BOX Elements Modulate Gene Expression in *Streptococcus pneumoniae*: Impact on the Fine-Tuning of Competence Development^{∇†}

Eivind Knutsen,¹ Ola Johnsborg,¹ Yves Quentin,² Jean-Pierre Claverys,² and Leiv Sigve Håvarstein^{1*}

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway,¹ and
Laboratoire de Microbiologie et Génétique Moléculaires, UMR 5100 CNRS-Université Paul Sabatier,
118 route de Narbonne, 31062 Toulouse Cedex 9, France²

Received 14 June 2006/Accepted 14 August 2006

More than 100 BOX elements are randomly distributed in intergenic regions of the pneumococcal genome. Here we demonstrate that these elements can affect expression of neighboring genes and present evidence that they are mobile. Together, our findings show that BOX elements enhance genetic diversity and genomic plasticity in *Streptococcus pneumoniae*.

The BOX elements of *Streptococcus pneumoniae* are short repeated sequences whose function and origin are unknown. They were first discovered more than a decade ago in the vicinity of genes involved in virulence and natural competence, and it was therefore speculated that BOX elements could be involved in regulating the expression of these genes (13). A total of 127 BOX elements are randomly distributed in intergenic regions of the TIGR4 strain (15, 20), whereas 115 elements are present in strain R6 (7). Most of these elements consist of three different modules (boxA, boxB, and boxC), and boxB (45 bp) is flanked by boxA (59 bp) and boxC (50 bp) (13). Most often, one to four copies of boxB are located between boxA and boxC, but occasionally up to eight copies have been found. Some BOX elements, however, consist of only boxA and boxC or, in a few cases, of single or tandem boxB sequences. The mechanism behind this heterogeneity and the possible biological significance of it remain to be elucidated. BOX elements containing a boxA module and a boxC module have the potential to form a stable stem-loop structure. This structure appears to be functionally important as compensating base changes have been observed in several cases where the sequences of boxA-boxC pairs differ from the consensus sequence (13). The presence of such stem-loop structures in transcribed regions could modulate the expression of neighboring genes. We therefore decided to investigate whether the BOX elements associated with the early competence (*com*) genes *qsrAB* and *comAB* affect their level of expression. Competence for genetic transformation in *S. pneumoniae* is regulated by competence-stimulating peptide (CSP) (2, 5), a secreted peptide pheromone, through a signal transduction pathway consisting of the histidine kinase ComD and the cognate response regulator ComE (6, 17). When the external concentration of CSP in a pneumococcal culture reaches about 10 ng/ml, early and late *com* genes are expressed, resulting in

development of competence (5). The early *com* genes are regulated by ComE, which initiates transcription from promoters (P_E) containing a conserved direct-repeat motif (21). The *comAB* genes encode an ABC transporter (ComA) and its accessory protein (ComB), which together constitute the CSP secretion apparatus (8), whereas the *qsrAB* genes encode an ABC transporter having an unknown function.

The BOX elements associated with the *qsrAB* and *comAB* genes are located in the promoter regions of these operons, between P_E and the first downstream gene (Fig. 1A and B), and they are therefore present at the 5' end of transcripts originating from P_E . We have previously shown that expression of *qsrAB*, which in noncompetent cells depends on a housekeeping σ^A promoter (P_A) situated between boxC and the *qsrA* gene (Fig. 1A and B), increases about twofold when competence is induced (10, 21). Interestingly, insertion of a *lacZ* reporter gene immediately downstream of P_E strongly impaired CSP-induced expression of the reporter compared with a location of *lacZ* in the *qsrB* gene (10). This observation suggested that the presence of a BOX element at the 5' end of transcripts somehow enhances expression of downstream genes. To further explore this possibility, we reintroduced boxAB^{2C} at its original position between P_E and *lacZ*. This was done by using PCR to specifically amplify the *qsr* boxAB^{2C} element with primers containing BamHI sites at their 5' ends. Subsequently, the PCR fragment was ligated in either orientation into plasmid pOE4144 (10) at a unique BamHI site located exactly at the original position of the *qsr* boxAB^{2C} (see Tables S1 and S2 in the supplemental material). Both constructs were transformed into strain EK100 and integrated into its genome by single-crossover insertion-duplication at the homologous target site upstream of boxAB^{2C} (see the supplemental material for a more detailed description). The resulting strains, OE4144-AB^{2C} and OE4144-CB^{2A}, were assayed for CSP-induced production of β -galactosidase (β -Gal) as described previously (10). Strain OE4144-AB^{2C} exhibited a fourfold increase in β -Gal activity compared to OE4144, demonstrating that the presence of the BOX element had a strong positive effect on the expression of *lacZ* (Fig. 2). The level of CSP-induced expression of *lacZ* was even higher with the inverted BOX element (strain OE4144-CB^{2A}), indicating that the proposed secondary structure rather than the orientation

* Corresponding author. Mailing address: Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, 1432, Ås, Norway. Phone: 47-64965883. Fax: 47-64965901. E-mail: sigve.havarstein@umb.no.

