JOURNAL OF BACTERIOLOGY, Oct. 1997, p. 6311–6317 Vol. 179, No. 20
0021-9193/97/$04.00+0

Characterization of the *oriC* Region of *Mycobacterium smegmatis*

MING-HUI QIN,[1] MURTY V. V. S. MADIRAJU,[1] SUNNY ZACHARIAH,[2] AND MALINI RAJAGOPALAN[1,2]*

Department of Biochemistry[1] and Center for Pulmonary and Infectious Disease Control,[2] The University of Texas Health Center at Tyler, Tyler, Texas 75710

Received 19 May 1997/Accepted 14 August 1997

A 3.5-kb DNA fragment containing the *dnaA* region of *Mycobacterium smegmatis* has been hypothesized to be the chromosomal origin of replication or *oriC* (M. Rajagopalan et al., J. Bacteriol. 177:6527–6535, 1995). This region included the *rpmH* gene, the *dnaA* gene, and a major portion of the *dnaN* gene as well as the *rpmH-dnaA* and *dnaA-dnaN* intergenic regions. Deletion analyses of this region revealed that a 531-bp DNA fragment from the *dnaA-dnaN* intergenic region was sufficient to exhibit *oriC* activity, while a 495-bp fragment from the same region failed to exhibit *oriC* activity. The *oriC* activities of plasmids containing the 531-bp sequence was less than the activities of those containing the entire *dnaA* region, suggesting that the regions flanking the 531-bp sequence stimulated *oriC* activity. The 531-bp region contained several putative nine-nucleotide DnaA-protein recognition sequences [TT(G/C)TCCACA] and a single 11-nucleotide AT-rich cluster. Replacement of adenine with guanine at position 9 in five of the putative DnaA boxes decreased *oriC* activity. Mutations at other positions in two of the DnaA boxes also decreased *oriC* activity. Deletion of the 11-nucleotide AT-rich cluster completely abolished *oriC* activity. These data indicate that the designated DnaA boxes and the AT-rich cluster of the *M. smegmatis dnaA-dnaN* intergenic region are essential for *oriC* activity. We suggest that *M. smegmatis oriC* replication could involve interactions of the DnaA protein with the putative DnaA boxes as well as with the AT-rich cluster.

The *dnaA* gene flanking region in many bacteria is the origin of chromosomal DNA replication, or *oriC* region (10, 26, 27). The *oriC* region is generally AT nucleotide rich, containing repeats of AT-rich nucleotide sequences varying in length from 13 (e.g., *Escherichia coli* and pseudomonads) to 16 (e.g., *Bacillus subtilis*) together with several 9-nucleotide DnaA protein recognition sequences called DnaA boxes (10, 17, 26, 27). Biochemical studies with purified proteins and the *oriC* region of *E. coli* revealed that the first step in the initiation of DNA replication involved binding of 20 or more DnaA molecules to the DnaA boxes (3). This step is presumed to be followed by opening of the double-stranded *oriC* DNA at 13-mer AT-rich cluster sites for the entry of DnaC-directed DnaB protein. Additional events at the *oriC* region then result in unwinding of double-stranded DNA, priming of DNA synthesis, and formation of replication forks. Genetic studies with mutant DnaA proteins and mutant origin regions further clarified the role(s) of the DnaA protein and the mechanism of replication initiation in *E. coli* (10, 15).

The mycobacteria exhibit varied growth rates, with the doubling times ranging from 2 to 3 h (*Mycobacterium smegmatis* and *M. fortuitum*) to 10 to 12 h (*M. avium-intracellulare* complex) to 22 to 24 h (*M. tuberculosis* and *M. bovis* BCG) (29). The genetic and biochemical aspects of replication, especially initiation, in mycobacteria are not well understood. As a first step toward identifying the key players, i.e., *cis*- and *trans*-acting factors, in mycobacteria, we cloned DNA fragments containing the *dnaA* gene region, i.e., *rpmH*, *dnaA*, and *dnaN* and the intergenic regions, from *M. smegmatis* (22) and *M. tuberculosis* (21) and determined their nucleotide sequences (GenBank accession numbers are U17833 for *M. smegmatis* and U38891 for *M. tuberculosis*). We further showed that the *dnaA* region from *M. smegmatis* exhibited functional replication origin activity; i.e., when cloned in pUC18-based plasmids, it rendered the plasmids capable of autonomous replication in *M. smegmatis* (22). Plasmid pUC18 is characteristically nonreplicative in *M. smegmatis* (22, 28). The *dnaA* region of *M. smegmatis* is referred to in this report as *oriC*. Nucleotide sequence analysis of this region revealed one 11-nucleotide AT-rich cluster and several putative DnaA protein recognition sequences (DnaA boxes) with one to two mismatches with the consensus sequence of TT(G/C)TCCACA. These studies, however, did not localize *oriC* activity to a defined region of the 3.5-kb DNA fragment.

Recently Salazar et al. reported comparative nucleotide sequence analyses of the *dnaA* regions from *M. tuberculosis*, *M. leprae*, and *M. smegmatis* (23) and, consistent with our results, showed that the *M. smegmatis dnaA* region exhibits *oriC* activity. Their studies localized *oriC* activity to an approximately 1.1-kb DNA fragment of the *dnaA* region that contained the 3′ and 5′ ends of the *dnaA* and *dnaN* genes, respectively, together with the intergenic region. While these and our earlier studies identified the *oriC* region, there are several unanswered questions with regard to the sequences that are essential for *oriC* activity. For example, (i) what is the minimal *dnaA* region fragment length that exhibits *oriC* activity, (ii) which of the putative DnaA boxes are essential for *oriC* activity, and (iii) is the single AT-rich nucleotide cluster essential for *oriC* activity? Answers to these questions form the basis for this report. Our results indicated that a 531-bp fragment from the *dnaA-dnaN* intergenic region is sufficient to exhibit *oriC* activity. Point mutations in five of the designated DnaA boxes decreased the *oriC* activity of *M. smegmatis*. Further, deletion of the AT-rich cluster to the left of the DnaA boxes abolished *oriC* activity. The significance of these results with respect to *M. smegmatis* *oriC* replication is discussed.

* Corresponding author. E-mail: malini@uthct.edu.

