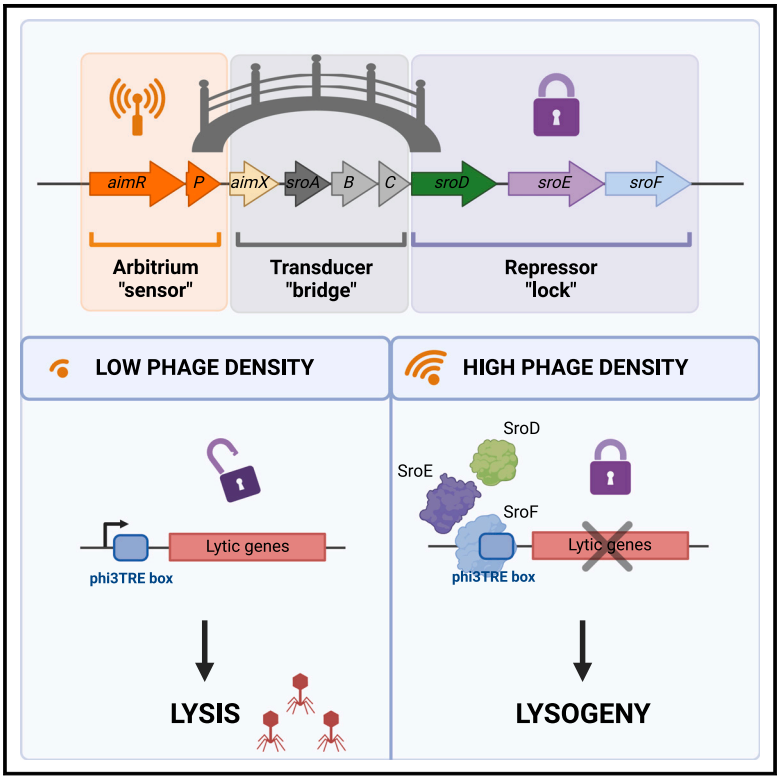
Article

Cell Host & Microbe

Characterization of a unique repression system present in arbitrium phages of the SPbeta family

Graphical abstract



Authors

Aisling Brady, Elena Cabello-Yeves, Francisca Gallego del Sol, ..., Nuria Quiles-Puchalt, Alberto Marina, José R. Penadés

Correspondence

amarina@ibv.csic.es (A.M.), j.penades@imperial.ac.uk (J.R.P.)

In brief

SPbeta *Bacillus* phages use arbitrium to coordinate lysis-lysogeny decisions via the *sro* operon. Brady et al. show that the operon encodes two master repressors that bind multiple sites of the phage genome to promote lysogeny. Three accessory proteins connecting the arbitrium signal to the repression module constitute the transducer module.

Highlights

- SPbeta phages encode two master repressors, SroF and SroE, in the *sro* operon
- Multiple SroF binding sites required for repression and maintenance of lysogeny
- The *sro* operon consists of conserved repressor and variable transducer modules
- The transducer module bridges the arbitrium signal to the repression module

Brady et al., 2023, Cell Host & Microbe *31*, 2023–2037
December 13, 2023
https://doi.org/10.1016/j.chom.2023.11.003

CellPress

Cell Host & Microbe

CellPress
OPEN ACCESS

Article

Characterization of a unique repression system present in arbitrium phages of the SPbeta family

Aisling Brady,[1,2,7] Elena Cabello-Yeves,[3,4,7] Francisca Gallego del Sol,[3,4] Cora Chmielowska,[1] Javier Mancheño-Bonillo,[3,4] Sara Zamora-Caballero,[3,4] Shira Bendori Omer,[5] Manuela Torres-Puente,[3] Avigdor Eldar,[5] Nuria Quiles-Puchalt,[1,6] Alberto Marina,[3,4,*] and José R. Penadés[1,8,*]

[1]Centre for Bacterial Resistance Biology, Imperial College London, London SW7 2AZ, UK
[2]Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow G12 8TA, UK
[3]Instituto de Biomedicina de Valencia (IBV-CSIC), 46010 Valencia, Spain
[4]CIBER de Enfermedades Raras (CIBERER), 46010 Valencia, Spain
[5]Shmunis School of Biomedicine and Cancer Research, Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv, Israel
[6]Department of Biomedical Sciences, Faculty of Health Sciences, Universidad CEU Cardenal Herrera, CEU Universities, Alfara del Patriarca 46115, Spain
[7]These authors contributed equally
[8]Lead contact
*Correspondence: amarina@ibv.csic.es (A.M.), j.penades@imperial.ac.uk (J.R.P.)
https://doi.org/10.1016/j.chom.2023.11.003

SUMMARY

Arbitrium-coding phages use peptides to communicate and coordinate the decision between lysis and lysogeny. However, the mechanism by which these phages establish lysogeny remains unknown. Here, focusing on the SPbeta phage family's model phages phi3T and SPβ, we report that a six-gene operon called the "SPbeta phages repressor operon" (*sro*) expresses not one but two master repressors, SroE and SroF, the latter of which folds like a classical phage integrase. To promote lysogeny, these repressors bind to multiple sites in the phage genome. SroD serves as an auxiliary repressor that, with SroEF, forms the repression module necessary for lysogeny establishment and maintenance. Additionally, the proteins SroABC within the operon are proposed to constitute the transducer module, connecting the arbitrium communication system to the activity of the repression module. Overall, this research sheds light on the intricate and specialized repression system employed by arbitrium SPβ-like phages in making lysis-lysogeny decisions.

INTRODUCTION

Bacteriophages, or phages, are viruses that target bacteria and are governed by two primary life cycles: lytic and lysogenic. In the lytic cycle, phages infect bacteria, hijack cellular machinery, replicate, and cause the bacterial cell to burst (lysis). The liberated phages can then infect other bacteria, perpetuating the lytic cycle. Although some phages exclusively follow the lytic cycle, temperate phages can choose between lytic and lysogenic cycles. During lysogeny, their genetic material integrates into the bacterial genome, as a dormant prophage, and is replicated passively with the host. This decision between lysis and lysogeny profoundly impacts both the phage and the bacterial host, influencing population dynamics and bacterial evolution.[1,2]

In classic phages like λ (*Escherichia coli*) and 80α (*Staphylococcus aureus*), the lytic-lysogenic choice relies on two primary regulators, CI and Cro.[1,3] CI promotes lysogeny by suppressing lytic gene expression, whereas Cro promotes the lytic cycle by antagonizing CI.[1,3] This regulation responds to environmental cues, especially the SOS response triggered by DNA damage or stress, which activates RecA, leading to CI's cleavage and prophage induction.[3] Various factors influence temperate phages' lifecycle decisions, such as pH, temperature, salt levels, exposure to exogenous DNA, or other prophages.[4] Recently, it was discovered that phages can employ small-molecule signaling, akin to quorum sensing systems in bacteria, to coordinate lysis-lysogeny decisions. One such system, called arbitrium, was found in the *Bacillus*-infecting temperate phage phi3T.[5,6] Arbitrium allows phages to communicate via small molecules and coordinate their life cycle decisions. Notably, the SPbeta phage family, encompassing phi3T, lacks a classical CI-like repressor, adding to the mystery of how they establish lysogeny.

The arbitrium system comprises three phage genes, *aimR*, *aimP*, and *aimX*.[5] AimR function is controlled by AimP, a short peptide that is secreted extracellularly and internalized by host *Bacillus* cells.[5] When a threshold concentration is reached following multiple infections, AimP binds to AimR, inactivating its function. AimR was initially considered a transcriptional activator but has recently been identified as an antiterminator.[7] It inhibits a transcriptional terminator (TT) between *aim*P and *aim*X during early infection stages, enabling the concurrent expression of these genes. In phi3T, AimX is a peptide that promotes lysis by binding to MazF, an mRNA-cleaving toxin, and to phi3T_93, a phage protein.[7,8] As AimP concentration increases following

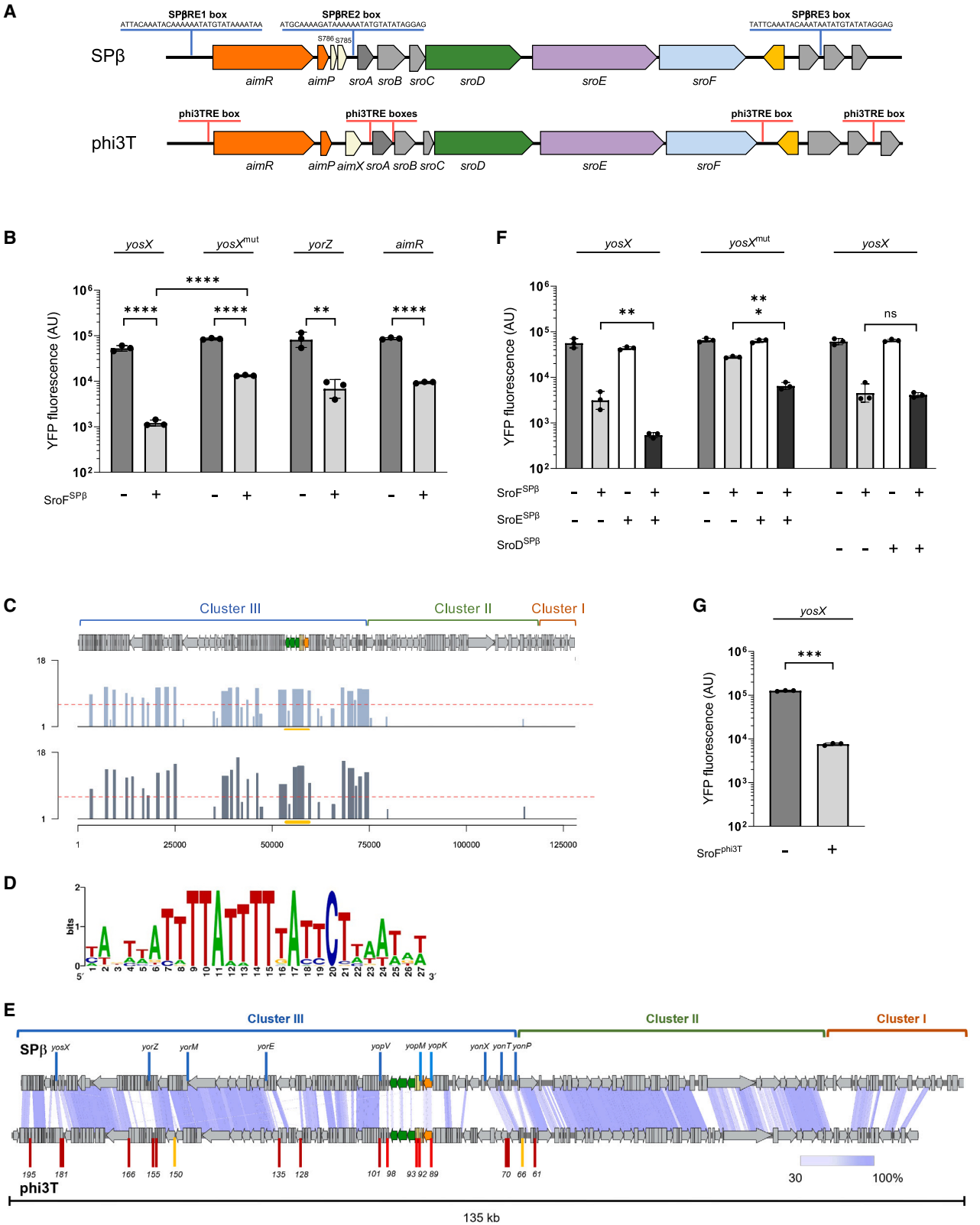


Figure 1. SroE and SroF are the master repressors of phages phi3T and SPβ

(A) Map of the *sro* operons present in SPβ and phi3T, and location of the SPβRE and phi3TRE boxes.

(B) Expression of YFP from *yosX*, *yosX*^{mut}, *yorZ*, and *aimR* promoters in strains containing a xylose-inducible *sroF*^{SPβ} gene.

(legend continued on next page)

multiple phage infections, it inhibits AimR's antitermination function, resulting in loss of *aimX* transcription. In absence of AimX, phi3T_93 binds to MazE, releasing MazF and promoting lysogeny by cleaving lytic transcripts.[7,9]

The *phi3T_93* gene is part of an operon essential for the arbitrium system's functionality.[10] Here, we name this operon the "SPbeta repressor operon" (*sro*), which is conserved in SPbeta group phages encoding arbitrium systems. It contains six genes in phi3T, now renamed *sroA*^{phi3T} to *sroF*^{phi3T} (Figure 1A),[7] including *sroB* (phi3T_93) and the phage master repressor *sroF* (phi3T_97). In SPβ, these operon genes were originally named *yopM* to *yopR* (now renamed *sroA*$^{SP\beta}$ to *sroF*$^{SP\beta}$; Figure 1A). A previous study demonstrated that SroF$^{SP\beta}$ (YopR) is the master repressor of the SPβ phage,[10] whereas SroA$^{SP\beta}$ (YopM), as occurred with SroB^{phi3T}, binds to MazE blocking the lytic cycle.[9]

This study aimed to elucidate the functions of the master repressors SroF^{phi3T} and SroF$^{SP\beta}$ and their connection to the arbitrium system. Note that arbitrium prophages also require, in addition to the arbitrium system, the SOS response for their induction.[10–12] Contrary to classical phages with a single main repressor, our study reveals that SPbeta phages encode two master repressors and three accessory proteins pivotal in regulating lysis-lysogeny decisions, emphasizing the complex control of arbitrium-carrying phages' persistence in nature.

RESULTS

SroF folds like a classical phage integrase

We previously identified SroF^{phi3T} and SroF$^{SP\beta}$ as master repressors in phages phi3T and SPβ, respectively. Mutants lacking SroF (Δ*sroF*$^{SP\beta}$ or Δ*sroF*^{phi3T}) exhibited a lytic phenotype, unable to produce lysogens.[7,10] SroF$^{SP\beta}$ and SroF^{phi3T} proteins exhibit 30% protein-level identity (Table S1) and lack sequence homology with classical phage repressors. Intriguingly, they were annotated as integrase/recombinase proteins. To further characterize the structure and function of this family of repressors, we successfully resolved the structure of SroF$^{SP\beta}$. The crystal structure of SroF$^{SP\beta}$ was solved in two different space groups, P321 and C2, containing one and three SroF$^{SP\beta}$ molecules in their asymmetric units, respectively (Table S2). Structure analysis (PDBePISA) indicated no probable quaternary assemblies (interface score 0.0), as confirmed by size exclusion chromatography coupled to multi-angle light scattering (Figure S1A). The four SroF$^{SP\beta}$ independent molecules from both crystals are almost identical with low values (>0.43 Å) of root-mean-square deviation (RMSD) for the superimposition of the 320 Cα carbon atoms of each molecule (Figures 2 and S1B). Recently, the SroF$^{SP\beta}$ structure was published by another group (PDB: 8A0A).[13] This structure, which was produced under different conditions, shows two monomers in the asymmetric unit, identical to those found in our reported crystalline forms.

