Transcription analysis of the *dnaA* gene and *oriC* region of the chromosome of *Mycobacterium smegmatis* and *Mycobacterium bovis* BCG, and its regulation by the DnaA protein

Leiria Salazar, Elba Guerrero, Yveth Casart, Lilia Turcios and Fulvia Bartoli

Structural Biology Department, Instituto Venezolano de Investigaciones Científicas (IVIC), Apartado 21827 Caracas, 1020A Venezuela

Correspondence
Leiria Salazar
lsalazar@ivic.ve

The regions flanking the *Mycobacterium dnaA* gene have extensive sequence conservation, and comprise various DnaA boxes. Comparative analysis of the *dnaA* promoter and *oriC* region from several mycobacterial species revealed that the localization, spacing and orientation of the DnaA boxes are conserved. Detailed transcriptional analysis in *M. smegmatis* and *M. bovis* BCG shows that the *dnaN* gene of both species and the *dnaA* gene of *M. bovis* BCG are transcribed from two promoters, whereas the *dnaA* gene of *M. smegmatis* is transcribed from a single promoter. RT-PCR with total RNA showed that *dnaA* and *dnaN* were expressed in both species at all growth stages. Analysis of the promoter activity using *dnaA*–*gfp* fusion plasmids and DnaA expression plasmids indicates that the *dnaA* gene is autoregulated, although the degree of transcriptional autorepression was moderate. Transcription was also detected in the vicinity of *oriC* of *M. bovis* BCG, but not of *M. smegmatis*. These results suggest that a more complex transcriptional mechanism may be involved in the slow-growing mycobacteria, which regulates the expression of *dnaA* and initiation of chromosomal DNA replication.

INTRODUCTION

The *dnaA* and *dnaN* genes, encoding the initiator protein DnaA and the β subunit of DNA polymerase III, respectively, are essential for DNA chromosome replication in eubacteria. *Escherichia coli* DnaA protein binds to four 9 bp sequences known as DnaA boxes within the *E. coli* origin of replication (*oriC*) and mediates open complex formation by making secondary contacts with three 13mer motifs within an A + T-rich region. DnaA also recruits DnaB helicase to the open complex, where it unwinds the origin and commits the chromosome to bidirectional replication (for review see Kaguni, 1997). The DNA polymerase III holoenzyme, the major bacterial replicase, directs the bidirectional replication of the chromosome. In *E. coli* it is a 900 kDa complex that contains several components: a catalytic core that includes the α subunit plus accessory subunits. The β subunit is a sliding DNA clamp responsible for tethering the polymerase to the DNA and endowing it with high processivity (for review see Kelman & O'Donnell, 1995).

The *dnaA* gene has been identified in many eubacteria and comparison at the amino acid sequence level has revealed

significant conservation (for review see Skarstad & Boye, 1994). The *dnaA* regulatory region of *E. coli* consists of two promoters, which are separated by one consensus DnaA box (Hansen *et al.*, 1982). Two functional promoters have also been mapped for the *dnaA* gene from *Pseudomonas putida* (Ingmer & Atlung, 1992), while only one promoter has been identified upstream of the *dnaA* gene from *Bacillus subtilis* (Moriya *et al.*, 1992), *Micrococcus luteus* (Fujita *et al.*, 1990), *Caulobacter crescentus* (Zweiger & Shapiro, 1994), *Streptomyces lividans* (Zakrzewska-Czerwinska *et al.*, 1994), *Mycoplasma capricolum* (Seto *et al.*, 1997) and *Thermus thermophilus* (Nardmann & Messer, 2000). In exponentially growing *E. coli* cells, *dnaN* is expressed predominantly from transcripts starting at the *dnaA* promoters (Pérez-Roger *et al.*, 1991); however, four promoters for *dnaN* have been detected in the second half of the *dnaA* structural gene (Quiñones & Messer, 1988; Armengod *et al.*, 1988), while in *B. subtilis dnaA* and *dnaN* constitute an operon (Ogura *et al.*, 2001).

Apart from its primary function as a replisome organizer, the DnaA protein acts as a regulatory protein. *In vivo* and *in vitro* studies have suggested that in *E. coli* the expression of the *dnaA* gene is negatively regulated by the interaction of its own protein product with the DnaA box in the promoter region (Atlung *et al.*, 1985). Within the *S. lividans dnaA* promoter region, two DnaA boxes have been found (Zakrzewska-Czerwinska *et al.*, 1994) and autoregulation of

---

the *dnaA* gene has also been demonstrated (Jakimowicz *et al.*, 2000). In *B. subtilis*, there are eight DnaA boxes in the *dnaA* promoter region, which are involved in the auto-repression of *dnaA* (Ogura *et al.*, 2001). However, mutations introduced into the DnaA boxes in the *dnaA* promoter region of *E. coli* (Smith *et al.*, 1997) and *Streptomyces coelicolor* (Jakimowicz *et al.*, 2000) did not have the expected effect of *dnaA* derepression. On the other hand, in *B. subtilis* the addition of extra DnaA boxes did not derepress the *dnaA–dnaN* operon (Moriya *et al.*, 1999), thereby suggesting a more complex and restrictive control for the regulation of the *dnaA* gene.

The two major obligate pathogens of the genus *Mycobacterium* are *Mycobacterium tuberculosis* and *Mycobacterium leprae*, the causative agents for tuberculosis and leprosy, respectively. In the past decades, development of effective antimicrobial therapy has significantly reduced the incidence of leprosy but tuberculosis (TB) still remain leading cause of death from any single infectious agent. According to the World Health Organization (2002) TB kills approximately 2 million people each year. In 1995 the global TB incidence was estimated at 8·8 million cases, while the projections suggest that TB incidence might be as high as 11·9 million by 2005 (Pio & Chaulet, 1998). On the other hand, one third of the world's population is currently infected with the TB bacillus, and individuals with latent tuberculosis carry a 2 to 23 % lifetime risk of developing reactivation of the disease later in life. The risk of reactivation dramatically increases ($\sim$5–10 % per year) under immuno-suppressive conditions, including HIV infection (Antonucci *et al.*, 1995). In countries with low or moderate tuberculosis endemicity, most cases of tuberculosis result from the reacti-vation of latent infection (Canetti *et al.*, 1972; van Rie *et al.*, 1999; Lillebaek *et al.*, 2002). Although there is evidence for the presence of tubercle bacilli in a nonreplicating persistent state in mammalian hosts (Parrish *et al.*, 1998), the nature of the tuberculosis bacterium in the latent infection state as well as the factors and stimuli that contribute to its reactivation are poorly understood. The enlightening of the molecular genetic aspects of *M. tuberculosis* chromosome replication, specifically its initiation and regulation, is important considering that in the latent state the tubercle bacillus is believed to persist in a metabolically active but non-growing state which can resume bacterial replication at an opportune time later in life (Bloom & MacKinney, 1999).

The genus *Mycobacterium* is composed of species with widely differing growth rates ranging from approximately 3 h in *Mycobacterium smegmatis* to 24 h in *M. tuberculosis*. The chromosomal region surrounding the origin of DNA replication in *M. smegmatis*, *M. tuberculosis*, *M. leprae* and *Mycobacterium avium* has been sequenced (Salazar *et al.*, 1996; Qin *et al.*, 1997; Madiraju *et al.*, 1999; Qin *et al.*, 1999), revealing an extensive sequence conservation in the intergenic regions flanking the *dnaA* gene. The *dnaA–dnaN* intergenic region has seven DnaA boxes arranged in a 165 bp segment while the *dnaA* regulatory region has three conserved DnaA boxes localized approximately 100 bases upstream of the *dnaA* start codon (Salazar *et al.*, 1996).

In this work, using *M. smegmatis* and *M. bovis* BCG as model systems of fast and slow-growing mycobacteria respectively, we report the characteristic features of the *dnaA* and *dnaN* regulatory regions. We have also determined the transcrip-tion of the *dnaA* and *dnaN* genes as well as the *oriC* region. In addition, analysis of promoter activity using DnaA box deletion mutants and quantitative determination of pro-moter repression by overexpression of the DnaA protein have revealed details of the regulation of the *dnaA* gene by the DnaA protein.

METHODS

Media, bacterial strains and growth conditions. *E. coli* XL-1 Blue cultures were grown in Luria–Bertani (LB) broth or on LB agar plates at 37 °C. *M. smegmatis* mc²155 (Snapper *et al.*, 1990) and *M. bovis* BCG Pasteur (ATCC 35734) were grown at 37 °C using Middlebrook 7H9 broth or 7H10 agar supplemented with 0·5 % (v/v) glycerol and 10 % (v/v) Middlebrook OADC (Difco). Tween 80 (0·05 %) was added to liquid media. The following concentrations of antibiotics were added when appropriate: carbeni-cillin, 50 µg ml^{-1}; kanamycin, 50 µg ml^{-1} (*E. coli*) or 25 µg ml^{-1} (mycobacteria).