TABLE 1. Strains and plasmids used

Strain or plasmid	Genotype or description	Source or reference
Bacterial strains		
E. coli TOP 10	F′ (*lac*I^q Tet^r) *mcr*A Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX74 deoR recA1 araD139* Δ(*ara-leu*)*7697 galU galK rpsL endA1 nupG*	Invitrogen, Inc.
M. smegmatis mc^{2}155	Transformation-proficient strain	W. R. Jacobs, Jr.
Plasmids		
pUC18	Amp^r	Lab stock
pMR40	4-kb *Pvu*II fragment containing the *dnaA* gene region of *M. smegmatis* cloned into pUC18; Amp^r	22
pMR41	Same as pMR40 except that the 4-kb *Pvu*II fragment was cloned in the opposite orientation; Amp^r	22
pMR43	A 456-bp *Bam*HI fragment of pMR40 was replaced with a 1.3-kb fragment containing the *aph* gene; Amp^r Km^r	22
pMQ4	A 1.2-kb *Xho*I fragment of pMR40 (nt 790–2051[a]) cloned into the *Sal*I site of pUC18; Amp^r	22
pMQ59[b]	A 1-kb *Eco*RI fragment deleted from pMR41; Amp^r Km^r	This study
pMQ61[b]	A 1-kb *Eco*RI fragment from pMR41 was ligated with the large *Eco*RI fragment of pMQ4; orientation was confirmed; the *aph* gene was inserted at the *Bam*HI site of the insert; Amp^r Km^r	This study
pMQ62[b]	A 0.8-kb *Xho*I fragment of pMR40 (nt 2183–3038[a]) cloned into the *Sal*I site of pUC18; Amp^r Km^r	This study
pMQ63[b]	A 3.0-kb *Xho*I-*Hin*dIII fragment deleted from pMR40; Amp^r Km^r	This study
pMQ64	A 1.9-kb *Sma*I fragment deleted from pMR43; Amp^r Km^r. Note that the resulting construct retained most of the *rpmH* gene	This study
pMQ65	Same as pMQ64 except that the *Xho*I site was inactivated by a 4-bp fill-in reaction; Amp^r Km^r	This study
pMQ131[b]	*Bam*HI-*Eco*RI-digested PCR fragment (1 kb, nt 2421–3425[a]) was cloned into the same sites of pUC18; Amp^r Km^r	This study
pMQ67[b]	Same as pMQ131 with inactivated *Xho*I site; Amp^r Km^r	This study
pMQ68[b]	A 0.7-kb *Ava*I fragment of pMQ67 (nt 2673-3386[a]) cloned into the same site of pUC18; Amp^r Km^r	This study
pMQ69[b]	A 549-bp PCR fragment (nt 2710–3259[a]) cloned into the *Sma*I site of pUC18; Amp^r Km^r	This study
pMQ71[b]	*Dra*I-*Eco*RI-digested fragment from pMQ121 (531 bp, nt 2728–3259[a]) cloned into the *Hin*cII-*Eco*RI site of pUC18; Amp^r Km^r	This study
pMQ72[b]	A 519-bp (nt 2740–3259[a]) PCR fragment cloned into pUC18, Amp^r Km^r	This study
pMQ74[b]	*Eco*RI-digested PCR fragment (495 bp, nt 2740–3235[a]) cloned into the same site of pUC18; Amp^r Km^r	This study
pMQ75[b]	*Ava*I-*Eco*RI digested PCR fragment (676 bp, nt 2710–3386[a]) cloned into the *Sal*I-*Eco*RI site of pUC18; Amp^r Km^r	This study
pMQ76[b]	*Sma*I-*Eco*RI digested PCR fragment (584 bp, nt 2675–3259[a]) cloned into *Hin*cII-*Eco*RI site of pUC18; Amp^r Km^r	This study

[a] Nucleotide (nt) numbering was based on the published *dnaA* gene region sequence (22).
[b] A 1.3-kb DNA fragment carrying *aph* gene was cloned into the multiple cloning site of the pUC18 vector.

MATERIALS AND METHODS

Bacterial strains and plasmids. Table 1 lists the strains of mycobacteria and *E. coli* used in this study. The pUC18 multipurpose cloning vector was used for cloning various DNA fragments. Details of the recombinant plasmid constructions are described in Table 1.

Media, growth, and transformation conditions. *E. coli* cultures were grown in Luria-Bertani (LB) broth (16). Transformation-proficient *E. coli* was prepared by treatment of cultures with rubidium chloride (22, 24). Recombinants were selected on LB agar plates containing ampicillin (50 μg/ml), kanamycin (50 μg/ml), or both (22). Mycobacterial strains were grown in 7H9 broth containing 0.05% Tween 80, 0.5% albumin, 0.2% dextrose, and 0.02% NaCl as described previously (22). Electrocompetent *M. smegmatis* cells were prepared as described previously (9). Electroporation was carried out in a Bio-Rad Gene Pulser (Bio-Rad Laboratories, Richmond, Calif.) at 2.5 kV, 25 μF, and 1,000 ohms. Following electroporation, 0.8 ml of 7H9 broth was added, the cell suspension was incubated at 37°C for 3 h, and the recombinants were selected at 37°C on 7H10 agar plates containing 50 μg of kanamycin per ml. The plates were incubated for up to 1 week, the transformants were counted, and transformation efficiencies were calculated. Approximately 0.2 to 0.5 μg of DNA was used for each transformation. Unless otherwise stated, the higher amount of DNA (0.5 μg) per transformation assay was used for those *oriC* plasmids that gave lower transformation efficiencies. Transformation efficiencies in *E. coli* of all *oriC* plasmids tested were comparable (approximately 10^7 colonies/μg of DNA). An *E. coli-Mycobacterium* shuttle plasmid, pMV206, and the *oriC* plasmid, pMR43, containing the 3.5-kb *dnaA* region (22) were used as positive controls in all *M. smegmatis* transformation experiments.

Recombinant DNA techniques. Plasmid DNA from *E. coli* was isolated and manipulated as described by Sambrook et al. (24). Chromosomal DNA from *M. smegmatis* was prepared as described by Husson et al. (8). Plasmid DNA was recovered from *M. smegmatis* into *E. coli* in a two-step process. First, *E. coli* Top 10 cells were transformed with plasmid DNA from *M. smegmatis* either by electroduction (1) or by standard transformation with an aliquot of the crude lysate from *M. smegmatis* cells lysed by minibead beating with zirconium beads (20a). Plasmid DNA was then recovered from *E. coli* cells by standard protocols (24). DNA sequencing was performed in a Pharmacia ALF DNA sequencer, using the Pharmacia autocycle sequencing kit and protocol (Pharmacia Biotech, Piscataway, N.J.). Oligonucleotide primers used in this study were synthesized in a Pharmacia Gene Assembler Special. Restriction digestion, labeling of DNA with radioactive nucleotides, transfer of DNA to nitrocellulose membranes, and Southern hybridization were carried out as described elsewhere (24).

***oriC* activity.** *oriC* activity was defined as the ability of DNA fragments from the *dnaA* region of *M. smegmatis* to render pUC18-based plasmids capable of autonomous replication in *M. smegmatis* and is expressed as the total number of kanamycin-resistant transformants obtained per microgram of input DNA (22).

