The Sinorhizobium meliloti chromosomal origin of replication

Christopher D. Sibley, Shawn R. MacLellan and Turlough Finan

The predicted chromosomal origin of replication (oriC) from the alfalfa symbiont Sinorhizobium meliloti is shown to allow autonomous replication of a normally non-replicating plasmid within S. meliloti cells. This is the first chromosomal replication origin to be experimentally localized in the Rhizobiaceae and its location, adjacent to hemE, is the same as for oriC in Caulobacter crescentus, the only experimentally characterized alphaproteobacterial oriC. Using an electrophoretic mobility shift assay and purified S. meliloti DnaA replication initiation protein, binding sites for DnaA were mapped in the S. meliloti oriC region. Mutations in these sites eliminated autonomous replication. S. meliloti that expressed DnaA from a plasmid lac promoter was observed to form pleomorphic filamentous cells, suggesting that cell division was perturbed. Interestingly, this cell phenotype is reminiscent of differentiated bacteroids found inside plant cells in alfalfa root nodules.

INTRODUCTION

Bidirectional replication of the bacterial chromosome requires the specific interaction of the ATPase DnaA at a precise region on the replicon, termed the chromosomal origin of replication (oriC) (Fuller et al., 1981, 1984). oriC is the location of the initial strand opening event that precedes replication and the site of replisome assembly (Bramhill & Kornberg, 1988b). Many of the molecular events that occur at the Escherichia coli oriC have been well characterized, providing a general model of the essential biochemical process which results in the formation of two replication forks and duplication of oriC (Bramhill & Kornberg, 1988a). While E. coli oriC is the most extensively characterized chromosome origin, sequences capable of autonomous replication have been isolated from many bacterial chromosomes (Harding et al., 1982; Zyskind et al., 1983; Yee & Smith, 1990; Zakrzewska-Czerwinska & Schrempf, 1992; Fujita et al., 1990, 1992; Calcutt & Schmidt, 1992; Schaper et al., 2000; Moriya et al., 1992; Salazar et al., 1996; Jakimowicz et al., 1998). In most sequenced bacterial chromosomes, oriC has been predicted on the basis of sequence analysis where a switch in the asymmetry of the two strands of DNA coincides with the origin and terminus of the chromosome (Lobry, 1996). The architecture of the bacterial chromosome origin is nevertheless diverse with respect to the physical sequence as well as the number and organization of binding sites for proteins that control the timing of origin firing. Binding sites for the replication initiator DnaA (DnaA boxes) appear to be one of the hallmark features of bacterial chromosome origins and because of this the location of DnaA boxes in the bacterial genome has been used as a method for predicting oriCs (Mackiewicz et al., 2004). Another feature of many functional origins is that adjacent to the DnaA boxes is an exceptionally AT-rich region which is the site of initial strand separation. Bacterial chromosome origins are commonly located in close vicinity to the dnaA gene and attempts have been made to exploit this linkage in the cloning of oriCs (Smith et al., 1991; Fujita et al., 1990, 1992; Schaper et al., 2000; Moriya et al., 1992; Salazar et al., 1996; Jakimowicz et al., 1998). Examples do exist in which hallmarks of oriCs have not been found to be linked to the dnaA replication initiation gene, such as the chromosomes of Synechocystis spp. (Richter & Messer 1995), Prochlorococcus marinus (Richter et al., 1998) and Sinorhizobium meliloti (Margolin et al., 1995).

S. meliloti is a Gram-negative bacterium that forms root nodules on alfalfa. The colonization of alfalfa roots by S. meliloti provokes a complex differentiation program resulting in morphological changes in both organisms. Nitrogen-fixing bacteroids have been reported to contain more nucleic acid per cell than the free-living form and thus may undergo a process of endoreduplication within alfalfa (Paau et al., 1977). It is unclear what role DNA replication of the S. meliloti genome has during differentiation; however, the origins of replication in the S. meliloti genome may serve as important elements that coordinate this process within plant cells. The tripartite genome of S. meliloti (Galibert et al., 2001) is composed of a circular chromosome (3.6 Mb) and two megaplasmids, pSymA (1.35 Mb) and pSymB (1.68 Mb). Previously, the chromosomal origin was predicted to be located adjacent to the hemE gene (encoding uroporphyrinogen decarboxylase) on the basis of DNA
strand asymmetry (Capela et al., 2001). Here we show that a DNA fragment encompassing this region can confer autonomous replication to a non-replicating plasmid and that these minichromosomes are maintained in *S. meliloti* at copy numbers of less than one per host chromosome. Using a combined bioinformatic, genetic and biochemical approach we have mapped essential DnaA-binding sites within oriC. We previously characterized the replication origins from both megaplasmids carried by *S. meliloti* (Chain et al., 2000; MacLellan et al., 2005) and this work completes our initial characterization of the third origin of replication carried by this bacterium.

**METHODS**

**Bacterial strains and plasmids.** The bacterial strains used in this study are listed in Table 1. *E. coli* was grown at 37 °C in LB broth and *S. meliloti* was grown at 30 °C in LB supplemented with 2.5 mM MgSO$_4$ and 2.5 mM CaCl$_2$ (LBmC) or on LB solidified with agar (16 g l$^{-1}$).

### Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genotype or characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET431a</td>
<td>Ap$^R$, NusAHis6 tag expression vector, inducible with IPTG</td>
<td>Novagen</td>
</tr>
<tr>
<td>pRK600</td>
<td>pRK2013 npt::Tn9, Cm$^R$</td>
<td>Finan et al. (1986)</td>
</tr>
<tr>
<td>pBBR1MCS-5</td>
<td>Gm$^R$, broad host-range cloning vector, LacZa peptide</td>
<td>Kovach et al. (1995)</td>
</tr>
<tr>
<td>pUCP30T</td>
<td>Gm$^R$, ColE1 oriV cloning vector, oriV</td>
<td>Schweitzer et al. (1996)</td>
</tr>
<tr>
<td>pTH838</td>
<td>pUCP30T with 3 kb oriC (AB24853–AB24854) cloned via EcoRI, hemE</td>
<td>This work</td>
</tr>
<tr>
<td>pTH879</td>
<td>710 bp HindIII–SalI oriC fragment cloned into pUCP30T</td>
<td>This work</td>
</tr>
<tr>
<td>pTH880</td>
<td>875 bp BamHI–SalI oriC fragment cloned into pUCP30T</td>
<td>This work</td>
</tr>
<tr>
<td>pTH881</td>
<td>861 bp SacI oriC fragment cloned into pMB439 suicide plasmid</td>
<td>This work</td>
</tr>
<tr>
<td>pTH1081</td>
<td>pET431a with dnaA (AB26340–AB26341) cloned via BamHI/EcoRI</td>
<td>This work</td>
</tr>
<tr>
<td>pTH1091</td>
<td>PBBR1MCS-5 with dnaA and ribosome-binding site (AB29744–298675)</td>
<td>This work</td>
</tr>
<tr>
<td>pTH105</td>
<td>cloned via XbaI/SacI, same orientation as Plac</td>
<td></td>
</tr>
<tr>
<td>pTH1205</td>
<td>pTH838 with 3 bp deletion in DnaA box #1 TCATCCGCGC→TCGCGCC,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>created via site-directed mutagenesis (AB31447–AB31448)</td>
<td></td>
</tr>
<tr>
<td>pTH1206</td>
<td>pTH838 with 4 bp deletion in DnaA box #2 TCATCCACA→TCACA,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>created via site-directed mutagenesis (AB31449–AB31450)</td>
<td></td>
</tr>
<tr>
<td>pTH1207</td>
<td>pTH838 with 4 bp deletion in DnaA box #3 TTGTTCCACA→TTGTA,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>created via site-directed mutagenesis (AB31451–AB31452)</td>
<td></td>
</tr>
<tr>
<td>pTH1245</td>
<td>pUCP30T with 483 bp hemE–Y02793 intergenic region (AB32323–32324)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1253</td>
<td>pTH838 with 4 bp deletion in DnaA box #4 TTATCAACA→CAACA,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>created via site-directed mutagenesis (AB32750–32751)</td>
<td></td>
</tr>
<tr>
<td>pTH1416</td>
<td>pUCP30T with 2289 bp oriC (ML1817–ML1818) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1442</td>
<td>pUCP30T with 2472 bp oriC (ML1817–AB24853) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1443</td>
<td>pUCP30T with 2823 bp oriC (ML1818–AB24854) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1451</td>
<td>pUCP30T with 2402 bp oriC (ML2444–ML1818) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1452</td>
<td>pUCP30T with 2421 bp oriC (ML2443–ML1818) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1453</td>
<td>pUCP30T with 1783 bp oriC (ML2444–ML2445) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td>pTH1454</td>
<td>pUCP30T with 1802 bp oriC (ML2443–ML2445) cloned via EcoRI</td>
<td></td>
</tr>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> DH5x</td>
<td>F$^{-}$ endA1 hsdR17 (rK, mK) supE44 thi-1 recA1 gyrA96 relA1 Δ(argF-lacZYA) U169 Φ80d lacZ Δ (AMB)</td>
<td>BRL</td>
</tr>
<tr>
<td><em>E. coli</em> MT616</td>
<td>MM294A recA-56 (pRK600)</td>
<td>Finan et al. (1986)</td>
</tr>
<tr>
<td><em>E. coli</em> MT620</td>
<td>MM294A recA-56 R$^R$</td>
<td>Stratagene</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 STAR</td>
<td>F$^{-}$ ompT hsdSB (rK, mK) gal dcm rne131 (DE3)</td>
<td></td>
</tr>
<tr>
<td><em>S. meliloti</em> Rm1021</td>
<td>SU47 str-27</td>
<td>Meade et al. (1982)</td>
</tr>
<tr>
<td><em>S. meliloti</em> Rm5004</td>
<td>Rm1021 recA::Tn5</td>
<td></td>
</tr>
<tr>
<td><em>S. meliloti</em> Rm5004</td>
<td>Rm1021 recA::Tn5</td>
<td></td>
</tr>
<tr>
<td><em>S. meliloti</em> K1012</td>
<td>Rm5004(pTH838) small colony</td>
<td>This work</td>
</tr>
<tr>
<td><em>S. meliloti</em> K1012</td>
<td>Rm5004(pTH838) large colony</td>
<td>This work</td>
</tr>
</tbody>
</table>
Bacterial matings. Plasmids were transferred from *E. coli* (donor) cells to *S. meliloti* (recipients) in triparental matings using MT616, a helper strain that carries transfer functions on pRK600. Strains used in the triparental mating were grown overnight and then washed twice with saline. The *E. coli* donor and recipient cells were adjusted to have twice the cell density as the *S. meliloti* recipient. Equal volumes of cells (20 µl) were spotted onto an LB agar plate and incubated overnight at 30 °C. The mating spot was resuspended in 0-85% NaCl, serially diluted and 100 µl volumes were plated on LB with antibiotic to select for the *S. meliloti* recipient cells (200 µg streptomycin ml⁻¹) and on LB plates containing antibiotics used to select for the recipient and the transferred test plasmid (200 µg streptomycin ml⁻¹ and 60 µg gentamicin ml⁻¹). *E. coli* MT620 recipients were selected for on LB plates containing 20 µg rifampicin ml⁻¹ and transconjugants on LB plates containing 20 µg rifampicin ml⁻¹ and 10 µg gentamicin ml⁻¹.

