Phage-Borne Factors and Host LexA Regulate the Lytic Switch in Phage GIL01

Nadine Fornelos,1,2 Jaana K. H. Bamford,2* and Jacques Mahillon1

Laboratory of Food and Environmental Microbiology, Université Catholique de Louvain, Croix du Sud, 2, Box L7.05.12, B-1348 Louvain-la-Neuve, Belgium,1 and Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä, P. O. Box 35, 40014 Jyväskylä, Finland2

Received 20 June 2011/Accepted 25 August 2011

Temperate phages are defined by their ability to infect a susceptible host cell and opt between either of two developmental pathways, the lytic or the lysogenic cycle. In the lytic cycle, the phage genome is replicated and expressed to produce a large number of virions that are released from the host cell. Once engaged, this pathway inevitably ends with the death of the host cell and the release of a crop of newly formed virions in the environment. In contrast, the lysogenic cycle is defined by the tight repression of lytic behavior, and the phage genome replicates along with that of the host cell, either physically integrated into the host genome or as an independent plasmid. The quiescent prophage is thereby efficiently inherited as cells divide, until the lytic cycle is reactivated in response to alterations in host cell physiology. This transition requires the efficient interaction of regulatory components that control the passage from a highly stable standby mode to multiple rounds of genome replication, vigorous gene expression, virion assembly, and host lysis.

In the well-studied bacteriophage λ, establishment of lysogeny relies on the synthesis of CI, a transcriptional regulator that directs the expression of the lytic cycle repressor CI (15). Once produced, CI dimers bind cooperatively at operator sites in the so-called switch region to prevent transcription from the early promoters $P_L$ and $P_R$ and, subsequently, due to the lack of early gene products, from the late lytic promoters as well.

Expression of CI alone is sufficient to maintain lysogeny once it has been established (26). Despite an intrinsically stable quiescent state, the λ prophage can efficiently switch to an alternate lytic cycle. The λ CI repressor functions analogously to the bacterial SOS response master repressor LexA, with which it shares significant sequence similarity (21, 45). LexA binds as a dimer to specific operator sites and negatively regulates the expression of a host of genes that are essentially involved in DNA repair and cell survival (10, 18). Upon DNA damage, such as that inflicted by UV irradiation or mitomycin C (MMC) treatment, the host recombinase RecA is converted to an activated form when it binds single-stranded DNA. Activated RecA then acts as a coenzyme to stimulate LexA autoregulation and, through a similar process, CI autoregulation as well (43, 59). As a consequence, LexA dissociates from its consensus sites to initiate the SOS response. Likewise, CI is no longer able to bind its operator sites in the λ genome, leading to lytic gene expression and viral proliferation.

Alternative mechanisms of regulation have been described for several SOS-inducible phages. The repressors of coliphages 186 and N15, for example, are not RecA sensitive but are instead antagonized by phage-borne antirepressors that are under LexA control (47, 61). In CTXΦ, the filamentous phage that encodes cholera toxin in Vibrio cholerae, host LexA and the phage repressor function as both activators and repressors of gene expression to ensure the permanent production and secretion of CTXΦ (36).

The Bacillus thuringiensis prophage GIL01 has been shown to be induced by DNA-damaging treatments such as UV irradiation and MMC (69). The 15-kb linear prophage does not integrate into the bacterial chromosome but instead replicates as a stable autonomous replicon within the cell. Similar to that of the lambdoid prophages, the lytic cycle of GIL01 is induced as part of the cellular SOS response to DNA damage. However, no CI-like maintenance repressor has been detected in the phage genome, suggesting that GIL01 uses a novel mechanism to maintain lysogeny. To gain insights into the GIL01 regulatory circuit, we isolated and characterized a set of $cp$ mutants that are unable to lysogenize. Two phage-encoded proteins, gp1 and gp7, are required for stable lysogen formation. Analysis of $cp$ mutants also identified a 14-bp palindromic dinBox1 sequence within the $P1$-$P2$ promoter region that resembles the known LexA-binding site of Gram-positive bacteria. Mutations at conserved positions in dinBox1 result in a $cp$ phenotype. Genomic analysis identified a total of three dinBox sites within GIL01 promoter regions. To investigate the possibility that the host LexA regulates GIL01, phage induction was measured in a host carrying a noncleavable lexA (Ind−) mutation. GIL01 formed stable lysogens in this host, but lytic growth could not be induced by treatment with mitomycin C. Also, mitomycin C induced β-galactosidase expression from GIL01-lexA promoter fusions, and induction was similarly blocked in the lexA (Ind−) mutant host. These data support a model in which host LexA binds to dinBox sequences in GIL01, repressing phage gene expression during lysogeny and providing the switch necessary to enter lytic development.

* Corresponding author. Mailing address: Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä, P. O. Box 35, 40014 Jyväskylä, Finland. Phone: 358 14 260 4160. Fax: 358 14 260 2221. E-mail: jaana.bamford@juu.fi.

1 Published ahead of print on 2 September 2011.

* Corresponding author. Mailing address: Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä, P. O. Box 35, 40014 Jyväskylä, Finland. Phone: 358 14 260 4160. Fax: 358 14 260 2221. E-mail: jaana.bamford@juu.fi.
independently within its host and occasionally experiences induction of its lytic cycle. GIL01 shares a similar genome organization and size with PRD1, the model tectivirus infecting Gram-negative hosts. However, unlike PRD1 and its close siblings, which are all lytic phages, GIL01 is able to enter a dormant state during which its lytic functions are turned off. Two additional tectiviruses infecting B. thuringiensis have also been described: B. cereus, which is almost identical to GIL01 (55, 64) and has been the object of many studies on the internal membrane that defines tectiviruses (20, 39–41), and B. anthracis, which shares considerable sequence identity (83.6%) with GIL01 (67). A more distant relative, PBclin15, has been sequenced alongside the B. cereus reference strain ATCC 15479 genome but is not yet known to be a prophage (29). Recently, the sequencing of AP50, a temperate phage highly specific to B. anthracis strains, added a new member to the subgroup of Tectiviridae that proliferates among members of the B. anthracis group (62).

By analogy to other temperate phages, it has been assumed that GIL01 controls its lysogenic cycle by expressing a lytic gene repressor. A potential candidate for this function was established by the lysogenic cycle—were previously mapped to the N-terminal LexA-like DNA-binding domain (67). However, unlike the X phage CI repressor, ORF6 lacks any recognizable protease linker required for repressor cleavage or the C-terminal dimerization domain. Furthermore, spontaneous GIL01 clear plaque (cp) mutants—which are unable to establish the lysogenic cycle—were previously mapped to ORF7, but no such mutations were found to affect ORF6 (67). In this study, we demonstrated that ORF7 is indeed involved in the regulation of the phage cycle and that it does so along with ORF1. Importantly, we also showed that GIL01 is induced in response to SOS signals due to direct control by the host LexA repressor. LexA likely binds to a conserved operator, termed dinBox1 (for damage-inducible Box), that when mutated commits GIL01 to propagate exclusively as a virulent phage. dinBox1 lies within two promoters, P1 and P2, that direct the expression of regulation and replication functions that are likely important for lysogenic development and/or maintenance. Two additional LexA-binding sites, dinBox2 and dinBox3, were identified at promoter P3 regulating the downstream structure and lysis module. Curiously, no cp mutations were ever isolated that carried mutations within these operator sites, possibly indicating that dinBox2 and dinBox3 are not required for lysogenic regulation but are important for lytic growth. Altogether, we identified one cis-acting element (dinBox1), two phage-borne genes (ORF1 and ORF7), and a host regulator (LexA) that determine passage from the latent state to lytic development in GIL01. The identification of these factors will ultimately help elucidate the gene circuit that regulates GIL01 development.

