INTRODUCTION

How diverse bacteria control chromosome replication remains a challenging problem. The chromosome replication origins (oris) of distantly related bacteria such as *Bacillus subtilis* and *Caulobacter crescentus* are clearly significant variations of the established *Escherichia coli* model (Mott & Berger, 2007; Zakrzewska-Czerwinska et al., 2007). DnaA protein binds and initiates replication at most eubacterial oris, and like *E. coli*, both *B. subtilis* (Ogasawara et al., 1991) and *C. crescentus* (Gorbatyuk & Marczynski, 2001) also require DnaA to initiate chromosome replication. However, to cite one significant difference, most bacteria lack the DNA adenine methylation (Dam) (GATC) methylation and SeqA proteins that *E. coli* employs to repress (sequester) its ori (Kaguni, 2006). The dimorphic growth of *C. crescentus* requires cell-type-specific controls of chromosome replication. *C. crescentus* divides asymmetrically, producing distinct swarmer and stalked cell types, whose chromosomes follow distinct replication programs (McAdams & Shapiro, 2003). The progeny swarmer cell maintains a non-replicating chromosome, while the progeny stalked cell initiates chromosome replication. The swarmer cell does not initiate chromosome replication until it differentiates into a stalked cell. The CtrA protein contributes significantly to replication control. The *C. crescentus* (CB15) chromosome origin (*Cori*) has five CtrA binding sites (Siam & Marczynski, 2000), and CtrA protein is present in swarmer cells but absent in stalked cells (Domian et al., 1997; Bowers et al., 2008). The simplest hypothesis proposes that CtrA binds and represses *Cori* in swarmer cells but not in stalked cells.

Whether other bacteria use CtrA or different regulators to control chromosome replication is only one of many issues that can be addressed by comparative ori analysis. When related DNA sequences are aligned, functionally important DNA motifs are conserved against a background of random variations. When applied to bacteria closely related to *E. coli*, this approach reveals oris with conserved DnaA
binding sites and conserved Dam methylation sites (Zyskind et al., 1983), whose importance has later been confirmed by functional analysis (Mott & Berger, 2007; Zakrzewska-Czerwinska et al., 2007).

Such considerations motivated our isolation and comparative analysis of oris from freshwater Caulobacter (FWC) and marine Caulobacter species (MCS). Unlike E. coli and other bacteria, the freshwater CB15 Cori is positioned between divergently transcribed operons beginning with the hemE biosynthetic gene on one side and an ORF (CC0001/COG1806/Duf299, domain of unknown function) on the other side (Marczynski & Shapiro, 1992; Nierman et al., 2001). This genetic combination is shared by related alpha-proteobacteria and we have proposed that these genes predict oris in the intervening DNA (Brassinga et al., 2001). Our present studies support this hypothesis, and DNA alignments suggest a consensus FWC ori. We describe how comparative analysis constrains previous hypotheses on replication control. Our work reveals that marine Caulobacter MCS10 ori is also repressed by CtrA. But surprisingly, our analysis also shows that repression by CtrA evolved independently in the MCS and FWC oris. We discuss a specific evolution and its relevance to this apparent convergent

RESULTS

Autonomous plasmid replication assays for Caulobacter replication origins

We isolated presumed replication origins (oris) between hemE and Duf299 (Fig. 1a) by a direct PCR cloning strategy, described in Methods. To test for autonomous replication (AR), oris plasmids were introduced by electroporation into FWC and MCS cells and spread on antibiotic selection plates (summarized in Table 1). Cell electroporation with empty vector plasmids (pDrive and pGM1585) provided negative controls, since they replicate in E. coli but not in Caulobacter. The broad-host-range plasmid pRK290 provided a positive AR control (Ditta et al., 1980). AR was first scored by observing at least 100-fold more colonies on the test plasmid plates compared with the negative controls. Next, presumed positive colonies were taken to inoculate liquid cultures and the cells were treated using a modified alkaline-lysis plasmid extraction protocol (Marczynski & Shapiro, 1992). When plasmids were detected by agarose gel electrophoresis with ethidium bromide staining, they were next tested with diagnostic restriction endonuclease digestions. These included Hinfl digestions, which test for alpha-proteobacteria-specific DNA methylation (at GANTC sites) that is absent in E. coli (Wright et al., 1997). The input plasmids were sensitive to Hinfl digestion, and they gained resistance to Hinfl digestion upon replication in Caulobacter cells (data not shown). A positive AR score in Table 1 indicates that all of these criteria were met. Control plasmid pRK290 was easily introduced and recovered (Table 1), indicating that all five

METHODS

Strains and growth conditions. Bacterial strains and plasmids used in this study are summarized in Supplementary Table S1. Caulobacter growth conditions have been described elsewhere (Abraham et al., 1999). All Caulobacter strains were grown at 28 °C in peptone yeast extract (PYE) or in PYE plus 1% sea salts (Sigma), as described elsewhere (Abraham et al., 1999). For antibiotic selection, 20 μg ampicillin ml⁻¹, 50 μg kanamycin ml⁻¹ or 2 μg tetracycline ml⁻¹ were added where appropriate for the corresponding plasmid (Amp-R, Kan-R or Tet-R resistance genes).

