Identification of a New Regulator in *Streptococcus pneumoniae* Linking Quorum Sensing to Competence for Genetic Transformation

MYEONG S. LEE AND DONALD A. MORRISON*

Laboratory for Molecular Biology, Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois 60607

Competence for genetic transformation in *Streptococcus pneumoniae* is regulated by a quorum-sensing system encoded by two genetic loci, *comCDE* and *comAB*. Additional competence-specific operons, *cilA*, *cilB*, *cilC*, *cilD*, *cilE*, *cinA-recA*, *coiA*, and *cfl*, involved in the DNA uptake process and recombination, share an unusual consensus sequence at −10 and −25 in the promoter, which is absent from the promoters of *comAB* and *comCDE*. This pattern suggests that a factor regulating transcription of these transformation machinery genes but not involved with *comCDE* and *comAB* expression might be an alternative sigma factor. A search for such a global transcriptional regulator was begun by purifying pneumococcal RNA polymerase holoenzyme. In preparations from competent pneumococcal cultures a protein which seemed to be responsible for *cilA* transcription in vitro was identified. The corresponding gene was identified and found to be present in two copies, designated *comX1* and *comX2*, located adjacent to two of the repeated rRNA operons. Expression of transformation machinery operons, such as *cilA*, *cilD*, *cilE*, and *cfl*, but not that of the quorum-sensing operons *comAB* and *comCDE*, was shown to depend on *comX*, while *comX* expression depended on ComE but not on ComX itself. We conclude that the factor is a competence-specific global transcription modulator which links quorum-sensing information transduced to ComE to competence and propose that it acts as an alternate sigma factor. We also report that *comAB* and *comCDE* are not sufficient for shutoff of competence-stimulating peptide-induced gene expression nor for the subsequent refractory period, suggesting that these phenomena depend on one or more ComX-dependent genes.

Streptococcus pneumoniae (pneumococcus) is naturally competent for genetic transformation. Competence, the state of cells able to take up DNA, develops suddenly in response to a cell-cell signal at some point during exponential growth phase, reaches a maximum about 20 min after its induction (induction phase), disappears abruptly (shutoff phase), and remains off for about 40 to 60 min during a period in which cells are refractory to the signal (7, 10, 20, 34, 48). During the period of competence, double-stranded DNA encountered by pneumococcus cells is bound to the cell surface and one strand of the DNA is degraded into short oligonucleotides in the medium while part of the other strand is imported inside the cell (29). The imported single strand of DNA is protected from nuclease activity by a single-stranded DNA binding protein (SSB) (32) and is finally incorporated into the chromosome by homologous recombination involving general recombination machinery, such as RecA (35). The quorum-sensing signal responsible for competence induction is a heptadecapeptide, named CSP (competence-stimulating peptide) (17), which derives from a precursor (ComC) by cleavage and transport into the medium by an ATP-binding cassette (ABC) transporter, ComAB (54). CSP acts through a receptor (ComD) and a response regulator (ComE) to activate both *comAB* and *comCDE* operons (39), establishing a positive feedback loop ensuring an abrupt rise in CSP levels, making all cells in a culture competent simultaneously. But how this quorum-sensing circuit (*comCDE* and *comAB*) evokes the expression of the genes for the machinery

of genetic transformation is not understood, although it has been proposed that ComE (the quorum-sensing transducer) might act as a transcription factor to induce both the competence machinery genes and those of the CSP circuit (4). How competence is shut off after induction is also unknown, although it has been proposed that ComE has dual functions—activation at low doses of the CSP stimulus and repression at high doses—for the regulation of *comCDE*, which could account for the successive induction and suppression of competence in response to increasing levels of CSP (4).

Recently several competence-specific operons which are probably involved with the DNA uptake process and recombination were identified, such as *cilA* (*ssb2*, a gene for SSB), *cilB* (*dal*, like *dprA* in *Haemophilus influenzae* (22), *cilC* (*ccl*, like *comC* in *Bacillus subtilis*), *cilD* (*cglABCDE*), *cilE* (*celAB*), and *coi* (5, 40). These operons and additional competence-related operons, such as *cinA-recA* and *cfl* (like *comF* in *B. subtilis*) (24, 25), all contain an unusual perfectly conserved consensus sequence, TACGAATA (cin-box), at position −10 from the transcription start and a T-rich region at −25 (5). Campbell et al. showed that this consensus sequence is important for the transcription of the *cinA-recA* operon by measuring the activities of mutated promoters in vivo (5). The fact that this consensus sequence is not present in the promoters of *comCDE* and *comAB* suggests that a factor regulating transcription of these transformation machinery genes is different from the factor (ComE, apparently) responsible for *comCDE* regulation. Such a factor for expression of transformation machinery genes could be an alternative sigma factor, since the consensus sequences overlap the RNA polymerase holoenzyme binding site (15). Therefore we sought the putative global transcriptional modulator linking quorum-sensing information to the

---

* Corresponding author. Mailing address: Laboratory for Molecular Biology, Room 4110 MBR, 900 S. Ashland Avenue, Chicago, IL 60607. Phone: (312) 996-6839. Fax: (312) 413-2691. E-mail: DAMorris@uic.edu.

TABLE 1. Pneumococcal strains and DNA constructs

Strain, plasmid or amplicon	Relevant characteristics	Description	Source[a] or reference
Strains			
CP1250	*hex malM511 str-1 bgl-1*	Rx derivative, low β-galactosidase background	39
CP1500	*hex nov-r1 bry-r str-1 ery-r1 ery-r2*	Donor of point markers	6
CPM1	CP1250 *rpoC::C-his-10*::pEVP3	10× His tagged at C terminus of β′ of RNA polymerase with pEVP3; Cm^r	CP1250 × pMSL1
CPM2	CP1250 Δ*comX1*::PcEm	Insertion-deletion replacement of *comX1* with synthetic amplicon, aMSL2, Em^r	CP1250 × aMSL2
CPM3	CP1250 *comX1′*::(pEVP3)::*comX1*^+	*lacZ* fusion to *comX1* by using pEVP3, Cm^r ComX1^+	CP1250 × pMSL2
CPM4	CPM2 *comX2′*::(pEVP3)::′*comX2*	*lacZ* fusion to *comX2*, Δ*comX1*, Δ*comX2*, Cm^r Em^r	CPM2 × pMSL3
CPM5	CP1250 Δ*comX2*::PcTet	Insertion-deletion of *comX2* with synthetic amplicon, aMSL5, Tet^r	CP1250 × aMSL5
CPM6	CP1250 *comC′*::(pEVP3)::*comC*^+	*lacZ* fusion to *comC* by using pEVP3, Cm^r ComC^+	CP1250 × pXF520
CPM7	CP1250 *ssb2′*::(pEVP3)::*ssb2*^+	*lacZ* fusion to *ssb2* by using pEVP3, Cm^r Ssb2^+	CP1250 × pMSL4
CPM8	CPM2 Δ*comX2*::PcTet	Δ*comX1*, Δ*comX2*, Em^r Tet^r	CPM2 × CPM5
CPM9	CP1250 *comX2′*::(pEVP3)::′*comX2*	*lacZ* fusion to *comX2* by using pEVP3, Cm^r ComX2^+	CP1250 × pMSL3
CPM10	CPM8 *comC′*::(pEVP3)::*comC*^+	*lacZ* fusion to *comC*, Δ*comX1*, Δ*comX2*, ComC^+ Em^r Tet^r Cm^r	CPM2 × CPM5 × CPM6
CPM11	CPM8 *comA′*::(pEVP3)::′*comA*	*lacZ* fusion to *comA*, *comA*, Δ*comX1*, Δ*comX2*; Em^r Tet^r Cm^r	CPM2 × CPM5 × CP1649
CPM12	CPM8 *ssb2′*::(pEVP3)::*ssb2*^+	*lacZ* fusion to *ssb2*, Δ*comX1*, Δ*comX2*, Ssb2^+ Em^r Tet^r Cm^r	CPM2 × CPM5 × CPM7
CPM13	CPM8 *cglA′*::(pEVP3)::′*cglA*	*lacZ* fusion to *cglA*, *cglA*, Δ*comX1*, Δ*comX2*; Em^r Tet^r Cm^r	CPM2 × CPM5 × CP1548
CPM14	CPM8 *celB′*::(pEVP3)::′*celB*	*lacZ* fusion to *celB*, *celB*, Δ*comX1*, Δ*comX2*; Em^r Tet^r Cm^r	CPM2 × CPM5 × CP1601
CPM15	CPM8 *cflA′*::(pEVP3)::′*cflA*	*lacZ* fusion to *cflA*, *cflA*, Δ*comX1*, Δ*comX2*; Em^r Tet^r Cm^r	CPM2 × CPM5 × CP1506
CPM16	CP1250 *comX2′*::(pEVP3)::*comX2*^+	*lacZ* fusion to *comX2* by using pEVP3, Cm^r ComX2^+	CP1250 × pMSL5
CPM17	CPM3 Δ*comE*::PcEm	*lacZ* fusion to *comX1*, Δ*comE*, Em^r Cm^r	CPM3 × aMSL6
Plasmids			
pEVP3		Nonreplicative vector (Cm^r *lacZ*)	8
pMSL1	pEVP3::′*rpoC*::*C-His-10*	Vector tagging *rpoC* with 10 × His	pEVP3::a(DAM211, DAM254)
pMSL2	pEVP3::′*ftsH comX1′*	*comX1* targeting *lacZ* fusion insertion	pEVP3::a(MSL13, MSL14)
pMSL3	pEVP3::′*comX2′*	Mutation of *comX* and *lacZ* fusion	pEVP3::a(MSL27, MSL28)
pMSL4	pEVP3::*ssb2′*	*ssb2* targeting and *lacZ* fusion	pEVP3::a(DAM214, DAM215)
pMSL5	pEVP3::′*nusG comX2′*	*lacZ* fusion to *comX2*, ComX2^+	pEVP3::a(MSL42, MSL28)
pXF520	pEVP3::′*orfL-comC′*	*lacZ* fusion to *comC*, ComC^+	39
Amplicons			
aMSL1	PcEm	Em^r marker with synthetic promoter	a(DAM212, DAM213)
aMSL2	ComX1up::PcEm::comX1dw	For insertion-deletion replacement for *comX1*	See Fig. 3
aMSL3	Pc	Constitutive promoter	a(DAM212, MSL31)
aMSL4	PcTet	New synthetic marker	a(DAM212, MSL33)
aMSL5	ComX2up::PcTet::comX2dw	For insertion-deletion replacement for *comX2*	See Materials and Methods
aMSL6	ComEup::PcEm::comEdw	For insertion-deletion replacement for *comE*	See Materials and Methods

[a] Strain construction by a cross carried out by using transformation is indicated as recipient × DNA donor × additional DNA donor.

expression of the genes for transformation by purifying RNA polymerase holoenzyme from competent pneumococcal cultures. We report here identification of a competence-specific transcription modulator, ComX (probably an alternative sigma factor), which (i) is induced through ComE by the competence-stimulating peptide, (ii) is required for the expression of competence-specific operons which contain a cin-box but not of those of the quorum-sensing operons or of *comX* itself, and (iii) is present in two copies in the Rx derivative strain studied. We also report that the quorum-sensing system alone causes neither competence shutoff nor the CSP refractory period, and

we propose that a ComX-induced gene may be responsible for these postcompetence phenomena.

