B. subtilis*, the β subunit ORF lies just upstream of β′. A similar structure probably exists in *S. gordonii*, since we could demonstrate a high rate of cotransformation between Rf^r and Em^r by using genomic DNA from Ch2 (Rf^r) clones transformed with pCC93 (13.5% cotransformation linkage compared to 0.4% linkage for negative control DNA). *orfE* could potentially encode a 122-amino-acid peptide with some similarity to DNA-binding proteins, specifically a 40-amino-acid region of the URE-B1 protein of rat brain (10) and the entire 167-amino-acid LktC activator protein from the leukotoxin operon of *Pasteurella haemolytica* (4). In the latter case, overall similarity was 39%, and identity was 22%. The role of this ORF in transformation is under investigation, and results will be presented elsewhere.

The putative *comY* operon is currently composed of four ORFs designated *comYA* to Δ*comYD*. If similarity to the *comG* locus is maintained downstream of our sequenced region, the operon will probably contain additional coding regions. ComYA is the ComGA homolog, and the putative peptides exhibited 57% similarity and 33% identity (Fig. 2). ComYB and ComGB were 55% similar and 26% identical (Fig. 3). ComGA and ComGB probably constitute a transport system with high similarity to components of the general secretory pathway. ComYA was found to contain only the A box of the Walker motif found in ATP-binding proteins (positions 138 to 146) (Fig. 2, open boxes) (28), together with aspartic acid boxes Asp-1 (positions 163 to 169) and Asp-2 (positions 196 to 210) (Fig. 2, shaded boxes) conserved in some secretory pathway components (24). ComGC, ComGD, and ComGE are most likely the substrates for this secretion system. These peptides contain leader sequences with high similarity to the type IV *N*-methylphenylalanine-type pilins as well as some secretory system components of gram-negative bacteria (12). ComYC-ΔComYD and ComGC-ComGD alignments are shown in Fig. 4. The type IV prepilins are composed of peptides generally having short leader sequences with a basic charge. Cleavage by a leader peptidase, specifically the ComC protease in *B. subtilis* (7), occurs on the C-terminal side of a conserved

```
  1 ..MVQKIAQAIVRQAKEECAQDIYFVPKDDCYELYMRIGDERRFIQTYDF ComYA
    ..:|:...:: :|  . | ||.:||::    :.:|:::.  :::. |:
  1 MDSIEKVSKNLIEEAYLTKASDIHIVPRERDAIIHFRVDHA..LLKKRDM ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
 49 DQLAAV..ISHFKFLAGMNVGEKRRSQLGSCDYRYDDKETSIRLSTVGDY ComYA
    .. ..|  |||||||.:|::||:|:.| || ..:.:: :. :|:||:...
 49 KKEECVRLISHFKFLSAMDIGERRKPQNGSLTLKLKEGNVHLRMSTLPTI ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
 97 RGYESLVIRLLHDEETELKFWFTHFPELREKF.....KDRGLYLFSGPVG ComYA
    .: |||||::.: :..  .:. ||.  ..:     ..:|:.:.|.|||
 99 NE.ESLVIRVMPQYNIPSIDKLSLFPKTGATLLSFLKHSHGMLIFTGPTG ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
142 SGKTTLMHQLAQLKFK..GQQVMSIEDPVEIKQEDMLQLQLNETIGLTYE ComYA
    ||||| :. |.|.  |  ...:.:.|||||.::||:|||:|:||. |:||.
148 SGKTTTLYSLVQYAKKHFNRNIVTLEDPVETRDEDVLQVQVNEKAGVTYS ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
190 SLIKLSLRHRPDLLIIGEIRDSETARAVVRASLTGATVFSTIHAKSIPGV ComYA
    . :| |||| ||::|:||||.||  .|||.:||  |:..:|.:. .|.
198 AGLKAILRHDPDMIILGEIRDAETAEIAVRAAMTGHLVLTSLHTRDAKGA ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
240 YERLLELGVSEEELKIVLQGICYQRLIGGGGVI..DFASDNYQEHEPTVW ComYA
    . ||||:|:. :|:. .: :|. |||:: :.:  : .|. |  :....:
248 IYRLLEFGINMNEIEQTVIAIAAQRLVDLACPFCENGCSSVYCRQSRNTR ComGA
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
288 NQQIDQLLAAGHIHPEQAEAEKIRNQQAKTSY..................... ComYA
    .. : :|| : :::.  .||  :..:|.  |
298 RASVYELLYGKNLQQCIQEA...KGNHANYQYQTLRQIIRKGIALGYLTT ComGA
```