DNA isolation, manipulation and analysis. To isolate ori DNA, we designed degenerate PCR primer pairs (Supplementary Table S2) based on conserved coding sequences. Chromosomal DNA was prepared from saturated cultures, as previously described (Marczynski & Shapiro, 1992), and used with PCR primers (Supplementary Table S2) to amplify ori DNA with Pfu polymerase (Stratagene) as specified by the manufacturer. A preliminary titration with magnesium chloride (0–10 mM) was required. These PCR DNA bands were ligated directly into open (U overhang) pDrive plasmids (Qiagen), and selected clones were sequenced from both ends as recommended by the manufacturer. The Sheldon Sequencing Center (McGill University) provided the DNA sequencing and raw data files. Identified ori DNA molecules were also mobilized by their flanking BamHI sites into pGM1585. Site-directed mutagenesis was performed with the QuickChange kit (Stratagene) using the primers specified in Supplementary Table S3. The ori DNA sequences of strains CB15, NA1000 (CB185N) and CB2 were identical, and these are treated synonymously in our analysis. Likewise, our MCS10 ori DNA sequences were identical to those in the recently published Maricaulis maris MCS10 genome. All DNA sequence sources are listed in Supplementary Table S4, and they were made available through www.ncbi.nlm.nih.gov and www.jgi.doe.gov. The conditions for performing autonomous replication assays by cell electroporation, including plasmid extraction and endonuclease analysis, have been described previously (Marczynski et al., 1995). Plates usually required 2–4 days’ growth depending on the Caulobacter species, but were always incubated over 7 days to establish a negative result.

Footprint analysis. The conditions for standard DNase I protection experiments, for CtrA protein binding and for preparing CTP [³²P] end-labelled fragments have been described previously (Siam & Marczynski, 2000). Preparation of the MCS10 ori [³²P] end-labelled fragment required a further subcloning from the pMCS10 ori plasmid (Supplementary Table S1) into the BamHI and EcoRI sites of pKS BluescriptII (Stratagene). This step used the PCR primers 5’-GAT CGG ATC TTC GAG ACC ATG TG (new BamHI end) and 5’-GAT CGA ATT CCG TAT GGG CTC GAG AGT C (new EcoRI end). The MCS10 ori was [³²P] end-labelled at the new BamHI site (Figs 5 and 6). The glutathione S-transferase (GST)-tagged CtrA protein was produced by first ligating the BamHI–EcoRI (ctra) fragment from pGM1397 into pGEX-2T (Amersham). This GST-CtrA protein was then expressed and purified as recommended by the manufacturers. We used experimental conditions established for E. coli DnaA (Li & Crooke, 1999) to purify His-tagged C. crescentus DnaA protein and bind it to MCS10 ori (Fig. 5). The E. coli dnaA gene, on plasmid pZL411 (Li & Crooke, 1999), was replaced by the C. crescentus dnaA gene. The dnaA gene was PCR-amplified from pGM2195 (Gorbatyuk & Marczynski, 2005) using primers 5’-GCT TGA TAT GGG CTG CAG AGT C and 5’-TCT AGA GGA TCC CCT GTC TCC AGA AC. This DNA was then cut and ligated between the Nhel and BamHI of pZL411.
Caulobacter species could receive and maintain replicating plasmids.

FWC CB15 Cori plasmids showed AR in three FWC but not in two MCS species (Table 1). The other FWC17, FWC18 and FWC42 ori plasmids demonstrated AR in the CB15 model. However, except for FWC42 ori plasmid AR in FWC42 cells, the other FWC cells did not support AR of their corresponding ori plasmids. For example, we could not detect FWC18 ori plasmid AR in FWC18 cells. Since FWC42 is most closely related to CB15 (Fig. 1), these results suggest a common physiology that is more permissive for AR. The more distantly related MCS10 and MCS18 ori plasmids failed to show AR in MCS10 and MCS18 cells. However, MCS10 ori plasmids did show vigorous AR in CB15 cells. MCS10 ori plasmid yields were comparable to those of CB15 Cori plasmids, indicating that 5–10 ori plasmids were maintained per cell (data not shown). We failed to detect MCS18 ori plasmid AR (Table 1), so its ‘ori’ designation is based on significant DNA sequence similarities to MCS10 ori, described below.

