Bacteriophage SPP1 DNA replication strategies promote viral and disable host replication in vitro

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Nucleic Acids Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>NAR-02822-H-2012.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>1 Standard Manuscript</td>
</tr>
<tr>
<td>Key Words:</td>
<td>DNA replication, Gram-positive, SPP1, Origin binding protein, single-stranded DNA binding protein</td>
</tr>
</tbody>
</table>
Bacteriophage SPP1 DNA replication strategies promote viral and disable host replication in vitro.

Elena M. Seco¹, John C. Zinder², Carol M. Manhart², Ambra Lo Piano¹, Charles S. McHenry²,* and Silvia Ayora¹,*

¹Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain
²Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309, USA

*To whom correspondence should be addressed. Tel: +34 91585 5450; Fax +34 91585 4506; E-mail: sayora@cnb.csic.es

Correspondence may also be addressed to Charles S. McHenry. Tel: +1 303 6206357; Fax: +1 303 7351347; E-mail: charles.mchenry@Colorado.EDU
Abstract

Complex viruses that encode their own initiation proteins and subvert the host’s elongation apparatus have provided valuable insights into DNA replication. Using purified bacteriophage SPP1 and *Bacillus subtilis* proteins, we have reconstituted a rolling circle replication system that recapitulates genetically defined protein requirements. Eleven proteins are required: phage-encoded helicase (G40P), helicase loader (G39P), origin binding protein (G38P) and G36P single-stranded DNA binding protein (SSB); and host-encoded PolC and DnaE polymerases, processivity factor (β2), clamp loader (τ-δ-δ’) and primase (DnaG). This study revealed a new role for the SPP1 origin binding protein. In the presence of phage SSB it is required for initiation on replication forks that lack origin sequences, mimicking the activity of the PriA replication restart protein in bacteria. The SPP1 replisome is supported by both host and viral SSBS, but phage SSB is unable to support *B. subtilis* replication, likely due to its inability to stimulate the PolC holoenzyme in the *B. subtilis* context. Moreover, phage SSB inhibits host replication, defining a new mechanism by which bacterial replication could be regulated by a viral factor.
Introduction

Double-stranded (ds) DNA viruses may encode most if not all the components for their own replication, as in the case of the T4, Φ29 or HSV-1 viruses (1-3), or may encode a subset of proteins, including an origin-specific initiation protein, and recruit the host DNA replication machinery to achieve efficient viral replication. The study of the replication mechanisms of the latter type of viruses has provided significant insight into cellular DNA replication processes. For example, SV40, a virus that encodes its own origin binding protein and helicase within the T antigen has provided a viral window into eukaryotic DNA replication (4,5). Its use enabled the only full reconstitution of a eukaryotic DNA replication system with purified proteins and revealed the special roles of Pol α-primase and the Pol δ holoenzyme in the process (5,6). Bacteriophage λ has provided similar insight into the replication of Gram-negative bacteria (7). λ encodes its own origin protein and a helicase loader that subverts the host DnaB helicase, leading to the acquisition of the cell’s elongation apparatus. Through this system, the role of heat shock proteins in freeing the DnaB helicase from tightly bound λ O and P proteins was discovered, providing one of the initial observations of chaperone function (7,8). The mechanism used by these viruses to recruit host proteins could provide significant insight into viral and host processes.

*Bacillus subtilis* SPP1 is a virulent dsDNA phage whose mature genome is a linear 45.4-kb dsDNA. The ends of the packaged DNA are terminally redundant and are permuted to facilitate circularization after DNA injection into cells. SPP1 replication starts with the circle-to-circle replication mode (θ replication), but after one or a few rounds, it switches to concatemeric replication (termed σ replication or rolling circle replication) by a process driven by homologous recombination (9,10). This switch in the
mode of replication is a strategy used by many viruses to produce linear head-to-tail concatemers that are used by the packaging machinery. This late-phase DNA replication, which is believed to be independent of an origin of replication, has been reconstituted in vitro for viruses that encode their own polymerase [e.g., HSV-1, T7, T4 and Φ29, (11-14)] and for bacteriophage λ (15).

Genetic analyses showed that SPP1 DNA replication is independent of the host origin binding protein (DnaA), the replicative DNA helicase (DnaC), primosomal proteins DnaB and PriA, and RNA polymerase. These studies also showed that SPP1 replication requires the host DnaG primase and PolC DNA polymerase (16-18).

The SPP1 phage possesses two origins of replication, oriL and oriR, which are 32.1 kb apart in a linear map of the SPP1 genome. Replication proteins are encoded by two operons. The first one, which is under the control of the early promoter PE2, codes for proteins that have been shown to be required for θ replication: the G38P origin binding protein, the G39P helicase loader, and the G40P helicase. G38P, which does not belong to the AAA+ family, is widely conserved in phages (19,20). G38P binds with high affinity to the two origins of replication (21) and forms a complex with G39P (17). G39P does not share homology with other studied helicase loaders, but performs a similar role: it delivers G40P, upon interacting with G38P, to the origin of replication (22,23). G40P is a widely studied helicase that belongs to the DnaB family (24-26).

Genes required for the recombination-dependent σ replication mode are under the control of the early promoter PE3. These include a recombinase, G35P (18,27), and a 5’ → 3’ exonuclease, G34.1P (28). In this operon there is also a gene (gene 36) that encodes a single-stranded DNA binding protein (SSB), G36P, whose role in replication has not yet been analyzed. G36P is 48% identical to the essential host SSB (B. subtilis...
SsbA, Supplementary Figure S1), and 38% identical to the competence-specific SSB [B. subtilis SsbB, (29)].