† Supplemental material for this article may be found at <http://jbb.asm.org/>.

[∇] Published ahead of print on 22 September.

FIG. 1. Map of early competence genes *qsrAB* and *comAB* (A) and their promoter regions (B). The CSP-inducible promoter, P_E, consists of two binding sites for ComE, designated DR1 and DR2, and the cognate -10 hexamer. P_A is a strong constitutive housekeeping σ^A promoter that drives expression of *qsrAB* in noncompetent pneumococci. The BOX element in the *qsrAB* promoter region contains two boxB modules (shaded) sandwiched between boxA and boxC modules, whereas the BOX element upstream of *comAB* contains only a single boxB module. RBS, ribosome binding site.

of the BOX element is important. We then investigated the effects of different combinations and orientations of the boxA, boxB, and boxC modules (Fig. 2). The constructs shown in Fig. 2 were made by using plasmid pCR2.1-TOPO::boxAB²C as the template in PCRs with specific primer pairs annealing to regions flanking the box modules to be deleted (see Table S2 in the supplemental material). As the primers were complementary to opposite strands and contained unique restriction sites

at their 5' ends, copies of pCR2.1-TOPO::boxAB²C containing different permutations of the box modules were created. The various box combinations were excised from their pCR2.1-TOPO plasmids with BamHI and ligated in either orientation into the corresponding site downstream the P_E promoter in pOE4144. The resulting constructs were introduced into the genome of EK100 as described above (see the supplemental material for details). The constructs lacking boxB modules,

FIG. 2. Effects of the 5'-proximal *qsr* BOX element and various derivatives of this element on the level of expression of a cotranscribed *lacZ* reporter gene. The complete boxAB²C element and downstream P_A promoter were removed in the OE4144 parental strain, and expression of the *lacZ* gene was therefore driven exclusively by the CSP-inducible *qsr*-P_E promoter (see Fig. 1). In the AB²C and CB²A mutants, the original boxAB²C element was reintroduced in the forward and inverted positions, respectively. Similarly, mutants containing only boxA and boxC modules (AC) or boxB modules (B²), inserted in either orientation, were constructed. The BOX elements of the AB² and the B²C mutants lacked the boxC and boxA modules, respectively, whereas seven tandem boxB modules were inserted between boxA and boxC modules in the AB⁷C and CB⁷A mutants. The β-Gal activity is expressed in Miller units, and the data are means ± standard errors of three independent experiments.

OE4144-AC and OE4144-CA, both exhibited a fourfold increase in *lacZ* expression compared to OE4144 (Fig. 2), demonstrating that the presence of the boxA-boxC combination alone is sufficient to stimulate expression of downstream genes and that boxAC works equally well in both orientations. This finding supports the idea that the stimulatory effect of boxAB²C is due mainly to the formation of a stem-loop structure in the *qsrAB* transcript. The level of expression of *lacZ* was significantly higher in mutant OE4144-CB²A than in mutant OE4144-AB²C, indicating that the boxB modules might have an inhibitory effect that depends on their orientation. The properties of a series of mutants having various orientations and numbers of boxB modules support this view. The level of expression of the reporter gene in the OE4144-B² inverted mutant, which contained two boxB modules in the inverted orientation, was not significantly different from the level of expression in the strain lacking the boxAB²C element (OE4144). In contrast, in mutants OE4144-B², OE4144-AB², and OE4144-B²C, in which the boxB modules were incorporated in the forward orientation, there was a clear reduction in *lacZ* expression compared to the expression in the OE4144 parental strain (Fig. 2). The number of boxB modules inserted in tandem between boxA and boxC in BOX elements dispersed throughout the genome was found to range from zero to eight. To determine whether the inhibitory effect observed with forward-oriented boxB modules increased with a larger number of tandem repeats, we constructed mutants OE4144-AB⁷C and OE4144-CB⁷A containing seven boxB modules in the forward and inverted orientations, respectively. This was done by amplifying a boxAB⁷C motif located upstream of the spr1604