MATERIALS AND METHODS

Bacterial strains, media, and DNA. All pneumococcal strains, plasmids, and amplicons (DNA products amplified by PCR) used in this study are described in Table 1. CP1500 was used as the source of the template for in vitro transcription and the donor for transformation assays, and CP1250 (39) was a recipient strain and the source for targeting fragments in mutagenic plasmids and *lacZ* fusion plasmids. The insertion vector, pEVP3 (it carries a robust chloramphenicol resistance [Cm^r] marker that is expressed as a single copy independent of target gene transcription, and it has a promoterless *lacZ* downstream of cloning sites

TABLE 2. Oligonucleotides used in this study

Name	Sequence (5′ → 3′)[a]
DAM138	CTTCCACAGTAGTTCACCACCTT
DAM139	TAGCTCTAGACACGCGTAGCAT
DAM140	TAATGCTACGCGTGTCTAGAGCTA
DAM145	ATCCCACTTTATCCAATTTTCGT
DAM206	CTGACTTTCTCAAGATAAAAAGCC
DAM207	TCAACCTCGTCGTGTTCAGA
DAM208	GACTTCCCAATCCTCTGCAAC
DAM209	TTTGTGCCAACAAGCCTAAAT
DAM211	CCAATGCAT-AATCGGTGACAAACACATCG
DAM212	CCGGGCCCAAAATTTGTTTGAT
DAM213	AGTCGGCAGCGACTCATAGAAT
DAM214	ATCAAACAAATTTTGGGCCCGG-AGACGTTAAACGTCCAATCAAGA
DAM215	TCGCAACGGATATCAAGGTGAC
DAM250	TTTTTGTTCATGTAATCACTCCTTC
DAM251	CACGCCAAAGTAAACAATTTAAG
DAM254	CTTAGTGATGGTGATGGTGATGGTGATGGTGATG-TTCTACTGTTTCTTCCACAGTTTCA
MSL13	ATCAAACAAATTTTGGGCCCGG-CACAGTCCCTTGGACTTCTTCAT
MSL14	TGCTATGCTTGGTGCACAGAGTC
MSL15	AATTCTATGAGTCGCTGCCGACT-TTTCGAGGGCGTCAAAGAGTATT
MSL16	ACAACTTACGTTGTAGCGCCCTG
MSL17	CGTGCAATGGTTACAGAGTACGG
MSL18	CTGAGCCAGGATCAAACTCTCAT
MSL23	ATTGTCGGATTGGGAGCAAGAAG
MSL24	TTCATTGCTTAAGACGCGTTCTA
MSL27	GGACTGGTAGACGATATTCCACG
MSL28	TGAAAGAGATAATAATCATCTAGCCA
MSL31	CATGTAATCACTCCTTCTTAA
MSL32	GTAATTAAGAAGGAGTGATTACATG-AATACATCCTATTCACAATCGAA
MSL33	TCCCAAAGTTGATCCCTTAACGA
MSL34D	ACWCCWGGIGTIACIGGWTTYGTIGG
MSL35	ACGCGTAGCATTACGTATTATTGATGG
MSL36	TCGTTAAGGGATCAACTTTGGGA-TTTCGAGGGCGTCAAAGAGTATT
MSL37	CCTGCACATTGGTTCGTCTTGTT
MSL38	TTTGGTTGTGTCGTAAATTCGAT
MSL39	AGTTGGCTGGTTACCTTGAATGT
MSL42	TTGGAACAAGAAATTCGTGACAT
MSL51	ATCAAACAAATTTTGGGCCCGG-CTCACTTGATGTTCAATAACATCTT
MSL52	TGCTCAGTCAATTGTCTATGCTCG
MSL53	AATTCTATGAGTCGCTGCCGACT-TTTGTATATTGAAACAACAGGGG
MSL54	ACCAACGGACCTTCTATCTGTAGC
Q1	CTTTACGAAACAGGAAAGATGCC
Q2	GTGCTTTTGCAGACTACACTGGT
Q3	CTTGACCAACGGACCTGAGTTGT

[a] -, border of 5′ extension.

allowing transcriptional fusion of a reporter to a targeted gene after integration [8]), was used for cloning pneumococcal targeting fragments in *Escherichia coli* DH10B (Gibco BRL). Plasmid pLS10, provided by S. A. Lacks, and the synthetic erythromycin resistance (Em[r]) marker cassette described previously (8) were used for developing new resistance markers. The sequences of all oligonucleotide primers used in this study are listed in Table 2. Luria-Bertani (LB) medium (45) was used for culture of *E. coli*, and the complete CAT medium (24) and its modified forms were used for that of pneumococcus (see below for details).

Transformation of pneumococcal cells and selection of transformants. Cells were exposed to DNA (0.2 μg/ml) for 45 min after treatment with CSP as described previously (24). Transformants were selected on CAT and 1.5% agar solid media with appropriate drugs at the following concentrations (see details in reference 33): novobiocin, 2.5 μg/ml; chloramphenicol, 2.5 μg/ml; erythromycin, 0.25 μg/ml; and tetracycline, 0.25 μg/ml.

Construction of new pneumococcal strains. To construct the pneumococcal strain CPM1, containing a His-tagged RNA polymerase, we chose to tag the carboxyl terminus of RpoC (13, 41). A blunt-ended 530-bp fragment containing the 3′ end of *rpoC* (gene encoding the β′ subunit of RNA polymerase) and a 10-His-codon extension was amplified by PCR from CP1250 DNA with *Pfu* DNA polymerase (Stratagene), using primers DAM211 and DAM254. DAM211 includes the sequence of *rpoC* extending from −487 to −468 relative to its stop codon and a CCAATGCAT 5′ extension; DAM254 includes the sequence of *rpoC* from −1 to −25, with 10 His codons and a stop triplet as a 5′ extension. After insertion of the PCR product into the *Sma*I site of pEVP3, an *E. coli* DH10B clone carrying a chimeric plasmid having *rpoC* and *lacZ* on the same

strand was identified. The plasmid (pMSL1) was integrated into CP1250 by insertion duplication at the *rpoC* locus by transformation, resulting in strain CPM1. The integrated structure was confirmed by sequencing PCR products at both junctions.

To construct the strain CPM2 (ΔcomX1), an Em[r] marker (PcEm) was amplified by PCR with DAM212 and DAM213 from the synthetic Em[r] cassette (8). comX1up (450 bp), which contained part of both *comX1* and *ftsH* plus a sequence complementary to the end of PcEm, was amplified with MSL14 and MSL13 from CP1250 DNA. comX1dw (430 bp), which contained part of *comX1* and a part of downstream sequence plus a terminus complementary to the other end of PcEm, was amplified with MSL15 and MSL16 from CP1250 DNA. The three PCR products were used as a mixed template for PCR with MSL14 and MSL16 to produce aMSL2. After transformation of CP1250 with aMSL2, the structures of the insertion-deletion mutations in Em[r] clones were confirmed by PCR, which demonstrated the expected junction fragments (620, 630, and 1,900 bp), using the primer pairs DAM250-MSL17, DAM251-MSL18, and MSL17-MSL18.

To construct the strain CPM3, with a *comX1-lacZ* fusion, the amplicon comX1up was inserted into the *Sma*I site in pEVP3, and the construct (pMSL2) was introduced into CP1250 by transformation. The correct orientation and integrity of the *lacZ* insertion in Cm[r] clones were confirmed by observing the expected product (570 bp) of PCR with MSL14 and DAM138.

To construct the strain CPM4, with a *comX1* deletion and a *comX2-lacZ* fusion, the targeting fragment for *comX2* was prepared by PCR from DNA of CPM2 (ΔcomX1) by using MSL27 and MSL28, which are complementary to an

internal region of *comX1*. After insertion into the *Sma*I site of pEVP3, transforming CPM2 with the chimeric plasmid (pMSL3) yielded the double mutant CPM4, whose structure was confirmed by reading the sequences flanking *comX1* and *comX2* (see below).

To construct the CPM5, with a *comX2* deletion, the coding region of the structural gene for tetracycline resistance (Tet[r]) was amplified by PCR with MSL32 and MSL33 from the template pLS10. The fragment Pc (constitutive promoter) was amplified with DAM212 and MSL31 from the synthetic Em[r] cassette (8). Both products were connected by PCR by using DAM212 and MSL33. ComX2up, which contains a *comX2*-specific upstream sequence (part of *nusG*) and a small 5′ segment of *comX* with an extension complementary to an end of PcTet, was amplified with MSL35 and MSL13 from a CP1250 template. comX2dw, which contains a 3′ fragment of *comX2* and adjacent downstream sequence with an extension complementary to the other end of PcTet, was amplified with MSL36 and MSL37 from the CP1250. The three DNA fragments, ComX2up, PcTet, and comX2dw, were used as a template for PCR with primers MSL35 and MSL37 to produce aMSL5. After transformation of CP1250 with aMSL5 and selection of Tet[r] colonies, the gene replacement structure was confirmed by PCR with the primer pairs MSL38-MSL42 and MSL39-MSL18.

To construct the strain CPM7, with a reporter of *ssb2* gene expression, a fragment encompassing the putative promoter of *ssb2* gene was amplified by PCR with DAM214 and DAM215. After the product was cloned in the *Sma*I site of pEVP3, the chimeric plasmid was introduced into CP1250 by transformation. The correct integration structure was confirmed by observing the expected product after PCR with primers specific for vector (DAM138) and for a site upstream of *ssb2* (DAM214). The strains CPM8 through CPM15 were produced by crosses between the strains as described in Table 1.

To construct strain CPM16, which contains a *lacZ* fusion at *comX2* at the same site as in CPM9 and CPM4 and also an intact *comX2* allele, a *comX2*-specific targeting fragment extending from *nusG* into *comX2* was amplified with MSL42 and MSL28. The amplified product was incorporated into the *Sma*I site of pEVP3. The correct chimeric plasmid (pMSL5) was identified by observing a product of the expected size after PCR with primers MSL42 and DAM138. After introduction into CP1250 by transformation, the correct integrated structure was confirmed by obtaining a product of the expected size by PCR with MSL42 and DAM138. To construct strain CPM17, which contains both a deletion of *comE* and a *lacZ* fusion at *comX1*, we assembled amplicon aMSL6, which contains PcEm flanked with terminal *comE* fragments, following a strategy similar to that outlined in Fig. 3. comEup and comEdw fragments were amplified with MSL51 and MSL52 and with MSL53 and MSL54, respectively. aMSL6 was synthesized by PCR with MSL52 and MSL54 by using comEup, comEdw, and PcEm as templates. After introduction of aMSL6 into CPM3 by transformation, the correct integrated structures in Em[r] Cm[r] colonies were confirmed by observing the expected products of PCR with MSL52 and DAM250 and with MSL54 and DAM251.