**MATERIALS AND METHODS**

**Bacterial strains, phages, and plasmids.** The natural host of phage GIL01 is the insect pathogen *B. thuringiensis* subspecies *israelensis* strain NB31 (2). The permissive plasmid-free strain GBJ002 (31), derived from strain G02 and related to NB31 (32), was used to propagate GIL01. In order to work in identical genetic backgrounds, we generated GBJ002 lysogens as follows. GIL01 filter-sterilized suspensions were spotted onto 0.7% top agar lawns seeded with GBJ002. Plates were incubated overnight at 37°C. The center of a lyed spot was then recovered with a sterile inoculation loop and streaked on LB agar. After 16 h at 37°C, individual colonies had grown and were screened by PCR with GIL01-specific primers to confirm lysogeny. A positive colony (GBJ338) was selected and retained for the study.

GIL01 cp mutants were obtained spontaneously by infecting GBJ002 with wild-type GIL01 phage suspensions. Briefly, phage preparations from serial 10-fold dilutions were added to 200 μl of mid-log-phase GBJ002 and incubated at room temperature for 30 min. The infected cells were then plated on LB agar with 4 ml of molten 0.7% top agar. The plates were incubated overnight at 37°C and visually scored for the presence of clear plaques. Single clear plaques, isolated from independent experiments, were picked with microneedle tips and subjected to three rounds of single-plaque purification on host GBJ002. The wild-type pGIL01 genome sequence is available from the NCBI GenBank database under the accession number AJ586073.

**Strain GBJ499 (a recA null mutant)** was constructed by disrupting the recA gene of strain GBJ002 with the integrative vector pMutin4 (Bacllince Genetic Stock Center). pMutin4 is unable to replicate in *Bacillus* and is therefore particularly well suited for insertional mutagenesis in this host (66). Briefly, an internal fragment of recA (coordinates 3,719,072 through 3,720,089 in B. cereus strain G9842, GenBank accession number NC_011772) was amplified from strain GBJ002 with primers RecAmpf (5’-GGGCGGCCGCTCTTCACCAATTCCGTGAATGTTACAATGCCGCTTG-3’) and RecAmpr (5’-GAATTCATTTGGTAAAGGTTCAATTA-3’). The PCR product was digested with NotI and EcoRI (restriction sites in underlined), purified, and cloned into NotI- and EcoRI-digested pMutin4 vector. The ligated was transformed into *Escherichia coli* (ER2925), and recombinant plasmids were extracted and introduced into competent GBJ002. *Bacillus* transformants, having integrated the vector by homologous recombination, were selected on 0.3 μg of erythromycin ml⁻¹. The insertion of pMutin4 in the desired locus was confirmed by sequence analysis of recA. The downstream gene, pcrA, encodes a phosphodiesterase (B1422), is separated from recA (B1421) by a 500-bp noncoding region, including a Rho-independent terminator. It is therefore unlikely that the insertion of pMutin4 in recA disrupts an operon.

**Strain GBJ396 [lexA4(A96D)] mutant** was constructed by replacing an Ala (GCT) residue at position 96 in *lexA* with Asp (GAT). A 1.6-kb segment encompassing *lexA* (coordinates 3,623,569 through 3,625,169 in B. cereus strain G9842, GenBank accession number NC_011772) was amplified from strain GBJ002 with primers MutLex1 (5’-GCGATCTCGTTATATCTTCCACAAACT-3’) and MutLex2 (5’-AACGTTGTGGTTAGACACCTTGCTGA-3’) and cloned into the pCR4-TOPO vector (Invitrogen). The ligated was transformed into *E. coli* (TOP10), and recombinant plasmids were extracted and introduced into competent GBJ002. Similarly, *Bacillus* transformants, having integrated the vector by homologous recombination, were selected on 0.3 μg of erythromycin ml⁻¹. One of these recombinants was grown for several consecutive cycles of 4 h in liquid LB, and samples were recovered at each cycle and plated on LB agar at different dilutions. Single colonies were subjected to UV irradiation (wavelength, 254 nm) for 10 s and left at room temperature for 2 days. UV-sensitive colonies were amplified with primers MutLex1 and MutLex2, and the resulting PCR products were sequenced in order to screen for clones that had undergone a second recombination event resulting in the *lexA4(A96D)* genotype. Lysogens GBJ500 and GBJ406 were generated by infecting the mutant strains GBJ499 and GBJ396 with GIL01 as explained above.

*E. coli* strain ER2925 (New England Biolabs), a dam and dem mutant K12 derivative, was used to propagate unmethylated constructs prior to transformation into GBJ002. pCR4-TOPO-derived constructs were transformed into *E. coli* strain TOP10 (Invitrogen) or XL1-Blue (Stratagene). ORF1 and ORF7 were amplified with primers MutLex1 (5’-CTCCTGTAATTGGTAAACCA-3’) and MutLex2 (5’-CTCCTGTAATTGGTAAACCA-3’), and the resulting PCR products were sequenced in order to screen for clones with different mutagenic results (63). One of these recombinants was grown for several consecutive cycles of 4 h in liquid LB, and samples were recovered at each cycle and plated on LB agar at different dilutions. Single colonies were subjected to UV irradiation (wavelength, 254 nm) for 10 s and left at room temperature for 2 days. UV-sensitive colonies were amplified with primers MutLex1 and MutLex2, and the resulting PCR products were sequenced in order to screen for clones that had undergone a second recombination event resulting in the *lexA4(A96D)* genotype. Lysogens GBJ500 and GBJ406 were generated by infecting the mutant strains GBJ499 and GBJ396 with GIL01 as explained above.
Growth conditions and phage induction. *B. thuringiensis* and *E. coli* strains were grown in modified lysogeny broth (LB) (1% tryptone, 0.5% yeast extract, 0.5% NaCl) (6) at 37°C. *B. thuringiensis* liquid cultures were always grown overnight with shaking at 200 rpm before being diluted 1:100 in LB and allowed to grow exponentially. Transformation of *B. thuringiensis* and *E. coli* strains with the different constructs was performed by electroporation (46, 57), and transformants were selected on LB agar supplemented with 25 μg of kanamycin ml⁻¹ (pDG18-derived constructs), 50 μg of kanamycin ml⁻¹ (pCR4-TOPO-derived constructs), 25 μg of erythromycin ml⁻¹, and 100 μg of ampicillin ml⁻¹ (pHT304-18Z-derived constructs in *B. thuringiensis* and *E. coli*, respectively) or 0.3 μg of erythromycin ml⁻¹ and 100 μg of ampicillin ml⁻¹ (pMutin4-derived constructs in *B. thuringiensis* and *E. coli*, respectively).

GIL01 was induced from GB338, GB500, and GB406 by UV irradiation and MMC treatment. In both cases, exponential lysogenic cultures were pelleted at 3,000 × g for 7 min at 4°C, and the supernatants were recovered and filtered (0.45 μm). For MMC induction, MMC was added to 5 ml of resuspended cells at a concentration of 50 ng ml⁻¹. The plates were incubated for one additional hour, and samples (20 μl) from each different condition were taken and estimated by infecting GBJ002 in soft agar (0.7%).

