

The SinR·SlrR Heteromer Attenuates Transcription of a Long Operon of Flagellar Genes in *Bacillus subtilis*

Ayushi Mishra, Abigail E. Jackson, Xindan Wang, and Daniel B. Kearns*

Department of Biology, Indiana University, Bloomington, IN, USA

Correspondence to Daniel B. Kearns: dbkearns@iu.edu (D.B. Kearns)

<https://doi.org/10.1016/j.jmb.2025.169123>

Edited by Eric Cascales

Abstract

During growth, *Bacillus subtilis* differentiates into subpopulations of motile individuals and non-motile chains, associated with dispersal and biofilm formation, respectively. The two cell types are dictated by the activity of the alternative sigma factor SigD encoded as the penultimate gene of the 27-kb long *fla/che* flagellar operon. The frequency of SigD-ON motile cells is increased by the heteromeric transcription factor SwrA·DegU that activates the *fla/che* promoter. Conversely, the frequency of motile cells is decreased by the heteromeric transcription factor SinR·SlrR, but the mechanism and location of inhibition is poorly understood. Here, using ChIP-Seq analysis, we determine the binding sites of the SinR·SlrR heteromer on the genome. We identified two sites within the *fla/che* operon that were necessary and sufficient to attenuate transcript abundance by causing premature termination upstream of the gene that encodes SigD. Thus, cell motility and the transition to biofilm formation depend on the expression of a long operon governed by two opposing heteromeric transcription factors that operate at two different stages of the transcription cycle. More broadly, our study serves as a model for transcription factors that control transcriptional elongation and the regulation of long operons in bacteria.

© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Introduction

Bacillus subtilis grows as a mixed population of two different cell types; some cells are motile and grow as individuals, while others are non-motile and grow in long chains.^{1–3} The two subpopulations are differentiated by the level and activity of the alternative sigma factor SigD.^{3–6} Motile cells have high levels of SigD protein and express a regulon containing late flagellar structural proteins and peptidoglycan (PG) lyases that promote cell separation after division (SigD^{ON} cells).^{3,6–8} Conversely, chaining cells have low levels of SigD protein and fail to express both the flagellar proteins and PG lyases, such that cells fail to separate from one another after division (SigD^{OFF} cells).^{4–6} The two cell types are the product of a developmental, epigenetically-inherited switch,^{6,9,10} and likely evolved as a

“bet-hedging” strategy to compensate for the fact that the assembly of functional flagella takes two to three generations at high growth rates.¹¹

A number of proteins regulate motile cell development. SigD is important; as without it, the SigD-regulon is inactivated and all cells in the population grow as non-motile chains.^{3,5,7,12} The gene encoding SigD, *sigD*, is found near the 3' end of the 32-gene, 27-kb long *fla/che* operon that also encodes proteins required early in flagellar biosynthesis.^{6,12–14} SigD activity is antagonized by its cognate anti-sigma factor FlgM, which is secreted out of the cells once the hook-basal body (HBB) of the flagellum is assembled.¹⁹ Mutation of any of the flagellar structural genes in the *fla/che* operon disrupts assembly of HBB, resulting in failure of FlgM secretion, accumulation of FlgM in the cytoplasm and inhibition of SigD activity.^{15–19} The

promoter of the *fla/che* operon is activated by a heteromeric complex comprised of the response regulator DegU and small protein SwrA, to increase SigD levels and increase the frequency of motile SigD^{ON} cells.^{20–26} Finally, another heteromeric complex of two paralogous DNA binding proteins, SinR and SlrR, has the opposite effect and decreases the frequency of the motile SigD^{ON} state.^{8,37}

SinR is tetrameric DNA binding repressor protein that binds to and inhibits the expression of operons that promote biofilm formation.^{27–30} One target of SinR is the *eps* promoter that drives expression of 15 gene products including EpsH, an enzyme involved in the synthesis of the biofilm extracellular polysaccharide (EPS), and EpsE, a bifunctional EPS synthase and inhibitor of flagellar rotation.^{31–33} At the same location upstream of *eps*, SinR also

represses the oppositely-oriented *slrR* gene that encodes SlrR, a SinR-paralog^{34,35} (Figure 1). SinR is stoichiometrically antagonized by interaction with either of two proteins, SinI or SlrA,^{35,36,60} and when SinR binding to DNA is antagonized, SlrR is derepressed and forms a heteromer with SinR.^{37,38} SlrR is not known to bind DNA on its own but is thought to reprogram the binding of SinR to other sites in the chromosome, including those responsible for the inhibition of SigD-dependent gene expression and population heterogeneity.^{8,37}

Antagonism of SinR and the formation of the SinR-SlrR heteromer is central to the transition from motility to biofilm formation. How the heteromer inhibits SigD-dependent gene expression is poorly understood and there are two models for how inhibition might occur. One model for SinR-SlrR inhibition of motility gene expression

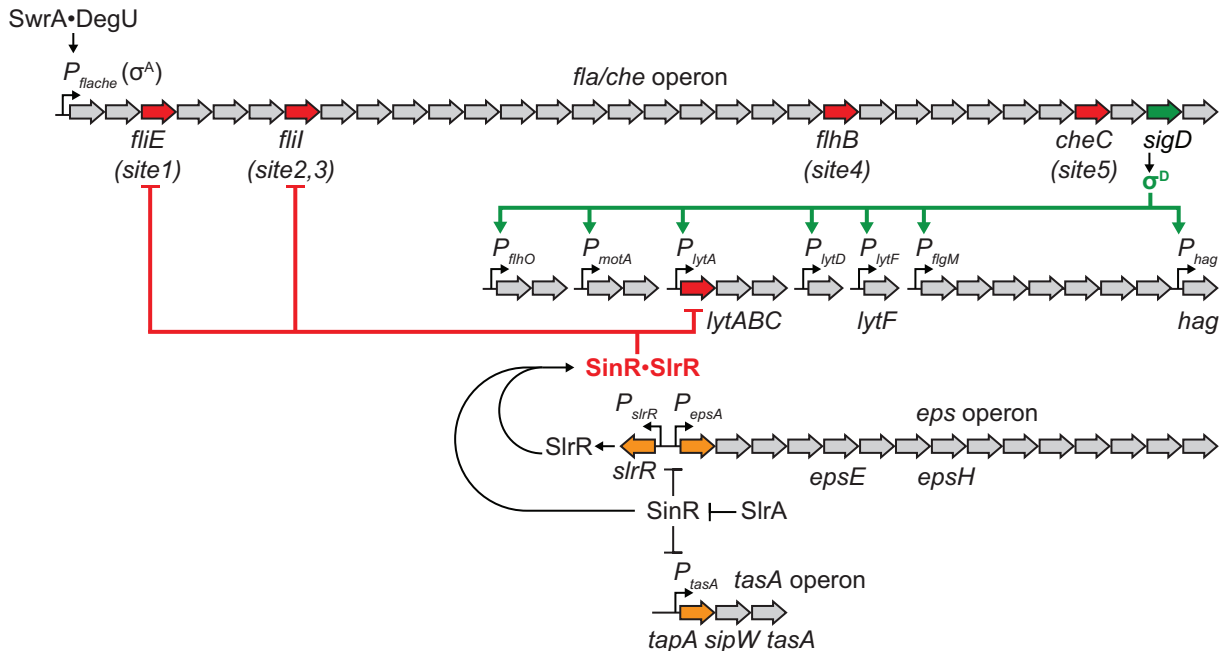


Figure 1. A model for inhibition of the SigD regulon by the SinR-SlrR heteromer. A model for SinR-SlrR-mediated inhibition of SigD-dependent gene expression. A cartoon diagram of the flagellar and biofilm regulons is presented in the figure. SinR represses the promoter of the *tasA* operon that encodes for protein components of biofilm formation and the promoter of the *eps* operon that encodes genes for the synthesis of the biofilm extracellular polysaccharide (EPS). SinR simultaneously inhibits a divergent promoter that expresses the paralog, SlrR. Genes directly downstream to the promoters inhibited by SinR are highlighted in orange. Elevated levels of SlrA (or SinI) antagonize SinR, promoting formation of the SinR-SlrR heteromer. The heteromer binds to multiple sites within the *fla/che* operon, reducing flagellar gene expression and the levels of SigD, causing a failure in SigD-dependent gene expression. Inhibition by the heteromer occurs in a subpopulation of cells and can be overridden by SwrA·DegU-dependent activation of the *P_{fla/che}* promoter. Bent arrows indicate promoters. Block arrows indicate genes, and those containing binding sites for SinR-SlrR are highlighted in red. Other arrows indicate activation and T-bars indicate inhibition. The names of relevant genes mentioned in the text are indicated below their corresponding gene location. We have chosen to use the general terms homomer and heteromer to describe SinR and SinR-SlrR respectively (despite the fact that SinR has been shown to bind DNA as a tetramer,^{29,30} and SinR and SlrR have been shown to interact *in vitro* as a dimer³⁸) because the ChIP seq analysis used here cannot distinguish protein stoichiometry when bound to DNA *in vivo*.

is direct, in which the heteromer inhibits gene expression directly at individual promoters recognized by SigD-containing RNA polymerase. In support of this model, the heteromer was shown to bind to promoter regions of multiple SigD-dependent genes *in vitro* including those that code for the flagellin protein Hag, and the peptidoglycan lyases LytF and LytABC.³⁷ Thus, the heteromer directly blocks expression of motility and cell separation genes at their individual promoters. The “direct inhibition model”, however, cannot explain why SigD fails to accumulate in the presence of the heteromer,⁶ or explain the observation that seemingly all of the SigD regulon is repressed,⁸ unless the heteromer binds every SigD-controlled promoter. Another model for SinR-SlrR inhibition of motility gene expression is indirect, in which the heteromer inhibits the expression of SigD itself. In support of this model, the heteromer inhibits accumulation of SigD by inhibiting transcript levels of the *fla/che* operon somewhere between the $P_{fla/che}$ promoter and the *sigD* gene.⁸ The “indirect inhibition model”, however, cannot explain where the heteromer might bind within the *fla/che* operon, and even if it did, how transcriptional inhibition might occur. The sequence to which the SinR-SlrR heteromer binds is unknown.