Table 1. Cell electroporation assays for autonomous ori plasmid replication

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Freshwater caulobacters</th>
<th>Marine caulobacters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CB15</td>
<td>FWC42</td>
</tr>
<tr>
<td>pCB15 Cori</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pFWC42 ori</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pFWC18 ori</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pFWC17 ori</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pMCS10 ori</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>pMCS18 ori</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pRK290</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Fig. 1. Phylogeny of Caulobacter oris. (a) Alpha-proteobacteria oris presumably lie between the conserved hemE and Duf299 coding sequences. The most N-terminal amino acid signatures that are always found in these coding sequences are also shown. (b) Phylogenetic tree created from ori DNA alignments. The corresponding DNA sequences, starting and ending at the N-terminal amino acid signatures, were analysed by multi-way DNA alignment (scoring matrix: mismatch=9, open gap=1, extended gap=9, Clone Manager software version 7). (c) Phylogenetic trees created from DNA alignments of 16S rRNA sequences. The same software and scoring parameters were applied. The scale bars show the percentage of DNA base pair mismatches.

Alignment and phylogenetic grouping of Caulobacter oris

N-terminal amino acid signatures of hemE and Duf299 (Fig. 1a) were used for alignment. This strategy is important, because it is difficult to unambiguously identify the start codons actually used in vivo, so these alignments prevent the arbitrary assignment of the ori ends. More importantly, DNA required for CB15 Cori replication overlaps the start of the hemE coding DNA (Marczynski et al., 1995), so replication sequences might be missed if only the non-coding sequences were considered.

A phylogenetic tree of ori DNA, aligned between conserved N-terminal amino acid signatures (Fig. 1a), is presented in Fig. 1(b). In addition to those sequences isolated by our PCR strategy, we also included related sequences [Caulobacter sp. K31 (Cb sp. K31) and Oceanicaulis alexandrii (O alex)] recently made available by genomic sequencing projects. The five FWC oris clearly group separately from the three MCS ones, and the FWC oris present a tighter phylogenetic grouping (Fig. 1b). This tree is consistent with taxonomic studies (Abraham, et al., 1999; Stahl et al., 1992). As a further control, we also analysed the corresponding 16S rRNA sequences with the same phylogenetic algorithm (Fig. 1c). The ori DNA sequences are significantly less constrained than the 16S rRNA DNA sequences. However, despite a roughly 100-fold difference
in the percentages of base pair changes (note the different scale bars), both trees show the same groupings and they both show exactly the same branch topologies. This congruence argues that both the rapidly changing ori and the slowly changing 16S rRNA DNA sequences have experienced the same historic patterns of evolution.

**DnaA boxes in *Caulobacter oris***

An ori is often presumed from detection of a clustering of DnaA boxes that resemble the *E. coli* TTATNCACA consensus (Mott & Berger, 2007; Zakrzewska-Czerwinska et al., 2007). CB15 ori is unusual, since it has only one similar DnaA box (a 1 bp mismatch, TGATNCACA) near the 5’ end of Duf299 (Fig. 2a). Genetic experiments have demonstrated that CB15 chromosome replication absolutely requires DnaA (Gorbatyuk & Marczynski, 2001), and site-directed mutations have shown that this DnaA box is absolutely required for CB15 *C. oris* plasmid replication (Marczynski & Shapiro, 1992). DnaA is also a global transcription regulator (Zakrzewska-Czerwinska et al., 2007). A whole-genome analysis of DnaA-regulated transcription has suggested that CB15 uses a degenerate *E. coli* consensus (HBVTNMCA) as a DnaA box (Hottes et al., 2005). According to these relaxed criteria, CB15 *C. oris* still contains only two DnaA boxes (Fig. 2a). Most significantly, both CB15 DnaA boxes are conserved among all five FWC *oris* (consensus DnaA boxes 1 and 2). However, CB15 is exceptional, since the other FWC *oris* have one or two additional DnaA boxes (Fig. 2a).

By these criteria, the MCS oris have four to five DnaA boxes (Fig. 2b) and three of these DnaA boxes align (consensus DnaA boxes 1, 2 and 3 in Fig. 2b), implying homology (shared DNA sequences by common descent). However, comparing FWC and MCS oris shows that only consensus DnaA box 2 (near the 5’ end of Duf299) aligns without major distortions. Therefore, DnaA boxes, in both their numbers and positions, are evolutionarily very flexible.