A search for structural homologs confirmed the integrase/recombinase fold in SroF$^{SP\beta}$, yielding good similarity (*Z* score higher than 10) with other integrases (Figures S1C–S1E). As in other integrases, the SroF$^{SP\beta}$ protomer can be divided into N-terminal (residues 1–102) and C-terminal domains (residues 110–320), corresponding to the core binding (CB) and catalytic (CAT) domains in tyrosine recombinases, connected by a short linker (103–109). The CB domain is all helical (α1–α5), and the CAT domain is also mostly helical (α6–α16), except for the presence of four short β strands (β1–β4), which are arranged in pairs (β1-β2 and β3-β4) of β-hairpins (Figures 2 and S1B–S1E). Modeling of SroF^{phi3T} with AlphaFold2 showed a high structural similarity with SroF$^{SP\beta}$ (RMSDs of 1.3 and 1.8 Å for the superimposition of CB and CAT domains, respectively), supporting the conserved activity for both proteins (Figures 2A and S1F–S1H). Members of the tyrosine recombinase family show low sequence similarity, but the CB domain is characterized by a conserved R-K-H-R-Y/H pentad of catalytic residues in addition to the tyrosine nucleophile that gives name to the family. Structural alignment of SroF$^{SP\beta}$ and SroF^{phi3T} with model integrase/recombinases showed only the first Lys (K169 and K170 for SroF$^{SP\beta}$ and SroF^{phi3T}, respectively) of this catalytic pentad is conserved in both proteins, and only SroF$^{SP\beta}$, not SroF^{phi3T}, maintains the putative nucleophilic Tyr (Y304) (Figure 2A). This poor conservation of essential residues would suggest both proteins are catalytically inactive as integrase/recombinases, supporting their evolution toward a new function as phage master repressors. The structural alignment also revealed two insertions in the SroF$^{SP\beta}$ CAT domain, the first comprising residues 192–211, which includes the β3-β4 hairpin, and the second from residues 269 to 279, which enlarges α13 (Figure 2). These insertions are conserved in SroF^{phi3T}, and AlphaFold2 predicted an almost identical fold to SroF$^{SP\beta}$ (Figures S1F–S1H). DALI analysis revealed that the first insertion is located in the same place as the unique insertion in integron integrases (Figure 2A).[14] The integron integrases present a β-hairpin in their core (Figures S1E and S1I), which is used to recognize a flipped-out nucleotide that dictates the position of the integrase on the DNA.[15]

The relative arrangement of the CB and CAT domains in the structure of SroF$^{SP\beta}$ differs from what is typically observed in other integrase/recombinases bound to target DNA. Structural comparisons indicated a relative 60° rotation between domains, facilitated by the linker, which would allow SroF$^{SP\beta}$ to adopt a DNA-binding-competent conformation (Figure S1J). The SroF$^{SP\beta}$ structures

(C) SroF^{phi3T} ChIP-seq peaks along the phi3T phage genome (2 replicates). The region corresponding to SPβ repressor operon is highlighted in yellow.
(D) Sequence motif obtained from the ChIP-seq data using MEME.
(E) Localization of the *sroF* binding sites in the SPβ and phi3T genomes. SPβRE and phi3TRE boxes are indicated as blue (SPβ) and red (phi3T) lines. Yellow lines in the phi3T genome indicate phi3TRE boxes with 2 mismatches in the sequence. The adjacent gene, following the operon expression sense, is indicated as a reference to locate the SroF box. Arrows indicate gene direction. ORFs are represented in gray, and the SPβ repressor operon is represented with the sensor, transducer, and repressor modules in orange, yellow, and green, respectively. Lilac shading between genomes indicates regions of pairwise nucleotide identify calculated with txBLASTn.
(F) Expression of YFP from *yosX* and *yosX*^{mut} promoters in strains containing a xylose-inducible *sroF*$^{SP\beta}$ gene and IPTG-inducible SroE$^{SP\beta}$ or SroD$^{SP\beta}$. Expression was measured in the presence or absence of xylose and/or IPTG as indicated by the ± symbols.
(G) Expression of YFP from the *yosX* promoter in phi3T in strains containing a xylose-inducible *sroF*^{phi3T} gene.
In (B), (F), and (G), the geometric means and SDs are presented (n = 3). A two-tailed t test on $\log_{10}$ transformed data was performed to compare mean differences. p values are indicated above the plots: $p \geq 0.05$ not significant (ns), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

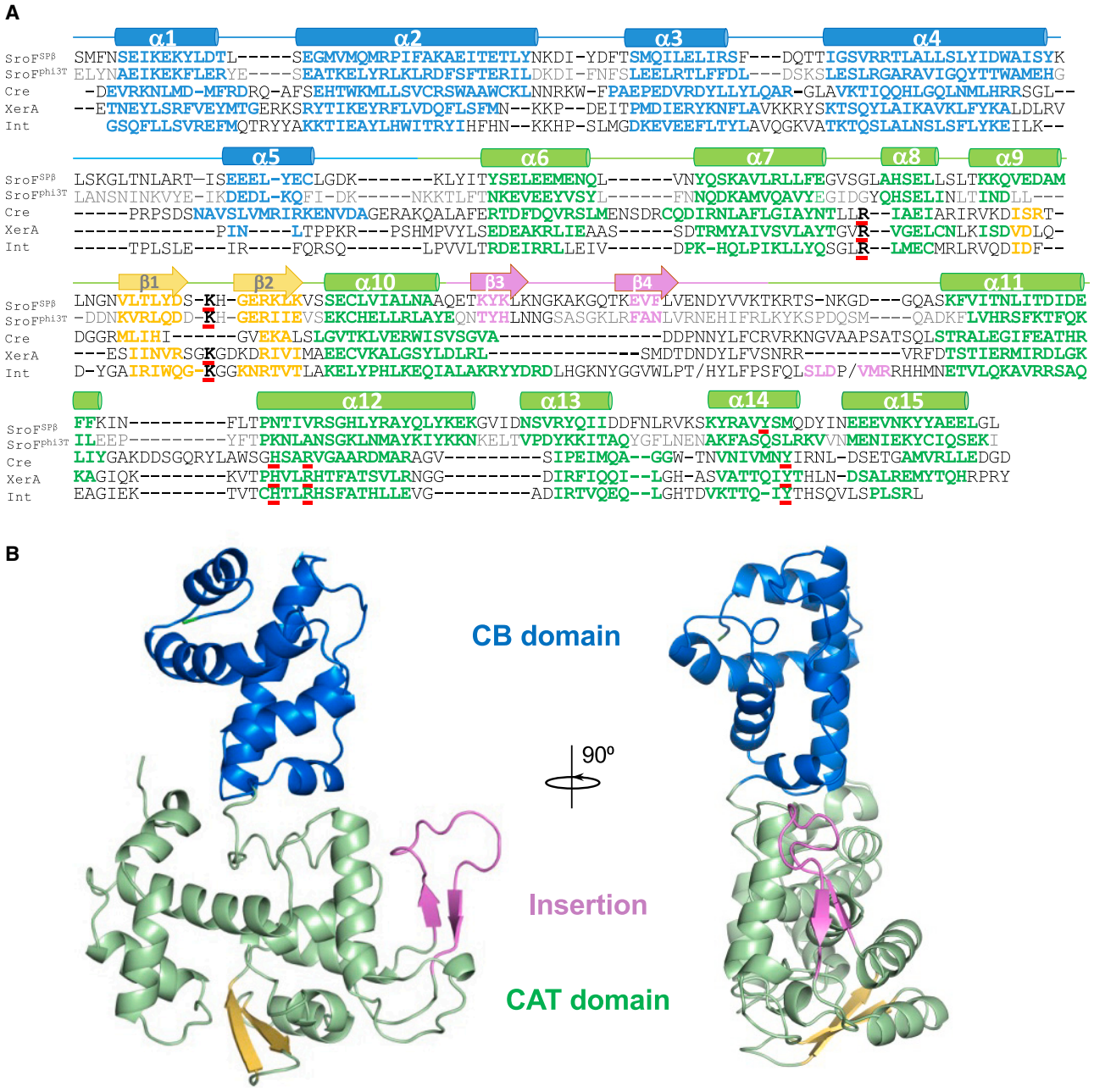


Figure 2. SroF$^{SP\beta}$ shows a tyrosine recombinase fold

(A) Sequence structural alignment of SroF$^{SP\beta}$ and SroF^{phi3T} (AlphaFold2model). Secondary structural elements of SroF$^{SP\beta}$ are shown above the sequence alignment with α helices represented as cylinders (blue and green for core binding (CB) and catalytic (CAT) domains, respectively) and β strands as arrows (orange). The insertion is colored in magenta. Residues in the multiple sequence alignment are colored following the color code of the secondary structure element. Conserved catalytic residues in tyrosine recombinases are underlined in red.

(B) SroF$^{SP\beta}$ structure. Two orthogonal views of SroF$^{SP\beta}$ in cartoon representation and colored as in (A) are shown.

from different conditions and crystal groups, including an additional recently obtained structure, all display identical conformations (Figure S1B), which suggests the observed conformation is biologically relevant and not a result of packaging or crystallization conditions. Consequently, it indicates that the repressor must possess a high degree of plasticity in its repression process to adopt the DNA-binding conformation observed in other integrase proteins, or it performs its function in a novel way.

Conserved residues relevant for integrase activity are not required for repression

Our structure-guided comparison identified residues K170 in SroF^{phi3T} and K169 and Y304 in SroF$^{SP\beta}$ as putative catalytic residues required for the recombination process in tyrosine integrases (Figure 2A).[16] To assess their importance for the repression function, we substituted each of these residues with alanine. These modified versions were cloned into the integrative

expression vector under an isopropyl-β-D-thio-galactopyranoside (IPTG)-inducible promoter and introduced into the *B. subtilis* 168 Δ6 strain, which served as our host strain in most experiments due to its lack of all prophages and other mobile genetic elements.[17] Recipient strains expressing wild-type (WT) or mutant versions of SroF^{phi3T} or SroF^{SPβ} were tested for their ability to complement the phi3T Δ*sroF*^{phi3T} and SPβ Δ*sroF*^{SPβ} mutants upon infection. Note that, in all experiments, both the WT and mutant phages carry a kanamycin resistance marker used for selecting prophage integration. As shown in Figures S2A and S2B, the mutated versions of SroF^{phi3T} and SroF^{SPβ} were capable of complementing lysogenization in their respective mutants, demonstrating that mutations in residues required for the integrase activity of the ancestral protein are no longer necessary for their role as master repressor of the phage. Finally, we did not observe cross-complementation when repressor proteins were used to block infection by non-cognate phages (Figures S2C and S2D). This implies that these repressor systems have evolved to avoid cross-activity and are specific for individual phage lysogeny management.

Identification and characterization of the SroF binding sites

In SPβ, the master repressor binds to a 31 bp sequence region known as SPβRE, distributed across the phage genome.[13,18] Interestingly, three SPβRE boxes are located near the arbitrium repressor operon (Figure 1A): (1) upstream of *aimR* (SPβRE1); (2) upstream of *sroA*^{SPβ} (SPβRE2); and (3) upstream of *yopV*, situated approximately 1 kb downstream of *sroF*^{SPβ} (SPβRE3). We investigated the interaction of SroF^{SPβ} with these SPβRE boxes using electrophoretic mobility shift assays (EMSAs) and bio-layer interferometry (BLI). These methods revealed that SroF^{SPβ} binds to DNA with high affinity. However, the binding appeared non-specific, as no significant differences were observed between the negative control and DNAs containing the SPβRE boxes (Figures S2E and S2F).

Therefore, we employed a transcriptional reporter assay to analyze the effect of SroF^{SPβ}'s binding to SPβRE boxes. We first analyzed SroF^{SPβ} binding to the SPβRE site between the *yosW* and *yosX* genes (Figure 1B). YosX is the first gene of an operon presumably associated with phage lytic induction, suggesting SroF^{SPβ} binding may repress this operon during lysogeny. To test this, we cloned the intergenic region between the two genes into a yellow transcriptional reporter (YFP) and integrated it into a strain containing a xylose-inducible *sroF*^{SPβ} gene. Without xylose, YFP was highly expressed, but the addition of a saturating concentration of xylose, which induces *sroF*^{SPβ} expression, led to a ∼100-fold reduction in YFP expression (Figure 1B). A similar repression pattern was observed in two other YFP reporters of the SPβRE-containing promoters of *aimR* (SPβRE1 box) and *yorZ* (Figures 1A and 1B), indicating SroF^{SPβ} directly controls SPβRE promoters. To further validate this interaction, we mutated the *yosX* SPβRE site (*yosX*^{mut}), without interfering with the putative σ^A binding motif (Figure 1B). This mutation had a minor effect on YFP expression in the absence of xylose and only ∼5-fold reduction in the presence of xylose. These results support the role of SPβRE as the SroF^{SPβ} binding sites. The residual repression suggests that either these mutations are insufficient for full elimination of SroF^{SPβ} binding or, less likely, a weaker SPβRE site is present in another position on the promoter.

We next analyzed whether SroF^{phi3T} represses phi3T by binding to specific boxes across its genome. Chromatin immunoprecipitation with sequencing (ChIP-seq) using a 3xFLAG-tagged SroF^{phi3T} revealed a strong enrichment of reads in the sample lysogenized with phi3T versus insignificant enrichment in the phage-free strain. Among the 24 significantly enriched peaks identified in the lysogenic strain, all but one were located in the phi3T genome (Figure 1C). The phi3T genome has been divided into three gene clusters (I–III) based on transcription orientation, with clusters II and III housing late and early phage genes, respectively. Most peaks mapped to cluster III, linked to phage induction.[19] Their summits were located in intergenic regions, as expected for repressors. Using the MEME algorithm, we identified a 27 bp motif (p value 1.8e^{-74}) and a highly conserved central 16 bp sequence ATTTTATTTTTATTCT, which we designated as the phi3TRE box (Figure 1D). There were 16 fully conserved phi3TRE boxes and 19 additional boxes with only one (10 sites) or two (9 sites) mismatches present in the phi3T genome (Figure 1E; Table S3). All these sites were found in the enriched peaks in the ChIP-seq assay, with several presenting multiple phi3TRE boxes with variable (32–424 bp) distances between them (Figures 1C and 1E; Table S3). Therefore, in both SPβ and phi3T, multiple SroF binding sites seem to be required for repression. Testing the phi3T *yosX* promoter region using the YFP reporter system showed a ∼10-fold reduction in YFP expression in the presence of xylose (presence of SroF^{phi3T}), confirming conserved binding sites in phi3T (Figure 1F). Despite sequence differences, SPβRE and phi3TRE boxes were mostly present in the same genomic regions (cluster III), including the arbitrium and *sro* operons as well as early lytic genes (Figure 1E), supporting a conserved mechanism of action.

In classical λ-like phages, CI expression in recipient cells leads to the selection of lytic phages with mutations in CI binding sites. Due to multiple SroF binding sites in SPbeta phages, we investigated whether these phages could evade the repressive effects of SroF proteins in recipient cells. We challenged strains overexpressing SroF^{phi3T} and SroF^{SPβ} with their respective Δ*sroF*^{phi3T} and Δ*sroF*^{SPβ} lytic phage lysates to isolate escape phages using plaque assays. Although we could not obtain escape mutants for SPβ Δ*sroF*^{SPβ} phage challenged against SroF^{SPβ}, we isolated 4 distinct evolved phi3T Δ*sroF*^{phi3T} mutants insensitive to SroF^{phi3T}. As expected, all these mutant phages were unable to generate lysogens even in the presence of SroF. Sequencing revealed two of these evolved phages had significant genomic deletions, including the loss of all the arbitrium system and the repressor operon genes (*aimP-aimR* and *sroABCDEF*; Table S4). The third evolved phage had a deletion in the *sroBCDEF* region, suggesting one of the *sroBCDE* genes is also important for lysogeny. In support of this, the final mutant phage had a single nucleotide mutation in *sroE* (A22P). This led to the hypothesis that SroE may be an important component of the phi3T repression system. In summary, these findings suggest SroF^{phi3T} and SroF^{SPβ} execute their roles as phage repressors in a manner distinct from conventional phage repressors.