General molecular biology. Plasmid DNA isolation, genomic DNA isolation, restriction analysis, PCR and site-directed mutagenesis were all performed according to standard protocols. PCR products were purified using QIAquick spin columns (Qiagen). Sequencing and primer synthesis was performed at the MobiX Central Facility (McMaster University, Hamilton, Ontario, Canada) using the ABI PRISM 3100 Genetic Analyser using the BigDye terminator chemistry.

Copy number determination. Total DNA was prepared from cultures grown to an OD₆₀₀ of 0-5-0-6, according to a standard protocol. Genomic DNA (10 µg) was digested in a 30 µl reaction volume for 4 h. Digested genomic DNA was loaded onto a 0-8% agarose gel and electrophoresed at 15 V overnight. Southern hybridization was done using either linearized plasmid DNA or purified PCR products as probes which were radioactively labelled with [γ⁻³²P]ATP using the Roche Random Primed DNA Labelling Kit. Following hybridization, the membrane was exposed to a Storage Phosphor Screen (Amersham Biosciences) for 1 h. After exposure, the screen was scanned on a Storm 820 Phosphimager (Molecular Dynamics) at a pixel size of 50 µm. Band intensities were calculated using the Image Quant 5.2 program (Molecular Dynamics). Plasmid copy number was calculated as a ratio of plasmid signal to chromosome signal. Copy number was determined for triplicate samples and a mean copy number is reported.

Protein purification. The *dnaA* gene was PCR-amplified from genomic DNA using the AB26340 and AB26341 primers (Table 2) and cloned via engineered BamHI and EcoRI restriction sites into the pET43a NusAHis₆ tag expression vector (Novagen) to create *S. meliloti* dnaA-His6 strain. This strain was grown at 37 °C and cloned via engineered BamHI and EcoRI restriction sites into the pET43a NusAHis₆ tag expression vector (Novagen) to create *S. meliloti* dnaA-His6 strain. This strain was grown at 37 °C and a mean copy number is reported.

Electrophoretic mobility shift assay. Probe DNA (~100–200 bp purified PCR product) was quantified visually on a 1-8% agarose gel. One picomole of 5’ terminus was then end-labelled with [γ⁻³²P]ATP using T4 polynucleotide kinase (New England Biolabs). Following the labelling reaction, probes were purified with a QIAquick PCR Purification Kit (Qiagen). Specific activity of the labelled probe was calculated using a liquid scintillation counter.

The binding reaction was set up as described by Schaper & Messer (1995). In a microtube on ice the following were mixed (in this order): 4 µl 5 x binding buffer (100 mM HEPES-KOH, pH 8-0, 25 mM magnesium acetate, 5 mM Na₂EDTA, 20 mM DTT, 25 mg BSA ml⁻¹, 1% Triton X-100, 25% glycerol), ddH₂O, 0-4 µl ATP (50 mM), probe DNA (50 mM), 1 µl poly dI : dC (100 ng ml⁻¹) and purified DnaA (100–500 nM) in a total reaction volume of 20 µl. The reaction was incubated on ice for 10 min and then at room temperature for 20 min. The reactions were loaded onto a 4% polyacrylamide gel and electrophoresed at 14 V cm⁻¹ (252 V) for 10 min and then 9 V cm⁻¹ (162 V) for 2-5 h at room temperature. The gel was dried and exposed to Kodak Scientific Imaging Film and a Storage Phosphor Screen for quantification.

Environmental scanning electron microscopy. For glutaraldehyde fixation, several colonies were used to inoculate 2 ml LBcm with 30 µg gentamicin ml⁻¹ and the culture was grown to an OD₆₀₀ of 0-5. The culture was pelleted in a microtube and the supernatant was removed. The pellet was then resuspended in 1 ml 0-2 M sodium cacodylate buffer (pH 7-4). One millilitre of 0-2 M sodium cacodylate buffer containing 5% glutaraldehyde (pH 7-4) was added, thus changing the effective concentration of glutaraldehyde to 2-5%, and the tubes were inverted a couple of times and left for 1 h at room temperature.

For slide preparation, a cover glass was mounted on an aluminium ESEM stub with conductive glue (equal parts white Elmers glue and colloidal graphite) making sure that a line of conductive glue was made from the edge of the glue spot to the edge of the cover glass and just around to the sample side of the glass. The mounted cover glass was air-dried for 30 min and then coated with a 5 nm layer of gold using a Sputter Coater. The glutaraldehyde fixed sample was washed six times in 2 ml ddH₂O to remove all traces of salt. Then, 1 µl resuspended sample was spotted onto the gold-coated glass slide and the spot was allowed time to air dry. For visualization and image capturing the stub was placed into an Electroscan 2020 Environmental Scanning Electron Microscope and set to Wet mode. Samples were viewed at 2-4–4-0 Torr with an accelerating voltage of 20–30 keV. Various magnifications were used and images were saved as TIFF files.
RESULTS AND DISCUSSION

Sequence analysis of predicted oriC reveals hallmark features of chromosomal replication origins

The origin of replication on the S. meliloti chromosome has been predicted to be approximately 400 kb from dnaA and adjacent to hemE on the basis of GC skew (Capela et al., 2001) and this corresponds to the same location as the well studied chromosome origin from C. crescentus (Marczynski & Shapiro, 1992; Marczynski et al., 1995; Brassinga & Marczynski, 2001). The C. crescentus chromosome origin is located ~ 5 kb from the dnaA gene region, a sequence with a similar genetic arrangement to the S. meliloti dnaA region (Brassinga et al., 2001). This yidB1-rpsT-dnaA region is unusual in that it deviates from the rnpA-rmpH-dnaA-dnaN-recF-gyrB-rnpA structure linked to many bacterial chromosome origins (Yoshikawa & Ogasawara, 1991; Ogasawara & Yoshikawa, 1992). Interestingly, the parB-parA-gidA-gidB-thdF gene cluster located ~ 70 kb from the C. crescentus oriC is within 8 kb of the predicted S. meliloti oriC. We examined the predicted S. meliloti oriC region for features characteristic of bacterial chromosome origins. The most obvious feature found was an exceptionally AT-rich (~ 30 mol% G+C over 80 nt) region contrasting with the

