Plasmids derived from Gifsy-1/Gifsy-2, lambdoid prophages contributing to the virulence of Salmonella enterica serovar Typhimurium: implications for the evolution of replication initiation proteins of lambdoid phages and enterobacteria

Bartosz Słomiński, Joanna Całkiewicz, Piotr Golec, Grzegorz Węgrzyn and Borys Wróbel

1Department of Molecular Biology, University of Gdańsk, Kładka 24, 80-822 Gdańsk, Poland
2Department of Genetics and Marine Biotechnology, Institute of Oceanology, Polish Academy of Sciences, Powstańców Warszawy 55, 81-712 Sopot, Poland

Gifsy-1 and Gifsy-2 are lambdoid prophages which contribute to the virulence of Salmonella enterica serovar Typhimurium. The nucleotide sequence of the replication region of both prophages is identical, and similar in organization to the replication region of bacteriophage λ. To investigate the replication of the Gifsy phages and the relationship between Gifsy and host chromosome replication, a plasmid which contained all the genes and regulatory sequences required for autonomous replication in bacterial cells was constructed. This plasmid, pGifsy, was stably maintained in Escherichia coli cells. The helicase loader of the Gifsy phages is very similar to the DnaC protein of the host, a feature characteristic of a large group of prophages common in the sequenced genomes of pathogenic enterobacteria. This DnaC-like protein showed no similarity to the helicase loader of bacteriophage λ and closely related phages. Interestingly, unlike plasmids derived from bacteriophage λ (λ plasmids), pGifsy did not require a gene encoding the putative helicase loader for replication, although deletion of this gene resulted in a decrease in plasmid copy number. Under these conditions, it was shown that the plasmid utilized the helicase loader coded by the host. On the other hand, the viral protein could not substitute for DnaC in bacterial chromosome replication. The results of the current study support the hypothesis that the enterobacterial helicase loader is of viral origin. This hypothesis explains why the gene for DnaC, the protein central to both replication initiation and replication restart in E. coli, is present in the genomes of Escherichia, Shigella, Salmonella and Buchnera, but not in the genomes of related enterobacteria.

INTRODUCTION

Lambdoid bacteriophages have been basic models in molecular biology studies for many years (Friedman & Court, 2001; Węgrzyn & Węgrzyn, 2005). Recent observations that many toxins produced by pathogenic bacteria are encoded by prophages related to λ (for recent reviews, see Brussow et al., 2004; Herold et al., 2004) have rekindled the interest in this group of viruses.

The properties of the replication–regulation region of lambdoid prophages can be conveniently investigated using plasmids derived from these viruses. Indeed, plasmids derived from bacteriophage λ have been found to be very useful models in studies of replication in general (for a review, see Węgrzyn & Węgrzyn, 2002). These plasmids are constructed by ligation of an antibiotic-resistance gene with a fragment of phage DNA containing all genes and regulatory sequences required for autonomous replication in bacterial cells. The replication–regulation region of bacteriophage λ consists of the pr promoter and cro, cII, O and P genes. The function of cro and cII is regulatory: cro is a repressor of the pr promoter (Murotsu & Matsubara, 1980; Ptashne, 1992), while cII is involved in the regulation of the lysogenic development of λ, and has no importance for the replication and maintenance of λ plasmids.

Replication of λ plasmid DNA is initiated at the oriλ region, located in the middle of the O gene (Fig. 1). This gene encodes the replication initiator protein, which binds to the replication origin, forming the nucleoprotein structure called the ‘O-some’. The host-encoded DnaB helicase is delivered to the O-some by another λ replication protein, the P gene product. The oriλ/O/P/DnaB structure,
called the ‘pre-primosome’, is stable but inactive in promoting DNA replication, as strong interactions between P and DnaB proteins prevent helicase activity of the latter component (unlike an analogous DnaC/DnaB complex that operates during initiation of *Escherichia coli* chromosome replication at oriC, which is quite unstable). Action of the heat-shock proteins DnaK, DnaJ and GrpE is necessary to liberate DnaB from P-mediated inhibition (for a review, see Taylor & Węgrzyń, 1995), although the P protein seems to still be present in the complex, and DnaK also remains bound to it (Potrykus et al., 2002).

The process of chaperone-dependent pre-primosome remodelling is coupled with transcriptional activation of *oriλ* (transcription proceeding in the replication origin region; for a review see Taylor & Węgrzyń, 1995). Transcriptional activation of the *oriλ* is necessary for efficient initiation of λ DNA replication *in vivo*, even if all the replication proteins are provided. It seems that changes in DNA topology caused by movement of RNA polymerase during transcription may play a crucial role in stimulation of replication initiation. The *pλ* promoter is a natural start site of transcription that produces mRNA for synthesis of λ replication proteins (O and P), and acts to activate *oriλ*. The final step in the initiation of λ DNA replication is binding of DNA polymerase III holoenzyme and accessory replication proteins (DNA gyrase, single-stranded DNA-binding protein, and other proteins), encoded by the host, to the *oriλ* region (for a review see Węgrzyń & Węgrzyń, 2002).

Lambdoid prophages Gifsy-1 and Gifsy-2 both contribute to the virulence of *Salmonella enterica* serovar Typhimurium, in which they were discovered (Figueroa-Bossi & Bossi, 1999). Various virulence factors are associated with Gifsy phages (Figueroa-Bossi et al., 2001; Lawley et al., 2006), e.g. *gipA* (Gifsy-1; Stanley et al., 2000), *gggB* (Gifsy-1; Coombes et al., 2005), *gtgE* (Gifsy-2; Ho et al., 2002), *gggB*sseI (Gifsy-2; Miao & Miller, 2000; Worley et al., 2000) and *sodCI* (Gifsy-2; De Groote et al., 1997; Farrant et al., 1997; Fang et al., 1999; Sansone et al., 2002; Krishnakumar et al., 2004). Gifsy prophages contain regions with an organization similar to that of the bacteriophage λ replication region. In this work, we constructed a series of plasmids containing the replication region of Gifsy prophages, analogous to λ plasmids, capable of autonomous replication in *E. coli* cells, in order to study the relationship between Gifsy and host chromosome replication.

