The Mycobacterium tuberculosis oriC (the origin of chromosomal replication) region contains 13 non-perfect DnaA boxes. The M. tuberculosis initiator protein, DnaA, was overexpressed in Escherichia coli as a soluble His-tagged fusion protein. The purified protein His6-MtDnaA was investigated for its binding properties to DnaA boxes from the oriC region. Gel retardation demonstrated that the DnaA from M. tuberculosis requires two DnaA boxes for efficient binding. Electron microscopy as well as Dnase I footprinting showed that the His6-MtDnaA protein binds to four specific regions, which correspond to the location of 11 out of 13 previously identified DnaA boxes within the M. tuberculosis oriC. Probably, in M. tuberculosis, DnaA molecules by co-operative binding of numerous ‘non-perfect’ DnaA boxes assemble along the oriC region and subsequently form a massive nucleoprotein complex.

Key words: DnaA, initiation of chromosome replication, Mycobacterium tuberculosis, origin of chromosomal replication (oriC).

INTRODUCTION

Replication of chromosomes is strictly regulated to ensure that all copies are replicated exactly once per cell cycle [1,2]. In all branches of life, the replication is controlled at the initiation stage. Therefore the events that occur at the replication origin play a central role in the cell cycle. In bacteria, replication starts at the unique site, oriC (origin of chromosomal replication), and continues bidirectionally around the chromosome until the two replication forks meet in the terminus region opposite to the oriC or reach chromosomal ends. Initiation of bacterial chromosome replication is mediated by the initiator DnaA protein, which interacts with repetitive, non-palindromic, nonamer sequences, the DnaA boxes, which are localized within the oriC region [3,4]. Among bacteria, the initiation of replication is best understood in Escherichia coli. Binding of 10–20 DnaA protein molecules to five DnaA boxes of the E. coli oriC region promotes unwinding within the AT-rich region of the oriC [4].

TB (tuberculosis) kills 2 million people each year. This global epidemic is growing and becoming more dangerous. The World Health Organization (http://www.who.int/en/) estimates that between 2002 and 2020, approx. 1000 million people will be newly infected, over 150 million people will get sick and 36 million will die of TB, if control is not further strengthened. In man, TB is usually caused by Mycobacterium tuberculosis. The genus Mycobacterium includes numerous species ranging from rapidly-growing organisms, M. smegmatis and M. fortuitum (generation time is 3–4 h), to slow-growing organisms, such as M. tuberculosis (24 h) and M. leprae (2 weeks) [5]. Assuming that the rate of chromosomal DNA synthesis in mycobacteria is approximately the same as in other bacterial organisms (such as E. coli), the mycobacterial DNA replication machinery must be completely repressed for most of the cell cycle, particularly in slow growers. Recently, the genome sequences of four different mycobacteria species including two M. tuberculosis strains [6,7] were determined. However, our knowledge of mycobacterial cell cycle and initiation of chromosome replication is scarce. In spite of the fact that the DNA fragments containing the dnaA–dnaN intergenic region from M. avium, M. leprae, M. smegmatis and M. tuberculosis were shown to promote autonomous replication [8–12], little is known about interactions between the two key elements of the initiation of mycobacterial chromosome, oriC and DnaA protein. In contrast with E. coli, the structure of mycobacterial oriC region is more complex; the oriC is longer than the one in E. coli and possesses numerous ‘non-perfect’ DnaA boxes. Each of the 13 DnaA boxes located within the M. tuberculosis oriC [12] region differs at least by one base from the perfect E. coli DnaA box (TTATCCACA). Thus it is important to understand more of the initiation of DNA replication in these slow-growing organisms. In the present study, we describe the procedure for purification of the soluble M. tuberculosis DnaA protein and we characterize binding requirements of this protein.