Construction of *oriC* plasmids. Plasmids containing various lengths of the *dnaA* region as indicated in Fig. 1 were constructed either by deletions of indicated regions from pMR40, pMR41, and pMR43 or by gel purification of appropriate restriction digestion products followed by recloning into pUC18 vectors (Table 1). In some cases, desired DNA fragments were amplified by PCR, gel purified, and cloned into pUC18 vectors. To select transformants in mycobacteria, a 1.3-kb DNA fragment containing the aminocyclitol phosphotransferase (*aph*) gene, which confers resistance to kanamycin, was also cloned in the multiple cloning site of the pUC18 vector.

The oligonucleotides used were Q3, Q18, and Q19 (see below), Q8 (5′-GGG**GAATTC**GTCGTCGCCACGTCCTATG-3′), Q9 (5′-GGG**GAATTC** GAGCAACTTCTGTGACAC-3′), Q10 (5′-GGG**GAATTC**GACGACGCCTG

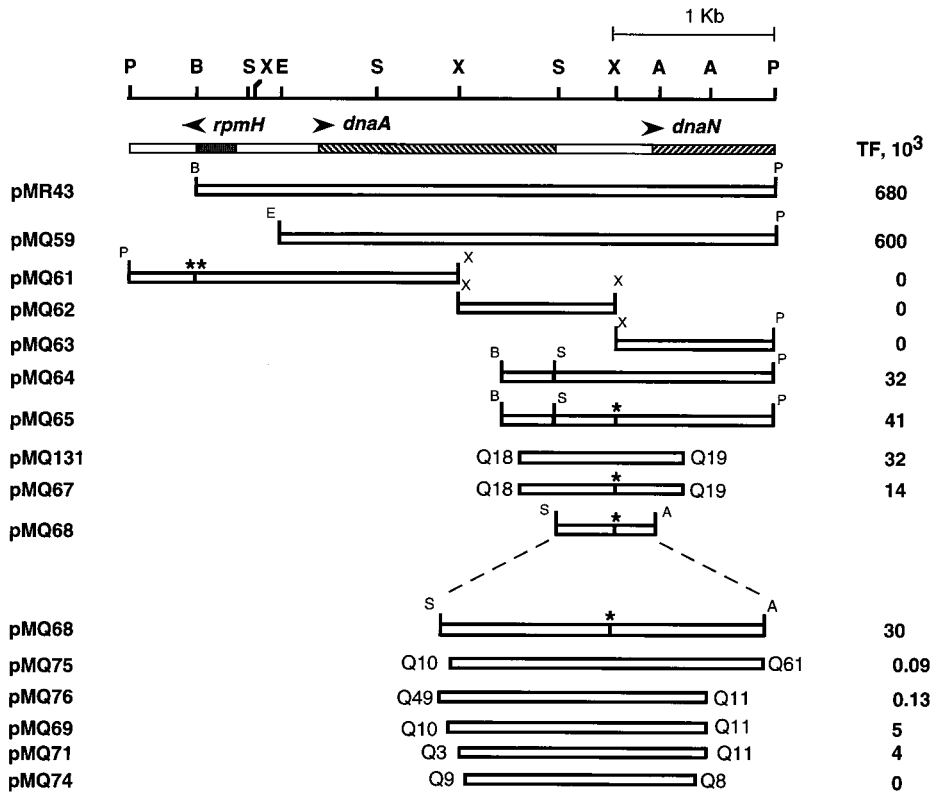

FIG. 1. Determination of the *M. smegmatis* minimal *dnaA* region required for *oriC* activity. The 4-kb *dnaA* gene region of *M. smegmatis* is shown. Shaded boxes indicate the open reading frames of the indicated genes. The directions of transcription of *rpmH*, *dnaA*, and *dnaN* genes are indicated by arrowheads. Restriction endonuclease sites: B, *Bam*HI; P, *Pst*I; S, *Sma*I; E, *Eco*RI; X, *Xho*I; A, *Ava*I. The names of various *oriC* plasmids and the corresponding *oriC* fragments (open boxes) are shown on the left. The total numbers of transformants obtained per microgram DNA are shown on the right. Approximately 0.2 to 0.5 μg of DNA per transformation assay was used. In the cases of pMQ75 and pMQ76, 0.6 μg of DNA was used in transformation assays. Following electroporation, 7H9 broth was added, cells were incubated as described in Materials and Methods, and appropriate dilutions were made and plated on 7H10 agar plates containing kanamycin. This resulted in typically 20 to 500 transformants per plate. No transformants were obtained with pMQ61, pMQ62, pMQ63, and pMQ74 even when 0.5 μg of plasmid DNA per transformation assay was used. The oligonucleotide primers used to amplify indicated *oriC* DNA fragments are shown on either side of the amplified product; e.g., pMQ131 was constructed by cloning a 1-kb DNA fragment amplified by using Q18 and Q19 oligonucleotide primers, etc. *, inactivated *Xho*I site; **, location of the *aph* gene in pMQ61; TF, number of transformants obtained per microgram of DNA.

ACCG-3′), Q11 (5′-GGGGCCCTGCAGGGATCC**GAATTC**AAGTCGGTC AGCCCAGCC-3′), Q49 (5′-TGTGCCGTGAGCTCACCGA-3′), and Q61 (5′-GCCGAGACCTCGTAGTCGAA-3′). The boldface letters indicate the sequence of the *Eco*RI restriction enzyme recognition site. The following combination of oligonucleotides were used to amplify the *oriC* region and subsequently derive the indicated plasmids: pMQ69, Q10-Q11; pMQ71, Q3-Q11; pMQ74, Q8-Q9; pMQ75, Q10-Q61; pMQ76, Q11-Q49; and pMQ131, Q18-Q19.

Amplification reaction mixtures contained the following in a 25-μl volume: 20 mM Tris-HCl buffer (pH 8.5), 10 mM ammonium sulfate, 2 mM magnesium sulfate, 100 μg of bovine serum albumin per ml, 0.1% Triton X-100, 10 to 30 ng of DNA template, oligonucleotides at a final concentration of 1 μM, and 1.5 U of *Pfu* polymerase (Stratagene Corporation, San Diego, Calif.). The following amplification conditions were used: initial 5-min heating at 95°C followed by 30 cycles of 1 min each at 95, 60, and 72°C. The PCR products were digested with the appropriate restriction enzymes and cloned into pUC18 vectors (Table 1). In some cases, blunt-ended PCR fragments were cloned into the *Sma*I site. The nucleotide sequences of all PCR-generated inserts were verified by DNA sequencing.