**METHODS**

**Bacterial strains and plasmids.** pGifsy was constructed by PCR amplification using the host (*S. enterica* serovar Typhimurium LT2) genomic DNA as a template, and primers Gifsy-up 5′-CGG CCG AAT TCT AGA TAT CTC GAG AAC TTG CTG GAT ACG C-3′ and Gifsy-down 5′-CTG GAT CAT CCG CTA CCG AGG AGC GCC TAT TGG ACA ATC CCG AGG TG-3′. The following thermal profile was used: 94 °C for 5 min, then 30 cycles of 94 °C for 30 s, 64 °C for 30 s, 72 °C for 1 min and final extension at 72 °C for 5 min. Prophages Gifsy-1 and Gifsy-2 are identical at the nucleotide level in the amplified region, which corresponds to nt 2 767 129–2 769 658 (Gifsy-1 prophage) and 1 105 429–1 107 958 (Gifsy-2 prophage) in the *S. enterica* serovar Typhimurium LT2 genome (GenBank accession no. NC_003197; McClelland et al., 2001). The PCR fragment, after digestion with XbaI and BamHI, was ligated to a fragment carrying the tetracycline-resistance gene obtained by digesting pACYC184 with XbaI and BclI. pGifsyNoLoader was derived from pGifsy by digestion with restriction endonucleases *SacI* and *BamHI*.

Plasmids derived from Gifsy prophages
each primer pair: Ecoli-dnaC-up 5'-GCC GAA GGA TCC CAT TAC ATG AAA AAC GTT GCC GAC C-3', Ecoli-dnaC-dn 5'-GGG CCG TTA GCC TAG GAT GTT AAT ACT CTG TTA TTA CCC-3' (94 °C for 5 min, 30 cycles at 94 °C for 30 s, 62 °C for 30 s, 72 °C for 1 min, and final extension at 72 °C for 5 min); Rac-up 5'-GGG GCG CGG GGT GAT CAA GGA TCC CAT TAC ATG AAA AAT ATT GG-3', Rac-dn 5'-GCC AAG CTA GCT TGG ATG CCT ACC AAT AAA AAA CGC CCG GTC AAG ATC TTA CTT CGC AAT TCC'TGG CTG-3' (94 °C for 5 min, five cycles of 94 °C for 30 s, 48 °C for 30 s, 72 °C for 1 min, 25 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min 50 s, and final extension at 72 °C for 5 min); GifsyDnaC-up 5'-GGG GAT CAA GGA TCC CAT TAC ATG AAA AAT ATT GG-3', GifsyDnaC-dn 5'-CTC GAG CTA GCC TAG GAT GTG GCC AAG AAT CAG CCG TT-3' (94 °C for 5 min, five cycles at 94 °C for 30 s, 48 °C for 30 s, 72 °C for 1 min, 25 cycles at 94 °C for 30 s, 59 °C for 30 s, 72 °C for 1 min 30 s, and final extension at 72 °C for 5 min). The structure of all plasmids was confirmed by restriction analysis and sequencing.

The plasmids carrying the bacterial dnaC (pLNC) and dnaB (pLNB) genes under the control of the Bam promoter have been described by Allen & Kornberg (1991). DnaC or DnaB overproduction in bacteria carrying these plasmids was induced by addition of 0.1% arabinose.

pCB104 (Boyd & Sherratt, 1995) is a standard λ plasmid bearing chloramphenicol resistance. pKB2π is a λ plasmid bearing a G410T (Arg137Gly) mutation in the P gene (Kur et al., 1987; Wegrzyn et al., 1996).

Derivatives of E. coli K-12 strains were used in all the experiments. MG1655 (wild-type strain; Jensen, 1993) was used to determine plasmid copy number and in the experiments on plasmid co-maintenance. XXL-Blue (F: Tn10 proxB B Δ lacP Δ(lacZ)M15 recA1 endA1 gyrA96 (Nal1) thi hsdR17 rK mK supE44 relA1 lac; Bullock et al., 1987) was used in overexpression experiments to allow for better control of the Bam promoter. The effects of the dnaC1(ts) mutation on the replication of Gifsy plasmids were measured in D765519 (leuB6 thyA47 deoC3 rpsL) and its otherwise isogenic dnaC1(ts) derivative PC15030 (Kędzierska et al., 2003). The effects of the dnaA66(ts) mutation on plasmid replication have been investigated using WAM106 (Thomas & Glass, 1991) and its otherwise isogenic dnaA66(ts) derivative (Slomińska et al., 2003). The dependence of pGifsy on molecular chaperones has been tested using W3101 (galE sup+, lacIΔ) (Lipińska et al., 1989), and its otherwise isogenic derivatives bearing dnaK756 (Georgopoulos & Herskovitz, 1971; Georgopoulos, 1977; strain W3101K) or dnaJ (Yochem et al., 1978; strain W3101J) mutations. In the experiments on plasmid replication in amino acid-starved cells, the strains CP78 [leu arg thr his thi (λ30)]; Fiil & Friesen, 1968; Wroβel et al., 1998b] and CP79 (same as CP78, but relA2; Fiil & Friesen, 1968; Wroβel et al., 1998b) were used.

Estimation of the efficiency of transformation, plasmid copy number and relative plasmid amount after temperature upshift and in amino acid-starved bacteria. The procedures described by Wegrzyn et al. (1996) were employed. Bacterial growth was monitored by measurement of culture OD575. The plasmid amount relative to bacterial cell mass (determined by OD measurement) was measured densitometrically after isolation from samples taken at indicated time intervals, after temperature upshift in the E. coli dnaC1(ts) and dnaA66(ts) strains, or after the onset of isoleucine starvation in the E. coli relA and relAΔ strains. We report the mean results from three independent experiments. Isoleucine starvation of bacteria growing in a minimal medium was induced by addition of L-valine to a final concentration of 1 ng ml⁻¹, as described previously (Wegrzyn et al., 1991).