Transcriptional fusion to *gfp* and fluorescence measurement. The shuttle plasmid pFPV27 (Valdivia *et al.*, 1996) was used to clone fragments fused to the *gfp* gene (Table 1). The *rpmH–dnaA* and *dnaA–dnaN* intergenic regions were obtained by PCR amplifica-tions. The *rpmH–dnaA* intergenic region was amplified using the primers LS60B (5′-GCGGATCCTGGAAGGTCCGGTTGCCCTTG-3′) and Sm15B (5′-GCGGATCCGGACGATTACCCCCTTTGAGG-3′) for *M. smegmatis*, and Mb19B (5′-GCGGATCCGTCTCCTCGCTATGT-CTG-3′) and Mb11B (5′-CCGGATCCGGTCAACGACGTATCTC-3′) for *M. bovis* BCG. The *dnaA–dnaN* intergenic region was amplified using the primers Sm11B (5′-AAGGATCCACGCTCGGCGGCTGT-GGA-3′) and Sm10B (5′-TTGGATCGCCCCTTCGATAATCCCCGCA-3′) for *M. smegmatis*, and ForiMb (5′-AAGGATCCTTCCGACAAC-GTTCTTAAAAA-3′) and RoriMb (5′-TTGGATCCTTTCACCTC-ACGATGAGTTC-3′) for *M. bovis* BCG. Genomic DNA or the pIV101 plasmid (Table 1) was used as template in the PCR reac-tions. The resulting fragments were cloned into the *Bam*HI site of pFPV27 generating plasmids pGFP85, pGFP11, pGFP61, pGFP8, pGFPS5, pGFPB7, pGFPS12 and pGFPB11. The pGFP87 and pGFP71 plasmids were derived from pGFP85 and have been pre-viously described (Salazar, 2000). The pGFP22 and pGFP16 plasmids were obtained by subcloning from pGFP11. Fragments containing shorter regions from the upstream *dnaA* region of *M. bovis* were obtained by PCR amplification and used in the construction of addi-tional transcriptional fusions. In a similar way, fragments containing shorter regions from the upstream *dnaN* region of *M. smegmatis* and *M. bovis* were obtained by PCR amplification and used in the construction of the plasmids shown in Table 1. The direction of the inserts was confirmed by mapping with restriction endonucleases and sequencing.

M. smegmatis mc²155 and *M. bovis* BCG cells bearing the transcrip-tional fusion to *gfp* were obtained by electroporation (Snapper *et al.*, 1990) and grown at 37 °C in 7H9 medium containing kanamycin. Aliquots (150 µl) of the cultures were taken at exponential and stationary growth phase for fluorescence measurements using a Spectrafluor Tecan (Microplate Reader).

Table 1. Plasmids used in this study

Plasmid	Relevant features	Reference or source
pFPV27	Km^r, shuttle vector for operon and gene fusion to *gfp* gene	Valdivia *et al.* (1996)
pOS239	3·3 kb *Bam*HI–*Bgl*II fragment containing *rpmH–dnaN* region of *M. smegmatis* cloned in pIJ963, Cb^r Hyg^r	Salazar *et al.* (1996)
pOS246	Deletion of 270 bp *Hin*dIII–*Eco*RI containing P$_{dnaA}$ from pOS239	Salazar (2000)
pIV101	~40 kb fragment from *M. smegmatis* mc^{2}6 containing *dnaA–gyrA–gyrB* genes cloned in pYUB18	Salazar *et al.* (1996)
pDNA6	1512 bp PCR fragment from pIV101 containing *dnaA* gene cloned in pGEX-4T1	This work
pDNA7	1521 bp PCR fragment from *M. bovis* BCG containing *dnaA* gene cloned in pGEX-4T1	This work

Fragments of the *rpmH−dnaA* intergenic region fused to *gfp* in the pFPV27 vector

Plasmid	Cloned region	Reference or source
pGFP85	540 bp PCR fragment from pOS239 (nt −540 to −1) cloned in the direction of *dnaA* gene	Salazar (2000)
pGFP61	540 bp PCR fragment from pOS239 (nt −540 to −1) cloned in the direction of *rpmH* gene	This work
pGFP87	396 bp PCR fragment from pOS239 (nt −540 to −145) cloned in the direction of *dnaA* gene	Salazar (2000)
pGFP71	265 bp PCR fragment from pOS246 (nt −540 to −421 and −145 to −1) cloned in the direction of *dnaA* gene	Salazar (2000)
pGFP11	608 bp PCR fragment from *M. bovis* (nt −601 to +7) cloned in the direction of *dnaA* gene	This work
pGFP8	608 bp PCR fragment from *M. bovis* (nt −601 to +7) cloned in the direction of *rpmH* gene	This work
pGFP22	Deletion of 155 bp *Nru*I–*Bam*HI (nt −151 to +7) fragment from pGFP11	This work
pGFP16	Deletion of 453 bp *Bam*HI–*Nru*I (nt −601 to −151) fragment from pGFP11	This work
pGFP30	430 bp PCR fragment from *M. bovis* (nt −423 to +7) cloned in the direction of *dnaA* gene	This work
pGFP9	430 bp PCR fragment from *M. bovis* (nt −423 to +7) cloned in the direction of *rpmH* gene	This work
pGFP22-4	Deletion of 158 bp *Nru*I–*Bam*HI (nt −151 to +7) fragment from pGFP30	This work
pGFP16-3	Deletion of 272 bp *Bam*HI–*Nru*I (nt −423 to −151) fragment from pGFP30	This work
pGFP6	223 bp PCR fragment from *M. bovis* (nt −216 to +7) cloned in the direction of *dnaA* gene	This work
pGFP7	223 bp PCR fragment from *M. bovis* (nt −216 to +7) cloned in the direction of *rpmH* gene	This work

Fragments of the *dnaA−dnaN* intergenic region fused to *gfp* in the pFPV27 vector

Plasmid	Cloned region	Reference or source
pGFPS5	446 bp PCR fragment from pIV101 (nt −455 to −10) cloned in the direction of *dnaN* gene	This work
pGFPS12	446 bp PCR fragment from pIV101 (nt −455 to −10) cloned in the direction of *dnaA* gene	This work
pGFPB5	252 bp PCR fragment from pIV101 (nt −13 to −264) cloned in the direction of *dnaN* gene	This work
pGFPB10	252 bp PCR fragment from pIV101 (nt −13 to −264) cloned in the direction of *dnaA* gene	This work
pGFPB16	212 bp PCR fragment from pIV101 (nt −456 to −245) cloned in the direction of *dnaN* gene	This work
pGFPB12	212 bp PCR fragment from pIV101 (nt −456 to −245) cloned in the direction of *dnaA* gene	This work
pGFPB7	511 bp PCR fragment from *M. bovis* (nt −516 to −6) cloned in the direction of *dnaN* gene	This work
pGFPB11	511 bp PCR fragment from *M. bovis* (nt −516 to −6) cloned in the direction of *dnaA* gene	This work
pGFPR2	271 bp PCR fragment from *M. bovis* (nt −276 to −6) cloned in the direction of *dnaN* gene	This work
pGFPR9	271 bp PCR fragment from *M. bovis* (nt −276 to −6) cloned in the direction of *dnaA* gene	This work
pGFPF12	261 bp PCR fragment from *M. bovis* (nt −516 to −256) cloned in the direction of *dnaN* gene	This work
pGFPF14	261 bp PCR fragment from *M. bovis* (nt −516 to −256) cloned in the direction of *dnaA* gene	This work

RNA extraction and primer extension analysis. The RNA isolation from *M. smegmatis* and *M. bovis* BCG and the primer extension reactions were performed according to Gónzalez-y-Merchand *et al.* (1996) with slight modifications. Briefly, exponential-phase cells were ruptured by four pulses of 45 s each (4 m s^{-1}), in a cell disrupter (FastPrep FP120, Bio 101-Savant). Four additional 15 s pulses at 5 m s^{-1} were applied to the *M. bovis* BCG cells. The lysate was extracted three to four times with 2 vols chloroform/isoamyl alcohol (24:1). The total RNA was precipitated by the dropwise addition of 0·5 vols cold ethanol and redissolved in the appropriate volume of DEPC-treated dH$_2$O. At least three synthetic oligonucleotides complementary to each strand of the upstream *dnaA* and *dnaN* sequences were 5′ end labelled with [γ–^{32}P]ATP and T4 polynucleotide kinase and used for the extension reactions. Each labelled primer (100 fmol) and 5–20 µg total RNA were annealed at 52 °C for 30 min. After cooling at room temperature, the primer extension
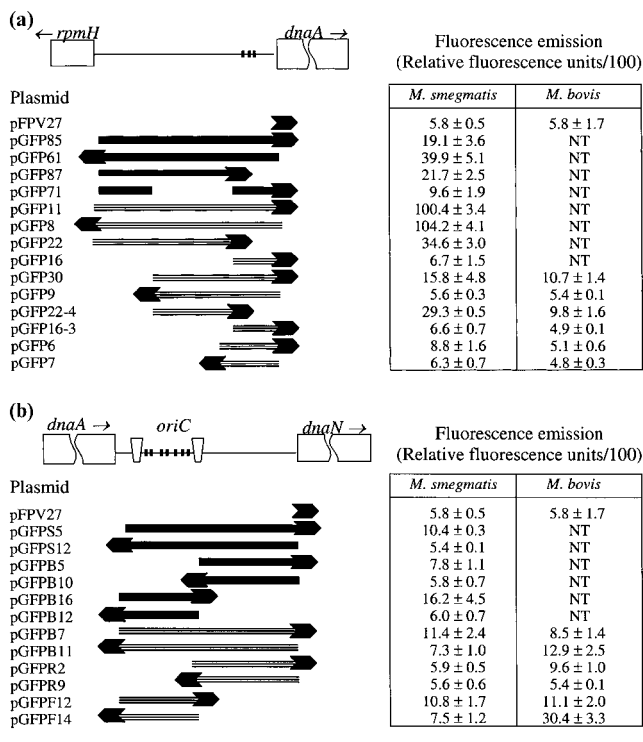
Fig. 1. dnaA–gfp (a) and dnaN–gfp (b) transcriptional fusion of *M. smegmatis* (solid rectangles) and *M. bovis* (striped rectangles), and measurement of the fluorescence emission. Transcriptional fusions were generated as described in Methods. Fluorescence was determined by spectrofluorometry and the specific promoter activity is expressed as relative fluorescence units at 535 nm (emission filter) corrected for the fluorescence emission of untransformed cells. The fluorescence activity was measured in the host cells *M. smegmatis* mc²155 or *M. bovis* BCG bearing transcriptional fusions with *gfp*. NT, not tested. The DnaA boxes (■) and A+T-rich regions (▯) are indicated. The pFPV27 plasmid was used as control and the black arrows represent the *gfp* gene. All measurements were carried out at least on triplicate cultures.