SroE is the second master repressor

Since the previous results suggested additional proteins encoded by the phi3T *sro* operon could be involved in establishing lysogeny, we engineered phi3T prophages bearing individual

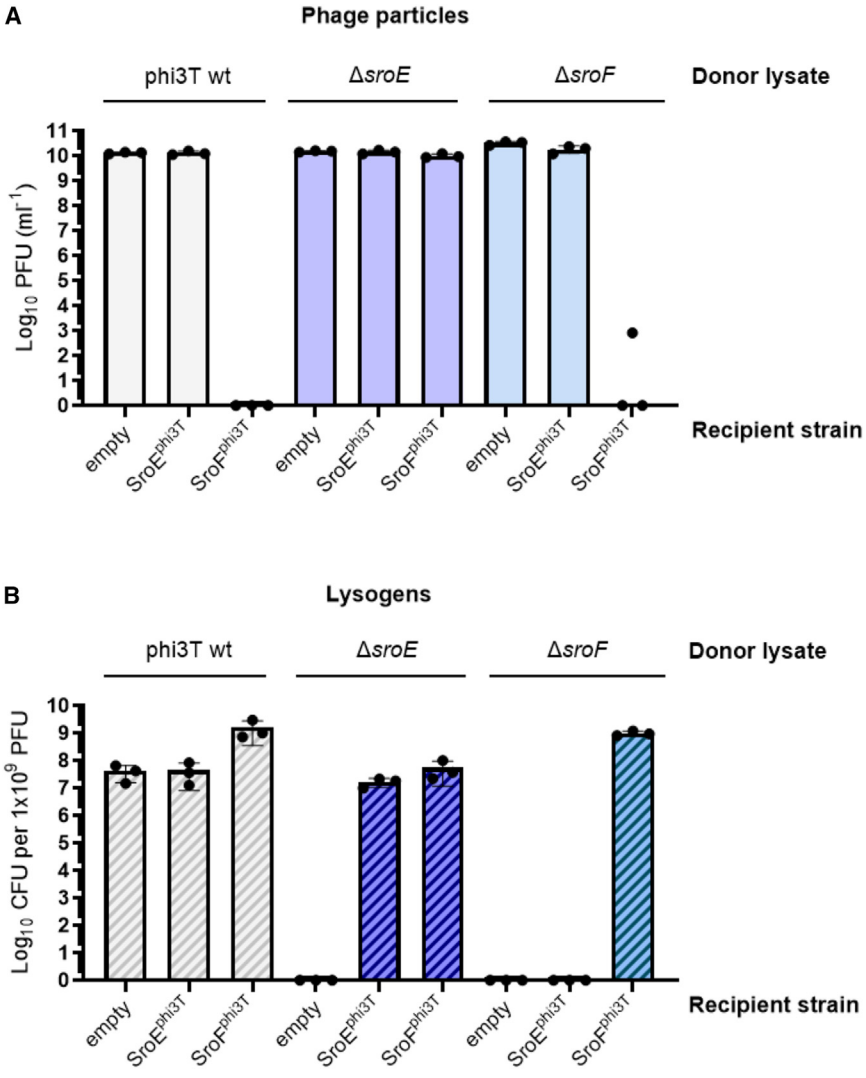


Figure 3. Characterization of the two phi3T master repressors

Strains carrying phages phi3T WT, phi3T Δ*sroE*-^{phi3T} *amyE*::P_{spank}-*sroE*^{phi3T}, or phi3T Δ*sroF*^{phi3T} *amyE*::P_{spank}-*sroF*^{phi3T} were MC induced. The number of plaques (A) or lysogens (B) was quantified using 168 Δ6 *amyE*::P_{spank}(empty), 168 Δ6 *amyE*::P_{spank}-*sroE*^{phi3T} or 168 Δ6 *amyE*::P_{spank}-*sroF*^{phi3T} as recipient strains. Results are represented as log PFUs mL^{-1} (A) or colony-forming units (CFUs) mL^{-1} normalized by PFUs mL^{-1} and represented as the log CFU of an average phage titer (1 × 10^9 PFU) (B). Geometric means and SDs are presented (n = 3). Values below detection limit (1 × 10^1 PFUs mL^{-1} or CFUs mL^{-1}) are marked on axis.

mutations in the different genes of the operon. These mutant prophages were induced with mitomycin C (MC) to activate the SOS response, and the ability to generate infective particles was assessed. Concurrently, we quantified their capacity to generate lysogens after infection. In parallel, we individually overexpressed the *sro* operon genes in non-lysogenic cells and evaluated the capacity of the expressed proteins to block phage infection (results summarized in Figure S3).

It was not possible to mutate *sroE*^{phi3T} (*phi3T_96*) in the phi3T prophage, suggesting SroE also plays a crucial role in maintaining lysogeny. Therefore, we complemented the lysogenic strain with an ectopic copy of *sroE*, enabling us to generate a phi3T Δ*sroE* mutant strain. After MC induction of this mutant prophage, the phage titer was identical to the WT phage (Figure 3A; empty plasmid). However, the plaques generated by the Δ*sroE* mutant were larger and clearer than those observed for the WT phage, resembling the characteristics of the Δ*sroF* mutant.

Next, lysates containing phi3T WT, Δ*sroE*, or Δ*sroF* mutant phages were used to infect cells expressing SroE^{phi3T}, SroF^{phi3T}, or empty plasmid. Overexpression of SroF^{phi3T} blocked plaque formation and increased lysogenization of the WT and Δ*sroF* mutant phages (Figure 3). By contrast, SroF^{phi3T} expression did not impede the Δ*sroE* mutant phage's ability to infect the cells (Figure 3A) but largely restored its ability to generate lysogens (Figure 3B). Intriguingly, in contradiction to the typical role of a master repressor, overexpression of SroE^{phi3T} did not block infection of any of the tested phages (Figure 3A) and did not augment lysogenization of the phi3T WT or Δ*sroF* phages (Figure 3B). Although SroE^{phi3T} expression increased lysogenization of the Δ*sroE* mutant phage, this increment was similar to that observed when this mutant was complemented with SroF^{phi3T} (Figure 3B). In summary, these results confirm the existence of two crucial repressors, with SroF^{phi3T} playing a more prominent role.

Since SPβ encodes a SroE protein that shows 38% identity to that encoded by phi3T (Table S1), we explored whether this protein also functioned as a second repressor. Indeed, it was not possible to delete the *sroE*^{SPβ} gene without ectopic complementation. Induction of the complemented Δ*sroE*^{SPβ} mutant resulted in a higher titer and sharp plaques (Figure S4B). Further, overexpression of either SroE^{SPβ} or SroF^{SPβ} was required to restore the ability of the Δ*sroE*^{SPβ} mutant to form lysogens (Figure S4C). These results confirm that SroE^{SPβ} and SroF^{SPβ} are functional homologs of the repressor module in phi3T.

To investigate whether SroE^{SPβ} could directly bind to the SPβRE sites or promote lysogeny by enhancing SroF^{SPβ}'s binding to these regions, we cloned an IPTG-inducible *sroE*^{SPβ} and integrated this construct into a strain harboring the P_{xyl}-*sroF*^{SPβ} and P_{yosX}-YFP constructs. We found that IPTG induction had no effect on *yosX* reporter expression, suggesting that SroE^{SPβ} does not repress this promoter directly (Figure 1E). To examine whether SroE^{SPβ} facilitates SroF^{SPβ} function, we expressed SroF^{SPβ} with a low level of xylose, resulting in mild repression of the *yosX* reporter. Addition of IPTG (SroE^{SPβ} expression) significantly increased the level of repression mediated by SroF^{SPβ}, suggesting SroE^{SPβ} enhanced SroF^{SPβ} binding to the SPβRE sites (Figure 1F).

These results suggested SroE and SroF could work together by binding DNA in a specific manner. However, EMSA assays showed that at high concentrations, SroE$^{SP\beta}$ displayed non-specific DNA binding, as was the case with SroF$^{SP\beta}$ (Figure S3C). Intriguingly, no differences in DNA affinity were observed when both SroE$^{SP\beta}$ and SroF$^{SP\beta}$ were present (Figure S2H). Accordingly, we were not able to detect the SroE$^{SP\beta}$-SroF$^{SP\beta}$ interaction *in vitro* using BLI.

We next used AlphaFold2 to generate structural models of SroE^{phi3T} and SroE$^{SP\beta}$. Both models present a similar three-domain structure connected by flexible linkers (Figure S4). The N-terminal domain (residues 1–105) harbors homology to DNA-binding POU domains, the middle domain (residues 110–275) shares similarity with the nuclease domain in the chromosome-partitioning protein ParB, and the C-terminal domain (residues ~290-end) resembles domains found in proteins involved in oligomerization. These results suggest SroE might possess a distinctive global folding, potentially enabling interactions with DNA, aligning with the repressor function inferred from *in vivo* experiments.

SroD is also required to efficiently establish lysogeny

SroD^{phi3T} (phi3T_95) and SroD$^{SP\beta}$ (YopP) share 34% identity (Table S1), and AlphaFold2 models revealed a characteristic integrase/tyrosine recombinase fold (Figure S4). After induction, the phi3T Δ*sroD* mutant generated slightly fewer infective particles and smaller plaques compared with the WT phage. Intriguingly, the lysogen formation was significantly reduced, and colony morphology was altered (Figures 4A–4C). Whereas phi3T produced normal lysogens, the Δ*sroD* mutant consistently produced a mixture of small, more translucent colonies alongside normal ones (Figure 4B). Complementation in recipient cells restored these phenotypes and did not block WT phi3T infection (Figures 4D and 4E). An overview of the various phenotypes observed for both phi3T and SPβ *sro* mutants, including this mutant, is illustrated in Figure S3.

Subsequent investigation of these lysogens' morphologies showed the colony phenotypes remained stable after restreaking. Sequencing analysis confirmed the presence of phi3T in both normal and small colonies but unveiled substantial deletions in the small colonies, affecting both the prophage and chromosomal regions adjacent to the *attR* site (Table S4). These large (70–137 kb) chromosomal deletions may be responsible for the observed phenotype of the small colonies. Therefore, we hypothesize that SroD is required for phage integration. The normal colonies might be a result of low-frequency correct integration of the Δ*sroD* mutant phage, whereas the small colonies result from aberrant integration or unstable lysogen formation, which is stabilized through recombination between the phage genome and the chromosome.

SroD$^{SP\beta}$ shows sequence homology to SroD^{phi3T} and seems to have a similar function. Induction of the Δ*sroD*$^{SP\beta}$ mutant resulted in fewer plaque-forming units (PFUs) and smaller plaques (Figure S5). As with the Δ*sroD*^{phi3T} mutant, impaired lysogen formation and two types of lysogenic colonies were observed (Figure S5D). Expression of either *sroD*^{phi3T} or *sroD*$^{SP\beta}$ in recipient cells trans-complemented the ability of the phi3T or SPβ Δ*sroD* mutant phages to form lysogens (Figures 4F and 4G), yielding normal colony morphologies. Finally, we analyzed whether SroD$^{SP\beta}$ binds directly to the SPβRE boxes repressing expression or whether, conversely, it facilitates SroF$^{SP\beta}$ binding to these regions. Our results indicate that neither of these scenarios occurred *in vivo* (Figure 1F). Taken together, our results indicate the SroD proteins promote lysogeny through a conserved yet enigmatic mechanism.

Characterization of the remaining genes in the repression operon

We continued characterizing the genes within the arbitrium operon. Although SroDEF proteins from phi3T and SPβ showed significant sequence homology, SroABC proteins did not (Table S1). However, we hypothesized analogous functions in their respective phages. The *sroC*^{phi3T} gene (*phi3T_94*) encodes a 39 aa putative protein adopting a helix hairpin-like structure (AlphaFold2 model; Figure S5). Given its small size and the prevalence of this motif in diverse proteins, such as TATA box-binding proteins and zinc transporters, inferring a function from its putative structure proved challenging. SroC$^{SP\beta}$, a 71 aa protein, is confidently predicted to adopt a helix-turn-helix structure, common in DNA-binding proteins (Figure S4). In both phages, the Δ*sroC* mutant showed no significant differences in phage titer or lysogen formation following MC induction compared with the WT phage (Figures S5E, S5F, and S6). Overexpressing SroC^{phi3T} or SroC$^{SP\beta}$ in recipient cells had no effect on plaque formation of cognate phages during titering (Figure S5G). Therefore, the role of the SroC proteins remains unclear.

SroB^{phi3T} (phi3T_93) is a dimeric protein that has been extensively characterized elsewhere.[7] Its expression promotes lysogeny by binding to MazE to facilitate MazF release.[7,9] By contrast, during the lytic cycle of the phage, AimX^{phi3T} binds to MazF and SroB^{phi3T}, inactivating their function.[7] However, the Δ*sroB*^{phi3T} mutant exhibited no significant difference in phage titer or lysogen formation following induction compared with the WT phage.[7] In phage SPβ, SroB$^{SP\beta}$ is predicted to adopt a three-helix bundle, a different structure from the experimentally resolved SroB^{phi3T} (Figure S4). Deleting the *sroB*$^{SP\beta}$ gene also did not impact the titer of the mutant after induction, but it produced larger and clearer plaques compared with the WT phage.[10] Moreover, deletion of *sroB*$^{SP\beta}$ in the *aimR* mutant background resorted the defect generated by the *aimR* mutation.[10]

SroA^{phi3T} (phi3T_92) is a 68 aa protein with a low-confidence helix-loop-helix fold prediction (AlphaFold2; Figure S4). After induction, the Δ*sroA*^{phi3T} mutant resulted in similar PFUs but generated 100 times fewer lysogens than the WT phi3T (Figures 5A and 5B). This suggested that SroA^{phi3T} plays a role in promoting lysogeny, and the overexpression of SroA^{phi3T} in recipient cells might block phi3T propagation. However, attempts to transform a vector carrying the *sroA*^{phi3T} gene with its native upstream promoter into *B. subtilis* Δ6 were unsuccessful. Only one clone was obtained, which carried a single nucleotide variation (SNP) located 3 bp upstream of the *sroA*^{phi3T} start codon (Figure 5C), predicted to target either the ribosomal binding site (RBS) spacer or the RBS itself, potentially affecting the translation levels of SroA^{phi3T}. A plasmid with this SNP (*sroA*^{mut}) was successfully transformed into *B. subtilis* 168 Δ6, and SroA^{phi3T} expression blocked phi3T phage propagation drastically (Figure 5D).