Nucleotide modifications. Nucleotide modification enzymes were purchased from Roche and Fermentas. Phusion Flash high-fidelity PCR master mix (Finnzymes) and GoTaq Flexi DNA polymerase (Promega) were used in PCR amplifications. Oligonucleotides, antibiotics, mitomycin C, IPTG, and GenElute plasmid miniprep kits were purchased from Sigma-Aldrich. QIAquick gel extraction and PCR purification kits were purchased from Qiagen.

**Determination of transcription start sites.** Total RNA was extracted from mid-log-phase GB338 cultures with an Ambion RiboPure-bacteria kit and treated with DNase DNA-free solution (Ambion). Transcription start sites were determined by primer extension using 5′-end-labeled primers and processed as described (16), with some modifications. Bacterial RNA was pretreated with Roche and Fermentas. Phusion Flash high-fidelity PCR master mix (Finnzymes) and GoTaq Flexi DNA polymerase (Promega) were used in PCR amplifications. Oligonucleotides, antibiotics, mitomycin C, IPTG, and GenElute plasmid miniprep kits were purchased from Sigma-Aldrich. QIAquick gel extraction and PCR purification kits were purchased from Qiagen.

**lacZ fusions and β-galactosidase assays.** The region encompassing promoters P1 and P2 as well as dinBox1 (coordinates 39 through 413) was amplified using primers pH1 (5′-AAGCTTGAACATATAAATTATATTGAT and pH4 (5′-GGAATTCCTTCTTGTGTTAAGCGA). The corresponding product was purified and cloned into the selection vector pCR4-TOPO. The ligated product was transformed into competent DH5α and cultured on LB plates with 25 μg of kanamycin ml⁻¹. Control cultures were otherwise treated equally. GIL01 titers from induced and uninduced cultures were estimated by infecting GBJ002 in soft agar (0.7%).

**Sequencing and analysis sequences.** Sequences were sequenced at the Flanders Intenuniversity Institute for Biotechnology (VIB) Genetic Service Facility (University of Antwerp, Belgium) by double-strand primer walking, with individual quality control of sequence reads included. Sequences were analyzed with the footprint discovery program from RSAT (65) and followed the protocol as described for study case 2 (17). This algorithm was run on the lexA regulon of the reference strain ATCC 14579 (GenBank accession number AE016877) and the GIL01-specific primer pairs designed to cover the 15-kb genome in segments of 1 kb with 100-bp overlaps. PCR products were purified and sequenced with a BigDye Terminator v3.1 cycle sequencing kit on an ABI Prism 3100xl genetic analyzer instrument (both from Applied Biosystems). GenBank accession numbers for GIL01 sequences are NC_011772, with which GBJ002 shares most sequence identity.

**Phylogenetic footprinting.** To discover evolutionary conserved elements, we used the footprint discovery program from RSAT (Regulatory Sequence Analysis Tools) (65) and followed the protocol as described for study case 2 (17). This strategy has been systematically evaluated on *E. coli* and gives good results at different taxonomical levels for well-conserved genes such as lexA (30). The algorithm was run on the lexA and recA promoter sequences from the *B. cereus* reference strain ATCC 14579 (GenBank accession number AE016877) and the GIL01 promoters.

**RESULTS**

**The GIL01 genome has two major functional domains.** Like most phage genomes sequenced to date, the GIL01 genome contains closely packed genes with few intergenic spaces. All GIL01 genes are transcribed in the same rightward direction, 980 μl of buffer Z (containing 50 mM ß-mercaptoethanol) and 10 μl of toluene. After 5 min at 37°C, 200 μl of buffer Z supplemented with ß-nitrophenyl-ß-D-galactopyranoside (ONPG; 4 mg ml⁻¹) were added to each sample. The treated samples were incubated at 37°C for a time period ranging from 20 to 30 min, and reactions were stopped by adding 500 μl of NaCO₃ 1 M. The samples were centrifuged for 7 min at 13,000 rpm, and the supernatants were collected. Optical densities (ODs) were measured at a wavelength of 420 nm. The β-galactosidase activity values are expressed as Miller units and were calculated with the formula (1,000 × OD₅₄₀)/(OD₅₄₀ × time × [min] × [volume ml]).

**pDG1, pDG7, and pDG17 plasmid constructions and complementation tests.** ORF1 (coordinates 356 to 532) was amplified from GIL01 DNA with primers DG1F (5′-GCTCTAGAGTGAATAAATCTACTGCG) and DG1R (5′-CTGGCATCATTAATGCTGATTTT) and digested with XbaI and SphI and ligated into XbaI- and SphI-digested pDG148 plasmid DNA to generate pDG1. pDG7 was constructed similarly by using primer pairs DG7F (5′-GCTCTAGATAAGCAAGAATTGCTG) and DG7R (5′-GCTCTAGATCTGCATCATTCTTCC) (coordinates 4,564 through 4,716). An ORF7 from ORF7 transcriptional unit was generated by amplifying ORF1 with primers DG1_rbsF (5′-GAAAGCTTCAAGGAGGGATATG) and DG1_rbsR (5′-GCTCTAGATTATGTCATTTAATACTTTGTTTATTG) (coordinates 336 through 533, including the ORF1 ribosome-binding site shown in Fig. 1C in lowercase), restricting the PCR product with HindIII and XbaI, and cloning the purified product into HindIII- and XbaI-digested pDG148 plasmid, resulting in construct pDG1_rbs. ORF7 was amplified with primers DG7_rbsF (5′-GCTCTAGATAAGCAAGAATTGCTG) and DG7_rbsR (5′-GCTCTAGATCTGCATCATTCTTCC) (coordinates 4,549 through 4,717, including the ORF7 ribosome-binding site shown in lowercase). After digestion with XbaI and gel purification, the fragment was ligated into XbaI-digested pDG1_rbs plasmid, resulting in construct pDG1/7. pDG1, pDG7, and pDG17 ligation were transformed into competent *E. coli* (ER2925), and recombinant plasmids were extracted and confirmed by DNA sequence analysis. Competent GBJ002 was transformed with unmethylated pDG1, pDG7, and pDG17 (containing ORF7 in the correct orientation) to yield the clones GBJ250, GBJ223, and GBJ467, respectively. Protein expression was induced by adding IPTG to a final concentration of 1 mM to one-half of the liquid cultures in the early exponential growth phase. Control cultures were grown without the addition of IPTG. After an additional 3-h incubation at 37°C with shaking, 200 μl of induced cultures were infected with three 10-fold serial dilutions of *cp* mutant phage stocks and plated with molten top agar (0.7%) on LB plates supplemented with IPTG to a final concentration of 1 mM. The same treatment was followed with uninduced cultures, but they were plated on LB plates devoid of IPTG. The plates were incubated overnight at 37°C, and plaque counts were determined. Plaque pictures were taken with the Bio-Rad Molecular Imager Gel Doc XR system using Quantity One version 4.6.3. imaging software.