**CtrA binding sites in *Caulobacter oris***

CB15 *C. oris* has five CtrA binding sites (Siam & Marczynski, 2000), labelled ‘a–e’ (Fig. 2a). CB15 sites ‘b–e’ perfectly match the TTAA-N7-TTAA consensus, while site ‘a’ is a single base pair mismatch. Although CtrA has been proposed to block chromosome replication in swarmer cells (Quon et al., 1998), the necessity and the significance of all five binding sites remain unclear. CtrA protein phosphorylation stimulates binding to 7 bp-spaced TTAA motifs, and to adjacent sites ‘a’ and ‘b’ that control a strong cell-cycle-regulated transcription promoter (Siam & Marczynski, 2000). This transcription may stimulate replication, and it does not significantly contribute to HemE protein synthesis (Marczynski et al., 1995). Site ‘c’ overlaps an integration host factor (IHF) binding site (Siam et al., 2003), and the above sequence analysis showed that it also overlaps a consensus DnaA box 1 (Fig. 2a). Site-directed mutations in the right TTAA of site ‘c’ abolish *C. oris* plasmid replication if they also change this DnaA box (Siam et al., 2003). However, mutations in the left TTAA of site ‘c’ only reduce *C. oris* plasmid replication. Presumably these mutations are less severe because they change the IHF binding site (described below) but not the DnaA box. These published results combined with our present sequence analysis suggest that CtrA interacts with both IHF and DnaA proteins at site ‘c’. A comparable hypothesis is not available for site ‘d’, but published site-directed mutations have yielded increased *C. oris* plasmid copy numbers (Quon et al., 1998). That CB15 CtrA site ‘e’ is 4 bp adjacent to an essential DnaA box (Marczynski & Shapiro, 1992) suggests yet another interaction with DnaA.

Comparing CB15 *C. oris* with the other FWC *oris* reveals that CtrA sites ‘b’, ‘d’ and ‘e’ are the most conserved (Fig. 2a). These are also the three strongest *in vitro* CtrA binding sites (Siam & Marczynski, 2000). These observations support several hypotheses. First, that CtrA binding at site ‘b’ is probably sufficient (without cooperative binding to ‘a’) for transcription regulation and its influence on replication. Second, that site ‘d’ is important for an as yet unidentified interaction that limits replication. Third, that CtrA binding at site ‘e’ is important, perhaps for its association with DnaA. CtrA binding site ‘c’ is less conserved, since the left-hand side TTAA motif is missing in the relatively distant FWC *Cb* sp. K31. However the right-hand side TTAA motif plus its overlapping DnaA box are conserved. Interestingly, *Cb* sp. K31 has a third DnaA box only 7 bp away from CtrA site ‘b’, perhaps to compensate for this site ‘c’ variation. Another TTAA motif that is completely conserved in FWC *oris* lies over the ‘weak’ transcription promoter (−10) site of *hemE*. This second promoter is not under cell cycle control (Marczynski et al., 1995), and CtrA protein binding was not detected to this TTAA motif *in vitro* (Siam & Marczynski, 2000). However, this motif is required for CB15 *C. oris* plasmid replication (Marczynski et al., 1995), and its conservation indicates that its role in replication needs further analysis.

In contrast to FWC *oris*, the three MCS oris, despite being relatively AT rich, have only four TTAA motifs: MCS10 and MCS18 *oris* each have only one TTAA motif, and the *O. alex* ori has two. However, these resemble established CtrA binding sites because they are paired with single-base-pair-mismatch TTAA motifs and the 7 bp spacing needed for strong CtrA binding (Ouimet & Marczynski, 2000; Siam & Marczynski, 2000). Similar CtrA binding sites are seen at CB15 *C. oris* site ‘a’ and at many other CtrA-regulated promoters (Wu et al., 1998; Ouimet & Marczynski, 2000). Interestingly, the MCS10 and MCS18 TTAA motifs overlap a near consensus DnaA box (TTATTAACA), in which the fifth position N is T and the sixth position (permissible C or A) is A. These CtrA and DnaA box overlapping sequences also lie within the most conserved DNA region 5’ to Duf299, and so they became the focus of our analysis, described further below.
IHF binding sites in *Caulobacter oris*

The IHF protein binds *E. coli* oriC selectively at the start of chromosome replication (Mott & Berger, 2007; Zakrzeska-Czerwinska *et al.*, 2007). The IHF protein also binds in vitro to CB15 Cori over the central CtrA site ‘c’, and we proposed that CtrA in swarmer cells excludes IHF while the absence of CtrA in stalked cells allows IHF