FIG. 2. GAP alignment of ComYA and ComGA. Solid bars represent exact sequence conservation while colons and single dots represent decreasing levels of functional conservation corresponding to GAP comparison values greater than or equal to 0.50 or 0.10, respectively. The open box denotes the Walker A motif. The shaded boxes are Asp-1 and Asp-2, respectively.

glycine residue at position −1 to produce a mature pilin-like protein with an N-terminal phenylalanine at position +1 followed by a hydrophobic stretch of 20 to 30 residues that probably specify membrane localization. In fact, leader cleavage is required for subsequent membrane insertion (7). Sequence similarity then diverges distal of these conserved N-terminal domains. As expected, alignments of ComYC and ΔComYD with ComGC and ComGD indicated that the N-terminal regions were generally similar (Fig. 4). The streptococcal leaders were longer, however (11 and 15 amino acids compared to 5 and 10 for the *B. subtilis* proteins, respectively), but contained the conserved lysine residue at −2. Most surprising was the presence of an alanine residue at position −1 in both ComYC and ΔComYD. The glycine residue normally present at position −1 is conserved in every known example of a prepilin or prepilin-like protein previously examined (12).

RNA analysis. Total RNA was isolated from cultures of the CF-defective mutant ComA1 and wild-type strain Ch2 that had

```
  1 ....................MRQG.........LANGQAFSEIMASLGFS ComYB
                        .||:          | :|.:| ::: ||:|
  1 MTAGGYTLLDGLRLMELQMNKRQAADLTDSVTCLREGAPFYQVLKSLSFH ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
 22 DAVVTQLSLAELHGNLSLALLKIEEYLDNLAKVKKKLIEVATYPMMLLGF ComYB
    ...|.   :|| ||:|. .::. :|.|:.     ..|  |  ||::|:
 51 KEAVGICYFAETHGELPASMIQSGELLERKIAQADQLKRVLRYPLFLIFT ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
 72 LVLIMIGLRNYLLPQLS....SQNFAT........QLIGHLPTIFLLTVL ComYB
    :.:::. |...::||:|    | |:.|        .::.|:. :::| ||
101 VAVMFYMLQSIIIPQFSGIYQSMNMETSRSTDMLFAFFQHIDLVIILLVL ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
110 MLLGLTGAIYLVFKGQKRIPVYSFLARLPFVGSFVRIYLTAYYAREWGNM ComYB
    :  |:.  .:|||| ... .  :  |:|:||.:|::: . :: :::.:
151 FTAGIGIYYWLVFKKKSPARQMLICIRIPLVGKLVKLFNSYFFSLQLSSL ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
160 IGQGLELSQIFQIMQEQRSV.LFQEIGQDLGQALQNGQEFSDKIASYPFF ComYB
    :  ||.: : :. :..|  : ::.  :::| : |..|:.:.. |.:  |:
201 LKSGLSIYDSLNAFKHQTFLPFYRCEAEQLIERLKAGESIESAICGSLFY ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
209 KKELSLIIEYGEVKSKLGSELEIYALKTWEEFFGRVNRTMNLIQPLVFVF ComYB
    ..:|| :|..|::.::|:.|| .|.  .:: :  :..:  .::||:::.|
251 ETDLSKVISHGQLSGRLDRELFTYSQFILQRLEHKAQKWTGILQPMIYGF ComGB
    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
259 VALMIVLLYAAMLLPLYQNMEVHL ComYB
    || ||:|:| .||:|:|| |:
301 VAAMILLVYLSMLVPMYQMMNQM. ComGB
```

FIG. 3. GAP alignment of ComYB and ComGB. Essential features are as described in the legend to Fig. 2.