**Plasmid compatibility.** Plasmids bearing different antibiotic-resistance genes were introduced into E. coli cells (MG1655). After overnight growth in the presence of all the antibiotics, the bacteria were plated on plates lacking the antibiotic neutralized by one of the plasmids. Then, 100 colonies were tested for growth on plates containing the antibiotic lacking in the previous step.

**RESULTS**

**Analysis of the Gifsy regulation–replication module**

The replication–regulation region of bacteriophage λ consists of the pR promoter and four genes: cro, cII, O and P. Only O and P are necessary for replication, the two shorter genes code for DNA-binding regulatory proteins. The Gifsy-1 genome contains a region with a similar organization: a promoter, two short and two longer ORFs (Fig. 1). Although other symbols (ORF numbers) for these genes have been used during S. enterica serovar Typhimurium LT2 genome sequencing, in this work they are referred to as croGifsy, cIIGifsy (ORF STM2627 in Gifsy-1 region, STM1013 in Gifsy-2), OGifsy (STM2626 and STM1014, respectively) and dnaCGifsy (STM2625 and STM1015; see below).

The homology between the two putative regulatory proteins of Gifsy prophages (79 and 124 aa, respectively) and their counterparts in λ (67 and 94 aa) is not obvious. They do not show significant similarity at the nucleotide or amino acid level when the BLAST2 algorithm is used, and the expected values are calculated based on the size of the non-redundant GenBank database. However, when a more exact method to calculate protein similarity was used (Smith–Waterman algorithm implemented in PRSS3, part of the FASTA package; Pearson, 1996), and the expectancy values were calculated based on shuffling one of the sequences 1000 times, the similarity between the products of cIIGifsy and λ cII was significant (using the BLOSUM65 matrix, the penalty was 12 for gap opening and 2 for gap extension). The putative croGifsy and cIIGifsy products both contain a helix–turn–helix motif (Dodd & Egan, 1987, 1990; algorithm implemented in HELIXTURNHELIX, part of the EMBoss package; Rice et al., 2000), which suggests that they are DNA-binding proteins. Moreover, they show similarity with products of ORFs of other phages or prophages in the family, for example, the product of cII of Salmonella phage L (Schickmaier & Schmieger, 1997). We conclude that the two shorter reading frames in Gifsy prophages code for regulatory proteins.

Replication of λ plasmid DNA is initiated at the oriλ region. This region is located in the middle of the O gene, and contains four repeated iteron sequences. The O protein interacts with these 19 nt repeats with its N-terminal part; the C-terminal part interacts with the second replication protein P, coded by a downstream gene. P in turn interacts with the host DNA helicase DnaB. Analysis of the OGifsy

---

**Microbiology 153**

1886

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Wed, 12 Dec 2018 02:47:42
gene shows that it contains five repeats followed by an AT-rich region, and has a structure very similar to that of oriC (Tsurimoto & Matsubara, 1981a, b, Fig. 1). These repeats (ATCCGCGAAAGCAGGATA) are highly similar to λ iterons (ATCCGCGAAAGCAGGAt), which is expected considering the high similarity of the DNA-binding N-terminal part of Oγifsy to the corresponding protein region in λ O (Wróbel & Węgrzyn, 2002). However, the C-terminal part of the protein resembles the amino acid sequence of the host protein DnaT. This concerns especially the C-terminal 18 aa of these proteins, which show 68.4% identity and 73.7% similarity. We have speculated previously that this C-terminal sequence motif might be involved in the interaction with the helicase loader (Wróbel & Węgrzyn, 2002).

The second replication protein of Gifsy prophages (which we named DnaCGifsy) is similar to DnaC, the host helicase loader whose gene follows dnaT in the host genome (47.6% identity and 67.2% similarity in a global alignment obtained with the Needleman–Wunsch algorithm using default parameter values implemented in program NEEDLE, part of the EMBOSS package; Rice et al., 2000). DnaCGifsy shows no significant similarity to the bacteriophage λ protein P (Wróbel & Węgrzyn, 2002). This suggests that while the first steps of building the Gifsy replication complex probably resemble the assembly of the λ replication complex (binding of the O protein to the origin and O-some formation), the subsequent steps may be different, since Gifsy phages apparently use a helicase loader closely related to that of the host.

**Phylogenetic analysis of bacteriophage DnaC homologues**

In order to investigate the evolutionary relationships between the Gifsy replication module and (i) the replication region of other phages encoding a homologous helicase loader, and (ii) the dnaT–dnaC region in the genomes of enterobacteria, we searched the GenBank non-redundant database and the database of wholly sequenced genomes at the National Center for Biotechnology Information site for homologues of enterobacterial DnaC. Apart from the members of the genera *Escherichia*, *Shigella*, *Salmonella* and *Buchnera* (enterobacterial intracellular symbionts of aphids), no other γ-proteobacterial genome carried non-viral DnaC-like proteins. The only group of bacteria in which helicase loaders have been identified are Gram-positive bacteria related to *Bacillus* (Wróbel & Węgrzyn, 2002). In the preliminary phylogenetic analysis, the sequences of helicase loaders from γ proteobacteria and their phages were clearly separated from the sequences of helicase loaders present in the genomes of Gram-positive bacteria and their phages (not shown), and only the γ-proteobacterial sequences were used in further analysis. The sequences were aligned using ClustalW (the alignment was corrected manually). JTT (Jones et al., 1992), a substitution model optimal under the Akaike Information Criterion (program ProtTest, Abascal et al., 2005), was used in all analyses. Programs PROTDIST and NEIGHBOR (parts of the PHYLIP package; Felsenstein, 2005) were used to construct the neighbour-joining tree presented in Fig. 2. The support for the interior branches was measured using the bootstrap method with SEQBOOT and CONSENSE (parts of PHYLIP), and the weighted least-squares likelihood ratio test using program WeightLESS (Sanjuán & Wróbel, 2005).

