

# A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*

Bernard Martin, Odile Humbert, Miguel Camara<sup>1</sup>, Eric Guenzi<sup>2</sup>, John Walker<sup>1</sup>, Timothy Mitchell<sup>1</sup>, Peter Andrew<sup>1</sup>, Marc Prudhomme, Geneviève Alloing, Regine Hakenbeck<sup>2</sup>, Donald A. Morrison<sup>3</sup>, Graham J. Boulnois<sup>1</sup> and Jean-Pierre Claverys\*

Microbiologie et Génétique Moléculaires, CNRS-UPR 9007, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France, <sup>1</sup>Department of Microbiology, University of Leicester, PO Box 138, Medical Sciences Building, University Road, Leicester LE1 9HN, UK, <sup>2</sup>Max-Planck Institut für Molekulare Genetik, Ihnestr. 73, D-1000 Berlin 33, Germany and <sup>3</sup>Laboratory for Molecular Biology, Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL 60680, USA

Received February 10, 1992; Revised and Accepted June 12, 1992

EMBL accession no. X63602

## ABSTRACT

**We report the discovery of a group of highly conserved DNA sequences located, in those cases studied, within intergenic regions of the chromosome of the Gram positive *Streptococcus pneumoniae*. The *S. pneumoniae* genome contains about 25 of these elements called BOX. From 5' to 3', BOX elements are composed of three subunits (boxA, boxB, and boxC) which are 59, 45 and 50 nucleotides long, respectively. BOX elements containing one, two and four copies of boxB have been observed; boxB alone was also detected in one instance. These elements are unrelated to the two most thoroughly documented families of repetitive DNA sequences present in the genomes of enterobacteria. BOX sequences have the potential to form stable stem-loop structures and one of these, at least, is transcribed. Most of these elements are located in the immediate vicinity of genes whose product has been implicated at some stage in the process of genetic transformation or in virulence of *S. pneumoniae*. This location raises the intriguing possibility that BOX sequences are regulatory elements shared by several coordinately controlled genes, including competence-specific and virulence-related genes.**

## INTRODUCTION

Highly conserved, repeated DNA sequences constitute an important fraction of eukaryotic genomes. By comparison, the amount of repetitive DNA in prokaryotic genomes appears considerably lower although significant. The Palindromic Units (PU) (1), also called Repetitive Extragenic Palindromes (REP) (2), constitute the best characterized family of bacterial repetitive sequences. PU are present in about 500–1000 copies in the chromosome of *Escherichia coli* and of *Salmonella typhimurium*

(1, 2). PU sequences consist of a 35–40 bp inverted repeat and are found in clusters in which successive copies (up to six) are arranged in alternate orientation. A second family of repetitive elements, called IRU (Intergenic Repeat Units) (3) or ERIC (Enterobacterial Repetitive Intergenic Consensus) (4), has been described. IRU are 124–127 bp long and are present in about 30–50 copies in *E. coli* and 150 copies in *S. typhimurium* (3, 4). Although they resemble PU by several features, the nucleotide sequence is entirely different and IRU, contrary to PU, appear to occur singly. Both the PU and the IRU families are similarly located in non-coding, probably transcribed, regions of the chromosome and include a potential stem-loop structure.

Several roles have been reported for individual examples of PU elements. Binding of gyrase (5), of DNA polymerase I (6), and of an unidentified complex of five proteins (6) to DNA containing PU sequences supported the notion that these elements are involved in nucleoid folding (5, 2, 7). Effects on gene expression through stabilization of upstream mRNA (8, 9), on translational coupling (10), and on transcription termination (11) were also documented. Since no single function has been identified that could account for the conservation of PU sequences, the question of whether they have a specific function or represent a form of 'selfish' DNA propagating by gene conversion (12) was raised and is still open. Recently, it has been shown that each PU cluster contains, besides the PU themselves, a number of other conserved sequence motifs with a mosaic pattern of motif combination, hence the name of Bacterial Interspersed Mosaic Element (BIME) (13). It was suggested that BIME could belong to several functional classes depending on the nature of these other motifs (*ie* specific functions could be associated with specific motif combinations).