**Sequencing and analysis sequences.** Constructs were sequenced at the Flanders Intenuniversity Institute for Biotechnology (VIB) Genetic Service Facility (University of Antwerp, Belgium) by double-strand primer walking, with individual quality control of sequence reads included. Sequences were analyzed with the footprint discovery program from RSAT (65) and followed the protocol as described for study case 2 (17). This algorithm was run on the lexA regulon of the reference strain ATCC 14579 (GenBank accession number AE016877) and the GIL01-specific primer pairs designed to cover the 15-kb genome in segments of 1 kb with 100-bp overlaps. PCR products were purified and sequenced with a BigDye Terminator v3.1 cycle sequencing kit on an ABI Prism 3100xl genetic analyzer instrument (both from Applied Biosystems). GenBank accession numbers for GIL01 sequences are NC_011772, with which GBJ002 shares most sequence identity.

**Phylogenetic footprinting.** To discover evolutionary conserved elements, we used the footprint discovery program from RSAT (Regulatory Sequence Analysis Tools) (65) and followed the protocol as described for study case 2 (17). This strategy has been systematically evaluated on *E. coli* and gives good results at different taxonomical levels for well-conserved genes such as lexA (30). The algorithm was run on the lexA and recA promoter sequences from the *B. cereus* reference strain ATCC 14579 (GenBank accession number AE016877) and the GIL01 promoters.
and the ORFs are frequently overlapping, suggesting that GIL01 genes are organized into several operons (Fig. 1A). Interestingly, genes with similar functions cluster together, forming two functional modules. The left arm of approximately 4.8 kbp contains replication and regulatory genes, such as the predicted terminal protein (ORF4), the DNA polymerase (ORF5), and two proteins with predicted DNA-binding domains (ORF1 and ORF6). The larger downstream module (10 kbp) contains all the structural and assembly genes as well as two endolysins (ORF26 and ORF30) involved in cell wall degradation (68). Two noncoding segments stand out by their length and position in the genome as containing potential promoter/regulatory elements. The first one lies in the left extremity, upstream of ORF1 (355 bp), while the second one separates ORF8 from ORF9 (118 bp) (Fig. 1A). Based on this genetic organization, both noncoding segments were examined for the presence of promoters, using RACE PCR, a rapid method for the amplification of 5’-cDNA ends. The technique relies upon amplifying RNA transcripts between a defined internal site and the unknown 5’ extremity of interest, thus allowing for the identification of the transcription start site. Using RNA extracted from a lysogenic culture in the exponential growth phase, two transcription start sites were identified upstream of ORF1, at positions 175 and 278, and one initiation site was identified in the intergenic region between ORF8 and ORF9, at coordinate 4,915 (Fig. 1B). Centered approximately 10 and 35 nucleotides upstream of each transcription start site are conserved sigma A-dependent promoter sequences (Fig. 1A and C). The two potential promoters on the left were named \( P_1 \) and \( P_2 \), and the internal promoter, \( P_3 \). It is interesting to note that the –10 hexamer in \( P_1 \) is preceded by a TG dinucleotide, positioned one base upstream of the –10 sequence (Fig. 1C), a feature found in so-called extended –10 promoters. In \( E. coli \) and \( B. subtilis \), the TGn extension (where \( n \) is any nucleotide) allows for closer contacts between the RNA polymerase sigma subunit and its promoters, especially

![FIG. 1. Partial physical map of phage GIL01 with SOS-regulated promoters. (A) Predicted ORFs are depicted as open boxes, and left-to-right arrows indicate the direction of transcription of the left and right functional domains. Probable gene functions are indicated above certain ORFs. Putative Rho-independent transcription terminators are shown as open stem loops downstream of ORF8 and ORF30. Promoters are shown as filled arrows, and each promoter region is enlarged below the map. Boxes for 35 and –10 are depicted as black rectangles, and angled arrows represent transcription start sites. Gray boxes show the locations of conserved LexA-binding sites in relation to the three promoters. (B) Transcription start site mapping of GIL01 ORF1 and ORF9 by 5’ RACE. cDNAs were prepared from GIL01 mRNA, tailed with dCTP, and PCR amplified. The purified products were cloned into pCR4-TOPO and sequenced using the universal M13 primers. The electropherograms show the transcription start sites (indicated with arrows; sequences correspond to the GIL01 negative strand) fused to their oligo(dC) tails. The promoter from which transcription is initiated is indicated for each electropherogram. Note that for the transcription start at \( P_3 \), an additional thymine residue that is not found in the GIL01 sequence has been inserted upstream of the dC tail. (C) Detailed nucleotide sequences of promoters depicted in panel A. Transcription start sites detected in panel B are indicated with angled arrows (\( P_1 \), \( P_2 \), and \( P_3 \)). Putative ribosome-binding sites are shown in lowercase, start codons are shown in bold, and SOS boxes are shaded gray. A TG dinucleotide motif characteristic of –10 extended promoters is double underlined within \( P_1 \). Genome coordinates are indicated above each sequence.](http://jb.asm.org/)
TABLE 1. P1-P2 and P3 promoter activities in lexA* and lexA(496D) backgrounds

<table>
<thead>
<tr>
<th>Promoter fusion</th>
<th>lexA allele</th>
<th>Mean β-galactosidase activity (Miller units ± SD) in indicated growth conditions</th>
<th>−MMC</th>
<th>+MMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1-P2</td>
<td>lexA*</td>
<td>75 ± 13</td>
<td>303 ± 106</td>
<td>91 ± 5</td>
</tr>
<tr>
<td></td>
<td>lexA(496D)</td>
<td>91 ± 5</td>
<td>78 ± 5</td>
<td>78 ± 5</td>
</tr>
<tr>
<td>P1-P2 (dinBox1mut)</td>
<td>lexA*</td>
<td>345 ± 28</td>
<td>351 ± 10</td>
<td>339 ± 19</td>
</tr>
<tr>
<td></td>
<td>lexA(496D)</td>
<td>339 ± 19</td>
<td>326 ± 12</td>
<td>326 ± 12</td>
</tr>
<tr>
<td>P3</td>
<td>lexA*</td>
<td>3 ± 1</td>
<td>26 ± 10</td>
<td>3 ± 1</td>
</tr>
<tr>
<td></td>
<td>lexA(496D)</td>
<td>3 ± 1</td>
<td>6 ± 2</td>
<td>6 ± 2</td>
</tr>
</tbody>
</table>

*a lacZ fusions to P1-P2, P1-P2 mutated in dinBox1 (CGAAACagcTTTTT; boldface letter represents the mutated nucleotide), and P3 promoters were generated, and β-galactosidase activity was measured in both lexA* and lexA(496D) lysogens in normal growth (−MMC) and SOS-inducing (+MMC) conditions. Cultures were grown for 3 h at 30°C when MMC50 was added to one-half of the cells, and growth was allowed to proceed for one additional hour. Cells were then permeabilized and assayed for β-galactosidase. Activity values are the averages of three independent experiments.

when a well-conserved −35 box is absent (5, 11, 34, 37, 58, 70). This could be the case for P1, where the −35 box does not begin with T, the predominant base found at the beginning of the −35 hexamer in Gram-positive bacteria (70).

In addition to the reported promoter sequences, potential ribosome-binding sites were identified 10 and 6 nucleotides upstream of the start codons of ORF1 and ORF9, respectively. Also, Rho-independent transcription terminators were identified downstream of ORF8, where transcription of the replication/regulation operon would be expected to terminate, and downstream of ORF30, where transcription of the lysis/assembly operon is expected to terminate (Fig. 1A). We confirmed that transcription does not span the regulatory region between ORF8 and ORF9 by performing reverse transcription (RT)-PCR with primer sets that span this region (data not shown). Finally, lacZ reporter fusions were generated to the regions containing P1-P2 and P3 in order to verify that these are biologically active promoters. As shown in Table 1, both promoter regions directed the expression of β-galactosidase in the GIL01 B. thuringiensis host. Such transcriptional organization and the corresponding gene distribution further strengthen the idea that the GIL01 genome is transcribed as two major operons, one responsible for genome replication and regulation, the other for virion structure, assembly, and host cell lysis.