FIG. 3. Comparison of spontaneous competence development in pneumococcal strains containing (OE4170) or lacking (OE4171) a BOX element in the *comAB* promoter region. The OE4180 parental strain, which contains a wild-type *comAB* locus with an intact boxABC element, was included as an additional positive control. (Upper panel) Monitoring of competence (i.e., luciferase expression from a transcriptional fusion between *luc* and the late *com* gene *ssbB*) in strains OE4170 (◆), OE4171 (□), and OE4180 (△) grown at 37°C with an Anthos Lucy 1 luminometer for 7.5 h. (Lower panel) Corresponding growth curves (optical densities). The number of relative light units (RLU) (luminescence/optical density at 492 nm) and optical density at 492 nm (OD₄₉₂) were determined at 10-min intervals. The experiment was repeated a number of times with similar results.

gene by PCR, ligating it in either orientation into the BamHI site of pOE4144, and transforming the resulting constructs into the EK100 strain as described above. Our results showed that the level of β-Gal produced was reduced more than threefold in the mutant containing boxAB⁷C, strongly indicating that forward-oriented boxB modules residing in the 5' end of transcripts downregulate expression of cotranscribed genes. In accordance with this finding, a comparison of mutants OE4144-AB²C and OE4144-AB⁷C showed that considerably less β-Gal was produced in the mutant containing the highest number of boxB modules (Fig. 2). Together, these data indicate that forward-oriented boxB modules have an inhibitory effect on the

TABLE 1. Comparison of BOX elements at 13 different loci

BOX element located close to pneumococcal gene ^a	Function or putative function of gene	Composition and orientation of BOX element in:				
		<i>S. pneumoniae</i> strain TIGR4	<i>S. pneumoniae</i> strain 670	<i>S. pneumoniae</i> strain 23F	<i>S. pneumoniae</i> strain R6	<i>S. mitis</i> NCTC12261
SP1697 (<i>mmsA</i> or <i>recG</i>) (BOX element starts 18 bp downstream of stop codon)	DNA helicase involved in homologous recombination	CB ² A	CB ³ A	Absent	Absent	CA
SP0097 (BOX element starts 5 bp downstream of stop codon)	Function unknown	AC	Absent	Absent	Absent	ABC
SP0095 (BOX element starts 50 bp downstream of stop codon)	Predicted sulfur transferase, function unknown	CBA	CBA	CBA	Absent	Absent
SP0042 (<i>comA</i>) (BOX element starts 17 bp upstream of start codon)	CSP transporter	ABC	ABC	ABC	ABC	ABC
SP1180 (BOX element starts at stop codon)	Ribonucleoside diphosphate reductase, small subunit	AC	Absent	AC	Absent	Absent
SP0293 (BOX element starts at stop codon)	Function unknown	CA	Absent	Absent	Absent	ABC
SP0965 (<i>hlyB</i>) (BOX element starts 9 bp downstream of stop codon)	Glucosaminidase involved in cell separation	Absent	Absent	Absent	Absent	ABC
SP0782 (BOX element starts at stop codon)	Induced during competence, function unknown	ABC	ABC	ABC	ABC	Absent
SP1872 (<i>pit1A</i>) (BOX element starts 18 bp downstream of stop codon)	Periplasmic binding protein involved in uptake of iron	CBA	CBA	CBA	CBA	Absent
SP0561 (BOX element starts 21 bp upstream of start codon)	Function unknown	AB ² C	AB ³ C	AB ⁴ C	AB ² C	ABC
SP1717 (<i>qsrA</i>) (BOX element starts 115 bp upstream of start codon)	ABC transporter presumably involved in Na ⁺ transport	AB ² C	AB ² C	AB ² C	AB ² C	QsrA absent
SP2184 (<i>glpF</i>) (BOX element starts 1 bp downstream of stop codon)	Glycerol uptake facilitator protein	AC	Absent	AC	Absent	Absent
spr0072 (BOX element starts 31 bp upstream of start codon)	Glycosyl transferase family	CBA	CB ² A	CB ² A	CBA	Absent

^a The prefixes SP and spr refer to *S. pneumoniae* strains TIGR4 and R6, respectively.

expression of downstream genes and that the level of inhibition increases with the number of tandemly repeated boxB units. Inverted boxB modules, on the other hand, apparently had no such effect.