Purification of His-tagged RNA polymerase. CPM1 was grown in 4 liters of CAT with 6 mM HCl to optical density (OD) of 0.1 at 37°C (24, 33). For some preparations, competence was induced by adding NaOH (11 mM) and CSP (100 ng/ml). After 10 min at 37°C, the cells were chilled to 4°C and harvested by centrifugation at $10,000 \times g$ for 15 min. All further purification steps were done at 0 to 4°C unless otherwise specified. The cell pellet was washed with rinse buffer (10 mM Tris HCl [pH 8.0], 300 mM NaCl, 20% glycerol, 10 mM MgCl$_2$, 5 mM β-mercaptoethanol) (28, 41) once and stored at −80°C for at least 1 h. After thawing, resuspension in 10 ml of rinse buffer, addition of 1 mM phenylmethylsulfonyl fluoride, 0.1% Triton X-100, and 5 μg of DNase I/ml, and incubation for 5 to 10 min at 37°C to lyse the cells and break the chromosomal DNA, the lysate was clarified (14,000 × g, 40 min) and applied to a 0.25-ml Ni-nitrilotriacetic acid agarose (Qiagen) affinity column at the rate of 0.4 ml/min. The Ni-nitrilotriacetic acid resin was previously equilibrated with lysis buffer (rinse buffer with 1 mM phenylmethylsulfonyl fluoride and 0.1% Triton X-100) in a 0.7 cm by 10 cm column. The adsorbed material was sequentially washed with lysis buffer containing 5 mM imidazole (three times with 7 ml each time) and 45 mM imidazole (one or three times with 5 ml each time). Bound proteins were eluted with 1 ml of 105 mM imidazole in lysis buffer and 1 ml of 205 mM imidazole in lysis buffer, sequentially. The combined eluates were dialyzed twice against 200 ml of storage buffer (50 mM Tris HCl [pH 8.0], 250 mM NaCl, 50% glycerol, 10 mM MgCl$_2$, 0.1 mM EDTA, and 1 mM dithiothreitol) overnight and stored at −20°C. Protein concentrations were determined by comparison of bands stained with Colloidal Blue stain (Novex) to known amount of bovine serum albumin (BSA) on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels.

In vitro transcription assay. Templates were amplified by PCR by using *Pfu* DNA polymerase and CP1500 DNA. A blunt-ended 369-bp fragment of the *amiA* locus (GenBank accession no. X17337) containing the promoter and a truncated *amiA* gene was amplified with primers DAM208 (positions 262 and 242) and DAM209 (positions −105 and −85), and a 605-bp fragment of *ssb2* (The Institute for Genomic Research [TIGR]) containing the promoter and the complete gene was prepared with primers DAM206 (positions 447 and 425) and DAM207 (positions −159 and −140) (the parenthetical numbers refer to positions of primer ends relative to the first base of each gene). Expected run-off transcript sizes are 300 bp for *amiA* and 420 bp for *ssb2*. After purification with Millipore ultrafilter units, these PCR products were used for transcription assays

in vitro. The transcription reaction (final volume, 20 μl) consisted of 0.4 pmol of the PCR product, 0.5 μg of pneumococcal RNA polymerase, 20 U of RNasin (RNase inhibitor; Promega), 10 nmol each of ATP, CTP, and GTP, 0.2 nmol of UTP, and 10 μCi of [α-^{32}P]UTP (400 Ci/mM) (Amersham) in transcription buffer (40 mM Tris HCl [pH 8.0], 250 mM KCl, 10 mM MgCl$_2$, 1 mM dithiothreitol, and 50 μg of BSA/ml (11, 43). Initiation complexes were allowed to form by incubation of the reaction mixture at 37°C for 10 min with neither nucleoside triphosphates nor [α-^{32}P]UTP. After addition of the nucleoside triphosphates and labeled UTP, the reaction was continued for 50 min at 37°C and stopped by addition of 10 μl of loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF). Samples were denatured in boiling water for 5 min and electrophoresed in a 6% polyacrylamide gel containing 7 M urea. Dried gels were exposed to Kodak X-Omat film.

Protein sequencing. After competence-specific RNA polymerase prepared from eight 4-liter cultures was dialyzed against distilled water, denatured by heat, and resolved by SDS-PAGE on a 10% NuPAGE Bis-Tris gel following the recommendation by the supplier (NOVEX), the resolved components were electrotransferred to polyvinylidene difluoride membrane and then chemically sequenced by The Protein Research Laboratory at the University of Illinois at Chicago.

PCRs. All PCRs were carried out with the procedure described previously (24) except for variations in the annealing temperature, the extension time, and the enzyme composition. The annealing temperature chosen was 3°C below the annealing temperature predicted by the primer supplier (Operon Technology), and extension time was at least 1 min per kilobase of amplified product. For most analytical PCRs, cloned *Taq* DNA polymerase (Gibco BRL) was used. For preparative PCRs, cloned *Pfu* polymerase (Stratagene) was used, except for PCR to connect templates, where *Taq* (Gibco BRL) and Vent polymerases (New England Biolabs) were mixed 1:1 (50). Ligation-mediated PCR (LMPCR) (37) for obtaining upstream sequence of *comX2* was carried out as described previously (24) by using *Csp*6I and a linker containing DAM139 and DAM140, followed by PCR amplification of the ligated fragments with DAM138 and DAM139. Annealing temperature for PCR with MSL34D and DAM138 to obtain the sequence of the *comX2* upstream region was 40°C.

Southern analysis. Chromosomal DNA from each strain was digested with *Eco*RI (New England Biolabs). Digested DNA (1 μg) was loaded on each lane and was electrophoresed for 3 h at 30 V in a 0.7% agarose gel in 0.5× TBE (Tris-borate-EDTA) buffer. All image-developing processes were based on the protocol provided by the supplier (Stratagene) of the kits for Southern analysis. DNA in the gel was depurinated by incubation in 0.25 N HCl for 10 min and then denatured in 1.5 M NaCl and 0.5 M NaOH for 30 min. After neutralization in 1.5 M NaCl and 0.5 M Tris HCl (pH 8.0), DNA was transferred to a Duralon-UV membrane in 10× SSC buffer (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) by capillary blotting for 16 h (45). After cross-linking by UV irradiation, prehybridization and hybridization reactions were done sequentially in 6× SSC buffer, 5× Denhardt's reagent, 0.5% SDS, and 100 μg of sheared denatured salmon sperm DNA/ml without probe and with probe for 3 and 16 h, respectively, at 65°C. The probe was prepared by using Prime-It Fluor Fluorescence Labelling kit from the PCR-amplified 190-bp internal fragments of *comX1* and boiled for 5 min before hybridization. After hybridization, the membrane was washed three times with 0.1× SSC buffer and 0.1% SDS at 65°C for 15 min. The bands were detected by the Illuminator Chemiluminescent Detection System.

Recovery of sequences flanking *comX1* and *comX2*. The fragments flanking *comX1* were obtained by PCR by using PcEm-specific primers and *comX1*-specific primers (MSL17-DAM250 and MSL18-DAM251) from CPM4 chromosomal DNA. PCR with DAM145-MSL18 yielded a product of the size expected if the duplicated region of *comX2* extended downstream beyond MSL18 from CPM4. As PCR with MSL17-DAM138 failed, a sequence upstream of *comX2* was recovered by LMPCR as described above. As the product exhibited significant homology to *nusG* of *B. subtilis*, PCR using a degenerate primer (MSL34D) complementary to a conserved region of *nusG* and DAM138 was used to obtain an 800-bp product, which extended the homology to *nusG* further. To get more reliable sequences, both *comX1* and *comX2* regions were sequenced forward and backward twice by using sequence primers (Q1, MSL27, Q3, and MSL28 for *comX1* and Q2, MSL27, MSL28, and Q3 for *comX2*) after obtaining fragments from both loci from CP1250 by PCR with primers MSL17-MSL18 (*comX1*) and MSL42-MSL18 (*comX2*).

β-Galactosidase assay. Cultures grown in CAT plus 6 mM HCl at 37°C were split for treatment of one half with 200 ng of CSP/ml and 11 mM NaOH. After incubation for 40 min at 37°C for competence induction, cells were chilled, collected by centrifugation at $12,000 \times g$ for 10 min at 4°C, resuspended in 40 μl of 0.1 M sodium phosphate buffer–0.1% Triton X-100, and lysed for 10 min at 37°C. A β-galactosidase assay with ONPG (*o*-nitrophenyl-β-D-galactopyranoside) as a substrate followed the protocol described previously (45); the activity is expressed in Miller units with respect to the OD of the culture when CSP was added (30).

RESULTS

Identification of a factor associated with RNA polymerase from competent cells and structure of its gene. Because the

putative global competence transcriptional modulator might act as an alternative sigma factor, we chose to search for a factor present specifically in RNA polymerase preparations from competent cultures. Since a procedure for purifying pneumococcal RNA polymerase had not been reported, we employed a His tagging approach (13) to simplify the purification. *rpoC*, the gene for the largest subunit (β′) among the RNA polymerase components, positioned at the end of the putative *rpoBC* operon, was tagged with a sequence encoding 10 histidine residues at the C terminus in CP1250, as described in the Materials and Methods section. To prepare the His-tagged RNA polymerases from both competent and noncompetent cultures, several cultures of strain CPM1 were treated with CSP, but other parallel cultures were not. RNA polymerases were prepared from those cultures as described in the Materials and Methods section. Figure 1A shows six such RNA polymerase preparations, three from independent competent cultures and three from independent untreated, noncompetent cultures. Figure 1B shows that they were active in RNA synthesis. While some polypeptide components of the RNA polymerase preparations obtained varied among the preparations, the seven principal bands in all preparations were similar in molecular weight (MW) to seven which were also reported to be the primary components of similar partially purified RNA polymerase preparations from *B. subtilis* (14). As those bands appear to be the major components of the RNA polymerase holoenzyme, we designate them β, β′, σ, α, δ, ω1, and ω2, following the terminology used for bacillus. Three of these identifications were verified by N-terminal sequence determination (σ, α, and δ) and comparison of the corresponding genes with those of *B. subtilis* (data not shown). An additional band (MW, 19,000) was observed in most RNA polymerase preparations from competent cultures (8 of 10 independent preparations, 3 of which are shown in Fig. 1A) but was absent from all of 10 independent preparations from uninduced cultures. To determine whether RNA polymerase preparations with this particular additional band (e.g., lanes 1 and 2 in Fig. 1A) could direct transcription of a known competence-specific gene, we performed in vitro transcription assays with two linear templates (see Materials and Methods), i.e., an *amiA* fragment which contains the constitutive promoter of the *amiA* gene (2) and a *cilA* (*ssb2*) fragment which contains the cin-box promoter of the *ssb2* gene, which is expressed specifically at competence (5). Polymerase prepared from noncompetent cultures directed transcription from the *amiA* fragment (300 bp) but not at all from the *ssb2* fragment (420 bp) (one pair of results is shown in Fig. 1B). In contrast, polymerase prepared from competent cultures and displaying the additional protein transcribed both the *ssb2* fragment and the *amiA* fragment (two pairs of results are shown in Fig. 1B). These results suggested that the 19-kDa additional RNA polymerase component might be responsible for the transcription of the competence-specific gene, *ssb2*, since the presence of no other protein observed in the RNA polymerase preparations correlated with successful in vitro transcription of *ssb2*.