FIG. 4. GAP alignment of ComYC and ΔComYD with ComGC and ComGD. (A) ComYC-ComGC alignment. (B) ΔComYD-ComGD alignment. Essential features are as described in the legend to Fig. 2. The alanine or glycine residues present at position −1 of the deduced leaders are shown in boldface with the downward arrows indicating the probable point of cleavage.

either been induced with pure CF or grown for periods of time in which competence is either maximal or nonexistent in our culture system. Results of Northern analysis of the samples are shown in Fig. 5. Each panel represents hybridization with a *comYA*-specific probe. Figure 5A contains RNA from ComA1 cells pulsed with CF (lane 1) or buffer control (lane 2). Figure 5B contains RNA from Ch2 cells at maximum competence (lane 3) or noncompetent 18-h cells (lane 4). Clearly evident is the strict dependence on exogenous CF for the expression of a single 6.0-kb transcript containing *comYA* sequences. Furthermore, temporal expression of this transcript followed the competent state of the cultures. Four-hour cells were maximally competent, with the production of a transcript, and 18-h cells were noncompetent, with essentially no detectable transcription.

DISCUSSION

The only transformation mutant previously described for *S. gordonii* is the chemically derived recombination-defective strain CipA11 (27). This strain, while being sensitive to UV light and methylating agents, is impaired for chromosomal transformation but still capable of reduced levels of transformation with plasmid DNA. For this reason, conventional transposon mutagenesis was utilized to generate a collection of transformation-defective strains with which to identify new genes involved in the development of competence. Traditionally, Tn*916* has been utilized for mutagenesis in the viridans streptococci. Tn*916* is an *Enterococcus faecalis* conjugative transposon capable of encoding both transposition and DNA mobilization functions (8). We decided to utilize the *Staphylococcus aureus* composite-type transposon Tn*4001* instead because of its small size, stability, and lack of any DNA mobilization machinery that could complicate analysis of transformation functions in host cells. Tn*4001* is a 4.5-kbp element composed of a central resistance gene encoding a single protein with both aminoglycoside acetyltransferase and phosphotransferase activities flanked by dual copies of IS*256* (16). Our present Tn*4001* insertion collection represents several new classes of mutants defective in competence regulation, CF production, and DNA processing. The *comYA* mutant described here was unique in that it involved the first streptococcal competence locus with significant structural and sequence similarity to a *B. subtilis* operon essential for transformation.

The putative *comY* operon currently consists of four reading frames with significant similarity to the *B. subtilis comG* locus (1). Additional *comG* homologs are expected in Challis as DNA sequence analysis progresses in this region. The Tn*4001* insertion in m8G2 was found to reside within *comYA*. ComYA and ComYB are similar to ComGA and ComGB. ComYA possesses signature domains indicative of a protein involved with secretion. A single 6.0-kb transcript containing *comYA* sequences is clearly induced by CF in the nonproducing mutant ComA1. Similar inducibility was noted in wild-type cultures grown under natural conditions of either competence or noncompetence. We also have noticed strict CF-dependent synthesis of a 2.4-kb *comD*-containing transcript in ComA1 cells as well as in maximally competent Ch2. No obvious message was detected in buffer-pulsed ComA1 or in noncompetent Ch2 cells (data not shown). These observations for *S. gordonii* are in sharp contrast to studies with *Streptococcus pneumoniae* indicating either partial inducibility (23) or constitutive expression (6) of the *comCDE* locus.

The ComGC, ComGD, and ComGE proteins of *B. subtilis* represent small peptides of 98, 143, and 115 amino acids, respectively, that contain leader sequences of high similarity to those present on several gram-negative type IV prepilins and components of protein secretion systems (12). ComYC and ΔComYD are similar to ComGC and ComGD, respectively. ComGC is proteolytically processed by the ComC leader peptidase and subsequently translocated to the outer surface of the *B. subtilis* cytoplasmic membrane via an unknown mechanism (7). This step occurs independent of ComGA and ComGB, but optimal processing appears to require ComGA. The ComGA and ComGB complex may function to eject ComGC dimers from the outer membrane surface. We predict that *S. gordonii* also will be found to contain a ComC-like peptidase activity. ComYC and ΔComYD are unique in that a conserved glycine residue present at position −1 of the leader in all other examples of prepilins and prepilin-like molecules is replaced with an alanine.