Panel (a) — *rpmH* / *dnaA*

Plasmid	Fluorescence emission (Relative fluorescence units/100)	
	M. smegmatis	*M. bovis*
pFPV27	5.8 ± 0.5	5.8 ± 1.7
pGFP85	19.1 ± 3.6	NT
pGFP61	39.9 ± 5.1	NT
pGFP87	21.7 ± 2.5	NT
pGFP71	9.6 ± 1.9	NT
pGFP11	100.4 ± 3.4	NT
pGFP8	104.2 ± 4.1	NT
pGFP22	34.6 ± 3.0	NT
pGFP16	6.7 ± 1.5	NT
pGFP30	15.8 ± 4.8	10.7 ± 1.4
pGFP9	5.6 ± 0.3	5.4 ± 0.1
pGFP22-4	29.3 ± 2.5	9.8 ± 1.6
pGFP16-3	6.6 ± 0.7	4.9 ± 0.1
pGFP6	8.8 ± 1.6	5.1 ± 0.6
pGFP7	6.3 ± 0.7	4.8 ± 0.3

Panel (b) — *dnaA* / *oriC* / *dnaN*

Plasmid	Fluorescence emission (Relative fluorescence units/100)	
	M. smegmatis	*M. bovis*
pFPV27	5.8 ± 0.5	5.8 ± 1.7
pGFPS5	10.4 ± 0.3	NT
pGFPS12	5.4 ± 0.1	NT
pGFPB5	7.8 ± 1.1	NT
pGFPB10	5.8 ± 0.7	NT
pGFPB16	16.2 ± 4.5	NT
pGFPB12	6.0 ± 0.7	NT
pGFPB7	11.4 ± 2.4	8.5 ± 1.4
pGFPB11	7.3 ± 1.0	12.9 ± 2.5
pGFPR2	5.9 ± 0.5	9.6 ± 1.0
pGFPR9	5.6 ± 0.6	5.4 ± 0.1
pGFPF12	10.8 ± 1.7	11.1 ± 2.0
pGFPF14	7.5 ± 1.2	30.4 ± 3.3

reaction was carried out with AMV reverse transcriptase (Promega) at 42 °C for 45 min. The extension products were separated on an 8 % polyacrylamide/urea gel, alongside a sequencing reaction generated using PCR fragments corresponding to the analysed sequence and the oligonucleotide used in the primer extension reaction as primers. The nucleotide sequence of the primers used can be given upon request.

Detection of mRNA by RT-PCR. Total RNA (0·5 μg) was reverse transcribed in a total volume of 20 μl containing 10 mM each dATP, dCTP, dGTP and dTTP, 2·5 μM reverse primer, 5 mM MgCl₂, 1×PCR buffer (100 mM Tris/HCl pH 8·3, 50 mM KCl), 20 U RNase inhibitor (Pharmacia) and 50 U MuLV reverse transcriptase (Roche). The RNA was denatured at 65 °C for 10 min and chilled on ice. After addition of the reaction mixture, the RT reaction was carried out at 42 °C for 30 min. The PCR reaction was performed in a final volume of 25 μl containing 5 μl cDNA template, 0·5 μM forward primer and *Taq* DNA polymerase (Gibco). The amplification was carried out for 30 cycles (95 °C for 1 min, 58 °C for 2 min and 72 °C for 2 min); each RT-PCR was

repeated three times. A 10 μl PCR sample from each reaction was subjected to electrophoresis on a 1·8 % agarose gel containing ethidium bromide. Non-reverse-transcribed PCR controls indicated the absence of contaminating genomic DNA and that the PCR products derived from mRNA.

Northern hybridization. Blot hybridization was performed following published protocols (Ausubel *et al.*, 1999). All solutions were prepared with DEPC-treated water. Briefly, 10 μg total RNA in each lane was separated in a denaturing agarose (1 %) gel containing formaldehyde (2·2 M) followed by partial hydrolysis (0·05 M NaOH, 1·5 M NaCl) and neutralization (0·5 M Tris/HCl pH 7·4, 1·5 M NaCl). The RNA was then transferred overnight by capillary action to Hybond-N+ (Amersham) and immobilized to the membrane by UV cross-linking. The membranes were then incubated in prehybridization solution (50 % formamide) at 42 °C for at least 3 h before the addition of probe [1–5 × 10⁵ c.p.m. (ml probe)⁻¹] labelled with [α-³²P]dCTP by random priming (Amersham). The probes were obtained by PCR amplification of coding regions of the *dnaA* and *dnaN* genes of *M. bovis* BCG with lengths of 1521 bp and 1197 bp respectively. The membranes were washed at high stringency and exposed for 2–10 days at −70 °C.

Regulation by DnaA protein. To investigate whether the *dnaA* and *dnaN* genes are subject to transcriptional regulation by the DnaA protein, the *dnaA* genes of *M. smegmatis* and *M. bovis* were expressed under the control of the P_{tac} promoter. Amplification by PCR was used to generate fragments encoding the DnaA protein. Considering that the first codon of *dnaA* of mycobacteria is a leucine (TTG; Salazar *et al.*, 1996), it was exchanged with ATG with the aim of improving the translation efficiency. The *dnaA* gene of *M. smegmatis* was amplified using the primers LS51 (5′-CGGGATCCATGACTGCTGACCCCGACCCAC-3′) and LS52 (5′-TAGCGGCCGCTCAGCGTTTGGCGCGCTGGC-3′) and DNA from pIV101 as template, while the *dnaA* gene of *M. bovis* BCG was amplified using the primers LS53 (5′-TAGCCGCCGCTAGCGCTT-GGAGCGCTGACG-3′) and LS54 (5′-CGGGATCCATGACCGAT-GCCCCGGTTCAG-3′) from genomic DNA. The PCR products were cloned into the *Bam*HI/*Not*I sites of pGEX-4T1 (Pharmacia Biotech). The resulting pDNA6 and pDNA7 plasmids were each co-transformed into *E. coli* XL-1 Blue with plasmids containing transcriptional fusion between the *dnaA* and *dnaN* promoter and the *gfp* gene (pGFP85, pGFP87, pGFP30 and pGFP22-4 plasmids; see Fig. 1). Transformed colonies were selected for kanamycin and carbenicillin resistance. *E. coli* cells harbouring both plasmids were grown in LB media with the appropriate antibiotics until exponential growth phase (OD₅₈₀ 0·7–0·8) was reached and 0·1 mM IPTG was added. The fluorescence emission was measured to assess the levels of *dnaA* or *dnaN* transcription with and without induction of the DnaA protein.

Other molecular techniques. Digestions, ligations, filling-in of protruding ends and plasmid DNA isolation were performed according to standard procedures. Amplified fragments and plasmid DNAs were sequenced with Sequenase 2.0 (USB, Amersham) and [α-³⁵S]dATP or with a dye terminator cycle sequencing kit and an ABI 377 sequencer (PE Biosystems), using the appropriate primers.

RESULTS

Determination of promoter activity in the *rpmH*, *dnaA* and *dnaN* regulatory regions

To identify the promoters responsible for the transcription of the *rpmH*, *dnaA* and *dnaN* genes, we cloned fragments of the *rpmH–dnaA* and *dnaA–dnaN* intergenic regions in the

pFPV27 vector upstream of the *gfp* reporter gene (Table 1). The constructs were tested for fluorescence emission in *M. smegmatis* mc²155, and selected *M. bovis* constructs were also tested in *M. bovis* BCG. The analysis of fluorescence emission of the cloned fragments is summarized in Fig. 1.

In *M. bovis* BCG, the nucleotide sequence upstream of the *dnaA* gene was obtained by PCR amplification. It was found to be highly similar to the corresponding region of *M. tuberculosis* H37Rv (accession nos X92504, ALO21426, AE007194 and U38891) and *M. bovis* strain AF2122/97 (spoligotype 9); the same region in *M. smegmatis* mc²6 has been previously reported under the accession no. X92503 (Salazar *et al.*, 1996).

The fragments containing the full-length *rpmH–dnaA* intergenic region emitted fluorescence regardless of the direction of cloning (pGFP85, pGFP11, pGFP61 and pGFP8 plasmids), suggesting that these regions carry the *rpmH* and *dnaA* promoter sequences.

Analysis of the subclones derived from pGFP85 (pGFP87 and pGFP71) showed that the *dnaA* promoter activity in *M.*

smegmatis was confined to the region between nt −540 and −145. Subclones and deletions derived from pGFP11 and pGFP30 (pGFP22 and pGFP22-4) showed that in *M. bovis* BCG, the plasmids whose DNA region extends from nt −601 to nt −151 have the majority of the transcriptional activity. However, fragments covering the 200 nt immediately upstream of the *dnaA* gene (pGFP6) emitted a fluorescence level slighter higher than those emitted by the cells carrying the control vector plasmid. The fluorescence emitted by pGFP6 was relatively weak but highly reproducible, and is abolished when the first 65 nt are deleted (pGFP16-3 plasmid).