<table>
<thead>
<tr>
<th>Oligonucleotide name</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB24853</td>
<td>AAGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB24854</td>
<td>CGGAATTCACCAAGAAAGTGAAGTGAAG</td>
</tr>
<tr>
<td>AB26340</td>
<td>CGCGGATATGGCGATTAATGGCG</td>
</tr>
<tr>
<td>AB26341</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB27527</td>
<td>GCGAAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB27528</td>
<td>GCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB27526</td>
<td>GCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>dnaA32AS</td>
<td>AATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB29675</td>
<td>GAGCTCTATTGTTGATCGTGCCTTC</td>
</tr>
<tr>
<td>AB29744</td>
<td>ATCTAGAATACGTGAATGGAGCCAG</td>
</tr>
<tr>
<td>AB30423</td>
<td>GTAGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31447</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31448</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31449</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31450</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31451</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB31452</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB32750</td>
<td>GCTTTTTTGGTCGCTGCTCTCCAGG</td>
</tr>
<tr>
<td>AB32751</td>
<td>GCGAAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB32323</td>
<td>TCGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>AB32324</td>
<td>ATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML700</td>
<td>GTTTGCCGATGATGTAGTGAGG</td>
</tr>
<tr>
<td>ML701</td>
<td>CTCCGCGGATGATGTAGTGAGG</td>
</tr>
<tr>
<td>ML702</td>
<td>GATCGGTTGCTACGTGATGAGG</td>
</tr>
<tr>
<td>ML703</td>
<td>GAACTCCTCCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML1182</td>
<td>GTTTGGCGATGATGTAGTGAGG</td>
</tr>
<tr>
<td>ML1183</td>
<td>CTCCGCGGATGATGTAGTGAGG</td>
</tr>
<tr>
<td>ML1187</td>
<td>TGAAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML1188</td>
<td>ATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML2444</td>
<td>ATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML2443</td>
<td>ATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML2445</td>
<td>ATGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML2796</td>
<td>CGCGCAACGTGGAGAGTACAGCC</td>
</tr>
<tr>
<td>ML2797</td>
<td>CCTCGCGGATGATGTAGTGAGG</td>
</tr>
<tr>
<td>ML3257</td>
<td>CTTTCAGAATCGCTGTGATGAGG</td>
</tr>
<tr>
<td>ML3258</td>
<td>AATCGAATTCACCGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML3852</td>
<td>GCAATGATGTGGAGCGGCAAGCCAG</td>
</tr>
<tr>
<td>ML3853</td>
<td>CGCGGAATTCACCGCGGCAAGCCAG</td>
</tr>
</tbody>
</table>

Table 2. Oligonucleotides used in this study

Restriction sites that were engineered into the primers are shown in bold.
flanking DNA and the rest of the chromosome which has a G+C content of 62 mol% (Fig. 1a). In *E. coli*, Bramhill & Kornberg (1988b) showed that duplex opening occurred in the AT-rich region of oriC and presumably the lower double-stranded thermodynamic stability in this region promotes strand dissociation.

DnaA-binding sites are a consistent feature of bacterial oriCs. Sequences recognized by DnaA have been most extensively studied in *E. coli* and these 9 bp sequences seem relatively conserved amongst bacteria. We scanned the predicted *S. meliloti* oriC region for matches based upon the expanded *E. coli* consensus sequence (T/C)(T/C)(A/T/C)T(A/C)C(A/G)(A/C/T)(A/C) (Schaefer & Messer, 1991). As shown in Fig. 1(b), five putative DnaA boxes were identified in the 3 kb oriC region. An additional box was also found in the dnaA gene region. This DnaA box (box 5, Fig. 1b) is an 8/9 bp match to the *E. coli* consensus sequence and is an exact match to an experimentally identified DnaA box found in the *C. crescentus* chromosome origin (Marczynski & Shapiro, 1992). DnaA box 3 (also an 8/9 bp match to the consensus) is identical to the sequences of DnaA boxes found in chromosome origins from *Streptomyces* spp. (Majka et al., 2001) and *Micrococcus luteus* (Fujita et al., 1990).

Since the initiation of replication of the megaplasmid origins may also be subject to regulation by DnaA, the repABC replicator regions for the pSymA and pSymB megaplasmids were also scanned for DnaA-binding sites. As shown in Fig. 1(c), five putative DnaA boxes were identified in the pSymA repABC region. Interestingly, DnaA box 2, found in the *hemE–Y02793* intergenic region, is found twice in the pSymA replicator region, once 158 bp upstream of the translational start site of repA2 and again inside the repA2 ORF. Eleven DnaA-binding sites were predicted in the pSymB repABC locus (Fig. 1d). It has been documented that only the repC ORF and sequence downstream of repC is necessary for repABC plasmid replication (Ramirez-Romero et al., 2000) and thus the origin of replication must be encoded either within repC or just downstream of the gene. This work predicts a single DnaA box in the sequence downstream of repC in both megaplasmid replicator regions and this box may therefore serve a role in recruiting the cellular replication machinery to the megaplasmid origins. However, the distribution of these putative DnaA boxes appears to be biased to upstream of or within the repA1 gene (8 of 11 predicted sites in the region). DnaA may possibly be involved in regulating repABC gene expression. There is no experimental evidence for the involvement of DnaA in repABC plasmid replication, but this work might point to such a role.