MATERIALS AND METHODS

Bacterial strains, media and culture conditions

Epicurian Coli™ XL1-Blue [endA1, gyrA46, hsdR17, lac, recA1, supE44, thi, F’ lac: lacIq, Δ(lacZ)M15, Tn10, proA, proB] and DH5α [F–, F80lacZΔM15, recA1, endA1, gyrA96, thi-1, hsdR17, (r–, mK–)], supE44, relA1, deoR, Δ(lacZYA-argF)U169 [13] served as host for cosmids and plasmids. E. coli BL21 [F− dcm ompT hsdS(r− m− s−)gal] was the host for overproduction of the fusion protein His6–DnaA. E. coli strains were grown in Luria–Bertani medium at 30 or 37 °C. Antibiotics were used at the following concentrations: ampicillin (100 µg/ml) for E. coli, kanamycin (30 µg/ml) for plasmids and tetracyclin (12.5 µg/ml) for Epicurian Coli™ XL1-Blue.

DNA manipulations

Cosmids, plasmids and DNA fragments were purified using kits according to the manufacturer’s instructions (Qiagen, Chatsworth,
The pellet was frozen at −80 °C and kept until further purification. The cell suspension was incubated on ice for 30 min. The cells were lysed by sonication and centrifuged at 34,500 g for 50 min. The supernatant was precipitated with 0.34 g/ml ammonium sulphate. The precipitated proteins were centrifuged at 27,000 g for 30 min, the pellet was suspended in 2.5 ml of LG buffer [45 mM Hepes/KOH (pH 7.6), 200 mM potassium glutamate, 10 mM magnesium acetate, 0.5 mM EDTA and 5 mM 2-mercaptoethanol] and desalted on a PD-10 column (Amersham Biosciences) according to the manufacturer’s instructions. The sample was loaded on to the Ni-NTA (Ni2+-nitrilotriacetate)–agarose column, previously equilibrated with LG buffer. After washing with LG buffer, the His6MtDnaA protein was eluted with an imidazole gradient (1–100 mM imidazole) and analysed by SDS/PAGE.

**DnaA purification**

The *M. tuberculosis* dnaA gene PCR amplified by pDNAABamHI primer and pDNAANotI primer, (using cosmid I441 as a template) was cloned into pGEM-T Easy. The authenticity of the pGEM-T Easy derivative containing the *M. tuberculosis* dnaA gene (this study) was checked by sequencing of both strands. The plasmid was digested with restriction enzymes KpnI and EcoRI, transformed into *E. coli* strain BL21 and grown in LB medium containing ampicillin and 1 mM IPTG (isopropyl β-D-thiogalactoside) at 30 °C for 30 min in Marians’ binding buffer [20 mM Hepes/KOH (pH 7.6), 5 mM magnesium acetate, 1 mM EDTA, 4 mM dithiothreitol, 0.2% Triton X-100, 3 mM ATP and 50 μg/ml BSA] [16]. The bound complexes were separated by PAGE (4 or 6% gel) [0.25 × TBE (22.25 mM Tris/borate/0.25 mM EDTA)], at 4 V/cm, 4 °C. Gels were dried and analysed by a Typhoon 8600 Variable Mode Imager and ImageQuant software.

**Gel retardation assay**

For binding assays, 32P-labelled DNA was incubated with DnaA protein in the presence of non-specific competitor [poly(dI/dC); Roche] at room temperature (20 °C) for 30 min in Marians’ binding buffer [20 mM Hepes/KOH (pH 7.6), 5 mM magnesium acetate, 1 mM EDTA, 4 mM dithiothreitol, 0.2% Triton X-100, 3 mM ATP and 50 μg/ml BSA] [16]. The bound complexes were separated by PAGE (4 or 6% gel) [0.25 × TBE (22.25 mM Tris/borate/0.25 mM EDTA)], at 4 V/cm, 4 °C. Gels were dried and analysed by a Typhoon 8600 Variable Mode Imager and ImageQuant software.