GIL01 cp mutants map to three distinct loci. A chief objective of this study was to identify the components of the genetic circuit governing GIL01 development. The only obvious candidates for transcription regulation were ORF1 and ORF6. PSI-BLAST searches identified a potential helix-turn-helix (HTH) DNA-binding motif in ORF1 similar to the HTH found in the MerR superfamily of transcription regulators. In ORF6, the same database searches revealed a potential winged HTH DNA-binding domain similar to that found in the N terminus of LexA, the master repressor of the SOS regulon. We reasoned that since GIL01 is induced in DNA-damaging conditions, most likely by way of the cellular SOS response, it is potentially under the transcriptional control of a cleavable repressor analogous to CI in the phage λ. However, neither ORF1/ORF6 nor any other protein-coding region in GIL01 displays the typical structure of the CI/LexA class of repressors: an N-terminal DNA-binding domain, followed by an Ala-Gly or Cys-Gly cleavage site and a C-terminal protease domain. Yet we did not rule out the existence of a distinct type of SOS-dependent induction mechanism, such as the LexA-regulated antirepressors found in coliphages 186 (8, 38) and N15 (47), or that GIL01 induction relies on a yet-unknown SOS pathway.

To investigate these possibilities, we took advantage of knowledge acquired from a previous study of spontaneous GIL01 cp mutants defective for the lysogenic cycle (67). Since ORF6 was a strong candidate for encoding a repressor of the lytic cycle, the first cp mutants had been sequenced within and in the vicinity of ORF6. Unexpectedly, no mutations were detected in ORF6. Instead, mutations were found in the downstream ORF7, the function of which was unknown. ORF7 cp mutants displayed either a deletion or insertion of 11 nucleotides in an almost perfect direct repeat (DR) (5′-CAAGTCGGTAaG-3′; the lowercase letters represent the only mismatch) at coordinates 4,615 through 4,636.

In this study, a more extensive collection of cp mutants was assembled from multiple independent experiments and analyzed. GIL01 cp mutants arose at a frequency of approximately 10−4 and displayed otherwise normal growth properties. In the expected direction of DR mutations in ORF7 (deletions in cp04 and cp18, insertions in cp25 and cp29), a mutant (cp08) carrying a single nucleotide substitution in the second repeat, resulting in a premature terminator codon, was also isolated. Another ORF7 mutant, cp27, has a single-nucleotide insertion downstream of the DR (Fig. 2A). In every instance, the mutation altered the predicted protein sequence: cp04, cp08, cp18, cp25, and cp29 all resulted in truncation of ORF7, while cp27 resulted in a modified C terminus (Fig. 2B).

Additionally, a class of GIL01 cp mutants that lacked mutations in ORF7 was isolated, suggesting that mutations in other regions of the GIL01 genome might account for their cp phenotype. This class of mutants all contained single-nucleotide substitutions, deletions, or insertions within ORF1 (Fig. 2A) or they contained mutations immediately upstream of ORF1 between promoters P1 and P2 (Fig. 1C and 3A). Interestingly, mutations in ORF1 cluster in two regions: the predicted HTH DNA-binding motif (cp32, cp33, and cp36) or the most C-terminal residues (cp06 and cp10) (Fig. 2B). Mutations in the noncoding region upstream of ORF1, hereafter termed dinBox1, resulted in base substitutions (cp16, cp19, cp23, and cp26) or single-nucleotide deletions (cp17 and cp37) at either of two nucleotide positions within a 14-bp sequence that bears strong similarity to the LexA-binding site in Gram-positive bacteria (Fig. 3B). As outlined below, a single mutation in dinBox1 is sufficient to abolish the lysogenic pathway of GIL01.

To rule out the possibility that additional mutations, outside the sequenced areas, contributed to, or were responsible for, the observed cp phenotype, we sequenced the entire GIL01 genome of cp26 (mutated in dinBox1), cp27 (cp29) (mutated in ORF7), and cp32/cp33 (mutated in ORF1). cp26 (dinBox1), cp27 (ORF7), and cp33 (ORF1) displayed only the initially identified nucleotide substitutions (cp26 and cp33) and insertion (cp27). cp29 and cp32, however, each contained one ad-
FIG. 2. cp mutants in ORF1 and ORF7. Gene (A) and protein (B) sequences of cp mutants were aligned and compared to their respective wild-type sequences. Nucleotide coordinates (A) and the corresponding protein sizes (B) are indicated for each sequence. Start and stop codons are shown in bold, and mutations appear on a light gray background. (A) The direct repeat in ORF7 is highlighted against a black background in the wild-type sequence. (B) ORF1 harbors a putative helix-turn-helix DNA-binding motif in its N-terminal domain, and the two helices, H1 and H2, are shown boxed in black in the wild-type sequence. H1 and H2 were predicted using the software programs Phyre (35) and ProteDNA (14).
ditional mutation. In cp29, a single-nucleotide substitution changed a Gln to Lys (CAA→AAA) at the ORF29 residue 19, the function of which is currently unknown. cp32 contained an additional deletion of three nucleotides at codon 64 within ORF19 (function unknown), resulting in the in-frame deletion of a single Ile residue. Yet cp29 and cp32 display the same infectious behavior as their cognate cp27 and cp33 mutants, which are mutated solely in ORF7 and ORF1, respectively (see complementation experiments below and Table 2). This strongly suggests that ORF18 and ORF29 are not involved in the regulatory network of GIL01. Taken together, our results indicate that dinBox1, ORF1, and ORF7 are critical for the establishment or maintenance of the GIL01 lysogenic cycle.

Ultimately, we wished to classify cp mutants according to their degree of virulence on lysogen and nonlysogen hosts. However, none of the cp mutants tested produced plaques on lawns of the lysogen strain GBJ338, indicating that GIL01 lysogens are immune to infection by cp mutants. This could be due to the activity of phage- and/or host-encoded immunity factors that shift GIL01 toward lysogenic development. Alternatively, we cannot rule out the possibility that infection was barred by a yet-unidentified phage-borne superinfection exclusion system.

![FIG. 3.](image-url)

**TABLE 2. Functions of defective gp1 and gp7 can be complemented**

<table>
<thead>
<tr>
<th>Host (expressed protein[s])</th>
<th>cp23/cp26 (dinBox1)</th>
<th>cp32/cp33 (ORF1)</th>
<th>cp27/cp29 (ORF7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBJ220 (gp1)</td>
<td>0.5</td>
<td>0.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.2</td>
</tr>
<tr>
<td>GBJ223 (gp7)</td>
<td>1</td>
<td>1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.5</td>
</tr>
<tr>
<td>GBJ467 (gp1/gp7)</td>
<td>0.2</td>
<td>&lt;10&lt;sup&gt;-8&lt;/sup&gt;</td>
<td>0.2&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> The GIL01-cured host strain (GBJ002) was transformed with plasmids expressing either gp1 (GBJ250), gp7 (GBJ223), or gp1 and gp7 (GBJ467). The resulting strains were grown exponentially before 1 mM IPTG was added to one-half of the culture, and growth was allowed to proceed for three additional hours. Uninduced and induced cultures were then infected with GIL01 cp mutants in dinBox1 (cp23/cp26), ORF1 (cp32/cp33), or ORF7 (cp27/cp29). Two different cp mutants at each locus were tested independently, and similar infection patterns were obtained for each pair of mutants. The results shown are representative of four independent experiments and give the ratio between titers produced on induced cells and titers counted on the same uninduced strains (PFU on host + IPTG/PFU on host – IPTG).