The presence of the *dnaT–dnaC* operon was detected only in two enterobacterial lineages: one led to *Escherichia*/*Shigella* and *Salmonella*, and the other to the enterobacterial symbionts of aphids (*Buchnera*) (Fig. 2). The separation between these two lineages has been estimated to have occurred 200 million years ago (Clark et al., 1999; Lerat et al., 2003). It is particularly interesting that there was a lack of *dnaC* in bacteria proposed to be closely related to the free-living branch of enterobacteria (in particular, *Versinia*) or the endosymbiont branch (in particular, *Wigglesworthia* and *Blochmannia*). However, if the hypothesis of the acquisition of the helicase loader from the lambdoid phages is true (Wróbel & Węgrzyn, 2002), it was probably a one-off event. All genomes of free-living enterobacteria that contain a *dnaC*-like gene also contain *dnaT*, whose product is involved in the mechanisms for reloading the replication machinery after replication-fork breakdown (Sandler, 2000; Boonsombat et al., 2006; Heller & Marians, 2005). This is also true for the *Buchnera* genomes, with the exception of *Buchnera* from *Baizongia pistaciae* (which suggests a recent loss rather than independent acquisition in the *Buchnera* lineage). Thus, the evidence indicates that the adaptation of the *dnaT–dnaC* system for host replication occurred once in enterobacteria. It is likely that it involved adaptive changes in other genes encoding the members of both the *oriC* initiation and replication-restart complexes.

The enterobacterial DnaC sequences form a monophyletic cluster with viral helicase loaders from a large group of prophages present in enterobacterial genomes, and are likely to be involved in the virulence of their hosts (Wróbel & Węgrzyn, 2002). Within the cluster of lambdoid helicase loaders, DnaC-like proteins from Gifsy phages formed a group separate from those similar to the helicase loader of the Rac prophage present in the *E. coli* K-12 genome.

Helicase loaders of viruses infecting other groups of bacteria (*Pseudomonas* and *Burkholderia*) are quite distant from this group of DnaC-like sequences. The other more distantly related group of viruses related to phage P27 carry in their replication region three genes which probably code for an iteron-binding protein, a helicase loader and a helicase, in that order. The helicase encoded by these viruses is closely related to the helicases of the lambdoid phages that carry two genes only (for an iteron-binding protein and a helicase; data not shown).

Three amino acids suspected to be involved in the interaction with DnaB (Nakayama et al., 1987; Ludlam et al., 2001; Wróbel & Węgrzyn, 2002) are conserved in the
DnaC-like proteins belonging to the monophyletic cluster mentioned above: Phe/Tyr23, Trp32 and Cys69 (coordinates correspond to the amino acid numbers in the E. coli protein). These residues are not conserved in the sequences of two viruses infecting Pseudomonas and Burkholderia, and only Cys corresponding to Cys69 is conserved in DnaC-like proteins in four phages/prophages related to P27. This suggests that these proteins interact with the viral helicase, not the host protein, although, of course, such a question cannot be decisively answered by sequence analysis alone.

**Construction of plasmids derived from Gifsy phages**

Using the host genomic DNA as a template, the Gifsy-1 replication–regulation region (which is identical to the corresponding Gifsy-2 sequence) was amplified and, after digestion with restriction endonucleases, ligated with a gene conferring resistance to tetracycline (see Methods for details). The construct efficiently transformed E. coli cells and was named pGifsy. We found that pGifsy maintained in E. coli cells in Luria–Bertani medium at 37 °C with aeration, at a mean copy number of ~50–55 in wild-type bacteria (MG1655), which is a value comparable to that observed for λ plasmids in cells cultured under similar conditions (also ~50; Węgrzyn, 1995).

**DnaCgifsy-independent replication of pGifsy**

O and P are the only two genes necessary for replication and maintenance of λ plasmids in E. coli cells (for a review see Węgrzyn & Węgrzyn, 2002). However, as noted above, Gifsy phages use a helicase loader very similar to that of their host. Could the host DnaC be able to deliver the host helicase to the Gifsy origin? To answer this question, we constructed a plasmid devoid of the functional dnaCgifsy gene. We found that this plasmid, pGifsyNoLoader, in which most of the dnaCgifsy gene was deleted, could transform E. coli cells and was stably maintained, though at a copy number lower (~30–35 in wild-type strain MG1655) than that of a plasmid with a functional dnaCgifsy gene (~50–55).

This experiment suggests that the cellular DnaC protein can substitute for DnaCgifsy. On the other hand, the
activity of the viral helicase loader delivered from pQE-GifsyDnaC was not sufficient to suppress the dnaC1(ts) mutation, or to allow the survival of the host at 43 °C, even in the presence of 1 mM IPTG. In contrast, the presence of the E. coli gene on pQE-EcoliDnaC (even without promoter induction) allowed for continued growth of the dnaC1(ts) bacteria in liquid culture after the temperature upshift (data not shown; delivery of DnaC from a plasmid does not allow growth on plates at 43 °C; Allen & Kornberg, 1991).