**Autonomous replication of the DNA region encompassing the predicted oriC**

Exploiting the ability of a cloned sequence to support autonomous replication of a normally non-replicating plasmid in the host cell has been a successful strategy for isolating the chromosome origins from *Enterobacter aerogenes*, *Klebsiella pneumoniae* (Harding et al., 1982), *Erwinia carotovora* (Takeda et al., 1982), *Vibrio harveyi* (Zyskind et al., 1987).
et al., 1983), *Pseudomonas aeruginosa*, *Pseudomonas putida* (Yee & Smith, 1990) and *Streptomyces lividus* (Zakrzewska-Czerwinska & Schrempf, 1992). We used this strategy to examine whether the previously predicted *S. meliloti* oriC region could support autonomous replication. For this purpose a 3 kb region encompassing the AT-rich region, the *hemE* and Y02793 genes, and the five putative DnaA boxes was amplified by PCR from *S. meliloti* genomic DNA using primers AB24853 and AB24854. This DNA was cloned into pUCP30T (a plasmid that cannot replicate in *E. coli*) to form plasmid pTH838 (see Fig. 3b). pTH838 was transferred from *E. coli* donor cells into a recA derivative of *S. meliloti* (Rm5004) via conjugation. The ability of pTH838 to promote the formation of transconjugant colonies on medium containing an antibiotic (60 μg gentamicin ml⁻¹) to select for its presence is indicative of autonomous replication. The pTH838 plasmid transferred into both Rm5004 and wild-type Rm1021 cells and transconjugants were obtained at a frequency of 10⁻¹ per recipient cell. The high transfer frequency into both recipient strains reflects an ability of pTH838 to autonomously replicate because the recA mutation in Rm5004 prevents homologous recombination of the plasmid with the *hemE* locus on the chromosome. To our surprise, every time this mating experiment was performed, both small and large *S. meliloti* transconjugant colonies arose on selective medium after incubation. Both small and large Rm5004(pTH838) transconjugants colonies were purified three times on selective medium, then tested for the presence of an autonomously replicating plasmid by Southern blot hybridization. A single small colony and a single large colony (done in triplicate) was used to inoculate LB and total DNA was prepared following growth for 36 h. As a control, total DNA was prepared from a pUCP30T cointegrant strain RmK569 (a wild-type Rm1021 derivative in which a single copy of pUCP30T has integrated at the *pstS* locus). Total transconjugant genomic DNA was restricted and probed in a Southern blot with labelled pUCP30T. pTH838 DNA was detected as a single restriction fragment and this is indicative of the plasmid being maintained as a closed circular molecule in the transconjugant cells. Two restriction fragments hybridized with the pUCP30T probe in the control strain RmK569. Similarly, integration of pTH838 into the chromosomal *hemE* locus would result in two restriction fragments, whereas only a single fragment was observed (data not shown), revealing that pTH838 is an autonomously replicating plasmid.

The conjugal transfer of pTH838 into Rm5004 resulted in both small and large transconjugants with approximately 50 times more small than large colonies. The rate and extent to which colonies form upon selection medium is at least in part dependent on the phenotypic level of resistance conferred by the antibiotic resistance gene, and the resistance gene dosage (i.e. copy number of plasmid carrying the gene) might influence this parameter. To investigate whether the difference in colony size was due to differences in the copy number of the pTH838 plasmid, we carried out a Southern blot hybridization to assess the relative amounts of the autonomously replicating pTH838 plasmid in both the small and large transconjugants. Large and small colonies were streak-purified to single colonies three times and total DNA was prepared in triplicate from cultures inoculated with either a small or large transconjugant colony. Lanes: 1, small #1; 2, small #2; 3, small #3; 4, large #1; 5, large #2; 6, large #3; 7, RmK569(pTH838); 8, Rm1021 (wild-type); 9, Rm5004; 10, HindIII-restricted pTH838 plasmid DNA. A PCR product of the 477 bp *hemE*-Y02793 intergenic region (using primers AB32323 and AB32324) was randomly prime-labelled and used as a probe. The 5-8 kb band is probe-hybridized with plasmid DNA and the 3-4 kb band corresponds to the chromosomal signal. Intensities of the bands were determined and relative copy number was calculated as a ratio of plasmid to chromosome signal. The mean copy number was calculated from the three independent colonies tested.
hemE–Y02793 intergenic region (Fig. 2). The intensities of the plasmid signal (5–8 kb) and the chromosome signal (3–5 kb) were quantified with a phosphoimager. A ratio of plasmid signal : chromosome signal was determined for each transconjugant sample and then the three ratios for each of the small and large strains were used to calculate a mean number of plasmid molecules per chromosome. The mean copy number of the three small transconjugants was determined to be 0·053 plasmids per chromosome and the mean copy number of the three large transconjugants was 0·135 plasmids per chromosome. Thus the large colonies contained approximately twice as much autonomously replicating pTH838 plasmid DNA as the small colonies.

It is not uncommon to observe altered colony morphologies in bacterial cells containing plasmid copies of the host chromosome origin. Colony morphologies consistent with bacterial cells containing plasmid copies of the host chromosome origin. Colony morphologies consistent with the delayed colony development phenotype reported for S. meliloti in this work have been observed in C. crescentus cells harbouring minichromosomes (Marczynski et al., 1995). Mutations that impair the replication of C. crescentus minichromosomes do so in a way that leads to a reduction of the copy number of the autonomously replicating plasmid and results in a delayed growth phenotype on selective medium. This parallels our observation that Rm5004(pTH838) transconjugant colony size directly correlates with the copy number of the pTH838 minichromosomes.