Here we explore the mechanism of SigD inhibition by determining where in the genome SinR, SlrR, and the SinR-SlrR heteromer bind, using chromatin immunoprecipitation and DNA sequencing (ChIP-Seq). Our ChIP-Seq data support known SinR binding sites and reveal that SlrR binds to DNA in the absence of SinR. The data also support the idea that SinR and SlrR form a heteromer that binds to sites different from either homomer alone, with a predicted consensus sequence that resembles directly adjacent half-sites for each protein. Many of the heteromer binding sites were within open reading frames, including five locations within the *fla/che* operon. Genetic analysis indicated that two of the heteromer binding sites within *fla/che* were responsible for decreasing the frequency of SigD-dependent gene expression both when SlrA was

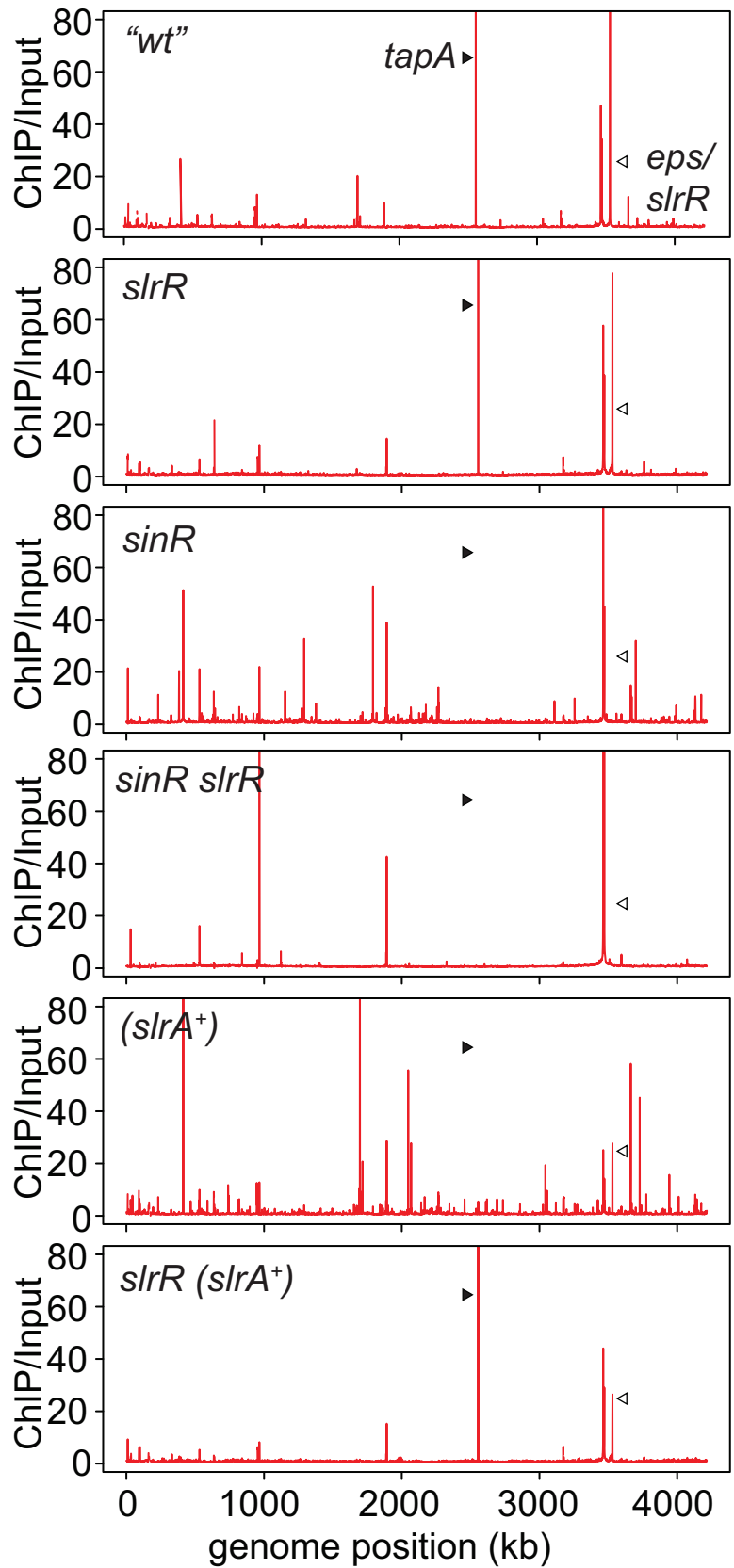
expressed in extra copy and when SwrA was absent. Moreover when bound, SinR-SlrR inhibited transcript abundance both at and downstream of the binding sites, likely by impeding elongation and/or promoting premature termination, and each site was sufficient for doing so when integrated in artificial reporter systems. In sum, our data support the indirect inhibition model of motile cell development in *B. subtilis*, in which the SinR-SlrR heteromer inhibits SigD levels. Impairment of transcription elongation by transcription factor binding is considered rare in bacteria, but may be more common than appreciated, and long operons may be particularly susceptible to such attenuation.

Results

Homomers and heteromers of SinR and SlrR bind to different sites

To explore the relative regulatory contributions of the paralogs SinR and SlrR, we performed ChIP-Seq analysis on exponentially growing wild type cells using a polyclonal antibody that reacts to both proteins.⁸ Sequences from the experimental IP are normalized to the control and projected as sequence abundance relative to chromosomal location. Thus, site enrichment increases with bound protein proximity such that the peak intensity has the highest likelihood of containing a specific binding sequence. As SinR represses the expression of SlrR during growth, any peaks obtained were expected to be largely due to binding by the SinR homomer (Figure 1). Consistent with expectations, relatively few peaks were observed, two of which were centered on known SinR binding sites, located in the intergenic regions upstream of the gene that encodes SlrR, the *eps* operon, and *tasA* operon^{31,39} (Figure 2, Figure 3A). Also consistent with expectation, mutation of SlrR produced a peak pattern very similar to that observed for wild type (Figure 3B), while mutation of SinR abolished the peaks upstream of the *eps* and *tasA* operons (Figure 2: Figure 3A, Table S1). MEME analysis of 200 base

Figure 2. ChIP-Seq analysis indicates different enrichment profiles for SinR, SlrR and the SinR-SlrR heteromer *in vivo*. ChIP-Seq was performed using a primary antibody to SinR that also cross-reacts with SlrR.⁸ For each sample, the number of normalized reads from samples treated with α -SinR was divided by the normalized reads for the corresponding untreated samples to determine fold enrichment (ChIP/Input). ChIP/input reads were plotted in 1 kb bins against genome position in kilobases. Each strain used in this experiment was deleted for *epsH* to abolish cell-clumping that occurs in the absence of SinR. Therefore, an *epsH* mutant was considered “wt” for this experiment. The genotype of the strain used to generate each panel is indicated in the top left corner. The following strains were used to generate the data in the indicated panel: “wt” (DS6776), *slrR* (DK9313), *sinR* (DK9090), *sinR slrR* (DK9314), (*slrA*⁺) (DK9093), and *slrR* (*slrA*⁺) (DK9332). The genotype (*slrA*⁺) indicates a “wt” strain with an additional copy of P_{slrA} -*slrA* integrated at an ectopic locus. Two carets indicate peaks that correspond to the promoter regions of the *eps* operon and the *tapA-slpW-tasA* operon and labelled with the first gene downstream to the promoter region, *epsA* and *tapA* respectively.