The replication of pGifsyNoLoader in the dnaC1(ts) strain stopped after temperature upshift (Fig. 3b, c), but this effect could be mitigated in the presence of a plasmid that carries either the co-specific dnaCGifsy gene or a dnaC-like gene from the E. coli prophage Rac (which we called dnaCRac in the present work; Fig. 3b). Overexpression of these genes did not result in more efficient replication of pGifsyNoLoader (Fig. 3c). On the other hand, when the cells carried the E. coli dnaC expression plasmid, the effect on pGifsyNoLoader replication could only be seen when expression of the helicase loader was induced (Fig. 3c). However, when the helicase loaders were delivered from the pQE-based plasmids, the replication of pGifsyNoLoader continued for only about one round (Fig. 3b, c) and then stopped. In the control experiments, at 30 °C, replication of pGifsy was also stopped after transfer of the dnaC1(ts) host to 43 °C (Fig. 4). We investigated whether or not the arrest of pGifsy replication in dnaA46(ts) bacteria was due to the general arrest of plasmid DNA replication after termination of host replication. Such a general arrest was unlikely, since the replication of Gifsy plasmids continued even 200 min after the upshift (Fig. 5), while it stopped in the dnaC1(ts) strain after ~60 min (Fig. 4). Moreover, if such a general arrest occurred, it should have affected other lambdoid plasmids. Although wild-type λ plasmids could not be introduced into the dnaC1(ts) strain, a so-called π derivative (pKB2π) could be introduced (data not shown). This is perhaps because the presence of the wild-type λ P is toxic for these cells. It is possible that wild-type λ P binds the DnaB protein so strongly that the mutated DnaC protein cannot deliver the helicase to the host origin. The π plasmid carries a mutation in the P gene which allows plasmid replication in strains deficient in the Hsp70 chaperone machine: the mutated P protein competes less efficiently with the host DnaC protein for binding to DnaB.

Interestingly, replication of pGifsy was also stopped after transfer of the dnaC1(ts) host to 43 °C (Fig. 4). We investigated whether or not the arrest of pGifsy replication in dnaC1(ts) strains after temperature upshift was due to the general arrest of plasmid DNA replication after termination of host replication. Such a general arrest was unlikely, since the replication of Gifsy plasmids continued after host replication was halted by transferring dnaA46(ts) bacteria to 43 °C (Fig. 5). The effect of the temperature upshift on these bacteria should be similar to that on the dnaC1(ts) strain: inability to initiate chromosomal replication. In the dnaA46(ts) strain, however, pGifsy replication continued even 200 min after the upshift (Fig. 5), while it stopped in the dnaC1(ts) strain after ~60 min (Fig. 4). Moreover, if such a general arrest occurred, it should have affected other lambdoid plasmids. Although wild-type λ plasmids could not be introduced into the dnaC1(ts) strain, a so-called π derivative (pKB2π) could be introduced (data not shown). This is perhaps because the presence of the wild-type λ P is toxic for these cells. It is possible that wild-type λ P binds the DnaB protein so strongly that the mutated DnaC protein cannot deliver the helicase to the host origin. The π plasmid carries a mutation in the P gene which allows plasmid replication in strains deficient in the Hsp70 chaperone machine: the mutated P protein competes less efficiently.
with the host DnaC protein for DnaB binding (Konieczny & Marszalek, 1995). Our experiments showed that although the replication of pGifsy was arrested in dnaC1 (ts) cells after temperature upshift, replication of pKB2p continued (Fig. 4).

The fact that replication of both pGifsyNoLoader and pGifsy stopped at 43°C in the dnaC1(ts) bacteria suggests that DnaCGifsy forms heterocomplexes with the host DnaC, and these heterocomplexes are inactive for helicase loading after temperature upshift. If this is so, when the host DnaC is overexpressed, only a minority of heterocomplexes contain the protein that is inactivated at high temperature. Interference of the DnaC1(ts) protein in the heterocomplexes with the proper loading of the helicase at the Gifsy origin might explain why the suppression of the effects of the dnaC1 mutation on pGifsyNoLoader and pGifsy replication by helicase loaders produced from another plasmid is only transient. This explanation should, however, be treated as highly speculative, since it is not clear how one can explain that replication of pGifsyNoLoader remains transient even when DnaCGifsy is overexpressed. Moreover, if this explanation is true, it does not apply to host replication initiation, since dnaC1(ts) bacteria grow normally when E. coli DnaC is delivered from a plasmid, as mentioned above.

Transient replication of pGifsyNoLoader in cells over-expressing the phage helicase loader might, at least in part, result from the toxicity of overproduction of phage DnaC-like proteins in E. coli cells. It has been shown that overproduction of E. coli DnaC inhibits bacterial growth (Allen & Kornberg, 1991). λ P protein, when over-produced, is also toxic to E. coli, via a mechanism that involves not only the inactivation of the DnaB helicase, but also impairment of the function of DnaA, the initiator of DNA replication from oriC (Datta et al., 2005a, b). We found that overproduction of DnaCGifsy from an over-expression plasmid pQE-GifsyDnaC (Fig. 6) was also toxic to E. coli and so was the overexpression of its homologue, DnaCRac, from pQE-RacDnaC; data not shown). This result supports the hypothesis that in bacteria over-producing DnaCGifsy, the host helicase is trapped in DnaCGifsy-DnaB and possibly DnaCGifsy-DnaCEcoli-DnaB complexes, which are inactive in oriC replication. The toxicity could be alleviated in the presence of limited amounts of E. coli DnaB or DnaC proteins (Fig. 6). Overproduction of these host proteins did not affect the results (data not shown), suggesting that even low-level production is sufficient to mitigate the toxic effects of DnaCGifsy.

Compatibility of plasmids derived from bacteriophage λ and Gifsy prophages

Sequence analysis suggests that the structure of iterons carried by the origin of replication of Gifsy phages is very similar to that of the iterons carried by phage λ. Also, both

![Fig. 5. Replication of Gifsy plasmids in dnaA46(ts) cells. The relative plasmid content in the cells was measured densitometrically at 60 min intervals in a dnaA46(ts) strain harbouring pGifsy (□), pGifsyNoLoader (△), pKB2p (a λ replicon, ○) or pQE30 (a pMB1 replicon, ×), before and after temperature upshift (time 0).](image)