To determine whether the boxABC element residing in the *comAB* promoter (Fig. 1) affected the expression level of these genes, a mutant strain lacking this element was constructed. In brief, a PCR fragment generated by amplifying a 1,302-bp fragment encompassing the 5' end of *comA*, the *comAB* promoter, and its upstream region was ligated into pLitmus 28 (New England Biolabs), producing recombinant plasmid pLicom1. boxABC was removed from pLicom1 by the PCR procedure described previously using primers flanking the boxABC element (see Table S2 in the supplemental material). The resulting plasmid, pLicom2, and parental plasmid pLicom1 were digested with HindIII and BamHI to excise the cloned inserts. Then the excised fragments were ligated into the corresponding restriction sites upstream of the promoterless *lacZ* gene of the nonreplicating pEVP3 vector (see Table S1 in the supplemental material). Finally, the constructs were introduced into the genome of the EK100 strain by natural transformation, followed by single-crossover homologous recombination (see the supplemental material for details). Deletion of the boxABC element (strain OE4151) resulted in a modest but significant reduction (26%) in CSP-induced β -Gal production compared to the production in the strain in which the *lacZ* reporter gene was placed behind the wild-type *comAB* promoter (strain OE4150) (data not shown). Even though the reduction was small compared to that observed for the *qsrAB* locus, we thought that it was worthwhile to investigate whether deletion of the *comAB* BOX element affected spontaneous competence development. Since it has been demonstrated pre-

viously that the CSP export capacity of wild-type cells is rate limiting for competence development (14), we hypothesized that competence would be delayed (i.e., would develop at a higher cell density) in cultures of the mutant lacking a BOX element compared to that in the wild-type parent. To check this hypothesis, we compared spontaneous competence induction in strains OE4171 (*comAB* BOX element deleted), OE4170 (identical to OE4171 but BOX element present), and OE4180 (wild-type *comAB* locus) by monitoring luciferase activity from a transcriptional fusion between the *luc* gene and the late *com* gene *ssbB*. To generate the OE4170 and OE4171 strains, we first constructed strains OE4061 (*comAB* BOX element deleted) and OE4060 (identical to OE4061 but *comAB* BOX element intact) by using insertion-duplication mutagenesis. The procedure used was essentially the same procedure that was used for construction of strains OE4150 and OE4151 (see above), except that a 1-kb PCR fragment containing the complete *comAB* promoter and the 5' half of the *comA* gene was used as a starting point. The *luc* reporter gene was inserted behind the *ssbB* promoter by transforming OE4160, OE4161, and CP1200 with a derivative (pR459) of plasmid pR424 (1), resulting in strains OE4170, OE4171, and OE4180, respectively (see the supplemental material for details). Luciferase activity was detected as described previously (1, 18). The results showed that the OE4171 mutant strain developed competence at a higher cell density than the positive controls developed competence (Fig. 3), strongly suggesting that the BOX element plays an important role in the fine-tuning of development of spontaneous competence in *S. pneumoniae*.

Altogether, the data described above clearly show that BOX elements situated at the 5' end of mRNA transcripts have the potential to change the expression pattern of operons (e.g.,

FIG. 4. (A) BOX element insertion generates a single-nucleotide duplication. Comparison of chromosomal sequences flanking BOX elements in the TIGR4 genome (open reading frames identified by designations beginning with SP) with the corresponding BOX element-free sequences in the R6 genome (open reading frames identified by designations beginning with SPR or the gene designations) (as listed in Table 1) revealed the presence of a single additional nucleotide (large type) flanking the BOX element on one side or the other. Note that the additional nucleotide is always identical to the nucleotide (underlined) present on the other side of the BOX element, which is highly suggestive of an insertion mechanism that generated a single nucleotide duplication. (B) Diagram of ISSpn2 and ISStso1, two iso-ISs detected in the genomes of *S. pneumoniae* and *S. sobrinus*, respectively. These ISs are thought to be responsible for generating and mobilizing BOX elements in *S. pneumoniae*. (C) Alignment of the extremities of ISSpn2 and of ISStso1 with the extremities of boxA and boxC. The IRs of each IS are underlined. IRL and IRR, left and right inverted repeats, respectively.

qsrAB) and even entire regulons (the competence regulon in the case of *comAB*). It has been firmly established for both gram-negative and gram-positive bacteria that the 5' end of transcripts is a major determinant of mRNA stability and that mRNAs are stabilized by 5'-proximal stem-loop structures (3, 4, 9, 19). It is therefore highly plausible that the 5'-proximal