To validate the calculated pTH838 copy number of less than one, K1012 and K1013 and the control strain RmK569 (Rm1021 with a single copy of pUCP30T that has integrated at the pstS locus) were grown in LB with antibiotic selection. Two hundred colonies were then patched to mid-exponential phase and plated onto LB plates without antibiotic. Two hundred colonies were then patched back onto selective medium and incubated at 30 °C for 4 days. A total of 20 of the 200 (0·1) patches and 13 of the 200 (0·065) patches were gentamicin-resistant from the K1013 and K1012 cultures, respectively. As expected, all 200 of the RmK569 patches maintained gentamicin resistance because of the stability of the chromosomally integrated copy of pUCP30T. These values are very close to the mean copy number calculated from the Southern blot (Fig. 2), 0·053 vs 0·065 for the small transconjugants and 0·135 vs 0·1 for the large transconjugants. These ratios represent the fraction of cells in the cell population growing in the presence of gentamicin that actually contain a pTH838 plasmid molecule or the gentamicin 3′-acetyltransferase protein encoded by the pTH838 plasmid.

We employed a second indirect method to establish that pTH838 was autonomously replicating in Rm5004. The method is based on the premise that autonomously replicating plasmids, but not integrated plasmids, can be readily transferred from S. meliloti to E. coli. Thus we tested strains K1012 [Rm5004(pTH838) small transconjugant], K1013 [Rm5004(pTH838) large transconjugant] and strain Rm-K569 (as a control) for their ability to transfer pTH838 (as monitored by gentamicin resistance) to E. coli. The results from these experiments (Table 3) clearly showed a high frequency of transfer of gentamicin resistance from the Rm5004(pTH838) transconjugants K1012 and K1013 to E. coli (>10−1 per donor). In contrast no transfer of the integrated copy of pUCP30T in RmK569 to E. coli was detected (<10−8 per donor). Restriction analysis of plasmid DNA prepared from gentamicin-resistant E. coli transconjugant cells confirmed that the transferred plasmid was pTH838 (data not shown). These data demonstrate that pTH838 is capable of autonomous replication and therefore represents a mini-derivative of the chromosomal basic replicon.

**Table 3. Transconjugation frequencies per recipient and donor resulting from the transfer of pTH838 from S. meliloti into E. coli**

Three S. meliloti strains were used as donors; E. coli MT620 was used as the recipient and MT616 was used as the mobilizing strain in the triparental mating. The mating spot was resuspended, serial-diluted and plated onto selective medium. Transconjugants were selected for on LB agar with 20 μg rifampicin ml−1 and 10 μg gentamicin ml−1; recipients were selected for on LB agar with 20 μg rifampicin ml−1 and donors were selected for on LB agar with 200 μg streptomycin ml−1 and 60 μg gentamicin ml−1. The transconjugation frequency for S. meliloti K569 and E. coli MT616 was 0 (<10−8 per donor).

<table>
<thead>
<tr>
<th>Donor</th>
<th>Transconjugation frequency</th>
<th>Transconjugants/ E. coli recipient</th>
<th>Transconjugants/ S. meliloti donor</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. meliloti K1013</td>
<td>0·23</td>
<td>0·13</td>
<td></td>
</tr>
<tr>
<td>S. meliloti K1012</td>
<td>0·28</td>
<td>0·13</td>
<td></td>
</tr>
</tbody>
</table>

Delineation of the minimal sequence that supports autonomous replication

As described above, the nucleotide sequence coincident with the predicted oriC encodes hallmark features of chromosomal origins, including an AT-rich sequence and five putative DnaA boxes, and a 3 kb region encompassing these features is able to autonomously replicate in S. meliloti cells. To delineate the oriC region required to allow the suicide plasmid pUCP30T to replicate in S. meliloti, various subfragments from the insert in pTH838 were subcloned into pUCP30T. The resulting plasmids were examined for their ability to replicate in Rm5004 (Fig. 3b). The sequence encompassing the AT-rich region and DnaA boxes 2, 3 and 4 (pTH1245) was not sufficient for autonomous replication (transconjugant colony formation). The 5′ limit of the sequence required for autonomous replication was mapped to within 19 bp (the difference between pTH1451 and pTH1452, and the difference between pTH1453 and pTH1454) and within this essential 19 bp is the predicted DnaA box 5. The 3′ limits required for transconjugant colony formation (pTH1454) extends 250 bp into the Y02793 ORF and thus the plasmid copy of Y02793 is not...
required for autonomous replication. The minimal size of the *S. meliloti* oriC is 1802 bp and this is much larger than the 437 bp *C. crescentus* minimal chromosome origin (Marczynski & Shapiro, 1992). It may be that the *hemE* ORF itself is not required for replication, but only the DnaA box located downstream of *hemE* and thus deletion of the intervening 956 bp *hemE* gene would reduce the size of the *S. meliloti* oriC to 846 bp. The minimal nucleotide sequence required to support autonomous replication of the suicide plasmid pUCP30T suggests that the functional *S. meliloti* oriC sequence may be defined by the distribution of DnaA boxes that flank the AT-rich region.