Apart from transposons and insertion sequences, to date none of the reports of repetitive sequences in prokaryotes concern Gram positive bacteria (7, 14). In a recent study of the distribution

\* To whom correspondence should be addressed

of PU (REP) and IRU (ERIC) sequences in eubacteria, specific probes were shown to hybridize preferentially to genomic DNA from Gram negative enteric bacteria and related species, but several species of Gram positive bacteria, including *Streptococcus pneumoniae*, failed to yield hybridization signals (14).

In this paper we present the first description of a highly conserved DNA element located, in those cases studied, within intergenic regions of the chromosome of *S. pneumoniae*. We discuss the possibility that this element has been recruited to play a regulatory role for several coordinately controlled genes.

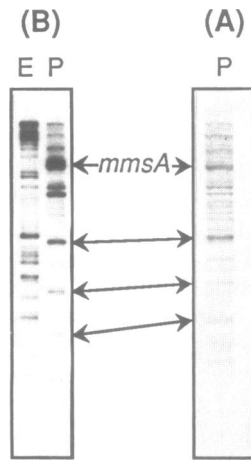
**RESULTS AND DISCUSSION**

**A repeated sequence in the genome of *S. pneumoniae***

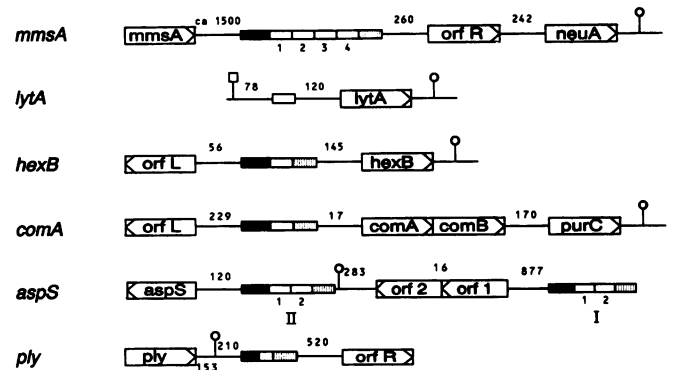
In the course of the study of a cloned 6 kb *Bcl*I fragment carrying the *mmsA* gene of *S. pneumoniae* which is involved in DNA repair and recombination, it was observed that this fragment hybridizes with a number of fragments from a total chromosomal digest, in addition to those fragments which cover the *mmsA* region itself (Martin, Humbert and Claverys, unpublished data). The region responsible for these multiple hybridizations was mapped to a *Hpa*I-*Pvu*II 340 bp fragment located ca 1.5 kb

downstream from the *mmsA* gene. When used as a probe, this 340 bp fragment hybridized to ca 25 different chromosomal fragments (Fig. 1A).

DNA sequencing of this 340 bp fragment revealed the presence of four direct repeats of a 45 bp segment (boxB in Fig. 2). A computer search of the EMBL DNA sequence database (release