We have exploited the apparatus required for the σ mode of replication to establish a robust rolling circle replication system that requires four phage proteins and seven host elongation proteins. These studies revealed surprising new roles for the G38P origin binding protein in the initiation of DNA replication on forks that do not contain origin sequences. In addition, they show the versatility of the SPP1 replication fork, where both the viral and the host SSB may be used, in contrast to the B. subtilis replication fork, which uses only its own SSB (the SsbA protein). Moreover, B. subtilis replication is inhibited by the viral SSB (G36P), a mechanism that is likely exploited by the phage to shut down host DNA replication synthesis and foster its own replication.
Materials and methods

Rolling circle assays

Standard reactions consisted of 30 nM G40P, 300 nM G39P, 300 nM G38P, 8 nM DnaG, 15 nM DnaE, 20 nM PolC, 25 nM τδ, 25 nM δ, 24 nM β, 30 nM G36P, 5 nM mini-circular DNA template, 350 μM ATP, 100 μM CTP, GTP and UTP, 48 μM dNTPs (except 18 μM dCTP or dGTP for the leading and lagging strand DNA synthesis respectively), and 0.2 μCi/reaction [α-32P]dCTP or [α-32P]dGTP. The DNA template was a 409 nt circle containing a 396 nt tail described in (30), but prepared by an alternative procedure that included a PCR amplification step (Yuan and McHenry, in preparation). The reactions were carried out in 12.5 μl of buffer BsRC (40 mM Tris-acetate (pH 7.8), 12 mM magnesium acetate, 3 μM ZnSO4, 1 mM dithiothreitol, 0.02% (w/v) Pluronic F68, (30)) that contained 500 mM potassium glutamate and 1% (w/v) polyethylene glycol (PEG-8K). The buffer also contained 4% glycerol, 19 mM NaCl and 4 mM Tris-HCl that was contributed by the addition of protein solutions. Incubations were conducted for 10 min at 37 ºC. An enzyme mix containing all protein components except SSB (G36P, or SsbA as indicated) was prepared in buffer BsRC. Two different substrate mixes containing template DNA, rNTPs, dNTPs, SSB (G36P or SsbA) and either [α-32P]dCTP or [α-32P]dGTP for measurement of leading and lagging strand synthesis, respectively, were prepared. Reactions were initiated by mixing the enzyme mix and a substrate mix. After incubation, reactions were stopped by addition of an equal volume of stop mix [40 mM Tris-HCl (pH 8.0), 0.2% SDS, 100 mM EDTA, and 50 μg/ml proteinase K]. Samples were treated for 20 min at 37 ºC, then applied onto Sephadex G-50 columns to eliminate non-incorporated dNTPs. The extent of DNA synthesis in leading and lagging strands was quantified by scintillation counting.
For the analysis of the size of leading and lagging strand products, samples were brought to 50 mM NaOH, 5% (v/v) glycerol and 0.05% bromphenol blue and fractionated on alkaline 0.5% agarose gels for approximately 3 h at 80 V. Alkaline agarose gel buffer consisted of 30 mM NaOH and 0.5 mM EDTA. Gels were fixed in 7% (w/v) trichloroacetic acid, dried, autoradiographed on storage phosphor screens and analyzed with Quantity One (Bio-Rad) software.

For calculating the rate of SPP1 fork progression, aliquots were removed, quenched and processed as described above. The molecular weight of the longest leading strand product at each time was extrapolated from labeled DNA size standards and plotted as a function of time. The elongation rate was determined by calculating the slope of this curve (31).


**Extension of DNA primers annealed to M13**

Templates were prepared by mixing 50 pmol single-stranded M13*ori* DNA (32) with 60 pmol synthetic DNA primer (5’AGGCTGGCTGACCTTCATCAAGAGTAATCT) in 70 µl of a buffer consisting of 40 mM Tris-HCl (pH 7.8), 150 mM NaCl, and 1 mM EDTA, heating to 95°C, cooling to room temperature over 1 h and diluting the resulting mixture to 29 nM as circles. Holoenzyme reactions (25 µl each) contained 2.3 nM template; 2 nM PolC or 3 nM DnaE; 25 nM β2, if present, 15 nM τ4, 20 nM δ, 20 nM δ’, and variable concentrations of B. subtilis SsbA4 or SPP1 G36P4; 48 µM dATP, dGTP and dCTP; 18 µM [3H]dTTP (specific activity 113 cpm/pmol); and 250 µM ATP in buffer BsM13 (40 mM Tris-acetate (pH 7.8), 340 mM potassium glutamate, 10 mM
magnesium acetate, 4 µM ZnSO₄, 0.015% (w/v) Pluronic F68). Reaction mixtures were prepared on ice, initiated by incubation at 30 °C, stopped after 3 min (for PolC) or 5 min (for DnaE) with 2 drops of 0.2 M sodium pyrophosphate, and incorporated nucleotides were precipitated with 0.5 ml 10% (w/v) trichloroacetic acid. Unincorporated nucleotides were removed and reaction products were quantified as described (33).

**Helicase assays**

Oligonucleotides were obtained from Biosearch Technologies. The substrate diagrammed in Figure 6A was assembled from the following HPLC-purified oligonucleotides. Leading strand template 90-mer- 5’-TET-CGCCGTATAGATCATTATAACATGTAGATTCATGATAATATAAGAGATGACGAATATGATTTTGCTGTAATGTAAGAATCTTCAA contained fluorescent tetrachlorofluorescein (TET) at the 5’ terminus. Lagging strand template 90-mer- 5’-TT(biotin)₄ATATTATCATGAATCTAACATGTTATAGTAATGATCTATACGCG-BHQ-1-3’ contained biotin conjugated to preceding thymidine and Black Hole Quencher-1 (BHQ-1) at 3’ terminus that quenches fluorescent TET dye. Primer 35-mer- 5’-TTGAAGATTCTTACATTAGTTGACAAAATCATATT, when annealed to the leading strand template, created a 10 nt gap. Trap oligo 45-mer 5’-TATATTATCATGAATCTAACATGTTATAGTAATGATCTATACGCG-3’ was used to capture helicase-displaced leading strand so that it did not reanneal to the lagging strand template that contained the fluorescence quencher. The substrate was formed by annealing 1 µM leading strand template with 1 µM lagging strand template and 1 µM primer in a buffer containing 10 mM Tris (pH 7.75), 50 mM NaCl, and 1 mM EDTA in a final volume of 25 µl. The sample was heated to 95 °C for 5 min and cooled to 25 °C at 1 °C/min.
For FRET experiments, 20 nM oligonucleotide substrate was combined with 100 nM trap oligo, 200 nM streptavidin, and protein components in a buffer containing 50 mM Hepes (pH 7.5), 10 mM magnesium acetate, 10 mM dithiothreitol, 20% (v/v) glycerol, 0.02% (v/v) Nonidet-P40 detergent, 200 µg/ml bovine serum albumin, 100 mM potassium glutamate, and 10 mM ATP in a round-bottomed black 96-well plate in a final volume of 50 µl. Samples were incubated at 30 °C for 15 min. Fluorescence emission was detected at 535 nm using an Envision plate reader with an excitation of 485 nm. Using concentrations of unannealed fluorescent leading strand template that are in the linear range of the assay, fluorescence units were converted to molarity using a standard curve.