<sup>b</sup> cp32 and cp33 formed small clear plaques on the cured host strain. There was no systematic change in plaque morphology between induced and uninduced cells infected with ORF1 mutants.

<sup>c</sup> The plaques observed on induced lawns were turbid and significantly smaller than those produced on the uninduced lawns (Fig. 4).
gp1 and gp7 regulate lysogenic development. Since mutations in either ORF1 or ORF7 commit GIL01 to multiply exclusively as a lytic phage, we investigated more closely the roles of their respective gene products, gp1 and gp7, in GIL01 development. The corresponding genes were cloned into the E. coli-B. subtilis shuttle vector pDG148, placing them under the transcriptional control of the IPTG-inducible Pspac promoter. Three different plasmids, whereby gp1 and gp7 were expressed individually or together as a single transcriptional unit, were generated. The permissive B. thuringiensis host GBJ002 was transformed with the resulting plasmids, generating strains GBJ250 (gp1), GBJ223 (gp7), and GBJ467 (gp1/gp7). Each of these strains was induced for protein expression and subjected to infection with GIL01 cp mutants in dinBox1, ORF1, or ORF7. The results of these experiments are summarized in Table 2. Expression of gp1 alone reduced the infective titers of ORF1 mutants cp32 and cp33 approximately 10-fold. Surprisingly, gp1 expression also reduced the plating efficiency of dinBox1 and ORF7 mutants, although to a lesser extent (2-fold and 5-fold PFU reductions, respectively). This was not the case when gp7 was expressed in the host strain. In this case, a reduction in phage titers was observed only when induced GBJ223 cells were infected with the ORF7 mutants cp27 and cp29. Although titers showed only a 2-fold decrease, the plaques were now turbid and considerably smaller than those seen on the uninduced strain GBJ223 (Fig. 4). This observation was less remarkable when ORF1 mutants were plated on induced cultures of GBJ250, as ORF1 mutants tend to produce smaller plaques in general.

Infection of these same induced cell lines with wild-type GIL01 yielded titer reductions that varied between 2-fold and 10-fold of those of infections of uninduced cultures, while plaque morphology remained unaltered (data not shown). This noticeable reduction in titers, as well as the considerable experiment-to-experiment variability in titers, could be explained by variations in the cellular concentrations of gp1 and gp7 generated with our plasmid-borne, IPTG-inducible system. Wild-type GIL01 probably expresses balanced levels of gp1 and gp7 inside the host cell, and disrupting that balance is likely to affect phage behavior in many ways. These data are consistent with the idea that gp1 and gp7, and possibly the proper expression rates of these two factors, play important roles in the choice between lytic and lysogenic development.

Given the above results, we investigated whether there was any effect due to the simultaneous expression of gp1 and gp7 in the cell. There was indeed a substantial reduction in phage titers when induced GBJ467 cells were infected with dinBox1 mutants, and no plaques at all were observed with ORF1 mutants (Table 2). In contrast, no difference in plating efficiency was observed between infections of induced GBJ250 and GBJ467 with ORF7 mutants. Hence, coexpressing gp1 and gp7 in the same strain enhances complementation of dinBox1 and ORF1 mutants but does not result in enhanced rescue of ORF7 mutants.

Since complementation was associated with sharp reductions in infective titers, we wished to verify that expression of gp1 and/or gp7 conferred immunity by preventing the infecting mutant from propagating, rather than by blocking infection altogether. In order to investigate the fate of the cp mutants postinfection, bacterial samples were recovered from infected spots and streaked on fresh plates containing IPTG. PCRs with GIL01-specific primers were performed on isolated colonies in which gp1 and/or gp7 was expressed and which had been infected with mutants cp26 (dinBox1), cp27 (ORF7), and cp33 (ORF1). Induced GBJ223 colonies were found to be positive for GIL01 only when infected with cp27, showing that expression of gp7 complements the missing immunity function of the corresponding cp mutant. GIL01-positive colonies were also obtained from GBJ250 spot infections with cp33. Similar results were obtained when induced GBJ467 was infected with each of the three cp mutants. These results are in agreement with the infection patterns described above (Table 2), and altogether, they support the idea that gp1 and gp7 act in immunity against infection and, together with the dinBox1 element, play important roles in the regulatory network of GIL01.

dinBox1 is a conserved LexA box. dinBox1 showed strong similarity to the consensus LexA-binding site (CGAA Cn,GTTGC, where n is any nucleotide) in B. subtilis (13, 30, 72). Its sequence, CGAACaagcGTTT, is centered at positions +17 and −86 with respect to the P1 and P2 transcription start sites, respectively (Fig. 1C). cp mutants at dinBox1 were unable to enter lysogeny, and mutations always involved changes to strongly conserved positions within the SOS box. Interestingly, Bam35c, a cp mutant of phage Bam35 (55), differs from GIL01 at only a few nucleotide positions, one of which is located within the critical bases for LexA binding to dinBox1 (Fig. 3A). In agreement with this finding, experimental data for the LexA box sequence requirements in B. subtilis show that changes to any one of the five outer nucleotides can considerably reduce or prevent LexA binding (24). In addition
to conserved dinBox1, two motifs with similar characteristics were found near promoter P3. dinBox2 and dinBox3 are centered at −22 and +11 relative to the start of transcription (Fig. 1C). It is worth noting that, while numerous mutations in dinBox1 were isolated, none of the mutations isolated in this study carried mutations in dinBox2 or dinBox3.

To further consolidate our observations, we searched for SOS box sequences in the upstream regions of SOS genes in B. cereus strain G9842, which is closest to strain GBJ002 in DNA sequence. We compared the LexA operators found upstream of genes involved in transcriptional regulation (lexA), recombination (recA and ravAB), and DNA repair (uvrAB) to the consensus sequence for LexA binding in B. subtilis and to the three dinBox sequences identified in GIL01. As shown in Fig. 3B, the B. cereus SOS boxes and the three GIL01 dinBox sequences closely match the B. subtilis consensus SOS box. Additionally, an in silico approach using the program RSAT (17) identified the same three dinBox sites found by visual analysis of the GIL01 promoter sequences (Fig. 3B). Lastly, alignment of the B. cereus LexA protein sequence with the Gram-positive consensus LexA sequence revealed strong conservation of the N-terminal DNA-binding domain (Fig. 3C). Taken together, these observations strongly suggest that the dinBox sequences identified in GIL01 function as LexA-binding sites in the B. cereus group of bacteria.