**Figure 1.** Hybridization of *S. pneumoniae* strain R800 (33) genomic DNA with the BOX-*mmsA* fragment probe (A) or the boxB consensus oligonucleotide probe (B). 1 µg *Eco*RI (E) or *Pst*I (P) digested chromosomal DNA, electrophoresed on 1% agarose gels, was transferred (34) on Biotrans B (PALL Industrie s.a.) membranes. (A): The *Hpa*I-*Pvu*II 340 bp *mmsA* fragment probe was uniformly digoxigenin-labeled according to the manufacturer (Boehringer Mannheim) and used at 10 ng/ml in the hybridization solution. Hybridization was performed in 5×SSC-50% formamide-0.1% N-lauroylsarcosine-0.02% SDS-5% blocking reagent (Boehringer Mannheim) at 30°C for 18 hrs. Membranes were washed at room temperature twice for 5 min with 2×SSC-0.1% SDS, then twice for 15 min with 0.1×SSC-0.1% SDS. Digoxigenin nonradioactive detection was performed as recommended by the supplier (Boehringer Mannheim). (B): 2.5 pmol of the 22-base boxB oligonucleotide [sequence as follows: 5'-TAT(C/T)(T/C)ACA(A/G)(C/T)(C/G)TCAAAA(C/T)AGTG-3'] was end-labeled with 2.5 U T4 polynucleotide kinase (Bethesda Research Laboratory) and 2.5 pmol [ $\alpha$ -<sup>32</sup>P]ATP (5000 Ci/mmol; Amersham) as described by Maniatis et al. (34). Hybridization was performed in 6×SSC-10×Denhardt's reagent (34)-0.1% SDS-1% blocking reagent (Boehringer Mannheim) at 40°C for 20 hrs. Membranes were washed at room temperature twice for 5 min. with 6×SSC-0.1% SDS. Autoradiograms were exposed on Trimax (3M) film with an intensifying screen at -75°C for 45 hrs. In both (A) and (B), membranes were prehybridized for 1.5 hrs at the hybridization temperature.



**Figure 2.** Schematic diagram (not to scale) showing the organization of BOX sequences and their position relative to identified flanking genes or orfs and their transcriptional signals. BoxA, boxB, and boxC are indicated by black-, open-, and shaded-rectangles, respectively. Each copy of boxB at the same locus is numbered (below the line) for identification in Fig. 3. Leftside and rightside BOX elements in the *aspS* fragment are referred to in Fig. 3 as II and I, respectively. Coding or putative coding regions and their direction of transcription are denoted by open rectangles with an arrow. Vertical bars with an open circle on top denote a proposed rho-independent terminator. The *lytA* promoter is indicated by a vertical bar with an open square. Distances between coding regions, elements, and transcription initiation and termination proposed signals are indicated in bp above each line.



**Figure 3.** Alignment of BOX sequences. Sequences of boxA (top), boxB (middle), and boxC (bottom) which are presented separately for convenience are directly adjacent at each locus. Sequences are labelled as in Fig. 2 and in Table 1. Nucleotides are shown in lower case where they do not match the consensus. The consensus (cons.) is based on the nucleotides present in at least 50% of the sequences. Nucleotides present in all sequences are boxed. The *Asu*II site is underlined in the consensus.

17; all entries) using the FASTA program (15) (with a word size of 6) identified a unique sequence that showed a high degree of similarity to this 45 bp repeat. This sequence is located on the *lytA* fragment of *S. pneumoniae*, immediately upstream of the *lytA* gene. A similar sequence was further identified in the center of a 150 bp region of high homology between *hexB* and *comA* upstream regions. Comparison of DNA sequences surrounding the 45 bp repeats in the *mmsA* fragment with that in *hexB* and *comA* revealed that both 5' (boxA, 59 bp) and 3' (boxC, 50 bp) flanking regions were also highly conserved. The whole element (including some combination of boxA, boxB, and/or boxC) was named BOX. A region of the *comA* fragment has been reported to hybridize to 6 to 10 different fragments of the *S. pneumoniae* chromosome (16). The multiple hybridizing sequence (MHS) was tentatively located to the left of the *AsuII* site upstream of the *comA* gene (17). Our observations suggest that this MHS is indeed BOX and that it is located immediately to the right of this *AsuII* site which is part of boxA (see Fig. 3).

A DNA sequence containing four direct repeats of 45 bp was also detected in the course of the sequencing of the *neuA* gene (Camara and Boulnois, unpublished data) which encodes one of the two neuraminidase genes (18). Direct sequence comparison revealed that the *mmsA* fragment and the *neuA* fragment were in fact overlapping (Fig. 2). Similar sequences were further identified on a cloned fragment possibly containing a penicillin-resistance determinant (Guenzi and Hakenbeck, unpublished results). This fragment, *aspS*, harbors two copies of BOX that are ca 2.5 kb apart, each of them containing a tandem repeat of boxB (Fig. 2). Finally, a BOX element was also identified 386bp downstream of the *ply* gene (Walker, Mitchell and Boulnois, unpublished data) which encodes pneumolysin (19). The organization of BOX sequences and their position relative to the flanking genes are shown in Fig. 2.