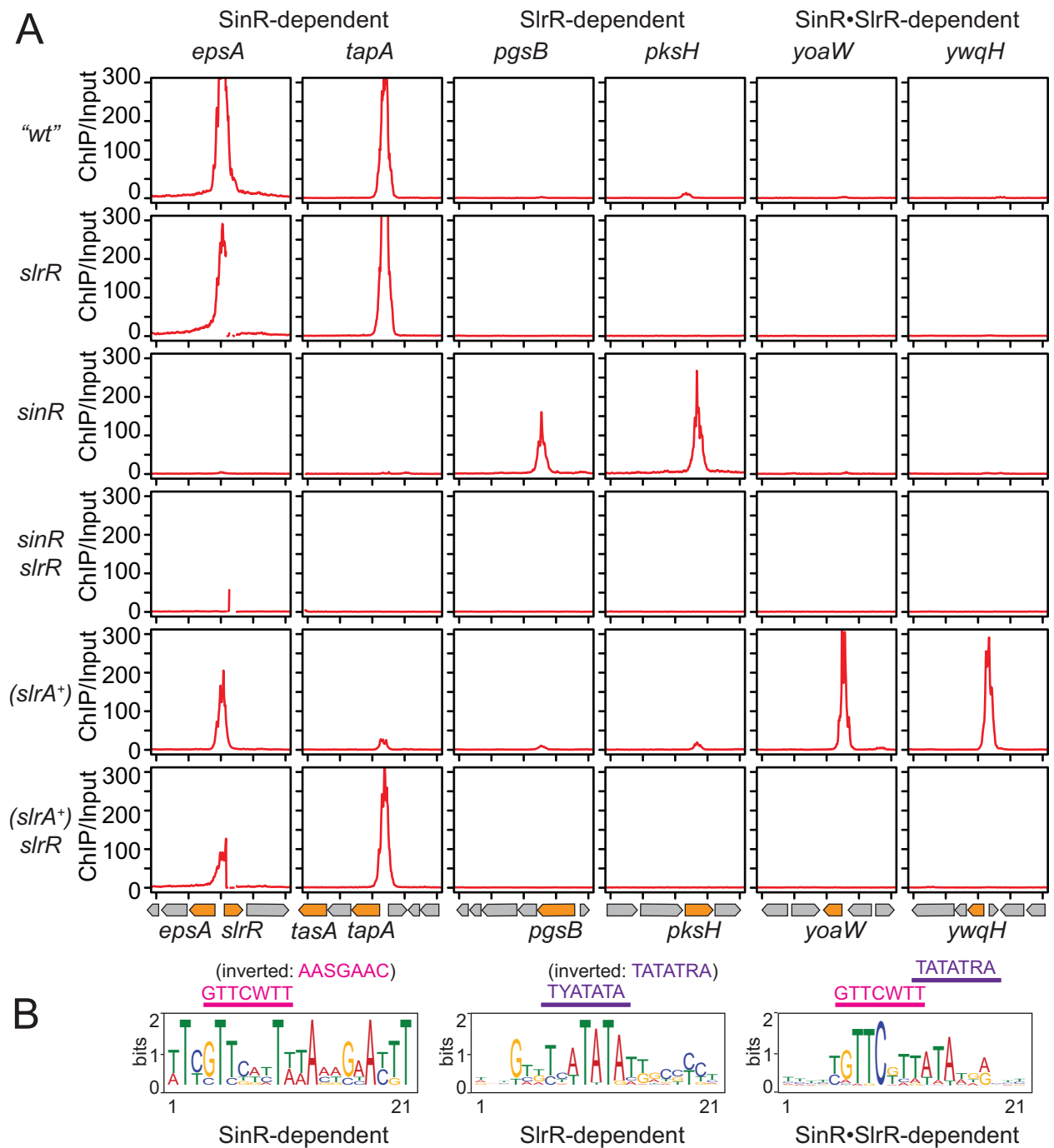


Figure 3. MEME analysis indicates distinct sequence patterns for SinR homomer, SlrR homomer and SinR-SlrR heteromer-dependent enrichment. A) Enlarged view of regions of interest from data presented in Figure 2. Analysis was performed the same way as in Figure 2 and the ChIP/input (Y-axis) was plotted in 10 bp bins over a 4 kb range. Ticks on the X-axis represent 1 kb intervals. The neighboring genome architecture is cartooned below each vertical panel and relevant genes of interest are colored orange. Each vertical panel is labelled after the gene or promoter region within which the peak was found. The following strains were used to generate this figure. “wt” (DS6776), *slrR* (DK9313), *sinR* (DK9090), *sinR slrR* (DK9314), (*slrA*⁺) (DK9093) and *slrR* (*slrA*⁺) (DK9332). Note, a discontinuity is observed in some of the peaks in the left-most column due to deletion of *slrR*. **B)** A 200 bp region around the center of peaks in each category was subjected to MEME analysis and a 21 bp MEME was generated. The predicted SinR half-site is marked in pink and is consistent with reported binding sites³⁴. The predicted half-site of SlrR is marked in purple. The following abbreviations were used to describe base conservation: W (A or T), S (G or C), Y (C or T), and R (A or G).

pair sequences surrounding the peak centers of SinR-dependent peaks identified binding sites for SinR, which were similar to the consensus half-site of GTTCTYT that was previously-identified³⁹ (Figure 3B, Table S1). We conclude that our genome-wide analysis supports pre-existing data and models where SinR is the predominant repressor that binds to and represses both SlrR and the loci required for biofilm formation.

In the wild type, peaks were enriched by ChIP-Seq that could not be attributed to SinR binding, and new peaks were enriched in a SinR mutant (Figure 2). Some of the peaks in the SinR mutant were possibly due to de-repression of SlrR, but to the best of our knowledge SlrR has not been reported to bind DNA or regulate gene expression on its own. ChIP-Seq analysis of a strain doubly mutated for SinR and SlrR abolished 20 peaks, which were deemed to be SlrR-dependent (Figure 2, Figure 3A, Table S2). Unlike the intergenic SinR-dependent peaks, the SlrR-dependent peaks were found to be overwhelmingly located within open reading frames (Table S2), and MEME analysis of 200 base pair sequences surrounding the peak centers indicated a putative consensus half-site of TYATATA (Figure 3B). Finally, some peaks remained in the absence of both SinR and SlrR and we wondered whether these might be due to one or more of the seventeen other Xre transcription factor family paralogs encoded by *B. subtilis*. Simultaneous mutation of SinR, SlrR, and the paralog YgzD, abolished a single additional peak upstream of the gene that encodes YgzD and the *ygzd* promoter was found to be auto-repressed (Figures S1, S2). We conclude that the polyclonal antibodies originally raised against SinR³¹ and known to cross-react with SlrR,⁸ also cross-react with YgzD and potentially other members of the family. We further conclude that SlrR and other SinR-paralogs bind to specific locations in the chromosome.

Previously reported genetic and biochemical data suggest that SinR and SlrR can form a heteromer that reprograms SinR to bind new sites in the chromosome when the small antagonist protein SlrA is in excess.^{6,8} Extra SlrA is thought to disrupt a subpopulation of SinR homomers, partially de-repress expression of SlrR and facilitate heteromer formation. To test for the binding of the heteromer, ChIP-Seq was performed on a strain that encoded an extra copy of the *slrA* gene expressed from its native promoter and integrated at an ectopic site in the chromosome (*slrA*⁺) (Figure 2). When *slrA* was present in extra-copy, SinR-dependent peaks were diminished, perhaps consistent with partial SinR antagonism, and a number of new peaks that were not previously attributed to either SinR or SlrR alone, were observed (Figure 2, Figure 3A, Table S3). As with the putative SlrR homomer, the peak sites for the heteromer were largely intragenic,

and mutation of SlrR abolished the additional peaks (Figure 2, Figure 2A). MEME analysis of 200 base pair sequences in the centers of the putative heteromer peaks indicated an elongated consensus sequence of GTTCWTTATATRA (Figure 3B). We note that the consensus appears to be the half-sites of SinR and SlrR respectively, directly juxtaposed. We conclude that the SinR-SlrR heteromer binds to a new set of genes that are not direct targets of either homomer.

SinR-SlrR binds within *lytA* and multiple sites within the *fla/che* operon

One model to explain how the SinR-SlrR heteromer inhibits SigD-dependent gene expression is by direct repression of the SigD-dependent genes *lytABC*, *lytF* and *hag*.³⁷ No peaks were detected near *lytF* or *hag* in any of the ChIP-Seq experiments, but a peak was detected near the *lytA* promoter region (Figure 4A). MEME analysis suggested a putative heteromer binding site was located within the *lytA* open reading frame near the 5' end, and mutation of SlrR but not SinR, abolished the peak (Figure 4A). To test the role of the putative binding site in the regulation of *lytA*, two promoter fusions were generated to the *lacZ* gene encoding β -galactosidase, one which included just the intergenic region upstream of *lytA* (P_{lytA} -*lacZ*) and one which included the intergenic region plus the putative intragenic binding site (P_{lytA}^{ext} -*lacZ*) (Figure 5A). Mutation of SinR or both SinR and SlrR together, did not alter expression of either reporter suggesting that the ChIP-Seq enrichment by SlrR alone was inconsequential (Figure 5B). In the presence of an extra copy of the *slrA* gene, however, expression of both reporters was reduced, and was restored when SlrR was also mutated, consistent with heteromer-dependent repression (Figure 5B). We conclude that the heteromer inhibits expression of the P_{lytA} promoter but whether the inhibition is direct or indirect is unclear as it appeared to be independent of the putative binding sequence.

Another model to explain how the SinR-SlrR heterodimer inhibits SigD-dependent gene expression is by inhibiting expression within the *fla/che* operon somewhere downstream of the $P_{fla/che}$ promoter and upstream of the gene encoding SigD.⁸ Consistent with the previous work suggesting that the $P_{fla/che}$ promoter was not a direct target, no peaks were detected near the $P_{fla/che}$ promoter in any of the ChIP-Seq experiments.⁸ Instead, four peaks were observed within the *fla/che* operon, and MEME analysis indicated five heteromer binding sites centered within the ChIP-Seq peaks. Thus, we named the putative sites: *site1* (within *fliE*), *site2* and *site3* (within *fliI*), *site4* (within *flhB*) and *site5* (within *cheC*). One way in which the heteromer could impair *sigD* expression is by binding to the sites and repressing internal promoters. No β -galactosidase activity was detected, however,

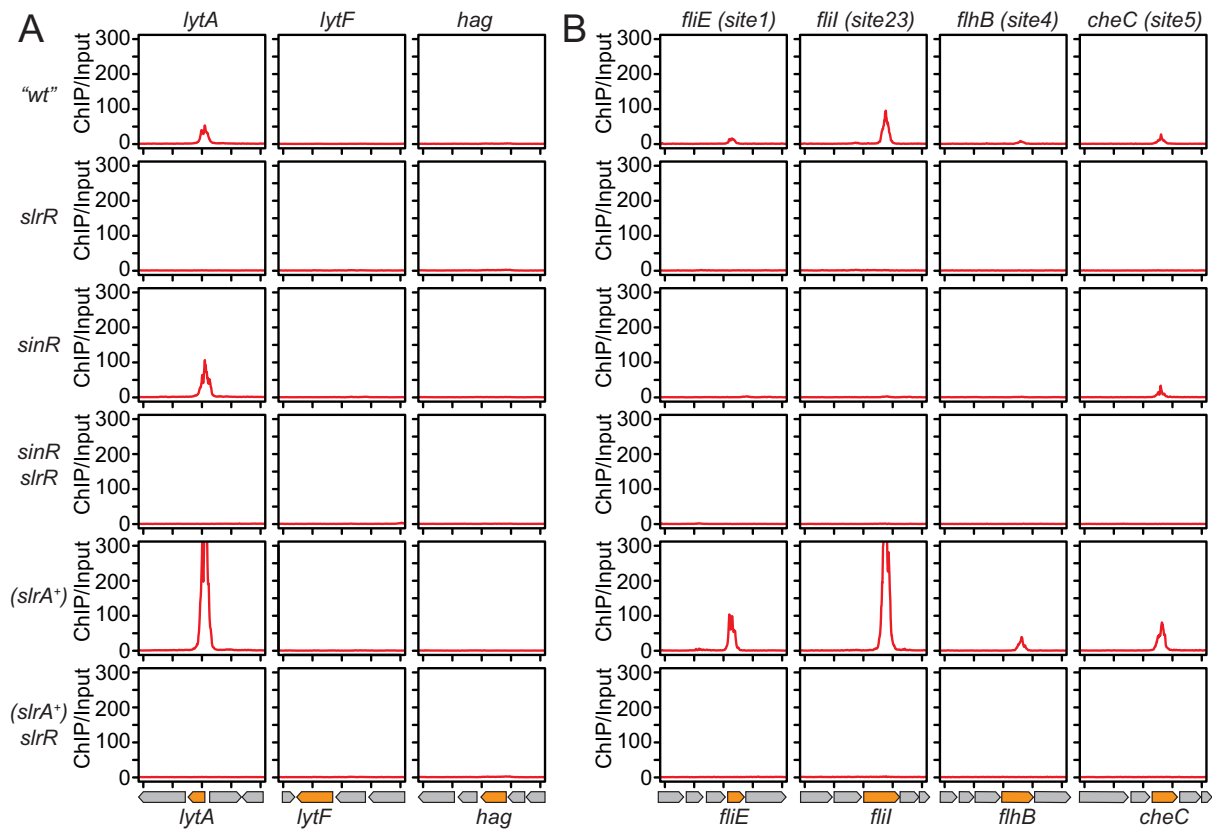


Figure 4. The SinR-SlrR heteromer is enriched at four locations within the *fla* operon. (A) Enlarged view of regions surrounding previously reported SinR-SlrR heterodimer target genes: *lytA*, *lytF* and *hag*. Analysis was performed the same way as in Figure 2 and the ChIP/input (Y-axis) was plotted in 10 bp bins over a 4 kb range. Ticks on the X-axis represent 1 kb intervals. Each vertical panel corresponds to a particular gene or promoter region indicated at top. The neighboring genome architecture is drawn below each vertical panel and relevant genes of interest are colored orange. **(B)** Enlarged view of regions enriched by the heteromer within the *fla/che* operon. A total of five putative binding sites were identified in four locations, and the sites were named *site1* through *site5*. Each vertical panel corresponds to a particular gene indicated at top. The following strains were used to generate this figure. "wt" (DS6776), *slrR* (DK9313), *sinR* (DK9090), *sinR slrR* (DK9314), (*slrA*⁺) (DK9093) and *slrR* (*slrA*⁺) (DK9332).

