The replicative polymerases PolC and DnaE are required for theta replication of the *Bacillus subtilis* plasmid pBS72

Marina Titok,1 Catherine Suski,2 Bérengère Dalmais,3 S. Dusko Ehrlich3 and Laurent Jannière3

1Belarussian State University, Biological Faculty, Department of Genetics and Biotechnology, Minsk 220050, 4 Scorina Avenue, Belarus
2Memorial Sloan-Kettering Cancer Center, NY 10021, USA
3Laboratoire de Génétique Microbienne, Bâtiment des Biotechnologies, INRA, 78352 Jouy en Josas, France

Plasmids are the tools of choice for studying bacterial functions involved in DNA maintenance. Here a genetic study on the replication of a novel, low-copy-number, *Bacillus subtilis* plasmid, pBS72, is reported. The results show that two plasmid elements, the initiator protein RepA and an iteron-containing origin, and at least nine host-encoded replication proteins, the primosomal proteins DnaB, DnaC, DnaD, DnaG and Dnal, the DNA polymerases DnaE and PolC, and the polymerase cofactors DnaN and DnaX, are required for pBS72 replication. On the contrary, the cellular initiators DnaA and PriA, the helicase PcrA and DNA polymerase I are dispensable. From this, it is inferred that pBS72 replication is of the theta type and is initiated by an original mechanism. Indirect evidence suggests that during this process the DnaC helicase might be delivered to the plasmid origin by the weakly active DnaD pathway stimulated by a predicted interaction between DnaC and a domain of RepA homologous to the major DnaC-binding domain of the cellular initiator DnaA. The plasmid pBS72 replication fork appears to require the same functions as the bacterial chromosome and the unrelated plasmid pAM1. Most importantly, this replication machinery contains the two type C polymerases, PolC and DnaE. As the mechanism of initiation of the three genomes is substantially different, this suggests that both type C polymerases might be required in any Cairns replication in *B. subtilis* and presumably in other bacteria encoding PolC and DnaE.

**INTRODUCTION**

DNA replication is a ubiquitous biological process carried out by proteins that have evolved profoundly since the last common ancestor. For key elements like initiators, helicases, primases, catalytic DNA polymerases and processivity factors, their evolution has been so extensive that they are postulated to originate from peptides that diverged beyond recognition or from different proteins (Edgell & Doolittle, 1997; Giraldo, 2003; Leipe et al., 1999). Surprisingly, this diversity, obvious when comparing Eubacteria to Archaea and Eukarya, is also manifest among bacterial species. For instance, in the well characterized eubacteria *Escherichia coli* and *Bacillus subtilis*, the primosomes required for initiating replication at origins or collapsed replication forks have been shown to greatly differ in their constitution and mechanism of helicase loading (Bruand et al., 1995; Cox et al., 2000; Davey & O’Donnell, 2003; Lemon et al., 2002; Marsin et al., 2001; Polard et al., 2002; Velten et al., 2003). Differences have also been documented in the DNA polymerase III (Pol III) holoenzyme, a subassembly of the replication machinery where DNA synthesis is catalysed. In *E. coli* this element contains 10 different subunits of which four (θ, γ, χ and ψ) are missing in *B. subtilis* (Bruck & O’Donnell, 2000; Bruck et al., 2002; Lemon et al., 2002; Martinez-Jimenez et al., 2002; McHenry, 2003).