**Electron microscopy**

Electron microscopy analysis of DnaA–DNA interactions was performed by the method described previously [18]. Briefly, the *M. tuberculosis* oriC region was PCR-amplified using pDNAABamHI primer and pGEM-T EasyNotI primers (Table 1). The PCR DNA fragment was purified by filtration through Amicon Microcon-PCR Centrifugal...
Filter Devices (Millipore) and then DNA was incubated with streptavidin (Promega) for 5 min at room temperature. Excess of streptavidin was removed by filtration as above. *M. tuberculosis* DnaA protein was incubated at room temperature with the biotin/streptavidin DNA fragment in binding buffer (as indicated above). After fixation in 0.2% glutaraldehyde (15 min, room temperature), samples were prepared for electron microscopy by adsorption on mica [19]. They were analysed with a transmission electron microscope (Philips CM100) at 60 kV. The lengths of DNA fragments and the position of complexes were measured by using an electronic digitizer and evaluated with a computer program [14].

**RESULTS**

**Expression and purification of a soluble *M. tuberculosis* DnaA protein**

The vector pET-28a(+) was chosen to overexpress the *M. tuberculosis* DnaA as a His-tagged protein (see the Materials and methods section for details). To isolate the recombinant His6MtDnaA protein in its native form, the best conditions for expression of the soluble protein in *E. coli* were established. *E. coli* BL21 cells harbouring the pET-28a(+)His6MtDnaA vector. Proteins were analysed by SDS/PAGE. Tracks: M, molecular-mass standard (protein molecular-mass standard; MB, Fermentas); 1, whole extract from uninduced cells; 2, whole extract from the cells after 3 h of induced expression; 3, soluble cell lysate from induced cells; 4, non-soluble cell debris from induced cells; 5, protein pellet after (NH4)2SO4 precipitation; 6, soluble protein fraction after (NH4)2SO4 precipitation; 7, (NH4)2SO4 protein pellet dissolved in LG buffer; 8, protein fraction desalted on PD-10 column; 9–17 fractions eluted from Ni-NTA–agarose column with LG containing an increased amount of imidazole: 9, 0 mM imidazole; 10, 1 mM imidazole; 11–12, 10 mM imidazole; 13–16, 50 mM imidazole; 17, 100 mM imidazole.

The 5′-end-radiolabelled DNA fragments (approx. 10 fmol) were incubated with different amounts of DnaA protein in binding buffer (see above) at room temperature for 30 min. Then DNase I digestion was performed as described by Majka et al. [17]. The DNase I cleavage products were separated in an 8% polyacrylamide/urea sequencing gel. Gels were dried and analysed by a Typhoon 8600 Variable Mode Imager and ImageQuant software.

**Figure 1** Expression and purification of the recombinant *M. tuberculosis* DnaA

DnaA protein was isolated from *E. coli* BL21 cells harbouring the pET-28a(+)His6MtDnaA vector. Proteins were analysed by SDS/PAGE. Tracks: M, molecular-mass standard (protein molecular-mass standard; MB, Fermentas); 1, whole extract from uninduced cells; 2, whole extract from the cells after 3 h of induced expression; 3, soluble cell lysate from induced cells; 4, non-soluble cell debris from induced cells; 5, protein pellet after (NH4)2SO4 precipitation; 6, soluble protein fraction after (NH4)2SO4 precipitation; 7, (NH4)2SO4 protein pellet dissolved in LG buffer; 8, protein fraction desalted on PD-10 column; 9–17 fractions eluted from Ni-NTA–agarose column with LG containing an increased amount of imidazole: 9, 0 mM imidazole; 10, 1 mM imidazole; 11–12, 10 mM imidazole; 13–16, 50 mM imidazole; 17, 100 mM imidazole.

Thus taking into account the results of gel retardation assays, it could be concluded that the *M. tuberculosis* DnaA protein requires at least two DnaA boxes for efficient binding.
Binding requirements of the His₆MtDnaA protein

The gel retardation assays were performed using ³²P-labelled DNA fragments and increasing amounts of the His₆MtDnaA. (A) The double-stranded oligonucleotide containing the single ‘perfect’ DnaA box (TTGTCCACA, 21 bp). Tracks: 1, 0; 2, 2.5; 3, 5; 4, 10; 5, 25; 6, 50; 7, 100; 8, 250 nM His₆MtDnaA. (B) The double-stranded oligonucleotide containing two DnaA boxes (TTGTCCACA and TTGTCCCCA, 33 bp). Tracks: 1, 0; 2, 1; 3, 5; 4, 10; 5, 25; 6, 50; 7, 100 nM His₆MtDnaA. (C) Non-DnaA box DNA fragment (348 bp), tracks: 1, 0; 2, 2.5; 3, 5; 4, 10; 5, 25; 6, 50; 7, 100 nM His₆MtDnaA. (D) The oriC fragment (541 bp). Tracks: 1, 0; 2, 1; 3, 2.5; 4, 5; 5, 10; 6, 25; 7, 50; 8, 100 nM His₆MtDnaA.