Site-directed mutagenesis. Unless indicated otherwise, all mutations were created in the 1-kb *oriC* region of pMQ131. The PCR mutagenesis method of Nelson and Long (20) was used to create specific mutations in DnaA boxes. This method entails the use of four oligonucleotides; one is the mutagenic primer used to direct mutagenesis, and the other three are selection primers that enable selective amplification of the mutated sequence. The selection primers were chosen such that they can be used for amplification of all target sequences. The selection primers used were Q12 (5′-GGGGCCCTGCAGGGATCC**GAATTC**-3′), Q18 5′-GCCG**GGATCC**AGGGCGCTGGCCC-3′), and Q19 (5′-GGGG CCCTGCAGGGATCC**GAATTC**AGCGTCCACCTTGACCTC-3′). Boldface letters indicate the sequence of either *Eco*RI or *Bam*HI restriction enzyme recognition site introduced for cloning purposes. The following mutagenic primers were used in combination with selection primers to obtain mutant *oriC* plasmids: Q3 (5′-TTTCGACGACGCCT**TTTTT**AAATTTTTTGTGAG-3′) for pMQ121, Q5 (5′-GATGACTACGGAATC**A**CTC**A**ACACCTGCGCGAC-3′) for pMQ123, Q6 (5′-GATGACTACGGAATC**AT**TCCACACCTGCGCGAC-3′) for pMQ124, Q13 (5′-CAACCGCTCTCCAC**G**CACCAC-3′) for pMQ126; Q14 (5′-GTGGATACCGAAGTGCAGAACCTGCGC-3′) for pMQ127, Q15 (5′-GAATCTCTCCAC**C**CCTGCGCGACG-3′) for pMQ128, Q17 (5′-CAAC CGCTCT**A**CACACACGAC-3′) for pMQ130, Q30 (5′-CTACGGAATC**A**CT CCACACCT-3′) for pMQ132, Q31 (5′-AATCTCTCCAC**G**CCTGCGCGA-3′) for pMQ133, Q32 (5′-TGGATACCGA**C**GTGCAGAACC-3′) for pMQ134, and Q36 (5′-CACTCACAC**G**GCGACACACCGA-3′) for pMQ138. The underlined boldface letters indicate the substituted bases. DNA sequencing was performed to confirm each mutation prior to use of the mutant plasmids in transformation experiments. The *Xho*I site in plasmids pMQ65, pMQ67, and pMQ68 was inactivated by a fill-in reaction (24).

Copy number determination. *M. smegmatis* isolates containing the appropriate *oriC* plasmids were harvested from 7H10 plates containing kanamycin. The total DNA was isolated (8), digested with *Pvu*II, electrophoretically separated, transferred to nitrocellulose membranes, and probed with ^{32}P-labeled 531-bp *oriC* DNA fragment by Southern hybridization (24). The copy number of *oriC* plasmids relative to the chromosomal *oriC* region was determined by comparing the relative intensities of the chromosomal DNA band with that of the plasmid DNA band by densitometric scans of autoradiographs on a Bio Image densitometer (Millipore Corporation, Bedford, Mass.).

Stability experiments. *M. smegmatis* containing *oriC* plasmids were inoculated into 7H9 broth without kanamycin. An aliquot of the culture after appropriate dilution was plated on 7H10 agar plates with and without kanamycin, and the plates were incubated at 37°C. The remaining culture was incubated at 37°C with shaking at 170 rpm. Aliquots of culture were removed at 6, 12, 24, 36, 48, and 60 h of incubation, diluted with 7H9 broth, and plated on 7H10 agar plates with

and without kanamycin, and incubated at 37°C. Colonies from both sets of plates were counted, and the ratio of colonies obtained in the presence and absence of kanamycin was determined. In addition, plasmid DNA was recovered from the cells cultured on antibiotic plates following 60 h of growth in antibiotic-free 7H9 broth, digested with *Bam*HI, electrophoresed, transferred to nitrocellulose membrane, and probed with an *oriC*-specific radiolabeled DNA fragment by Southern hybridization (24).

M. smegmatis transformed with *oriC* plasmids produced heterogeneous colony sizes ranging from normal (wild type) to small colonies. The size of the small colonies approached that of the normal colonies with increased incubation up to the time at which no new colonies appeared on the plates. Transformants containing the shuttle vector pMV206, on the other hand, gave homogeneous colony sizes. These findings are consistent with those of Salazar et al. (23). We found that when the primary transformants (containing either *oriC* plasmids or pMV206) were directly resuspended in 7H9 broth and plated, the ratios of colonies on antibiotic plates to those on antibiotic free plates varied. However, these ratios remained constant during various growth periods examined, suggesting that the kanamycin-resistant cells within the primary transformants do not lose their plasmids (data not shown). On the other hand, when the transformants were streaked on 7H10 agar plates containing kanamycin prior to resuspension and plating, the ratios of colonies on antibiotic plates to those on antibiotic-free plates was always 1. Mycobacteria have a tendency to clump together, thereby producing a heterogeneous population in which some plasmid-free cells remain associated with plasmid-containing cells as background. The plasmid-free cells were eliminated upon subsequent streaking and colony purification. Hence, *M. smegmatis* mc^2155 transformants, unless stated otherwise, were colony purified on 7H10 agar plates containing 50 μg of kanamycin per ml prior to use in all experiments.

RESULTS AND DISCUSSION

Identification of minimal size DNA fragments from the *dnaA* gene region of *M. smegmatis* that support *oriC* activity. Deletion analyses of the 4.0-kb *dnaA* region of *M. smegmatis* from the 5′ and 3′ ends followed by determination of *oriC* activity of the resultant plasmids (pMQ59, pMQ61, pMQ62, pMQ63, and pMQ131) located the *oriC* activity to a 1-kb DNA fragment. Plasmids containing a 1.9-kb internal deletion spanning a region upstream of the *rpmH* gene to the end of the *dnaA* gene also exhibited *oriC* activity (pMQ64 in Fig. 1; Table 1). The 1-kb *oriC* region contained the 3′ and 5′ ends of the *dnaA* and *dnaN* genes, respectively, together with the intergenic region (see Fig. 1 and Table 1 for details). These results are consistent with those of Salazar et al. (23). A notable finding, however, was a 20-fold reduction in transformation frequency in these constructs (compare pMQ131 and pMQ64 with pMQ59 or pMR43). The 4-kb *dnaA* region of *M. smegmatis* contained the entire *rpmH* and *dnaA* genes as well as part of the *dnaN* gene (22). Removal of the *rpmH* and *dnaN* coding regions did not reduce *oriC* activity further (compare pMQ131 with pMQ64 [Table 1]). Together these data indicate that a DNA fragment containing the *dnaA-dnaN* intergenic region functions as *oriC* and that its 5′ flanking region containing the *dnaA* gene, its promoter, and possibly the *rpmH* promoter function in the regulation of *oriC* activity. The putative promoters of *dnaA* and *rpmH* have not been mapped (22). The individual contributions of putative promoter regions and that of the *dnaA* coding region remain to be established. It is pertinent to note that elimination of the *rpmH* promoter region in *B. subtilis* resulted in stimulation of *oriC* activity (17). Mutation at the *Xho*I site of the *dnaA-dnaN* intergenic region did not reduce *oriC* activity significantly, suggesting that the sequence in this region is not essential for *oriC* activity (compare pMQ65 and pMQ67 with pMQ64 and pMQ131, respectively, in Fig. 1). pMQ65 was derived from pMQ64 (see above), whereas pMQ67 was derived from pMQ131 (Table 1).