Based on systems for which a Pol III holoenzyme has been isolated and/or reconstructed *in vitro*, it is generally thought that both DNA strands of genomes are polymerized by a single replicative polymerase of the C (prokaryotes) or B (eukaryotes) families, dimerized in a holoenzyme structure (for reviews, see Benkovic et al., 2001; Bullock, 1997). However, this scheme might not be universal as replication in numerous organisms depends on two different polymerases. For instance, in *B. subtilis* and related bacteria of the *Bacillus/Clostridium* and *Thermotogales* groups, two type C polymerases are required.
polymersases (PolC and DnaE) are essential for chromosome replication and colocalize with the replication machinery (Dervyn et al., 2001; Inoue et al., 2001; Lemon & Grossman, 1998, 2000). In the replication fork of the B. subtilis plasmid pAMβ1, PolC might be mainly involved in leading strand synthesis while DnaE might be required for efficient lagging strand polymerization (Dervyn et al., 2001). This polymerase specialization detected in a plasmid system might also apply to chromosomal forks as (i) initiation of pAMβ1 replication depends on the cellular machinery that initiates DNA synthesis at collapsed replication forks (Bruand et al., 1995, 2001; Marsin et al., 2001; Polard et al., 2002) and (ii) a bioinformatic study showed that bacteria containing both PolC and DnaE have a much stronger gene strand bias than DnaE+ and PolC− species (Roche, 2002). This latter point suggests that asymmetrical replication forks could have a major role in organizing bacterial chromosomes. DNA replication also depends on two different replicative polymerases (Polδ and Polε, B family) in numerous eukaryotic cells (for reviews, see Burgers, 1998; Hubscher et al., 2000; Johnson & O’Donnell, 2005; Kawasaki & Sugino, 2001; Sugino, 1995). As in bacterial forks, these polymerases might be specialized as several data suggest that Polδ would elongate leading strand synthesis while Polδ would synthesize the lagging strand (see, for instance, Fukui et al., 2004; Garg et al., 2004; Karthikeyan et al., 2000; Morrison & Sugino, 1994; Shcherbakova & Pavlov, 1996). Recently, it was proposed that chromosomal replication in archaea might also depend on two distinct DNA polymerases (Henneke et al., 2005).

To further test the requirement for PolC and DnaE in Gram-positive bacteria, we undertook a genetic study of the replication of B. subtilis plasmid pBS72 (Titok et al., 2003). Results showed that pBS72 replication (i) is of the theta-type, (ii) depends on two plasmid elements, an initiator protein and an origin of replication, (iii) initiates by an original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork original mechanism independent of the cellular initiators type, (ii) depends on two plasmid elements, an initiator protein and an origin of replication, (iii) initiates by an original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork. Results showed that pBS72 replication (i) is of the theta-type, (ii) depends on two plasmid elements, an initiator protein and an origin of replication, (iii) initiates by an original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork. Results showed that pBS72 replication (i) is of the theta-type, (ii) depends on two plasmid elements, an initiator protein and an origin of replication, (iii) initiates by an original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork. Results showed that pBS72 replication (i) is of the theta-type, (ii) depends on two plasmid elements, an initiator protein and an origin of replication, (iii) initiates by an original mechanism independent of the cellular initiators DnaA and PriA and (iv) is carried out by a replication fork.

DNA manipulation and plasmid construction. DNA preparation and manipulation were carried out essentially according to standard procedures (Bron, 1990; Sambrook et al., 1989). Previously described plasmids used were pMTL20C, pMTL21E (Swinfield et al., 1991) and pMTLBS72 (a pMTL20C derivative containing a BgII fragment carrying the PBS72 replica; Fig. 1, top line) (Titok et al., 2003). Derivatives of pMTLBS72 carrying a truncated or mutated PBS72 replica (Fig. 1) were generated in E. coli by inactivating the parental plasmid with restriction enzymes cleaving in the cloned fragment and, when appropriate, in the flanking polylinker (segments were blunt-ended with the Klenow enzyme when required). The plasmids yielded were termed pLJH16-18 and pLJH67-76. The DnaA requirement was investigated with the pBS72 derivative pLJH13 composed of the PBS72 BgII fragment of pMTLBS72 cloned at the BgII site of the erythromycin-resistant E. coli vector pMTL21E.

Copy number (CN) measurement. To determine CN variation in replication mutants, total DNA was extracted, restricted, run in an agarose gel and analysed by Southern analysis using as probes radio-labelled plasmid and chromosomal sequences. Molecular Dynamic tools were used to quantify CN variations as described previously (Titok et al., 2003). CN values presented in Table 1 were calculated from three independent measurements. Control experiments showed that the integrated plasmid replicon (pLS2) that directs chromosomal replication in the ßdnaA chromosome (Hassan et al., 1997) does not affect PBS72 derivative CN values (not shown). CN measurement in the ßPriA strain was carried out in LB with fresh transformants isolated in a poor medium. Since the strain is genetically unstable (Bruand et al., 2001; Polard et al., 2002), the ßPriA phenotype was confirmed genetically at the same time as DNA preparation (not shown).

