JOURNAL OF BACTERIOLOGY, Sept. 1997, p. 5632–5635 Vol. 179, No. 17
0021-9193/97/$04.00+0

Identification and Characterization of Cell Wall-Cell Division Gene Clusters in Pathogenic Gram-Positive Cocci

MICHAEL J. PUCCI,* JANE A. THANASSI, LINDA F. DISCOTTO, ROBERT E. KESSLER, AND THOMAS J. DOUGHERTY

Department of Microbiology, Bristol-Myers Squibb Pharmaceutical Research Institute, Wallingford, Connecticut

Received 28 March 1997/Accepted 26 June 1997

Clusters of peptidoglycan biosynthesis and cell division genes (DCW genes) were identified and sequenced in two gram-positive cocci, *Staphylococcus aureus* and *Enterococcus faecalis*. The results indicated some similarities in organization compared with previously reported bacterial DCW gene clusters, including the presence of penicillin-binding proteins at the left ends and *ftsA* and *ftsZ* cell division genes at the right ends of the clusters. However, there were also some important differences, including the absence of several genes, the comparative sizes of the *div1B* and *ftsQ* genes, and a wide range of amino acid sequence similarities when the genes of the gram-positive cocci were translated and compared to bacterial homologs.

Biosynthesis of peptidoglycan in bacteria is a complex process involving numerous enzymes, most of which have been shown to be essential in pathogenic bacteria (3, 16). These enzymes have been studied most extensively in *Escherichia coli*, and the genes which encode them have been cloned and the proteins have been characterized over the past several years (29, 30). Many of these genes had been previously mapped in this organism due to the availability of temperature-sensitive lethal mutants. In *E. coli*, it was discovered that a number of these genes, including the one encoding penicillin-binding protein 3 (PBP3), were organized in a cluster located at 2 min on the chromosomal map (29, 34). Also present in this cluster were other genes important in cell division, including *ftsA* and *ftsZ* (4, 19, 31). All of these genes were found to be tightly packed, with reading frames sometimes overlapping, all oriented in the same direction of transcription, and there has been speculation about possible regulation of cellular growth and division within this region. Peptidoglycan biosynthetic genes were not exclusively in this 2-min region, however, as other genes, such as *murA* (69.3 min; 2), *murB* (89.9 min; 23), and *murI* (*dga*; 89.8 min; 10, 12), were found elsewhere on the chromosome.

A cluster similar to the 2-min region of *E. coli* was found in the gram-positive rod *Bacillus subtilis* (6, 27). Interestingly, with a few exceptions (for example, sporulation-specific genes), the gene arrangement was quite similar to that seen in *E. coli*, including the tight arrangement of genes and a similar direction of transcription (7, 20). However, missing from this region and located elsewhere on the chromosome were the *murC*, *murF*, and *ddl* genes found in the *E. coli* cluster. Recently, the genome sequence of another gram-negative rod, *Haemophilus influenzae*, was reported and the gene order in this region was found to be identical to that of *E. coli* (13). These findings suggest a possible evolutionary relationship for genes involved in peptidoglycan biosynthesis and cell division.

The existence of similar gene clusters in gram-positive cocci has not been reported. This work describes clusters of peptidoglycan biosynthetic and cell division genes in *Staphylococcus aureus* and *Enterococcus faecalis*, and the results presented show both similarities to the genetic organization seen in *E. coli*, *B. subtilis*, and *H. influenzae* and some significant differences. In contrast to the gram-negative rods *E. coli* and *H. influenzae* and the gram-positive rod *B. subtilis*, the gram-positive cocci in this study contained fewer genes in these cell wall-cell division or DCW clusters (7, 34). Clearly, several essential genes found in the DCW clusters of the gram-positive and gram-negative rods are located elsewhere in the chromosomes of the gram-positive cocci.