Since SroB^{phi3T} binds to MazE liberating the cellular toxin MazF,[7,9] the fact that overexpression of SroA^{phi3T} was lethal

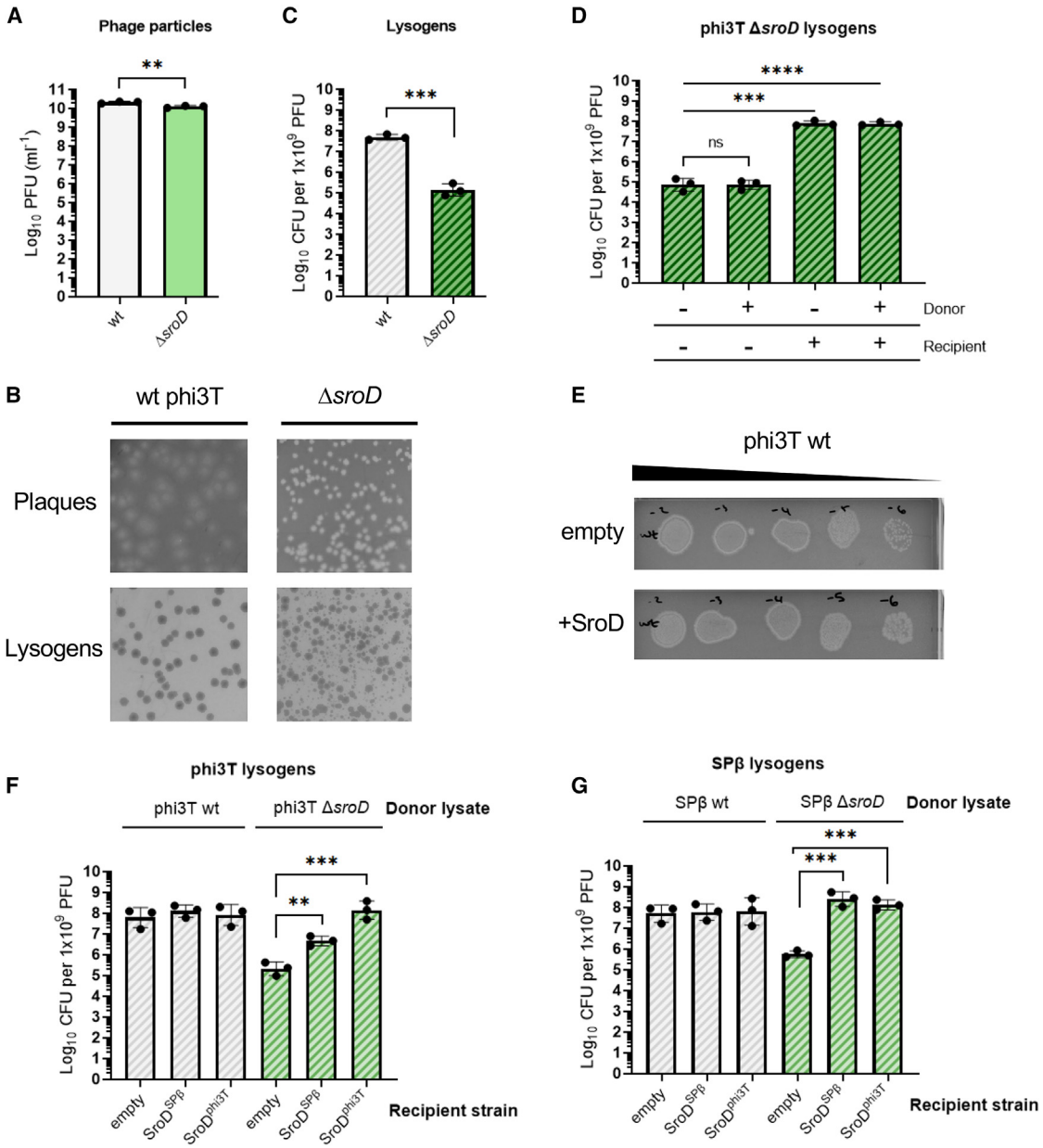


Figure 4. phi3T Δ*sroD* mutation impacts lysogenization and morphologies of plaques and colonies

Strains carrying phages phi3T WT or Δ*sroD* were MC induced. The number of plaques (A) or lysogens (B) was quantified using 168 Δ6. The plaque and colony morphologies were photographed (C). Complementation (*amyE*::P_{spank}-*sroD*^{phi3T}) was performed in the donor and/or recipient cells as indicated by the ± symbols on the x axis (D). phi3T WT lysate dilutions were spotted on lawns of Δ6 *amyE*::P_{spank}(empty) and Δ6 *amyE*::P_{spank}-*sroD*^{phi3T} recipient cells (E). To test cross-complementation, strains carrying phages phi3T and SPβ (WT or Δ*sroD*) were MC induced. The number of lysogens was quantified using 168 Δ6 *amyE*::P_{spank} (empty), 168 Δ6 *amyE*::P_{spank}-*sroD*^{SPβ} or Δ6 *amyE*::P_{spank}-*sroD*^{phi3T} as recipients (F and G). Results are represented in (A) as log PFUs mL^{-1} and in (C), (D), (F), and (G) as CFUs mL^{-1} normalized by PFUs mL^{-1} and represented as the log CFU of an average phage titer (1×10^9 PFU). Geometric means and SDs are presented (n = 3). In (A), (C), (D), (F), and (G), two-tailed t tests were performed on $\log_{10}$-transformed data to compare mean differences. p values are indicated above the plots: $p \geq 0.05$ not significant (ns), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

for the cell suggested a possible interaction between SroA^{phi3T} and the MazEF module. Unfortunately, we were unable to conduct these experiments due to the insolubility of the SroA^{phi3T} protein. Therefore, we introduced a plasmid expressing the WT *sroA* construct into recipient cells with deletions in either *mazF* or both *mazE-mazF*. This rendered SroA^{phi3T} non-toxic, and we successfully obtained the required clones, verified through PCR and sequencing. This suggests that SroA^{phi3T} might also interact with the MazEF module to facilitate lysogenization. The specific mechanism, whether it involves direct binding to MazF or interaction via MazE (as occurred with SroB^{phi3T}), remains to be determined.

In the SPβ phage, SroA^{SPβ} (66 aa) was predicted to fold as a single long helix (Figure S4). Following induction, a

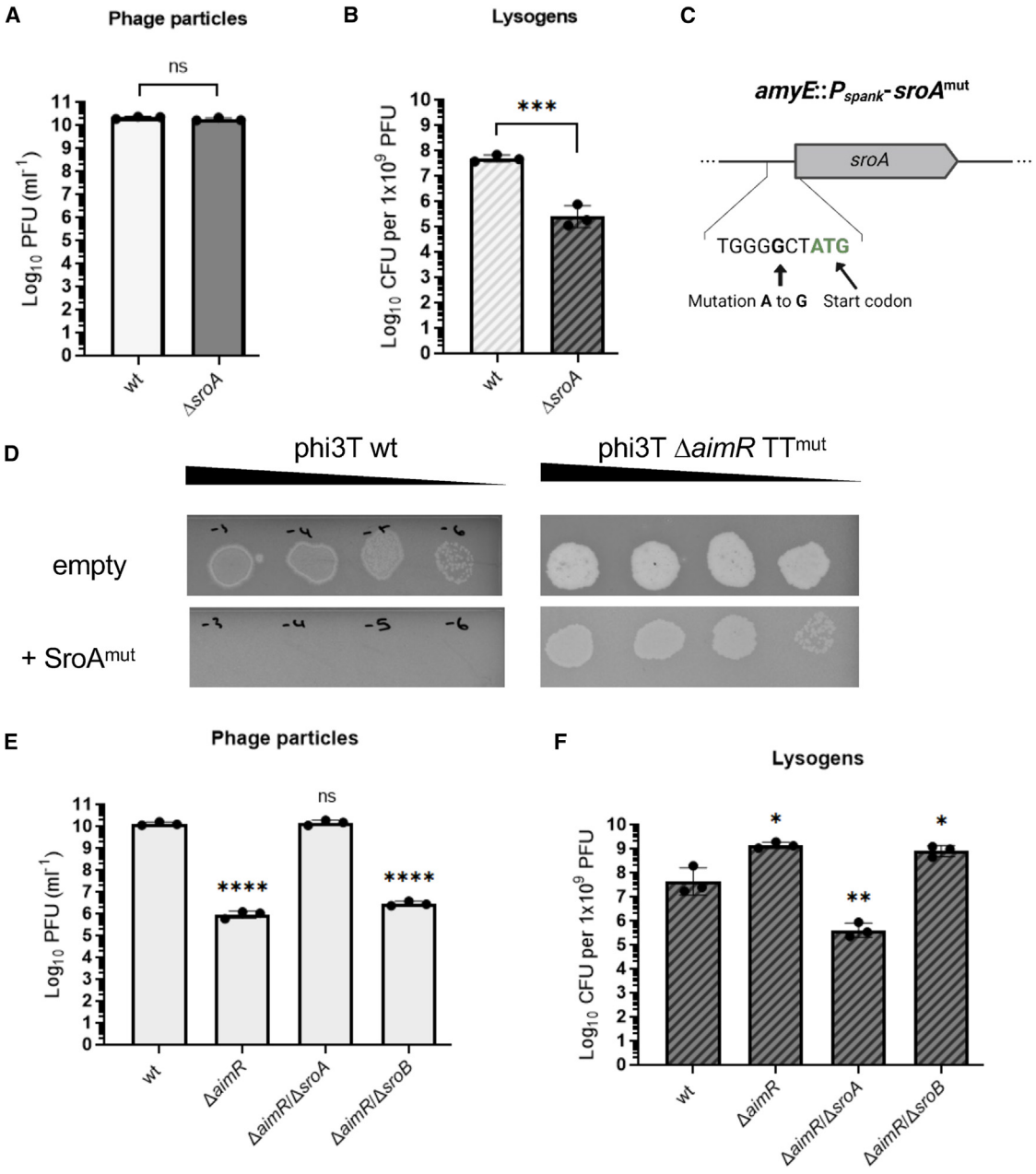


Figure 5. Characterization of phi3T sroA

(A and B) Strains carrying phages phi3T WT or Δ*sroA* were MC induced. The number of plaques (A) or lysogens (B) was quantified using 168 Δ6.

(C) Schematic of the genetic layout where the mutation was identified following transformation of the *sroA* vector. The DNA sequence at the bottom shows the mutated base denoted with an arrow and the *sroA* start codon in green.

(D–F) (D) Phage phi3T WT and phi3T Δ*aimR* TT^{mut} lysate dilutions were spotted on lawns of Δ6 *amyE*::P_{spank}(empty) and Δ6 *amyE*::P_{spank}-*sroA*^{mut} recipient cells. Strains carrying phages phi3T WT, phi3T Δ*aimR*, and double mutants phi3T Δ*aimR*/Δ*sroA*, phi3T Δ*aimR*/Δ*sroB* were MC induced, and the number of plaques (E) or lysogens (F) was quantified and represented as log PFUs mL^{-1} in (A) and (E) and as CFUs mL^{-1} normalized by PFUs mL^{-1} and represented as the log CFU of an average phage titer (1 × 10^9 PFU) in (B) and (F). Geometric means and SDs are presented (n = 3). A two-tailed t test was performed on $\log_{10}$-transformed data to compare mean differences between WT and mutant phages. p values are indicated above the plots: $p \geq 0.05$ not significant (ns), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

Δ*sroA*$^{SP\beta}$ mutant produced similar levels of infective particles but generated fewer lysogens upon infection than the WT SPβ (Figures S6A and S6B). Overexpression of *sroA*$^{SP\beta}$ in recipient cells significantly blocked SPβ phage propagation (Figure S6C). These similar results for phi3T and SPβ suggest a generalized function for SroA proteins in promoting the lysogeny or suppressing the lytic cycle in SPbeta phages. In fact, a recent report also indicates SroA$^{SP\beta}$ binds to MazE, activating MazF's activity and promoting lysogeny.[9]

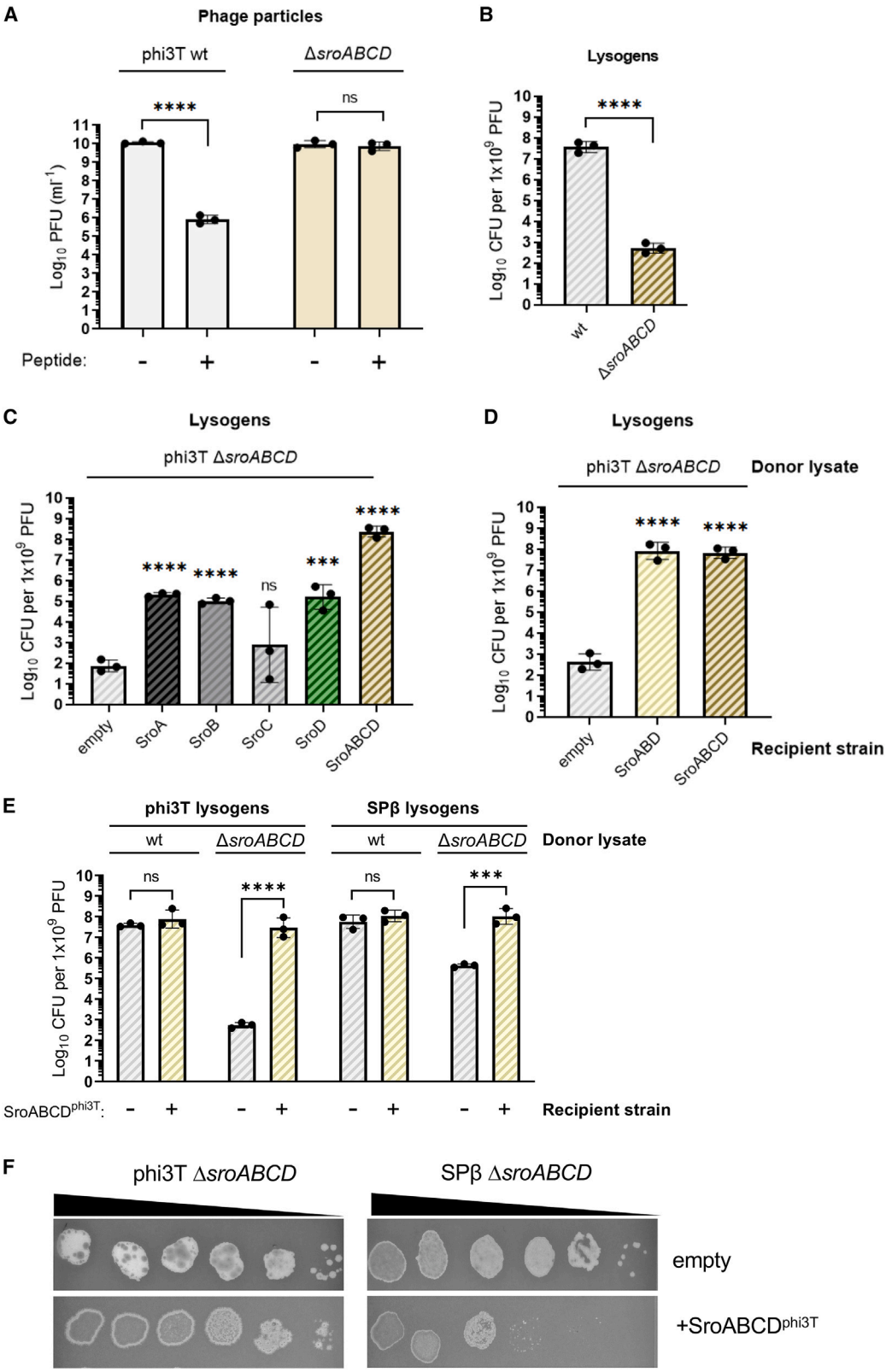


Figure 6. Characterization of phi3T and SPβ *sroABCD* mutant

(A) Strains carrying phages phi3T WT or Δ*sroABCD* were MC induced in the presence (+) or absence (–) of AimP^{phi3T} peptide (5 μM). The number of phages was quantified and is represented as log PFUs mL^{-1}.

(legend continued on next page)

Characterization of a phi3T mutant prophage expressing SroE and SroF

Since we were able to remove all individual arbitrium repressor operon genes, except SroE^{phi3T} and SroF^{phi3T}, we also generated a phi3T Δ*sroABCD*^{phi3T} mutant prophage. Induction of this mutant resulted in similar PFUs as the WT phi3T prophage (Figure 6A). However, this deletion substantially impaired lysogen formation, suggesting a preference for the lytic cycle (Figure 6B).