**Fig. 2.** Schematic *Caulobacter oris*, emphasizing DnaA boxes and CtrA binding site motifs. (a) FWC oris, individually and as a proposed consensus. (b) MCS oris, individually and as a proposed consensus. This figure is based on DNA sequence alignments presented in Supplementary Figs S1 and S2. The filled arrows indicate a perfect or a 1 bp mismatch to the *E. coli* DnaA box consensus (TTATNCACA). The striped arrows indicate the degenerate DnaA boxes (HNVTNMACA) that are used by CB15, see text (H, not G; V, not T; M, A or C). The filled circles indicate perfect CtrA half-site (TTAA) motifs. Since a 7 bp spacing is required for a strong CtrA binding site, the open circles mark a single base pair mismatch to a TTAA motif only when it is paired with a 7 bp-spaced TTAA motif. Conspicuously AT-rich DNA, an exact purine track GGGGAAAA and the flanking coding sequences are also shown approximately to scale.
binding to Cori at the start of chromosome replication (Siam et al., 2003). The IHF consensus (TAA-N5-TGWTW) lies under the IHF protein footprint in CB15 Cori. This predictive consensus is present in closely related FWC ori's, and a second consensus (but pointing in the opposite direction) also overlaps CtrA site 'b' in the FWC42 ori (Fig. 3). However, this IHF sequence is absent in the Cb sp. K31 ori and it is absent in two MCS ori's. Interestingly, this IHF consensus is present in the centre of MCS18 ori, but it is not homologous to the Cori IHF site, because it points in the opposite direction. We also scanned for a more degenerate IHF consensus (WAWCA-N5-WTR) that is adjacent to the first CB15 IHF site but pointing in the opposite direction (Fig. 3). This IHF consensus is also present in FWC42 ori and FWC18 ori, but its significance is unclear, since it does not position under the IHF footprint in CB15 Cori (Siam et al., 2003). However, even when scanning with this more degenerate IHF consensus, there are no matches in the Cb sp. K31, O alex or MCS10 ori. Presumably, these three bacteria use a substitute for IHF which bends DNA and thereby facilitates other protein interactions (Goodman et al., 1999). This analysis also argues against the general validity of our regulatory hypothesis (Siam et al., 2003). If CtrA does regulate IHF binding, then this mechanism operates only in closely related FWC ori's.

DNA methylation sites in Caulobacter oris

E. coli Dam regulates chromosome replication, and Dam (GATC) methylation sites are conspicuously clustered inside ori's from E. coli and related species (Smith et al., 1985). Caulobacters lack Dam, but instead have the CcrM enzyme that methylates the A positioned inside the GANTC sites (Reisenauer et al., 1999). The CB15 ccrM gene is essential, and GANTC sites participate in the cell cycle transcription regulation of the ccrM and ctrA genes (Stephens et al., 1996; Reisenauer et al., 1999). If CcrM regulates chromosome replication, then GANTC sites should be concentrated inside ori's and their positions should be conserved.

CB15 Cori contains five GANTC sites (lightning bolts in Fig. 3). This is more than the two sites expected at random. The GANTC site marked by the unfilled lightning bolt should be excluded from analysis, because it lies inside a conserved Duf299 amino acid signature. Accordingly, this Duf299 GANTC site is present in all Caulobacter oris (Fig. 3). The four remaining CB15 ori GANTC sites are also found at the corresponding positions in FWC17, FWC18 and FWC42 ori's (Fig. 3). However, the FWC Cb sp. K31 ori contains only one GANTC, whose position does not correspond to any of the four CB15 sites. Likewise, the O alex and MCS10 ori each contain only two non-

![Fig. 3. Schematic FWC and MCS ori's, showing the IHF binding site motifs and the CcrM (GANTC) DNA methylation sites. This figure is based on DNA sequence alignments presented in Supplementary Figs S1 and S2. The filled triangles indicate the proposed TAANNNTGWTW (W, A or T) IHF binding site motifs. In FWC42 this motif overlaps CtrA site 'b' but points in the opposite direction. The striped triangles indicate more degenerate WAWCA-N5-WTR (R, A or G) IHF binding site motifs. In CB15 this motif points in the opposite direction. The filled lightning bolts mark the CcrM DNA methylation sites (GANTC). The open lightning bolts mark a GANTC that overlaps the SDXXXXT amino acid signature (Fig. 1).](image-url)
corresponding GANTC sites, and the MCS18 ori lacks GANTC sites. These observations indicate either that CcrM methylation is not important for replication or that its importance is limited to FWC. We favour the latter hypothesis, because the precise conservation of the four GANTC sites among CB15, FWC17, FWC18 and FWC42 oris is in fact statistically significant (P<3 × 10⁻³). Also, studies of the ccrM and of the ctra transcription promoters indicate that even a single GANTC can provide an input into cell-cycle control (Stephens et al., 1995; Reisenauer et al., 1999).