Cloning and sequencing of DCW gene clusters in *S. aureus* and *E. faecalis*. When this work was initiated, there was little data available concerning the existence of DCW gene clusters in gram-positive cocci analogous to the major ones reported for *E. coli*, *B. subtilis*, and *H. influenzae*. However, there was a DNA sequence entered in the GenBank database for the *ftsZ* gene, which is found in all of the above reported clusters, from *S. aureus* (1). Chromosomal libraries from *S. aureus* ATCC 8325-4 and *E. faecalis* A24836 (Bristol-Myers Squibb Culture Collection) were constructed in pRDD40 (pLC4 shuttle vector from which the promoterless *xylE* gene has been removed; 25). Restriction endonuclease *Tsp*509I (AATT recognition site; New England Biolabs, Beverly, Mass.) was used to partially digest the chromosomal DNA. Because of the low G+C content of the genomes (32 to 38 mol% G+C; 17, 21), the choice of this enzyme was critical for this work because it generated much more randomly cut fragments than other enzymes commonly used in library construction, such as *Sau*3AI (data not shown). DNA fragment sizes of 3 to 5 kb were selected by sucrose density gradient centrifugation as inserts for cloning into pRDD40, and four pools of plasmids from approximately 1,000 colonies (representing a total of about 4,000 colonies) constituted the genomic libraries. Specific primers were synthesized and used to isolate a 450-bp internal *S. aureus ftsZ* fragment by PCR for use as a probe against the *S. aureus* library, while degenerate primers FOR (5′GGTATGGGNGG NGGWACNGGWACWGGNGCNGCACCNGT3′) and REV (5′AAANCCNGTTGCAATNACTGTNACNACAATC TGATC3′) were used to obtain a 600-bp internal *ftsZ* fragment by PCR for use as a probe against the *E. faecalis* library. Clones carrying plasmids with inserts containing the *ftsZ* gene and adjoining DNA were identified by colony hybridizations for each organism. Next, chromosome walking experiments were initiated. These were done first by using PCR primers from known chromosomal sequences and the pRDD40 vector to

* Corresponding author. Mailing address: Department of Microbiology (Dept. 104), Bristol-Myers Squibb Pharmaceutical Research Institute, 5 Research Parkway, Wallingford, CT 06492. Phone: (203) 284-6372. E-mail: pucci@bms.com.

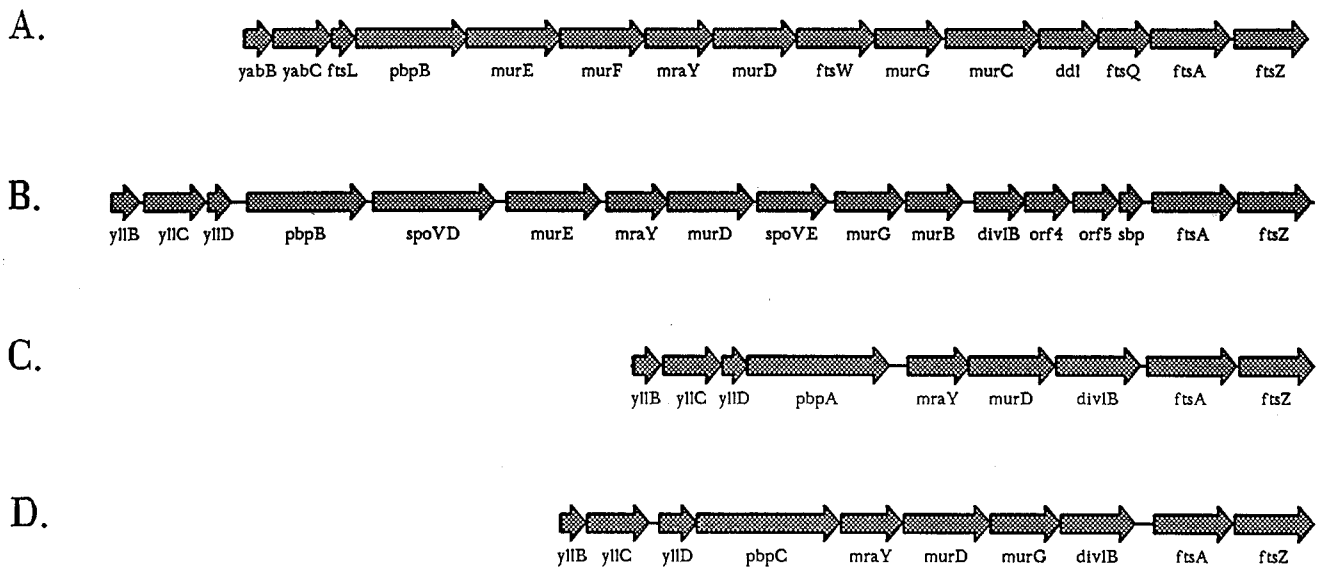

FIG. 1. Comparison of the DCW gene clusters of four bacterial species. A, *E. coli*; B, *B. subtilis*; C, *S. aureus*; D, *E. faecalis*. The arrows indicate ORFs and directions of transcription. Gene designations are under the ORFs. Unidentified ORFs are designated orfs.

obtain sequences from library inserts which contained additional adjacent DNA. Subsequently, the use of the GenomeWalker Systems (Clontech, Palo Alto, Calif.) allowed efficient identification and sequencing of adjacent DNA regions from PCR fragments generated directly from genomic DNAs. A total of 12.1 or 14.0 kb of DNA was sequenced from *S. aureus* or *E. faecalis*, respectively. Assignments of putative functions were made from the results of BLASTX searches of translations of all six potential reading frames against the GenBank database.