**In vivo replication of SPP1 in a ssbΔ35 background**

The *B. subtilis* FLB22 (ssbΔ35) and FLB23 (ssb3+) strains were a kind gift of P. Polard (CNRS, France). They were obtained by a single crossing-over integration procedure of pMUTIN-SPA derivatives; in those two strains, the essential *rpsR* gene, which is located immediately after *ssb*, is placed under the control of the IPTG-inducible Pspac promoter (34). FLB22 is a mutant strain, which expresses from its natural promoter an *SsbA* truncated of its last 35 amino acids. *ssbΔ35* cells are temperature-sensitive for growth above 47 °C in LB medium. FLB23 is an isogenic strain encoding a wild-type *SsbA* protein. FLB22 and FLB23 cells were grown at 30 °C in LB medium with 0.5 mM IPTG until OD560 = 0.2 and then shifted to 50 °C. After 15 min incubation, the cells were infected with a multiplicity of infection of 10 with the SPP1 phage and the cultures were incubated for 120 min at 50 °C. *Infection experiments at permissive temperature were performed in parallel.* After centrifugation, the supernatant that contained free phage particles was filtered through 0.45 µm filters. Titrations were carried out using *B.*
*subtilis* BG214 as the indicator strain.

**Results**

**Reconstitution of a SPP1 Replication Fork**

We reconstituted replication on a synthetic 409-bp circle containing a long flap that mimics a replication fork (Figure 1B). A 50:1 asymmetric G:C distribution in the synthetic template permits facile quantification of leading and lagging strand synthesis (Figure 1C). The structure of this template may mimic the intermediate that is formed in SPP1 once the D-loop, formed by recombination, is resolved to initiate concatemeric (σ mode) replication (35). SPP1 proteins were expressed and purified to homogeneity (Supplementary methods, Figure 1A) and added to various combinations of purified *B. subtilis* DNA replication proteins (30) to determine which combination was required for efficient replication. Consistent with genetic requirements, SPP1 G39P and G40P were required, as were host PolC and DnaG (Figure 1D and E). Additional components of the host replicase including the clamp loader [τ complex which consists of the τ, δ and δ’ subunits] and sliding clamp processivity factor (β2) were also found to be required. As observed in *B. subtilis*, DnaE was also necessary, and like primase, plays primarily a lagging strand role (30,36). The decreased level of leading strand synthesis in the absence of DnaE and DnaG may be due to a decreased efficiency when the replisome is incomplete—lacking these factors. To our surprise, the G38P origin binding protein was also essential, even though the origin sequence was not contained within the DNA template. G36P was also required, but some synthesis was observed in the absence of G36P, due to a non-physiological reaction where the helicase can ‘self-load’ by threading over the 5’-end of the template flap (Manhart and McHenry, in preparation).
The required proteins were individually titrated in the presence of optimal concentrations of the remaining proteins to optimize the replicative reaction (Supplementary Figure S2).

We examined the time course of DNA leading and lagging strand synthesis with the SPP1 replisome. Both leading and lagging strand synthesis exhibited a lag phase of 1 min, presumably the time required for loading of the helicase and assembly of the replication fork. After the lag, the synthesis rate remained linear for approximately 5 min (Figure 2A). We examined the elongation rate of reconstituted SPP1 replication forks by analyzing leading strand product formation after the first minute lag (Figure 2B and C). From these data, a rate of progression of 224±7 nucleotides/s was obtained. We performed as a control the same assays with a reconstituted *B. subtilis* replication fork and obtained a value of 200±6 nucleotides/s. These results show that both replisomes progress at a similar rate under the experimental conditions used.

**Primase, not DnaE, regulates the length of Okazaki fragments**

*B. subtilis* DnaG primase cycles on and off the replication fork through association with the DnaC helicase. Thus, higher concentrations of primase lead to more frequent associations with helicase and more frequent priming, resulting in shorter Okazaki fragments (30). We also observed a decrease in Okazaki fragment size with increasing DnaG primase concentration in our SPP1 system (Figure 3A). Okazaki fragment length varied from ca. 4 kb at 1.5 nM primase down to ca. 400 bp at 100 nM primase (Figure 3A). *B. subtilis* DnaE functions like eukaryotic DNA polymerase α, adding a stretch of deoxynucleotides to the RNA primer before handoff to the major replicase (30). We also investigated whether variations in DnaE concentration influence Okazaki fragment length. In the absence of DnaE, lagging strand synthesis was very low (Figure 1E and Figure 3B). Most of the synthesis observed in the absence of DnaE was due to
background incorporation of \([\alpha-^{32}\text{P}]\text{dGTP}\) into the leading strand product. The leading strand template contains 2% of the template C residues within our rolling circle template and the resulting incorporation yields a significant background only in the absence of lagging strand synthesis. The size of Okazaki fragments was similar over a wide range of concentrations (1.25-80 nM DnaE, Figure 3B).

**SPP1 replication forks can be reconstituted with SsbA from *B. subtilis* but the helicase loaders are not interchangeable**

In order to see if some SPP1 components could be replaced by their *B. subtilis* counterparts, we first determined whether host SsbA could replace G36P. *B. subtilis* SsbA supports an efficient reaction at both 30 nM [the optimal G36P concentration, that is also the amount of SSB needed to cover the 396-nt tail present in the DNA template in the 65-nt DNA binding mode (37) and at higher concentrations (Figure 4A, lanes 7 and 11). We observed a stronger dependence on G38P at 90 nM *B. subtilis* SsbA. The reactions were dependent on G39P under both conditions tested. In the absence of any SSB, efficient reactions are also observed, but dependence on G38P and G39P was lost. In the absence of SSB, helicases can self-assemble by threading over the exposed 5’-end of the flap of the forked substrates (Manhart and McHenry, in preparation). We found that G40P could not be assembled onto the fork by the host loading system (DnaB/DnaD/Dnal/PriA) in the presence of either SsbA or G36P (Figure 4A, lanes 2 and 15).