Plasmid replication functions

We previously cloned and sequenced a 3·1 kb replication determinant widely represented in large (≥90 kb) plasmids of a Belarus collection of B. subtilis strains (Titok et al., 2003). The segment, fused to a pUC-like plasmid to form pMTLBS72, contains three complete ORFs (Fig. 1, top). The product of ORF-1 (hereafter termed RepA) might correspond to the plasmid initiator as it contains a putative
DNA-binding motif (HTH motif between amino acids 111 and 132) and a C-terminal region (position 276–330) homologous (35% identity, 63% similarity) to the central part of the N-terminal domain I of the highly conserved chromosomal initiator DnaA. In *E. coli*, this internal region of domain I (position 13–68) is mainly involved in stable binding to the replicative helicase (Felczak et al., 2005; Seitz et al., 2000; Sutton et al., 1998). The ∼400 bp downstream sequence would include the plasmid origin as it contains typical hallmarks of such replication elements: repeated sequences (iterons), DnaA-binding sites and an AT-rich region (for reviews, see del Solar et al., 1998; Espinosa et al., 2000; Helinski et al., 1996). The function of the remaining two ORFs, which have no homologues in databases, is not known.

To genetically identify plasmid sequences required for replication, pMTLBS72 derivatives carrying parts of the 3 kb replicon were constructed and assessed for replication activity in *B. subtilis* by a transformation assay. Results (Fig. 1) showed that the minimal segment capable of autonomous replication is 1929 bp in length (pLJH76). It encompasses repA and the downstream ∼400 bp region. As expected, the repA product is essential for replication as a 5′ deletion of the coding frame and a truncation of the protein by frameshift mutations at the NsiI or SalI sites abolish transformation activity (pLJH16-18 and pLJH67-68). The ∼400 bp downstream sequence is also essential for replication since (i) its deletion (pLJH71 and pLJH75) causes a complete loss of plasmid pMTLBS72 and (ii) its alteration by a small (32 bp) deletion changing the spacing between two repeats (pLJH74) sharply reduces the transformation activity and the CN value of the replicon. These results suggest that RepA is the plasmid initiator and that the ∼400 bp includes the plasmid origin. Note that none of the functional derivatives (except pLJH74) displayed a marked change in CN (not shown). This indicates that ORF-2 and ORF-3 are not needed for replication control and that the CN of the replicon might be regulated at the level of repA expression and/or by origin handcuffing (del Solar et al., 1998; Espinosa et al., 2000; Helinski et al., 1996; Park et al., 2001).

Fig. 1. Genetic localization of the pBS72 replication determinants. (a) Mutation analysis. The 3081 bp replicon (BgII fragment) contains three complete ORFs and one truncated ORF (horizontal arrows in black and grey, respectively), conventional expression signals (promoters, bent arrows; transcription terminators, lollipops) and a putative origin of replication (ori). ORF1 (repA), ORF-2, ORF-3 and ORF-4 (C-terminal part of a truncated protein) encode proteins of 344, 168, 113 and 87 aa, respectively. Relevant restriction sites are indicated. Deleted or mutated (*) pBS72 replicons were isolated in *E. coli* and tested for transformation activity in *B. subtilis*. Constructs displaying a high transformation activity and CN similar to that of the original plasmid pMTLBS72 are indicated by ‘+’. The construct having a low transformation efficiency and a reduced CN is pinpointed by ‘±’. Derivatives that gave no transformants are indicated by ‘–’. The shortest segment required for optimal replication (CotI–EcoRI) is represented by a thicker line and is carried by plasmid pLJH76. (b) Salient features of the origin region. Arrows R1–R5 indicate putative RepA-binding sites. Black and grey bars indicate AT- and GC-rich sequences, respectively. The DnaA-binding site is boxed. Brackets delimit the sequence deleted in pLJH74.
Host-encoded replication functions