If ComYC and ComYD are processed in a similar fashion as ComGC and ComGD, these molecules would be available to associate with other ComY components into extracellular structures required for DNA uptake. It has been suggested that

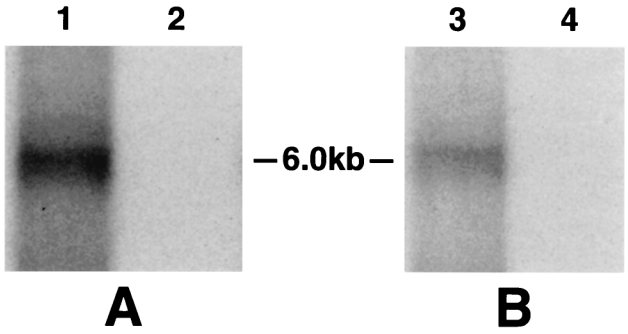

FIG. 5. Northern analysis of the *comY* locus. Ten-microgram samples of total cellular RNA isolated from either ComA1 (A) or Ch2 (B) induced either with CF (lane 1) or without CF (lane 2) or grown to maximum competence (lane 3) or no competence (lane 4). The 6.0-kb transcript is indicated.

these small peptides combine to form a pseudopilus-like structure that functions as an aqueous pore or tube spanning the cell wall and membrane to function as a direct route for DNA entry (12). In fact, ComGC mutants of *B. subtilis* are completely defective in competence and fail to bind DNA above background levels seen with noncompetent cells (3). However, the situation appears to be different in streptococci. We have previously described Mg^{2+}-independent DNA nicking and binding activities in surface extracts of both competent and noncompetent Challis cultures (19). Levels of these activities appear to be the same regardless of the competence state of the cells. Southwestern blot analysis of the preparations identified a single 56-kDa DNA-binding protein present in both competent and noncompetent cells that could function as a primary DNA receptor. In *S. pneumoniae*, the 30-kDa cell wall endonuclease EndA has been shown to be essential for competence (25). EndA also was found to be constitutively expressed. It would appear, therefore, that at least some components of the DNA recognition and uptake machinery in streptococci are not directly inducible by CF and that these components could associate with inducible subunits to produce a complete DNA uptake apparatus. It is tempting to speculate that the ComY structure may have a dual function: first, to provide the actual pathway for DNA transport by supplying a channel capable of spanning both the cell membrane and the relatively thick streptococcal cell wall; and second, to act as a scaffolding structure able to sequester constitutively expressed components such as the primary DNA receptor and nickase from the membrane and project them through the cell wall toward the extracellular milieu. Studies to determine the exact composition and topology of the *S. gordonii* DNA uptake system are in progress.

ACKNOWLEDGMENTS

Frank Robey is gratefully acknowledged for his assistance with high-pressure liquid chromatography purification of the CF peptide. Myra Jacobs and Jack London provided helpful comments on the manuscript.