We then tested whether G36P can replace SsbA in support of the *B. subtilis* replisome. *B. subtilis* chromosomal replication does not occur in the presence of G36P either in the presence of the natural helicase loading system (PriA, DnaD, DnaB, Dnal) (Figure 4B, lane 4) or in the presence of the phage G38P-G39P helicase loading system.
(Figure 4B, lane 6). Increasing G36P concentrations to levels found to be optimal for SsbA also did not stimulate the *B. subtilis* reaction (Supplementary Figure S3). In the presence of SsbA the bacterial helicase worked well with its own helicase loading proteins (Figure 4B, lane 1), but they could not be substituted with the viral helicase loading proteins (Figure 4B, lane 3).

**Elevated levels of G38P are required to reverse inhibition of DNA replication by high concentrations of G36P**

The above experiments showed that G36P and SsbA could work similarly on SPP1 replication forks. In titration experiments that we performed to optimize protein concentrations, we noted that increasing concentrations of G36P significantly reduced lagging strand DNA synthesis, whereas this effect was not observed with increasing SsbA (Figure 5A and B). We were concerned that the leading strand synthesis observed at elevated G36P may have been an inauthentic reaction, resulting from helicase independent strand displacement by the Pol C holoenzyme. Such a reaction is catalyzed by the *E. coli* Pol III holoenzyme at high SSB concentrations (31). However, dropout experiments where one protein was deleted from the reaction at a time confirmed that leading strand synthesis in the presence of elevated G36P concentrations retained a dependency on all of the leading strand replication proteins, including helicase (Supplementary Figure S4).

We suspected the inhibition by high G36P may have been caused by sequestering a component in a binary complex in solution, preventing its participation in the replicative reaction. In a search for proteins that reversed the G36P inhibitory effect at high concentrations, we found that G38P elicited this effect (Figure 5C). We also tested the other viral protein that participates in helicase loading, G39P but observed no effect (data not shown).
We also performed an independent fluorescence-based assay that detects helicase loading and ensuing strand separation (see Figure 6A for a description of the substrate). This permits analysis of the SSB effect on helicase activity independent of its influence on priming and polymerase activity. The 5'-end of the lagging strand template was blocked by streptavidin attachment, preventing helicase self-assembly by threading over the 5'-flap in the absence of SSB (Manhart and McHenry, in preparation). With the G40P helicase we also observed in these assays greater efficacy of G36P relative to SsbA at low concentrations (Supplementary Figure S5D). Furthermore, we observed an absolute requirement for G38P (Supplementary Figure S5A), consistent with a role for this protein at the helicase assembly step.

**SPP1 does not require SsbA in vivo**

The previous assays showed that the *in vitro* SPP1 replication system can use both viral G36P and host SsbA similarly. SPP1 rolling circle reaction requires an SSB with a C-terminal tail (Supplementary Figure S6) that is the site of interaction for almost all proteins that bind SSB (38). To test if the SPP1 phage requires the SsbA protein *in vivo*, we analyzed the levels of amplification of the SPP1 phage in a *B. subtilis* mutant strain, FLB22, which expresses from its natural promoter SsbA truncated of its last 35 amino acids. *ssbΔ35* cells and an isogenic control were infected at 30°C, and after two hours of infection the phage titer was determined. *ssbΔ35* cells are temperature-sensitive for growth above 47 °C in LB medium (34). Cultures were also grown at a permissive temperature until OD$_{560nm}$=0.2 and then shifted to 50°C. After 15 min incubation, the cells were infected with SPP1 and the cultures were incubated for 2 h at 50°C. Titrations of three independent experiments yielded a mean of 3x10$^9$ phage/ml and 1x10$^9$ phage/ml for infections of *ssbΔ35* cells at permissive and non-permissive temperature, respectively. 3x10$^9$ and 4x10$^9$ phage/ml titers were obtained when the isogenic strain
(ssb3+) was infected at the permissive and non-permissive temperature, respectively.

**The DnaC helicase unwinds DNA in the presence of G36P**

The preceding results show that G36P is not able to support replication in the full *B. subtilis* replication system. In an attempt to identify the defective step, we tested whether the DnaC helicase can be assembled onto DNA and unwind DNA efficiently in the presence of G36P. We exploited the fluorescence-based assay (Figure 6A). The results showed that SPP1 G36P and SsbA are interchangeable in the helicase assays (Figure 6B). So, the defect in the *B. subtilis* assay in the presence of G36P occurs after helicase loading and DNA unwinding.

**G36P stimulates synthesis by DnaE, but not by the PolC holoenzyme**

Continuing our search for the defect in *B. subtilis* chromosomal replication in the presence of G36P, we tested for its ability to stimulate reactions catalyzed by the DnaE and PolC holoenzymes. We observed that SsbA and G36P stimulate the DnaE holoenzyme with similar efficiency (Figure 7A). However, G36P failed to stimulate the PolC holoenzyme (Figure 7B). The same defect was observed in RNA primer extension reactions conducted with DnaE in the presence of PolC, which mimicked the reactions occurring at the lagging strand of the replication fork (Supplementary Figure S7). This suggests that the defect in the G36P-supported *B. subtilis* reaction likely resides in its inability to stimulate the PolC holoenzyme.

**G36P blocks host DNA replication**

The observation that G36P does not support *B. subtilis* replication led us to investigate if this could also occur in the physiological context, when both host and phage SSBs are present in the same reaction. *B. subtilis* replication forks were assembled in the presence
of saturating concentrations of SsbA₄ and increasing concentrations of G36P₄. Both leading and lagging strand synthesis were significantly inhibited even at the lowest G36P concentration tested (30 nM) (Figure 8).
Discussion

We have reconstituted an efficient rolling circle SPP1 reaction that recapitulates the concatemeric phase of SPP1 DNA replication. The reaction requires all proteins defined by genetic studies including the phage helicase (G40P), its loading protein (G39P) and the host DnaG primase and PolC DNA polymerase. We also observed a requirement for the G36P protein. Neither the \textit{B. subtilis} helicase loading proteins nor the helicase are required, but all of the polymerase elongation components are necessary, including \(\beta_2\), the \(\tau\) complex and a second DNA polymerase III, DnaE. The rate of elongation is similar to the rate of \textit{B. subtilis} replication under the conditions employed.