To search for host-encoded proteins required for plasmid replication, pMTLBS72 was introduced into wild-type and isogenic mutants affected in DNA polymerases PolC (dnaF69), DnaE (dnaETs2.6 and 2.10) or Pol I (polA5), the processivity factor DnaN (dnaG5), the replication machinery organizer DnaX (dnaX8132), the initiator proteins DnaA (ΔdnaA) and PriA (ΔpriA), the replicative helicase DnaC (dnaC14), the DNA loaders DnaB, DnaD and DnaI (dnaB19, dnaD23 and dnaI2), the primase DnaG (dnaE20) and the PcrA helicase required for rolling-circle replication and involved in Cairns replication (pcrA3) (for reviews, see Khan, 2005; Lemon et al., 2002; Moriya et al., 1999; Winston, 1982; Yoshikawa & Wake, 1993). Transformants were readily obtained at 30 °C with the complete set of strains. To measure the effect of T mutations on plasmid replication, total DNA was extracted before and after a short growth (about five generations) at the semi-restrictive temperature. The DNA was then restricted, run in agarose gels and analysed by Southern blotting and hybridization, using radiolabelled plasmid DNA and a chromosomal PCR fragment as probes. The ratios of plasmid/chromosomal signals were then determined before and after growth at the semi-permissive temperature and compared. In the case of null or constitutive mutations (ΔdnaA, ΔpriA, polA5 and pcrA3), plasmid/chromosomal ratios were measured at 37 °C and compared to those of the isogenic wild-type strains grown at the same temperature. Replica-plating analysis showed that the plasmid was maintained in the cell population throughout the experiment in most of the mutants. However, for the dnaB19 and, to a lesser extent, dnaI2 mutants, cured cells were detected at the end of growth in proportions reaching 0.5 and 0.3, respectively. The proportion of plasmid-free cells was taken into account for CN determination.

In most mutants, a moderate (two- to threefold) CN reduction was observed at semi-restrictive temperature (Fig. 2 and Table 1). Plasmid replication is thus more affected than chromosome replication in these strains, showing that the corresponding mutated factors are required for pBS72 replication. This applies to the polymerases PolC and DnaE, the processivity factor DnaN, the replication machinery organism.

![Fig. 2. Plasmid CN measurement in various replication mutants. Total DNA of wild-type (WT) and mutated strains grown in permissive or semi-permissive conditions was analysed by Southern analysis after PstI restriction (except in the case of dnaG5 and dnaX8132 where SspI was used) and hybridization with radiolabelled probes homologous to the plasmid (P) and a chromosomal (C) segments (see text for more details). Plasmid pMTLBS72 was used in all the strains except ΔdnaA and the wild-type isogenic cell where a related plasmid (pLJH13) was used. Results obtained within the 168 (a) and HVS325 (b) wild-type strains are shown as references for 168- and HVS325-derived strains.](image-url)
organizer DnaX and the primosomal proteins DnaB, DnaC, DnaD, DnaG and DnaI. In strains lacking DnaA or PriA (because of gene deletion) or encoding a replication-defective form of the polymerase Pol I (polA5 mutation; Laipis & Ganesan, 1972) or an inactive form of the PcrA helicase (pcrA3; Anand et al., 2005), the CN is either unchanged (Pol I and PcrA mutants) or increased (DnaA and PriA mutants). These factors are hence not required for pBS72 replication. As the DnaA and PriA null strains filament (Hassan et al., 1997; Polard et al., 2002; our observations), the CN increase observed in these contexts might result from dilution of negative factors in enlarged cells allowing plasmid amplification. However, in the DnaA context, it cannot be ruled out that the CN increase results from alleviation of a negative effect of DnaA on plasmid replication, because DnaA might bind to dnaA boxes of the plasmid origin or to host-encoded factors limiting for plasmid replication. Finally, it should be noted that (partial) inactivation of required replication factors did not always cause CN reduction. In several instances, the (partial) inactivation kept the CN unchanged. This occurs, for instance, when PolC inactivation is caused by the dnaF33 mutation or when cells are grown in the presence of the PolC-inhibitory compound HB-EMAU (not shown). A similar result was obtained in cells depleted of the wild-type PolC or DnaE enzymes because of the weak expression of the corresponding gene expressed from the IPTG-inducible spac-I promoter (not shown). This probably results from a simultaneous arrest of plasmid and chromosome replication. Finally, the lack of single-stranded chromosome replication. The second aspect of replication uncovered by this study concerns the mechanism of initiation of pBS72 replication. The goal of this replication step is to load the replicative helicase on the single-stranded DNA generated close to the initiator/DNA complex. In bacteria, two main initiation pathways have been described so far. The PriA-dependent pathway is the best characterized in B. subtilis. It is required to restart DNA synthesis at disassembled replication forks. Dissection of the pathway by genetic and biochemical means revealed the following scheme. First, PriA binds DNA forks carrying single-stranded DNA coated with single-stranded DNA-binding protein and recruits DnaD. The complex then assists the binding of DnaB which recruits DnaI and DnaC. This cascade of reactions yields helicase DnaC loading on the appropriate single-stranded DNA by a ring-assembly mechanism (Bruand et al., 2001, 2005; Marsin et al., 2001; Polard et al., 2002; Velten et al., 2003). The second initiation pathway, the DnaA-dependent pathway, operates at the chromosomal origin of replication. It requires a DnaA/origin interaction as an initial step followed by a cascade of reactions involving DnaB, DnaC, DnaD and DnaI (for reviews, see Lemon et al., 2002; Yoshikawa & Wake, 1993). It is speculated that this cascade is similar to that occurring in the PriA-dependent pathway (Bruand et al., 2005; Velten et al., 2003). It is interesting to note here that DnaD interacts with DnaA and PriA (Bruand et al., 2005; Ishigo-Oka et al., 2001; Marsin et al., 2001). This interaction might thus play a key role in bringing DnaB, DnaI and DnaC to the initiator/DNA complexes.