when a 500 bp region encompassing the peak center for each site was separately cloned upstream of the *lacZ* gene and inserted at an ectopic site in the wild type chromosome (Figure S3). We conclude that if the putative heteromer binding sites inhibit *sigD* gene expression, they do not do so by repressing nearby promoters.

To test the role of *sites 1* through 5, silent mutations were individually introduced in each of the SinR consensus half-sites such that the DNA binding sequence was altered but the protein code was not (Figure S4A). None of the silent mutations impaired swarming motility of the wild type, suggesting each was neutral on the effect of their respective flagellar genes (Figure S4B). Next, an extra copy of *slrA* was introduced into wild type and site mutant backgrounds to assess the effect of the heteromer on motility. Introduction of an extra copy of *slrA* in wild type resulted in complete loss of motility, and while mutation of

site1 in the same background caused a slight increase in swarming, none of the single point mutations was sufficient to restore motility to wild type levels (Figure S4B). Finally, reporters for SigD-dependent gene expression in which the promoter of the *hag* flagellin gene was cloned upstream of either the *lacZ* gene or the *gfp* gene encoding green fluorescent protein (GFP) were introduced (Figures 6, S5). Whereas wild type cells expressed high levels of β -galactosidase and GFP fluorescence, cells containing an extra copy of *slrA* were strongly inhibited for both reporters (Figure 6). Moreover, mutation of SlrR restored wild type levels of LacZ and GFP reporter expression in the presence of extra SlrA, but the single site mutations did not. We conclude that none of the binding site mutations were sufficient to restore wild swarming motility or SigD-dependent gene expression when inhibited by the SinR-SlrR heteromer. We note however, that

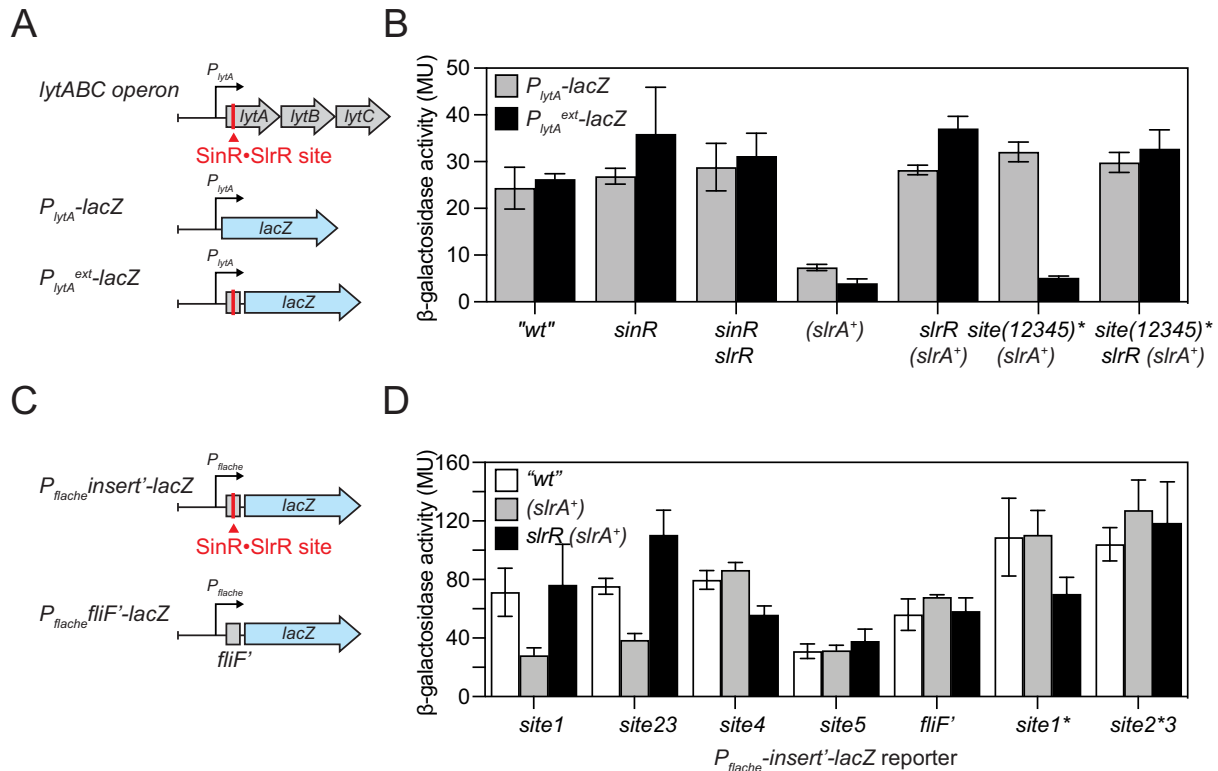


Figure 5. SinR-SlrR binding sites within *fla/che* and *lytA* are both necessary and sufficient for transcription inhibition. (A) A schematic of *lytA* genomic region and the two reporters, P_{lytA} -lacZ and P_{lytA}^{ext} -lacZ that were used in this experiment. P_{lytA} -lacZ includes the upstream promoter region of *lytA* transcriptionally fused to *lacZ*. P_{lytA}^{ext} -lacZ includes the upstream promoter region of *lytA* and the predicted SinR-SlrR binding site within *lytA* open reading frame transcriptionally fused to *lacZ*. (B) β -galactosidase activity in Miller units (MU) plotted on the Y-axis in linear scale. Each experiment was performed in three replicates and error bars indicate standard deviation of the three replicates. Gray bars represent expression of P_{lytA} -lacZ and black bars represents expression of P_{lytA}^{ext} -lacZ. The genetic background in which each reporter was tested is indicated on the X-axis. Each strain used in this experiment was deleted for *epsH* to abolish cell-clumping in the absence of SinR, and thus an *epsH* mutant was considered "wt" for this experiment. The following strains were used to generate this panel, "wt" (DB1829, DB1834), *sinR* (DB1830, DB1835), *sinR slrR* (DB1831, DB1836), (*slrA*⁺) (DB1832, DB1837), *slrR* (*slrA*⁺) (DB1845, DB1846), *site(12345)*^{*} (*slrA*⁺) (DB1833, DB1838) and *site(12345)*^{*} *slrR* (*slrA*⁺) (DB1880, DB1881). Raw data are presented in Table S4. (C) A schematic of transcriptional *lacZ* reporters of P_{flache} fused to 150 bp region surrounding the SinR-SlrR site within *fliE* (*site1*), *fliI* (*site23*), *fliH* (*site4*) and *cheC* (*site5*) and a site within *fliF* not associated with a heteromer binding peak in ChIP-seq analysis. Similar reporters were also constructed where P_{flache} was used to *site1*^{*} and *site2*^{*} mutants respectively. (D) β -galactosidase activity in Miller units (MU) plotted on the Y-axis in linear scale. Error bars are the standard deviation of three replicates. Each strain used in this panel was mutated for *epsE* to avoid clumping of cells in the absence of SinR and thus, an *epsE* mutant was considered as "wt" for this experiment. Each reporter was tested in "wt" (white bars), (*slrA*⁺) (gray bars) and *slrR* (*slrA*⁺) (black bars) genetic backgrounds. The following strains were used in this experiment, *site1* (DB1777, DB1765, DB1771), *site23* (DB1778, DB1766, DB1772), *site4* (DB1779, DB1767, DB1773), *site5* (DB1780, DB1768, DB1774), *fliF*' (DB1781, DB1769, DB1775), *site1*^{*} (DB1888, DB1889, DB1890) and *site2*^{*}3 (DB1891, DB1892, DB1893). Raw data are presented in Table S5.

minor rescue phenotypes may indicate additive effects of multiple sites within the operon.