Next, we MC induced the phi3T Δ*sroABCD*^{phi3T} mutant prophage in the presence of the AimP peptide. Unlike the WT phage, whose induction is hampered by AimP, the Δ*sroABCD*^{phi3T} mutant phage titer remained unaffected in AimP's presence (Figure 6A). These results indicated: (1) the SOS response can nullify the activity of the two master repressors in the system; (2) SroABCD proteins are dispensable for prophage maintenance but crucial for efficient lysogeny establishment post-infection; and (3) the induction of the Δ*sroABCD*^{phi3T} mutant phage is unaffected by the arbitrium peptide, suggesting the arbitrium system regulates the SroABCD proteins, rather than the master repressors.

Therefore, we tested the ability of the Δ*sroABCD*^{phi3T} mutant to generate lysogens in recipient cells expressing individual *sroABCD*^{phi3T} genes. Complementation with *sroA*^{phi3T}, *sroB*^{phi3T}, or *sroD*^{phi3T} led to partial restoration of lysogen formation, but no single gene could fully recover lysogeny to WT levels (Figure 6C). Full recovery was achieved by expressing either all four operon genes (*sroABCD*^{phi3T}) or the three genes with clear phenotypes (*sroABD*^{phi3T}) (Figures 6C and 6D).

The main role for AimR is to allow the expression of AimX, which inhibits MazF function. Since both SroA^{phi3T} and SroB^{phi3T} appear to activate MazF function, we hypothesized that these two proteins would be involved in linking the arbitrium system with the activity of the two master repressors. We propose the following mechanism of action: during infection, SroA, SroB, and AimX are expressed. SroA and SroB will liberate MazF, whereas AimX will block MazF function, promoting the lytic cycle. Once AimP concentration increases, SroA and SroB continue liberating MazF, but AimX is absent, thus allowing MazF activity to promote lysogeny.[7]

To test this, we obtained derivative phi3T prophages carrying double *aimR*/*sroA* or *aimR*/*sroB* mutations. We postulated that the defect in the induction of the phi3T *aimR* mutant, resulting from the absence of AimX, could be compensated for by the mutation in either *sroA* or *sroB*. Surprisingly, *sroB* deletion did not compensate for the *aimR* defect, but the double *aimR*/*sroA* mutant prophage generated 100× times more phage particles than the *aimR* single mutant, similar to WT phage induction (Figure 5E). However, as observed with the single *sroA* mutant, the *aimR*/*sroA* double mutant generated significantly fewer lysogens (Figure 5F).

We further validated this model using a phi3T Δ*aimR* TT mutant (phi3T Δ*aim*R TT^{mut}), which constitutively expresses *aimX*.[7] Since we hypothesize that SroA blocks phage propagation by promoting MazF activity, we tested if the inhibitory effect of SroA could be compensated by the overexpression of AimX. As anticipated, the phi3T Δ*aim*R TT^{mut} mutant phage was able to infect the recipient cells overexpressing SroA^{phi3T} (Figure 5D).

In silico analysis of the arbitrium and the arbitrium repressor operon present in SPbeta family

Recently, a taxonomic classification categorized 64 SPβ-like prophages from 10 *Bacillus* species into four main species clusters: *SPbeta*, *eta*, *bimanducare*, and *magnus*, plus three additional orphan phages.[20] To determine the operon genes conservation in the SPbeta phage family, we extracted the genomic regions encompassing the upstream *aimR* module and the downstream *sro* operon in 58 of those phages (full genomes for 6 phages were unavailable) (Table S5). We conducted open reading frame (ORF) distribution assessments and constructed alignments for the complete operon and individual ORFs, used in subsequent phylogenetic analyses (Figure 7A). Our bioinformatic analyses confirmed the ubiquitous presence of the arbitrium repressor operon across all prophages.

The operon and arbitrium system can be organized into three distinct modules based on sequence homology. The central region spanning 1,400–2,700 bp exhibited lower conservation compared with the rest, with pairwise global alignment revealing peaks of identity below 30% (Figure 7B). This region includes *sroA*, *sroB*, *sroC*, and *aimX*, although not every phage includes all these genes. By contrast, the upstream region containing *aimR* and *aimP*, and the downstream region containing *sroD*, *sroE*, and *sroF*, were more conserved.

An unrooted phylogenetic analysis of the *sro* operon and arbitrium system did not overlap with the previously proposed species clustering[21] (Figure 7A). In particular, phi3T and SPβ, the prototypical phages of the SPbeta cluster, appeared in distant branches. The phylogenetic differences in distribution of *sro* in respect to the entire phage genomes[21] and its internal variability suggest a modular organization of the *sro* operon, potentially arising from various recombination events. *In silico* analyses highlighted low sequence conservation within the region between *aimP* and *sroD* in the arbitrium repressor operons (Figure 7B). We delved deeper into this variability by focusing on seven representative prophages to explore their arbitrium repressor operon composition (Figure 7A). We observed variations in the number of ORFs within this region. Notably, only phages W1 and phi3T encode an AimX-like protein and showed similarities between SroABC-like proteins, whereas prophages KBN06P03352 and CSL2 lacked SroC (Figures 7B and S9).

(B) Strains carrying phages phi3T WT or Δ*sroABCD* were MC induced. The number of lysogens was quantified.
(C and D) Complementation of lysogeny was completed using 168 Δ6 *amyE*::P_{spank} harboring an empty control, the different individual operon genes (*sroA*, *sroB*, *sroC*, *sroD*) (C) or the region encoding *sroABD* and the full operon, *sroABCD* (D).
(E) To test cross-complementation, strains carrying phi3T WT, phi3T Δ*sroABCD*, SPβ WT or SPβ Δ*sroABCD* were MC induced. The number of lysogens was quantified using Δ6 *amyE*::P_{spank}(empty) and Δ6 *amyE*::P_{spank}-*sroABCD*^{phi3T} as recipient cells, as indicated by ± symbols under the x axis.
(F) phi3T Δ*sroABCD* and SPβ Δ*sroABCD* dilutions were spotted on lawns of Δ6 *amyE*::P_{spank}(empty) and Δ6 *amyE*::P_{spank}-*sroABCD*^{phi3T} recipient cells.
Results are represented in (A) as log PFUs mL^{-1} and in (B), (C), (D), and (E) as CFUs mL^{-1} normalized by PFUs mL^{-1} and represented as the log CFU of an average phage titer (1 × 10^9 PFU). Geometric means and SDs are presented (n = 3). Two-tailed t tests were performed on $\log_{10}$-transformed data to compare mean differences. In (C), (D), and (E) comparisons were made against the recipient harboring an empty vector. p values are indicated above the plots: $p \geq 0.05$ not significant (ns), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

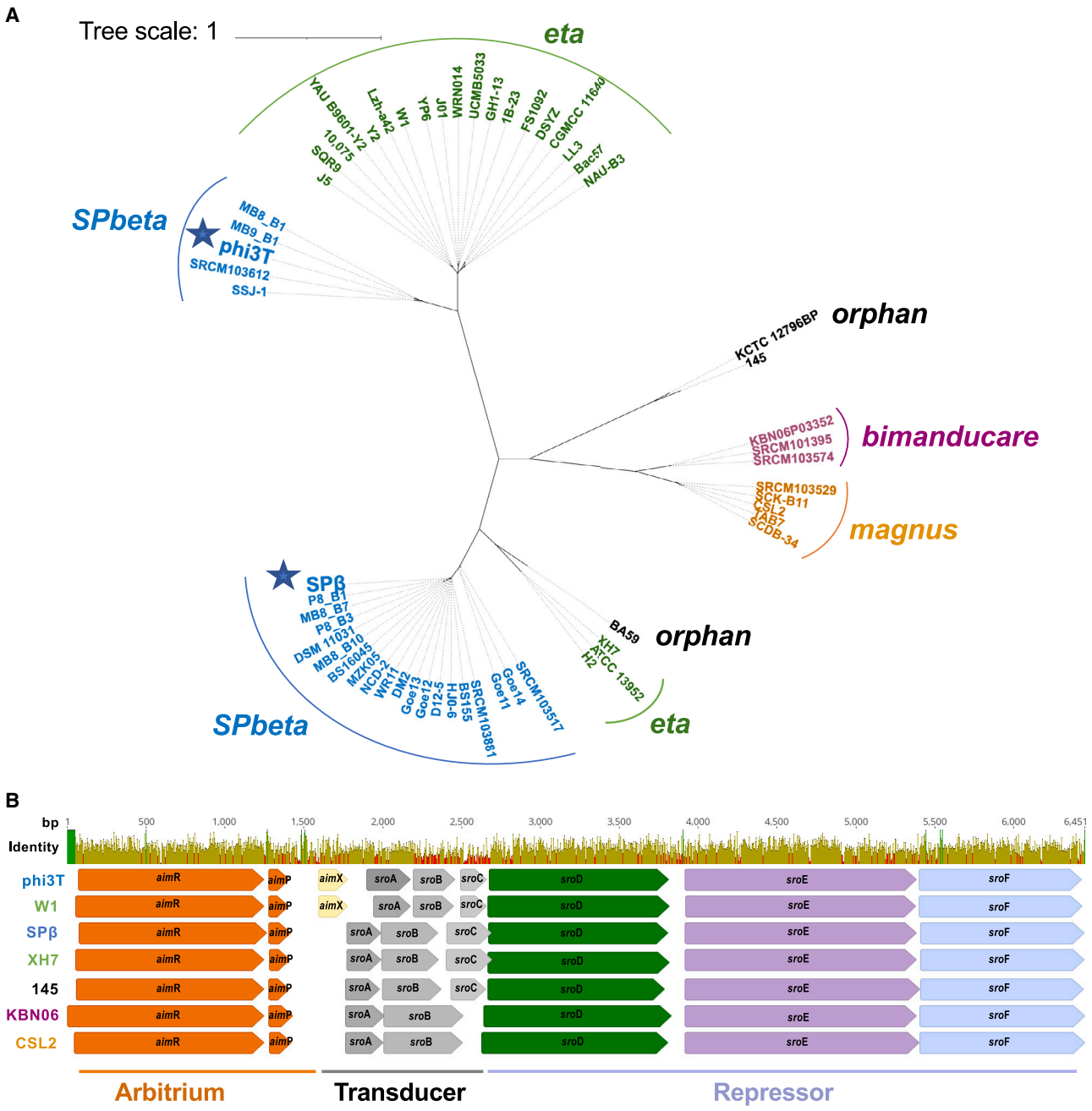


Figure 7. Arbitrium and repressor operon analysis in SPbeta-like phages

(A) An unrooted phylogenetic tree based on a multiple sequence alignment of the arbitrium and the repressor operon from 58 SPbeta-like phages.

(B) Pairwise global alignment of seven representative phages selected based on the clustering groups formed in the unrooted phylogeny. Mean pairwise identity over all pairs in the column is represented with colors in the top: green represents 100% identity, green-brown represents at least 30% and under 100% identity, and red represents below 30% identity. Encoded ORFs are displayed and labeled with the conserved proteins in the arbitrium and repressor modules shown in the same color. ORFs with low sequence similarity in the transducer module are displayed in different shades of gray. The ORF for AimX is colored in pale yellow.

Characterization of the arbitrium and the repressor operon modules

In summary, our *in silico* and experimental analyses suggest the arbitrium and the arbitrium repressor operon comprise three functional modules: a conserved sensor module (*aimR* and *aimP*), a conserved repression module (*sroD-F* genes), and a less conserved transducer module (*sroA-sroC*) that may also include *aimX*. Given that AimX can interact with SroB in phi3T and SroA and SroB may modulate MazF activity by interacting with MazE,[7,9] our results suggest the transducer module likely serves as a bridge connecting the sensor and repression modules. The presence of different transducer modules among phages may contribute to functional adaptation to control lysis-lysogeny.

To validate the presence of these three modules, we employed two strategies. First, we hypothesized that some phages might carry the repression module but lack the sensor and transducer modules. Our investigation of several phage genomes bearing *sroE-sroF* master repressor genes revealed that most SPβ-like phages indeed harbor the *sro* genes. However, we observed genetic variation, and in most cases, the absence of the arbitrium system (sensor module) correlated with the absence of the transducer module (*sroABC*). Nonetheless, the *sroD* gene consistently preceded *sroE-sroF* (Figure S7C).

Next, we hypothesized that since the transducer modules present in both SPβ and phi3T promote lysogenization by activating MazF activity, the lysogenization capacity of the phi3T Δ*sroABCD*^{phi3T} mutant phage could be restored by expressing the SroABCD^{SPβ} proteins from SPβ in recipient cells. However, we expected phage reproduction to be affected, since the transducer module could not be deactivated by a non-cognate sensor module. However, we were unable to clone the *sroABCD*^{SPβ} construct in *E. coli*. Therefore, we constructed a SPβ Δ*sroABCD* phage, which generated similar phage titer as the WT but was also affected in its ability to form lysogens (Figures S6F and S6G). As per our hypothesis, overexpression of *sroABCD*^{phi3T} increased the number of lysogens in both phages (Figure 6E). SroABCD^{phi3T} overexpression in recipient also inhibited plaque formation for SPβ Δ*sroABCD* phage, whereas it had no effect on the phi3T Δ*sroABCD* plaque formation (Figure 6F). These results strongly suggest the transducer region of these two phages controls lysis-lysogeny by modulating MazF activity; however, it seems to require the presence of the cognate sensor module to correctly regulate the life cycles.

DISCUSSION

Phages of the SPbeta family have played a pivotal role in uncovering the arbitrium system, a mechanism governing cell fate in temperate phages. This system is not limited to phages but is also present in other mobile genomic elements of Firmicutes.[5,19] The uniqueness of this communication system in controlling lysis-lysogeny is associated with a new repression mechanism for this phage family. We previously demonstrated that SPbeta phages lack the prototypical CI phage repressor, and the repression function predominantly relied on the master repressor SroF (formerly known as YopR), encoded in a six-gene operon downstream of the arbitrium system, here named *sro*.[10] Now, we have introduced a necessary co-repressor, SroE, and conducted an in-depth characterization of the *sro* operon using phages phi3T and SPβ as models. Based on our analyses, we proposed that the *sro* operon comprises two main modules: (1) a conserved repressor module significantly different from the classical CI repressors and (2) a more variable transducer module, which is required to connect the arbitrium communication system with the repression module.

Our structural data confirmed that SroF exhibits the prototypical tyrosine recombinase folding, shared by all SroF-like proteins. However, SroF repressors seem to lack endonuclease capacity since they retain few catalytic residues essential for this activity. By generating SroF mutants in putative catalytic residues, we have confirmed the endonuclease activity is not required for the repressor function. We have also identified the DNA-binding boxes for SroF^{phi3T} (phi3TRE boxes) and confirmed the previously proposed ones for SPβ (SPβRE boxes).[13,18] These boxes are different in sequence but located in similar positions in the phage genomes. We also demonstrated that SroF in conjunction with SroE forms a repression complex that is sufficient to maintain the phage in a lysogenic state.

Importantly, although these interactions were observed *in vivo*, *in vitro* SroF showed a high but non-specific affinity for DNA. Therefore, we speculate that the repression complex formed by SroE-F, and possibly SroD, recognizes target DNA by topology in addition to sequence. Thus, SroF may bind DNA in a manner similar to the integron integrase, which recognizes its target by the position of two extrahelical bases induced by the DNA conformation, with most interactions being non-specific with the DNA backbone.[15] For this topological recognition, this integrase presents a unique insertion. Since SroF shares this structural feature, it might allow SroF to topologically recognize phage DNA, particularly considering the AT-rich content in the SroF binding motif. SroF may initiate the formation of the repression complex by recruiting SroE and possibly SroD. SroD, a predicted recombinase, may form functional heterocomplexes, similar to the segregation recombinase XerC, with which SroD shares homology.[22] Although we could not confirm the SroD and SroF interaction *in vitro*, we propose a similar heterocomplex in the repression complex, with SroF playing a structural role since SroD is not essential for maintaining lysogeny. Additionally, the folding of SroE corresponds to the chromosome-partitioning protein ParB. In bacterial chromosome segregation, XerC-D work together with ParB to perform their function; however, direct protein-protein interactions have never been demonstrated.[22] It is not unreasonable to speculate that phages of the SPbeta family may have evolved these proteins to form the SroD-SroF repressor complex, relinquishing their endonuclease activity in the process. Understanding the molecular basis of this new repression mechanism, including the recognition of operators and the coordination of regulatory functions by different components, represents a crucial future challenge.