### Table 1. Plasmid CN measurement in various replication mutants: quantitative analysis

<table>
<thead>
<tr>
<th>Tested protein</th>
<th>Mutation</th>
<th>Relative CN*</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
<td>1 ± 0.18</td>
</tr>
<tr>
<td>PolC</td>
<td>dnaF69</td>
<td>0.57 ± 0.05</td>
</tr>
<tr>
<td>DnaE</td>
<td>dnaE2.6</td>
<td>0.33 ± 0.12</td>
</tr>
<tr>
<td>DnaE</td>
<td>dnaE2.10</td>
<td>0.27 ± 0.02</td>
</tr>
<tr>
<td>DnaX</td>
<td>dnaX8132</td>
<td>0.27 ± 0.05</td>
</tr>
<tr>
<td>DnaN</td>
<td>dnaG5</td>
<td>0.60 ± 0.08</td>
</tr>
<tr>
<td>DnaB</td>
<td>dnaB19</td>
<td>0.29 ± 0.13</td>
</tr>
<tr>
<td>DnaC</td>
<td>dnaC14</td>
<td>0.31 ± 0.08</td>
</tr>
<tr>
<td>DnaD</td>
<td>dnaD23</td>
<td>0.30 ± 0.1</td>
</tr>
<tr>
<td>DnaG</td>
<td>dnaE20</td>
<td>0.25 ± 0.05</td>
</tr>
<tr>
<td>DnaI</td>
<td>dnaI2</td>
<td>0.33 ± 0.17</td>
</tr>
<tr>
<td>DnaA</td>
<td>Null</td>
<td>2.41 ± 0.64</td>
</tr>
<tr>
<td>PriA</td>
<td>Null</td>
<td>6.43 ± 2.89</td>
</tr>
<tr>
<td>Pol I</td>
<td>polA5</td>
<td>1.11 ± 0.15</td>
</tr>
<tr>
<td>PcrA</td>
<td>pcrA3</td>
<td>1.08 ± 0.28</td>
</tr>
</tbody>
</table>