BOX elements studied here stimulate expression of downstream genes by increasing the half-lives of their mRNAs. As exemplified by the *qsrAB* operon, BOX elements do not need to be located immediately upstream of an open reading frame in order to be part of the cognate mRNA transcript. It is therefore difficult to estimate the prevalence of pneumococcal mRNA transcripts with BOX elements in their 5' ends. However, it is unlikely that the *qsrAB* and *comAB* mRNAs are unique. A comparison of 13 loci in the genomes of four different pneumococcal strains and *Streptococcus mitis* NCTC12261 demonstrated that BOX elements are only partially conserved at each specific site (Table 1). A typical example is the inverted boxABC element located 50 bp downstream of the SP0095 gene. This element is conserved in three pneumococcal strains but is missing in *S. pneumoniae* R6 and the *S. mitis* NCTC12261 strain. This implies that the BOX element was deleted from the R6 and *S. mitis* strains or that it was recently inserted into the genome of the common ancestor of the TIGR4, 670, and 23F strains. Interestingly, the boxABC element located upstream of the *comA* gene is conserved in all four pneumococcal strains, as well as in *S. mitis*. It is tempting to speculate that its stability is due to increased fitness of the cells carrying it and therefore that highly conserved BOX elements have acquired a function giving the host cell a selective advantage. In the examples listed in Table 1, all boxB modules are flanked by boxA and boxC modules. Curiously, the number of boxB modules sandwiched between boxA and boxC modules varies for different BOX elements in the genome of a particular strain and also for corresponding elements in different strains (Table 1). This variation in the number of sandwiched boxB modules might not necessarily result from mobilization of boxB modules but could be generated by a gene conversion mechanism (e.g., through homology-dependent repair of a damaged boxB motif using intact tandemly repeated boxB modules as the template). However, the presence of elements consisting of only single or tandem boxB modules (data not shown) suggests that the boxB module is able to move independent of the boxA and boxC modules.

The availability of BOX element-free and BOX element-occupied sites (Table 1) enabled us to compare sequences flanking BOX element insertion sites with the aim of identifying possible nucleotide duplications. Duplications are very frequently generated during transposition (12). A dinucleotide (TA) was previously detected at the borders of another repeated element of *S. pneumoniae*, RUP (16). This dinucleotide was suggested to result from duplication of the target generated by the transposase of IS630-*Spn1*, an insertion sequence (IS) belonging to the IS630 family identified in the pneumococcal genome and proposed to be responsible for formation and mobilization-amplification of RUP elements (16). Interestingly, comparison of BOX element-free and BOX element-occupied sites revealed the presence of a single additional nucleotide (A, T, or C) in every case examined (Fig. 4A). It is noteworthy that the same nucleotide flanks the other extremity of the BOX element (Fig. 4A), strongly suggesting that BOX element insertion generates a single-nucleotide duplication. This duplication event readily accounts for the previous observation that identical nucleotide pairs (either AA or TT) were present at the base of the predicted stem-loop structure of a BOX element (13). The detection of a nucleotide duplication

accompanying BOX element insertion led us to hypothesize that a BOX element could be derived from a mobile genetic element, as previously proposed for RUP (16). We therefore searched the pneumococcal genome for a candidate IS with inverted terminal repeats (IRs) homologous to BOX element extremities. We identified a new putative IS of *S. pneumoniae*, ISS*Spn2* (Fig. 4B). ISS*Spn2* is a 1,091-bp element (which has been deposited in the ISFinder database [http://www-is.biotoul.fr]) which harbors potential IRs (Fig. 4B) and has a calculated G+C content close to that of *S. pneumoniae* (37.7%). It appears to contain two potential open reading frames, *orf1* and *orf2* (Fig. 4B). A BLASTN search of streptococcal genomes carried out with the ISS*Spn2* sequence revealed the presence of an iso-IS only in *Streptococcus sobrinus* (70% identity in the last 825 nucleotides). This 1,095-bp element (G+C content, 39%), designated ISS*tso1*, is organized like ISS*Spn2*. A BLASTP search carried out with an ISS*tso1 orf1-orf2* (arbitrary) fusion product revealed that the protein belongs to the IS630 family of transposases, and the closest homologue is the transposase of ISC1048 from *Sulfolobus solfataricus* P2 (data not shown). A transposase sequence similarly assembled from ISS*Spn2 orf1-orf2* sequences exhibited similar clustering (~28% identity between ISS*Spn2* and ISC1048 transposases [see Fig. S1 in the supplemental material]). The comparison of the IRs of ISS*Spn2* with BOX element extremities revealed significant homology only between the left inverted repeat and boxC (Fig. 4C). However, significant homology was detected between both IRs of ISS*tso1* and boxA and boxB extremities (Fig. 4C). Interestingly, the level of identity between both BOX element extremities and the IRs of ISS*tso1* was very similar to the level of identity between the left and right inverted repeats of the IS. As these IRs are presumably both recognized by the putative ISS*tso1* transposase, this finding strongly suggests that the transposase readily recognizes the termini of the BOX element as well. These observations led us to propose that the BOX element could be transactivated by the ISS*tso1* transposase and possibly also by ISS*Spn2*, similar to the RUP elements in *S. pneumoniae* (16) and the Correia elements in *Neisseria* (11). We propose that through their mobilization BOX elements have contributed to shaping the genomes and transcriptomes of *S. pneumoniae* and *S. mitis* and could even have played a role in the evolution of *S. pneumoniae* as a human pathogen.