Identification of DnaA-binding sites within the *M. tuberculosis* oriC region

A total of 13 putative DnaA box motifs, each with at least one mismatch from the consensus sequence, were found within the oriC region of *M. tuberculosis* [10] (Figure 3). Such a relatively high abundance of ‘non-perfect’ DnaA boxes raised the question of how many of them are specifically bound by DnaA protein. To answer this question, the interactions between the His₆MtDnaA protein and multiple putative DnaA boxes of the oriC region were analysed by electron microscopy and DNase I footprinting.

For electron microscopy analysis, the *M. tuberculosis* oriC region was PCR-amplified using a pair of primers pMtoriCfw and biotinylated pGEM-T Easybiot. The PCR product was labelled with a streptavidin that served as reference point for measurements of the position of the nucleoproteins complex. At a ratio of 0.5:1 DNAA protein/DNAA box, the nucleoprotein complexes visualized by electron microscopy were small and their numbers did not exceed one or two per DNA molecule. The relative positions of the analysed nucleoprotein complexes corresponded to the position of most of the putative DnaA boxes along the oriC region (Figure 4). Within the analysed oriC region, the DnaA protein forms four separated complexes. The highest incidence of protein binding occurred at the two adjacent DnaA boxes (9th and 10th) and at the 3rd DnaA box, the only DnaA box with a single mismatch (Figures 3 and 4). Interestingly, the single 11th DnaA box separated from other boxes is also bound by the His₆MtDnaA protein. The lowest incidence of protein binding took place at the more distal DnaA boxes (12th and 13th). According to the electron microscopy analysis, 1st and 2nd DnaA boxes were not bound at all. However, if putative-binding motifs are located close to the ends of the analysed DNA fragment, protein binding is frequently not observed (protein slips off the ends).
DNA-binding properties of the Mycobacterium tuberculosis DnaA protein

DISCUSSION

Studies on bacterial initiation of chromosome replication were mostly focused on a model organism, E. coli. Little is known about the corresponding process in other bacteria, particularly in slow growing organisms such as mycobacteria. Knowledge about the interactions of M. tuberculosis DnaA protein with DNA may provide fresh insights into the function of this protein, and may further help in understanding the regulation of initiation of mycobacterial chromosome replication. In the present study, we describe the purification of the M. tuberculosis oriC region.

Recently, it has been reported that two mycobacterial DnaA proteins (M. avium and M. tuberculosis) [21,22] were isolated. However, both proteins were purified from inclusion bodies as His-tagged proteins on Ni\(^2+\)-affinity column under denaturing conditions. To avoid these problems, we designed the procedure for isolation of the M. tuberculosis DnaA protein in its native form from the E. coli BL21 cells harbouring pET 28a(+)His\(_6\)MtDnaA expression vector. Reduction of temperature and concentration of the inducer (IPTG) during expression of the recombinant His\(_6\)MtDnaA protein allowed us to obtain the protein in sufficient quantity for purification on affinity chromatography. The purified His\(_6\)MtDnaA protein was used for investigating its binding properties.

The oriC region of M. tuberculosis, unlike its E. coli counterpart, possesses 13 putative DnaA boxes, which differ significantly from the ‘perfect’ E. coli sequences (TTATCCACA). The consensus sequence for Mycobacterium DnaA box is TTG/CTCCACA [8]. Within the M. tuberculosis oriC, there is only one DnaA box with a single mismatch from the consensus sequence (DnaA box no. 3, Figure 3). The remaining DnaA boxes differ by two or even three bases from the Mycobacterium consensus sequence. The M. tuberculosis oriC region contains the same number of DnaA boxes as the oriC region from Thermus thermophilus [23]. However, within the latter, all but two of the 13 DnaA sequences match the E. coli ‘perfect’ DnaA box.