Gene linkage verification by PCR. The possibility existed that cloned fragments could represent scrambled chromosomal fragments. To verify that the library clones represent the true order of the genes found in the chromosomes, the order and linkage of the genes were checked by PCR. Based on DNA sequence information, a series of specific primer pairs were designed, each originating in and pointed toward the identified adjacent gene. Since the exact size of the PCR fragment which should be obtained from the chromosome with the primer pairs could be calculated, it would be possible to determine if the cloned adjacent genes were indeed next to each other in the chromosome. The results of these PCR experiments indicated that, in every case, the predicted-size PCR fragment was obtained (data not shown), thus verifying the gene order present in the chromosome of each organism.

Characterizations and comparisons of open reading frames (ORFs). The results revealed gene clusters with some similarities to those reported for *E. coli*, *B. subtilis*, and *H. influenzae* (Fig. 1). A summary of the genes and gene products found in the DCW clusters of *S. aureus* and *E. faecalis* compared with the homologous gene products found in *E. coli* and *B. subtilis* is shown in Table 1. The right ends of both the *S. aureus* and the *E. faecalis* clusters contained genes in the order *div1B* (protein homology with *B. subtilis*; Table 2), *ftsA* (*E. coli* and *B. subtilis*), and *ftsZ* (*E. coli* and *B. subtilis*). The *div1B* gene from *B. subtilis* was previously reported to encode a protein displaying homology to the *E. coli ftsQ* gene product (15), which would give the same *ftsQ-ftsA-ftsZ* gene order for both gram-

TABLE 1. Comparison of bacterial DCW cluster gene products

Gene(s)	No. of amino acids (molecular mass [Da])			
	E. coli	*B. subtilis*	*S. aureus*	*E. faecalis*
yllB[a]	152 (17,400)	147 (17,066)	145 (17,404)	143 (16,374)
yllC[b]	313 (34,900)	325 (36,942)	323 (36,899)	319 (36,211)
yllD[c]	121 (13,627)	121 (13,570)	134 (15,333)	135 (15,171)
pbp[d]	588 (63,877)	716 (79,305)	744 (82,718)	742 (81,803)
mraY	360 (39,874)	324 (35,589)	330 (36,423)	321 (35,839)
murD	437 (46,842)	451 (49,650)	451 (50,648)	456 (49,866)
murG	355 (37,771)	363 (39,936)	ND[f] (ND)	363 (39,920)
ftsQ, div1B[e]	276 (31,434)	262 (29,273)	440 (50,215)	385 (43,710)
ftsA	420 (45,330)	440 (48,102)	472 (53,340)	413 (45,638)
ftsZ	383 (40,297)	382 (40,355)	391 (41,040)	412 (44,360)

[a] *yabB* in *E. coli.*
[b] *yabC* in *E. coli.*
[c] *ftsL* in *E. coli.*
[d] PBP1 in *S. aureus*, PBP3(?) in *E. faecalis*, PBP3 in *E. coli*, and PBP2B in *B. subtilis.*
[e] *ftsQ* in *E. coli* and *div1B* in *B. subtilis.*
[f] ND, not determined.