![Fig. 6. Production of DnaB and DnaC alleviates the toxic effects of P_gifsy on bacterial growth. The growth of E. coli XL1-Blue strain harbouring pQE30 (vector control, ×), pQE-GifsyDnaC (dnaCGifsy under an IPTG-inducible promoter, ○) or pQE-GifsyDnaC with pINB (carrying dnaC under a p ara promoter, ■), or pQE-GifsyDnaC with pINC (carrying dnaB under a p ara promoter, △) was monitored by measuring OD575 relative to the optical density of the culture at the time of induction of dnaCGifsy overexpression with 1 mM IPTG (time 0). The figure shows results obtained in the absence of L-arabinose. Results obtained after induction of p ara were similar (data not shown).](image)
origins probably depend on a very similar repertoire of host factors for their replication. We therefore investigated if the plasmids derived from bacteriophage λ and Gifsy prophages could be stably co-maintained in bacterial cells. We found that pGifsy (which carries a tetracycline-resistance gene) was incompatible with pCB104, a λ plasmid carrying a gene for chloramphenicol acetyltransferase: out of 100 colonies formed by bacteria originally carrying both a λ plasmid and pGifsy, but selected for growth on a plate containing only one antibiotic (chloramphenicol or tetracycline), none grew on the plates that also contained the other antibiotic. However, in the presence of pQE-EcoliDnaC (which carries resistance for ampicillin; without addition of IPTG), both plasmids derived from lambdoid phages were compatible (100 colonies were tested). In cells carrying pQE-GifsyDnaC, on the other hand, the plasmids were weakly incompatible (without addition of IPTG, 77 % of the cells grown on chloramphenicol were tetracycline resistant and 32 % grown on tetracycline were chloramphenicol resistant). While the reasons for the incompatibility are not clear, these results provide an additional illustration of the interference between the helicase loading of the lambdoid plasmids and the host chromosome. Such interference might arise from the competition of the loaders for the DnaB helicase, and from the formation of heterocomplexes of helicase loaders inactive in the initiation of chromosomal initiation.

Involvement of molecular chaperones in pGifsy replication

λ P protein binds very strongly to the host DnaB helicase and inhibits all its activities. The action of molecular chaperones DnaK and DnaJ, supported by the activity of GrpE, is necessary for rearrangement of the pre-primerosomal complex and liberation of DnaB from λ P-mediated inhibition (Liberek et al., 1988; for review see Węgrzyn & Węgrzyn, 2002). DnaC–DnaB complexes are significantly weaker than λ P–DnaB complexes (Wickner, 1979; Wold et al., 1982; Mallory et al., 1990; Konieczny & Marszałek, 1995), and it appears that DnaC dissociates from DnaB without the assistance of molecular chaperones. The fact that host DnaC can act instead of DnaC_{Gifsy} in pGifsyNoLoader replication does not necessarily imply that the interaction of DnaC_{Gifsy} with DnaB is similarly weak. Indeed, selective pressure to redirect the host helicase to the viral origin has been put forward as an explanation for the strong λ P–host DnaB interaction, and all lambdoid phages are, conceivably, under similar pressure in this respect. We therefore investigated the requirement for molecular chaperones in replication of Gifsy-derived plasmid DNA.

We found that, unlike λ plasmids (Węgrzyn et al., 1996), pGifsy effectively transformed E. coli dnak756 and dnaJ259 mutants at 30 °C (data not shown). This suggests that the activity of the Hsp70 chaperone machine is not required for oriGifsy replication. It is still possible that molecular chaperones play some role in Gifsy replication, but this would be difficult to demonstrate in vivo, considering that host DnaC can function in the delivery of the helicase to the Gifsy origin.

Stringent control of pGifsy replication

Replication of plasmids derived from bacteriophage λ is inhibited in amino acid-starved wild-type cells (i.e. during the stringent response) but not in relA mutants (i.e. during the relaxed response; for a review see Węgrzyn, 1999). The replication can proceed in the absence of protein synthesis, due to formation of a stable replication complex, which is inherited by one of two daughter plasmid copies after each replication round, and can function in subsequent replication events (Węgrzyn & Węgrzyn, 2001). Under the stringent response, λ plasmid DNA replication is arrested due to a lack of activity of the λ P_{B} promoter, necessary for the transcriptional activation of the origin (Węgrzyn & Węgrzyn, 2001). The activity of this promoter is regulated by the specific nucleotide produced in amino acid-starved cells, guanosine 5′-diphosphate-3′-diphosphate (ppGpp; Wróbel et al., 1998a). In relA bacteria [which do not have (p)ppGpp synthetase I activity], ppGpp does not accumulate during amino acid starvation, and λ plasmid replication continues even though new replication complexes are not formed.

We found that, similar to λ plasmids and plasmids derived from the Rac prophage of E. coli K-12 (Potrykus et al., 2000), the replication of pGifsy continued in the amino acid-starved relA cells. We observed a moderate increase in the copy number of pGifsy during the relaxed response, but not during the stringent response (Fig. 7), which implies that replication of pGifsy is under stringent control, similar to the replication of λ plasmids. During the stringent response, plasmid replication appeared to be completely stopped, which resulted in the observed decrease in the amount of plasmid DNA relative to bacterial cell mass (determined by OD_{575} measurement; under these experimental conditions, the culture OD_{575} slowly increased, from ~0.2 at the onset of starvation to ~0.4 after 5 h).

DISCUSSION

In this report, we demonstrate that a DNA fragment of prophages Gifsy-1 and Gifsy-2 of S. enterica serovar Typhimurium, which in genetic organization is very similar to the replication region of bacteriophage λ, can replicate autonomously in E. coli cells. The origin of replication of pGifsy appears to be located in the middle of the O_{Gifsy} gene, which is necessary for plasmid replication, and likely encodes a replication initiator protein.

Helicase loaders of lambdoid phages

The analysis of the replication regions of lambdoid phages indicates that four types of these regions can be identified,
depending on the way the helicase is probably brought to the origin: (i) replicons using a homologue of phage λ P protein; (ii) those using a loader similar to the enterobacterial DnaC which transports the host helicase; (iii) those using such a protein but probably to transport a phage helicase encoded by a neighbouring gene; and (iv) phages that have a gene for a helicase but not a loader. The prophages carrying a dnaC-like gene are particularly common in the sequenced genomes of pathogenic bacteria (Fig. 2). It is unclear if the two types of helicase loaders are related: λ P and DnaC show no apparent sequence similarity, and the predicted secondary structure of the two proteins does not show any common pattern. Perhaps this question can be answered when the structure of both proteins is resolved.