The nucleotide sequences of the intergenic region *dnaA–dnaN* of *M. smegmatis* mc²6 (accession no. X92503) and *M. bovis* BCG (accession no. U75298) have been previously reported. Using specific primers we amplified the *dnaA–dnaN* intergenic regions of *M. smegmatis* and *M. bovis* BCG, and the fragments were cloned fused to the *gfp* reporter gene (see Methods and Table 1). The analysis of the fluorescence emission of the clones that carry the full-length *dnaA–dnaN* intergenic region of *M. smegmatis* (pGFPS5 and pGFPS12) showed the presence of promoter activity only when the fragment is fused to *gfp* in the *dnaN* transcription direction.
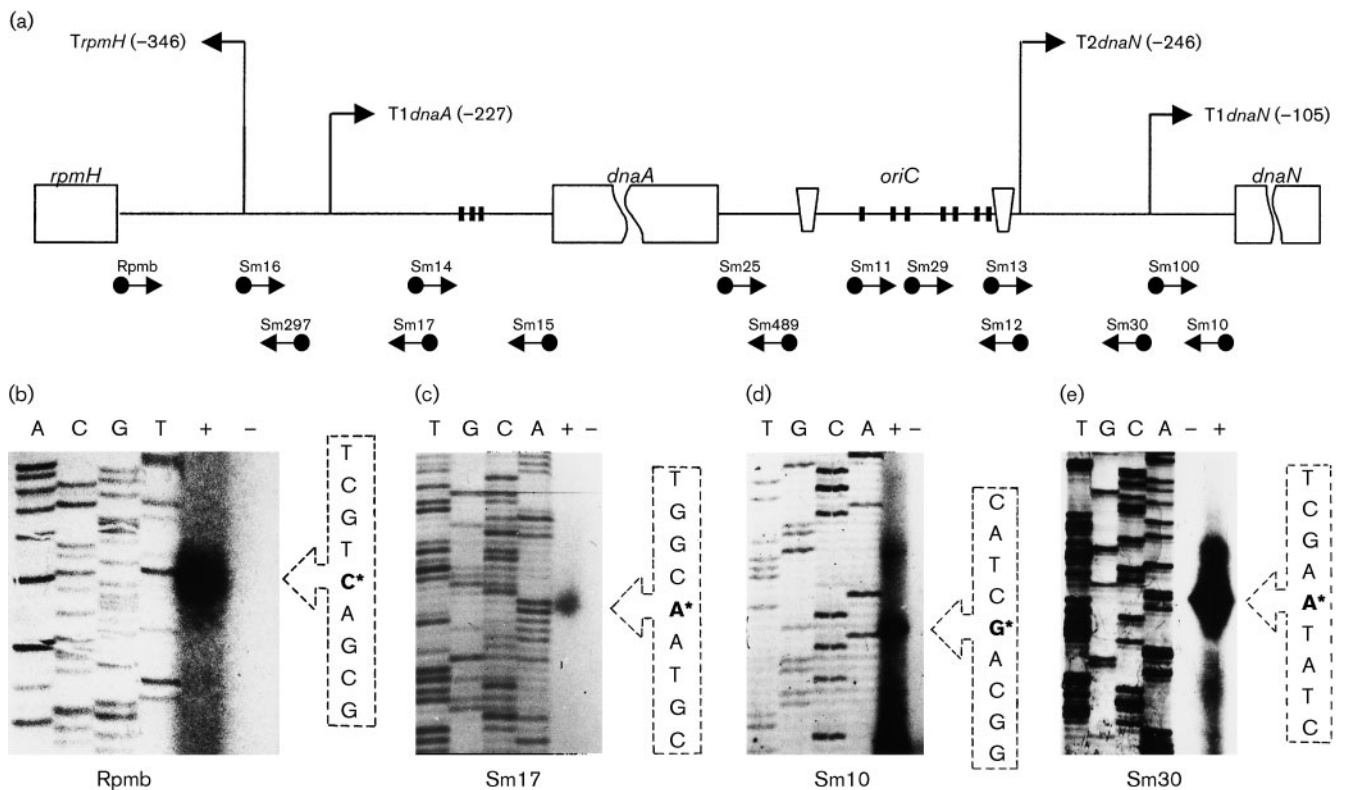
Fig. 2. Mapping the mRNA 5′ termini of the *rpmH–dnaA–dnaN* intergenic regions of *M. smegmatis* by primer extension. (a) Schematic representation of the *oriC* region showing the identified transcriptional start sites. The numbers in parentheses indicate the distance upstream from the *dnaA* or *dnaN* start. DnaA boxes (▮) and A+T-rich regions (▽) are indicated. (b–e) Primer extension using the oligos indicated. Asterisks show the transcription start points. Sequencing reactions with the same primer are also shown.

In contrast, the homologous region of *M. bovis* exhibited fluorescence activity independent of the direction of the cloned fragment. Fluorescence emission was also observed in shorter fragments of the *dnaA–dnaN* intergenic regions containing the first 250 nt upstream of *dnaN* (pGFPB5 and pGFPR2) as well as those fragments whose DNA region extends further upstream (pGFPB16 and pGFPF12),

suggesting the presence of more than one promoter sequence in this region. The pGFPB5 and pGFPR2 plasmids showed a fluorescence emission slighter higher than the control vector plasmid; these assays were repeated at least four times. However, plasmids pGFPB16 and pGFPF12, containing the seven DnaA boxes of *oriC*, showed a higher fluorescence activity than plasmids pGFPB5 and pGFPR2.
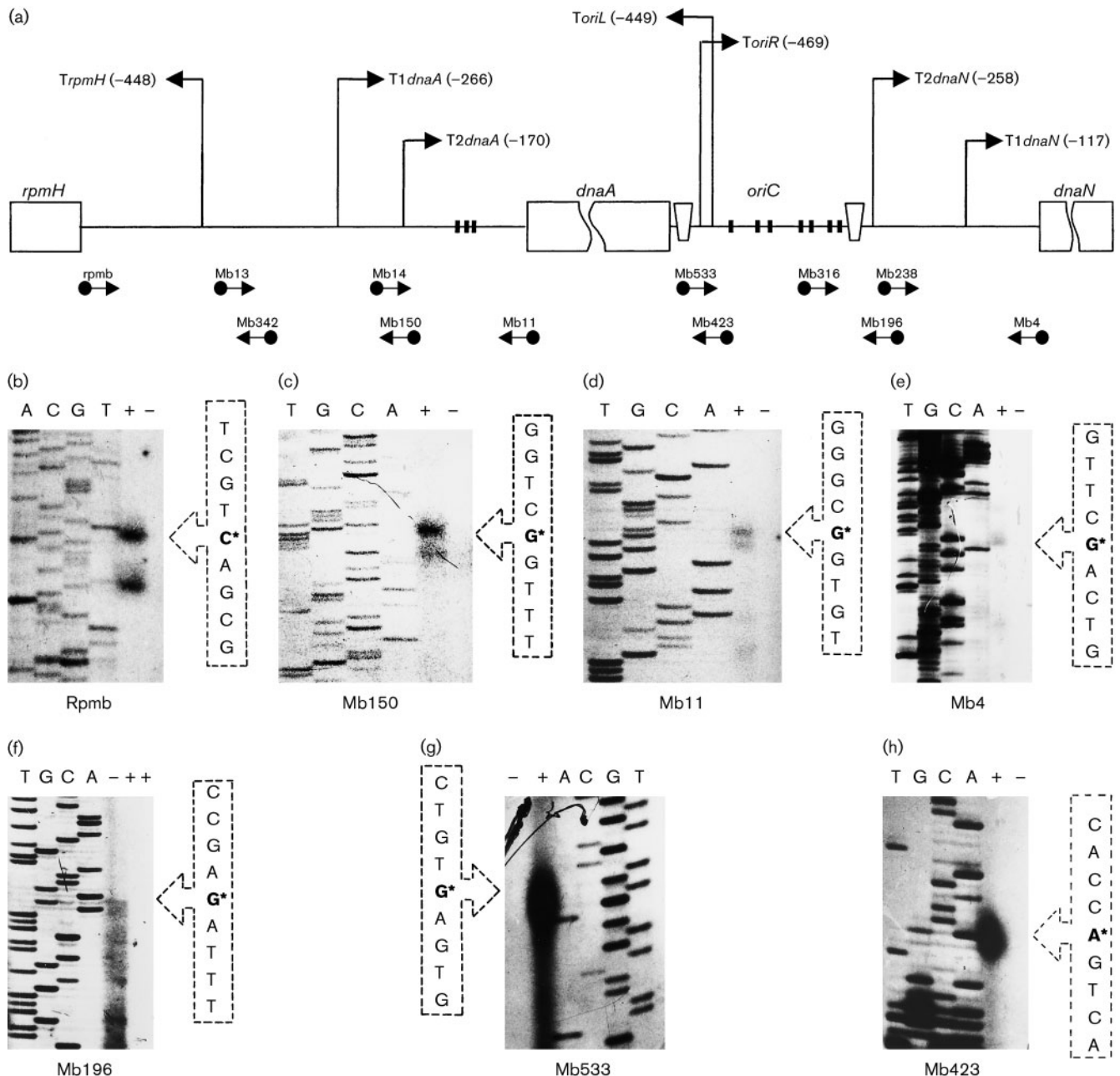
Fig. 3. Mapping the mRNA 5′ termini of the *rpmH–dnaA–dnaN* intergenic regions of *M. bovis* by BCG primer extension. (a) Schematic representation of the *oriC* region showing the identified transcriptional start sites. The numbers in parentheses indicate the distance upstream from the *dnaA* or *dnaN* start. DnaA boxes (▮) and A + T-rich regions (▽) are indicated. (b–h) Primer extension using the oligos indicated. Asterisks show the transcription start points. Sequencing reactions with the same primer are also shown.

Transcriptional analysis of the *rpmH–dnaA* and *dnaA–dnaN* intergenic regions

In an attempt to precisely localize the transcriptional start sites of the *rpmH*, *dnaA* and *dnaN* genes, several oligonucleotides were used with total RNA isolated from exponentially growing mycobacteria in primer extension experiments (Figs 2 and 3).

The divergent transcription in the *rpmH–dnaA* intergenic region was confirmed by the presence of mRNA 5′ ends for *rpmH* and *dnaA*. Using an oligonucleotide complementary to the first nucleotides of the *rpmH* coding sequence (oligonucleotide Rpmb), a unique 5′ end was identified (Figs 2b and 3b). This putative transcriptional start point (T*rpmH*) mapped to a conserved region in both *M. smegmatis* and *M. bovis* BCG at nt −346 and −448, respectively, 168 and 158 bases upstream of the translation start codon of *rpmH*. No other signals were identified upstream of *rpmH*, neither with these nor with oligonucleotides Sm16, Sm14, Mb13 and Mb14, nor by varying the annealing temperatures (data not shown). The mapped transcriptional start site is preceded by well conserved −35 (TTGACC) and −10 (c/aAGTACCCT) sequences, named P$_{rpmH}$ (Table 2), bearing a significant homology to the Group A *Mycobacterium* promoter recognition sequences (Gómez & Smith, 2000), similar to *E. coli* σ^{70}.

