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

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 (Quinones & 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 & Chaul, 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 (~5–10 % per year) under immunosuppressive conditions, including HIV infection (Antonacci et al., 1995). In countries with low or moderate tuberculosis endemicity, most cases of tuberculosis result from the reactivation of latent infection (Canetti et al., 1972; van Rie et al., 1999; Lillebaek et al., 2002). Although there is evidence for the presence of the 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 transcription 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 promoter 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 (0-0.5%) was added to liquid media. The following concentrations of antibiotics were added when appropriate: carbenicillin, 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 amplifications. The rpmH–dnaA intergenic region was amplified using the primers LS60B (5’-GCGGATCTGGAGGTTGTCCTTGGT-3’) and Sm15B (5’-GCGGATCCGGAAGATTACGCTTGGG-3’) for M. smegmatis, and Mb19B (5’-GCGGATCTGGCTTCTGCTATGG-3’) and Mb11B (5’-GCGGATCCGGTCAACGAGATCTG-3’) for M. bovis BCG. The dnaA–dnaN intergenic region was amplified using the primers Sm11B (5’-AAGAGATCCACGTCGCGGTGCTG-3’) and Sm10B (5’-TTGGATCCGGCTTGATATACTCGCGCA-3’) for M. smegmatis, and ForMb (5’-AAGGATCTCCGAGAACGCCGTCGTTCA-3’) and RoriMb (5’-TTGGATCCGGCTTCACCACTTGCATGGTTC-3’) for M. bovis BCG. Genomic DNA or the pIV101 plasmid (Table 1) was used as template in the PCR reactions. The resulting fragments were cloned into the BamHI site of pFPV27 generating plasmids pGFP85, pGFP11, pGFP61, pGFP8, pGFP85, pGFPB7, pGFP12 and pGFPB11. The pGFP87 and pGFP71 plasmids were derived from pGFP85 and have been previously 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 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 transcriptional 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

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant features</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pFPV27</td>
<td>Km’, shuttle vector for operon and gene fusion to gfp gene</td>
<td>Valdivia et al. (1996)</td>
</tr>
<tr>
<td>pOS239</td>
<td>3-3 kb BamHI–BglII fragment containing rpmH–dnaN region of M. smegmatis cloned in plP963, Cb′ Hyg′</td>
<td>Salazar et al. (1996)</td>
</tr>
<tr>
<td>pOS246</td>
<td>Deletion of 270 bp HindIII–EcoRI containing P_madA from pOS239</td>
<td>Salazar (2000)</td>
</tr>
<tr>
<td>pIV101</td>
<td>~40 kb fragment from M. smegmatis mc26 containing dnaA–gyrA–gyrB genes cloned in pYUB18</td>
<td>Salazar et al. (1996)</td>
</tr>
<tr>
<td>pDNA6</td>
<td>1512 bp PCR fragment from pIV101 containing dnaA gene cloned in pGEX-4T1</td>
<td>This work</td>
</tr>
<tr>
<td>pDNA7</td>
<td>1521 bp PCR fragment from M. bovis BCG containing dnaA gene cloned in pGEX-4T1</td>
<td>This work</td>
</tr>
</tbody>
</table>

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

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

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

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Cloned region</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGFP5</td>
<td>446 bp PCR fragment from pIV101 (nt –455 to –10) cloned in the direction of dnaN gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP12</td>
<td>446 bp PCR fragment from pIV101 (nt –455 to –10) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP5</td>
<td>252 bp PCR fragment from pIV101 (nt –13 to –264) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP10</td>
<td>252 bp PCR fragment from pIV101 (nt –13 to –264) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP16</td>
<td>212 bp PCR fragment from pIV101 (nt –456 to –245) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP12</td>
<td>212 bp PCR fragment from pIV101 (nt –456 to –245) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP7</td>
<td>511 bp PCR fragment from M. bovis (nt –516 to –6) cloned in the direction of dnaN gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP11</td>
<td>511 bp PCR fragment from M. bovis (nt –516 to –6) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP2</td>
<td>271 bp PCR fragment from M. bovis (nt –276 to –6) cloned in the direction of dnaN gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP9</td>
<td>271 bp PCR fragment from M. bovis (nt –276 to –6) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP12</td>
<td>261 bp PCR fragment from M. bovis (nt –516 to –256) cloned in the direction of dnaN gene</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP14</td>
<td>261 bp PCR fragment from M. bovis (nt –516 to –256) cloned in the direction of dnaA gene</td>
<td>This work</td>
</tr>
</tbody>
</table>