REFERENCES

1. **Albano, M., R. Breitling, and D. A. Dubnau.** 1989. Nucleotide sequence and genetic organization of the *Bacillus subtilis comG* operon. J. Bacteriol. **171:** 5386–5404.
2. **Boor, K. J., M. L. Duncan, and C. W. Price.** 1995. Genetic and transcriptional organization of the region encoding the β subunit of *Bacillus subtilis* RNA polymerase. J. Biol. Chem. **270:**20329–20336.
3. **Breitling, R., and D. Dubnau.** 1990. A membrane protein with similarity to the *N*-methylphenylalanine pilins is essential for DNA binding by competent *Bacillus subtilis*. J. Bacteriol. **172:**1499–1508.
4. **Burrows, L. L., E. Olah-Winfield, and R. Y. Lo.** 1993. Molecular analysis of the leukotoxin determinants from *Pasteurella haemolytica* serotypes 1 to 16. Infect. Immun. **61:**5001–5007.
5. **Byrne, M. E., D. A. Rouch, and R. A. Skurry.** 1989. Nucleotide sequence analysis of IS*256* from the *Staphylococcus aureus* gentamicin-tobramycin-kanamycin-resistance transposon Tn*4001*. Gene **81:**361–367.
6. **Cheng, Q., E. A. Campbell, A. M. Naughton, S. Johnson, and H. R. Masure.** 1997. The *com* locus controls genetic transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **23:**683–692.
7. **Chung, Y. S., and D. Dubnau.** 1995. ComC is required for the processing and translocation of ComGC, a pillin-like competence protein of *Bacillus subtilis*. Mol. Microbiol. **15:**543–551.
8. **Clewell, D. B., S. E. Flannagan, and D. D. Jaworski.** 1995. Unconstrained bacterial promiscuity: the Tn*916*-Tn*1545* family of conjugative transposons. Trends Microbiol. **3:**229–236.
9. **Deddish, P., and H. D. Slade.** 1971. Binding of deoxyribonucleic acid by cell walls of transformable and nontransformable streptococci. J. Bacteriol. **105:** 779–786.
10. **Gu, J., K. Ren, R. Dubner, and M. J. Iadarola.** 1994. Cloning of a DNA binding protein that is a tyrosine kinase substrate and recognizes an upstream initiator-like sequence in the promoter of the preprodynorphine gene. Brain Res. Mol. Brain Res. **24:**77–88.
11. **Havarstein, L. S., P. Gaustad, I. F. Nes, and D. A. Morrison.** 1996. Identification of the streptococcal competence pheromone receptor. Mol. Microbiol. **21:**863–869.
12. **Hobbs, M., and J. S. Mattick.** 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentious phage, and protein secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol. Microbiol. **10:**233–243.
13. **Leonard, C. G.** 1973. Early events in development of streptococcal competence. J. Bacteriol. **114:**1198–1205.
14. **Leonard, C. G., and R. M. Cole.** 1972. Purification and properties of streptococcal competence factor isolated from chemically defined medium. J. Bacteriol. **110:**273–280.
15. **Lindler, L. E., and F. L. Macrina.** 1986. Characterization of genetic transformation in *Streptococcus mutans* by using a novel high-efficiency plasmid marker rescue system. J. Bacteriol. **166:**658–665.
16. **Lunsford, R. D.** 1995. A Tn*4001* delivery system for *Streptococcus gordonii* (Challis). Plasmid **33:**153–157.
17. **Lunsford, R. D.** 1995. Recovery of RNA from oral streptococci. BioTechniques **18:**412–413.
18. **Lunsford, R. D., and F. L. Macrina.** 1986. Molecular cloning and characterization of *scrB*, the structural gene for the *Streptococcus mutans* phosphoenolpyruvate-dependent sucrose phosphotransferase system sucrose 6-phosphate hydrolase. J. Bacteriol. **166:**426–434.
19. **Lunsford, R. D., N. Nguyen, and J. London.** 1996. DNA binding activities in *Streptococcus gordonii*: identification of a receptor-nickase and a histone-like protein. Curr. Microbiol. **32:**95–100.
20. **Lunsford, R. D., and J. London.** 1996. Natural genetic transformation in *Streptococcus gordonii*: *comX* imparts spontaneous competence on strain Wicky. J. Bacteriol. **178:**5831–5835.
21. **Magnuson, R., J. Solomon, and A. D. Grossman.** 1994. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. Cell **77:**207–216.
22. **Ochman, H., M. M. Medhora, D. Garza, and D. L. Hartl.** 1990. Amplification of flanking sequences by inverse PCR, p. 219–227. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, Inc., San Diego, Calif.
23. **Pestova, E. V., L. S. Havarstein, and D. A. Morrison.** 1996. Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. Mol. Microbiol. **21:**853–862.
24. **Possot, O., and A. P. Pugsley.** 1994. Molecular characterization of PulE, a protein required for pullulanase secretion. Mol. Microbiol. **12:**287–299.
25. **Puyett, A., B. Greenberg, and S. A. Lacks.** 1990. Genetic and structural characterization of EndA, a membrane-bound nuclease required for transformation of *Streptococcus pneumoniae*. J. Mol. Biol. **213:**727–738.
26. **Raina, J. L., and A. W. Ravin.** 1980. Switches in macromolecular synthesis during induction of competence for transformation in *S. sanguis*. Proc. Natl. Acad. Sci. USA **77:**6062–6066.
27. **Raina, J. L., and F. L. Macrina.** 1982. A competence specific inducible protein promotes *in vivo* recombination in *Streptococcus sanguis*. Mol. Gen. Genet. **185:**21–29.
28. **Walker, J. E., M. Sarste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the α and β subunits of ATP synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
29. **Whittaker, C. J., C. M. Klier, and P. E. Kolenbrander.** 1996. Mechanisms of adhesion by oral bacteria. Annu. Rev. Microbiol. **50:**513–552.