Using oligonucleotides complementary to the first nucleotides of the *dnaA* gene (oligonucleotides Sm15 and Mb11), one 5′ end was identified at nt −170 in *M. bovis* BCG (T2*dnaA*, Fig. 3c) while in *M. smegmatis* no signal was identified at the homologous position. However, using the oligonucleotides Sm17 and Mb150 we identified a 5′ end at nt −227 of *M. smegmatis* (Fig. 2c) and a 5′ end at nt −266

of *M. bovis* BCG (Fig. 3d), named T1*dnaA*, on a region of conserved sequence in both species. No additional signals were observed further upstream of T1*dnaA* and T2*dnaA*. Examination of the nucleotide sequence upstream of T1*dnaA* and T2*dnaA* revealed motifs resembling the −10 (TAGCTT and TTGAAC) and −35 (TTGGCA and TCGACT) hexamers of the Group A *Mycobacterium* consensus promoters (Table 2).

Two signals were identified in the *dnaA–dnaN* conserved intergenic region by using primer extension with oligonucleotides complementary to both strands. In both cases, the 5′ ends indicate that the mRNA must be transcribed in the direction of the *dnaN* gene (Figs 2d, 2e, 3e and 3f). One of these transcriptional start points (T1*dnaN*) mapped at nt −105 and nt −117 of *M. smegmatis* and *M. bovis* BCG, respectively. The second one (T2*dnaN*) mapped 141 bases further upstream of the T1*dnaN* previously identified in both species. Sequence inspection of the region upstream of T1*dnaN* and T2*dnaN* showed the presence of potential −35 (TTCAAG, TCCCCA) and −10 (TACGGT, TACTGT) highly conserved sequences (Table 2). These data suggest that the *dnaN* genes in both *M. smegmatis* and *M. bovis* BCG are transcribed from two promoters, and support the results found with the transcriptional fusions to *gfp*.

The chromosomal origin of replication of *M. smegmatis* and *M. bovis* BCG has been precisely mapped on the *dnaA–dnaN* intergenic region. Only the *dnaA–dnaN* intergenic region (Salazar *et al.*, 1996) or the 5′ flanking region of the *dnaA–dnaN* intergenic region of *M. smegmatis* (Qin *et al.*, 1997) were shown to promote its *oriC* activity. This region includes seven 9 bp DnaA protein-binding sites (DnaA boxes) flanked by A + T rich regions. The A + T rich region was located upstream of the first DnaA box has been

Table 2. Sequences for *rpmH*, *dnaA* and *dnaN* mycobacterial promoters

	Promoter sequence			
	−35	SP*	−10	SP†
P$_{rpmH}$ *M. smegmatis*	TTGACC	14	CAGTACCCT	6
P$_{rpmH}$ *M. bovis*	TTGACC	14	AAGTACCCT	6
P1$_{dnaA}$ *M. smegmatis*	TTGGCA	14	TGTTAGCTT	5
P1$_{dnaA}$ *M. bovis*	TTGGCA	14	TGTTAGCTT	5
P2$_{dnaA}$ *M. bovis*	TCGACT	12	AACTTGAAC	6
P1$_{dnaN}$ *M. smegmatis*	TTCAAG	13	CTCTACGGT	8
P1$_{dnaN}$ *M. bovis*	TTCAAG	13	CTCTACGGT	8
P2$_{dnaN}$ *M. smegmatis*	TCCCCA	14	TATTACTGT	6
P2$_{dnaN}$ *M. bovis*	TCCCCA	14	TAATACTGT	6
	Consensus sequence‡			
M. smegmatis	T$_{73}$T$_{58}$G$_{68}$a$_{26}$C$_{57}$a$_{36}$		t$_{47}$g$_{42}$t$_{47}$T$_{94}$A$_{84}$T$_{63}$a$_{42}$a$_{42}$T$_{78}$	
M. tuberculosis	T$_{62}$t$_{42}$G$_{76}$A$_{66}$C$_{71}$a$_{33}$		c$_{42}$G$_{52}$g$_{38}$T$_{76}$A$_{81}$K$_{66}$R$_{66}$a$_{42}$T$_{81}$	

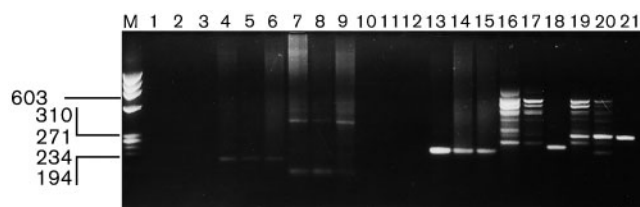*Length of the spacer between the −35 and −10 hexamers.

†Length of the spacer between the −10 hexamer and the transcriptional start point.

‡According to Gómez & Smith (2000).

proposed to be the site at which the local unwinding of DNA begins at initiation of replication. The oligonucleotides Mb423 and Mb533 permitted us to identify two additional mRNA 5′ ends in the *dnaA–dnaN* intergenic region, precisely on the origin of replication of *M. bovis* BCG, between the left A + T rich region and the first DnaA box (Fig. 3g and 3h). These putative transcriptional start points showed divergent transcription and were named T*oriL* and T*oriR*, mapping at nt −449 and at nt −469 respectively, upstream of *dnaN*. Although an exhaustive analysis was done on the *M. smegmatis* homologous region, using selected oligonucleotides and assaying at different annealing temperatures, no signals were observed (data not shown). These results are consistent with the fluorescence emission observed from plasmids containing the first half of the *dnaA–dnaN* intergenic region of *M. bovis* BCG and *M. smegmatis* (pGFPF12, pGFPF14, pGFPB16 and pGFPB12, Fig. 1). We did not detect promoter sequences resembling the *E. coli* σ^{70} consensus upstream T*oriL* and T*oriR*, suggesting that these transcripts must be expressed by σ factors other than σ^A or σ^B.

dnaA and dnaN are expressed at all growth phases

Northern blots were performed with separate gene-specific probes on RNA that was isolated from *M. smegmatis* and *M. bovis* BCG cells at various growth phases. Unfortunately, a hybridization smear was observed for both *dnaA* and *dnaN* probes, suggesting that the RNA transcripts were unstable. Within the smear, a pattern of at least five bands was

consistently found for the *dnaA* probe in the RNA from *M. bovis*, whose lengths were calculated as 892 ± 33, 1258 ± 95, 2288 ± 170, 4165 ± 373 and 7197 ± 430 bp (data not shown).

As an alternative method, RT-PCR analysis was performed to determine if the *dnaA* and *dnaN* genes are expressed at different growth phases. cDNA molecules were amplified using specific primers bound to the *dnaA* and *dnaN* start codons and RNA from *Mycobacterium* cultures at exponential and stationary growth phases (data not shown).
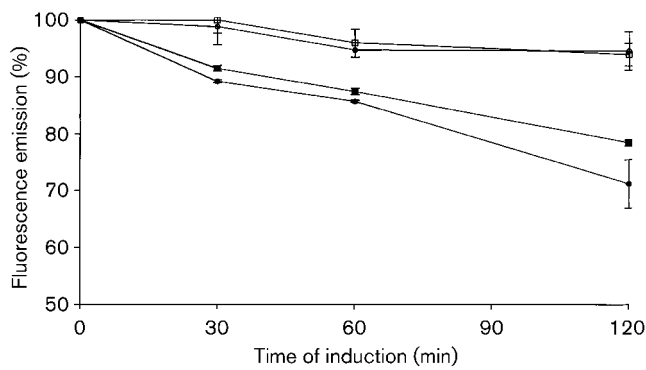
The RT-PCR products are shown in Fig. 4. cDNA molecules were obtained corresponding to the regions upstream of *dnaA* and *dnaN* of *M. smegmatis* and *M. bovis* BCG. Reverse transcriptase-dependent products of 160 bp (lanes 4–6) and 219 bp (lanes 13–15) for *dnaA*, and of 94 bp (lanes 7–9) and 252 bp (lanes 16–18) for *dnaN* were established, indicating that the *dnaA* and *dnaN* transcripts in both mycobacterial species were expressed during balanced growth and stationary phase. Additionally, a 330 bp reverse transcriptase-dependent product was observed with the Mb4 and Mb316 primers using RNA from *M. bovis* BCG (lanes 19–21), probably corresponding to T*oriR* transcripts. We have not detected PCR amplification products using the reverse primers Rpmb, Sm16, Sm13, Sm11, Mb13 and Mb533, confirming that there are no additional promoter sequences present further upstream of the transcripts previously located by primer extension. The additional unspecific amplifications observed in some cases could be attributed to the RT-PCR conditions.

dnaA transcription regulated by DnaA protein

To investigate whether the DnaA protein regulates transcription of the *dnaA* gene, we determined changes in the transcriptional activity driven by the *dnaA* promoter region under increasing levels of the intracellular DnaA protein.