In the well characterized \textit{E. coli} system, DnaG primase binds to the DnaB_6 helicase before the synthesis of each primer and then dissociates (39). Because of this equilibrium, high concentrations of primase lead to more frequent priming and shorter Okazaki fragments. A similar observation has been made in a reconstituted \textit{B. subtilis} DNA replication system (30) and in the SPP1 system, presumably for the same reasons. DnaE has been previously shown to be required to extend RNA primers a short distance before handing them off to PolC, analogous to the process in eukaryotes where Pol \(\alpha\) and the Pol \(\delta\) holoenzyme are required (5,6,30). Thus, it is possible that DnaE could also influence primer synthesis due to an interaction with primase. We tested this possibility and found it not to be the case. Okazaki fragment length did not vary over a wide range of DnaE concentrations in the SPP1 system, as may also occur in the bacterial systems.

These studies identified a requirement for the G38P origin binding protein for a replicative reaction initiating on forks that mimic the concatemeric stage of SPP1 replication. In SPP1 this type of replication \textit{initiates after recombinational events} [after
the processing of the D-loop formed by fork stalling, (10)]. Therefore, such reactions
might mimic the replication restart reactions that occur on stalled forks in chromosomal
systems that require PriA for their initiation (40). PriA was an essential component of
the in vitro rolling circle replication with the B. subtilis system (30), but SPP1
replication does not require PriA, consistent with genetic data (18). Thus, it appears that
G38P plays a dual role as an origin-specific and forked structure-specific initiation
protein. Future studies directed at common features of PriA- and G38P-supported
replication initiation may provide insight regarding the mechanism used by these
proteins to drive the replication restart reaction in cellular systems. In B. subtilis, the
loading of the replicative helicase DnaC at oriC relies on the ordered associations of
DnaA, DnaB, DnaD and DnaI proteins (41) and outside of the origin relies on the PriA,
DnaB, DnaD and DnaI proteins, also through a cascade of protein interactions that are
crucial for the loading of the replicative helicase on SSB-coated ssDNA (42). The
results of the current study show that in SPP1 this reaction is restricted to two proteins
(G38P and G39P).

We found that both SPP1 G36P and B. subtilis SsbA supported the SPP1
replicative reaction, but with interesting differences. Under standard reaction
conditions, approximately 180 nM SsbA4 was required to achieve an optimal rate of
leading and lagging strand synthesis. However, with G36P4, an optimum rate is
achieved at a 6-fold lower concentration. Further increases of G36P concentration result
in marked inhibition of lagging strand synthesis. The most straightforward explanation
for this observation is that G36P makes specific protein-protein interactions that
sequester some viral component in inactive binary complexes. Since elevated levels of
G38P reverse the lagging strand inhibition induced by elevated G36P, it is the most
likely candidate as the sequestration target. Future work needs to be done to unravel the
cause of these differences. It is interesting that G38P is involved in both loading the
replicative helicase and reversing the inhibition by high G36P concentrations
specifically for the lagging strand. This dual activity suggests an additional function that
is required for ongoing lagging strand replication after helicase loading.

In contrast to the above observations, the phage-encoded SSB, G36P, will not
support a full host replisome-dependent replicative reaction. Dissection of the individual
reaction steps shows that G36P is fully competent to support PriA/DnaD/DnaB/Dnal-
dependent helicase loading and ensuing advancement to separate DNA strands. G36P is
also able to stimulate a DnaE holoenzyme activity as well as the host SSB. The only
defect detected was a failure of G36P to stimulate the major Gram-positive replicase,
the PolC holoenzyme. Yet, the PolC holoenzyme supports efficient replication in the
fully reconstituted SPP1 protein-dependent reaction. This suggests the presence of
additional or stronger PolC holoenzyme/phage protein interactions that either stimulate
PolC, protect it from inhibition, or bypass its requirement for lagging strand replication.

The lack of support by G36P of the *B. subtilis* replication fork and the efficient
use in the SPP1 replication fork of the two SSBs (the viral G36P and the host SsbA)
could be used *in vivo* as a strategy of the phage to amplify its own DNA while
inhibiting host chromosomal replication. This hypothesis is supported by our
observation that even at saturating amounts of SsbA, the *B. subtilis* replication system
was inhibited by low concentrations of G36P. Examples of viral inhibitors that block
host proliferation by binding to host proteins have been reported (43-45). But, a
strategy based on inhibition by a viral encoded SSB has not been reported to date. It is
interesting to note that many viruses encode SSB proteins (46). These include
*Staphylococcus aureus* phage 80α, which transfers pathogenicity islands between
staphylococci. This phage also encodes a G38P-like protein (47).
G36P is 48% identical to *B. subtilis* SsbA. Identity in the C-terminus, which is important for most SSB-protein interactions, is even higher (~75%). The most critical residues required for SSB-protein interactions are the Pro-Phe pair found at the extreme C-terminus (38). The C-terminal Phe of the *E. coli* homolog fits into a hydrophobic pocket of the $\chi$ subunit of the *E. coli* DNA polymerase III holoenzyme and the carboxyl group of this residue forms a salt bridge (48). Slightly more internal within SSB are three conserved acidic residues that form ionic bonds with positively charged residues in SSB binding sites (49). We note that the C-terminus of G36P contains two of these three acidic residues. In other systems, mutation of only one of the C-terminal three acidic residues of SSB results in only a modest decrease in binding to interacting proteins (38,49). Thus, explanation of the differential effects observed between G36P and SsbA on PolC holoenzyme stimulation and the specific inhibition of SPP1 replication at high G36P levels may require additional interactions outside of the prototypical C-terminal SSB interaction sequences.