#### Definition of the consensus and number of BOX elements

The existence of BOX elements containing either boxB alone (*lytA* locus) or a variable number of tandem repeats of boxB (*mmsA* and *aspS* fragments) suggest that boxB is the basic element around which the modular structure of BOX is organized. The alignment of all BOX sequences (Fig. 3) revealed a very high degree of sequence similarity for each subunit (above 90% for the entire BOX element, Table 1). The consensus BOX element is constituted, from 5' to 3', of three subunits, boxA (59 bp), boxB (45 bp), and boxC (50 bp), and is thus 154 bp long.

A 22-base oligonucleotide based on the consensus for boxB was used as a probe in hybridization experiments. The hybridizing sequences were found distributed on about 25 fragments of the *S. pneumoniae* chromosome (Fig. 1B), in agreement with results obtained with the shortest *mmsA* fragment probe (Fig. 1A).

#### Potential secondary structure of BOX sequences

Examination of the consensus 154-bp BOX element revealed several inverted repeat sequences that could lead to formation of a hairpin in DNA or a stable stem-loop in a putative RNA transcript of this molecule. The proposed secondary structure for the consensus BOX element is shown in Fig. 4. A stem-loop structure could be formed between boxA and boxC for all BOX elements. The structure of the central part of the element varies depending on the number of repeats, two or four, of boxB (data not shown).  $\Delta G$  values for these putative stem-loop structures are shown in Table 1. Significantly, in several cases where BOXs differ from the consensus, a compensating base change is seen

in the complementary arm of the predicted stem-loop (see Fig. 4), implying that maintenance of the secondary structure is important. For example, of the eleven nucleotides in the *comA* BOX that depart from the consensus, four maintain the consensus structure (Fig. 4) and five increase helix stability, by addition of one (bottom of helix VI) or three (top of helix II) base pairs. In both *aspS* BOX elements, a single nucleotide change results in the presence of two additional base pairs at the top of helix V. Also possibly significant is the fact that all the sequences show a high

Table 1. Location and properties of BOX sequences

Locus (min.)	Accession number	Position (bp)	Percent identity	Free energy	References
<i>comA</i> (7')	M36180	344-496	92.9	-43.0	(37)
<i>hexB</i> (10-11')	M29686	201-353	90.9	-24.4	(22)
<i>lytA</i> (21-24')	M13812	38-84	-	-	(38)
<i>mmsA</i> (24-26')	X63602	3-291	93.8	-71.1	a
<i>aspSI</i> (0-6')	-	-	93.5 (-12 bp)	-47.6	b
<i>aspSII</i> (0-6')	-	-	95.5	-51.3	b
<i>ply</i>	-	-	92.2 (-38 bp)	-24.4	c

Where known, the map location in minutes (min.) of the locus is indicated. For convenience, the circular chromosomal map of *S. pneumoniae* (29) has been arbitrarily divided in 60 min., with the 0/60 position at the top and a clockwise increase. The percentage of identity to the consensus (as defined in Fig. 3) was calculated from the sequences in Fig. 3. The deletion in *aspSI* and *ply* BOX elements which are indicated between parentheses, were not taken into account in this calculation. Free energy ( $\Delta G$ ) values ( $\text{kcal mol}^{-1}$ ) were calculated using the estimations of Turner et al. (36). Reference a: Martin, Humbert and Claverys (unpublished data); b: Guenzi and Hakenbeck (unpublished data); reference c: Walker, Mitchell and Boulnois (unpublished data).