To analyze the function of this 19-kDa protein further, we obtained its N-terminal amino acid sequence. From 2 μg of gel-purified material, a partial N-terminal sequence, N-XXKE-LYXXVQXXVY-C, was obtained by Edman degradation and microsequencing (see Materials and Methods). In the available partial genome sequence of pneumococcus (the type 4 strain, [1, 47]), a single match to this sequence was found, in contig 4125. Translation beginning at the matched region could yield a 159-amino-acid protein with a translational start site 10 bp downstream of an apparent ribosome binding (Shine-Dalgarno) sequence. The size of the predicted protein, 19.9 kDa,
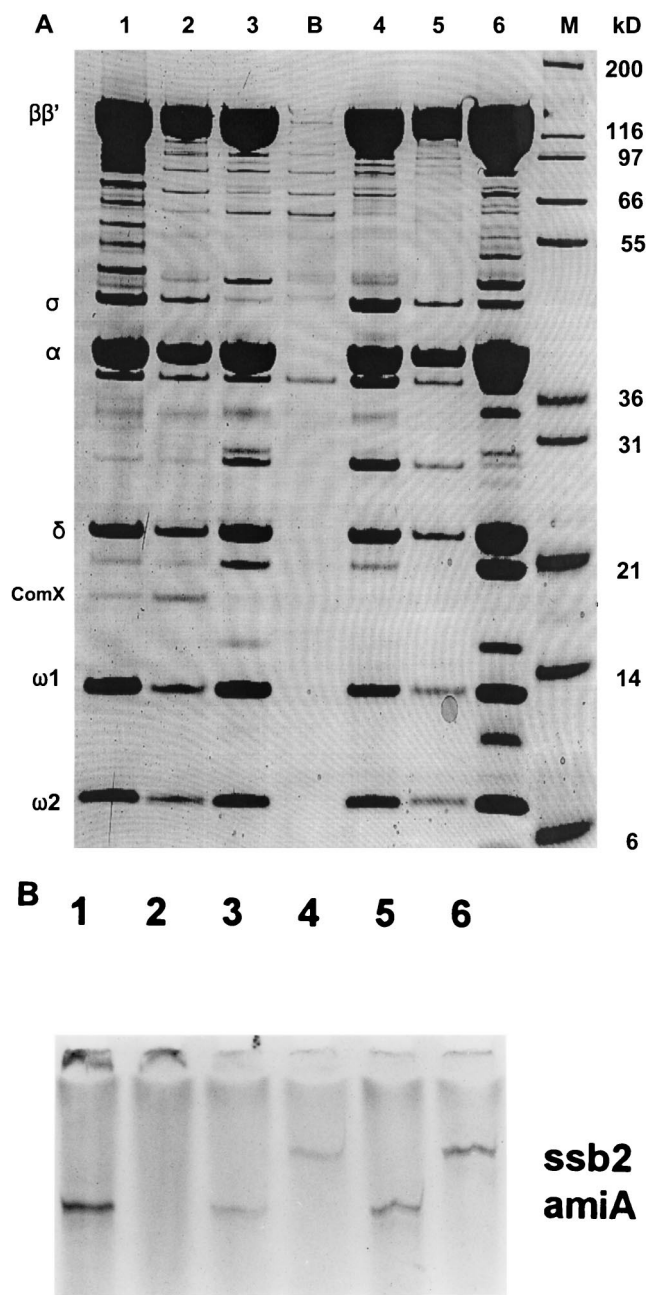

FIG. 1. Specific transcription of a competence gene, *ssb2*, by preparations of His-tagged RNA polymerase with ComX. (A) SDS-PAGE analysis of RNA polymerase preparations from competent and noncompetent cultures is shown. RNA polymerase was prepared from three competent cultures (lanes 1, 2, and 3) and from three independent noncompetent cultures (lanes 4, 5, and 6). Lane B was blank (a control preparation with no His tag from CP1250), and lane M contained standards. The amounts of protein loaded were approximately 25, 10, 25, 25, 10, and 50 μg, from left to right, respectively. The gel was stained with Colloidal Blue stain. (B) Transcriptional specificity of RNA polymerase preparations. In vitro transcription of two genes, *amiA* and *ssb2*, was performed with three RNA polymerase preparations as described in the Materials and Methods section. The reaction products were analyzed on a denaturing 6% polyacrylamide gel. Products are shown for two preparations from competent cultures (lanes 3 and 4, enzyme displayed in panel A, lane 1; lanes 5 and 6, enzyme displayed in panel A, lane 2) and one from a noncompetent culture (lanes 1 and 2, enzyme displayed in panel A, lane 6). Templates were *ssb2* for lanes 2, 4, and 6 and *amiA* for lanes 1, 3, and 5. The predicted transcript sizes are 420 bp for *ssb2* and 300 bp for *amiA*.
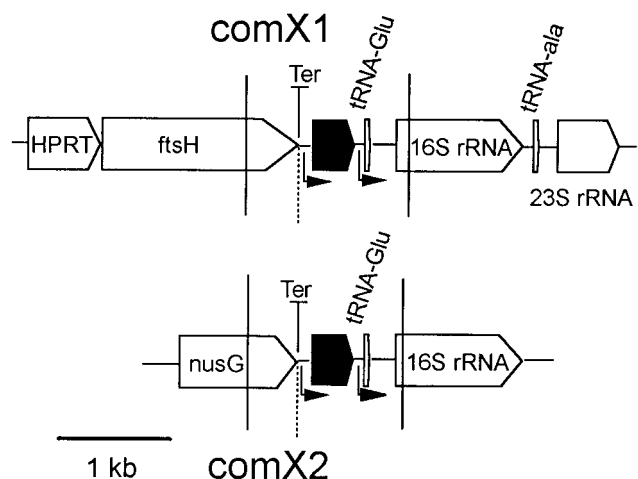
FIG. 2. Map of regions neighboring *comX1* and *comX2*. Solid vertical lines represent the ends of sequence read for each locus, and the dotted vertical line marks the border of structural identity between the loci. Open pentagons represent putative gene assignments; filled pentagons indicate *comX* copies. Ter represents rho-independent terminator; bent arrows represent promoters. The complete map at the top is derived from the partial type 4 genome sequence. The organization of the sequence flanking *comX1* in CP1250 was the same as that of contig 4125 from type 4. The *comX2* locus was absent from the type 4 genome sequence database and was determined in CP1250 as described in the text.

was consistent with the gel position of the competence-specific band shown in Fig. 1A. Since the additional polymerase component appears to be encoded by this open reading frame (ORF), and since the gene proved to be involved in competence regulation (see below), we name the corresponding gene *comX1* (to distinguish it from an additional copy present at a different locus in strain CP1250, as described below).

Since *comX1* is positioned in the middle of contig 4125, we could map the gene (shown in Fig. 2). *ftsH*, a cell division gene, is positioned upstream of *comX1* and is separated from it by a

strong rho-independent terminator. There was no obvious consensus −35 promoter signal, but a putative pneumococcal extended −10 sequence, TcTGtTAgAcT (consensus sequence appears in capital letters), sufficient for promoter activity in several documented cases (44), is found upstream of *comX1*. Although no obvious terminator structure is recognized downstream of *comX1*, a perfect −35 and −10 consensus sequence is followed by tRNA genes and rRNA genes. Therefore, *comX1* seems to be in a monocistronic operon.

comX1 is induced during competence but is not required for competence. To begin characterization of this gene, it was mutated by replacement with an Em[r] marker by using natural transformation and it was fused to a promoterless *lacZ* reporter gene by using pEVP3 (Fig. 3) to produce CPM2 and CPM3, respectively. Enzyme assays of the reporter gene product showed that the expression of *comX1* was turned on in response to treatment with CSP (Table 3). However, in contrast to the other genes known to be expressed specifically during competence, the ΔcomX1 strain CPM2 was transformed at normal levels under the same conditions (Table 4). These observations suggested the possibility of an additional, functional, copy of *comX* in strain Rx.

comX is present in two copies which are flanked by different upstream genes but by identical downstream genes. To test for the presence of an additional, possibly functional, copy of *comX* in strain CP1250, CPM2 (ΔcomX1) was subjected to PCR with primers internal to *comX1* (MSL23 and MSL24 or MSL27 and MSL28) and to Southern blot analysis by using an amplified internal *comX1* fragment as a probe. PCR amplification products were observed (320 and 190 bp) corresponding to internal regions of *comX1* (data not shown). Furthermore, Southern blot analysis revealed two *Eco*RI fragments hybridizing to *comX* for CP1250 and one for CPM2 (data not shown). These data together implied that there was an additional copy of *comX* present in the wild-type genome of Rx. Indeed, as subsequent deletion of a second *comX* gene (see below) produced a strain, CPM8, that was negative on PCR with the same primers used for detecting the presence of *comX1* and *comX2*
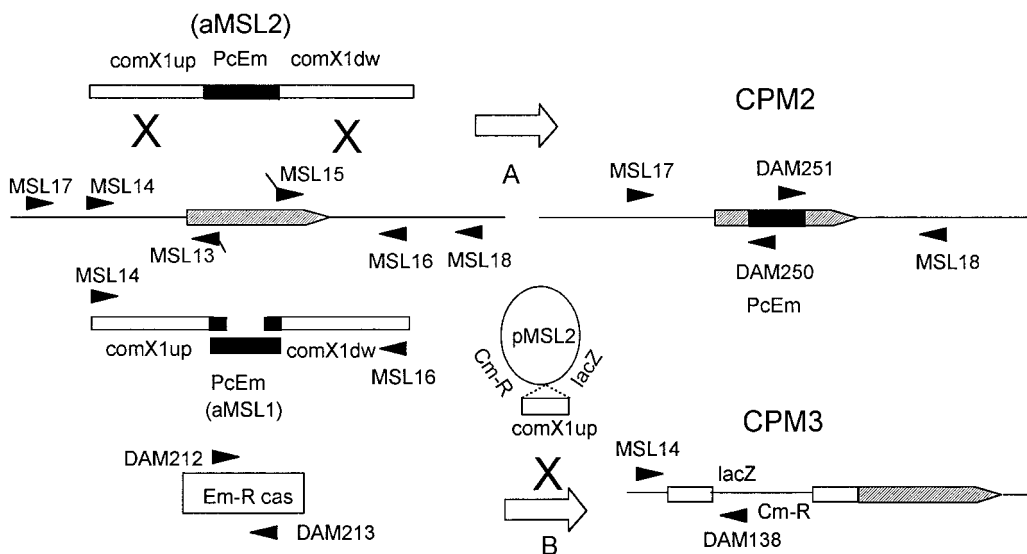
FIG. 3. Strategies for construction and confirmation of an insertion-deletion mutation of *comX1* and a *comX1::lacZ* fusion. Solid arrows and arrowheads mark primers for PCR; hollow arrows represent transformation processes. Hatched pentagon represents *comX1* gene and its direction of transcription. Boxes and lines indicate double-stranded DNA. Black boxes in DNA represent PcEm markers. (A) Construction of CPM2 by transformation of CP1250 with the amplicon aMSL2, assembled by PCR from an Em[r] cassette and two DNA fragments flanking *comX1*. (B) Construction of CPM3 by transformation of CP1250 with pMSL2, an insertion plasmid targeting *comX1*. Details of construction of the strains and the verification of marker integration are described in the Materials and Methods section.