It has been observed by fluorescence microscopy that the *dnaA* promoters of *M. smegmatis* and *M. bovis* BCG are expressed well in *E. coli* (data not shown); therefore the effect of the DnaA protein on the *dnaA* promoter activity was determined in *E. coli*. Plasmids containing *dnaA* promoter region of *M. smegmatis* and *M. bovis* BCG, with (pGFP85 and pGFP30) or without (pGFP87 and pGFP22-4) the DnaA box sequences, fused to the GFP reporter marker were each co-transformed into *E. coli* with plasmids containing the *dnaA* gene of the respective species under the control of the P_{tac} promoter (pDNA6 and pDNA7, see Methods). The fluorescence emission of cells bearing both plasmids (P_{dnaA}–*gfp* fusion and IPTG-induced DnaA clones) was determined at different concentrations of the DnaA protein. Changes in the intracellular concentrations of the DnaA protein were obtained by induction of *E. coli* cultures with IPTG, as confirmed by Western blot using anti-DnaA serum raised in rabbits (data not shown). The *dnaA* transcription levels were expressed as the percentage of

Fig. 4. RT-PCR analysis of the *dnaA* and *dnaN* genes. Total RNA was reverse transcribed with specific primers bound to the *dnaA* and *dnaN* start codons and amplified with selected primers complementary to the region upstream of each gene. The total RNA used was isolated from cultures at different growth phases. For *M. smegmatis* OD_{640} 0·5 (lanes 1, 4 and 7), 1·2 (lanes 2, 5 and 8) and 2·7 (lanes 3, 6 and 9), and for *M. bovis* OD_{640} 0·4 (lanes 10, 13, 16 and 19), 0·8 (lanes 11, 14, 17 and 20) and 1·3 (lanes 12, 15, 18 and 21) were used. M, molecular mass marker: ϕX174 RF DNA/*Hae*III. Lanes: 1–3, control PCR amplifications without reverse transcriptase using RNA samples from *M. smegmatis* and primers Sm15/Sm14; 4–6, RT-PCR using primers Sm15/Sm14; 7–9, RT-PCR using the primers Sm10/Sm100; 10–12, control PCR amplifications without reverse transcriptase using RNA samples from *M. bovis* and primers Mb11/Mb14; 13–15, RT-PCR using primers Mb11/Mb14; 16–18, RT-PCR using primers Mb4/Mb238; 19–21, RT-PCR using primers Mb4/Mb316.

Fig. 5. Regulation of *dnaA* expression in the presence of DnaA. *E. coli* cells bearing the indicated plasmids were grown in liquid media to exponential growth phase; half the volume was taken and induced with IPTG. Aliquots were taken at the intervals indicated to determine the percentage of fluorescence emission of the *dnaA–gfp* transcriptional fusion in the presence of inducible DnaA. ●, pDNA6 + pGFP85; ○, pDNA6 + pGFP87; ■, pDNA7 + pGFP30; □, pDNA7 + pGFP22-4. The pGFP and pDNA plasmids are described in Table 1 and the text.

fluorescence emission with induction/non-induction. As a control, the fluorescence emission of each P_{dnaA}–*gfp* plasmid (singly transformed) was determined and showed that the levels of fluorescence were indistinguishable with and without induction of DnaA expression. As shown in Fig. 5, *dnaA* promoter activity was clearly reduced after 30 min DnaA induction. This decrease in the percentage of fluorescence emission was observed when the *dnaA* promoter region cloned included the three DnaA boxes located within the *dnaA* promoter region (pDNA6 + pGFP85 and pDNA7 + pGFP30). *dnaA* promoter-driven fluorescence emission decreased nearly 25 % after 2 h DnaA induction. However, the percentage of fluorescence emission was unaltered by induction of the DnaA protein in the absence of the DnaA boxes (pDNA6 + pGFP87 and pDNA7 + pGFP22-4). These experiments suggest that the DnaA protein is able to transcriptionally repress expression of the *dnaA* promoter, and that the DnaA boxes are involved in this regulation. The influence of the DnaA boxes in the P_{dnaA} region on the expression level of the *dnaA* gene is supported by the results found with the reporter gene analysis. Deletion of the three DnaA boxes (compare pGFP85 with pGFP87, and pGFP30 with pGFP22-4) resulted in an increase in the fluorescence emission (Fig. 1 and Salazar, 2000).
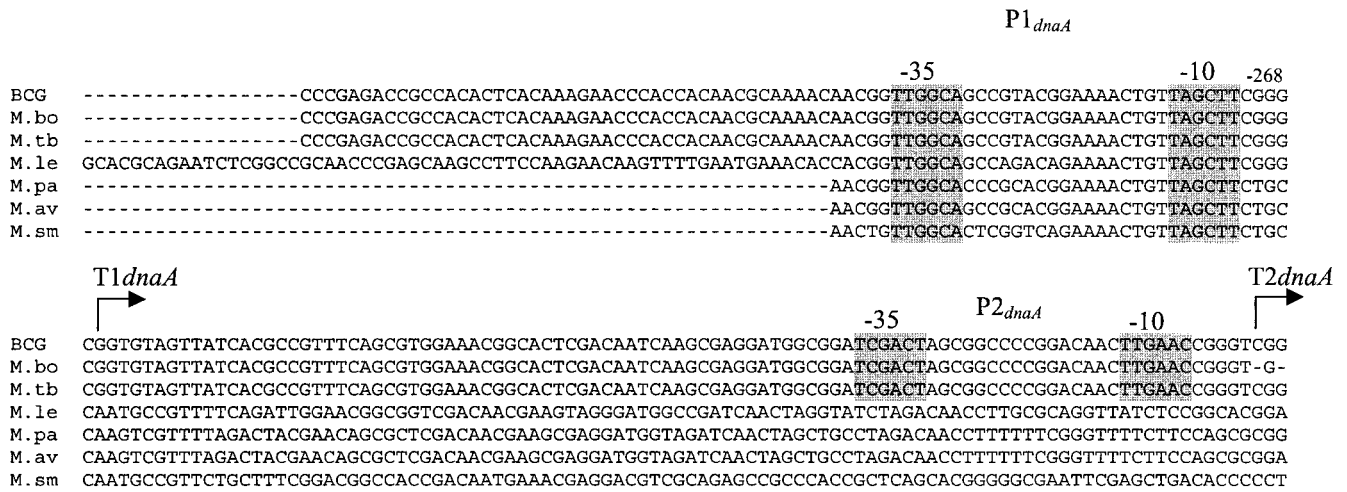
DISCUSSION

The major components of the *E. coli* DNA replication machinery have been identified and characterized (see Kornberg & Baker, 1992) and the presence of DnaA, the initiator protein, in many eubacteria suggests a conserved mechanism. However, many questions concerning the cell cycle regulation of initiation at the origin of chromosomal replication remain unanswered. The mycobacterial *dnaA* and *dnaN* genes are located flanking *oriC*, in a gene order

that is well conserved among other Gram-positive organisms. In this study, we have found that *M. smegmatis* and *M. bovis* BCG, species representing the fast and slow-growing mycobacteria respectively, have clear differences in the transcriptional pattern of the *dnaA* gene and at *oriC*. This conclusion is based on the results obtained from reporter gene expression, primer extension analysis and RT-PCR of the region.

All the transcriptional start sites (TSPs) identified for the *rpmH*, *dnaA* and *dnaN* genes are preceded by a well conserved −35 ($T_{100}T_{67}G_{56}A_{56}C_{78}N$) and −10 ($T_{100}A_{89}C_{67}C_{67}NT_{89}$) promoter region with characteristic features of σ^A and σ^B *Mycobacterium* promoters, which has homology to the *E. coli* σ^{70} sequence consensus (Table 2). The nucleotide initiating at the TSP is most frequently A, with a distance of 5–8 bp between the TSP and the −10 hexamer, and with a spacing of 12–14 nt between the −35 and the −10 regions. This coincides with our observations that the $P1_{dnaA}$ of *M. smegmatis* and *M. bovis* BCG were well expressed in *E. coli* (Fig. 5). Moreover, the close similarity of the P_{rpmH} and P_{dnaA} promoters to the *Mycobacterium* σ^A consensus and the high fluorescence emission observed for P_{rpmH}–*gfp* and P_{dnaA}–*gfp* transcriptional fusions (Fig. 1) suggests that these promoters, if not subjected to any regulatory constraints, would act as strong promoters *in vivo*. Although $P2_{dnaA}$ of *M. bovis* does not have the conserved T in position two of the −35 hexamer and A and T in positions two and six of the −10 hexamer (Table 2), we propose that all promoter sequences identified can be recognized *in vivo* by the mycobacterial housekeeping sigma factor, homologous to *E. coli* σ^{70}.

Examination of the nucleotide sequence of the *dnaA* regulatory region of *M. tuberculosis*, *M. leprae* (Salazar *et al.*, 1996), *M. avium* (Madiraju *et al.*, 1999) and *M. avium* subsp. *paratuberculosis* (accession no. AF222789) shows that the −35 and −10 sequences of $P1_{dnaA}$ are also conserved at homologous positions and are located within a region of more extensive homology between these species (Fig. 6), suggesting that $P1_{dnaA}$ corresponds to the main mycobacterial *dnaA* promoter. However, we have found that *dnaA* gene of *M. bovis* BCG could be expressed from two different promoters (Fig. 3a, c, d) and that $P2_{dnaA}$ must contribute substantially to the *dnaA* expression (Fig. 1). When the search for additional transcriptional factor binding sites was extended to the nucleotide sequence of mycobacterial homologous regions already published, a $P2_{dnaA}$ homologous sequence was found only in the *dnaA* regulatory region of strains belonging to the *M. tuberculosis* complex (Fig. 6).