**Ctra protein shows selective and high-affinity binding to MCS10 ori**

We argued above (Fig. 2b) that MCS10 and MCS18 oris contain one important Ctra binding site. We therefore performed DNase I protection 'footprint' experiments to examine C. crescentus Ctra protein binding to MCS10 ori. The Ctra footprint in MCS10 ori is very similar in both size and intensity to the footprints in Cori, and the overlapping MfeI cut site permitted precise footprint alignment with the MCS10 ori DNA sequences (Fig. 4). As shown in Fig. 6(b), the protected region is 27 bp long and lies directly over the TTAA-N7-TAAA motif. To accurately measure the binding affinity of Ctra for this MCS10 ori site, increasing concentrations of Ctra were added to footprint reactions and the degree of protection was measured. The Ctra protein concentrations that gave half maximal DNase I protection were identical for the MCS10 ori DNA sequences (Fig. 4). As shown in Fig. 6(b), the protected region is 27 bp long and lies directly over the TTAA-N7-TAAA motif. To accurately measure the binding affinity of Ctra for this MCS10 ori site, increasing concentrations of Ctra were added to footprint reactions and the degree of protection was measured. The Ctra protein concentrations that gave half maximal DNase I protection were identical for the MCS10-ori DNA sequences (Fig. 4). Therefore, although MCS10 ori has only one Ctra binding site, this has a similar binding affinity to that of the strongest Ctra binding site in Cori (Siam & Marczynski, 2000).

**DnaA protein binding in MCS10 ori**

We also argued above (Fig. 2b) that MCS10 and MCS18 oris contain a DnaA box that overlaps the Ctra binding site. We therefore also performed footprint experiments with the C. crescentus DnaA protein and the same [³²P] end-labelled MCS10 ori DNA, which demonstrated strong Ctra binding. DnaA footprints are problematic, because long stretches of DNA, and not just the DnaA box, are affected by the DnaA protein (Fuller et al., 1984). A recent report argues that the DnaA protein forms a helix, around which the DNA is wrapped (Mott & Berger, 2007). Consequently, when we added DnaA to our footprint experiments, a long stretch of periodic DNase I-hypersensitive cut sites appeared (Fig. 5). Despite this complication, DnaA protects only one site within a ~200 bp span near the [³²P] labelled end of MCS10 ori. The MfeI cut site permitted precise alignment with the MCS10 ori DNA. As shown in Fig. 6(b), this DnaA-protected region is 16 bp long and lies directly over the DnaA box (TTATTAACA). This DnaA footprint also lies completely inside the longer Ctra footprint, suggesting that the two proteins cannot bind simultaneously to this overlapping site.

**Mutations at the overlapping Ctra and DnaA binding sites of MCS10 ori**

We tested the biological relevance of binding sites by site-directed mutations and AR assays. As described above in Table 1, the wild-type MCS10 ori plasmids showed vigorous AR in CB15. However, when the same plasmid received 5 bp changes inside the DnaA box and the left TTAA Ctra motif (MCS10-1 ori plasmid, Fig. 6a), these mutations completely abolished AR. This result indicates that DnaA, Ctra or both are required for AR. To distinguish these alternatives, we directed mutations selectively to the DnaA box (MCS10-2 ori plasmid, Fig. 6a), and DnaA binding site, the right TTAA motif (MCS10-3, Fig. 6a). Both left and right TTAA motifs are required for strong Ctra binding (Siam & Marczynski, 2000). The MCS10-2 ori plasmid showed poor AR, as judged by tiny colonies (Supplementary Fig. S3) and negligible plasmid yields (Supplementary Fig. S4). Conversely, the MCS10-3 ori
plasmid showed enhanced AR, as judged by larger colonies (Supplementary Fig. S3) and higher plasmid yields (Supplementary Fig. S4) compared with the wild-type MCS10 ori plasmid. These results demonstrate that while DnaA is essential for replication, the lack of a strong CtrA binding site increases replication, which we interpret as a release from repression.