Multiple SinR-SlrR binding sites within the *fla/che* operon are necessary for inhibiting SigD-dependent gene expression

To test the possibility that each of the putative SinR-SlrR binding sites within the *fla/che* operon had an additive effect on SigD inhibition,

mutations were sequentially added until all five sites had been disrupted. Simultaneous mutation of *site1*^{*} and *site2*^{*} (*site(12)*^{*}) restored partial swarming motility when an extra copy of *slrA* was present, and motility was further improved by additional mutations such that the quintuple mutant exhibited wild type swarming rates, albeit with an extended lag period (Figure S4B). Likewise, mutation of *site(12)*^{*} in cells containing an extra copy of *slrA* increased P_{hag} expression to

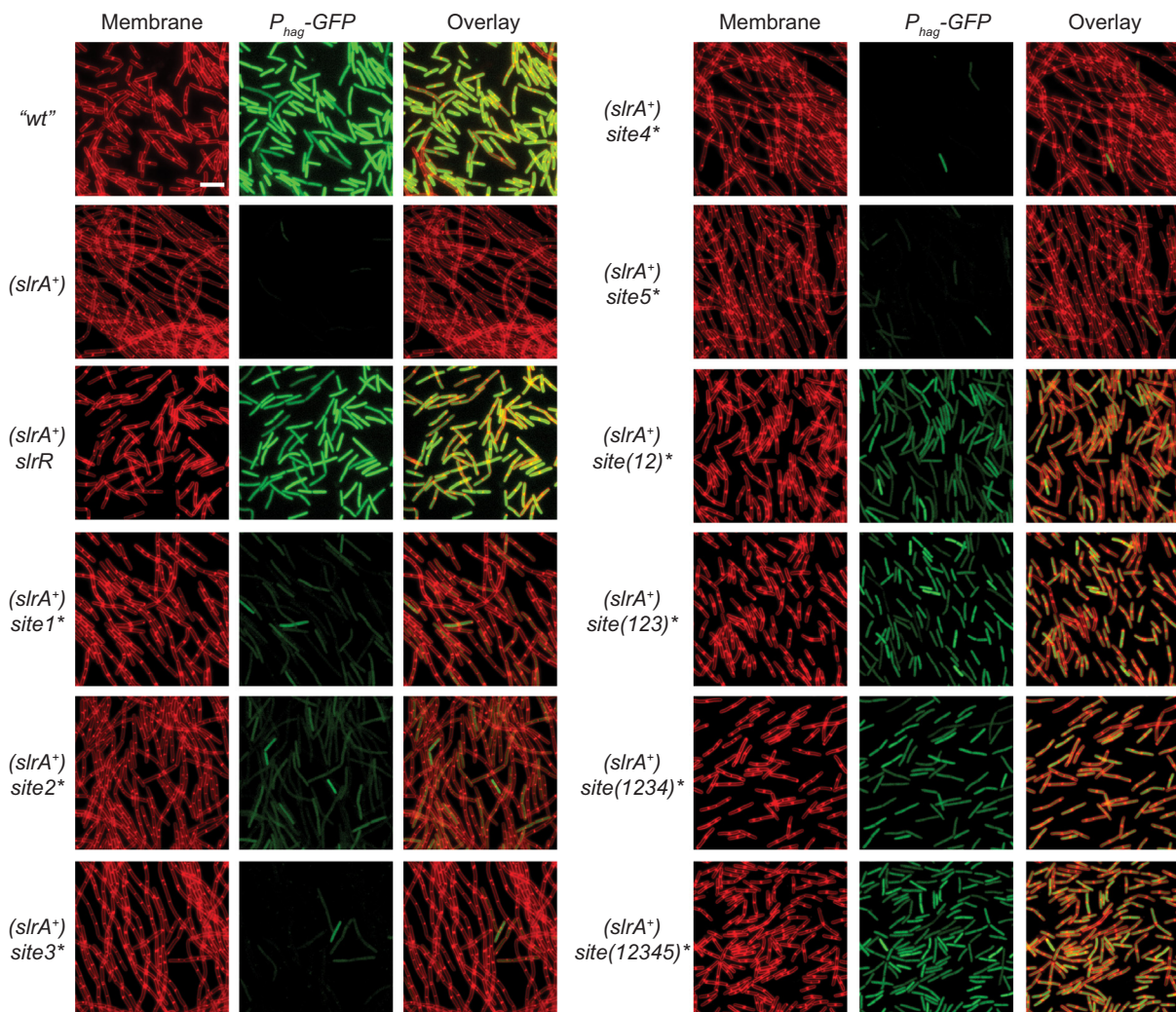


Figure 6. Mutation of SinR-SlrR binding sites 1 and 2 together is necessary to restore P_{hag} -GFP expression and cell separation in cells expressing an extra copy of $slrA$. Fluorescent micrographs of cells that contain a P_{hag} -GFP reporter for SigD-dependent gene expression (GFP, false colored green) and stained with FM 4–64 (membrane, false colored red). All strains used in this panel are mutated for $epsE$ (“ wt ”) and thus an $epsE$ mutant was considered as “ wt ” for this experiment. The following strains were used to generate the panels: “ wt ” (DB457), ($slrA^+$) (DB1306), $slrR$ ($slrA^+$) (DB1305), $site1^*$ ($slrA^+$) (DB1449), $site2^*$ ($slrA^+$) (DB1450), $site3^*$ ($slrA^+$) (DB1451), $site4^*$ ($slrA^+$) (DB1447), $site5^*$ ($slrA^+$) (DB1448), $site(12)^*$ ($slrA^+$) (DB1549), $site(123)^*$ ($slrA^+$) (DB1680), $site(1234)^*$ ($slrA^+$) (DB1675), $site(12345)^*$ ($slrA^+$) (DB1710). Scale bar is 8 μ m.

near wild type levels, but mutation of additional sites did little to improve expression further (Figures 6, S5). Finally, ChIP-Seq analysis in the quintuple mutant in cells containing an extra copy of $slrA$ indicated that mutation of first four sites abolished enrichment at their respective locations but mutation of $site5^*$ did not (Figure S6). We conclude that $site1$ and $site2$ are required for SinR-SlrR binding *in vivo*, have an additive effect when mutated in tandem, and together play a predominant role in inhibiting both swarming motility and SigD-dependent gene expression.

To determine if any of the SinR-SlrR predicted sites within the *fla/che* operon were sufficient to inhibit transcription, reporters were generated in

which 200 base pairs surrounding the peaks corresponding to $site1$, $site23$, $site4$, $site5$, and a control fragment from the gene *flif*, were cloned between the P_{flache} promoter and the *lacZ* gene (Figure 5C). Expression from the reporters with an intervening sequence containing either $site1$ or $site23$ were reduced in expression when $slrA$ was present in extra copy, but the remaining reporters were unaffected (Figure 5D). Moreover, the expression levels of both the $site1$ and $site2$ containing reporters was restored either in the absence of SlrR or when mutations that altered the putative binding site ($site1^*$ and $site2^*$, respectively) were introduced into the intervening sequence (Figure 5D). We conclude that while not

all heteromer binding sites are necessarily relevant, the *site1* and *site2* cis-elements specifically attenuate transcription when cloned between a promoter and reporter gene. We further note that while *site1* and *site2* were sufficient to inhibit transcription outside of the *fla/che* operon and away from linear proximity to the other sites, the effect might be magnified or otherwise altered by their cis arrangement at the native location.

To directly observe the effect of *site1* and *site2* on transcript abundance of the *fla/che* operon, RNAseq was performed in wild type and a variety of mutants. Transcript per million (TPM) values of each gene was calculated and normalized to the TPM of *sigA*, gene encoding the housekeeping sigma factor SigA. Similar to a previous report,⁸ transcript levels of the *fla/che* operon were high in the wild type, but decreased in abundance in when *slrA* was provided in extra copy, specifically near *site1*, with a further decrease that was observed near *site2* (Figure 7). Mutation of *site1** in the strain containing an extra copy of *slrA* raised transcript levels to that of wild type early in the operon, but transcript levels decreased near *site 2* and persisted at a low level (Figure 7). Finally, either mutation of both *site1* and *site2* simultaneously (*site(12)**), or mutation of SlrR restored wild type transcript levels to the *fla/che* operon and nearly all of the SigD-regulon in the presence of an extra copy of *slrA* (Figure 7, Figure S7). We conclude that heteromer binding to either *site1* or *site2* is both necessary and sufficient to attenuate transcript abundance downstream of promoter initiation. We further conclude that transcriptional attenuation within the *fla/che* operon contributes to the reduction in SigD protein levels to impair expression of the SigD-regulon.

We note that of the genes in the SigD regulon, transcript abundance of the *lytABC* operon was unique in that it was restored by mutation of SlrR but not the *site(12)** mutation (Figure S7). As the SinR-SlrR heteromer inhibits SigD accumulation, and expression from P_{lytA} is SigD-dependent, we wanted to determine why transcript abundance of *lytA* failed to increase when SigD activity was restored by mutation of the heteromer binding sites within *fla/che*. Simultaneous mutation of *sites (12345)**, restored expression of P_{lytA} -*lacZ* when an extra copy of *slrA* was present, consistent with the indirect model in which the heteromer acts through inhibiting SigD levels, but inconsistent with the low level of *lytA* transcript expression observed by RNAseq (Figure 5B, Figure S7). Mutation of *sites(12345)** however, did not restore expression to the reporter that included the intragenic heteromer binding site (P_{lytA} -*lacZ*) in the presence of an extra copy of *slrA*, thereby supporting a model where the heterodimer also directly inhibits *lytA* (Figure 5B). Finally, mutation of SlrR restored wild type levels of expression to the extended reporter when *slrA* was in extra copy and the *fla/che* operon binding sites were mutated (Figure 5B). We conclude that the SinR-SlrR heteromer primarily represses the SigD regulon by binding within the *fla/che* operon, but heteromer binding can also have direct effects within individual target genes. Thus, both models for transcriptional inhibition of the SigD regulon are at work, at least for the *lytA* gene.

To promote the formation of the SinR-SlrR heteromer, a strain that expresses an extra copy of the *slrA* gene has been used that switches the population heavily in favor of the SigD-OFF state.