**GIL01 is induced in DNA-damaging conditions in a LexA-dependent manner.** In a previous study, we reported that GIL01 was induced by DNA-damaging agents such as MMC, nalidixic acid, or exposure to UV irradiation (69). Such treatments are known to elicit the cellular SOS response in a RecA-dependent manner (71). To investigate if RecA is involved in the induction cycle of GIL01, we disrupted the B. thuringiensis recA gene through insertional mutagenesis (66). The resulting strain, GBJ499, was particularly sensitive to UV and MMC treatment and displayed a slower growth rate than the wild-type strain (data not shown). GBJ499 was infected with GIL01 to generate the lysogen GBJ500. Phage titers produced by spontaneous induction of GBJ500 were reduced approximately 10-fold compared to those of the wild-type lysogen GBJ338 (Table 3). In addition, whereas GIL01 was strongly induced after UV or MMC treatment of the wild-type lysogen strain, there was no increase in phage titers when GBJ500 was similarly induced (Table 3, row 2). These results indicate that GIL01 induction is RecA dependent and potentially linked to the B. thuringiensis SOS response pathway. To investigate this hypothesis, the role of cellular LexA in the lytic switch was assessed. We were unable to isolate a lexA null mutant of B. thuringiensis. This result was not entirely unexpected, considering that LexA represses the expression of cell division inhibitors in other bacterial groups (sulA in E. coli and yneA in B. subtilis). In E. coli, lexA null mutants are not viable unless sulA is first inactivated (52). In B. subtilis, lexA-deficient cells grow into long filaments (25) as a consequence of YneA upregulation (33). The B. cereus LexA is 70% identical to its counterpart in B. subtilis, and most of the shared residues lie in the N-terminal DNA-binding domain (Fig. 3C). Most importantly, B. cereus LexA contains an Ala96-Gly97 peptide sequence that, in both B. subtilis and E. coli (at Ala91-Gly92 in both organisms) is the site of RecA-stimulated autoproteolysis (28, 51). Certain amino acid changes at the Ala-Gly cleavage site are known to abolish LexA proteolysis (19, 42). We used a similar approach to construct a noncleavable lexA mutant by allele exchange in B. thuringiensis, replacing the Ala codon with an Asp codon [lexA(A96D)]. As expected, the resulting strain, GBJ396, was hypersensitive to DNA damage conditions, displaying a 300-fold reduction in survival after treatment with MMC, compared to a 9-fold reduction in survival of MMC-treated wild-type cells. Furthermore, a lexA-lacZ reporter fusion, when introduced into GBJ002 (lexA−), showed an expected 4-fold increase in lacZ expression after treatment with MMC. However, there was no increase in lacZ expression after SOS treatment of GBJ396 [lexA(A96D)] carrying the same lexA-lacZ reporter plasmid, indicating that LexA was not proteolyzed and therefore not released from its operator sites by SOS induction signals (data not shown). The lexA(A96D) mutant was infected with GIL01 to generate the lysogen strain GBJ406. Similar to what we observed for the recA mutant lysogen, GIL01 induction from GBJ406 was strongly reduced in comparison to that of the wild-type host strain (Table 3, row 3). The data obtained with both recA and lexA mutants show that GIL01 is induced as part of the host SOS response pathway.

<table>
<thead>
<tr>
<th>Strain [relevant genotype]</th>
<th>Phage titer without MMC</th>
<th>Phage titer with MMC</th>
<th>Ratio</th>
<th>Phage titer without UV</th>
<th>Phage titer with UV</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBJ338 [GBJ002 (GIL01)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBJ500 (GBJ338 recA::pmutin4)</td>
<td>4.0 × 10⁶</td>
<td>3.0 × 10⁴</td>
<td>0.8</td>
<td>1.1 × 10⁶</td>
<td>2.0 × 10⁶</td>
<td>1.8</td>
</tr>
<tr>
<td>GBJ406 [GBJ338 lexA(A96D)]</td>
<td>1.5 × 10⁵</td>
<td>1.8 × 10⁴</td>
<td>1.2</td>
<td>6.7 × 10⁴</td>
<td>5.3 × 10⁴</td>
<td>0.8</td>
</tr>
</tbody>
</table>

* Phage titers were determined by plating serial dilutions of GIL01 on the cured host GBJ002. Induction ratios were determined by dividing phage titers of induced cultures by the titers of uninduced cultures. Results are averages from three independent experiments.

The MMC concentration used was 50 ng ml⁻¹, and cultures were induced for 1 h.

* Cells were irradiated at 254 nm for 10 s.

---

TABLE 3. GIL01 induction is recA and lexA dependent*
Moreover, β-galactosidase expression from a P1-P2-lacZ fusion in which dinBox1 was mutated at a critical residue was no longer repressed in either GBJ338 (lexA') or GBJ338 [lexA(A96D)], and MMC treatment did not further increase LacZ expression. Taken together, these data strongly indicate that LexA regulates P1 and P2 and that dinBox1 is necessary for this LexA-mediated regulation. Parallel experiments with promoter P3, in which a 226-bp fragment that includes dinBox2 and dinBox3 was fused to lacZ, yielded a 9-fold increase in β-galactosidase activity in the wild-type host after MMC treatment. Once again, the increase in lacZ expression levels after MMC treatment was eliminated when the same experiment was performed in a host carrying the lexA(A96D) mutation. These experiments demonstrate that host LexA regulates promoters P1, P2, and P3, most likely by binding to the dinBox sequences and preventing transcription initiation.

**DISCUSSION**

In this study, we used a genetic approach to identify central components of the system that regulates the temperate state in phage GIL01. We provided evidence that the host SOS repressor LexA is directly involved in the switch but that phage-borne factors gp1 and gp7 are also necessary for lysogeny. LexA is likely to act by binding to dinBox1, a conserved SOS box lying between two newly identified promoters upstream of lysogeny and replication genes. Two additional conserved LexA-binding sites were also identified upstream of structure and lysis genes, further suggesting that LexA plays a central role in GIL01 development. Our results led to the conclusion that GIL01 uses a combination of host- and phage-derived regulators to control its cycle.

At first glance, GIL01 does not appear to display the characteristic modular organization usually encountered in temperate phages. Its 15-kbp genome is smaller than most sequenced prophage genomes, and all 30 ORFs are transcribed in the same rightward direction, a rarely observed feature. In addition, only a few genes have attributed functions, making it difficult to define the boundaries of gene clusters involved in related functions. Most notably, a lysogeny module is not readily apparent, as there are only two gene candidates, ORF1 and ORF6, predicted to code for DNA-binding proteins that could function as transcription regulators. Nevertheless, the identification of promoters in large noncoding segments and Rho-independent terminators at the end of transcriptional units revealed a functional organization in two major modules (Fig. 1A). Two adjacent promoters, P1 and P2, at the left extremity of the phage genome, control transcription of lysogeny and replication functions, whereas a probable single internal promoter, P3, directs the expression of structural and lytic genes. Such transcriptional organization typically allows for complex, multilayered regulation of gene transcription. It would, for instance, permit the expression of regulatory factors and phage DNA polymerase to levels that are required for lysogeny while keeping lytic genes fully repressed. Conversely, upon perception of an inducing signal, synthesis of terminal protein and DNA polymerase would be significantly enhanced along with induction of structural and lytic genes.