TABLE 3. β-Galactosidase production in various strains showing effects of *comX* mutations on CSP-dependent regulation of transformation genes

Strain	*lacZ* reporter fusion sites	Phenotype[b]	Clone	β-Galactosidase activity[a]							
				Expt 1		Expt 2		Expt 3		Expt 4	
				−CSP	+CSP	−CSP	+CSP	−CSP	+CSP	−CSP	+CSP
CPM6	*comC*	WT		330	360	150	150				
CPM10	*comC*	ComX1⁻ ComX2⁻	1				150	300	390		
			2							240	430
CP1649	*comA*	ComA⁻		2.2	130	2.3	130				
CPM11	*comA*	ComX1⁻ ComX2⁻ ComA⁻	1			2.0	150	2.7	210		
			2							0.9	270
CPM3	*comX1*	WT		0.9	100			3.2	50		
CPM16	*comX2*	WT	1					3.5	130	4.1	190
			2					3.2	140	4.1	180
CPM17	*comX1*	ComE⁻	1					2.7	2.6	2.5	3.4
			2					3.6	3.5		
CPM9	*comX2*	ComX2⁻	1	8.7	260			3.2	150	4.6	170
			2							2.4	190
CPM4	*comX2*	ComX1⁻ ComX2⁻	1	3.1	570			3.3	330	3.1	390
			2							2.7	390
CPM7	*ssb2*	WT		0.9	130	2.2	110			1.7	60
CPM12	*ssb2*	ComX1⁻ ComX2⁻	1			1.8	2.0	3.3	3.2		
			2							3.9	4.5
CP1548	*cglA*	CglA⁻		3.4	370		160	3.5	190		
CPM13	*cglA*	ComX1⁻ ComX2⁻ CglA⁻	1			1.8	1.8	2.9	2.9		
			2							3.9	3.1
CP1601	*celB*	CelB⁻		1.5	90	1.8		2.5	50		
CPM14	*celB*	ComX1⁻ ComX2⁻ CelB⁻	1			2.2	2.1	2.9	3.2		
			2							3.5	4.4
CP1506	*cflA*	CflA⁻		0.5	90		70				
CPM15	*cflA*	ComX1⁻ ComX2⁻ CflA⁻	1				1.6	3.0	3.2		
			2							3.7	4.1
CP1250	None	WT		0.4	1.5			2.9	2.8	3.1	2.5

[a] Two independent clones of each strain were grown in CAT containing 6 mM HCl to an OD_{550} of 0.05 in several experiments. After 1 ml of culture was either not treated (−CSP) or treated (+CSP) with CSP and NaOH, incubation was continued for 40 min. Enzyme activity in crude cell extract at 40 min was measured by the ONPG assay and is expressed in Miller units (30).
[b] WT, wild type.

(data not shown), we conclude that *comX* is present in two copies in strain CP1250, an Rx derivative.

To map both *comX* copies more exactly, their flanking sequences were obtained by LMPCR and direct PCR from strain

TABLE 4. One copy of *comX* is sufficient and required for transformation

Strain	Relevant genotype	Clone	Competence[a] (Novr transformants/µl)		
			Expt 1	Expt 2	Expt 3
CP1250	WT		830	770	1,560
CPM2	ΔcomX1 comX2⁺	1	400	540	1,480
		2	2,000		
		3	570		
CPM5	comX1⁺ ΔcomX2			810	1,350
CPM8	ΔcomX1 ΔcomX2			<0.01[b]	<0.01
CPM4	ΔcomX1 comX2	1		<1	
		2		<1	

[a] Several independent clones of each strain were grown in CAT containing 6 mM HCl to an OD_{550} of 0.025 in several experiments. 5MC (Nov^r) DNA was added to 10 µg/ml with CSP and NaOH. Incubation was continued for 40 min. Competence was measured as the number of Novr transformants in 1 µl (10^5 cells) of transformed culture except for CPM8, where it was measured in 100 µl of the culture. Where transformants were observed, at least 100 colonies were counted.
[b] No transformants were observed when the designated portion of culture was challenged.

CPM4 (see below), using the inserted marker sequences to provide copy-specific tags, as described in the Materials and Methods section. The results showed that the region containing *comX1* in strain Rx had exactly the same structure, except for several single-base substitutions, as that of the *comX* allele contained in the type 4 genome database, while *comX2* (absent from the partial genome database) was flanked upstream by *nusG*, an antiterminator homologue, and downstream by tRNA and rRNA genes were identical to those adjacent to *comX1*. The detailed sequences read at least twice in each strand for each *comX* region (GenBank accession no. AF161700 and AF161701). The results showed that the sequences of *comX1* and *comX2* in CP1250 were perfectly identical to one another. Furthermore, their putative promoter regions were identical up to the terminal triplets of the upstream genes. Several base differences were observed between the regions downstream of *comX1* and *comX2*. Finally, although the sequence of the *comX1* gene in the TIGR database had several base changes compared to *comX1* in CP1250, the predicted protein products differed at only one position (Glu31Asp).

A comparison of the deduced amino acid sequence of *comX* with other known protein sequences was done by using BLAST to search both the GenBank database and the unfinished microbial genome databases at the National Center for Biotechnology Information (NCBI [36]) and at WIT (52). It identified two genes of unknown function encoding very similar proteins:

```
SPN ComX       1  ----------------------------------MIKEL EEVQGTVYKCRNEYY LHLW L
SMU ORF        1  -------------------------------ME DFEIVFNKVKPIVWKLSRYYFIKMWTR
SPY ORF        1  ------------------------------MSI TRAAFEKVKPILKLKRHYYIQLWDR
LLA hSigH      1  ------------------------------ FEGLFSGMKPIIRKLMKQIRIKAWDI
EFA hSigH      1  MLEQ------------------IQQALAGDQ QFQAIGQYQPVYKCMKKYY KDYDT
BSU SigH       1  MNLQNNKGKFNKEQFCQLEDEQVIEKVHVGDS ALDYLITKYRNFVRAKARSYLIGADR
Consensus      1                                    lf    v k  r yyl  wd


SPN ComX      28  SDWEQEGMVCLHELISR--EEGLVDLIPRLRKYFKTKFRNRILDYIRKQESQKRRYDKE-
SMU ORF       31  EDWQQEGMVILHQLLRE--HPELEEDTKLYIYFKTRSNYIKDVLRQQESQKRRFNRM-
SPY ORF       31  DDWQQEGHIILLQLLER--YPELIEEERLYRYFKTKFSSYLKDLLRRQESQKRQFHKL-
LLA hSigH     28  EDYQEGMIILHILLEEN-HPSTN-----IYTKFKVKIHQHLIDELRHSYAKKRLHDIF-
EFA hSigH     42  EDWIQEGRIVFFRTLERF-DSEKQ---ASLGKIFKRNFENHVHSLIRHQCAYKRRTDTLS
BSU SigH      61  EDIVQEGMIGLYKSIRDFKEDKLTSFKAFAELCITRQIITAIKTATRQKHIPLNSYASLD
Consensus     61  eDw QEG i l  l               l   fk f  i  lR     kr


SPN ComX      85  -----PYEEVG-EISHRISEGGLW-LDDYYIFHETLRSYRNKQ----SKEKQEELERVLS
SMU ORF       88  -----SYEEVG-EIEHCISGGMQ-LDEYILFRDSLLAYKQGL----STEKQEIFERIVA
SPY ORF       88  -----AYEEVG-EVAHAIPSRGLW-LDDYIAKQEVIASLENQL----NSQERMQFQALIR
LLA hSigH     81  -----VGLDIY-ECSDWIDAGGST-PESEIVFIHLLAVYEGL----SAHYQEILVRQMR
EFA hSigH     98  ISLEQRLESAGEEIVDYIGTAITD-PLGQIMIQEKLAAFPTRL----SSLEQIILNAYL
BSU SigH     121  KP--IFDEESDRTLLDVISGAKTLNPEEMLINQEEFDIIEMKVGELLSDLERKVLVYLD
Consensus    121         ee   ei  i  gg    d il e l       l    s         l


SPN ComX     134  N---ERFRGRQ-------RVLRDIRIVFKEFTIRTH----
SMU ORF      137  G---EHFLGRQ-------SILKDIRKILSDFKEK------
SPY ORF      137  G---ERFKGRR-------AILRKISPYFKEFAQQL-----
LLA hSigH    130  G---EELTRME-------RYRLREKIKNILFSRDD----
EFA hSigH    153  GYTLEEIAKVRGGTFSAVRSAYDRAKAKMKYTLLNE----
BSU SigH     179  GRSYEEISDELNRHVKSIDNALQRVKIKLEKYLEIREISL
Consensus    181  g   e                         f
```
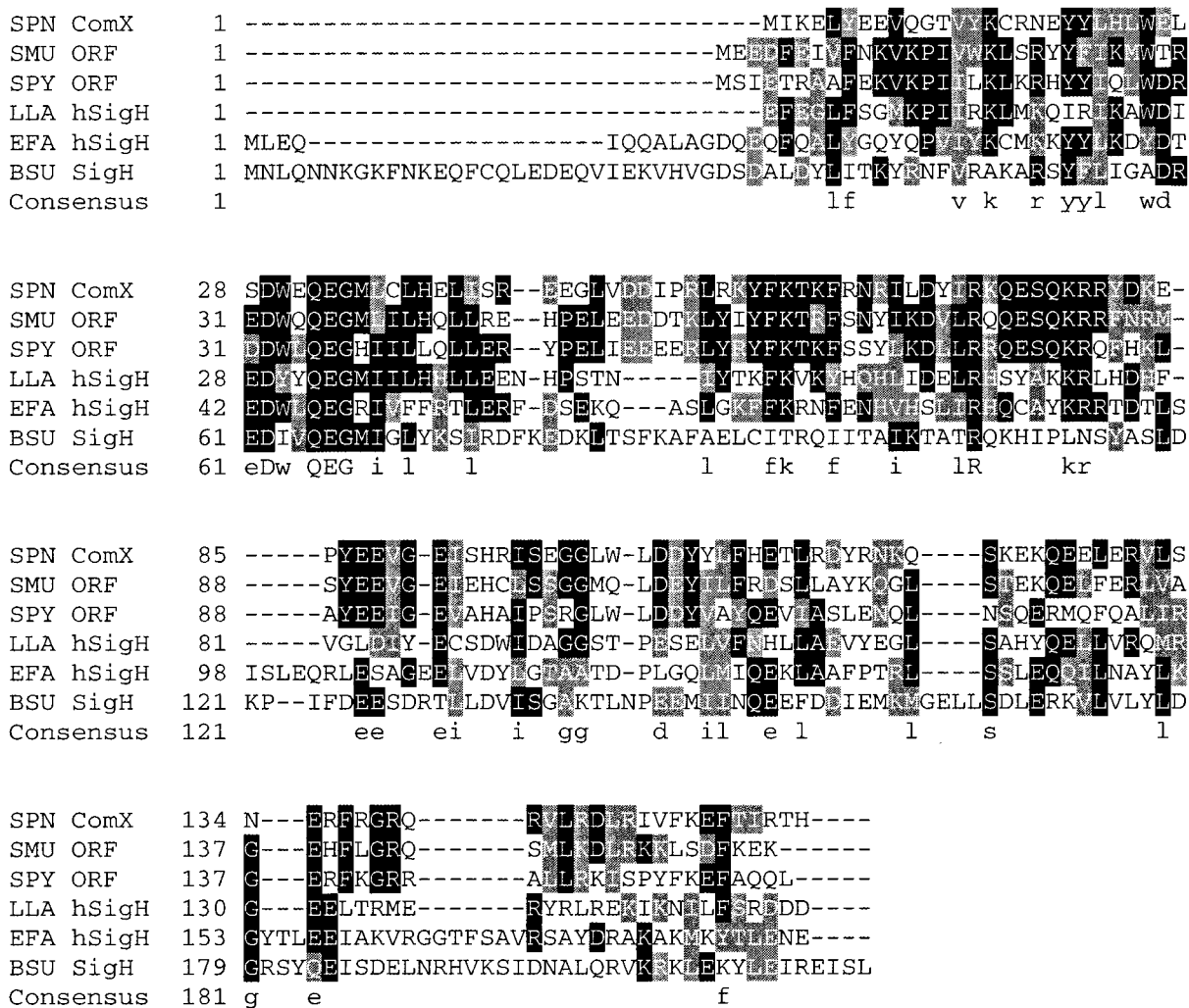
FIG. 4. Sequence alignment of ComX homologues. The sequences of the homologues were aligned with the program Clustal W (9). Amino acid similarity is indicated by highlighting (black background, ≥50% identity; shaded background, ≥50% similarity). In the consensus sequence, residues with ≥80% similarity are shown in lowercase and 100% matches are shown in uppercase. The most highly conserved region of the sigma 70 family (subregion 2.2) is located between residues 56 and 75 of *B. subtilis sigH*. Protein abbreviations and sources are as follows: SPN ComX, ComX in *S. pneumoniae*, accession no. RPN00272 in the WIT database (52); SMU ORF, *S. mutans*, contig 746 (bp 2275 through 1796 of the sequence in the NCBI database [36]); SPY ORF, *S. pyogenes* accession no. RST00265 in the WIT database; LLA hSigH, *L. lactis* tr/Q48591; EFA hSigH, *E. faecalis* REF02274 in the WIT database; BSU SigH, *B. subtilis* RBS00098 in the WIT database.