DISCUSSION

Our previously published results indicated that alphaproteobacteria oris are located between hemE and Duf299 (Brassinga et al., 2001). This argument was based on similarities between C. crescentus (hemE C0001/Duf299) and Rickettsia prowazekii (hemE RP001/Duf299) DNA sequences. This strategy recently identified the Sinorhizobium meliloti ori (Sibley et al., 2006). A presumptive ori should be supported by functional criteria such as AR. However, this is not always possible, because an ori plasmid must insinuate into a regulatory system that rigorously restricts chromosome replication. Tolerance of extra ori DNA is variable. For example, E. coli oriC plasmids are tolerated by E. coli (Leonard & Helmstetter, 1986), while the B. subtilis oriC plasmids are not (Moriya et al., 1992). Therefore, while AR strongly supports an ori hypothesis, its absence is not a strong counter-argument.

The data in Table 1 indicate that CB15 cells are exceptionally tolerant, and that CB15Cori plasmids are also well tolerated by other FWC cells. In contrast, MCS18 does not accept ori plasmids, nor are MCS18 ori plasmids tolerated by other caulobacters. We interpret our results in terms of ori tolerance and not DNA exclusion, because all of the caulobacters in Table 1 were efficiently transformed by the broad-host-range plasmid pRK290. This control indicates that the negative results were not due to transformation barriers such as restriction endonucleases. Tolerance is therefore ori replicon-specific and could for example result from limiting replication proteins. Despite strong sequence similarities between MCS10 and MCS18

![Fig. 5. Footprint assays reveal selective DnaA protein binding to the MCS10 ori. As described in Methods, the 5' [32P] end-labelled MCS10 ori DNA fragment was incubated for 5 min at room temperature with 2, 5, 10 or 20 nM concentrations of His-tagged C. crescentus DnaA protein (inside lanes) or without protein (outside lanes). These binding reactions were next treated with DNase I and processed as in Fig. 4. The DnaA footprint (protected zone) is indicated by the open bar starting at the MfeI cut site (see also Fig. 6b). The positions of enhanced DNase I cleavage, in the presence of DnaA, are indicated by arrows.](image)

![Fig. 6. Consequences of directed mutations at the overlapping CtrA and DnaA protein binding sites of MCS10 ori. (a) Wild-type (Wt) DNA, directed mutations (MCS10-1, 10-2 and 10-3) and corresponding MCS10 ori plasmid autonomous replication results. +, Standard colonies; −, no colonies, following MC10 ori plasmid electroporation of GM1609 (CB15) cells (Table 1). ±, Tiny colonies; ++, significantly larger colonies under the same plating conditions. Supplementary Figs S3–S6 further substantiate these replication results. (b) Summary of CtrA and DnaA footprint results (from Figs 4 and 5) superimposed over MCS10 ori DNA sequences. The DnaA box (TTATTAACA) and the CtrA binding motif (TTAA-N7-TAAA) align with the corresponding footprints. The bottom strand of the indicated DNA sequence was 5' [32P] end-labelled.](image)
ori (Fig. 2b), especially around the essential DnaA box and CtrA binding site (Figs 4, 5, and 6), without positive AR we could not confirm that MCS18 ori initiates replication in vivo. We believe that ori plasmid tolerance is mediated at least partly by CtrA. Comparing Table 1 with Fig. 2 shows that tolerance roughly correlates with the number of CtrA binding sites. Interestingly, when CtrA binding sites are mutated on the CB15 chromosome, these strains become less tolerant of Cori plasmids (D. P. Bastedo & G. T. Marczynski, unpublished results).

**CtrA probably regulates DnaA activity at Caulobacter ori**

We focused our studies on DnaA and CtrA because both proteins are essential receivers of regulatory signals. Of the 16 DnaA boxes observed in five FWC ori (Fig. 2a), seven DnaA boxes overlap CtrA binding sites while the remaining nine are all positioned adjacent to CtrA binding sites. This pattern indicates that DnaA activity is a major target of CtrA regulation. In contrast, MCS ori have single CtrA binding sites (Fig. 2b). We chose MCS10 ori for our studies because it presented a simple system to study DnaA and CtrA interactions. Remarkably, the DnaA box and CtrA overlap sequence (TTATTTAACA) that we studied (Fig. 6) is absolutely unique in the whole 3,368,780 bp MCS10 genome. Therefore, CtrA binding at this site provides an exceptionally selective mechanism for regulating DnaA. A similar DnaA box overlap also occurs in freshwater caulobacters at CtrA site ‘c’ (Fig. 2a). *In vitro* binding experiments show that CtrA displaces DnaA protein from DnaA box 1 of CB15 Cori (M.-C. Ouimet & G. T. Marczynski, unpublished results). Interestingly, FWC42 ori appears to use this overlap mechanism in three places, at CtrA sites ‘b’, ‘c’ and ‘d’.