Viruses that encode their own initiation machinery but depend on the host for elongation functions have provided significant insight into cellular processes (4,5,7,50). With the simple single-stranded DNA coli-phages, the only viral protein that is required for replication is a nicking/religation activity that functions after duplex formation to shift SS -> RF to duplex DNA replication (51). The bacteriophage $\lambda$ encodes its own origin binding protein and helicase loader, and subverts the host helicase (7). Thus, it differs from SPP1 in that it does not encode a helicase or an SSB. However, important similarities are also noted. $\lambda$, like SPP1, encodes an origin-specific initiation protein (O protein) that is not a DnaA homolog. The O protein, like SPP1 G38P, can serve to initiate DNA synthesis on artificial replication forks that lack origin sequences (15). SV40, like SPP1, encodes its own origin binding protein and helicase, both contained
within the same polypeptide chain as the T-antigen (52). This system has been invaluable in revealing the mechanisms of eukaryotic DNA replication (4,5).

The best biochemically defined Gram-positive bacteriophage is Φ29 (53). It encodes its own replication machinery and has provided a wealth of tools and insight, varying from basic models for viral replication that extend to eukaryotic adenoviruses (54) to tools for biotechnology (55). Before this work, an *in vitro* replication system for a Gram-positive phage that is dependent on host proteins had not been established.

Thus, the present study provides a new viral window into Gram-positive replication processes. The availability of a hybrid assay that exploits the efficiency and simplicity of initiation by phage-encoded proteins and a requirement for host proteins for elongation provides a powerful tool to support studies of fork dynamics, macromolecular interactions and regulation or replicative processes in Gram-positive organisms. Having two SSBs that can be used interchangeably with differential effects should help unravel the importance of SSB interactions in replication fork processes. And, having the G38P protein that supports initiation on both defined duplex origins and replication forks may provide an opportunity to learn more about the mechanistic features common to these two important processes.

**Supplementary data**

Supplementary Data are available at NAR online: Supplementary Materials and Methods, Supplementary Reference [56], and Supplementary Figures S1-S7.

**Acknowledgements**
We thank Gloria E. García-Baca, and Begoña Carrasco for providing us with purified G36P and SsbA, Maria López for excellent technical assistance, and Patrice Polard and Juan C. Alonso for providing us with strains. ALP received a contract from the Community of Madrid. CSM sabbatical leave was supported, in part, by the Spanish Ministry of Education (SAB20111-0026) and by the Burroughs Wellcome Fund 2012 Collaborative Research Travel Grant. JZ received partial stipend support from NIH Molecular Biophysics Training Grant T32 GM65013.

Funding

This work was partially supported by Spanish Ministry of Economy and Competitiveness [BFU2009-09520 to SA], and the National Science Foundation [MCB-0919961 to CSM].
REFERENCES


FIGURE LEGENDS

Figure 1. Reconstitution of SPP1 rolling circle DNA replication with B. subtilis and SPP1 purified proteins. (A) 15% SDS-PAGE gel of purified SPP1 bacteriophage proteins used to reconstitute rolling circle DNA replication: G40P (49.8 kDa), G39P (14.6 kDa), G38P (29.9 kDa) and G36P (18 kDa). The purification of these proteins is described under Supplementary Materials and Methods. (B) Diagram of the DNA template used. It has an asymmetric G:C ratio (50:1) between the two strands, permitting quantification of leading and lagging strand synthesis by measuring radioactive dCMP and dGMP incorporation. (C) Addition of [α-32P]dCTP or [α-32P]dGTP allows the detection of products corresponding to leading and lagging strands synthesis, respectively, resolved on an alkaline agarose gel. (D) Activity and source of proteins involved in SPP1 rolling circle replication. (E) Protein requirements for SPP1 rolling circle DNA replication. The values represented are the mean of three independent experiments.

Figure 2. Progression of the reconstituted SPP1 replication fork. (A) Time course of the leading strand and lagging strand DNA synthesis. The values represented are the mean of three independent experiments. (B) Rate of SPP1 replication fork progression. Standard reactions were scaled up to a volume of 125 µl and incubated at 37 °C with [α-32P]dCTP to detect leading strand synthesis. After the first minute (indicated as time 0), aliquots of 12.5 µl were removed and quenched every 10 s. (C) Quantification of the rate of SPP1 fork movement. Plot of the largest DNA fragment present versus time to determine replication fork rate.

Figure 3. Effect of increasing DnaG and DnaE on the size of Okazaki fragments. (A) Size of Okazaki fragments varies inversely with primase concentration. Rolling circle reactions were performed in the presence of [α-32P]dGTP, and were stopped after 10
min. The reactions were run on an alkaline agarose gel. Primase concentrations assayed varied in a 2-fold dilution series from 100 nM down to 0.4 nM. (B) Variations in DnaE concentrations do not alter the size of the Okazaki fragments. Reactions were performed changing the DnaE concentration in a 2-fold dilution series from 80 nM to 1.25 nM.

**Figure 4.** Interchangeable components of the replication machinery of *B. subtilis* and SPP1. (A) The SPP1 replication fork can use SsbA or G36P. *In vitro* replication of reconstituted SPP1 replication forks assembled in the presence of 30 nM G36P₄ (green box), 30 nM or 90 nM SsbA₄ (grey boxes), and in the absence of an SSB (white box). The presence of G39P (helicase loader) and G38P (origin binding) or the *B. subtilis* loading system (formed by DnaB-DnaD-DnaI-PriA) is indicated. (B) The *B. subtilis* replication fork strictly requires SsbA and its own helicase loading system. *In vitro* replication of reconstituted *B. subtilis* replication forks assembled in the presence of 30 nM G36P₄ (green box) or 90nM SsbA₄ (grey boxes) and the indicated components. The data shown are the mean of at least three independent experiments.

**Figure 5.** Increasing concentrations of G36P protein inhibit lagging strand DNA synthesis; synthesis can be restored by increasing G38P. (A) G36P₄ titration in SPP1 replication fork reconstitution. Standard reactions to measure DNA synthesis at leading and lagging strands were performed at the indicated G36P₄ concentrations. (B) High SsbA₄ concentrations do not inhibit lagging strand synthesis. (C) Effect on lagging strand synthesis of increasing G38P at high G36P₄ (180 nM). The values represented are the mean of three independent experiments.