an ORF in *Streptococcus pyogenes* (expectation value [E] = 10^{-31}, 40% identical and 65% similar residues; Fig. 4) and another ORF in *Streptococcus mutans* (E = 10^{-36}, 46% identical and 66% similar residues). In addition, the search revealed two rather more-distant homologues; one from *Enterococcus faecalis* (E = 10^{-9}, 32% identical and 46% similar residues) and the other from *Lactococcus lactis* (E = 10^{-8}, 26% identical and 45% similar residues). The two distant homologues of ComX are classified as sigma H homologues within the sigma 70 family of sigma factors on the basis of weak similarity to sigma-H of *B. subtilis*, an accessory sigma factor which is involved in sporulation and stationary-phase transcription (12, 21). Furthermore, this set of six putative sigmas exhibited their greatest similarities in a highly conserved region, subregion 2.2 (26, 27), which is the most highly conserved region among sigma 70 family members; 11 of 20 residues in the region were more than 80% similar among the six sigmas, and four residues were identical in all six (Fig. 4).

Either copy of *comX* is sufficient for competence. To test whether ComX is required for transformation, we mutated the second copy of *comX* carried by the Δ*comX1* strain CPM2 by insertion-duplication mutagenesis using 190 bp internal to *comX* as a targeting fragment. In the resulting double mutant, CPM4, transformation was reduced to less than 0.2% of wild-type levels (Table 4). In the *comX* double deletion strain, CPM8, constructed as described below, transformation was reduced still further, to less than 0.002% of wild-type level (Table 4). Therefore, the *comX2* gene is required and sufficient for genetic transformation in the Δ*comX1* background.

The results presented above establish that *comX1* is induced by CSP and that *comX2* is sufficient for competence. To confirm that *comX1* is also active, as suggested by the identity of sequences in the promoters of the two *comX* copies and by its induction by CSP, we constructed a tagged *comX2* deletion, using a new synthetic marker, PcTet (aMSL4), composed of a Tet[r] gene under the control of a synthetic pneumococcal con-
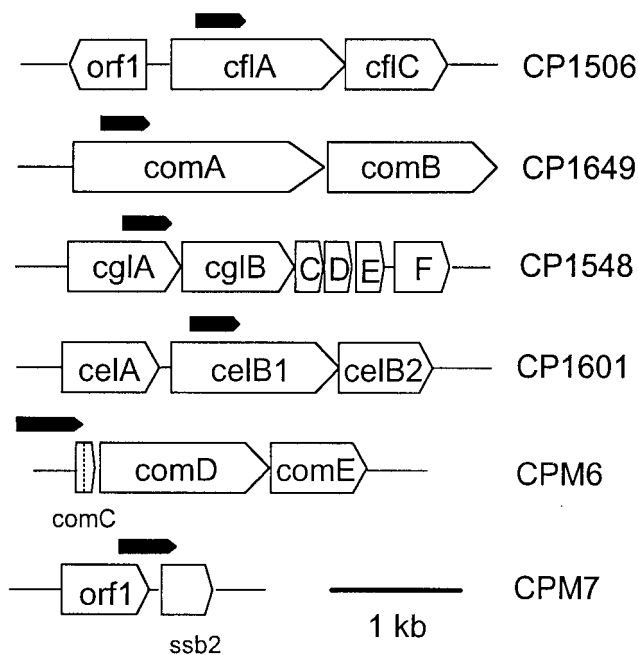
FIG. 5. Map of *lacZ* fusion sites at six competence loci. Each targeting fragment (filled pentagons) directed the nonreplicative vector, pEVP3, to introduce a *lacZ* fusion at the site marked by the point of the pentagon. Fusion constructs except that of CPM7 (see the Methods and Materials section) were described previously (24). The exact positions of the targeting fragments in the competence loci in the mutants are as follows: strain CP1506, contig 4155 (bp 6235 through 5935); CP1649, contig 4105 (bp 4920 through 5220); CP1548, contig 4194 (bp 16895 through 16595); CP1601, contig 4139 (bp 4806 through 4506); CPM6, U33315 (bp 332 through 721); CPM7, contig 4219 (bp 836 through 476). Gene positions are as follows: *cflA*, contig 4155 (bp 6454 through 5156); *celB1*, contig 4139 (bp 5029 through 3506); *comA*, contig 4105 (bp 4693 through 6852); *cglA*, contig 4194 (bp 17495 through 16554); *ssb2*, contig 4219 (bp 514 through 118).

stitutive promoter (8). The new marker was substituted for *comX2* in CP1250 by the strategy described in the Materials and Methods section. As the resulting ΔcomX2 strain, CPM5, transformed at normal levels (Table 4), we conclude that *comX1* is also sufficient for transformation and thus that both copies of *comX* are active, with either one being sufficient for wild-type levels of competence.

comX is a regulatory gene, upstream of cin-box genes but downstream of comE. *comX* is induced by CSP and required for genetic transformation, properties shared with many competence-related genes. Since ComX was associated with RNA polymerase, seemed to be responsible for specific in vitro transcription of *ssb2*, and was similar only to hypothetical sigma H homologues among genes whose function is known, it was a reasonable candidate for the hypothetical global transcription modulator of the transformation machinery genes mentioned in the introduction. To determine independently the functional position of ComX in the hierarchy of competence regulation, we carried out a series of epistasis experiments with *lacZ* reporter fusions, examining both the dependence of *comX* expression on other competence genes and the dependence of other competence genes' expression on *comX*. First, six strains, which contained each of the competence genes *comC*, *comA*, *ssb2*, *cglA*, *celB*, and *cflA* fused to *lacZ* (a fusion map is shown in Fig. 5), were crossed into the *comX* double-mutation background, by transforming CPM2 simultaneously with the *comX2* deletion construct from CPM5 and individual *lacZ* fusion con-

structs, to make CPM10 (*comC*::*lacZ*), CPM11 (*comA*::*lacZ*), CPM12 (*ssb2*::*lacZ*), CPM13 (*cglA*::*lacZ*), CPM14 (*celB*::*lacZ*), and CPM15 (*cflA*::*lacZ*). All these reporters exhibited the expected low or negligible expression in noncompetent cultures (except for those of CPM6 and CPM10). As shown in Table 3, *comAB*, an operon of the quorum-sensing system, was induced by CSP equally well in the *comX* double-mutation background and in the *comX*-positive background (compare data for CPM11 and CP1649). The expression of *comCDE*, the other quorum-sensing operon, was not affected by *comX* mutation (compare data for CPM10 and CPM6; CPM6 was previously noticed to exhibit higher endogenous *comCDE* expression in competence-inhibiting acidic media for unknown reasons [39]). In contrast, induction of four competence operons containing the cin-box (*ssb2*, *cgl*, *cel*, and *cfl*) was abolished in the *comX* double-mutation background, but it was clearly observable in the *comX*-positive background (compare data for CPM12, CPM13, CPM14, and CPM15 with those for CPM7, CP1548, CP1601, and CP1506). Thus, *comX* is required for the expression of competence-specific genes containing the consensus (cin-box) but not for the expression of the quorum-sensing operons *comCDE* and *comAB*. To test whether *comX* gene products affect the expression of *comX* itself, we made isogenic *comX2*::*lacZ* fusion strains that were ComX positive (CPM16), ComX2 negative (CPM9), and negative for both ComX1 and ComX2 (CPM4). In all these strains, *lacZ* expression was found to be induced by CSP (Table 3), showing that the induction of *comX2* is not dependent on the *comX* product. Thus, *comX* is a regulatory gene that divides the class of genes induced by CSP into two groups: those dependent on *comX* function and those that are inducible without *comX* function. The former appear to be principally those thought of as structural genes of DNA processing; the latter appear to be quorum-sensing genes and *comX*. These data are perfectly consistent with the idea that ComX could act on the cin-box directly for the expression of the structural genes since it affects only cin-box genes among genes known to be induced by CSP.

Since *comX* was induced by CSP and expression of the *comCDE* operon was not affected by *comX* loss, the quorum-sensing system seemed likely to be upstream of ComX in the process of competence induction. To locate ComX more precisely, we examined the expression of *comX* in a *comE* (quorum-sensing transducer) deletion background. The results (Table 3) show that ComE is upstream of ComX for competence induction, since CSP elicited no detectable β-galactosidase activity in CPM17 (ΔcomE, comX1::lacZ). Since ComE is upstream of ComX and there is a report that ComE binds a direct repeat region common to the putative promoters of the *comCDE* and *comAB* operons, we searched for any similar conserved sequences in the putative promoters of *comX* and *comAB*. Indeed, similar direct repeats are observable upstream of all three operons, with constant separation and a constant distance from an apparent −10 consensus. They are -68TTG gGAGAA-60 and -48TTGaGAGAA-40 at *comCDE*, -68AtGg agaAA-60 and -48AgGagctAA-40 at *comX*, and -68TGGaggG ag-60 and -48TGGgaaGga-40 at *comAB*. The consensus among the six repeat units was (T70A30)(G50T50)G(G50A50) (G70A30)(A50G30)(G70) (A80G20)(A80G20). Thus, the mechanism of regulation of *comX* may share some components with that for regulation of *comAB* and *comCDE*, and ComE may act at the direct repeats in the promoters of all three operons.