This work was supported by grants from the Research Council of Norway.

REFERENCES

- Chastanet, A., M. Prudhomme, J. P. Claverys, and T. Msadek. 2001. Regulation of *Streptococcus pneumoniae* *clp* genes and their role in competence development and stress survival. *J. Bacteriol.* **183**:7295–7307.
- Claverys, J. P., and L. S. Håvarstein. 2002. Extracellular-peptide control of competence for genetic transformation in *Streptococcus pneumoniae*. *Front. Biosci.* **7**:d1798–d1814.
- Condon, C. 2003. RNA processing and degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* **67**:157–174.
- Deutscher, M. P. 2006. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.* **34**:659–666.
- Håvarstein, L. S., G. Coomaraswami, 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.
- 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.
- Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszcak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. Leblanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenry, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rostek, Jr., P. L. Skatrud, and J. I. Glass. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* **183**:5709–5717.
- Hui, F. M., and D. A. Morrison. 1991. Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. *J. Bacteriol.* **173**:372–381.
- Khemic, V., and A. J. Carpousis. 2004. The RNA degradosome and poly(A) polymerase of *Escherichia coli* are required *in vivo* for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Mol. Microbiol.* **51**:777–790.
- Knutson, E., O. Ween, and L. S. Håvarstein. 2004. Two separate quorum-sensing systems upregulate transcription of the same ABC transporter in *Streptococcus pneumoniae*. *J. Bacteriol.* **186**:3078–3085.
- Liu, S. V., N. J. Saunders, A. Jeffries, and R. F. Rest. 2002. Genome analysis and strain comparison of correia repeats and correia repeat-enclosed elements in pathogenic *Neisseria*. *J. Bacteriol.* **184**:6163–6173.
- Mahillon, J., and M. Chandler. 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**:725–774.
- Martin, B., O. Humbert, M. Camara, E. Guenzi, J. Walker, T. Mitchell, P. Andrew, M. Prudhomme, G. Alloing, R. Hakenbeck, D. A. Morrison, G. J. Boulnois, and J. P. Claverys. 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. *Nucleic Acids Res.* **20**:3479–3483.
- Martin, B., M. Prudhomme, G. Alloing, C. Granadel, and J. P. Claverys. 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. *Mol. Microbiol.* **38**:867–878.
- Mrázek, J., L. H. Gaynon, and S. Karlin. 2002. Frequent oligonucleotide motifs in genomes of three streptococci. *Nucleic Acids Res.* **30**:4216–4221.
- Oggioni, M. R., and J. P. Claverys. 1999. Repeated extragenic sequences in procaryotic genomes: a proposal for the origin and dynamics of the RUP element in *Streptococcus pneumoniae*. *Microbiology* **145**:2647–2653.
- 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.
- Prudhomme, M., and J. P. Claverys. There will be a light: the use of *luc* transcriptional fusions in living pneumococcal cells. In R. Hakenbeck and G. S. Chhatwal (ed.), *The molecular biology of streptococci*, in press. Horizon Scientific Press, Wymondham, United Kingdom.
- Sharp, J. S., and D. H. Bechhofer. 2005. Effect of 5'-proximal elements on decay of a model mRNA in *Bacillus subtilis*. *Mol. Microbiol.* **57**:484–495.
- Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**:498–506.
- Ween, O., P. Gaustad, and L. S. Håvarstein. 1999. Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. *Mol. Microbiol.* **33**:817–827.