*Relative mean CN values from three independent experiments ± SEM are shown. Values obtained in the strain 168 are given. In the HVS325 WT strain, the relative CN was 0 ± 0.09. The CN in the WT context is 5 ± 1 (Titok et al., 2003)."
The factors required for pBS72 initiation of replication are two plasmid elements, the origin and the RepA protein, and four host-encoded replication factors, namely DnaD, DnaB, DnaN, and DnaC. It is, however, surprisingly independent of the initiator proteins DnaA and PriA (so far, only oriF from R6K, R1, pLS32, pAMβ1 and RSF1010 have been shown to be DnaA-independent; Bernander et al., 1991; Espinosa et al., 2000; Hassan et al., 1997; Kelley & Bastia, 1992; Scherzinger et al., 1991; Tang et al., 1989). This suggests that the recruitment of the DnaD-driven cascade to the RepA(origin complex might differ from that occurring at the DnaA/oriC and PriA/DNA complexes. A simple model would be that the single-stranded DNA generated by RepA at the plasmid origin is used as a landing pad for DnaD which has a single-stranded DNA-binding activity (Bruand et al., 2005; Marsin et al., 2001). In support of this hypothesis, it was previously shown that the DnaC helicase can be loaded on long stretches of single-stranded DNA via a weakly efficient DnaA- and PriA-independent pathway that requires DnaD, DnaB and Dnal (the DnaD pathway) (Bruand et al., 2005; Marsin et al., 2001). However, similar to all initiator/DNA complexes studied in detail so far (for reviews, see del Solar et al., 1998; Helsinki et al., 1996), RepA must not generate a single-stranded region large enough to activate the DnaD pathway at the origin. Additional factors should come into play to compensate for this. Interestingly, several studies show that plasmid initiators interact with host-encoded replication proteins for initiating DNA synthesis [RepA of pSC101 (Datta et al., 1999; Sharma et al., 2001), TrfA of RK2 (Pacek et al., 2001) and π of R6K (Abhyankar et al., 2004; Lu et al., 1998; Ratanakar et al., 1996); for a review, see Konieczny, 2003]. Thus, it is conceivable that RepA interacts with replication factors to stimulate the DnaD pathway. One possible interacting protein would be DnaD itself as this enzyme is already known to interact with the initiators DnaA and PriA (Bruand et al., 2005; Ishigooka et al., 2001; Marsin et al., 2001). However, RepA carries a region homologous to domain 1 of DnaA-type proteins which, in the E. coli enzyme, strongly binds the replicative helicase and is required for stabilizing it in the prepriming complex at oriC (Felczak et al., 2005; Seitz et al., 2000; Sutton et al., 1998). Hence, instead of, or in addition to, interacting with DnaD, RepA may bind to DnaC, acting in place of DnaA to efficiently recruit and stabilize the helicase in the plasmid prepriming complex, increasing the efficiency of the DnaD pathway.

The final aspect of pBS72 replication uncovered by this analysis is that once the prepriming complex has formed, the DnaG primase is recruited for synthesizing RNA primers, opening a route for the assembly of the replication machinery. Actually, this machinery is indistinguishable from that acting on the chromosome and the plasmid pAMβ1 (Bruand et al., 1995; Ceglowski et al., 1993; Dervyn et al., 2001) for reviews, see Lemon et al., 2002; Winston, 1982; Yoshikawa & Wake, 1993]). It includes the two DNA polymerases PolC and DnaE, the replication machinery organizer DnaX and the processivity factor DnaN. According to studies carried out in Streptococcus pyogenes (Bruck & O’Donnell, 2000), it should also contain the δ and δ’ proteins. The lack of mutations in these genes has precluded their study in this work. Hence, three different genomes (pBS72, pAMβ1 and the B. subtilis chromosome) are replicated by related, and possibly identical, machinery requiring two different type C polymerases, PolC and DnaE. The substantial differences in the mechanism of DnaC loading in the three genomes (the DnaA pathway at the chromosomal origin, the PriA pathway in pAMβ1 and putatively the DnaD pathway in pBS72) suggest that this machinery operates in any Cairns replication in B. subtilis and presumably also in related bacteria of the Bacillus/Clostridium and Thermotogales groups.

ACKNOWLEDGEMENTS

We thank Shigeki Moriya for providing AdnaA and isogenic strains, Alexandre Galizzi for strain PB1669, Patrice Polard for the ΔpriA strain, and N. C. Brown and G. E. Wright for making available HB-EMAU. Nathalie Galleron is thanked for her excellent technical assistance in DNA sequencing and Emmanuelle Le Chatelier, Danièle Cancill and David Lane for stimulating discussions and critical reading of the manuscript. The work was supported, in part, by a research fund from INTAS (INTAS97-1464).

REFERENCES