To identify the smallest DNA fragment that supports *oriC* activity, further deletions in the 1-kb *dnaA* region (pMQ131) were generated by PCR amplification or restriction digestion (Table 1). *oriC* DNA fragments that were 711 bp (pMQ68), 549 bp (pMQ69), and 531 bp (pMQ71) but not 495 bp (pMQ74) in length exhibited *oriC* activity in pUC18 vectors (Fig. 1). The *oriC* activities of plasmids lacking either 35 bp from the 5′ end or 127 bp from the 3′ end of the 711-bp fragment (pMQ75 and pMQ76, respectively [Table 1]) were much lower than those of their controls (compare pMQ75 and pMQ76 with pMQ68 [Fig. 1]). However, simultaneous deletion of both regions resulted in only a sixfold decrease in *oriC* activity (pMQ69, [Table 1 and Fig. 1]). Similar types of analyses with pMQ71, however, did not produce the same effect (compare pMQ75, pMQ76, and pMQ74 with pMQ71). These results further suggested that the 531-bp *dnaA-dnaN* intergenic region of *M. smegmatis* is the minimal DNA fragment that exhibits *oriC* activity and that the regions flanking the 549-bp *oriC* DNA fragment appeared to be involved in regulation of *oriC* activity.

In contrast to our present results (Fig. 1) and those reported earlier (22), Salazar et al. showed that the presence of the *M. smegmatis dnaA* coding region in pIJ963, an *E. coli* plasmid with the hygromycin resistance marker, was inhibitory for *oriC* activity (23). Their *oriC* plasmids were also found to be unstable in the absence of selection. The *dnaA* gene region used in both our studies (this study and reference 22) and theirs (23) was obtained from the genomic DNA library cloned in pYUB18-derived cosmid vectors, not by PCR amplification. Furthermore, the high transformation efficiencies observed with our *oriC* plasmids containing the intact *dnaA* region (Fig. 1; see also reference 22) argues against the suggestion that the *dnaA* gene in our constructs was somehow mutated during transformation and subsequent growth (23). It is conceivable that the use of the hygromycin antibiotic marker gene and its location with respect to the *dnaA* gene region could have had an adverse effect on replication of their *oriC* plasmids. This possibility was partially supported by our current observation that *oriC* activities of pUC18-based plasmids containing a 1-kb *oriC* DNA fragment were severely decreased and the activities of those with the 549-bp fragment were completely abolished when the *aph* gene in these plasmids was located at the *Ssp*I site rather than the multiple cloning site adjacent to the *oriC* region (20a). The *Ssp*I site is located between the *lac* promoter and the *bla* gene of the pUC18 vector (24) and is approximately 580 bp away from the *oriC* sequence. Purification of the pMR43-encoded *dnaA* gene product and characterization of its interaction with *oriC* DNA as well as construction and characterization of *dnaA* mutant strains of *M. smegmatis* would help resolve the discrepancies in these experiments.

Stability of *oriC* plasmids. To investigate the stability of *M. smegmatis oriC* plasmids, isolates containing different *oriC* plasmids (pMR43, pMQ131, pMQ69, and pMQ71) were cultured for different time periods in 7H9 broth without kanamycin. At various time intervals, samples were removed, diluted with 7H9 broth, plated on 7H10 agar plates with and without kanamycin, and incubated at 37°C, and colonies were counted. The numbers of colonies on agar plates with and without kanamycin for pMR43 and pMQ131 *oriC* plasmids at all growth periods tested were found to be similar (Fig. 2). Similar results were also obtained with *M. smegmatis* cells containing pMQ69 and pMQ71 plasmids as well as those transformed with the recovered *oriC* plasmids (data not shown). Isolation of *oriC* plasmid DNA from cells grown on antibiotic plates, following 60 h incubation in broth without antibiotic, and localization of the *oriC* region by Southern hybridization (24) confirmed the presence of the *oriC* DNA fragment in the recovered *oriC* plasmids. Further, the sizes of input and recovered plasmids appeared to be the same (Fig. 3). Together these data suggest that *M. smegmatis oriC* plasmids are stably maintained. *oriC* plasmids from many other bacteria have been

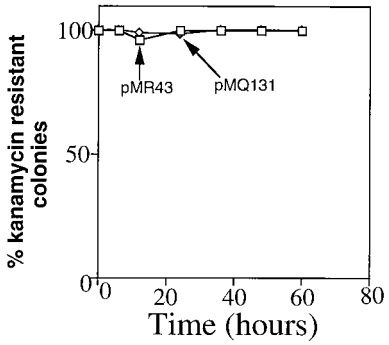

FIG. 2. Stability of *M. smegmatis oriC* plasmids pMR43 and pMQ131. *M. smegmatis* cultures containing *oriC* plasmids were grown in 7H9 broth; aliquots of the culture were removed at indicated time periods, diluted with fresh 7H9 broth, and plated on 7H10 agar plates with and without kanamycin. Colonies on both sets of plates were counted, and the percentage of kanamycin-resistant colonies was determined.

shown to be unstable (17, 27, 30). It is not clear why the *M. smegmatis oriC* plasmids are stably maintained. One possibility is that the *oriC* sequences in these plasmids contain appropriate partitioning signals that enable the stable maintenance of *oriC* plasmids.