We therefore examined the importance of the predicted DnaA boxes to replication by using a mutagenesis approach. The five DnaA boxes identified in pTH838 were mutated by oligonucleotide site-directed mutagenesis. The mutations in DnaA boxes 5, 2, 3 and 4 were 4 bp deletions and the mutation in DnaA box 1 was a 3 bp deletion (Fig. 3c). Plasmids carrying mutations in boxes 2, 3 and 4 lost the ability to autonomously replicate in *S. meliloti*, suggesting that these sequences are required for replication of plasmid-borne oriC (Fig. 3c). As expected, transconjugant colonies formed at a low frequency when mutant oriC plasmids were transferred into Rm1021, probably due to homologous

---

**Fig. 3.** (a) A genetic map of the *S. meliloti* chromosomal origin of replication showing the *hemE* and Y02793 loci, and the *dnaA* gene region of the *S. meliloti* chromosome which is located 400 kb away from oriC. A G+C content graph indicates the location of an exceptionally AT-rich region in the *hemE*–Y02793 intergenic region. DnaA boxes that match the expanded *E. coli* consensus sequence (T/C)(T/C)(A/T/C)T(A/C)C(A/G)(A/C/T)(A/C) (Schaefer & Messer, 1991) are shown with dark triangles and boxes that match the consensus 8/9 are represented with open triangles. (b) Subclones of the 3 kb origin in pTH838 are scored for their ability to generate transconjugant colonies when cloned into a suicide plasmid (pUCP30T) and transferred into an *S. meliloti* recA” strain (Rm5004). (c) Location of site-directed mutations in the *S. meliloti* oriC. The locations of the site-directed deletions are indicated with an inverted triangle. Transconjugation frequencies are provided in the accompanying table for transfer into Rm1021 and Rm5004.
recombination at the oriC genomic locus. Rm1021 was used as a control recipient in the replication assay to ensure that the oriC plasmids were still mobilizable after mutagenesis. The mutation in DnaA box 1 did not influence pTH838 replication and, surprisingly, the mutation carried on pTH1518 (mutation in DnaA box 5) also did not abolish replication even though our deletion analysis (Fig. 3b) demonstrated that this sequence was essential for replication. It is possible that the removal of the 4 bp in DnaA box 5 generated a sequence still capable of interacting with the replication initiator. DnaA box 5 (5'-TGATCCACA-3') is an 8/9 bp match to the expanded E. coli DnaA box consensus (Fig. 1); the site-directed mutation (deletion of GATC) in this predicted binding site generates 5'-TCA-CAGATA-3' and this sequence only deviates from the consensus sequence at positions 4 and 6. In other words, the 4 bp deletion in box 5 may result in the reconstitution of a sequence that can still bind DnaA.

DnaA binds to predicted binding sites in the hemE–Y02793 intergenic region

The ability of DNA fragments carrying the various putative DnaA-binding sites to bind purified DnaA protein was examined in electrophoretic mobility shift assays. We tested mutations in putative DnaA-binding sites that abolish replication of plasmid-borne oriC for their influence on DnaA binding. The six DNA sequences that were used in DnaA-binding experiments included a 190 bp probe (ML700/ML701) that contained all three predicted DnaA boxes in the hemE–Y02793 intergenic region (DnaA boxes 2, 3 and 4), a 178 bp probe (ML1182/ML1183) with the same sequence as the 190 bp probe except with 4 bp deletions in all three predicted DnaA boxes, a 117 bp probe (ML2796/ML2797) containing DnaA box 5 downstream of hemE, a 197 bp fragment (ML3257/ML3258) from the dnaA promoter that contains a predicted DnaA box overlapping the translational start site of the dnaA gene, a 233 bp repA2 promoter probe (ML702/ML703) which includes a DnaA box with an exact match to DnaA box 2 in oriC and a 125 bp repA1 promoter probe (AB27527/AB27526) containing two DnaA boxes.

Two complexes were resolved with the 190 bp oriC probe as seen in Fig. 4; however, in some cases using this probe resulted in the formation of three complexes (data not shown), consistent with the number of binding sites in the probe. Three rounds of site-directed mutagenesis of pTH838 were required to introduce the three mutations needed to generate a template that could be used for PCR amplification of the 178 bp probe. DnaA did not interact with this target as it did with the wild-type sequence and thus it appears that the mutations created in the hemE–Y02793 intergenic region represent bona fide deletions in DnaA boxes (Fig. 3c). Only a weak interaction was observed with the 117 bp DnaA box 5 probe. Complexes were not detected with the dnaA promoter probe or the repA1 promoter probe. DnaA did complex with the repA2 promoter probe, probably at the predicted DnaA box 158 bp upstream of the repA2 translational start site. DnaA has been implicated as a transcriptional factor as it autoregulates dnaA gene expression (Atlung et al., 1985) and either represses or activates expression from several other genes (Messer & Weigel, 1997). Thus the DnaA-binding site upstream of the repA2B2C2 operon may be biologically relevant such that expression of the pSymA genes encoded in the replicator region are transcriptionally coordinated with chromosome replication.

Expression of S. meliloti DnaA in E. coli and S. meliloti results in a block in cell division

DnaA seems to be conserved in bacteria as a regulator of replication initiation events at the chromosome origin and it has been reported that overexpression of DnaA in E. coli stimulates the initiation reaction at oriC (Atlung et al., 1987; Skarstad et al., 1989). This results in a loss of the correct timing in the cycle of replication initiation and leads to a block in cell division, thus generating filamentous cells (Pierucci et al., 1989). This block in cell division as a result of