TABLE 2. Comparisons of DCW gene products

S. aureus protein(s)	Similarity index[a]		
	E. coli	*B. subtilis*	*E. faecalis*
YllB[b]	34.9	62.1	57.2
YllC[c]	41.1	64.9	62.9
YllD[d]	14.9	26.5	24.3
PbpA[e]	24.4	39.8	31.4
MraY	30.7	56.5	48.5
MurD	30.7	44.6	46.0
MurG	ND[g]	ND	ND
FtsQ, Div1B[f]	15.0	26.6	25.6
FtsA	23.3	26.0	23.3
FtsZ	43.9	69.1	67.6

[a] Percentages of identical amino acids determined by Lipman-Pearson protein alignment are shown.
[b] YabB in *E. coli.*
[c] YabC in *E. coli.*
[d] FtsL in *E. coli.*
[e] PBP1 in *S. aureus*, PBP3(?) in *E. faecalis*, PBP3 in *E. coli*, and PBP2B in *B. subtilis.*
[f] FtsQ in *E. coli* and Div1B in *B. subtilis.*
[g] ND, not determined.

positive organisms. In *B. subtilis*, there are three ORFs between *div1B* and *ftsA*. Although the *S. aureus* and *E. faecalis div1B* gene products displayed low similarities at the amino acid level with the *E. coli* FtsQ protein (~15%), they did show similarities of 25 to 30% with the *B. subtilis* Div1B amino acid sequence (Table 2).

Upstream in the DNA sequence (Fig. 1), *E. coli* contains the gene order *mraY-murD-ftsW-murG-murC-ddl. B. subtilis* has *mraY*, *murD*, and *murG* but no *ftsW*, *murC*, or *ddl* homolog in this region. Likewise, the two gram-positive cocci lack several of the listed *E. coli* genes. *E. faecalis* also lacks *ftsW*, *murC*, and *ddl* and was found to contain the gene sequence *mraY-murD-murG* adjacent to the *div1B* homolog without any intervening ORFs. *S. aureus* displayed the same organization, except that no *murG* homolog was found in this area. Farther upstream, *E. coli* has the gene order *pbpB-murE-murF* while *B. subtilis* was found to have *spoVD* and *murE* in this region. Neither *E. faecalis* nor *S. aureus* contained a *murE* equivalent (this gene would add L-lysine rather than diaminopimelic acid in the third position of the peptidoglycan pentapeptide in these gram-positive cocci) or a *murF* homolog in this region of the chromosome. Both organisms, however, did contain PBP-encoding genes in this area: *pbpA* encoding PBP1 in *S. aureus* (32) and a previously unreported PBP-encoding gene, designated *pbpC*, in *E. faecalis* encoding a protein of 82 kDa with a tentative assignment of PBP3 based only on the putative size of the gene product (33).

The intergenic spacing in the gram-positive coccus clusters more closely resembles that of *B. subtilis* (Fig. 1), where there can be >100 bp between adjacent genes. This is particularly evident between the *pbpA* and *mraY* genes in *S. aureus* (293 bp) and the *div1B* and *ftsA* genes in both *E. faecalis* and *S. aureus* (231 and 108 bp, respectively). This is in contrast to the organization in *E. coli*, where several of the reading frames of the genes overlap. In all of these homologous gene clusters, all of the genes have the same direction of transcription (depicted as left to right in Fig. 1).

Discussion. All six of the DCW clusters discussed above have been found to contain a PBP-encoding gene at one end, along with three homologous ORFs with unknown functions immediately upstream of the PBP-encoding genes. In *E. coli*, there is evidence that PBP3 in the DCW cluster plays a role in septum formation and cell division (28). Models have been proposed in which PBP3 is part of a membrane complex with cell division proteins such as *ftsA* and *ftsZ* to mediate septum formation and subsequent cell division (14, 31). When these PBP genes are used to search protein databases, the highest similarity scores obtained are with the other PBPs in DCW clusters. The presence of PBP genes in all of these DCW clusters may indicate similar functions in other bacteria, rods as well as cocci. The DNA sequence of the putative *E. faecalis* PBP3-encoding gene discovered in this work had not been previously reported, while the sequences of the genes encoding *S. aureus* PBP1 and *S. pneumoniae* PBP2x were previously entered into public databases. It is notable that with the exception of *ftsW*, all of the genes missing in the gram-positive cluster encode cytoplasmic precursor biosynthetic enzymes. Since most of these missing genes presumably encode functions essential in the gram-positive cocci, they must be present elsewhere on the chromosome. In that regard, there is evidence of at least one additional DCW cluster in *S. aureus* containing the *ftsW*, *murF*, and *ddlA* genes (24). In all of the organisms sequenced to date, the *ftsA-ftsZ* relationship appears to be conserved. There is evidence of coordinated regulation of these two genes to maintain a proper balance of expression levels (5, 8). Also, many of the genes in the *E. coli* DCW cluster were found to overlap adjacent genes, further fueling speculation of some type of coordinate regulation of peptidoglycan biosynthesis and/or cell division (9). However, unlike that of *E. coli*, analysis of the *B. subtilis* DCW cluster, as well as those of two gram-positive cocci, shows no widespread gene overlaps with intergenic spacing of >100 bp in many instances.