Analysis of the sequence in the region surrounding the *dnaA* promoters of *M. bovis* has revealed some interesting features. Immediately upstream of $P1_{dnaA}$ of *M. bovis* BCG, as well as in *M. tuberculosis* H37Rv and *M. leprae* (Salazar *et al.*, 1996), there is a short non-conserved sequence (55–73 nt), which is not present in *M. smegmatis* or *M. avium*. This short sequence might well be a rich playground for the interaction

P1$_{dnaA}$

```
                                                  -35                    -10    -268
BCG  ------------------CCCGAGACCGCCACACTCACAAAGAACCCACCACAACGCAAAACAACGGTTGGCAGCCGTACGGAAAACTGTTAGCTTCGGG
M.bo ------------------CCCGAGACCGCCACACTCACAAAGAACCCACCACAACGCAAAACAACGGTTGGCAGCCGTACGGAAAACTGTTAGCTTCGGG
M.tb ------------------CCCGAGACCGCCACACTCACAAAGAACCCACCACAACGCAAAACAACGGTTGGCAGCCGTACGGAAAACTGTTAGCTTCGGG
M.le GCACGCAGAATCTCGGCCGCAACCCGAGCAAGCCTTCCAAGAACAAGTTTTGAATGAAACACCACGGTTGGCAGCCAGACGAAAACTGTTAGCTTCGGG
M.pa ----------------------------------------------------------------AACGGTTGGCACCCGCACGGAAAACTGTTAGCTTCTGC
M.av ----------------------------------------------------------------AACGGTTGGCAGCCGCACGGAAAACTGTTAGCTTCTGC
M.sm ----------------------------------------------------------------AACTGTTGGCACTCGGTCAGAAAACTGTTAGCTTCTGC
```

T1$dnaA$ T2$dnaA$

```
                                              -35        P2$_{dnaA}$       -10
BCG  CGGTGTAGTTATCACGCCGTTTCAGCGTGGAAACGGCACTCGACAATCAAGCGAGGATGGCGGATCGACTAGCGGCCCCGGACAACTTGAACCGGGTCGG
M.bo CGGTGTAGTTATCACGCCGTTTCAGCGTGGAAACGGCACTCGACAATCAAGCGAGGATGGCGGATCGACTAGCGGCCCCGGACAACTTGAACCGGGT-G-
M.tb CGGTGTAGTTATCACGCCGTTTCAGCGTGGAAACGGCACTCGACAATCAAGCGAGGATGGCGGATCGACTAGCGGCCCCGGACAACTTGAACCGGGTCGG
M.le CAATGCCGTTTTCAGATTGGAACGGCGGTCGACAACGAAGTAGGGATGGCCGATCAACTAGGTATCTAGACAACCTTGCGCAGGTTATCTCCGGCACGGA
M.pa CAAGTCGTTTTTAGACTACGAACAGCGCTCGACAACGAAGCGAGGATGGTAGATCAACTAGCTGCCTAGACAACCTTTTTTCGGGTTTTCTTCCAGCGCGG
M.av CAAGTCGTTTAGACTACGAACAGCGCTCGACAACGAAGCGAGGATGGTAGATCAACTAGCTGCCTAGACAACCTTTTTTCGGGTTTTCTTCCAGCGCGGA
M.sm CAATGCCGTTCTGCTTTCGGACGGCCACCGACAATGAAACGAGGACGTCGCAGAGCCGCCCACCGCTCAGCACGGGGGCGAATTCGAGCTGACACCCCCT
```
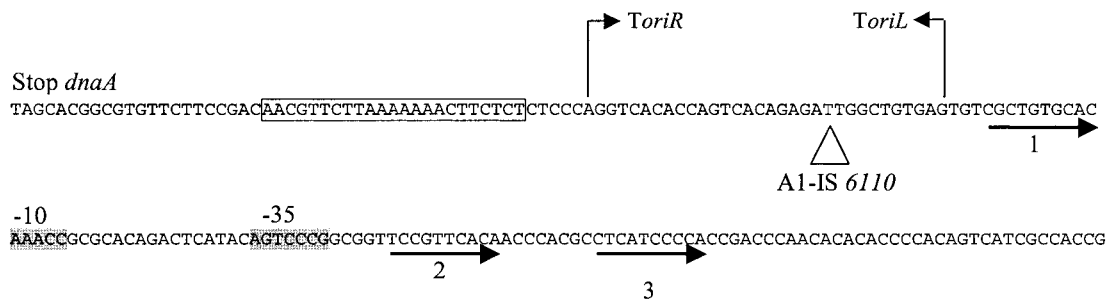
Fig. 6. Alignment of the *dnaA* regulatory region. Arrows showing the direction of transcription indicate the *dnaA* transcriptional start points (T1*dnaA* and T2*dnaA*) identified by primer extension. The proposed −10 and −35 sequences are shown in grey boxes. BCG, *M. bovis* BCG; M. bo, *M. bovis* strain AF2122/97; M. tb, *M. tuberculosis* H37Rv; M. le, *M. leprae*; M. pa, *M. avium* subsp. *paratuberculosis*; M. av, *M. avium*; M. sm, *M. smegmatis* mc^{2}6.

of DNA and architectural proteins, such as the bend-induced proteins.

In this work, we have observed a reduction in *dnaA* promoter activity when the intracellular concentration of the DnaA protein was increased (Fig. 5), suggesting that in mycobacteria the *dnaA* gene is autoregulated. The three DnaA boxes upstream of *dnaA* seem to be implicated in this regulation. Autoregulation of the *dnaA* gene by direct interaction of the DnaA protein with DnaA boxes has been demonstrated in *E. coli* (Atlung *et al.*, 1985), *S. lividans* (Jakimowicz *et al.*, 2000) and *B. subtilis* (Ogura *et al.*, 2001). However, no autoregulation of the *dnaA* gene has been observed in *P. putida* (Ingmer & Atlung, 1992), in spite of the fact that its regulatory region contains eight DnaA binding domains, nor in *Synechocystis* sp. (Richer & Messer, 1995), which has none.

The analysis of the *dnaA–dnaN* intergenic regions examined

here indicates that the *dnaN* gene is expressed from two different promoters (Figs 2 and 3). Despite the fact that the mycobacterial *dnaN* regulatory region presents a limited sequence homology, the −35 and −10 sequences identified (P1$_{dnaN}$ and P2$_{dnaN}$) are highly conserved (Table 2). The spatial conservation of these sequences in *M. tuberculosis*, *M. leprae* (Salazar *et al.*, 1996), *M. avium* (Madiraju *et al.*, 1999) and *M. avium* subsp. *paratuberculosis* (GenBank accession no. AF222789) raises the possibility that the mycobacterial *dnaN* gene could be expressed from the described promoters. Although we have not found protein binding motifs associated with the *dnaN* promoters, the overlap of the −35 region of the P2$_{dnaN}$ with DnaA box six of the *oriC* region, would suggest that this promoter may also be regulated by DnaA. Unfortunately, we cannot determine the influence of the DnaA boxes located in the *dnaA–dnaN* intergenic region on the transcription level of *dnaN* because *E. coli* cells bearing plasmids containing the *dnaN*

Tor*iR*   Tor*iL*

```
Stop dnaA
TAGCACGGCGTGTTCTTCCGACAACGTTCTTAAAAAAACTTCTCTCTCCCAGGTCACACCAGTCACAGAGATTGGCTGTGAGTGTCGCTGTGCAC
                                                                              A1-IS 6110         1

-10              -35
AAACCGCGCACAGACTCATACAGTCCCGGCGGTTCCGTTCACAACCCACGCCTCATCCCCACCGACCCAACACACACCCCACAGTCATCGCCACCG
          2                 3
```

Fig. 7. Transcription start points in the *oriC* region of *M. bovis* BCG. △, A1-IS*6110* insertional site (Kurepina *et al.*, 1998). The white boxed region indicates the left A+T rich region. Arrows indicate DnaA boxes.

promoter fused to *gfp* emitted a weak fluorescence, practically indistinguishable from cells bearing the control plasmid (pFPV27).

Interestingly, we have detected two additional transcripts in the *dnaA*–*dnaN* intergenic region of *M. bovis* BCG, between the left A + T-rich region and the DnaA box one at the *oriC* region, that initiate in opposite directions from each other (Figs 3 and 7). Examination of the nucleotide sequence surrounding these transcripts has not revealed feasible σ^A or σ^B promoter sequences. However, 12 and 21 nt upstream of T*oriL*, potential −10 (GGTTT) and −35 (CGGGAC) consensus sequences recognized by σ^H, were detected (Fig. 7). σ^H is a mycobacterial ECF (extra-cytoplasmic function) σ factor homologue of *S. coelicolor* σ^R, and is involved in the heat shock response and oxidative stress (Raman *et al.*, 2001; Manganelli *et al.*, 2002; Kaushal *et al.*, 2002). The functional significance of these transcripts is not known, nevertheless, it might be speculated that they are associated with regulation of the initiation of *oriC* in *M. bovis*. It has been suggested that the transcription of genes flanking the *E. coli oriC* participates in a positive–negative interplay during initiation. The *gidA* gene, located to the left of the *E. coli oriC*, is transcribed leftward away from *oriC* and plays a positive role in initiation (Asai *et al.*, 1990, 1992), while the mRNA transcribed from the *mioC* gene, located on the right side of the *E. coli oriC*, enters and goes through *oriC*, playing a negative role in initiation (Nozaki *et al.*, 1988; Tanaka & Hiraga, 1985). It will be interesting to determine whether the transcripts detected in the *oriC* region of *M. bovis* BCG are also present in other slow-growing mycobacteria, especially in the pathogenic species *M. tuberculosis*, and evaluate the relationship between their expression, the initiation of the DNA replication and pathogenicity. Kurepina *et al.* (1998) reported that in certain *M. tuberculosis* strain lineages, the *oriC* region is an IS*6110* hotspot, where at least ten different insertion sites have been identified. It will be important to determine if the relatively large IS*6110* insertion in the A1 site, which mapped between T*oriR* and T*oriL* to 15 nt of the first DnaA box (Fig. 7), affects *oriC* activity. Recently, it was reported that IS*6110* insertions in the A4 site, disrupting the DnaA box two, abolished *oriC* plasmid activity, although no effect on chromosomal replication in *M. tuberculosis* was observed (Dziadek *et al.*, 2002). Experiments in this direction are currently under way.

Combining all our evidence, we propose that the *dnaA* gene expression as well as the regulation of chromosomal replication initiation of the slow-growing mycobacteria, such as *M. bovis* and *M. tuberculosis*, seem to be subjected to a fine-tuned regulation.

ACKNOWLEDGEMENTS

REFERENCES

Antonucci, G., Girardi, E., Raviglione, M. C. & Ippolito, G. (1995). Risk factors for tuberculosis in HIV-infected persons: a prospective cohort study. *JAMA* **274**, 143–148.