---

**Fig. 4.** Electrophoretic mobility shift assay with target DNA containing predicted DnaA boxes and purified S. meliloti DnaA. Several target DNA fragments were used in a binding reaction with purified DnaA. Three concentrations of protein were added to the binding reaction (100, 300 and 500 nM) and the DNA fragments were also loaded without added protein. The size and description of each labelled DNA fragment is indicated above the lanes. The first left-hand lane for each DNA target is DNA without protein added and the next three lanes (left to right) are the DNA target incubated with 100, 300 and 500 nM purified DnaA, respectively.
increased DnaA levels has also been reported for *Mycobacterium smegmatis* (Greendyke *et al.*, 2002). We examined whether DnaA would have similar consequences on *S. meliloti* cell morphology if overexpressed. To this end the dnaA gene and 20 bp upstream of the predicted translational start (including a predicted ribosome-binding site) was PCR-amplified using primers AB29744 and AB29675. Primer AB29744 was designed with three engineered stop codons in all three reading frames to prevent a fusion to the Lac peptide encoded in pBBR1MCS-5. The PCR product was cloned into pBBR1MCS-5 such that the dnaA gene was under the transcriptional control of the *E. coli* lac promoter, generating plasmid pTH1091. The pTH1091 DnaA expression plasmid was transferred into wild-type *S. meliloti* Rm1021 and *E. coli* DH5α and transconjugant cells were visualized with an environmental scanning electron microscope. Many more of the *E. coli* cells expressing *S. meliloti* DnaA appeared filamentous compared to the parent DH5α (compare Fig. 5a and b), suggesting that *S. meliloti* DnaA has some form of cross-functionality and can perturb cell division in other Proteobacteria. *S. meliloti* cells expressing DnaA from the plasmid promoter also displayed an apparent block in cell division, with many of the cells growing up to 15 μm in length compared to the 1 μm length of Rm1021 wild-type cells (compare Fig. 5c and d). The *S. meliloti* filaments are complex and in most cases the cells display many branches with swollen areas flanked by regions that appear partially septated, as shown under higher magnification (Fig. 5e and f).

We cannot deduce the molecular mechanism underlying this phenotype; however, the morphological changes resulting from the plasmid-mediated expression of *S. meliloti* DnaA are very similar to the pleomorphic cell shape phenotype reported for *S. meliloti* cells overexpressing FtsZ or when treated with DNA-damaging agents (Latch & Margolin, 1997). It would seem that the default pathway for a block in cell division that arises from perturbations in

---

**Fig. 5.** Environmental scanning electron micrographs of *E. coli* and *S. meliloti* cells expressing *S. meliloti* DnaA from the *E. coli* lac promoter in the pBBR1MCS5 plasmid derivative, pTH1091. (a) shows *E. coli* DH5α harbouring pBBR1MCS5 and (b) shows *E. coli* DH5α carrying pTH1091. (c) shows wild-type Rm1021 harbouring the pBBR1MCS5 plasmid and (d), (e) and (f) are wild-type Rm1021 cells expressing DnaA from pTH1091. Bars, 25 μm (a, b), 30 μm (c, d) and 15 μm (e, f).
chromosome replication in this endosymbiont is a complex change in cell shape. Interestingly, many of the cells assume branched Y-shaped architectures that look very similar to the fully differentiated bacteroids found inside plant cells in root nodules.

Here we report the identification of a replication origin from the S. meliloti circular chromosome which represents the first such origin to be localized in the Rhizobiaceae and the second chromosome origin to be experimentally defined in the Alphaproteobacteria. This is the third autonomously replicating sequence to be identified in the S. meliloti genome. The other two sequences were isolated from the pSymB megaplasmid (Margolin & Long, 1993; Chain et al., 2000). The S. meliloti oriC possesses the hallmark features of a bacterial chromosome origin, such as sequences that interact specifically with the replication initiator DnaA (DnaA boxes) flanking an exceptionally AT-rich region. Interestingly, the S. meliloti chromosome origin is located greater than 400 kb from the dnaA gene which, in other bacteria, is often closely linked to the replication origin. This explains why an autonomously replicating sequence was not detected in the vicinity of dnaA in previous attempts to localize this replication origin (Margolin et al., 1995). S. meliloti transconjugant cells harbouring oriC plasmids display a delayed growth phenotype, perhaps explaining why a screen for autonomously replicating sequences encoded in the S. meliloti genome conducted by Margolin & Long (1993) did not detect the chromosome origin.

Previous reports have demonstrated that requirements for minichromosome replication may be very different than those for the chromosomal origin. The necessity of certain DnaA-binding sites and DNA-binding proteins HU and IHF for oriC plasmid replication, but not for chromosome replication (Weigel et al., 2001; Asai et al., 1998), suggests that the plasmid-borne oriC may adopt a very different DNA topology than the chromosomal origin. It is possible that constraints on DNA topology in S. meliloti minichromosomes may cause a severe reduction in the number of replication initiation events occurring at oriC in pTH838 and thus result in copy numbers of much less than one per chromosome.

The copy numbers of plasmids encoding bacterial oriC sequences is quite diverse. The best characterized bacterial minichromosome from E. coli has a copy number of approximately 38 per cell (Lobner-Olesen et al., 1987), but this does not represent the typical plasmid. Plasmids replicating from the cloned Pseudomonas chromosome origin have been reported to be present at as low as 0-7 copies per cell (Yee & Smith, 1990). The ColE1 replication origin has also been shown to exert a strong effect on the copy number of Mycobacterium tuberculosis minichromosomes and can reduce the copy number from approximately 17 to 0-6 plasmids per chromosome (Qin et al., 1999). The cause of this instability is unknown, but we do not rule out the possibility that the ColE1 replication origin present in pUCP30T may be the cause of the low S. meliloti minichromosome copy number.

While we refer to the region identified in this report as the oriC, we note that direct experimental proof that replication of the S. meliloti chromosome originates from this region has yet to be obtained. Similarly, the presumed replication origins for the pSymA and pSymB megaplasmids have yet to be experimentally verified. Such verification would seem to be very worthwhile and since this would require synchronization of cell division, these experiments could also address whether replication of the chromosome and two megaplasmids is coordinately regulated.

ACKNOWLEDGEMENTS

This work was primarily supported with grants from the Natural Sciences and Engineering Council of Canada to T. M. F. C. S. was also supported with funding from Genome Canada through the Ontario Genomics Institute and with funding from the Ontario research and development challenge fund to T. M. F.

REFERENCES


Fuller, R. S., Funnell, B. E. & Kornberg, A. (1984). The DnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* 38, 889–900.