Copy number of *oriC* plasmids. The *oriC* plasmids of many bacteria, including *B. subtilis*, *Streptomyces coelicolor*, and *Pseudomonas putida*, exist in low copy number (17, 27, 30). To determine the copy number of *M. smegmatis oriC* plasmids relative to the chromosomal *oriC*, total DNA from cells containing *oriC* plasmids was extracted, digested with *Pvu*II, gel electrophoresed, transferred to nitrocellulose membranes, probed with radiolabeled *oriC*-specific DNA fragments (24), and visualized by autoradiography (Fig. 4). Comparison of the intensity of the chromosomal DNA fragment with that of the input *oriC* plasmid revealed that the copy number of all *M.*

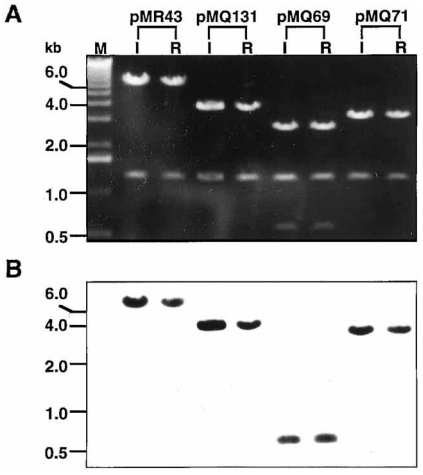

FIG. 3. Stability of *oriC* plasmids. (A) Agarose gel electrophoresis of *Bam*HI-digested *oriC* plasmids recovered from *M. smegmatis* grown for 60 h in the absence of kanamycin. Patterns obtained with input plasmid DNA are shown for comparison. Lanes: M, 1-kb ladder; I, input plasmid; R, recovered plasmid. The expected sizes of the fragments are as follows: with pMR43, 6.2 and 1.3 kb; with pMQ131, 3.7 and 1.3 kb; with pMQ69, 2.7, 1.3, and 0.5 kb; and with pMQ71, 3.2 and 1.3 kb. The size of the *aph* gene is 1.3 kb. (B) Autoradiogram of the agarose gel shown in panel A. The DNA in the agarose gel in panel A was transferred to a nitrocellulose membrane, probed with a radiolabeled *oriC*-specific probe by Southern hybridization, and visualized by autoradiography.

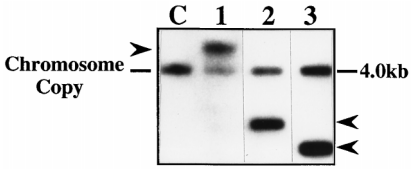

FIG. 4. Copy number of *oriC* plasmids. Total DNA from *M. smegmatis* cells containing pMR43, pMQ131, or pMQ71 was isolated, restricted with *Pvu*II, gel electrophoresed, transferred to a nitrocellulose membrane, and probed with a radiolabeled *oriC*-specific probe. Following washing, membranes were exposed to X-ray film and autoradiograms were scanned in a Millipore Bio Image densitometer. Lanes: C, chromosomal DNA from *M. smegmatis* cells; 1 to 3, total DNA preparations from *M. smegmatis* cultures containing plasmids pMR43, pMQ131, and pMQ71, respectively. Chromosomal copy of the *dnaA* gene on the 4-kb *Pvu*II fragment is marked. Arrowheads indicate plasmid-derived *oriC* fragments.

smegmatis oriC plasmids relative to the chromosomal *oriC* was approximately 2 (Fig. 4).

Mutagenesis of the *oriC* region. (i) The AT-rich cluster. One of the characteristic features of bacterial *oriC* is the presence of 13- to 16-bp AT-rich repeats, specifically to the left of the DnaA boxes (12, 13, 27). Studies with the *E. coli oriC* system indicated that the 13-mer AT-rich repeats to the left of the DnaA boxes were melted by the DnaA protein. This step is prerequisite for DnaC directed DnaB helicase loading (2, 3, 10, 15). To determine whether the single 11-nucleotide AT-rich cluster of the *M. smegmatis oriC* region located to the left the DnaA boxes is required for its activity, the AT-rich cluster was deleted from pMQ71. The resulting plasmid (pMQ72) exhibited no *oriC* activity (Table 2). To determine whether the AT cluster of a defined length, i.e., 11-nucleotide AT-rich sequence, in this region is important for *oriC* function, the GACCG nucleotides upstream of the AT-rich cluster (Fig. 5) were replaced with five T nucleotides (pMQ121 [Table 1]). The resultant mutant plasmid exhibited no *oriC* activity. Together these data indicated that the 11-nucleotide AT-rich cluster was important for *oriC* activity and that an increase in length of the AT-rich sequence in this region was inhibitory to *oriC* activity. Whether the function of the AT-rich cluster of *M. smegmatis* is the same as that in *E. coli* remains to be determined.

(ii) DnaA boxes. TT(G/C)TCCACA, the nine-nucleotide sequence present in the *M. smegmatis dnaA* gene region, has previously been defined as the consensus sequence for the putative DnaA boxes of *M. smegmatis* (22). Three DnaA boxes with one mismatch to this consensus sequence, referred to as B-1, B-2, and B-3, have been identified in the *dnaA-dnaN* intergenic region (Fig. 5). To determine which of the putative DnaA boxes are essential for *oriC* activity, each of the DnaA boxes was mutated individually and the mutational consequences were evaluated. Since *oriC* activities of plasmids containing a 1-kb *oriC* DNA fragment (i.e., pMQ131) were higher than activities of those containing the 549-bp (pMQ69) or 531-bp (pMQ71) fragments, all of the mutations were created in plasmid pMQ131. As a first step, the adenine nucleotide at the ninth position (A_9) of the DnaA box, which is very conserved in the DnaA boxes of all bacteria (3–7, 11, 12, 14, 15, 25, 26), was mutated by replacing it with guanine (G). The *oriC* plasmid with this mutation in DnaA box B-1, called pMQ134, exhibited a 110-fold reduction in the *oriC* activity (Table 2). Similar replacements in the other two boxes (B-2 and B-3) also severely decreased *oriC* activity (Table 2). While systematic replacement of A_9 with cytosine (C) and thymine (T) in all DnaA boxes was not carried out, an A_9-to-T change in the B-1 DnaA box abolished *oriC* activity, whereas an A_9-to-C change

TABLE 2. *oriC* activities in the plasmids used

Plasmid	Sequence		Base(s) or sequence mutated	No. of transformants/ μg of DNA (10^3)
	DnaA box[a]	AT cluster		
pMQ131	Wild type		None	32.0
pMQ134	TTCTGCACA (B-1)		A → G	0.29
pMQ133	CTCTCCACA (B-2)		A → G	0.13
pMQ126	CTCTCCACA (B-3)		A → G	0.21
pMQ128	CTCTCCACA (B-2)		A → C	8.6
pMQ127	TTCTGCACA (B-1)		A → T	0.0
pMQ130	CTCTCCACA (B-3)		C → A	0.0
pMQ132	CTCTCCACA (B-2)		T → A	3.4
pMQ123	CTCTCCACA (B-2)		T → A, C → A	1.5
pMQ124	CTCTCCACA (B-2)		T → A, C → T	0.6
pMQ138	ACTCACACC[b]		C → G	7.89
PMQ71		AAATTTTTTGT (wild type)	None	4.0
pMQ72		AAATTTTTTGT	Deletion	0.0
pMQ121		GACCGAAATTTTTTGT	GACCG → TTTTT	0.0

[a] B-1, B-2, and B-3 refer to the DnaA boxes shown in Fig. 5. Mutated bases are underlined.
[b] Mutation created outside the DnaA box (see also Fig. 5).

in the B-2 box had only a moderate effect (Table 2). A random point mutation outside the DnaA boxes, i.e., C → G (Table 2 and Fig. 5), did not produce major inhibition. As already noted, inactivation of the *Xho*I site located outside the DnaA boxes (pMQ65, pMQ67, and pMQ68 [Fig. 1]) also had no significant inhibitory effect. Whether any other sequences located outside the DnaA boxes influence *oriC* activity has not been tested.