When the relative activities of *comX2* promoters in the wild type (CPM16) and in the *comX* double-deletion background (CPM4) were compared by using identical reporter fusion sites, cumulative β-galactosidase activities after 40-min incubation with CSP were always found to be more than twice as
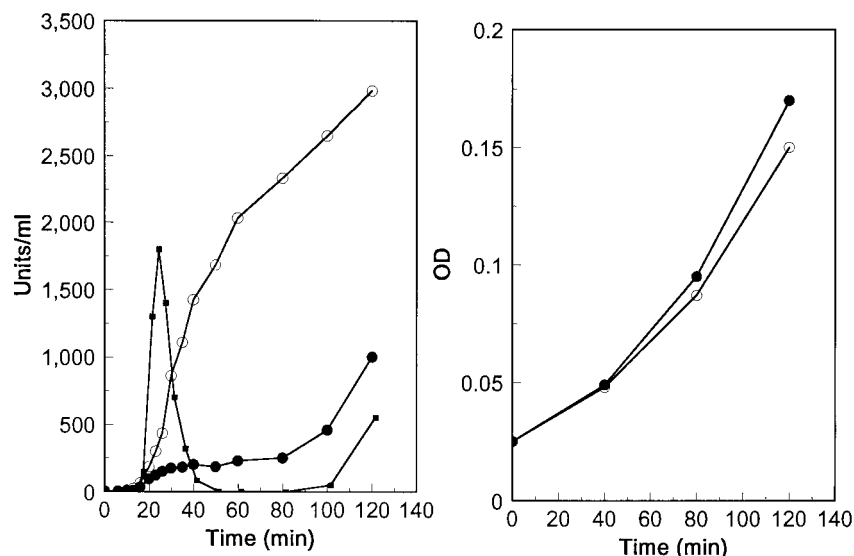
FIG. 6. Comparison of *comX* promoter activities in CPM16 and CPM4 after competence induction. Competence was induced by CSP and NaOH (at time 0) for each strain at OD_{550} of 0.025. Left panel: β-galactosidase activities (Miller units) were monitored for both strains CPM16 (●) and strain CPM4 (○). Transformation (1 unit = 100 Novr transformants) for CPM16 (■) was also monitored at the same time by determining the number of Novr transformants after exposing a portion of the culture to donor DNA for 90 s. Right panel: Growth of CPM16 (●) and CPM4 (○) was monitored by measuring OD_{550}.

high for CPM4 as for CPM16 (Table 3). This pattern suggests a negative feedback effect on expression of *comX* by *comX* gene products or by *comX*-regulated genes. To address the question of whether the increase in cumulative β-galactosidase in the *comX* mutant background arose via increased expression during the induction phase or from a failure of competence shutoff (34) at later stages in the response to CSP, we compared in detail the *comX* expression kinetics of these two strains from the moment of treatment with CSP through the shutoff phase and refractory period (Fig. 6). The results revealed a strong influence of *comX* on *comX* expression. In CPM4, *lacZ* expression was higher than in CPM16, both during the competence induction phase (15 to 25 min) and at the time of competence shutoff (25 to 40 min); *lacZ* synthesis continued even during the refractory period (40 to 90 min). In CPM16, in contrast, *lacZ* expression ceased during the shutoff phase and refractory period, reappearing only as competence reappeared, after 100 min. These observations suggest that during competence induction, the *comX* product inhibits its own synthesis to some degree and show that during the later competence shutoff phase and refractory period, either the *comX* gene product or the product of a *comX*-induced gene inhibits *comX* gene expression almost completely. Thus, *comX* affects its own expression levels during all phases of the response to CSP.

DISCUSSION

We have identified a new competence-induced gene (*comX*) in pneumococcus, present in two identical copies, that is required for competence. Epistasis experiments showed that cin-box operons depended on this gene for expression but that two known competence operons involved in quorum-sensing (*comCDE* and *comAB*), with different promoter characteristics, did not. Additionally, the *comX* gene product was not required for its own expression. Since the expression of *comX* did depend on *comE*, we conclude that ComX establishes a link between the quorum-sensing information transduced to *comE* and the later steps of competence induction.

ComX was initially identified as a protein present in preparations of RNA polymerase from competent cultures but absent from parallel RNA polymerase preparations from noncompetent cultures. His-tagged RNA polymerase containing the ComX protein directed transcription of a competence-specific late gene, *ssb2*, but the polymerase from noncompetent cultures did not, even though it was active in transcription from a constitutive gene, *amiA*. On the basis of the epistasis results, the association of ComX with competent polymerase, the cin-specific activity of ComX-containing polymerase, and the similarity of ComX to sigma H, we hypothesize that the competence-specific factor (ComX) induced by the quorum-sensing system is an alternative sigma factor which recognizes the cin-box and T-rich region in the promoter of transformation machinery genes and directs their transcription. It is interesting to note that the central four bases of the cin-box (GAAT) are identical to four of five bases of the −10 consensus sequence recognized by the *B. subtilis* sigma H (53).

It has long been recognized (31, 42) that streptococcal competence for genetic transformation can entail a transient but global shift of protein synthesis, occurring simultaneously throughout a growing culture, such that expression of many genes is shut down just when a specific set of genes is activated. This global protein shift accompanying competence induction would be explained by the proposed role of *comX* as a competitive sigma in which it could block transiently not only the transcription of quorum-sensing operons and the *comX* gene itself but also the large number of genes transcribed by the primary RNA polymerase.

We propose the following model for regulation of genetic transformation incorporating a role of *comX* in mediating the connection of quorum sensing to competence induction, based on results from previous studies and the present study, illustrated in Fig. 7. Environmental cues that affect competence, such as pH (7), phosphate concentration, and peptides (3) might modulate the basal level of *comCDE* and *comAB* expression or modulate the effectiveness of CSP interaction with its cognate receptor, ComD (18). As cell density in a growing
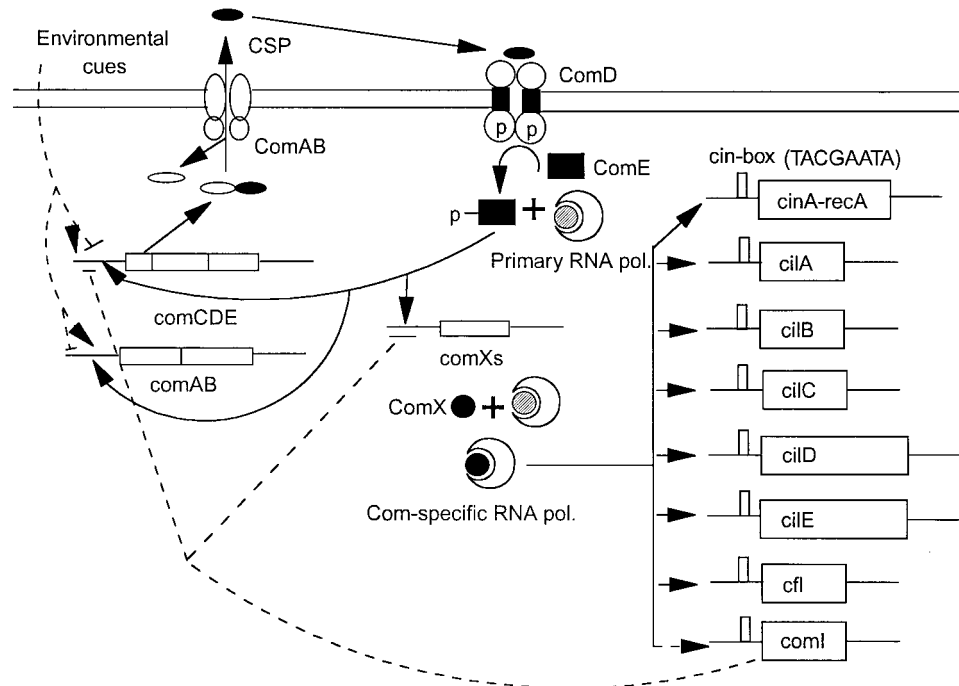
FIG. 7. Hypothetical model for the regulation of genetic transformation. Solid arrows indicate processing steps or transcriptional activation steps that have been shown to take place or for which supporting observations are described in the text. T bars indicate negative regulation. Dashed lines indicate hypothetical links. *comI*, a putative gene responsible for competence shutoff and refractory period.

culture increases, basal expression of these operons causes the absolute concentration of CSP outside the cells to increase, reaching a threshold concentration at which CSP can effectively activate the cognate receptor. Activated ComD (sensor of the two-component system), in turn, may activate ComE (response regulator) by phosphorylating it, as in many other bacterial two-component systems (46). Phosphorylated ComE may activate three competence-specific operons, *comCDE*, *comAB*, and *comX*, directly by interacting with the major RNA polymerase (primary sigma and core RNA polymerase) and the direct repeats in these promoters, since it has been reported that ComE binds to a direct repeat at the *comCDE* promoter strongly and to a similar site at the *comAB* promoter weakly (51) and since similar putative binding sites are identifiable in the promoter regions of these three operons. With ensuing accumulation of ComX, it replaces the primary sigma (sigma A homologue) from RNA polymerase holoenzyme, to create a competence-specific RNA polymerase holoenzyme, directing expression of cin-box-containing genes and thus allowing the synthesis of the genes for DNA transport and recombination machinery.

With the addition of ComX to the signal transduction chain linking CSP to competence, it becomes possible to begin to assign responsibility for other aspects of competence to either of two experimentally distinguishable sets of induced genes—those of the quorum-sensing circuit versus *comX* and those dependent on *comX* action. One such aspect is the transient nature of competence, which disappears rather rapidly despite the continued presence of high levels of CSP. The activity of the *comX* promoter measured during competence induction was increased by mutation of *comX*, and in the *comX* mutant background, the expression continued without interruption during the competence shutoff phase and refractory period, whereas the same gene was silenced in the *comX*-positive back-

ground. The first phenomenon suggests that the *comX* product inhibits its own synthesis to some degree, which could be explained simply on the basis of sigma factor displacement. Such a replacement of the sigma A homologue by *comX* is also likely to reduce expression of the two quorum-sensing operons, for which the major RNA polymerase also seems to be necessary, and could be the reason for the precocious reduction of *comCDE* expression during competence induction previously observed by Alloing et al. (4). But such a reduction of the expression of quorum-sensing operons and of *comX* through the role of ComX as a competitive sigma factor seems insufficient to shut off competence completely soon after induction or to keep it off for another generation, since the overall culture growth rate is not perturbed by exposure to CSP (Fig. 6). Therefore, the complete shutoff of competence and the long refractory period both seem to require an additional negative control mechanism. Alloing et al. suggested that the expression of *comCDE* induced by a phosphorylated ComE might be inhibited by a different, more highly phosphorylated form of ComE (4). However, the observation that *comX* expression persisted for a long time in *comX* double mutants suggests that the quorum-sensing system upstream of *comX* cannot accomplish the inhibition alone. Therefore, we conclude that one of the genes regulated by *comX* effects the complete shutoff of competence after induction and its continued suppression. Possible targets for the product of such a gene include the promoters of the *comCDE* operon or *comX* or the proteins ComD or ComE.

Many species (including *Streptococcus milleri*, *Streptococcus anginosus*, *Streptococcus constellatus*, *Streptococcus oralis*, *Streptococcus mitis*, *Streptococcus crista*, *Streptococcus intermedius*, *Streptococcus gordonii*, *Streptococcus sanguis*, and *S. pneumoniae*) in the genus *Streptococcus* are known to be naturally competent (19). *Streptococcus pyogenes*, although not known to

be competent naturally, appears to carry homologues for many transformation machinery genes, such as *celAB*, *cglAB*, *recA*, and *ssb*, and has homologues to the quorum sensor, ComD, and to the transducer, ComE (19). In addition, there are homologues to *dal* and *cfl* (data not shown). Thus, *S. pyogenes* has essentially all major components of the transformation machinery genes. Furthermore, there is also a copy of the cin-box in the promoter region of each of these apparent transformation machinery genes (19). As *S. pyogenes* also carries a homologue to ComX, the apparatus for transformation in *S. pneumoniae* and *S. pyogenes* seems to be broadly conserved, including elements for quorum sensing, for activation of transformation genes by ComX, and for transformation itself. This conservation suggests that even though a ComC homologue has not been found, some kind of quorum-sensing effector might be present in *S. pyogenes*. It would be interesting to learn whether expression of the ComX homologue in *S. pyogenes* makes the species competent, which could greatly facilitate the genetic manipulation of the species. *S. mutans* is also naturally transformable (38); although the sequencing of *S. mutans* is not far advanced, it has already revealed possible homologues to *celAB*, *comDE*, and *comX*. Therefore, it seems possible that *S. mutans* will also turn out to carry a transformation system much like that in pneumococcus.