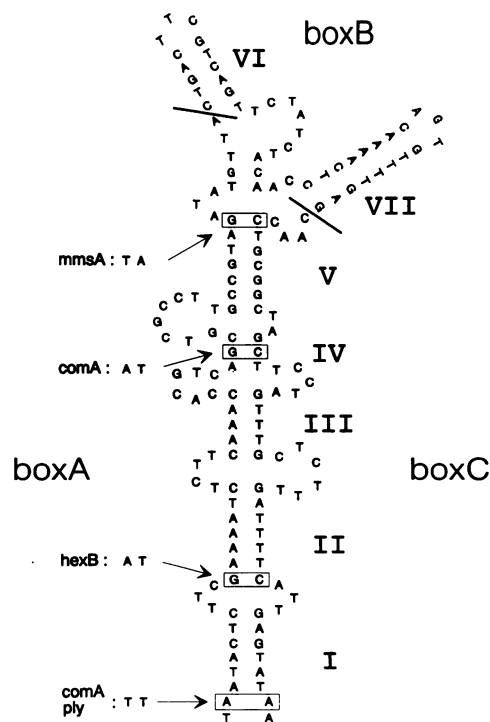


Figure 4. Predicted stem-loop structure of the consensus BOX sequence (as defined in Fig. 3) containing a single copy of boxB.  $\Delta G = -42.9 \text{ kcal mol}^{-1}$ . Secondary structure predictions were performed using the Zucker program (35) with the energies of Turner et al. (36). Arrows mark the location of compensating base changes maintaining the structure in the indicated sequence. Helices are numbered I to VI for reference. Straight lines separate boxA from boxB and boxC.

percentage (above 70%) of pyrimidines (or purines depending on the strand considered) in the single-stranded loop regions.

### BOX elements and flanking genes

The location of BOXs relative to neighboring genes is shown in Fig. 2. BOXs are located outside these genes, a not unexpected situation because of the constraints they would impose on protein structure, as pointed out for other highly repeated sequences (4). Comparison of the orientation of BOX elements on the *ply* and *aspS* fragments (Fig. 2) reveals that BOX can have either orientation with respect to flanking orfs. This implies that, if BOX acts at the mRNA level, what is important for BOX function (if any) is secondary structure rather than primary sequence.

In the *lytA* locus where transcripts have been mapped (20), BOX is located between the promoter and the structural gene. This is the only case in which the exact location of the transcription start point is known. Although the evidence is indirect, it is likely that the BOX element upstream of the *comA* gene is also transcribed: the spacing between the 3' end of the BOX element and the start codon of *comA* is 17 bp only which strongly suggests that the *comA* promoter is located either within or upstream of the BOX element. In addition, insertion of heterologous DNA at the *AsuII* site within boxA (Fig. 3) confers a Com<sup>-</sup> phenotype (16), which could mean that signals important for *comA* gene expression are located upstream of the BOX element.

### BOX, a regulatory element for coordinately controlled genes?

In search for a possible specific function of BOX elements, we noticed that several BOX elements are located in the vicinity of genes whose function is related to the process of genetic transformation. These include the *comA*, *hexB*, and possibly *mmsA* genes. The *comA* gene controls competence for genetic transformation (16). Competence is a transitory cell state accompanied by a global shift of protein synthesis to the production of a few competence-specific proteins (21). The *hexB* gene encodes an essential component of the Hex mismatch repair system (22) which eliminates base pair mismatches arising in heteroduplex DNA during transformation (23). Expression of the *mmsA* gene during competence is also not unlikely since this gene is involved in recombination.

Most of the other identified BOX elements are located close to genes possibly involved in the virulence of *S. pneumoniae*. These include the *neuA* gene which encodes neuraminidase, a putative pneumococcal virulence factor (18) and the *ply* gene which encodes pneumolysin, another putative virulence factor (19). Finally, the *lytA* gene function could be related to both transformation and virulence. First, it is known that cells developing competence exhibit increased susceptibility to lysis in certain buffers (24). This may result from increased *lytA* gene expression at the onset of competence. Second, inactivation of the *lytA* gene has been reported to severely reduce virulence of pneumococcus (25).