**Figure 6.** Either SsbA or G36P will support *B. subtilis* helicase loading and unwinding. (A) DNA substrate used in unwinding reactions. The fluorescence of tetrachlorofluorescein (TET) on the 5’ terminus increases upon being separated from a quencher (BHQ-1) on the opposing strand. The 5’-end of the lagging strand template is
blocked by attachment of streptavidin to an incorporated biotinylated thymidine to inhibit helicase self-loading by threading over a free 5’-end. (B) *B. subtilis* helicase activity in the presence of SsbA or G36P. DNA substrate was combined with varying amounts of either SsbA or G36P, and 150 nM PriA, 75 nM DnaD, 75 nM DnaB, 12 nM DnaC, and 50 nM DnaI.

**Figure 7.** SPP1 G36P stimulates DnaE holoenzyme but not PolC holoenzyme. (A) Extension of a DNA primer by the DnaE holoenzyme using either SsbA or G36P. SsbA or G36P were titrated in a DNA-primed M13ori reaction containing 3 nM DnaE in the presence (+) or absence (-) of β2 as indicated. (B) Extension of a DNA primer by the PolC holoenzyme (2 nM) using either SsbA or G36P. Total DNA synthesis was analyzed by quantification of [3H]dTTP.

**Figure 8.** G36P inhibits *B. subtilis* rolling circle DNA replication even when SsbA is present at saturating concentrations. *In vitro* replication of reconstituted *B. subtilis* replication forks assembled in the presence of increasing concentrations of G36P4 (30, 60, 90, 180 and 270 nM tetramer) and 90nM SsbA4. Both SSBs were added simultaneously to the substrate mix. The data shown are the mean of at least three independent experiments.
Figure 1
168x124mm (300 x 300 DPI)
Figure 2

86x222mm (300 x 300 DPI)
Figure 3
96x84mm (300 x 300 DPI)
Figure 4

82x138mm (300 x 300 DPI)
Figure 5

DNA synthesis (pmol)

A

Leading strand  Lagging strand

G36P (nM)

B

Leading strand  Lagging strand

SsbA (nM)

C

DNA synthesis (pmol)

G3BP (µM)

87x216mm (300 x 300 DPI)
Figure 6
Figure 7
Figure 8
79x58mm (300 x 300 DPI)
**Supplementary Information**

**Bacteriophage SPP1 DNA replication strategies promote viral and disable host replication**

Elena M. Seco¹, John Zinder², Carol M. Manhart², Ambra Lo Piano¹, Charles S. McHenry*² and Silvia Ayora*¹

¹Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain

²Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309, USA

**Supplementary Materials and Methods**

**Purification of SPP1 DNA Replication Proteins**

All proteins were expressed in *E. coli* overexpression vectors in which the encoding genes are the native genes without added tags. *B. subtilis* PriA, DnaD, DnaB, DnaI, DnaC helicase, DnaG primase, DnaE, PolC, τ, δ, δ', β and SSB were purified as described (30). G40P was purified using pCB367 plasmid (25). pCB367 is a pQE-11 (Qiagen) derivative containing wild-type gene 40 and the gene 39 carrying a hexa-histidine tag on its N-terminus, which allows the affinity purification of the native G40P. His₆-G39P stabilizes G40P and forms a complex with it only in the presence of ATP. Hence, in the presence of ATP, the co-expressed His₆-G39P-G40P complex is retained in a Ni-NTA column. Cells were grown at 37 °C to an OD₅₆₀ = 0.8, induced with 2 mM IPTG for 2 h, and harvested and stored at -20 °C. Five g of cells were suspended in 25 ml buffer A [50 mM sodium phosphate (pH 7.5), 10% (v/v) glycerol, 1
mM MgCl₂, 0.1 mM ATP, 1 mM b-mercaptoethanol, 300 mM NaCl], lysed by sonication and centrifuged at 18,000 rpm in an SS-34 rotor for 45 min. The supernatant carrying G₄₀P and G₃⁹P was brought to a concentration of 20 mM imidazole and loaded onto a Ni-NTA column (2 ml) equilibrated in buffer A carrying 300 mM NaCl and 20 mM imidazole. A gradient of fifty column volumes of buffer A containing 20–500 mM imidazole and in the absence of ATP was applied and the last fractions, which contained G₄₀P enriched with respect to G₃⁹P, were pooled and dialyzed against buffer B [50 mM Tris HCl (pH7.5), 10% (v/v) glycerol, 1mM MgCl₂, 50mM NaCl, 1mM b-mercaptoethanol] to be then loaded onto a Q-sepharose column. A linear gradient from 50 mM to 500 mM NaCl was applied and fractions enriched in G₄₀P protein were pooled and loaded again onto a new Ni-NTA column (0.5 ml) in which the rest of the G₃⁹P forming a complex with G₄₀P was retained. The flow-through containing pure G₄₀P was dialyzed against buffer B and loaded onto a Q-sepharose column (0.5 ml) to elute 1 mg of G₄₀P in the storage buffer [50 mM Tris HCl (pH 7.5), 50% (v/v) glycerol, 1 mM MgCl₂, 500 mM NaCl, 1 mM β-mercaptoethanol]. His₆-G₃⁹P was discarded because the His-tag affected the activity of the protein. To purify untagged G₃⁹P, plasmid pBT318 (17) was transformed into E. coli BL21(DE3)/pLysS, and cells were induced with 2 mM IPTG for 2 h, harvested and stored at -20 °C. A total of 7.5 g of cells were suspended in 37.5 ml buffer C [50 mM Tris HCl (pH 7.5), 10% (v/v) glycerol, 1 mM dithiothreitol] containing 250 mM NaCl, lysed by sonication and centrifuged at 18,000 rpm in an SS-34 rotor for 30 min. Polyethylenimine (10% v/v, pH 7.5) was slowly added to the supernatant containing G₃⁹P to a final concentration of 0.25% (A₂₆₀ = 120). The DNA and certain contaminating proteins were pelleted by centrifugation (12,000 rpm in an SS-34 rotor, 10 min) and the pellet was discarded. Then proteins of the supernatant were precipitated by addition of solid ammonium
sulfate to a final concentration of 80% saturation. The pellet containing G39P was resuspended in buffer C without salt and dialyzed against buffer C with 10 mM NaCl to be then loaded onto a Q-Sepharose column (2 ml). The protein eluted from the column in the 25 mM and 50 mM washes. These fractions were pooled and brought to 1.2 M ammonium sulfate to be then loaded onto a phenyl-Sepharose column. Serial washings of ten column volumes with buffer C and decreasing concentrations of ammonium sulfate from 1.2 to 0 were applied (1.2, 1, 0.8, 0.6, 0.4, 0.2 and 0 M ammonium sulfate) and the pure protein eluted in the 0.4 M and 0.2 M washing steps. Both fractions containing G39P were pooled and dialyzed against buffer B with 10 mM NaCl. A small Q-sepharose column (0.5 ml) was used to concentrate and elute 3 mg of G39P in the storage buffer (50 mM Tris HCl (pH 7.5), 50% (v/v) glycerol, 150 mM NaCl, 1 mM dithiothreitol).