As demonstrated by gel retardation analysis, the His\(_6\)MtDnaA does not bind a single DnaA box, not even the Mycobacterium consensus sequence DnaA box (TTATCCACA; Figure 2A). Recently, in contrast with our results, Dziadek et al. [24] showed by surface plasmon resonance technique that the M. tuberculosis DnaA protein binds a single DnaA box with three mismatches from the consensus sequence (DnaA box no. 4, CCGTTCACA). However, our DNase I footprinting as well as electron microscopy analyses demonstrated that this DnaA box, being in the vicinity of other DnaA boxes, is not efficiently protected by DnaA protein (see Figures 4 and 5). In addition, DnaA box 4 is not essential for oriC function; mutations in this DnaA box did not abolish chromosome replication [24].

Our gel retardation assay demonstrated that if another DnaA box accompanies the ‘perfect’ one, both of them are bound with a high affinity by the M. tuberculosis DnaA (Figure 2B). Thus, in contrast with E. coli, all other DnaA proteins from Helicobacter pylori, Streptomyces lividans as well as M. tuberculosis require two DnaA boxes for efficient binding if these contain mismatches. The DnaA protein of T. thermophilus requires even more DnaA boxes for efficient binding (at least three ‘perfect’) [23]. Our results suggest that interaction of the M. tuberculosis DnaA with two DnaA boxes exhibits co-operativity.

Electron microscopy as well as DNase I footprinting demonstrated that the purified His\(_6\)MtDnaA protein binds to four specific regions, which correspond to the location of 11 out of the 13 previously identified DnaA boxes within the M. tuberculosis oriC. Two DnaA boxes located at the 5th end of the oriC region are not bound by the His\(_6\)MtDnaA protein. Salazar et al. [9] have identified only seven putative DnaA boxes within the M. tuberculosis oriC region. Our electron microscopy studies demonstrated that the M. tuberculosis DnaA protein binds preferentially the...
fragment of the oriC region that contains these seven DnaA boxes (Figure 4, DnaA boxes 3–10).

Binding of the M. tuberculosis DnaA protein to the oriC region led to the formation of high-molecular-mass nucleoprotein complexes; a discrete nucleoprotein DnaA box was not observed (see the Gel retardation assay section; Figure 2D). In contrast, DnaA proteins from E. coli [14] and H. pylori [15] bind sequentially five DnaA boxes in their oriC regions; at least four distinct nucleoprotein complexes are formed. In S. lividans [25] and T. thermophilus [23] organisms, which also possess numerous DnaA boxes within their oriC regions (13 and 19 respectively), only high-molecular-mass complexes were observed. Probably, in M. tuberculosis, T. thermophilus and S. lividans, DnaA molecules by co-operative binding of numerous DnaA boxes assemble along the oriC region and subsequently form the initial complex. Formation of such a massive nucleoprotein complex suggests that the DnaA is involved not only in DNA–protein but also in protein–protein interactions. In S. lividans, two T, I, and III of DnaA protein are independently involved in protein–protein interactions [25].

Thus our results suggest that in M. tuberculosis, DnaA protein binds co-operatively to multiple ‘non-perfect’ DnaA boxes in oriC. Probably, the presence of a non-perfect DnaA box within the mycobacterial oriC region is compensated by their abundance.

This work was supported by the Ministry of Scientific Research and Information Research (The State Committee for Scientific Research, grant 3 PO4A 079 22). Cosmid K41 was kindly provided by R. Brosch (Institut Pasteur, Paris, France). We thank D. Kamińska (Institute of Pharmacology, Krakow, Poland) and B. Lis (Institute of Immunology and Experimental Therapy, Wroclaw, Poland) for the help in M. tuberculosis oriC cloning and isolation of the M. tuberculosis DnaA protein respectively. We also thank W. Messer for his comments on the paper.

REFERENCES