Although Gifsy is similar to λ in that it uses the host helicase and a phage-encoded helicase loader, the fact that the Gifsy loader can be substituted by the host-encoded DnaC protein is a significant difference between the two replicons. It is difficult to say at this point if this functional substitution allows for a drastically reduced requirement for molecular chaperones DnaK and DnaJ in Gifsy replication initiation. The other possibility is that the DNA–DnaB interaction itself resembles the DnaC–DnaB interaction, and that the activity of molecular chaperones is not necessary to free the helicase from the phage loader during recruitment at the origin of replication. Host DnaC leaves the initiation complex immediately after helicase loading, a process accompanied by ATP hydrolysis, freeing DnaB from inhibition (Wahle et al., 1989a, b). Whether phage DnaC-like proteins function similarly could be answered by in vitro experiments.

In λ replication, molecular chaperones are responsible for the rearrangement of the pre-primosomal complex and liberation of DnaB helicase from λ P-mediated inhibition (Węgrzyn & Węgrzyn, 2002). However, a substitution in the P gene (a π mutation) allows for λ replication in dnaK and dnaJ mutant hosts (Konieczny & Marszalek, 1995). The λ P π gene product interacts with DnaB, significantly more weakly than the wild-type λ P protein (Konieczny & Marszalek, 1995). Our experiments suggest that the π mutation allows the λ plasmid to be maintained in dnaCl(ts) bacteria. The fact that derivatives of plasmids encoding a loader with weaker binding to the helicase can be readily obtained suggests that the gene experiences selective pressure, so that the phage loader can compete more successfully with the host DnaC protein for DnaB (Konieczny & Marszalek, 1995).

Why is efficient binding of the host helicase advantageous for the virus? Perhaps it is important to prevent the replication of other replicons in the cell that depend on host DnaB (the chromosome, plasmids and competing viruses). Phage λ also uses other mechanisms to stall host replication (Kędzierska et al., 2003). This might explain why the viral DnaC-like proteins, despite high similarity to the host protein, are inactive in host replication, and indeed toxic when overproduced. The results of our experiments on the incompatibility between pGifsy and λ P plasmid suggest that moderate levels of the Gifsy helicase loader help to prevent plasmid loss (although not as efficiently as the host helicase loader), which is most likely caused by the competition for the host helicase between the lambdaoid plasmids and the bacterial chromosome. This is consistent with the hypothesis of formation of heterocomplexes between DnaCGifsy and DnaC which can deliver the helicase to oriC, as long as the host DnaC is present in the heterocomplex. According to this interpretation, DnaCGifsy homocomplexes (proposed to be abundant when the level of the protein is high, e.g., during the lytic cycle or when the protein is overexpressed from a plasmid) bound to DnaB are not active in the initiation of chromosomal replication. Additional mechanisms for DnaCGifsy toxicity (perhaps similar to the toxicity of λ P; Datta et al., 2005a, b) may come into play.

It would be interesting to see if the loader in the phages that encode a viral helicase also plays a role in binding a helicase of the host to stop replication of the competing DNA molecules. Such binding could be a pre-adaptation for the development of a lambdaoid replicon in which the host helicase is used instead of the phage-encoded protein, and provides a possible scenario for the origin of helicase loader proteins.

**Evolution of helicase loading in bacteria**

Although it is possible that helicase loaders are present in bacterial species whose genomic sequences have not yet
been obtained, so far, helicase loaders have only been identified in bacteria related to *E. coli* and *Bacillus subtilis*. In Gram-positive bacteria, two proteins are involved in loading the replicative helicase (which in *B. subtilis* has been named DnaC): DnaB and DnaI (Velten et al., 2003). *B. subtilis* DnaB does not seem to have known homologues outside the Firmicutes that could be found in BLAST searches of the GenBank database. *B. subtilis* DnaI, on the other hand, contains an ATP-binding domain homologous to that present in *E. coli* DnaC, and interacts with *Bacillus* helicase (Soutilanas, 2002), forming a helicase–loading complex (the same stoichiometry as *E. coli* DnaB6–DnaC6, Ioannou et al., 2006). Proteins similar to *B. subtilis* DnaI are encoded in replication modules of bacterial viruses infecting Gram-positive bacteria (data not shown).

In other words, analysis of fully sequenced genomes shows that numerous bacteria lack DnaC-like and DnaB/I-type helicase loaders. This may mean that: (i) these organisms encode a different type of helicase loader; (ii) that they have evolved so that such factors are no longer required for initiating DNA replication; or (iii) that the helicase-loading proteins are a late introduction into the bacterial replication initiation system.

Biochemical evidence suggests that the first possibility is unlikely, at least for some species related to *E. coli* in *Pseudomonas* species, DnaB protein apparently acts without any accessory protein (Caspi et al., 2001). The fact that most bacterial lineages are devoid of the DnaC homologue is not consistent with the hypothesis of an ancient origin for helicase loading; the wide-ranging loss of an important element of the cellular machinery is not likely. This reasoning leads to the proposal that the enterobacterial dnaT–dnaC genes and the gene for helicase loader in Gram-positive bacteria were acquired relatively recently from lambdoid replicons, most likely from a defective lambdoid prophage (Wróbel & Wegryzn, 2002). Low-level expression, perhaps after several mutations attenuating the toxicity, might have permitted the adaptation of the viral loader for the host replication initiation machinery. Subsequently, it might have allowed for at least a short-term competitive advantage against the toxic effect of viral helicase loaders. The ensuing molecular arms race could have resulted in the development of a λ P protein that binds DnaB very strongly, but requires the action of molecular chaperones to unblock helicase activity. The question whether viral DnaC-like proteins bind the helicase more strongly than the loader adapted by the host, as a result of such a molecular arms race, remains to be answered by in vitro experiments.