G38P was purified from *E. coli* BL21(DE3)/pLysS cells carrying the plasmid pBT320 as described (17). G36P was purified from *E. coli* BL21(DE3)/pLysS cells carrying the pCB596 plasmid. Plasmid pCB596 is a pET-3a (Novagen) derivative where gene 36 was cloned into the NdeI and BamHI restriction sites. Cells were grown to an OD$_{560}$ = 0.8 at 37 °C, 2 mM IPTG was added, and after 120 min cells were harvested by centrifugation and stored at -20 °C. Five g of cells were resuspended in 25 ml buffer D [50 mM Tris HCl (pH 7.5), 1 mM dithiothreitol, 15% (v/v) glycerol] containing 150 mM NaCl. Lysis was accomplished by sonication and cell debris was removed by centrifugation (18,000 rpm in an SS-34 rotor, 30 min). DNA and G36P were precipitated from the supernatant by addition of polyethylenimine (final concentration 0.25% v/v with A$_{260}$ = 120) and centrifugation at 12,000 rpm in an SS-34 rotor for 10 min at 4 °C. The protein was solubilized from the pellet in 25 ml buffer D containing 300 mM NaCl, and then precipitated by addition of ammonium sulfate to
30% saturation. Protein was redissolved in 25 ml of buffer D containing 300 mM NaCl and reprecipitated with 30% ammonium sulfate, rendering the protein almost pure. The pellet was resuspended in 25 ml buffer D containing 50 mM NaCl and extensively dialyzed. After dialysis, the sample was loaded onto a 1 ml Q-Sepharose column equilibrated with the same buffer. Serial washings of increasing concentrations of NaCl in ten column volumes were applied to the column and the pure protein eluted in the 150 and 200 mM NaCl washings steps. The pure protein was concentrated in another Q-Sepharose column and eluted in a single step with buffer D containing 500 mM NaCl. The G36P protein was dialyzed against storage buffer [50 mM Tris HCl (pH 7.5), 1 mM dithiothreitol, 50% (v/v) glycerol, 300 mM NaCl] and stored at -20 °C. The identity of the protein and the absence of *E. coli* SSB in the preparation were confirmed by MALDI-TOF. Plasmid pCB596 was used as a template to generate G36P variants lacking the last 9, 15 or 21 C-terminal residues that were purified following a protocol similar to the used for the wild type protein. The protein concentration was determined in all cases as previously described by measuring the absorbance of the peptide bond (56).

**Supplementary Reference**

Supplementary Figure S1. Alignment of the G36P and SsbA proteins. Identical residues have a black background and conserved residues gray.
Supplementary Figure S2. Optimization of SPP1 rolling circle replication. The reactions and the measurement of DNA synthesis on both leading and lagging strands were performed as described in Experimental Procedures. All the titrations were carried out in the presence of saturating levels of the other assay components and are the mean of at least two independent experiments. Titrations of (A) G39P, (B) G38P, (C) PolC, (D) DnaE, (E) and DnaG. An arrow indicates the concentration of protein chosen for subsequent experiments.
Supplementary Figure S3. The *Bacillus subtilis* replisome can not be reconstituted in the presence of G36P. *B. subtilis* replication reactions were assembled with all host components except that SsbA₄ was replaced by increasing concentrations of G36P₄.

Supplementary Figure S4. Protein requirements at high (180 nM) concentrations of G36P₄. Leading and lagging strand synthesis were quantified by [α-³²P]dCTP or [α-³²P]dGTP incorporation, respectively. The values represented are the mean of three independent experiments.
Supplementary Figure S5. Optimization of SPP1 helicase assay. The substrate (see Figure 6A for diagram) was combined at 20 nM with SPP1 helicase and helicase loading proteins using the conditions described under Materials and Methods for reactions with B. subtilis proteins. Helicase assembly proteins were titrated one at a time to optimize unwinding. (A) G38P titrated using 120 nM G39P, 15 nM G40P₆, and 100 nM G36P₄. (B) G39P titrated using 200 nM G38P, 15 nM G40P₆, and 100 nM G36P₄. (C) G40P titrated using 200 nM G38P, 500 nM G39P, and 100 nM G36P₄. (D) G40P requires an SSB to unwind forked substrates. G36P or SsbA titrated using 200 nM G38P, 500 nM G39P, and 10 nM G40P₆. The arrows indicate the concentration of protein chosen for subsequent experiments.
Supplementary Figure S6. The C-terminal tail of G36P is required for efficient in vitro replication. Leading and lagging strand synthesis were quantified by [α-\lowercase{32}P]dCTP and [α-\lowercase{32}P]dGTP incorporation, respectively, in the presence of the serial C-terminal deletion mutants at low (30 nM) or high (270 nM) SSB_4 concentrations. The values represented are the mean of three independent experiments.
Supplementary Figure S7. SsbA, but not G36P, stimulates RNA primer extension under conditions that require a handoff of the extended primer from DnaE to PolC in a reaction containing only *B. subtilis* proteins. Reactions were conducted as in Fig. 7 but contained RNA primers that cannot be elongated by PolC alone. (A) Extension of an RNA primer by both DnaE and PolC replicases using either SsbA or G36P. SsbA (blue and red lines) or G36P (green and black lines) was titrated in a reaction containing 0.5 nM DnaE and 2.5 nM PolC in the presence or absence of β2 as indicated. Reactions were incubated at 30 °C for 3 min. (B) Extension of an RNA primer by the PolC holoenzyme using either SsbA or G36P. Reactions were conducted as in (A) except DnaE was omitted from the reaction. (C) Extension of an RNA primer by the DnaE holoenzyme using either SsbA or G36P. Reactions were conducted as in (A) except PolC was omitted from the reaction.