**Independent evolution of CtrA binding sites in FWC and MCS ori**

Before these studies, we believed that CtrA binding sites would be present (conserved) due to descent from a common ancestor. Accordingly, we initially believed that the single CtrA binding site in the MCS10 and MCS18 ori corresponded directly to CtrA binding site ‘e’ of the FWC ori (Fig. 2a). However, we now believe that these CtrA binding sites evolved independently in FWC and MCS ori. Three arguments support this surprising hypothesis. First, while we observe that TTAA mutations increase MCS10 ori plasmid replication (Fig. 6a), the corresponding TTAA mutations in site ‘e’ of CB15 Cori plasmids have been shown to decrease replication (Marczynski et al., 1995). These results suggest that CtrA has opposite effects on DnaA activity in FWC and MCS ori. A second argument notes that site ‘e’ in FWC ori is 4 bp to the left of the DnaA box, while the CtrA binding site in MCS10 and MCS18 overlaps the DnaA box from the opposite side. It is hard to imagine simple common descent pathways of nucleotide changes that first form a CtrA binding site and then allow it to migrate to the other side of the DnaA box. Independent formation of CtrA sites is a simpler evolutionary pathway. A third argument notes that the O alex ori lacks the corresponding CtrA binding site that overlaps consensus DnaA box 2 (Fig. 2b). This homologous O alex DnaA box is a perfectly matching TTATTCACA consensus, but it completely lacks the extra CtrA binding site sequences seen in MCS10 and MCS18 (Supplementary Fig. S2c). Instead, the O alex ori has a centrally positioned CtrA binding site motif. This observation is particularly striking because O alex has a phylogenetic position between MCS10 and MCS18 and the FWCs (Fig. 1).

**Flexible and convergent ori evolution**

Our results imply an unexpectedly variable usage of DnaA, CtrA, IHF and CcrM (DNA methylation) for replication control. DnaA (Gorbatyuk & Marczynski, 2005), CtrA (Domian et al., 1997), IHF (Muir & Gober, 2005) and CcrM (Reisenauer et al., 1999) all have documented cell-cycle variations in *C. crescentus*. Therefore, each protein could in principle contribute to replication control. CtrA is the most prominent cell-cycle regulator, and CtrA is itself regulated by an exceptionally complex network, whose mechanisms include transcription regulators, DNA methylation, membrane-spanning kinases, phosphorelay proteins, adaptor proteins, subcellular positioning, and localized proteolysis (Bowers et al., 2008). As a two-component response regulator, CtrA is activated by a phosphorelay and CtrA, and thereby controls the transcription of ~25% of the cell-cycle-regulated genes in *C. crescentus* (Laub et al., 2002). CtrA is presumably a ‘global regulator’, because its binding site is particularly well suited for recruitment to promoters. In a previous publication we proposed a means of recruitment (binding site creation) based on the similarities between the CtrA TTAA motifs and the −35 TTGACA promoter elements (Ouimet & Marczynski, 2000). Following a G to A mutation, the second 7 bp-spaced TTAA motif is easily created because it need not be a perfect match (Ouimet & Marczynski, 2000). We also note the similarity between CtrA TTAA motifs and DnaA boxes (TTATNMCA, M=C or A). In evolutionary terms, the −35 promoter elements and DnaA boxes can be called ‘pre-adapted’ sequences, or ones capable of acquiring a ‘functional shift’, because they already resemble CtrA binding sites. The same concept may explain the IHF binding site overlap with CtrA sites ‘b’ and ‘c’ (Fig. 3). Similarly, a functional-shift concept suggests that CtrA binding sites could easily evolve independently in FWC and MCS ori.

These speculations are also well supported by published studies. A comparison of CtrA binding sites in *C. crescentus* and Brucella abortus suggests that the CtrA regulon is significantly different in this alpha-proteobacterium (Bellefontaine et al., 2002). In particular, the proposed Brucella abortus ori (hemE Dul299) lacks CtrA binding sites. However, we have previously demonstrated that the
ACKNOWLEDGEMENTS

We thank John Smit, University of British Columbia, Vancouver, BC, for supplying FWC and MCS strains. We also thank Elliot Crooke, Georgetown University Medical Center, Washington, DC, for pZL411, Boris Gorbatyuk, McGill University, for the His-tag C. crescentus DnaA protein. We thank John Smit, University of British Columbia, Vancouver, BC, for supplying FWC and MCS strains. We also thank Elliot Crooke, Georgetown University Medical Center, Washington, DC, for pZL411, Boris Gorbatyuk, McGill University, for the His-tag C. crescentus DnaA protein.

REFERENCES


Edited by: D. W. Ussery