Armengod, M. E., García-Sogo, M. & Lambíes, E. (1988). Transcriptional organization of the *dnaN* and *recF* genes of *Escherichia coli* K-12. *J Biol Chem* **263**, 12109–12114.

Asai, T., Takanami, M. & Imai, M. (1990). The AT richness and *gid* transcription determine the left border of the replication origin of the *E. coli* chromosome. *EMBO J* **9**, 4065–4072.

Asai, T., Chen, C. P., Nagata, T., Takanami, M. & Imai, M. (1992). Transcription *in vivo* within the replication origin of the *Escherichia coli* chromosome: a mechanism for activating initiation of replication. *Mol Gen Genet* **231**, 169–178.

Atlung, T., Clausen, E. S. & Hansen, F. G. (1985). Autoregulation of the *dnaA* gene of *Escherichia coli* K12. *Mol Gen Genet* **200**, 442–450.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smith, J. A. & Struhl, K. (1999). *Current Protocols in Molecular Biology*, 4th edn. New York: Wiley.

Bloom, B. R. & MacKinney, J. D. (1999). The death and resurrection of tuberculosis. *Nat Med* **5**, 872–874.

Canetti, G., Sutherland, I. & Svandova, E. (1972). Endogenous reactivation and exogenous reinfection: their relative importance with regards to the development of non-primary tuberculosis. *Bull Int Union Tuberc* **47**, 122–143.

Dziadek, J., Rajagopalan, M., Parish, T., Kurepina, N., Greendyke, R., Kreiswirth, B. N. & Madiraju, M. (2002). Mutations in the CCGTTCACA DnaA box of *Mycobacterium tuberculosis oriC* that abolish replication of *oriC* plasmids are tolerated on the chromosome. *J Bacteriol* **184**, 3848–3855.

Fujita, M. Q., Yoshikawa, H. & Ogasawara, N. (1990). Structure of the *dnaA* region of *Micrococcus luteus*: conservation and variations among eubacteria. *Gene* **93**, 73–78.

Gómez, M. & Smith, I. (2000). Determinants of mycobacterial gene expression. In *Molecular Genetics of Mycobacteria*, pp. 111–129. Edited by G. F. Hatfull & W. R. Jacobs, Jr. Washington, DC: American Society for Microbiology.

Gónzalez-y-Merchand, J. A., Colston, M. J. & Cox, R. A. (1996). The rRNA operons of *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: comparison of promoter elements and of neighbouring upstream genes. *Microbiology* **142**, 667–674.

Hansen, F. G., Hansen, E. B. & Atlung, T. (1982). The nucleotide sequence of the *dnaA* gene promoter and of the adjacent *rpmH* gene coding for the ribosomal protein L34 of *E. coli*. *EMBO J* **1**, 1043–1048.

Ingmer, H. & Atlung, T. (1992). Expression and regulation of a *dnaA* homologue isolated from *Pseudomonas putida*. *Mol Gen Genet* **232**, 431–439.

Jakimowicz, D., Majka, J., Lis, B., Konopa, G., Wegrzyn, G., Messer, W., Schrempf, H. & Zakrzewska-Czerwinska, J. (2000). Structure and regulation of the *dnaA* promoter region in three *Streptomyces* species. *Mol Gen Genet* **262**, 1093–1102.

Kaguni, J. M. (1997). *Escherichia coli* DnaA protein: the replication initiator. *Mol Cell* **7**, 145–157.

Kaushal, D., Schroeder, B. G. & Tyagi, S. (2002). Reduced inmunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative σ factor, SigH. *Proc Natl Acad Sci U S A* **99**, 8330–8335.

Kelman, Z. & O'Donnell, M. (1995). DNA polymerase III holo-enzyme: structure and function of a chromosomal replicating machine. *Annu Rev Biochem* **64**, 171–200.

Kornberg, A. & Baker, T. A. (1992). *DNA Replication*, 2nd edn. New York: WH Freeman.

Kurepina, N. E., Sreevatsan, S., Plikaytis, B. B., Bifani, P. J., Connell, N. D., Donnelly, R. J., van Soolingen, D., Musser, J. M. & Kreiswirth, B. N. (1998). Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS6110 elements in *Mycobacterium tuberculosis*: non-random integration in the *dnaA-dnaN* region. *Tubercle Lung Dis* **79**, 31–42.

Lillebaek, T., Dirksen, A., Baess, I., Strunge, B., Thomsen, V. O. & Andersen, A. B. (2002). Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. *J Infect Dis* **185**, 401–404.

Madiraju, M. V. V. S., Qin, M.-H., Yamamoto, K., Atkinson, M. A. L. & Rajagopalan, M. (1999). The *dnaA* gene region of *Mycobacterium avium* and the autonomous replication activities of its 5′ and 3′ flanking regions. *Microbiology* **145**, 2913–2921.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K., Dubnau, E., Gómez, M. & Smith, I. (2002). Role of the extracytoplasmic-function σ factor σ^{H} in *Mycobacterium tuberculosis* global gene expresion. *Mol Microbiol* **45**, 365–374.

Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H. & Ogasawara, N. (1992). Cloning of an autonomously replicating sequence *(ars)* from *Bacillus subtilis* chromosome. *Mol Microbiol* **6**, 309–315.

Moriya, S., Imai, Y., Hassan, A. K. M. & Ogasawara, N. (1999). Regulation of initiation of *Bacillus subtilis* chromosome replication. *Plasmid* **41**, 17–29.

Nardmann, J. & Messer, W. (2000). Identification and characterization of the *dnaA* upstream region of *Thermus thermophilus*. *Gene* **261**, 299–303.

Nozaki, N., Okazaki, T. & Ogawa, T. (1988). *In vitro* transcription of the origin region of replication of the *Escherichia coli* chromosome. *J Biol Chem* **263**, 14176–14183.

Ogura, Y., Imai, Y., Ogasawara, N. & Moriya, S. (2001). Auto-regulation of the *dnaA-dnaN* operon and effects of DnaA protein levels on replication initiation in *Bacillus subtilis*. *J Bacteriol* **183**, 3833–3841.

Parrish, N. M., Dick, J. D. & Bishai, W. R. (1998). Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol* **6**, 107–112.

Pérez-Roger, I., García-Sogo, M., Navarro-Aviño, J. P., López-Acedo, C., Macián, F. & Armengod, M. E. (1991). Positive and negative regulatory elements in the *dnaA-dnaN-recF* operon of *Escherichia coli*. *Biochimie* **73**, 329–334.

Pio, A. & Chaulet, P. (1998). Revising TB control programmes. In *Tuberculosis Handbook*, pp. 9–29. Geneva: Global Tuberculosis Programme, World Health Organization.

Qin, M. H., Madiraju, M., Zachariah, S. & Rajagopalan, M. (1997). Characterization of the *oriC* region of *Mycobacterium smegmatis*. *J Bacteriol* **179**, 6311–6317.

Qin, M. H., Madiraju, M. & Rajagopalan, M. (1999). Characterization of the functional replication origin of *Mycobacterium tuberculosis*. *Gene* **233**, 121–130.

Quiñones, A. & Messer, W. (1988). Discoordinate gene expression in the *dnaA–dnaN* operon of *E. coli*. *Mol Gen Genet* **213**, 118–124.

Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs, W. R., Jr & Husson, R. N. (2001). The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J Bacteriol* **183**, 6119–6125.

Richer, S. & Messer, M. (1995). Genetic structure of the *dnaA* region of the cyanobacterium *Synechocystis* sp. strain PCC6803. *J Bacteriol* **177**, 4245–4251.

Salazar, L. (2000). Inhibition of chromosome replication in *Mycobacterium smegmatis*: effect of the *rpmH–dnaA* promoter region. *Microbiology* **146**, 2199–2207.

Salazar, L., Fsihi, H., de Rossi, E., Riccardi, G., Rios, C., Cole, S. T. & Takiff, H. E. (1996). Organization of the origins of replication of the chromosome of *Mycobacterium smegmatis*, *Mycobacterium leprae* and *Mycobacterium tuberculosis* and isolation of a functional origin from *M. smegmatis*. *Mol Microbiol* **20**, 283–293.

Seto, S., Murata, S. & Miyata, M. (1997). Characterization of *dnaA* gene expression in *Mycoplasma capricolum*. *FEMS Microbiol Lett* **150**, 239–247.

Skarstad, K. & Boye, E. (1994). The initiator protein DnaA: evolution, properties and function. *Biochim Biophys Acta* **1217**, 111–130.

Smith, R. W. P., McAteer, S. & Masters, M. (1997). Autoregulation of the *Escherichia coli* replication initiator protein, DnaA, is indirect. *Mol Microbiol* **23**, 1303–1315.

Snapper, S. B., Melton, R. E., Mustapha, S., Kieser, T. & Jacobs, W. R. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**, 1911–1919.

Tanaka, M. & Hiraga, S. (1985). Negative control of *oriC* plasmid replication by transcription of the *oriC* region. *Mol Gen Genet* **200**, 21–26.

Valdivia, R. H., Hromockyj, A. E., Monack, D., Ramakrishnan, L. & Falkow, S. (1996). Applications for the green fluorescent protein (GFP) in the study of host-pathogen interactions. *Gene* **173**, 47–52.

van Rie, A., Warren, R., Richardson, M., Victor, T. C., Gie, R. P., Enarson, D. A., Beyers, N. & van Helden, P. D. (1999). Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *N Engl J Med* **341**, 1174–1179.

World Health Organization (2002). *Global Tuberculosis Control: Surveillance, Planning, Financing: WHO Report 2002*. WHO/CDS/TB/2002.295. Geneva: World Health Organization.

Zakrzewska-Czerwinska, J., Nardmann, J. & Schrempf, H. (1994). Inducible transcription of the *dnaA* gene from *Streptomyces lividians* 66. *Mol Gen Genet* **242**, 440–447.

Zweiger, G. & Shapiro, L. (1994). Expression of *Caulobacter dnaA* as a function of the cell cycle. *J Bacteriol* **176**, 401–408.