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 González-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 disruptor (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/isomyl 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 dH₂O. At least three synthetic oligonucleotides complementary to each strand of the upstream dnaA and dnaN sequences were 5′ end labelled with ³²P-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

http://mic.sgmjournals.org
transcriptase (Roche). The RNA was denatured at 65 °C for 10 min with 20 U RNase inhibitor (Pharmacia) and 50 U MuLV reverse transcriptase, and the nucleotide sequence of the primers used can be given on request. The extension products were separated on an 8 % polyacrylamide/urea gel, alongside a sequencing reaction generated using PCR fragments corresponding to the analysed sequence

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 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 45 min. The extension products were separated on a 1 % agarose (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 (1:8 %) 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 × 10⁶ 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 (TG; 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'-TACGGCCGGCTAGCGCTTGGCCGGCTGCCG-3') and DNA from pIV101 as template, while the dnaN gene of M. bovis BCG was amplified using the primers LS53 (5'-TACGGCCGGCTAGCGCTTGGCCGGCTGCCG-3') and LS54 (5'-CGGGATCCATGACTGCTGACCCCGACCCAC-3') from genomic DNA. The PCR products were cloned into the BamHI/NotI 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 reached 0:7–0:8) and reached 0:1 mM IPTG, which 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-32P]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 AE2122/97 (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 slightly 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* (pGFBB5 and pGFPR2) as well as those fragments whose DNA region extends further upstream (pGFBB16 and pGFPR12), suggesting the presence of more than one promoter sequence in this region. The pGFBB5 and pGFPR2 plasmids showed a fluorescence emission slightly higher than the control vector plasmid; these assays were repeated at least four times. However, plasmids pGFBB16 and pGFPR12, containing the seven DnaA boxes of *oriC*, showed a higher fluorescence activity than plasmids pGFBB5 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 (TrpmH) 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/aAGTACCCCT) sequences, named PrpmH (Table 2), bearing a significant homology to the Group A Mycobacterium promoter recognition sequences (Gómez & Smith, 2000), similar to E. coli s70.

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 (T2dnaA, 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 T1dnaA, on a region of conserved sequence in both species. No additional signals were observed further upstream of T1dnaA and T2dnaA. Examination of the nucleotide sequence upstream of T1dnaA and T2dnaA revealed motifs resembling the −10 (TAGCTT and TTGAAC) and −35 (TTGGCA and TGGACT) 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 (T1dnaN) mapped at nt −105 and nt −117 of M. smegmatis and M. bovis BCG, respectively. The second one (T2dnaN) mapped 141 bases further upstream of the T1dnaN previously identified in both species. Sequence inspection of the region upstream of T1dnaN and T2dnaN 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

<table>
<thead>
<tr>
<th>Promoter sequence</th>
<th>−35</th>
<th>SP*</th>
<th>−10</th>
<th>SP†</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrpmH M. smegmatis</td>
<td>TTGACC</td>
<td>14</td>
<td>CAGTACCT</td>
<td>6</td>
</tr>
<tr>
<td>PrpmH M. bovis</td>
<td>TTGACC</td>
<td>14</td>
<td>AAGTACCT</td>
<td>6</td>
</tr>
<tr>
<td>P1dnaA M. smegmatis</td>
<td>TTGCA</td>
<td>14</td>
<td>TGGTACCT</td>
<td>5</td>
</tr>
<tr>
<td>P1dnaA M. bovis</td>
<td>TTGCA</td>
<td>14</td>
<td>TGGTACCT</td>
<td>5</td>
</tr>
<tr>
<td>P2dnaA M. bovis</td>
<td>TGCAGT</td>
<td>12</td>
<td>AAGTACCC</td>
<td>6</td>
</tr>
<tr>
<td>P1dnaN M. smegmatis</td>
<td>TTCAG</td>
<td>13</td>
<td>CTCGTACCT</td>
<td>8</td>
</tr>
<tr>
<td>P1dnaN M. bovis</td>
<td>TTCAG</td>
<td>13</td>
<td>CTCGTACCT</td>
<td>8</td>
</tr>
<tr>
<td>P2dnaN M. smegmatis</td>
<td>TCCCCA</td>
<td>14</td>
<td>TGGTACCT</td>
<td>6</td>
</tr>
<tr>
<td>P2dnaN M. bovis</td>
<td>TCCCCA</td>
<td>14</td>
<td>TGGTACCT</td>
<td>6</td>
</tr>
</tbody>
</table>