As demonstrated here, this repression system differs significantly from classical CI repressors, both structurally and functionally. Intriguingly, it seems that this operon has recruited and evolved proteins with different primary functions, repurposed as repressors. However, both systems require SOS induction, a logical requirement since an activated SOS response likely results in the death of lysogenic cells. SPbeta phages use an additional regulatory system, the arbitrium system, which provides a fine-tuned control over lysis-lysogeny decisions.[5,10–12] We hypothesize here that the acquisition of this system necessitates an entirely different repression mechanism, one regulated not only by the SOS response but also by the transducer module. The size difference in SPbeta phages, nearly three times larger than classical CI phages, may justify the existence of multiple regions in the phage genome requiring repression by phage repressors. Furthermore, SPbeta phages have domesticated an anti-phage system, MazEF, to promote lysogeny.[8,9] The evolutionary forces driving this adaptation are unclear, but it is tempting to speculate that it enhances control over recipient cells in promoting lysogeny of incoming phages.

Recent reports indicate some phages encoding arbitrium systems (AimR/AimP/*aimX*) also encode classical CI repressors,[11,19] but not the arbitrium repression operon. This suggests some

phages carrying classical repressors have recruited the arbitrium sensing module or, vice versa, some arbitrium phages have replaced the transducer and repression modules with those encoding a classical repressor. In either case, the life cycle of these phages is controlled by a dual system (SOS and arbitrium). The exact mechanisms governing the interactions between these two systems remain to be determined.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Strain construction
- METHOD DETAILS
  - Plasmids and cloning
  - Measuring expression by fluorescent reporters
  - Bacteriophage induction assay
  - Bacteriophage Titering Assay
  - Lysogenisation assays
  - Recombinant protein expression and purification
  - Protein Crystallisation and data collection
  - Phase determination, model building and refinement
  - Structural prediction and similarities
  - SPbeta arbitrium and repressor operon extraction and annotation
  - Arbitrium and repressor operon phylogenies
  - Arbitrium repressor operon homology representation
  - EMSA assays
  - ChIP-seq data analysis
  - Biolayer Interferometry (BLI)
  - Size exclusion chromatography with multi-angle light scattering (SEC-MALS)
  - Whole genome sequencing and analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2023.11.003.

ACKNOWLEDGMENTS

We thank the IBV-CSIC Crystallogenesis Facility for protein crystallization screenings and the Tuberculosis Genomics Unit for their sequencing facilities. We thank Asier Zaragoza-Solas (University of Vienna) for his tool, https://github.com/asierzaragoza/flex2, used to create Figure 1E. The structural results and data collection experiments reported here derive from measurements made at the synchrotron DLS (Didcot, UK), ALBA (Cerdanyola del Valles, Spain), and ESRF (Grenoble, France). X-ray diffraction data collection was supported by block allocation group (BAG) DLS Proposal MX28394, ESRRF Proposal MX-2443 ALBA Proposal 2022075911. Some figures in this manuscript have been created with Biorender.com. This work was supported by grants PID2019-108541GB-I00 and PID2022-137201NB-I00 from the Spanish Government (Ministry of Science and Innovation) and PROMETEO/2020/012 by Valencian Government to A.M., and grants MR/V000772/1, MR/X020223/1, and MR/S00940X/1 from the Medical Research Council (UK), BB/V002376/1 and BB/V009583/1 from the Biotechnology and Biological Sciences Research Council (BBSRC, UK), and EP/X026671/1 from the Engineering and Physical Sciences Research Council (EPSRC, UK) to J.R.P. A.M. is part of the CSIC's Global Health Platform (PTI Salud Global). E.C.-Y. received an FPI predoctoral fellowship from the Spanish Ministry of Science and Innovation with identifier PRE2020-092531.

AUTHOR CONTRIBUTIONS

J.R.P. and A.M. conceived this study. A.B., E.C.-Y., F.G.d.S., C.C., J.M.-B., S.Z.-C., S.B.O., M.T.-P., N.Q.-P., and A.M. performed the experiments. A.B., E.C.-Y., F.G.d.S., C.C., J.M.-B., S.Z.-C., S.B.O., M.T.-P., A.E., N.Q.-P., A.M., and J.R.P. processed data. J.R.P. and A.M. wrote the manuscript with inputs from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 4, 2023
Revised: September 25, 2023
Accepted: November 3, 2023
Published: November 29, 2023

REFERENCES

1. Brady, A., Felipe-Ruiz, A., Gallego Del Sol, F.G., Marina, A., Quiles-Puchalt, N., and Penadés, J.R. (2021). Molecular basis of lysis–lysogeny decisions in Gram-positive phages. Annu. Rev. Microbiol. *75*, 563–581.
2. Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and Adhya, S. (2005). Switches in bacteriophage lambda development. Annu. Rev. Genet. *39*, 409–429.
3. Waldor, M.K., and Friedman, D.I. (2005). Phage regulatory circuits and virulence gene expression. Curr. Opin. Microbiol. *8*, 459–465.
4. Casjens, S.R., and Hendrix, R.W. (2015). Bacteriophage lambda: early pioneer and still relevant. Virology *479–480*, 310–330.
5. Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., Melamed, S., Leavitt, A., Savidor, A., Albeck, S., et al. (2017). Communication between viruses guides lysis-lysogeny decisions. Nature *541*, 488–493.
6. Silpe, J.E., and Bassler, B.L. (2019). A host-produced quorum-sensing autoinducer controls a phage lysis-lysogeny decision. Cell *176*, 268–280.e13.
7. Zamora-Caballero, S., Chmielowska, C., Quiles-Puchalt, N., Brady, A., del Sol, F.G., Mancheño-Bonillo, J., Felipe-Ruíz, A., Meijer, W.J.J., Penadés, J.R., and Marina, A. (2023). A complex network of antagonistic interactions between phage and host factors controls arbitrium lysis-lysogeny decision. Preprint at bioRxiv.
8. Guler, P., Bendori, S.O., Borenstein, T., Aframian, N., Kessel, A., and Eldar, A. (2023). Arbitrium communication controls phage lysogeny through non-lethal modulation of a host toxin-antitoxin defense system. Preprint at bioRxiv.
9. Cui, Y., Su, X., Wang, C., Xu, H., Hu, D., Wang, J., Pei, K., Sun, M., and Zou, T. (2022). Bacterial MazF/MazE toxin-antitoxin suppresses lytic propagation of arbitrium-containing phages. Cell Rep. *41*, 111752.
10. Brady, A., Quiles-Puchalt, N., Gallego Del Sol, F.G., Zamora-Caballero, S., Felipe-Ruíz, A., Val-Calvo, J., Meijer, W.J.J., Marina, A., and Penadés, J.R. (2021). The arbitrium system controls prophage induction. Curr. Biol. *31*, 5037–5045.e3.
11. Aframian, N., Bendori, S.O., Kabel, S., Guler, P., Stokar-Avihail, A., Manor, E., Msaeed, K., Lipsman, V., Grinberg, I., Mahagna, A., et al. (2021).

Dormant phages communicate via arbitrium to control exit from lysogeny. Nat. Microbiol. *7*, 145–153.

12. Bruce, J.B., Lion, S., Buckling, A., Westra, E.R., and Gandon, S. (2021). Regulation of prophage induction and lysogenization by phage communication systems. Curr. Biol. *31*, 5046–5051.e7.
13. Kohm, K., Jalomo-Khayrova, E., Krüger, A., Basu, S., Steinchen, W., Bange, G., Frunzke, J., Hertel, R., Commichau, F.M., and Czech, L. (2023). Structural and functional characterization of MrpR, the master repressor of the *Bacillus subtilis* prophage SPβ. Nucleic Acids Res. *51*, 9452–9474.
14. Messier, N., and Roy, P.H. (2001). Integron integrases possess a unique additional domain necessary for activity. J. Bacteriol. *183*, 6699–6706.
15. MacDonald, D., Demarre, G., Bouvier, M., Mazel, D., and Gopaul, D.N. (2006). Structural basis for broad DNA-specificity in integron recombination. Nature *440*, 1157–1162.
16. Duyne, G.D.V. (2001). A structural view of Cre-loxP site-specific recombination. Annu. Rev. Biophys. Biomol. Struct. *30*, 87–104.
17. Westers, H., Dorenbos, R., van Dijl, J.M., Kabel, J., Flanagan, T., Devine, K.M., Jude, F., Seror, S.J., Beekman, A.C., Darmon, E., et al. (2003). Genome engineering reveals large dispensable regions in *Bacillus subtilis*. Mol. Biol. Evol. *20*, 2076–2090.
18. Lazarevic, V., Düsterhöft, A., Soldo, B., Hilbert, H., Mauël, C., and Karamata, D. (1999). Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPbetac2. Microbiology (Reading) *145*, 1055–1067.
19. Stokar-Avihail, A., Tal, N., Erez, Z., Lopatina, A., and Sorek, R. (2019). Widespread utilization of peptide communication in phages infecting soil and pathogenic bacteria. Cell Host Microbe *25*, 746–755.e5.
20. Kohm, K., Floccari, V.A., Lutz, V.T., Nordmann, B., Mittelstädt, C., Poehlein, A., Dragoš, A., Commichau, F.M., and Hertel, R. (2022). The *Bacillus* phage SPβ and its relatives: a temperate phage model system reveals new strains, species, prophage integration loci, conserved proteins and lysogeny management components. Environ. Microbiol. *24*, 2098–2118.
21. Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M., et al. (2017). Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. Cell Syst. *4*, 291–305.e7.
22. Karaboja, X., Ren, Z., Brandão, H.B., Paul, P., Rudner, D.Z., and Wang, X. (2021). XerD unloads bacterial SMC complexes at the replication terminus. Mol. Cell *81*, 756–766.e8.
23. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat. Methods *19*, 679–682.
24. Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One *11*, e0163962.
25. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068–2069.
26. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.
27. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). Corrigendum to: IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 2461.
28. Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296.
29. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. *39*, W29–W37.
30. Saha, C.K., Sanches Pires, R.S., Brolin, H., Delannoy, M., and Atkinson, G.C. (2021). FlaGs and webFlaGs: discovering novel biology through the analysis of gene neighbourhood conservation. Bioinformatics *37*, 1312–1314.
31. Langdon, W.B. (2015). Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. BioData Min. *8*, 1.
32. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
33. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.
34. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
35. Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic Acids Res. *43*, W39–W49.
36. Deatherage, D.E., and Barrick, J.E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. In Engineering and Analyzing Multicellular Systems, *1151* (Humana Press), pp. 165–188.
37. Carver, T., Harris, S.R., Berriman, M., Parkhill, J., and McQuillan, J.A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics *28*, 464–469.
38. Carniol, K., Ben-Yehuda, S., King, N., and Losick, R. (2005). Genetic dissection of the sporulation protein SpoIIE and its role in asymmetric division in *Bacillus subtilis*. J. Bacteriol. *187*, 3511–3520.
39. Serrano, E., and Carrasco, B. (2019). Measurement of the length of the integrated donor DNA during *Bacillus subtilis* natural chromosomal transformation. Bio Protoc. *9*, e3338.
40. Skubák, P., Araç, D., Bowler, M.W., Correia, A.R., Hoelz, A., Larsen, S., Leonard, G.A., McCarthy, A.A., McSweeney, S., Mueller-Dieckmann, C., et al. (2018). A new MR-SAD algorithm for the automatic building of protein models from low-resolution X-ray data and a poor starting model. IUCrJ *5*, 166–171.
41. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.
42. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. *67*, 355–367.
43. Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. *66*, 22–25.
44. Holm, L. (2022). Dali server: structural unification of protein families. Nucleic Acids Res. *50*, W210–W215.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
trpC2	BGSC	168 (1A700)
CU1065 (phi3T) attSPβ *trpC2*	BGSC	1L1
trpC2; ΔSPβ; subclacin 168-sensitive; Δ*skin*; ΔPBSX; Δprophage 1; Δ*pks*::Cm; Δprophage 3; Cm^r	Westers et al.[17]; BGSC	Δ6 (1A1299)
trpC2 Δ*yopM*::*erm*	Koo et al.[21]; BGSC	BKE20840
trpC2 Δ*yopO*::*erm*	Koo et al.[21]; BGSC	BKE20820
trpC2 Δ*yopP*::*erm*	Koo et al.[21]; BGSC	BKE20810
trpC2 Δ*yopQ*::*erm*	Koo et al.[21]; BGSC	BKE20800
trpC2 Δ*yopR*::*erm*	Koo et al.[21]; BGSC	BKE20790
Δ6 lysogenic phi3T *phi3T_5*::*kan*	Brady et al.[10]	JP21870
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroF*^{phi3T}; amyE::P_{spank} *sroF*^{phi3T}	Zamora-Caballero et al.[7]	JP23811
Δ6 lysogenic SPβ *yokI*::*kan*	Brady et al.[10]	JP20866
Δ6 lysogenic SPβ *yokI*::*kan* Δ*sroF*^{phi3T}; *amyE*::P_{spank} *sroF*^{SPβ}	This study	JP24206
Δ6 *amyE*::P_{spank}	Brady et al.[10]	JP19679
Δ6 *amyE*::P_{spank} *sroF*^{phi3T}	Zamora-Caballero et al.[7]	JP23800
Δ6 *amyE*::P_{spank} *sroF*^{phi3T} K170A	This study	JP23416
Δ6 *amyE*::P_{spank} sroF^{SPβ}	Brady et al.[10]	JP21941
Δ6 *amyE*::P_{spank} sroF^{SPβ} K169A	This study	JP23248
Δ6 *amyE*::P_{spank} sroF^{SPβ} Y304A	This study	JP23249
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroE*^{phi3T}; *amyE*::P_{spank} *sroE*^{phi3T}	This study	JP23855
Δ6 lysogenic SPβ *yokI*::kan Δ*sroE*^{SPβ}; *amyE*::P_{spank} sroE^{SPβ}	This study	JP23853
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}	This study	JP23229
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}; amyE::P_{spank}	This study	JP24207
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}; amyE::P_{spank} *sroD*^{phi3T}	This study	JP23507
Δ6 lysogenic SPβ *yokI*::kan Δ*sroD*^{SPβ}	This study	JP23245
Δ6 lysogenic SPβ *yokI*::kan Δ*sroD*^{SPβ}; *amyE*::P_{spank}	This study	JP23506
Δ6 lysogenic SPβ *yokI*::kan Δ*sroD*^{SPβ}; *amyE*::P_{spank} *sroD*^{SPβ}	This study	JP24547
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroC*^{phi3T}	This study	JP23228
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroA*^{phi3T}	This study	JP23226
Δ6 lysogenic SPβ *yokI*::kan Δ*sroA*^{SPβ}	This study	JP23250
Δ6 lysogenic SPβ *yokI*::kan Δ*sroC*^{SPβ}	This study	JP21899
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroABCD*^{phi3T}	This study	JP24208
Δ6 *amyE*::P_{spank} *sroA*$^{mut\text{-}phi3T}$	This study	JP23417
Δ6 *amyE*::P_{spank} *sroB*^{phi3T}	Zamora-Caballero et al.[7]	JP23230
Δ6 *amyE*::P_{spank} *sroC*^{phi3T}	This study	JP23231
Δ6 *amyE*::P_{spank} *sroD*^{phi3T}	This study	JP23232
Δ6 *amyE*::P_{spank} *sroABCD*^{phi3T}	This study	JP23763