The above-described experiments focused on the DnaA boxes with one mismatch to the DnaA box consensus sequence. Nucleotide sequence analyses of the 550-bp *oriC* region revealed the presence of three DnaA boxes with two mismatches and several DnaA boxes with three mismatches to the consensus sequence for *M. smegmatis* DnaA boxes (22). The majority of these DnaA boxes are located in the first half of the *oriC* sequence, i.e., upstream of the *Xho*I site. Two of the boxes with a three-base mismatch to the consensus sequence are referred to as B-4 and B-5 in Fig. 5. Mutant plasmids containing the A_9-to-G change in either the B-4 or B-5 box resulted in an approximately 50-fold reduction in *oriC* activity, suggesting that these regions are also important for *oriC* activity (data not shown). Presumably they may function to some degree as DnaA boxes. In all of the above-described experiments, A_9 in the DnaA boxes was substituted either with G or with other bases. Point mutations at other locations in the DnaA boxes also affected *oriC* activity. Replacement of C_5 with A in DnaA box B-3 completely abolished *oriC* activity (pMQ130 [Table 2]), whereas substitution of T_2 with A in DnaA box B-2 resulted in a 10-fold decrease in *oriC* activity (pMQ132 [Table 2]). Comparable mutations, i.e., C_5 to A, in DnaA boxes R-1 and R-4 of *E. coli* had no effect on *oriC* activity (7). It is pertinent to point out that the deduced consensus sequence for the putative DnaA boxes of *M. smegmatis*, TT(G/C)TCCACA, is different from that reported for *E. coli*, TTATCCACA. Two base substitutions in the internal sequence of DnaA box B-2 also decreased *oriC* activity (pMQ123 and pMQ124 [Table 2]). Together these data indicate that the designated DnaA boxes and their precise nucleotide sequences are crucial for *oriC* activity. Assuming that *M. smegmatis* DnaA protein interacts with these presumptive DnaA boxes, as in *E. coli*, decreased *oriC* activity with mutant *oriC* plasmids could be due to defective interaction of DnaA protein with the mutant DnaA boxes.

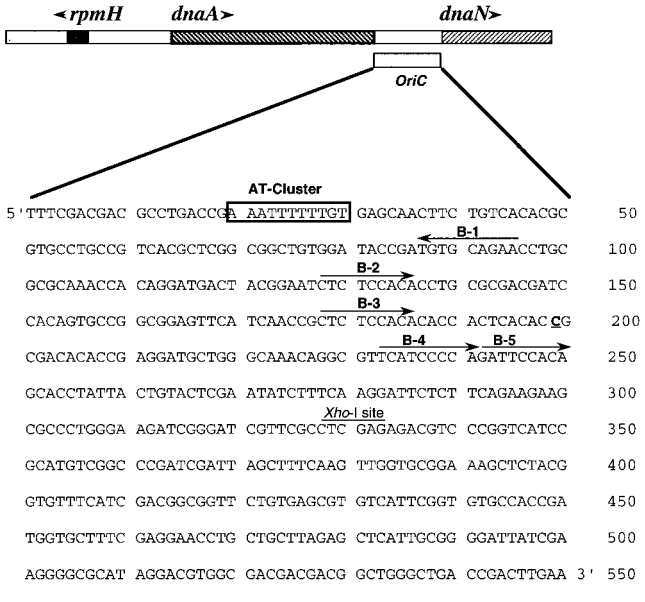

FIG. 5. Nucleotide sequence of the 550-bp *oriC* region. The putative DnaA boxes B-1, B-2, and B-3 containing one mismatch with the consensus sequence TT(G/C)TCCACA and B-4 and B-5 DnaA boxes containing three mismatches with the consensus sequence are marked. The positions and orientations of the DnaA boxes are marked by arrows above the sequence. The boldface underlined letter indicates the nucleotide outside the DnaA box region that was mutated (Table 2). The location of the *Xho*I site is indicated, and the AT-rich nucleotide cluster is boxed. The minimal *dnaA* region that exhibits *oriC* activity spans from the AT-rich cluster region to end of the sequence.

The majority of information about the events occurring at the origin of replication comes from the detailed genetic and biochemical studies carried out in *E. coli* and to a lesser extent in *B. subtilis* (10, 17–19). Although not all of the mutations examined were comparable to those studied in *E. coli*, our studies on *M. smegmatis oriC* replication revealed some differences from *E. coli*. Unlike in *E. coli*, one- or two-base site-specific mutations in any of the presumptive DnaA boxes resulted in decreased *oriC* activity in *M. smegmatis*. In *E. coli*, multiple mutations in two DnaA boxes designated R-1 and R-4 decreased *oriC* activity (7), and mutations in three boxes designated as R-1, R-2, and R-4 completely eliminated *oriC* activity (11). Next, optimal *oriC* activity in *M. smegmatis* requires

the *dnaA* coding region and its 5′ flanking sequences containing the promoters of the *rpmH* and *dnaA* genes. The specific sequences in this region that control *oriC* activity have not been identified. Finally, the overall sequence organization of *M. smegmatis oriC*, i.e., the size of *oriC*, total number of the AT-rich clusters, and the number and orientation of the putative DnaA boxes, is different from that of *E. coli*. These differences tend to suggest that while *oriC* replication in *M. smegmatis* could involve interactions of DnaA protein with the DnaA boxes, the precise mechanism of replication initiation and possibly its regulation could be different from that reported for *E. coli*. Our data suggested (Table 2) that interactions of *M. smegmatis* DnaA protein with all the DnaA boxes tested appeared to be crucial for replication initiation and possibly for subsequent organization of an initiation complex. We propose that following binding of *M. smegmatis* DnaA protein to the DnaA boxes, the DnaA protein molecules bound at one DnaA box interact with those present at other DnaA boxes. A consequence of this process would be the formation of a highly ordered initiation complex leading to subsequent replication initiation steps including interactions with the AT-rich cluster and possibly with other proteins. Thus, severe loss of *oriC* activity in mutant *oriC* plasmids could result from defective binding of DnaA protein to mutant DnaA boxes, thereby affecting the formation of an ordered initiation complex. Since DnaA box sequences are critical for determining the outcome of DnaA protein-DnaA box interactions, precise determination of the DnaA box consensus sequence is essential for understanding *M. smegmatis oriC* replication initiation. Characterization of interactions of DnaA protein with wild-type and mutant DnaA boxes including determination of binding constants would provide valuable data in this regard.