A similar overall genome organization can be found in the B. subtilis lytic phage φ29, with the notable difference that early and late genes are divergently transcribed. Tandem promoters A2b and A2c jointly direct the expression of the leftward regulation and replication genes, which are similarly distributed in GIL01 (49). Both promoters are strong, and their repression by two early factors coincides with late gene transcription activation from the flanking divergent promoter A3, a determining step in the switch from early to late transcription (11, 50). The coupling of tandem promoters to the same operator is frequently observed in bacteria (examples are the rRNA operons [23, 73], cp gene [22], and fis operon [53]) and offers a certain flexibility for differential gene regulation. In the autonomously replicating coliphage N15, for example, two adjacent promoters direct gene expression from the anti-immunity locus immA, and their differential control is important for the choice between lysis and lysogeny (56). Similarly arranged promoters regulate anti-immunity operons in distinct phases, such as P1, P4, and P7 (56), suggesting that this regulatory strategy is widespread among temperate phages. It is therefore conceivable that in GIL01, promoters P1 and P2 display distinct strengths, are subject to variable levels of regulation, or give rise to different transcripts.

GIL01 lytic development is induced by DNA-damaging treatments that elicit the cellular SOS response, and yet GIL01 does not code for a cleavable CI-like repressor. Instead, we found that regulation is mediated by host LexA, since the phage lytic pathway was no longer induced in a noncleavable lexA (Ind1) background. Accordingly, three putative LexA operators that bore a strong resemblance to the consensus LexA-binding site in Gram-positive bacteria were identified between promoters P1 and P2 (dinBox1) and at P3 (dinBox2 and dinBox3). We showed that the SOS induction of both promoter regions was blocked by a noncleavable lexA (Ind1) host mutation. Moreover, a single-nucleotide mutation at a strongly conserved position within dinBox1 completely eliminated transcription repression at P1-P2. Although additional experiments will be needed in order to confirm LexA binding to dinBox1, these data lend credence to the idea that host LexA is directly involved in the GIL01 control circuit.

Whereas numerous cp mutants were isolated in dinBox1, no mutations were ever detected in the downstream dinBox2 or dinBox3. This suggests that dinBox2 and dinBox3 are either essential for phage development or, if such mutants are indeed viable, they cannot be detected as cp formers. Also, dinBox2 and dinBox3 could perform an additional function in the regulatory system, such as in LexA-mediated DNA looping between dinBox2/3 and the dinBox1 promoter region. In this case, a cp mutant might not be obtained unless both boxes were simultaneously mutated, which is likely to occur at too low a frequency to be detected. In any event, LexA binding to both sites provides the basis for a plausible model in which a cleavable repressor would be immediately responsible for lytic gene expression. However, it is unlikely that LexA is the sole regulator of the switch, as this would lead to inefficient GIL01 induction. Due to the negative autoregulation of lexA transcription, normal LexA levels are restored once DNA damage inside the cell has been repaired (7, 44). LexA would then be able to bind its operators in GIL01 again and interrupt lytic development. In the λ phage, for instance, this is prevented by strong repression of CI synthesis once lytic development commences (60). It is therefore conceivable that LexA plays an ancillary role only in the regulatory network by controlling the
expression of phage-borne repressor(s) and activator(s), these in turn being responsible for the lytic switch per se. This hypothesis is strongly supported by the isolation of GIL01 cp mutants deficient for gp1 or gp7. Both proteins are required for lysogeny, and their absence can be complemented by expressing the missing functions in trans. Interestingly, cross-complementation is observed with gp1, whereas it is not seen with gp7. The observed differences could be attributed to distinct and/or independent roles of gp1 and gp7. gp1 potentially having a broader activity range. In addition, since the expression system used in this study does not provide any insight into the protein concentrations generated, it is quite probable that they differ from the intracellular levels produced by GIL01 during infection. This could also explain why wild-type GIL01 is not fully repressed when infecting cells that express both gp1 and gp7. One would expect the wild-type phage cycle to be hindered by the two factors known to be involved in lysogeny, and yet titers are only slightly reduced. This observation suggests that there is a need for precise levels of both factors and, probably, a determined timing of expression as well.

gp1 displays a DNA-binding domain similar to the one found in the MerR (repressor of mercury Resistance operons) superfamily of transcription regulators. MerR itself is known to mediate both repression and activation of transcription by virtue of its ability to distort the DNA helix (3, 4). A role for gp1 as a modulator of DNA structure is an attractive idea, as it is a common theme among transcription regulators. gp7, on the other hand, has no predicted function or recognizable sequence motif. It might cooperate with gp1 in the same regulatory process or it could act independently, at a completely different level of immunity regulation.

Temperate phages that are under the direct control of host LexA utilize a variety of regulatory mechanisms. Coliphages 186 and N15 code for LexA-regulated antirepressors (8, 47), and gene homologues that are preceded by LexA-binding sites have been identified in other phages (9, 12, 27). A unique example where LexA exerts a dual function is provided by the Vibrio cholerae dinBox filamentous phage CTX. The phage-encoded repressor RstR acts together with host LexA to activate and repress gene transcription (36, 54). Such a mechanism allows the phage to respond to small variations in RstR/LexA levels and switch from a low-particle-formation mode to higher yields and vice versa. This regulatory system appears to be well suited for a filamentous phage, the production of which is maintained throughout the cell cycle.

GIL01 cp mutants were isolated in three distinct loci—dinBox1, ORF1, and ORF7—each of which is necessary for the lysogenic cycle. Surprisingly, none of the isolated mutants produced plaques on lysogens, a result that we expected to see with at least the operator mutants in dinBox1. There are many possible explanations for the observed immunity effect. One is that GIL01 codes for an as-yet-unidentified superinfection exclusion system that blocks entry of newly infecting phages. A more attractive explanation, however, would be that dinBox1 is not the only operator needed for lysogeny. If other regulatory sites are present on the phage genome, potentially responding to gp1 and/or gp7, these sites could constitute themselves a LexA-independent pathway for regulation and compensate for the loss of the dinBox1 pathway. In this case, a mutant, superinfecting phage would still be sensitive to repression upon infection of a lysogen, provided that the appropriate immunity factors are readily available in the cell. It is noteworthy that even though the dinBox1, ORF1, and ORF7 alleles were independently isolated multiple times, they probably do not represent a saturated screen of all possible mutations leading to the cp phenotype. We therefore do not exclude the possibility that yet-unidentified factors also participate in GIL01 regulation.

GIL01 is regulated by a multicomponent system composed of at least host LexA, gp1, and gp7. The combination of host- and phage-encoded functions in the lytic switch has already been seen in other phages and enables a fine regulation of lytic functions. LexA is used as an internal sensor of stress and allows for the rapid escape from potentially deleterious damage to the host. The addition of phage-borne functions could modulate the switch by conferring more precision to the system or by adding a further layer of control. GIL01 has developed an elaborate network to control its development, and future studies will address how the different components identified in this study intervene in the switch.

ACKNOWLEDGMENTS

We are grateful to B. Clément and S. Degand for their contributions to the early stages of this work and to A. Toussaint and A. Aertsens for stimulating discussions. We thank R. Janky for continuous encouragement and skilful advice with RSAT and R. Leplae for assistance with protein analysis. We are particularly indebted to H. Kinsey for insightful observations and comments on the manuscript. We thank C. Condon for kindly providing the pDG148 vector.

This work was supported by grants from the National Fund for Scientific Research and the Université Catholique de Louvain (grants to N.F. and J.M.). Part of this work was also supported by the Finnish Centre of Excellence Program of the Academy of Finland (grant 1129648 to J.K.H.B.), an EMBO short-term fellowship, and the National Research Fund in Luxembourg, cofunded under the Marie Curie Actions of the European Commission (FP7-COFUND) (grants to N.F.).

REFERENCES


Downloaded from http://jb.asm.org/ on April 6, 2021 by guest