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Δ6 *amyE*::P_{spank} *sroA*$^{SP\beta}$	This study	JP22991
Δ6 *amyE*::P_{spank} *sroC*$^{SP\beta}$	This study	JP22992
Δ6 *amyE*::P_{spank} *sroD*$^{SP\beta}$	This study	JP22993
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroA*^{phi3T}::*erm*	This study	JP24209
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroB*^{phi3T}::*erm*	This study	JP24210
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*aimR*^{phi3T} Δ*sroA*^{phi3T}	This study	JP24202
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*aimR*^{phi3T} Δ*sroB*^{phi3T}	This study	JP24203
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*aimR*^{phi3T} TT^{mut}	Zamora-Caballero et al.[7]	JP23123
Δ6 lysogenic SPβ *yokI*::*kan* Δ*sroABCD*$^{SP\beta}$	This study	JP24486
Δ6 *amyE*::P_{spank} *sroABD*^{phi3T}	This study	JP24548
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, normal size colony from infection of Δ6	This study	JP24387
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, small colony from infection of Δ6	This study	JP24389
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, small colony from infection of Δ6	This study	JP24391
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, small colony from infection of Δ6	This study	JP24392
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, small colony from infection of Δ6	This study	JP24397
Δ6 lysogenic phi3T *phi3T_5*::*kan* Δ*sroD*^{phi3T}, small colony from infection of Δ6	This study	JP24399
PY79 Δ*xpf*	Aframian et al.[11]	AES6969
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *sacA*::P*yosX*-3xYFP cm	This study	AES9542
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *sacA*::P*yosX**-3xYFP cm	This study	AES9579
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *amyE*::P*yorZ*-3xYFP spec	This study	AES8417
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *amyE*::P*aimR*-3xYFP spec	This study	AES9556
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *amyE*::hs*sroE*$^{SP\beta}$ spec *sacA*::P*yosX*-3xYFP cm	This study	AES9545
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *amyE*::hs*sroD*$^{SP\beta}$ spec *sacA*::P*yosX*-3xYFP cm	This study	AES9562
PY79 Δ*xpf lacA*::P*xyl-sroF*$^{SP\beta}$ mls *amyE*::hs*sroE*$^{SP\beta}$ spec *sacA*::P*yosX**-3xYFP cm	This study	AES9651
Δ6 *amyE*::P_{spank} *SroF*^{phi3T}(FLAGx3)	This study	Δ6_SroF^{phi3T}(FLAGx3)
Δ6 lysogenic phi3T *phi3T_5*::*kan amyE*::P_{spank} *SroF*^{phi3T}(FLAGx3)	This study	Δ6_3T_SroF^{phi3T}(FLAGx3)
Chemicals, peptides, and recombinant proteins		
Lysogeny broth (LB), Miller	Sigma – Aldrich	Cat. # L3522-1KG
Lysogeny broth (LB), Lennox	Sigma – Aldrich	Cat. # L3022-1KG
Agar	Formedium	Cat. # AGA02
Spectinomycin dihydrochloride pentahydrate	Sigma – Aldrich	Cat. # S4014-5G
Erythromycin	Sigma – Aldrich	Cat. # E6376-25G
Kanamycin Sulfate	Sigma – Aldrich	Cat. # 60615-5G
Ampicillin Sodium Salt	Sigma – Aldrich	Cat. # A9518-25G
Isopropyl-β-D-thio-galactopyranoside (IPTG)	Melford	Cat. # 156000-5.0

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ammonium sulfate	Sigma – Aldrich	Cat. # A4915-500G
K_2HPO_4	Fisher scientific	Cat. # 10509263
KH_2PO_4	Fisher scientific	Cat. # 10573181
Tri-sodium citrate dihydrate	Fisher scientific	Cat. # 10396430
D-(+)-Glucose	Sigma – Aldrich	Cat. # G7021-1KG
Yeast extract	Fisher scientific	Cat. # 11407541
Casein hydrolysate	Sigma – Aldrich	Cat. # 22090-100G
Magnesium sulfate heptahydrate	VWR	Cat. # 25165.26
L-tryptophan	Sigma – Aldrich	Cat. # T8941-25G
L-methionine	Sigma – Aldrich	Cat. # M9625-25G
$CaCl_2$	VWR	Cat. # 190464K
Manganese II chloride dihydrate	Sigma – Aldrich	Cat. # 1059340100
Mitomycin C	Sigma – Aldrich	Cat. # M0503-5X2MG
NaCl	VWR	Cat. # 27810.295
Tris Base	Fisher scientific	Cat. # 10376743
Gen Elute Bacterial genomic DNA Kit	Sigma – Aldrich	Cat. # NA2120-1KT
UltraPure Agarose	Thermo Fisher	Cat. # 16500-500
Lysozyme	Sigma – Aldrich	Cat. # 10837059001
Proteinase K	Sigma – Aldrich	Cat. # P2308-500MG
Platinum Taq DNA polymerase HiFi	Thermo Fisher	Cat. # 11304029
D-(+)-Xylose	Fisher scientific	Cat. # BP708-250
GeneJet PCR Purification Kit	Thermo scientific	Cat. # K0702
Pierce™ Anti-DYKDDDK magnetic agarose	Thermo scientific	Cat. # A36797
SelenoMet medium base	Molecular Dimensions	Cat. # MD12-502 # MD12-501
HisTrap FF	Cytiva	Cat. # 10341287
2-β-mercaptoethanol	Acros organics	Cat. # 125472500
Deposited data		
RCSB Protein Data Bank		PDB codes 8BJ6, 8BJV and 8BPZ
European Nucleotide Archive (ENA)		PRJEB65757
Oligonucleotides		
See Table S7 for list of oligonucleotides used in this study	N/A	N/A
Software and algorithms		
Prism	Graph Pad	V9.5.0
AlphaFold2	Mirdita et al.[23]	ColabFold v1.5.2
BioRender	BioRender.com	N/A
SeqKit	Shen et al.[24]	V0.10.1
Prokka	Seemann[25]	1.14.6
MAFFT	Katoh and Standley[26]	v7.487
IQ-TREE	Minh et al.[27]	Version 2
iTOL	Letunic and Bork[28]	N/A
HMMER	Finn et al.[29]	N/A
webFlaGs	Saha et al.[30]	N/A
Geneious Prime software	(https://www.geneious.com/)	N/A
Bowtie2	Langdon[31]	Version 2.4.5
SAMtools	Li et al.[32]	N/A
MACS2	Zhang et al.[33]	v2.2.8

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Integrative Genomics Viewer (IGV)	Robinson et al.[34] (https://software.broadinstitute.org/software/igv/)	
MEME suite	Bailey et al.[35]	N/A
breseq	Deatherage and Barrick[36]	Version 0.38.1
ARTEMIS	Carver et al.[37]	N/A
R/RStudio	R Development Core Team	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, José R Penadés (j.penades@imperial.ac.uk).

Materials availability

All bacterial strains and plasmids generated during this work are freely available from José R. Penadés (j.penades@imperial.ac.uk). The study did not generate new reagents.

Data and code availability

Coordinates for atomic structures have been deposited at the RCSB Protein Data Bank (PDB codes 8BJ6, 8BJV and 8BPZ). The ChIP-seq data has been deposited in the European Nucleotide Archive (ENA) under project code PRJEB65757. This work did not use or generate new codes. All data reported in this paper will be shared by the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All bacterial strains used in this study belong to *B. subtilis* or *E. coli* species. *B. subtilis* strains were routinely grown at 37°C on LB (Miller) agar plates or in LB (Miller) broth liquid medium shaking at 200 rpm. *E. coli* DH5α was grown at 37°C on LB (Lennox) agar plates or in LB (Lennox) broth shaking at 180 rpm. When required, antibiotics were utilized at the following concentrations: erythromycin (1 μg ml^{-1}), kanamycin (10 μg ml^{-1}), ampicillin (100 μg ml^{-1}), tetracycline (10 μg ml^{-1}) or spectinomycin (100 μg ml^{-1}).[21]

Strain construction

Bacterial strains used in this study are listed in the key resources table. We generated the deletion mutants in phage phi3T by overlapping PCRs containing the erythromycin marker (including the *lox* sites) and 1 kb of flanking region for the desired gene. To generate the deletion mutants in phage SPβ we made use of the BKE collection,[21] amplifying the desired gene plus 1 kb of flanking region. Once purified, 1 μg of PCR products containing the erythromycin marker and 1 kb flanking regions of the gene of interest was transformed into the Δ6 phi3T strain or Δ6 SPβ strain. Once the insertion of the erythromycin cassette was confirmed by PCR and sequencing, the antibiotic resistance cassette was removed as previously described.[21] Briefly, plasmid pDR244 was transformed into strains harbouring the *loxP*-flanked antibiotic resistance cassette with selection for spectinomycin resistance at 30°C to allow for Cre-loxP-mediated loop-out of the cassette. Transformant colonies were then streaked onto LB plates and incubated overnight at 42°C for removal of the temperature-sensitive plasmid. Resulting strains were screened for plasmid curing (loss of spectinomycin resistance) and the antibiotic resistance cassette (loss of erythromycin resistance). Strains were streaked to single colonies and confirmation of the clean mutant was performed using PCR. To generate the *sroF*$^{SP\beta}$, *sroE*^{phi3T} and *sroE*$^{SP\beta}$ deletion mutants, first we introduced a copy of each gene under the control of the IPTG inducible promoter P_{spank} using a similar strategy to the one used to obtain the *sro*F^{phi3T} conditional mutant.[7]

METHOD DETAILS

Plasmids and cloning

Plasmids generated in this study are listed in Table S6. The different genes for overexpression were cloned into the *amyE* integration vector pDR110 under the control of the IPTG inducible promoter P_{spank}.[38] Cloning was performed after PCR amplification of the appropriate template DNA using primers listed in Table S7, followed by digestion with *Sal*I-HF and *Sph*I-HF, and ligation with pDR110 digested with *Sal*I-HF and *Sph*I-HF. *B. subtilis* competent cell preparation and transformation was performed as described.[39] Briefly, *B. subtilis* cells were grown in GM1 minimum medium to early stationary phase to induce natural competence and 1 μg of plasmid was added and incubated at 37°C for 1 h with shaking at 210 rpm. The culture was centrifuged at 6000 g for 1 min, 800 μl of the supernatant removed, and the pellet re-suspended in 400 μl and plated out onto the relevant antibiotic plates. Plates were incubated at 37°C for 24 h.

Plasmid constructs for protein production in *E. coli* were generated by cloning PCR products obtained with the primers listed in Table S7. SPβ gene *sroF* was amplified using primers SroF$^{SP\beta}$_fw/SroF$^{SP\beta}$_rv and genomic DNA from *B. subtilis* strain 168 as template. The *sroF*^{phi3T} and *sroE*^{phi3T} genes were amplified from *B. subtilis* strain 1L1 using primers SroF^{phi3}_fw/ SroF^{phi3}_rv and SroE^{phi3}_fw/ SroE^{phi3}_rv, respectively. The PCR products were purified and cloned into the pLicSGC1 plasmid using Ligation-Independent Cloning (LIC) system for *sroF*$^{SP\beta}$ and NEBuilder HiFi DNA Assembly master mix (New England Biolabs) for *sroF*^{phi3T} and *sroE*^{phi3T}. These plasmids express the full-length genes with an N-terminal 6xHistag followed by a TEV protease cleaving site.

For the strains carrying the reporters, pAEC2731 and pAEC2262 were constructed by amplifying the *yosX* and *yorZ* promoters using the SOB360/367 and SOB363/364 primer pairs, respectively. The PCR product was digested with *Eco*RI-HF and *Nhe*I-HF, and cloned into pAEC1003 digested with *Eco*RI-HF and *Nhe*I-HF. pAEC1779 and pAEC2669 were constructed by amplifying the *yorZ* and *aimR* promoters using the SOB363/364 and SOB206/208 primer pairs, respectively. The PCR product was digested with *Eco*RI-HF and *Nhe*I-HF, and cloned into pAEC277 digested with *Eco*RI-HF and *Nhe*I-HF. pAEC2748 was constructed by amplifying the whole pAEC2731 using SOB880/881 primer pair that contain a mutation in the putative SroF$^{SP\beta}$ binding site, changing the sequence from ACAAATAATATGTATA to CTGCCACCATGACCTA. The original plasmid in the PCR product was degraded using *Dpn*I. The PCR product was closed after adding phosphate using PNK enzyme, followed by ligation. pAEC2245 and pAEC2247 were constructed by amplifying pAEC1046 using the SOB466/467 primer pair and amplifying the *sroD*$^{SP\beta}$ and *sroE*$^{SP\beta}$ genes from an SPβ lysogen *B. subtilis* genomic DNA using the SOB472/473 and SOB474/475 primer pairs, respectively. The purified DNA fragments were then ligated using the Gibson assembly protocol. pAEC2263 was constructed by amplifying the *sroF*$^{SP\beta}$ gene using the SOB505/505 and SOB363/364 primer pairs, respectively. The PCR product was digested with *Bam*HI-HF and *Sac*II, and cloned into pAEC548 digested with *Bam*HI-HF and *Sac*II.

Plasmid construct for FLAGx3-tagged SroF^{phi3T} expression was generated by cloning the SroF^{phi3T} ORF into an *amyE* integration vector derived from pAND101, under the control of IPTG inducible promoter P_{spank}. The primers and steps followed are listed in Table S7. PCR products were purified and cloned using NEBuilder HiFi DNA Assembly master mix (New England Biolabs). Plasmid pAND_97_F3 was transformed, as described before, in *B. subtilis* strain 168 Δ6 to generate the strain Δ6_SroF^{phi3T}(FLAGx3). This new strain was lysogenized with phage phi3T to generate the strain Δ6_3T_SroF^{phi3T}(FLAGx3).

Measuring expression by fluorescent reporters

To determine expression levels of promoter reporters strains were grown overnight in LB at 37° with shaking at 220 RPM, then diluted 1:100 into fresh LB media. IPTG and xylose were added when indicated. Upon reaching OD_{600} = 0.3, YFP fluorescence was measured using a Beckman Coulter cytoflex flow cytometer.

Bacteriophage induction assay

For induction, an overnight culture was diluted 1:100 in LB media supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ and then grown at 37°C with 210 rpm shaking until reaching absorbance 0.2 at 600 nm. This step was repeated twice to ensure the cells were in exponential growth. After the second growth Mitomycin C (MC) at 0.5 μg ml^{-1} was added to the culture. Where experiments were performed to test the complementation of the mutants, 1 mM of IPTG was added at the same time as MC induction. Where experiments were performed to test the effect of the peptide on phage titer for phi3T and the Δ*sroABCD* mutant, 5 μM of AimP^{phi3T} (SAIRGA; ThermoFisher) was added before the second growth to ensure incorporation of the peptide by the cells. The induced cultures were incubated at 30°C with 80 rpm shaking for 4 h and then left overnight at room temperature. Following lysis, samples were filtered using 0.2 μm filters and lysates were stored at 4°C until use.