ACKNOWLEDGMENTS

This work was supported in part by NIH grant AI37956.

We thank Vincent A. Steingrube for helpful comments and careful reading of the manuscript. We thank William R. Jacobs, Jr., for providing the transformation proficient strain of *M. smegmatis.*

REFERENCES

1. **Baulard, A., C. Jourdan, A. Mercenier, and C. Locht.** 1992. Rapid mycobacterial plasmid analysis by electroduction between *Mycobacterium* spp and *Escherichia coli*. Nucleic Acids Res. **20:**4105.
2. **Bramhill, D., and A. Kornberg.** 1988. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *Escherichia coli* chromosome. Cell **52:**743–755.
3. **Bramhill, D., and A. Kornberg.** 1988. A model for initiation at origins of DNA replication. Cell **54:**915–918.
4. **Calcutt, M. J., and F. J. Schmidt.** 1992. Conserved gene arrangement in the origin region of the *Streptomyces coelicolor* chromosome. J. Bacteriol. **174:** 3220–3226.
5. **Fujita, M. Q., H. Yoshikawa, and N. Ogasawara.** 1989. Structure of the *dnaA* region of *Pseudomonas putida*: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli*, and *P. putida*. Mol. Gen. Genet. **215:**381–387.
6. **Fujita, M. Q., H. Yoshikawa, and N. Ogasawara.** 1990. Structure of the *dnaA* region of *Micrococcus luteus*: conservation and variations among eubacteria. Gene **93:**73–78.
7. **Holz, A., C. Schaefer, H. Gille, W.-R. Jueterbock, and W. Messer.** 1992. Mutations in the DnaA binding sites of the replication origin of *Escherichia coli*. Mol. Gen. Genet. **233:**81–88.
8. **Husson, R. H., B. J. James, and R. A. Young.** 1990. Gene replacement and expression of foreign DNA in mycobacteria. J. Bacteriol. **172:**519–524.
9. **Jacobs, W. R., Jr., G. V. Kalpana, J. D. Cirillo, L. Pascopella, S. B. Snapper, R. Udani, W. Jones, R. G. Barletta, and B. R. Bloom.** 1991. Genetic systems for mycobacteria. Methods Enzymol. **204:**537–555.
10. **Kornberg, A., and T. Baker.** 1991. DNA replication. W. H. Freeman and Company, New York, N.Y.
11. **Langer, U., S. Richter, A. Roth, C. Weigel, and W. Messer.** 1996. A comprehensive set of DnaA-box mutations in the replication origin, oriC, of *Escherichia coli*. Mol. Microbiol. **21:**301–311.
12. **Marczynski, G. T., and L. Shapiro.** 1992. Cell-cycle control of a cloned chromosomal origin of replication from *Caulobacter crescentus*. J. Mol. Biol. **226:**959–977.
13. **Marians, K. J.** 1992. Prokaryotic DNA replication. Annu. Rev. Biochem. **61:**673–719.
14. **Messer, W.** 1987. Initiation of DNA replication in *Escherichia coli*. J. Bacteriol. **169:**3395–3399.
15. **Messer, W., H. Hartmann-Kuhlein, U. Langer, E. Mahlow, A. Roth, S. Schaper, B. Urmoneit, and B. Woelker.** 1992. The complex for replication initiation of *Escherichia coli*. Chromosoma **102:**S1–S6.
16. **Miller, J. H.** 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Moriya, S., T. Atlung, F. G. Hansen, H. Yoshikawa, and N. Ogasawara.** 1992. Cloning of an autonomously replicating sequences (ars) from the *Bacillus subtilis* chromosome. Mol. Microbiol. **6:**309–315.
18. **Moriya, S., W. Firshein, H. Yoshikawa, and N. Ogasawara.** 1994. Replication of a *Bacillus subtilis oriC* plasmid in vitro. Mol. Microbiol. **12:**469–478.
19. **Moriya, S., K. Kato, H. Yoshikawa, and N. Ogasawara.** 1990. Isolation of a dnaA mutant of *Bacillus subtilis* defective in initiation of replication: amount of DnaA protein determines cells' initiation potential. EMBO J. **9:**2905–2910.
20. **Nelson, R. M., and G. L. Long.** 1989. A general method of site-specific mutagenesis using a modification of the *Thermus aquaticus* polymerase chain reaction. Anal. Biochem. **180:**147–151.

20a. **Qin, M.-H.** Unpublished data.

21. **Rajagopalan, M., M.-H. Qin, V. A. Steingrube, D. R. Nash, R. J. Wallace, Jr., and M. V. V. S. Madiraju.** 1995. Amplification and cloning of the *Mycobacterium tuberculosis dnaA* gene. Gene **163:**75–79.
22. **Rajagopalan, M., M. H. Qin, D. R. Nash, and M. V. V. S. Madiraju.** 1995. *Mycobacterium smegmatis dnaA* region and autonomous replication activity. J. Bacteriol. **177:**6527–6535.
23. **Salazar, L., H. Fsihi, E. de Rossi, G. Riccardi, C. Rios, S. T. Cole, and H. E. Takiff.** 1996. Organization of the origins of replication of the chromosomes of *Mycobacterium smegmatis*, *Mycobacterium leprae* and *Mycobacterium tuberculosis* and isolation of a functional origin from *M. smegmatis*. Mol. Microbiol. **20:**283–293.
24. **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.
25. **Schaefer, C., and W. Messer.** 1991. DnaA protein/DNA interaction: modulation of the recognition sequence. Mol. Gen. Genet. **226:**34–40.
26. **Skarstad, K., and E. Boye.** 1994. The initiator protein DnaA: evolution, properties and function. Biochim. Biophys. Acta **1217:**111–130.
27. **Smith, D. W., T. W. Yee, C. Baird, and V. Krishnapillai.** 1991. Pseudomonad replication origins: a paradigm for bacterial origins? Mol. Microbiol. **5:**2581–2587.
28. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. R. Hatful, S. B. Snapper, R. G. Barletta, W. R. Jacobs, and B. R. Bloom.** 1991. New use of BCG for recombinant vaccines. Nature **351:**456–460.
29. **Wheeler, P. R., and C. Ratledge.** 1994. Metabolism of *Mycobacterium tuberculosis*, p. 353–358. *In* B. R. Bloom (ed.), Tuberculosis: pathogenesis, protection, and control. American Society for Microbiology, Washington, D.C.
30. **Zakrzewska-Czerwinska, J., J. Majka, and H. Schrempf.** 1995. Minimal requirements of the *Streptomyces lividans* 66 *oriC* region and its transcriptional and translational activities. J. Bacteriol. **177:**4765–4771.