Natural competence in another gram-positive bacterium, *B. subtilis*, is known to depend on a quorum-sensing system (23). Accumulated small peptide pheromones (ComX and CSF) in the bacillus growing culture stimulate *comS* expression through a bacterial two-component system, *comPA*, at high cell density, and in turn, ComS activates ComK by releasing the inhibitor MecA. ComK is the key transcription factor for expression of all of the late competence genes that encode the DNA processing and uptake machinery (16, 49). Thus, cell density signals in *B. subtilis* lead to activation of a transcription factor which directs the major RNA polymerase to promoters of transformation machinery genes, while in *S. pneumoniae*, a cell density signal (CSP) induces expression of ComX, which appears to act as a competence-specific sigma factor to focus transcription on a regulon composed of transformation machinery genes.

ACKNOWLEDGMENTS

REFERENCES

1. **Aaberge, I. S., J. Eng, G. Lermark, and M. Løvik.** 1995. Virulence of *Streptococcus pneumoniae* in mice: a standardized method for preparation and frozen storage of the experimental bacterial inoculum. Microb. Pathog. **18:** 141–152.

2. **Alloing, G., M. C. Trombe, and J. P. Claverys.** 1990. The *ami* locus of the gram-positive bacterium *Streptococcus pneumoniae* is similar to binding protein-dependent transport operons of gram-negative bacteria. Mol. Microbiol. **4:**633–644.

3. **Alloing, G., C. Granadel, D. A. Morrison, and J. P. Claverys.** 1996. Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. Mol. Microbiol. **21:**471–478.

4. **Alloing, G., B. Martin, C. Granadel, and J. P. Claverys.** 1998. Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum-sensing by the oligopeptide permease. Mol. Microbiol. **29:**75–83.

5. **Campbell, E. A., S. Y. Choi, and H. R. Masure.** 1998. A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis. Mol. Microbiol. **27:**929–939.

6. **Cato, A., and W. R. Guild.** 1968. Transformation and DNA size. J. Mol. Biol. **37:**157–178.

7. **Chen, J. D., and D. A. Morrison.** 1987. Modulation of competence for genetic transformation in *Streptococcus pneumoniae*. J. Gen. Microbiol. **133:** 1959–1967.

8. **Claverys, J. P., A. Dintilhac, E. V. Pestova, B. Martin, and D. A. Morrison.** 1995. Construction and evaluation of new drug-resistance cassettes for gene disruption mutagenesis in *Streptococcus pneumoniae*, using an *ami* test platform. Gene **164:**123–128.

9. **Clustal W Website.** Program server. [Online.] http://www.clustalw.genome-.ad.jp. [12 April 1999, last date accessed.]

10. **Coomaraswamy, G.** 1996. Induction of genetic transformation in *Streptococcus pneumoniae* by a pheromone peptide and its synthetic analogues. Ph.D. thesis. University of Illinois, Chicago.

11. **Deora, R., and T. K. Misra.** 1996. Characterization of the primary σ factor of *Staphylococcus aureus*. J. Biol. Chem. **271:**21828–21834.

12. **Dubnau, E., J. Weir, G. Nair, L. Carter III, C. Moran, Jr., and I. Smith.** 1988. *Bacillus* sporulation gene *spo0H* codes for σ30 (σ^H). J. Bacteriol. **170:**1054–1062.

13. **Fujita, M., and Y. Sadaie.** 1998. Rapid isolation of RNA polymerase from sporulating cells of *Bacillus subtilis*. Gene **221:**185–190.

14. **Haldenwang, W. G., N. Lang, and R. Losick.** 1981. A sporulation-induced sigma-like regulatory protein from *B. subtilis*. Cell **23:**615–624.

15. **Haldenwang, W. G.** 1995. The sigma factors of *Bacillus subtilis*. Microbiol. Rev. **59:**1–30.

16. **Hamoen, L. W., A. F. Van Werkhoven, J. J. Bijlsma, D. Dubnau, and G. Venema.** 1998. The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. Genes Dev. **12:**1539–1550.

17. **Håvarstein, L. S., G. Coomaraswamy, and D. A. Morrison.** 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **92:**11140–11144.

18. **Håvarstein, L. S., P. Gaustad, I. F. Nes, and D. A. Morrison.** 1996. Identification of the streptococcal competence pheromone receptor. Mol. Microbiol. **21:**863–869.

19. **Håvarstein, L. S., and D. A. Morrison.** 1999. Quorum sensing and peptide pheromones in streptococcal competence for genetic transformation, p. 9–26. *In* G. M. Dunny and S. C. Winans (ed.), Cell-cell signaling in bacteria. ASM Press, Washington, D.C.

20. **Hotchkiss, R. D.** 1954. Cyclical behavior in pneumococcal growth and transformability occasioned by environmental changes. Proc. Natl. Acad. Sci. USA **40:**49–55.

21. **Jeong, S., H. Yoshikawa, and H. Takahashi.** 1993. Isolation and characterization of the *secE* homologue gene of *Bacillus subtilis*. Mol. Microbiol. **10:**133–142.

22. **Karudapuram, S., X. Zhao, and G. J. Barcak.** 1995. DNA sequence and characterization of *Haemophilus influenzae dprA*+, a gene required for chromosomal but not plasmid DNA transformation. J. Bacteriol. **177:**3235–3240.

23. **Lazazzera, B. A., T. Palmer, J. Quisel, and A. D. Grossman.** 1999. Cell density control of gene expression and development in *Bacillus subtilis*, p. 27–46. *In* G. M. Dunny and S. C. Winans (ed.), Cell-cell signaling in bacteria. ASM Press, Washington, D.C.

24. **Lee, M. S., B. A. Dougherty, A. C. Madeo, and D. A. Morrison.** 1999. Construction and analysis of a library for random insertional mutagenesis in *Streptococcus pneumoniae*: use for recovery of mutants defective in genetic transformation and for identification of essential genes. Appl. Environ. Microbiol. **65:**1883–1890.

25. **Londono-Vallejo, J. A., and D. Dubnau.** 1993. *comF*, a *Bacillus subtilis* late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases. Mol. Microbiol. **9:**119–131.

26. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843–3849.

27. **Lonetto, M. A., K. L. Brown, K. E. Rudd, and M. J. Buttner.** 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase σ factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. USA **91:**7573–7577.

28. **Losick, R., and M. Chamberlin.** 1976. RNA polymerase. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

29. **Mejean, V., and J. P. Claverys.** 1988. Polarity of DNA entry in transformation of *Streptococcus pneumoniae*. Mol. Gen. Genet. **213:**444–448.

30. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

31. **Morrison, D. A., and M. F. Baker.** 1979. Competence for genetic transformation in pneumococcus depends on synthesis of a small set of proteins. Nature **282:**215–217.

32. **Morrison, D. A., and B. Mannarelli.** 1979. Transformation in pneumococcus: nuclease resistance of deoxyribonucleic acid in the eclipse complex. J. Bacteriol. **140:**655–665.

33. **Morrison, D. A., S. A. Lacks, W. R. Guild, and J. M. Hageman.** 1983. Isolation and characterization of three new classes of transformation-deficient mutants of *Streptococcus pneumoniae* that are defective in DNA transport and genetic recombination. J. Bacteriol. **156:**281–290.

34. **Morrison, D. A.** 1997. Streptococcal competence for genetic transformation:

regulation by peptide pheromones. Microbial Drug Res. **3:**27–38.

35. **Mortier-Barriere, I., A. de Saizieu, J. P. Claverys, and B. Martin.** 1998. Competence-specific induction of *recA* is required for full recombination proficiency during transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **27:**159–170.

36. **National Center for Biotechnology Information Website.** Unfinished microbial genome database under BLAST server. [Online.] http://www.ncbi.nlm.nih.gov. [19 March 1999, last date revised.]

37. **Palittapongarnpim, P., S. Chomyc, A. Fanning, and D. Kunimoto.** 1993. DNA fingerprinting of *Mycobacterium tuberculosis* isolates by ligation-mediated polymerase chain reaction. Nucleic Acids Res. **21:**761–762.

38. **Perry, D., and H. K. Kuramitsu.** 1981. Genetic transformation of *Streptococcus mutans*. Infect. Immun. **32:**1295–1297.

39. **Pestova, E. V., L. S. Håvarstein, 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.

40. **Pestova, E. V., and D. A. Morrison.** 1998. Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a *lacZ* reporter insertion vector. J. Bacteriol. **180:**2701–2710.

41. **Qi, Y., and M. F. Hulett.** 1998. PhoP-P and RNA polymerase sigmaA holoenzyme are sufficient for transcription of pho regulon promoters in *Bacillus subtilis*: phoP-P activator sites within the coding region stimulate transcription in vitro. Mol. Microbiol. **28:**1187–1197.

42. **Raina, J. L., and A. W. Ravin.** 1980. Switches in macromolecular synthesis during induction of competence for transformation of *Streptococcus sanguis*. Proc. Natl. Acad. Sci. USA **77:**6062–6066.

43. **Rao, L., R. K. Karls, and M. J. Betley.** 1995. In vitro transcription of pathogenesis-related genes by purified RNA polymerase from *Staphylococcus aureus*. J. Bacteriol. **177:**2609–2614.

44. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended −10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144–155.

45. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

46. **Stock, J. B., M. G. Surette, M. Levit, and P. Park.** 1995. Two-component signal transduction systems: structure-function relationships and mechanism of catalysis, p. 25–51. *In* J. A. Hoch and T. J. Silhavy (ed.), Two-component signal transduction. American Society for Microbiology, Washington, D.C.

47. **TIGR Website.** Sequences on the Institute for Genomic Research web site. Nov. 21, 1997, release date. [Online.] http://www.tigr.org. [18 March 1999, last date accessed.]

48. **Tomasz, A., and R. D. Hotchkiss.** 1964. Regulation of the transformability of pneumococcal cultures by macromolecular cell products. Proc. Natl. Acad. Sci. USA **51:**480–487.

49. **van Sinderen, D., A. Luttinger, L. Kong, D. Dubnau, G. Venema, and L. Hamoen.** 1995. ComK encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. Mol. Microbiol. **15:**455–462.

50. **Wach, A.** 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. Yeast **12:**259–265.

51. **Ween, O., P. Gaustad, and L. S. Håvarstein.** Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. Mol. Microbiol., in press.

52. **WIT Website.** Genome annotations. [Online.] http://wit.mcs.anl.gov.

53. **Wösten, M. M. S. M.** 1998. Eubacterial sigma-factors. FEMS Microbiol. Rev. **22:**127–150.

54. **Zhou, L., F. M. Hui, and D. A. Morrison.** 1995. Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes. Gene **153:**25–31.