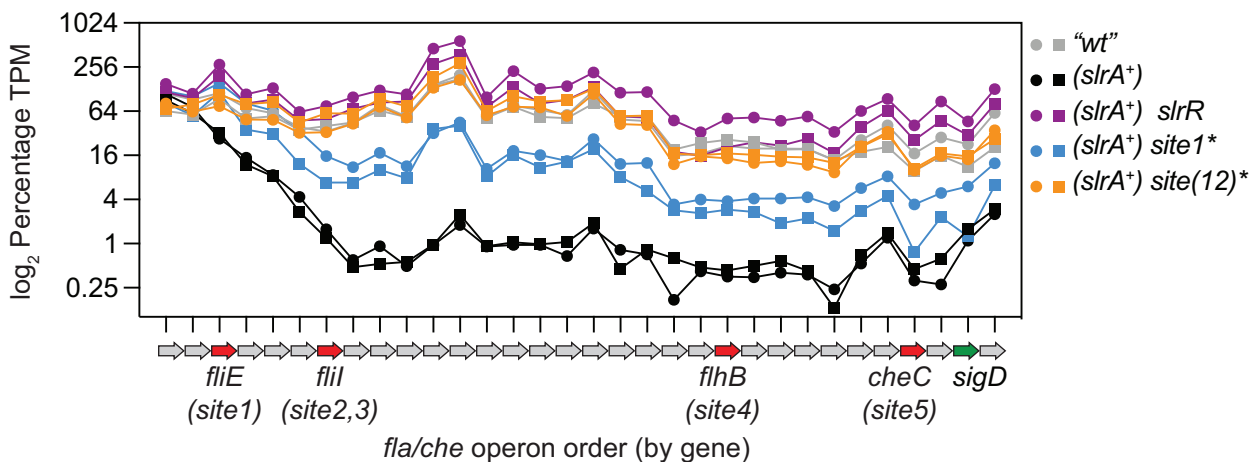
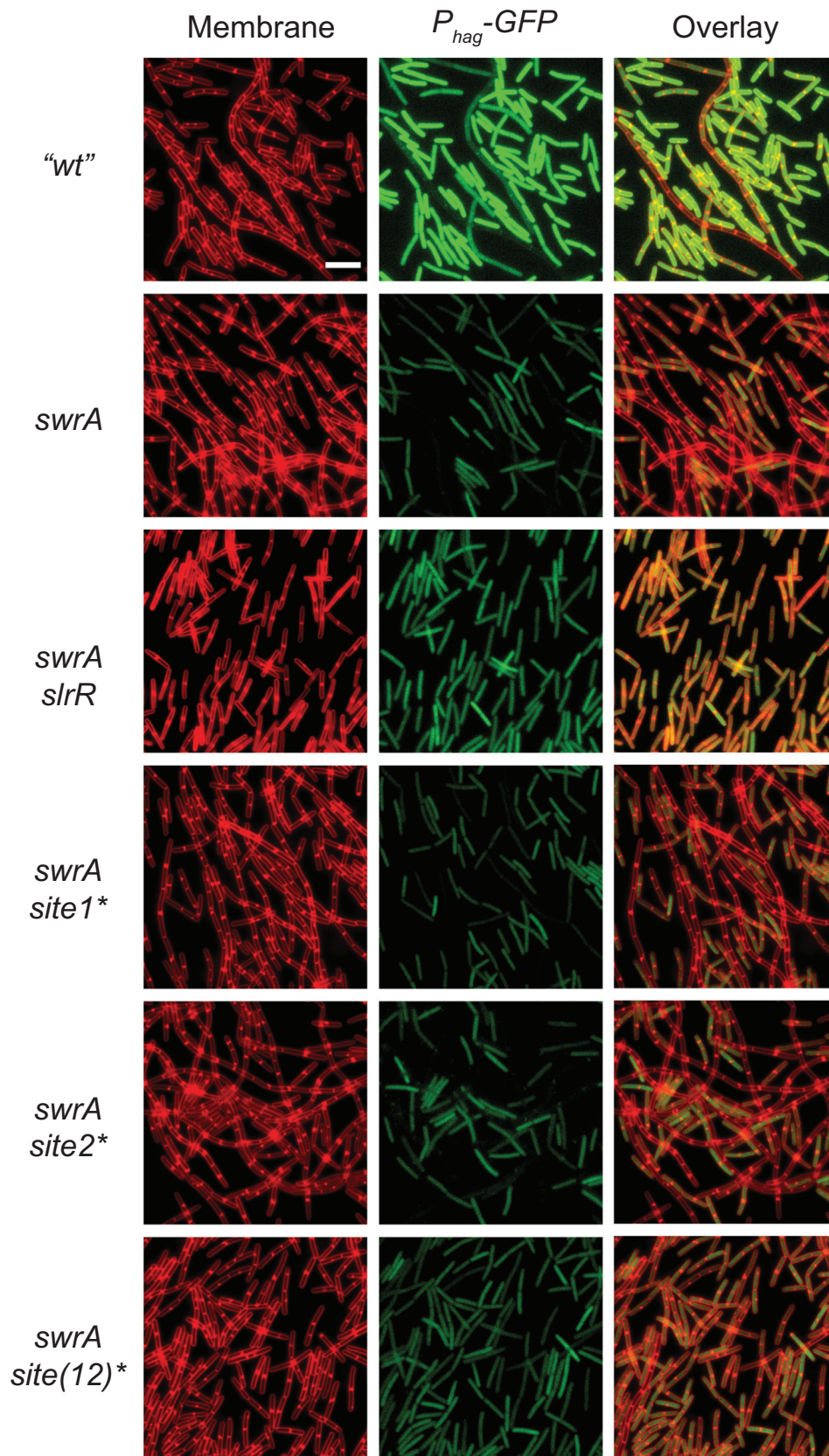


Figure 7. Binding of the SinR-SlrR heterodimer reduces transcript abundance within the *fla/che* operon. RNAseq transcriptomics were performed in duplicate (circles and squares) on the following strains. All strains were mutated for *epsE* (“wt”) for consistency with other experiments. “wt” (grey symbols, DK9699), (*slrA*⁺) (black symbols, DB154), (*slrA*⁺) *slrR* (purple symbols, DB141), (*slrA*⁺) *site1** (blue symbols, DB1438), and (*slrA*⁺) *site(12)** (orange symbols, DB1522). Transcript abundance by gene position along the *fla/che* operon is shown from 5' to 3' end on the (X-axis) as log₂ of the percentage transcript relative to the control transcript for the vegetative sigma factor SigA is shown on the (Y-axis). Raw data are available in Table S10.



Population heterogeneity with respect to motility, however, was first observed as both a reduction in fluorescence intensity and a reduction in the frequency of fluorescent SigD-ON cells in a strain mutated for the flagellar master activator protein SwrA and containing a P_{hag} -GFP reporter³ (Figure 8). Previous work indicated that the SinR-SlrR heteromer played a role as mutation of SlrR increased the frequency of SigD-ON cells in the absence of SwrA^{8,37} (Figure 8). Here we find that mutation of both *site1** and *site2** simultaneously, but not either site alone, was sufficient to increase the frequency of SigD-dependent gene expression in cells lacking SwrA, thereby phenocopying the absence of SlrR (Figure 8). Finally, the magnitude of SigD-dependent gene expression was not increased in either the SlrR mutant or the *site(12)** double mutant, consistent with the heteromer acting downstream of SwrA activation at the P_{flache} promoter.^{23,25,26,3} We conclude that the SwrA-DegU heteromer increases the magnitude of *fla/che* operon expression at the level of transcript initiation, while the SinR-SlrR heteromer attenuates transcript abundance within the operon. Together the two heteromeric systems calibrate the frequency at which the SigD-regulon is activated.

Discussion

Bacteria were once thought to be physiologically uniform during exponential growth, but growing *B. subtilis* spontaneously bifurcates into two phenotypically-distinct subpopulations: single motile cells and long non-motile chains.³ Each cell type is differentiated at the level of gene expression governed in part, by the DNA binding repressor protein SinR.^{8,37} SinR represses genes involved biofilm formation as a homomer,^{31,34,39} but transient antagonism relieves repression of a paralog called SlrR to form a SinR-SlrR heteromer.^{35,36} Genetic evidence indicates that the heteromer reprograms SinR to bind to new sites in the genome that repress a regulon for flagellar assembly and cell separation under the control of the alternative sigma factor SigD.^{8,37} Here we show that the heteromer binds to multiple sites within the long *fla/che* operon that are both necessary and sufficient for attenuating transcript abundance, likely by promoting RNA polymerase pausing and premature termination.

The gene encoding SigD is downstream of the binding sites, and we conclude that the SinR-SlrR heteromer indirectly inhibits SigD activity by preventing SigD accumulation above a threshold in a subpopulation of cells^{6,8} (Figure 1A).

Repressors commonly inhibit transcriptional activation by binding to sites that occlude access of RNA polymerase to promoter elements.^{40–42} Sometimes, repressors bind within open reading frames near the promoter and alter promoter access remotely by DNA looping or other changes in conformation. DNA binding repressors that bind within genes to inhibit elongation/promote premature termination like the SinR-SlrR heteromer are rare but one example is the global regulator CodY of *B. subtilis*.⁴³ CodY represses transcriptional initiation of many genes in response to cellular GTP and amino acids levels,^{44–46} but like SinR-SlrR, CodY also binds within the *fla/che* operon to antagonize SigD, perhaps as a “roadblock” to transcriptional elongation.^{47,48} How DNA binding proteins would inhibit RNA polymerase transcription bubble progression, however, is unclear as often, a dimer binds on the same surface of the dsDNA making impairment of unwinding unlikely. We note however that the SinR and SlrR inverted half-sites directly abut in the case of heteromer binding, spanning roughly 14 nucleotides. As 10 nucleotides constitute a helical turn, we speculate that binding of the heteromer might wrap all the way around the DNA and act as a clamp. While some RNA readthrough was observed at heteromer bound sites, the binding might pause RNA polymerase long enough to promote premature termination and cause polarity on downstream gene expression.

By whatever mechanism heteromer binding inhibits transcription, we note that not all binding sites were effective. For example, while *site1* and *site2* within the *fla/che* operon both induced a local decrease in transcript abundance at both the native site and in heterologous reporter assays, *site3*, *site4*, and *site5* did not. Moreover, ChIP-Seq analysis indicated that SlrR alone could bind DNA within open-reading frames, but seemed unable to inhibit transcription. In the case of the *lytA* gene, both SlrR-dependent heteromer and SlrR-dependent homomer enrichment was substantial, but the effect of transcription inhibition was only observed when SinR was present. Thus, we infer

Figure 8. Double mutation of SinR-SlrR binding sites(12)* restores frequency of P_{hag} -GFP expression in cells lacking SwrA. Fluorescent micrographs of cells that contain a P_{hag} -GFP reporter for SigD-dependent gene expression (GFP, false colored green) and stained with FM 4–64 (membrane, false colored red). All strains used in this panel are mutated for *epsE* and an *epsE* mutant was considered as “*wf*” for this experiment to maintain consistency with other figures in the manuscript. P_{hag} -GFP expression in the indicated genetic background is represented in the fluorescent micrographs. “*wf*” (DB457), *swrA* (DB1456), *swrA slrR* (DB1457), *swrA site1** (DB1444), *swrA site2** (DB1445), and *swrA site(12)** (DB1543). Scale bar is 8 μ m.

that transcript attenuation depends on both heteromer formation, and particular *cis*-element sequences, but how the combination of the two promotes transcriptional termination is unclear. It is also unclear why transcript attenuation is used to inhibit expression, instead of the more commonly observed mechanism of promoter inhibition. At least in this case, the attenuation works in the context of a long operon to decrease gene expression in a manner proportional to the distance from the binding site. Thus, perhaps partial expression of genes early in the operon is somehow beneficial, or there may be timing benefits of targeting a longer window of elongation rather than the instantaneous event of promoter initiation.⁴⁹ Finally, we note that activation by SwrA-DegU can override SinR-SlrR during swarming motility, and SinR-SlrR dampening can override SwrA-DegU during biofilm formation. Thus expression may be fine-tuned by differential regulation of two processes that act in opposition.