Bacteriophage Titering Assay

The number of phage particles contained in the phage lysate of interest were quantified by a titering assay. An overnight culture of the relevant recipient strain (normally *B. subtilis* Δ6 or with the corresponding integration vector) was diluted 1/100 in LB supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ and then grown at 37°C with 210 rpm shacking until reaching absorbance 0.2 at 600 nm. If needed 0.1 mM IPTG was added. Then, 100 μl of recipient bacteria was infected with 100 μl of serial dilutions of phage lysate in phage buffer (PhB; 1 mM NaCl, 0.05 M Tris pH 7.8, 0.1 mM $MnCl_2$, 5 mM $MgCl_2$) at room temperature for 10 min and 3 ml of phage top agar (LB media supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ and 0.7 % agar) at 55°C was added to the culture-phage mix and immediately poured over phage base agar plates (LB media supplemented with 0.1 mM $MnCl_2$ and 5 mM $MgCl_2$ and 1.5% agar). Plaques were counted after overnight growth at 37°C temperature and photographed.

Lysogenisation assays

To enable selection of phages that have successfully integrated into the bacterial genome, an antibiotic resistance gene was introduced into the genome of phi3T and SPβ. For this, a gene proven to be non-essential for the phage was replaced with a kanamycin cassette (*phi3T_5*::*kanR* and *yokI*::*kanR* for phi3T and SPβ respectively) by amplifying the marker without including the *lox* sites from one of the BKK Genome-Scale deletion library mutants (BGSC). To test and quantify the ability of wt and mutant phages to form lysogens in this study, we grew recipient strains to early exponential phase at 37°C for approximately 2 h to absorbance 0.2 at 600 nm. Phage lysates of interest that contain the kanamycin marker were serially diluted in PhB and 100 μl was added to 1 ml of the recipient bacteria in 12 ml tubes. This was completed for several different serial dilutions of the phage lysate. The bacteria-phage mixture was incubated at 37°C for 30 min to allow the phage to infect bacteria. The mixture was then transferred to 1.5 ml Eppendorf tubes and

centrifuged at 6,600 rpm for 1 min. The supernatant was removed, and the bacterial pellet was resuspended in 400 μl of fresh LB broth before plating onto selective antibiotic LB agar plates. Plates were incubated overnight at 37°C. The number of colony forming units (CFU) was calculated. The same lysate was used in titering assays to quantify the number of phage particles, and results are presented as CFUs of an average phage titer (1 x 10^9 PFU).

Recombinant protein expression and purification

For protein production a single colony of *E. coli* strain BL2_Codon plus (DE3) RIL (Agilent) carrying the expression plasmid for the selected protein was inoculated in a 200 ml flask containing 50 ml of LB medium supplemented with 100 mg ml^{-1} of ampicillin and 33 mg ml^{-1} of chloramphenicol and grown overnight at 37°C. One litre of LB medium supplemented with 100 mg ml^{-1} ampicillin and 33 mg ml^{-1} chloramphenicol were inoculated with 15 ml of overnight culture and were incubated at 37°C with shaking at 190 rpm. Cell growth was monitored until A_{600} reached 0.4, then temperature was lowered to 20°C and IPTG to a final concentration of 0.2 mM was added in order to induce protein expression. After 16 h, cells were pelleted by centrifugation at 10000 g for 45 minutes and the pellet was stored at -80°C. To produce selenomethionine (SeMet) derivative $SroF^{SP\beta}$ the expressing *E. coli* cells were grown in SelenoMet medium base supplemented with SelenoMet Nutrient Mix (Molecular Dimensions). When the culture reached absorbance 0.4 at 600 nm, the temperature was lowered to 20°C and L-selenomethionine at 30 mg ml^{-1}; isoleucine, leucine and valine at 50 mg l^{-1}; lysine, phenylalanine and threonine at 100 mg l^{-1} final concentration were added. After 30 min, IPTG (0.2 mM final concentration) was added to induce protein expression and the culture was maintained with shaking for 16 h at 20°C.

For protein purification, the pellets were suspended in lysis buffer (25 mM Tris-HCl pH 8, 250 mM NaCl) and disrupted by sonication on ice. The sonicated lysate was centrifuged at 10000 g for 1h. The supernatant was loaded onto a 5 ml HisTrap FF (GE Healthcare), washed with lysis buffer and eluted with lysis buffer supplemented with 500 mM imidazole. The eluted protein was digested with TEV protease (50:1 molar ratio protein:TEV) and dialysed against lysis buffer supplemented with 1 mM 2-β-mercaptoethanol. The dialyzed sample was loaded in a Hi-Load Superdex 200 16/60 (GE Healthcare) gel filtration column previously equilibrated in lysis buffer. Fractions containing the purest protein were pooled, concentrated in a centrifugal filter (Amicon ultra 30KDa) to 30 mg ml^{-1} and stored at -80°C. Typical yields were 20-30 mg of recombinant protein per litre of culture medium.

Protein Crystallisation and data collection

Commercial screens JCSG, JBS I, JBS II (JENA Biosciences) and MIDAS (Molecular Dimensions) were set up for $SroF^{SP\beta}$ at the Cristalogenesis service of the IBV-CSIC using 96-well plates (Swissci MRC2) and adding equal volumes of protein at 10mg ml^{-1} in lysis buffer and precipitant. A vapour-diffusion approach was used to grow the crystals in hanging drops at 21°C. $SroF^{SP\beta}$ crystallised into two distinct spatial groups. The $SroF^{SP\beta}$ crystals on the C2 spatial group were obtained using 20% (w/v) PEG 8000, 100 mM sodium phosphate dibasic/Citric acid pH 4.2 and 200 mM sodium chloride as precipitant. The $SroF^{SP\beta}$ crystals in P321 spatial group were obtained using 30 % (w/v) polyethylene glycol 4000, 100 mM tri-sodium citrate pH 5.6 and 100 mM ammonium sulfate as precipitant. Crystals grew in 2-3 days and were directly flash frozen in liquid nitrogen. Crystals were tested at ALBA, DLS and ESRF synchrotrons and diffraction data for solve the structures were collected at XALOC beamline of ALBA Synchrotron using single crystals at 100°K. Data sets were integrated with XDS and reduced using Scala (CCP4). Statistics for data collection are shown in Table S1.

Phase determination, model building and refinement

The crystal structure of $SroF^{SP\beta}$ from SPβ phage was solved by SAD phasing with CRANK[40] using a Selenothionine derivative protein that crystallised in space group C2. Initial CRANK model was manually rebuilt using COOT[41] and computationally refinement with Refmac[42] and after several rounds of rebuilding and refinement, the final refined model was generated (PDB 8B6J) and used as search model to solved the native $SroF^{SP\beta}$ crystal structure in space group C2 (8BJV) and P321 (8BPZ) by molecular replacement using MOLREP.[43] Statistics for refinement are summarised in Table S2.

Structural prediction and similarities

AlphaFold2 was used to predict structural fold of arbitrum repressor operon ORFs from phage phi3T. AlphaFold2 was run online in Google ColabFold v1.5.2 using default (auto) settings and entering the 1-letter sequence of the modelled protein as input.[23] The coordinates of the models were used as input to search with DALI server[44] structural similarities in PDB.

SPbeta arbitrium and repressor operon extraction and annotation

The SPbeta group prophages identified by Kohm and coworkers,[20] and classified genus SPbetavirus were used to generate a database. 58 genomic sequences of *Bacillus* strains carrying SPbeta-like prophages were downloaded from NCBI repositories (Table S5) and the phage sequences were extracted using SeqKit V0.10.1[24] by specifying the positions of each phage as defined by Kohm and coworkers.[20] Then, we utilized the Prokka pipeline 1.14.6[25] to annotate all the proteins of these phages, which resulted in consistent protein annotation across all genomes. Hidden Markov model (HMM) profiles were constructed for the AimR and SroF proteins to enable the identification of the start and end of the arbitrium and the arbitrium repressor operon within the SPbeta-like phage sequences annotated with Prokka. Once the specific CDS for the operon were located, SeqKit V0.10.1[24] was used to extract the 58 complete genomic sequences. Some sequences were reversed using the reverse complement function in SeqKit V0.10.1[24] to ensure that all the operon sequences were in identical 5'-3' orientation. Subsequently, the AimR-SroF region was reannotated using the Prokka pipeline 1.14.6.[25]

Arbitrium and repressor operon phylogenies

The arbitrium and repressor operon genomic sequences for each prophage were aligned using MAFFT v7.487 with its default parameters.[26] SroA, SroB and SroC proteins were also aligned separately using MAFFT v7.487 with default parameters.[26] Subsequently, phylogenetic trees were then inferred for the complete arbitrium systems and repressor operons and for the SroA-SroC proteins separately using the IQ-tree tool, with the parameters -bb 1000, -nt AUTO, and -alrt 1000.[27] Phylogenies were then visualized using the interactive Tree of Life (iTOL) platform.[28] HMM profiles were constructed for the AimR and SroF proteins to facilitate the identification of the start and end of the arbitrium and the arbitrium repressor operon within Prokka-annotated SPbeta-like phage sequences. Additionally, HMM profiles for the SroA, SroB, and SroC proteins from phi3T and SPβ were constructed to search for matches among the proteins of the seven representative phages. These HMM profiles were created using the hmmbuild option in the HMMER tool[29] and were matched against sequences using the hmmsearch option from the same tool. Further, a conservation analysis of the gene neighbourhoods was performed with webFlaGs tool using SroE as a query to define neighbouring genes.[30] The NCBI ID of SroD from SPbeta-like genes was used as input and the flanking gene-encoded proteins were clustered using HMM-based JackHMMER method with default parameters. The to-scale colour-coded representation and flanking genes description was obtained as web server output.

Arbitrium repressor operon homology representation

Geneious Prime software (https://www.geneious.com/) was used to generate a visual representation of homology identity as well for potential ORFs CDSs identification. The 58 genomic sequences of SPbeta-like phages were aligned using the global alignment algorithm of the software using the parameters free end gaps, 51% similarity threshold, 12 gap open penalty, gap extension penalty, and refinement iterations. A separate alignment of the 7 representative prophages was conducted with the same parameters.

EMSA assays

DNA binding capacity of $SroF^{SP\beta}$ and $SroE^{SP\beta}$ was analyzed by native agarose gel electrophoresis. Probes described in the manuscript were obtained by PCR using primers summarized in Table S7. Fragments were purified using the commercial kit GeneJet PCR Purification Kit (ThermoFisher®). Purified PCR product (0'05 pmol μl^{-1}) and variable quantities of protein depending on the assay were mixed in PBS buffer. The samples were incubated for 5 min at room temperature. Electrophoresis was then performed in 2% agarose gels in Tris-Borate-EDTA (TBE) buffer for about 50 min at 85 V at room temperature.

ChIP-seq data analysis

Overnight cultures of *B. subtilis* 168 Δ6 $SroF^{phi3T}$(FLAGx3) lysogenized or not with phi3T phage were diluted in 50 ml of LB (Miller) to OD_{600} = 0.05 and incubated at 37°C with 230 rpm shaking. When the OD_{600} reached 0.4-0.6, 1 mM IPTG was added to cultures to induce $SroF^{phi3T}$ expression and left three hours at 37°C with 230 rpm shaking. Cells were harvested and resuspended in PBS. Cross-linking was performed with 2% formaldehyde for 20 minutes at room temperature and then quenched with glycine and washed in TBS. Cultures were kept overnight at -20°C. Cells were digested using BugBuster® (Milipore™) combined with sonication on ice. Chromosomal DNA was sheared to an average size of 300 bp using Bioruptor® PICO (Diagenode™) sonication device. A portion of the sample was removed and store prior to immunoprecipitation as a DNA input control. The immunoprecipitation was performed with Pierce™ Anti-DYKDDDK magnetic agarose (Thermo Scientific™). After washing and elution, the samples were incubated at 65°C overnight with 800 rpm shaking to reverse the cross-linking and then, treated with proteinase K. DNA was recovered using GeneJet PCR purification kit (Thermo Scientific). For next-generation sequencing libraries preparation DNA concentration of the samples was measured using the Qubit HS DNA kit (Invitrogen). 10-40 ng of total DNA was employed to prepare libraries using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New England Biolabs), according to the manufacturer's protocol with 8 cycles for indexing PCR. Library quality control was evaluated with Qubit 3 fluorometer (Invitrogen) and TapeStation 4200 (Agilent Technologies) for concentration and size, respectively. Samples were sequenced in a MiSeq platform (2×150cycles paired-end run; Illumina).

Sequencing reads were mapped to the *B. subtilis* 168 Δ6_sroFphi3T reference genome using Bowtie2[31] with the following parameters: -N 1 -q -p 10 -S. Subsequently, BAM files were generated and further processed and sorted using SAMtools.[32] Peak calling was performed using MACS2 (Model-based Analysis of ChIP-Seq data) v2.2.8[33] with default parameters, except for the inclusion of –nomodel and –extsize 200 options. The identified binding peaks and their density were visualized using the Integrative Genomics Viewer (IGV) (https://software.broadinstitute.org/software/igv/).[34] For motif discovery, sequences of the ChIP-seq peaks were submitted to MEME suite[35] to identify enriched motifs with lengths between 4 and 50 bp. ChIP-seq data was analysed using custom R scripts to illustrate the genomic distribution of peaks.

Biolayer Interferometry (BLI)

The binding affinity constants between $SroF^{SP\beta}$ and DNA were measured by BLI using the BLITz system (FortéBio). Biotinylated DNA probes (IDT) (Figure S3B) were immobilised in Streptavidin biosensors (FortéBio) at 70 mg ml^{-1}. PBS buffer supplemented with 0.01% tween and 0.1% BSA was used for biosensor hydration, baselines and dissociation analysis. Proteins were diluted in PBS. At least four different dilutions of $SroF^{SP\beta}$ were used in a range between 0.035 and 5 μM. A control using a chip without loading DNA was used to subtract unspecific binding if any. Data analysis and kinetic calculations were performed employing a 1:1 model to fit the data using BLItz Pro 1.2 software.

Size exclusion chromatography with multi-angle light scattering (SEC-MALS)

A Wyatt DAWN HELEOS-II MALS instrument and a Wyatt Optilab rEX differential refractometer (Wyatt) coupled to a Shimadzu HPLC equipped with a manual injector was used to perform SEC-MALS experiments. Proteins were injected (20 μl of protein at 2 mg ml^{-1}) in a KW-803 (Shodex) column previously equilibrated with 25 mM Tris pH 8, 250 mM NaCl at a flow rate of 0.3ml per minute. Acquisition and analysis of the data was performed using the Astra 7.1.2 software from the manufacturer.

Whole genome sequencing and analyses

Colonies representing different morphologies (normal size/small colonies) were chosen from 3 independent lysogeny experiments using the phi3T Δ*sroD* mutant lysate. For evolved phi3T Δ*sroF* overcoming SroF overexpression in recipient cells, individual plaques were isolated. Total genomic DNA (GenElute, Bacterial Genomic DNA Kit, Sigma) or DNA extracted from the evolved phages was sent for whole-genome sequencing to SeqCenter facilities (Pittsburgh, USA). Samples were sequenced in a MiSeq platform (2× 150cycles paired-end run; Illumina). Mutation analyses of sequenced genomes was done by reference-based alignment of short reads using the breseq (version 0.38.1; Bowtie2 version 2.4.5) pipeline[36] followed by manual inspection of read alignment using ARTEMIS.[37] The following Genbank reference sequences were used: Δ 6 strain: CP015975; phi3T MT366945. The phi3T Δ*sroD* lysogens reported deletions were detected by comparing the sequencing results of the lysogens to strain JP24387 (Δ6 phi3T Δ*sroD*) from which the lysates were induced to be used in these experiments. For evolved phi3T Δ*sroF* phages, the reported changes are in comparison to the sequence of phi3T Δ*sroF*.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Statistical analysis was performed as indicated in the figure legends. All analyses were done using GraphPad Prism 9 software. The p-values represented in each figure are shown in the figure legends.