The locations of BOX elements in the immediate vicinity of genes involved in competence or in virulence raises the intriguing possibility of a regulatory connection between transformation and virulence. It is worth mentioning that the development of competence for genetic transformation can be modulated by changes in growth conditions such as temperature (26), or the initial pH of the culture (27), or magnesium and calcium concentrations (26). Although the possible regulation of virulence of *S. pneumoniae* is not documented, expression of virulence

determinants of several pathogenic bacteria is known to be controlled by environmental signals, including temperature, pH, and calcium (28). This leads us to speculate that BOX elements could be involved in coordinated expression of competence-specific and virulence-related genes. Both transformation and virulence would then be viewed as a global response of *S. pneumoniae* to changing environmental conditions.

If BOX elements have such a regulatory function, their modular organization with a variable number of boxB subunits immediately suggest the possibility that a protein binds to this subunit and that modulation of binding depends on the number of repeats of boxB. Should binding occur at the DNA level, the observed bias in base composition of the single stranded regions (see above) could constitute an important feature of BOX elements. Demonstration of the involvement of BOX elements in regulation of specific genes obviously requires deletion of some of these elements to study how it affects gene expression.

### Relation to other repeated DNA sequences in bacteria

BOX sequences are the first group of highly conserved repetitive DNA elements found in a Gram positive bacterium. Although the nucleotide sequence is entirely different, several characteristic of BOX are similar to those of the IRU (ERIC) and of the PU (REP) elements of enterobacteria (see INTRODUCTION). With respect to their size and estimated number of copies, BOX sequences are more similar to IRU than to PU. However, unlike IRU, they have a modular structure. Noncoding repetitive DNA is likely to be kept to a minimum in *S. pneumoniae* given its small genome size as compared to that of *E. coli* (2270 kb [29] and 4700 kb [30], respectively). Unless BOX sequences maintain themselves as 'selfish' DNA, a functional role of BOX as regulatory elements would explain their persistence and sequence conservation in *S. pneumoniae*. One may speculate that BOX represents a case of evolution of a repetitive DNA element recruited for a specific function.

Finally, it would be interesting to see if BOX elements are widespread among streptococcal species, as are IRU and PU in enterobacteria. BOXs could then be used as a phylogenetic index for oral streptococci, a group of bacteria the taxonomy of which is far from being well established (31, 32).

### ACKNOWLEDGEMENTS

We thank Anne-Marie Gasc for mapping the *aspS* and *mmsA* loci on the chromosome of *S. pneumoniae*. Work in Chicago was supported in part by a grant from the U.S. National Science Foundation Genetic Biology Program, in Leicester by grants from the Wellcome Trust, and in Toulouse by a Contrat de Recherche Externe 861005 from the Institut National de la Santé et de la Recherche Médicale. GJB is a Lister-Jenner Fellow, PA is a Nuffield Fellow and MC acknowledges support from the Spanish Ministry of Education (MEC). EG was the recipient of a fellowship from the European Community (SC1\*/0141-C).

### REFERENCES

1. Gilson, E., Clément, J.M., Brutlag, D. and Hofnung, M. (1984) *EMBO J.*, 3, 1417-1421.
2. Stern, M.J., Ames, G.F.-L., Smith, N.H., Robinson, E.C. and Higgins, C.F. (1984) *Cell*, 37, 1015-1026.
3. Sharples, G.J. and Lloyd, R.G. (1990) *Nucl. Acids Res.*, 18, 6503-6508.
4. Hulton, C.S.J., Higgins, C.F. and Sharp, P.M. (1991) *Mol. Microbiol.*, 5, 825-834.