Ultimately, we generate a comprehensive molecular model for both population heterogeneity and the transition from motility to biofilm formation (Figure 1). In conditions where SwrA is either mutated or otherwise low in cytoplasm, the SinR-SlrR heteromer increases the frequency of non-motile chains, and chaining cells have been thought to be a precursor to biofilm formation.^{36,50–52} We also find that hyperactivation of the SinR-SlrR heteromer can override motility gene expression even in the presence of SwrA, as in the artificial case where an extra copy of the *slrA* gene is provided. The *slrA* extra copy condition likely resembles the situation when biofilm formation is activated, where SinR repression fails and the *eps* operon is expressed to produce the EPS polysaccharide component that promotes cohesion.³¹ As a consequence, EpsE, a protein encoded within the *eps* operon, interacts with the flagellum to rapidly arrest rotation,^{32,33,53} and the SinR-SlrR heteromer attenuates the *fla/che* operon transcript so that motility-inhibited biofilm cells grow without synthesizing new flagella. Together, a complex system of functional and transcriptional inhibitors operate at fast and slow timescales, to promote and stabilize biofilm development respectively.

Materials and Methods

Strain and growth conditions

B. subtilis strains were grown in lysogeny broth (LB) (10 g tryptone, 5 g yeast extract, 5 g NaCl per liter) broth or on LB plates fortified with 1.5% Bacto agar at 37 °C. The following antibiotic concentrations were used when necessary: ampicillin 100 µg/ml (*amp*), kanamycin 5 µg/ml (*kan*), chloramphenicol 5 µg/ml (*cm*), spectinomycin 100 µg/ml (*spec*), tetracycline 10 µg/ml (*tet*), and erythromycin 1 µg/ml plus lincomycin 25 µg/ml (*mls*).

Strain construction

B. subtilis chromosomal DNA from indicated strains was used to amplify all PCR products. All constructs were transformed into the naturally competent DK1042 that carries a *comI*^{Q12L} mutation in the 3610 ancestral strain. SPP1 phage lysate of strains carrying constructs with selectable markers were prepared and transduced into desired genetic backgrounds using generalized transduction.

SPP1 phage transduction. Donor *Bacillus subtilis* strains were grown in TY broth (LB broth supplemented with 10 mM MgSO₄ and 100 µM MnSO₄). Serial dilutions of SPP1 phage stock were added to 0.2 ml of dense culture (OD₆₀₀ 0.6–1.0) and statically incubated at 37 °C for 15 mins. 3 ml of molten TY soft agar (TY supplemented with 0.5% agar) was added to each mixture, poured on top of fresh TY agar plates (TY supplemented with 1.5% agar) and incubated at 37 °C overnight. The top agar a plate containing near confluent plaques was scraped and collected in a 15 ml conical tube, vortexed and centrifuged at 5,000g for 10 mins. The supernatant that contained phage particles was passed through a 0.45 µm syringe filter to eliminate any bacterial contamination and stored at 4 °C. Recipient strains were grown to OD₆₀₀ 0.6–1.0 in TY broth at 37 °C and one ml of cells were mixed with 25 µl of SPP1 phage stock from the donor. 9 ml of TY broth was added to the mixture and the mixture was incubated at room temperature for 30 mins with gentle shaking on a rocker. The mixture was centrifuged at 5000g for 10 mins, supernatant was discarded, pellet was resuspended in the remaining volume and the 100 µl of the cell suspension was plated on LB plates fortified with 1.5% agar, supplemented with the appropriate antibiotics and 10 mM sodium citrate. The plates were incubated at 37 °C overnight. All strains used in this study are listed in Table 1. All primers used to build strains for this study are listed in Table S8 and all plasmids are listed in Table S9.

Transcriptional reporter constructs. *pAM58, 59, 60, 61, 81, 109.* DK1042 chromosomal DNA was used to amplify regions using primer pairs 7928/7929, 7930/7931, 7926/7927, 7924/7925, 8139/8140, 8141/8142 and ligated into the EcoRI/BamHI sites of pDG1663 containing the *lacZ* gene and the gene for *mls* resistance between arms of *thrC* to generate *pAM58, 59, 60, 61, 81* and *109* respectively. These were separately transformed into DK1042 and integration at the *thrC* locus was confirmed by the ability of the mutant to grown on *mls* and the inability of mutant to grow in defined media in the absence of threonine.

pAM112. DK1042 chromosomal DNA was used to amplify approximately 500 bp upstream of the *flache* ribosomal binding site using primers

Table 1 Strains.

Strain	Genotype ^a
3610	wildtype
DB141	$\Delta epsE$ <i>slrR::tet amyE::P_{slrA}-slrA cat</i>
DB143	$\Delta epsE$ <i>thrC::site1-lacZ mls</i>
DB144	$\Delta epsE$ <i>thrC::site23-lacZ mls</i>
DB145	$\Delta epsE$ <i>thrC::site4-lacZ mls</i>
DB146	$\Delta epsE$ <i>thrC::site5-lacZ mls</i>
DB154	$\Delta epsE$ <i>amyE::P_{slrA}-slrA cat</i>
DB225	$\Delta epsH$ <i>slrR::tet sinR::spec ygzD::kan</i>
DB402	$\Delta epsE$ <i>thrC::P_{hag}-lacZ mls</i>
DB643	$\Delta epsE$ <i>thrC::P_{flache}-lacZ mls</i>
DB995	$\Delta epsE$ <i>thrC::P_{ygzd}-lacZ mls</i>
DB996	$\Delta epsE$ <i>sinR::spec thrC::P_{ygzd}-lacZ mls</i>
DB997	$\Delta epsE$ <i>slrR::tet sinR::spec thrC::P_{ygzd}-lacZ mls</i>
DB998	$\Delta epsE$ <i>slrR::tet sinR::spec ygzD::kan thrC::P_{ygzd}-lacZ mls</i>
DB1305	$\Delta epsE$ <i>slrR::tet amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1306	$\Delta epsE$ <i>amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1357	$\Delta epsE$ <i>site4*</i>
DB1358	$\Delta epsE$ <i>site5*</i>
DB1360	$\Delta epsE$ <i>site1*</i>
DB1419	$\Delta epsE$ <i>site2*</i>
DB1420	$\Delta epsE$ <i>site3*</i>
DB1436	$\Delta epsE$ <i>site4* amyE::P_{slrA}-slrA cat</i>
DB1437	$\Delta epsE$ <i>site5* amyE::P_{slrA}-slrA cat</i>
DB1438	$\Delta epsE$ <i>site1* amyE::P_{slrA}-slrA cat</i>
DB1439	$\Delta epsE$ <i>site2* amyE::P_{slrA}-slrA cat</i>
DB1440	$\Delta epsE$ <i>site3* amyE::P_{slrA}-slrA cat</i>
DB1444	$\Delta epsE$ <i>site1* swrA::kan thrC::P_{hag}-GFP mls</i>
DB1445	$\Delta epsE$ <i>site2* swrA::kan thrC::P_{hag}-GFP mls</i>
DB1447	$\Delta epsE$ <i>site4* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1448	$\Delta epsE$ <i>site5* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1449	$\Delta epsE$ <i>site1* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1450	$\Delta epsE$ <i>site2* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1451	$\Delta epsE$ <i>site3* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1456	$\Delta epsE$ <i>swrA::kan thrC::P_{hag}-GFP mls</i>
DB1457	$\Delta epsE$ <i>swrA::kan slrR::tet thrC::P_{hag}-GFP mls</i>
DB1515	$\Delta epsE$ <i>site(12)*</i>
DB1522	$\Delta epsE$ <i>site(12)* amyE::P_{slrA}-slrA cat</i>
DB1543	$\Delta epsE$ <i>site(12)* swrA::kan thrC::P_{hag}-GFP mls</i>
DB1549	$\Delta epsE$ <i>site(12)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1598	$\Delta epsE$ <i>site(123)*</i>
DB1609	$\Delta epsE$ <i>site(123)* amyE::P_{slrA}-slrA cat</i>
DB1652	$\Delta epsE$ <i>site(12)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1661	$\Delta epsE$ <i>slrR::tet amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1662	$\Delta epsE$ <i>amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1663	$\Delta epsE$ <i>site1* amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1667	$\Delta epsE$ <i>site(1234)*</i>
DB1669	$\Delta epsE$ <i>site(1234)* amyE::P_{slrA}-slrA cat</i>
DB1675	$\Delta epsE$ <i>site(1234)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1680	$\Delta epsE$ <i>site(123)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1681	$\Delta epsE$ <i>site(123)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1702	$\Delta epsE$ <i>site(12345)*</i>
DB1705	$\Delta epsE$ <i>site(12345)* amyE::P_{slrA}-slrA cat</i>
DB1709	$\Delta epsE$ <i>site(12345)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1710	$\Delta epsE$ <i>site(12345)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-GFP mls</i>
DB1737	$\Delta epsE$ <i>site(1234)* amyE::P_{slrA}-slrA cat thrC::P_{hag}-lacZ mls</i>
DB1765	$\Delta epsE$ <i>ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site1-lacZ cat</i>
DB1766	$\Delta epsE$ <i>ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site23-lacZ cat</i>
DB1767	$\Delta epsE$ <i>ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site4-lacZ cat</i>
DB1768	$\Delta epsE$ <i>ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site5-lacZ cat</i>
DB1769	$\Delta epsE$ <i>ycgO::P_{slrA}-slrA kan amyE::P_{flache}-flIF-lacZ cat</i>
DB1771	$\Delta epsE$ <i>slrR::tet ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site1-lacZ cat</i>
DB1772	$\Delta epsE$ <i>slrR::tet ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site23-lacZ cat</i>
DB1773	$\Delta epsE$ <i>slrR::tet ycgO::P_{slrA}-slrA kan amyE::P_{flache}-site4-lacZ cat</i>