Another observation originating from this work involves the Div1B/FtsQ proteins. It appears that there are similarities and differences in these proteins based on whether the bacteria are rod or coccus shaped. Both cocci appear to have a longer hydrophilic region just preceding the largest hydrophobic region at the amino termini of these proteins (data not shown; 18, 22). This difference was particularly striking when the *E. coli* and *S. aureus* proteins were compared with N-terminal hydrophilic regions of about 30 and 150 amino acids, respectively. Although the functions of these Div1B/FtsQ proteins are largely unknown, it is possible that they play some role in shape determination in bacteria. Whether the extended N terminus is in some way involved in gram-positive coccal morphology remains to be determined.

Preliminary examination of the DCW cluster in *S. pneumoniae* (11) indicates that it is very similar to the same region in *S. pyogenes* (26). Both of these streptococci have *ftsA* and *ftsZ* at the distal end of the cluster and a PBP-encoding gene (PBP2x in pneumococci) and *mraY* at the proximal end, similar to the cases reported here. Interestingly, in both organisms, the *mraY* gene is followed by a gene which appears to be a DEAD box RNA helicase that is presumably unrelated to cell wall or cell division functions. Further characterization of this region is under way.

From the data presented here, certain evolutionary questions can be raised. The DCW clusters in *E. coli* and *H. influenzae* contain several additional genes absent in the equivalent chromosomal regions of the three gram-positive cocci. The *B. subtilis* gene cluster also lacks several of these genes but contains several additional genes, perhaps involved in sporulation, which are lacking in the other bacteria. It is unclear whether there was a common ancestor which contained all of these genes in one cluster and there was subsequent rearrangement and dispersal or whether the ancestor contained a simpler DCW region and organisms such as gram-negative rods rear-

ranged the distant genes into a cluster for transcriptional control or other regulatory purposes.

Nucleotide sequence accession numbers. The GenBank accession numbers of the *S. aureus* and *E. faecalis* DCW cluster nucleotide sequences determined in this work are U94706 and U904707, respectively.