Consensus sequence

M. smegmatis

| T7T6G3G6g3G5g3G7g3a36 |
| T7G3g3T4T9A4T3A3a3T3a3 |
| c4G3G3g6T7g6A6g6a4a2T8 |

M. tuberculosis

| T6g3G3G3G3G6G6G6g3G7a33 |

*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).
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 TorI\(\text{L}\) and TorI\(\text{R}\), 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 \(\sigma^70\) consensus upstream TorI\(\text{L}\) and TorI\(\text{R}\), suggesting that these transcripts must be expressed by \(\sigma\) factors other than \(\sigma^A\) or \(\sigma^B\).

**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 TorI\(\text{R}\) transcripts. We have not detected PCR amplification products using the reverse primers Rpmb, Sm16, Sm13, Sm11, Mb13 and Mb333, 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\(_{\text{arc}}\) promoter (pDNA6 and pDNA7, see Methods). The fluorescence emission of cells bearing both plasmids (P\(_{\text{dnaA}}\)-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...
fluorescence emission with induction/non-induction. As a control, the fluorescence emission of each $p_{dnaA}\rightarrow gfp$ 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 ($p_{DNA6}+p_{GFP85}$ and $p_{DNA7}+p_{GFP30}$). $dnaA$ 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 ($p_{DNA6}+p_{GFP87}$ and $p_{DNA7}+p_{GFP22-4}$). 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 $p_{GFP85}$ with $p_{GFP87}$, and $p_{GFP30}$ with $p_{GFP22-4}$) 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 $-10$ ($T_{100}A_{89}G_{67}C_{67}N_{78}$) promoter region with characteristic features of $\sigma^A$ and $\sigma^B$ Mycobacterium promoters, which has homology to the $E. coli$ $\sigma^70$ 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 $P_{1_{dnaA}}$ 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 $\sigma^A$ consensus and the high fluorescence emission observed for $P_{rpmH}\rightarrow gfp$ and $P_{dnaA}\rightarrow gfp$ transcriptional fusions (Fig. 1) suggests that these promoters, if not subjected to any regulatory constraints, would act as strong promoters in vivo. Although $P_{2_{dnaA}}$ 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 $E. coli\ \sigma^{70}$.

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 $P_{1_{dnaA}}$ are also conserved at homologous positions and are located within a region of more extensive homology between these species (Fig. 6), suggesting that $P_{1_{dnaA}}$ 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 $P_{2_{dnaA}}$ 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 $P_{2_{dnaA}}$ 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 $P_{1_{dnaA}}$ 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
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 (P1dnaN and P2dnaN) 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 P2dnaN 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

---

**Fig. 6.** Alignment of the dnaA regulatory region. Arrows showing the direction of transcription indicate the dnaA transcriotional start points (T1dnaA and T2dnaA) 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 mc².6.

---

**Fig. 7.** Transcription start points in the oriC region of M. bovis BCG. △, A1-IS6110 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 TorIL, potential −10 (GGTTT) and −35 (CAGGAC) consensus sequences recognized by σ^H, were detected (Fig. 7). σ^H is a mycobacterial ECF (extra-ribosomal function) σ factor homologue of S. coelicolor σ^N, 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 IS6110 hotspot, where at least ten different insertion sites have been identified. It will be important to determine if the relatively large IS6110 insertion in the A1 site, which mapped between TorIR and TorIL to 15 nt of the first DnaA box (Fig. 7), affects oriC activity. Recently, it was reported that IS6110 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

We thank J. Rivas for photographic work and A. Sánchez for technical support. We are grateful to J. Rosales and J. Concepción for the M. smegmatis and M. tuberculosis DnaA antiserum. This work was supported by grants from Fondo Nacional de Investigaciones Científicas y Tecnológicas – Venezuela (S1-97000023 and S1-200100176), and the European Commission through its INCO programme (ICA4-2201-10187).

REFERENCES