Perhaps the most interesting aspect of the hypothesis of dnaC acquisition from lambdoid prophages is the continued presence of the gene in the extremely reduced genomes of *Buchnera*, intracellular endosymbionts of aphids, as demonstrated by the results of our database searches. *Buchnera* genomes are about one-seventh of the *E. coli* genome and have experienced severe loss of coding sequences, including genes considered essential for bacterial cells, e.g. *recA* (Shigenobu et al., 2000). Reconstruction of the core genome of the last common ancestor of *E. coli*, *Klebsiella pneumoniae*, *Yersinia pestis* and *Vibrio cholerae* indicates that dnaC and dnaT may be the only two genes added to the genome of the *Escherichia–Buchnera* ancestor that have been subsequently conserved in the *Buchnera* lineage (Moran & Mira, 2001). The evolutionary scenario we propose is consistent with the view that the last common ancestor of *Buchnera* and *Escherichia* was more recent than the divergence of other enterobacterial lineages, in particular, the *Yersinia* lineage (Lerat et al., 2003).

However, it must be noted that the issue of the phylogeny of *Buchnera* (and other endosymbionts) is controversial (Lerat et al., 2003; Canback et al., 2004). The fact that, at least in one lineage of endosymbionts, reductive evolution did not affect the gene for the helicase loader, which is apparently not needed by closely related enterobacteria, such as *Yersinia*, suggests that the replication initiation machinery has been adapted to new members, the reversal of the process is difficult.

The hypothesis that cellular proteins involved in DNA metabolism have a viral or plasmid origin is not new. It has been proposed (Forterre, 1999) to explain the apparent lack of homology between the bacterial, archaean and eukaryotic proteins involved in the processing of DNA (as opposed, for example, to the genes involved in translation). This hypothesis postulates the replacement of genes that coded for such enzymes and originated before the divergence of the three lineages by viral/plasmid genes. In other words, the idea is to explain the very deep phylogenetic relationships (or the lack of them) between the genes of the three lineages as an alternative to the hypothesis according to which the lineages separated when RNA, not DNA, still carried out the information storage task in the cell (Mushegian & Koonin, 1996). As support for his hypothesis, Forterre (1999) has noted: (i) high similarity between many viral and plasmid proteins and bacterial replication proteins; and (ii) that non-orthologous gene displacement of many replication and recombination proteins has been observed in experimental conditions. We do not believe that the latter is very relevant (although it is suggestive), since the selection conditions in the laboratory are very different from those in the natural environment. Another argument (Iyer et al., 2005) against the displacement hypothesis is the lack of evidence that any gene central to DNA replication (e.g. those for the main replicative primase, helicase, DNA clamp, DNA ligase or DNA polymerase) has ever been displaced by the viral or plasmid gene within any lineage (the emphasis is ours). We agree that the central DNA replication machinery, once the individual components are fine-tuned, is most likely refractory to displacement. This is equivalent to saying that cells in which such co-opting occurs have no selective advantage over the wild-type competition. Of course, this view can only be falsified when such displacement is observed in at least one of the
lineages. However, the stress here should be put on the fact that the machinery first needs to be fine-tuned to be refractive, and the fact that the displacement has not been observed within the extant clades is not relevant to possible displacements of ancient, possibly less-adapted systems, at the time of lineage divergence.

Unfortunately, the hypothesis of the viral origin of helicase loaders in enterobacteria and Firmicutes is of little use to resolve this debate. It is our argument that the helicase loaders appear to be rather peripheral members of the DNA replication machinery. We propose a scenario of co-opting new peripheral players to the otherwise fine-tuned machinery under very strong selective pressure (first suppression of toxic effects, then competition between the host and virus replication origins for the host helicase), rather than a scenario of displacement.

The observed parallels between the mechanism of helicase loading in model organisms – bacteriophage λ, E. coli and B. subtilis – have attracted even more attention after the observation of the interaction between the Mcm2-7 helicase hexamer and a protein similar to DnaC of E. coli, Cdc6 (Davey et al., 2002; Maiorano et al., 2006), important for replication initiation in yeast. However, although Cdc6 and E. coli DnaC both belong to the AAA+ family of ATPases, this protein family is very large. In the case of the Firmicutes and enterobacteria, the same stoichiometry of the helicase-loader complex, and the relationships between the bacterial and viral proteins (with lambdoid phages as a bridge linking two bacterial lineages) suggest that the most recent common ancestor of these proteins could have been a helicase loader. It is not clear if this is the case for the ancestor of E. coli DnaC and yeast Cdc6. We are not aware of any results that would suggest that Cdc6 interacts with the helicase in a fashion similar to the bacterial helicase loaders (a hexamer of helicase loaders interacting with the helicasehexamer before DNA binding). Until this is demonstrated, it is doubtful whether the E. coli DnaC–helicase or Bacillus Dna–DnaB–helicase systems provide relevant bacterial models for replication initiation in eukaryotes.

The research on helicase loading in model organisms has led historically to many important discoveries (e.g. it helped in the discovery of molecular chaperones; Yochem et al., 1978; Liberek et al., 1988). It now appears possible that the current models of the initiation of chromosomal DNA replication may be relevant only to bacteria closely related to the model organisms E. coli and Bacillus. It is necessary to shift our attention to other bacterial species so that general mechanisms can be proposed.

ACKNOWLEDGEMENTS

This work was supported by the Polish Ministry of Science and Higher Education (grants 2 P04A 025 27, N301 122 31/3747 and 72-6PRUE-2005-7), the European Union (Structural Funds Z/2.22/II/2.6/002/05), and the Foundation for Polish Science (B. W.). We appreciate the insights and comments of the Editor (Dr Laurent Jannière) and the anonymous reviewer.

REFERENCES


Virulent *Salmonella Typhimurium* has two periplasmic Cu, Zn-superoxide dismutases. *Proc Natl Acad Sci U S A* 96, 7502–7507.


Jensen, K. F. (1993). The *Escherichia coli* ‘wild types’ W3110 and MG1655 have rph frame shift mutation that leads to pyrimidine starvation due to low pyrE expression levels. *J Bacteriol* 175, 3401–3407.