REFERENCES

1. **Alessi, D. M., and E. R. Olsen.** 1994. GenBank accession no. U06462.
2. **Brown, E. D., J. L. Marquardt, J. P. Lee, C. T. Walsh, and K. S. Anderson.** 1994. Detection and characterization of a phospholactoyl-enzyme adduct in the reaction catalyzed by UDP-*N*-acetylglucosamine enolpyruvoyl transferase, MurZ. Biochemistry **33:**10638–10645.
3. **Bugg, T. D., and C. T. Walsh.** 1992. Intracellular steps of bacterial cell wall peptidoglycan biosynthesis: enzymology, antibiotics, and antibiotic resistance. Nat. Prod. Rep. **9:**199–215.
4. **Dai, K., and J. F. Lutkenhaus.** 1991. *ftsZ* is an essential cell division gene in *Escherichia coli*. J. Bacteriol. **172:**3500–3506.
5. **Dai, K., and J. Lutkenhaus.** 1992. The proper ratio of FtsZ to FtsA is required for cell division to occur in *Escherichia coli*. J. Bacteriol. **174:**6145–6151.
6. **Daniel, R. A., and J. Errington.** 1993. DNA sequence of the *murE-murD* region of *Bacillus subtilis* 168. J. Gen. Microbiol. **139:**361–370.
7. **Daniel, R. A., A. M. Williams, and J. Errington.** 1996. A complex four-gene operon containing essential cell division gene *pbpB* in *Bacillus subtilis*. J. Bacteriol. **178:**2343–2350.
8. **Dewar, S. J., K. J. Begg, and W. D. Donachie.** 1992. Inhibition of cell division initiation by an imbalance in the ratio of FtsA to FtsZ. J. Bacteriol. **174:**6314–6316.
9. **Donachie, W. D.** 1993. The cell cycle of *Escherichia coli*. Annu. Rev. Microbiol. **47:**199–230.
10. **Doublet, P., J. van Heijenoort, and D. Mengin-Lecreulx.** 1992. Identification of the *Escherichia coli murI* gene, which is required for the biosynthesis of D-glutamic acid, a specific component of bacterial peptidoglycan. J. Bacteriol. **174:**5772–5779.
11. **Dougherty, T. J., and K. Kennedy.** 1997. Unpublished data.
12. **Dougherty, T. J., J. A. Thanassi, and M. J. Pucci.** 1993. The *Escherichia coli* mutant requiring D-glutamic acid is the result of two mutations in two distinct loci. J. Bacteriol. **175:**111–116.
13. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. Fitzhugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
14. **Ghuysen, J.-M.** 1994. Molecular structures of penicillin-binding proteins and β-lactamases. Trends Microbiol. **2:**372–380.
15. **Harry, E. J., S. R. Partridge, A. S. Weiss, and R. G. Wake.** 1994. Conservation of the 168 *divIB* gene in *Bacillus subtilis* W23 and *B. licheniformis*, and evidence for homology to *ftsQ* of *Escherichia coli*. Gene **147:**85–89.
16. **Holtje, J.-V., and U. Schwarz.** 1985. Biosynthesis and growth of the murein sacculus, p. 77–119. *In* N. Nanninga (ed.), Molecular cytology of *Escherichia coli*. Academic Press, Inc., New York, N.Y.
17. **Kloos, W. E., and K. H. Schleifer.** 1986. Staphylococcus, p. 1019. *In* P. H. A. Sneath (ed.), Bergey's manual of systematic bacteriology, 1st ed., vol. 2. The Williams & Wilkins Co., Baltimore, Md.
18. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
19. **Lutkenhaus, J.** 1990. Regulation of cell division in *E. coli*. Trends Genet. **6:**22–25.
20. **Miyao, A., A. Yoshimura, T. Sato, G. Theeragool, and Y. Kobayashi.** 1992. Sequence of the *Bacillus subtilis* homolog of the *Escherichia coli* cell-division gene *murG*. Gene **118:**147–148.
21. **Mundt, J. O.** 1986. Enterococci, p. 1064. *In* P. H. A. Sneath (ed.), Bergey's manual of systematic bacteriology, 1st ed., vol. 2. The Williams & Wilkins Co., Baltimore, Md.
22. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparisons. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
23. **Pucci, M. J., L. F. Discotto, and T. J. Dougherty.** 1992. Cloning and identification of the *Escherichia coli murB* DNA sequence, which encodes UDP-*N*-acetylpyruvoylglucosamine reductase. J. Bacteriol. **174:**1690–1693.
24. **Pucci, M. J., and J. A. Thanassi.** Unpublished data.
25. **Ray, C., R. E. Hay, H. L. Carter, and C. P. Moran.** 1985. Mutations that affect utilization of a promoter in stationary-phase *Bacillus subtilis*. J. Bacteriol. **163:**610–614.
26. **Roe, B. A., S. Clifton, M. McShan, and J. Ferretti.** 1997. Streptococcal Genome Sequencing Project. University of Oklahoma, Norman.
27. **Rowland, S. L., J. Errington, and R. G. Wake.** 1995. The *Bacillus subtilis* cell-division 135–137° region contains an essential *orf* with significant similarity to *murB* and a dispensable *sbp* gene. Gene **164:**113–116.
28. **Spratt, B. G.** 1977. Properties of the penicillin-binding proteins of *Escherichia coli* K-12. Eur. J. Biochem. **72:**341–352.
29. **van Heijenoort, J.** 1994. Biosynthesis of the bacterial peptidoglycan unit, p. 39–54. *In* J.-M. Ghuysen and R. Hakenbeck (ed.), Bacterial cell wall. Elsevier, Dordrecht, The Netherlands.
30. **van Heijenoort, J.** 1996. Murein synthesis, p. 1025–1034. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
31. **Vicente, M., and J. Errington.** 1996. Structure, function and controls in microbial division. Mol. Microbiol. **20:**1–7.
32. **Wada, A., and H. Watanabe.** 1994. GenBank accession no. D28879.
33. **Williamson, R., S. B. Calderwood, R. C. Moellering, Jr., and A. Tomasz.** 1983. Studies on the mechanism of intrinsic resistance to β-lactam antibiotics in group D streptococci. J. Gen. Microbiol. **129:**813–822.
34. **Yura, T., H. Mori, H. Nagai, T. Nagata, A. Ishihama, N. Fujita, K. Isono, K. Mizobuchi, and A. Nakata.** 1992. Systematic sequencing of the *Escherichia coli* genome: analysis of the 0–2.4 min region. Nucleic Acids Res. **20:**3305–3308.