5. Yang, Y. and Ames, G.F.-L. (1988) *Proc. Natl. Acad. Sci. USA*, 85, 8850–8854.
6. Gilson, E., Perrin, D. and Hofnung, M. (1990) *Nucl. Acids Res.*, 18, 3941–3952.
7. Gilson, E., Clément, J.M., Perrin, D. and Hofnung, M. (1987) *Trends Genet.*, 3, 226–230.
8. Newbury, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D. and Higgins, C.F. (1987) *Cell*, 48, 297–310.
9. Newbury, S.F., Smith, N.H. and Higgins, C.F. (1987) *Cell*, 51, 1131–1143.
10. Stern, M.J., Prossnitz, E. and Ames, G.F.-L. (1988) *Mol. Microbiol.*, 2, 141–152.
11. Gilson, E., Rousset, J.P., Clément, J.M. and Hofnung, M. (1986) *Ann. Inst. Pasteur Microbiol.*, 137B, 259–270.
12. Higgins, C.F., McLaren, R.S. and Newbury, S.F. (1988) *Gene*, 72, 3–14.
13. Gilson, E., Saurin, W., Perrin, D., Bachellier, S. and Hofnung, M. (1991) *Nucl. Acids Res.*, 19, 1375–1383.
14. Versalovic, J., Thearith, K. and Lupski, J.R. (1991) *Nucl. Acids Res.*, 19, 6823–6831.
15. Pearson, W.R. and Lipman, D.J. (1988) *Proc. Natl. Acad. Sci. USA*, 85, 2444–2448.
16. Chandler, M.S. and Morrison, D.A. (1988) *J. Bacteriol.*, 170, 3136–3141.
17. Morrison, D.A. (1989) In Butler, L.O., Harwood, C. and Moseley, B.E.B. (eds.), *Genetic Transformation and Expression*. Intercept Limited, Andover, Hants, pp. 3–11.
18. Camara, M., Mitchell, T.J., Andrew, P.W. and Boulnois, G.J. (1991) *Infect. Immun.*, 59, 2856–2858.
19. Walker, J.A., Allen, R.L., Falmagne, P., Johnson, M.K. and Boulnois, G.J. (1987) *Infect. Immun.*, 55, 1184–1189.
20. Diaz, E. and Garcia, J.L. (1990) *Gene*, 90, 157–162.
21. Morrison, D.A. and Baker, M. (1979) *Nature*, 282, 215–217.
22. Prudhomme, M., Martin, B., Méjean, V. and Claverys, J.P. (1989) *J. Bacteriol.*, 171, 5332–5338.
23. Claverys, J.P. and Lacks, S.A. (1986) *Microbiol. Rev.*, 50, 133–165.
24. Seto, H. and Tomasz, A. (1975) *J. Bacteriol.*, 121, 344–353.
25. Berry, A.M., Lock, R.A., Hansman, D. and Paton, J.C. (1989) *Infect. Immun.*, 57, 2324–2330.
26. Lacks, S.A. and Greenberg, B. (1973) *J. Bacteriol.*, 114, 152–163.
27. Chen, J.D. and Morrison, D.A. (1987) *J. Gen. Microbiol.*, 133, 1959–1967.
28. Mekalanos, J.J. (1992) *J. Bacteriol.*, 174, 1–7.
29. Gasc, A.M., Kauc, L., Baraillé, P., Sicard, M. and Goodgal, S. (1991) *J. Bacteriol.*, 173, 7361–7367.
30. Kohara, Y., Akiyama, K. and Isono, K. (1987) *Cell*, 50, 495–508.
31. Coykendall, A.L. (1989) *Clin. Microbiol. Rev.*, 2, 315–328.
32. Kilian, M., Mikkelsen, L. and Henrichsen, J. (1989) *Int. J. Sys. Bact.*, 39, 471–484.
33. Lefèvre, J.C., Claverys, J.P. and Sicard, A.M. (1979) *J. Bacteriol.*, 138, 80–86.
34. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
35. Zuker, M. and Steigler, P. (1980) *Nucl. Acids Res.*, 9, 133–148.
36. Turner, D.H., Sugimoto, N., Jaeger, J.A., Longfellow, C.E., Freier, S.M. and Kierzek, R. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, 52, 123–133.
37. Hui, F.M. and Morrison, D.A. (1991) *J. Bacteriol.*, 173, 372–381.
38. Garcia, P., Garcia, J.L., Garcia, E. and Lopez, R. (1986) *Gene*, 43, 265–272.