Table 1 (continued)

Strain	Genotype ^a
DB1774	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}site5\text{-}lacZ cat$
DB1775	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}fliF\text{-}lacZ cat$
DB1777	$\Delta epsE amyE::P_{flache}\text{-}site1\text{-}lacZ cat$
DB1778	$\Delta epsE amyE::P_{flache}\text{-}site23\text{-}lacZ cat$
DB1779	$\Delta epsE amyE::P_{flache}\text{-}site4\text{-}lacZ cat$
DB1780	$\Delta epsE amyE::P_{flache}\text{-}site5\text{-}lacZ cat$
DB1781	$\Delta epsE amyE::P_{flache}\text{-}fliF\text{-}lacZ cat$
DB1785	$\Delta epsE site4^* amyE::P_{slrA}\text{-}slrA cat thrC::P_{hag}\text{-}lacZ mls$
DB1786	$\Delta epsE site5^* amyE::P_{slrA}\text{-}slrA cat thrC::P_{hag}\text{-}lacZ mls$
DB1787	$\Delta epsE site2^* amyE::P_{slrA}\text{-}slrA cat thrC::P_{hag}\text{-}lacZ mls$
DB1788	$\Delta epsE site3^* amyE::P_{slrA}\text{-}slrA cat thrC::P_{hag}\text{-}lacZ mls$
DB1829	$\Delta epsE amyE::P_{lytA}\text{-}lacZ cat$
DB1830	$\Delta epsE sinR::spec amyE::P_{lytA}\text{-}lacZ cat$
DB1831	$\Delta epsE slrR::tet sinR::spec amyE::P_{lytA}\text{-}lacZ cat$
DB1832	$\Delta epsE ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}\text{-}lacZ cat$
DB1833	$\Delta epsE site(12345)^* ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}\text{-}lacZ cat$
DB1834	$\Delta epsE amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1835	$\Delta epsE sinR::spec amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1836	$\Delta epsE slrR::tet sinR::spec amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1837	$\Delta epsE ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1838	$\Delta epsE site(12345)^* ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1845	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}\text{-}lacZ cat$
DB1846	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1880	$\Delta epsE site(12345)^* slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}\text{-}lacZ cat$
DB1881	$\Delta epsE site(12345)^* slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{lytA}^{ext}\text{-}lacZ cat$
DB1888	$\Delta epsE amyE::P_{flache}\text{-}site1^*\text{-}lacZ cat$
DB1889	$\Delta epsE ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}site1^*\text{-}lacZ cat$
DB1890	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}site1^*\text{-}lacZ cat$
DB1891	$\Delta epsE amyE::P_{flache}\text{-}site2^*3\text{-}lacZ cat$
DB1892	$\Delta epsE ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}site2^*3\text{-}lacZ cat$
DB1893	$\Delta epsE slrR::tet ycgO::P_{slrA}\text{-}slrA kan amyE::P_{flache}\text{-}site2^*3\text{-}lacZ cat$
DK1042	$comI^{\Delta 12L}$
DK9090	$\Delta epsH sinR::spec$
DK9093	$\Delta epsH amyE::P_{slrA}\text{-}slrA cat$
DK9313	$\Delta epsH slrR::tet$
DK9314	$\Delta epsH slrR::tet sinR::spec$
DK9332	$\Delta epsH slrR::tet amyE::P_{slrA}\text{-}slrA cat$
DK9699	$\Delta epsE$
DS6776	$\Delta epsH$

^a All *B. subtilis* strains are in either 3610 or DK1042 genetic backgrounds.

8008/8423 and ligated into the EcoRI/ BamHI sites of pDG268 containing the *lacZ* gene and the *cat* gene for chloramphenicol resistance between arms of *amyE* to generate *pAM112*.

pAM113, 114, 115, 116, 117. DK1042 chromosomal DNA was used to amplify ~500 bp regions surrounding *sites 1, 23, 4, 5* and a region within *fliF* using primer pairs 8424/8425, 8426/8427, 8428/8429, 8430/8431 and 8432/8433 and ligated into the NheI/BamHI sites of *pAM112* to generate *pAM113*, 114, 115, 116 and 117 respectively. These plasmids were separately transformed into DK1042 and chromosomal integration into the *amyE* site was confirmed by resistance of the transformants to chloramphenicol and their inability to digest sucrose when grown on LB media supplemented with sucrose.

pAM118. Chromosomal DNA from *B. subtilis* strain DB1360 was used as a template to amplify 500 bp region surround *site1^** mutation using primer pairs 8424/8425 and ligated into the NheI/BamHI sites of *pAM112* to generate *pAM118*. *pAM118* was transformed into DK1042 and the transformants were confirmed as mentioned above.

pAM119. Chromosomal DNA from *B. subtilis* strain DB1419 was used as a template to amplify 500 bp region surround *site2^** mutation using primer pairs 8426/8427 and ligated into the NheI/BamHI sites of *pAM112* to generate *pAM119*. *pAM119* was transformed into DK1042 and the transformants were confirmed as mentioned above.

***slrA* complementation construct. *pAM103*.** DK1042 chromosomal DNA was used to amplify

approximately 500 bp upstream of the *slrA* ribosomal binding site using primers 8337/8338 and ligated into the EcoRI/BamHI sites of *pKM087* to generate *pAM103*. *pAM103* was transformed into DK1042 and integration into the chromosome was confirmed by the presence of a rugose colony morphology because of the presence of an extra copy of *slrA*.

Native site mutants. pDP620. DK1042 chromosomal DNA was used to amplify ~1000 bp flanking fragments surrounding *site1* using primer pairs 8273/8274 and 8275/8276 that contained the *site1** mutation. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

pDP621. DK1042 chromosomal DNA was used to amplify ~1000 bp flanking fragments surrounding *site23* using primer pairs 8277/8278 and 8279/8280 that contained the *site2** mutation. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

pDP622. DK1042 chromosomal DNA was used to amplify ~1000 bp flanking fragments surrounding *site23* using primer pairs 8281/8282 and 8283/8284 that contained the *site3** mutation. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

pDP623. DK1042 chromosomal DNA was used to amplify ~1000 bp flanking fragments surrounding *site4* using primer pairs 8285/8286 and 8287/8288 that contained the *site4** mutation. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

pDP624. DK1042 chromosomal DNA was used to amplify ~1000 bp flanking fragments surrounding *site5* using primer pairs 8289/8290 and 8291/8292 that contained the *site5** mutation. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

pDP638. Chromosomal DNA from *B. subtilis* strain DB1419 was used to amplify ~1000 bp flanking fragments using primer pairs 8281/8397 and 8398/8284 that contained both *site2** and *site3** mutations. *pminiMAD3* was linearized by digesting with SmaI and the two flanking fragments were assembled into *pminiMAD3* by Gibson assembly.

The plasmids were passaged individually through *recA + E. coli* strain TG1, transformed into DK1042 and plated at restrictive temperature for plasmid replication (37 °C) on LB agar supplemented with *spec* to select for transformants with single crossover plasmid integration. Plasmid eviction was ensured by growing the strains for 14 h at a

permissive temperature for plasmid replication (22 °C) in the absence of *spec* selection. Cells were serially diluted, plated on LB agar plates in the absence of *spec* and individual colonies were replica patched on LB agar plates with and without *spec* to identify *spec* sensitive colonies that have successfully evicted the plasmid. Chromosomal DNA was isolated from the colonies that had excised the plasmid and allelic replacement was confirmed by sequencing.

Swarm expansion assay

1 mL of mid-log phase cells (OD₆₀₀ 0.3–1.0) grown at 37 °C in LB were harvested and resuspended to an OD₆₀₀ of 10 in pH 8.0 PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄) containing 0.5% India ink (Higgins). Freshly prepared LB plates fortified with 0.7% bacto agar (25 mL per plate) was dried for 10 min in a laminar flow hood, centrally inoculated with 10 µL of the cell suspension, dried for another 10 min, and incubated at 37 °C. The India ink demarks the origin of the colony and the swarm radius was measured relative to the origin every 30 min. For consistency, an axis was drawn on the back of the plate and swarm radii measurements were taken along this transect.

β-galactosidase assay

B. subtilis strains were grown in LB broth at 37 °C with constant rotation to OD₆₀₀ 0.7–1.0. One mL of cells was harvested by centrifugation and resuspended in 1 mL of Z-buffer (40 mM NaH₂PO₄, 60 mM Na₂HPO₄, 1 mM MgSO₄, 10 mM KCl and 38 mM β-mercaptoethanol). To each sample, lysozyme was added to a final concentration of 0.2 mg/mL and incubated at 30 °C for 15 min. Each sample was diluted appropriately in 500 µL of Z-buffer and the reaction was started with 100 µL of start buffer (4 mg/ml 2-nitrophenyl β-D-galactopyranoside (ONPG) in Z-buffer) and stopped with 250 µL 1 M Na₂CO₃. The OD₄₂₀ of the reaction mixtures were recorded and the β-galactosidase specific activity was calculated according to the equation: (OD₄₂₀/time × OD₆₀₀) × dilution factor × 1000.

Chromatin immunoprecipitation sequencing (ChIP-Seq)

Bacillus subtilis cultures were grown to an OD₆₀₀ of 1.0 at 37 °C with constant rotation. 20 mL of cells were cross-linked for 30 min at room temperature using 3% formaldehyde (Sigma), quenched with 125 mM glycine, washed with PBS, and then lysed. DNA was sheared to an average fragment size of 170 bp using Qsonica sonicator (Q8000R), and then incubated overnight at 4 °C with α-SinR³⁴. Immunoprecipitation was performed

using Protein A Magnetic Sepharose beads (Cytiva #45002511), washed, and DNA was eluted in TES (50 mM Tris pH8, 10 mM EDTA and 1% SDS). Crosslinks were reversed overnight at 65 °C. DNA samples were treated with a final concentration of 0.2 mg/ml RNaseA (Promega #A7973) and 0.2 mg/ml Proteinase K (NEB #P8107S) respectively, and subsequently extracted using phenol/chloroform/isoamyl (25:24:1). DNA samples were then used for library preparation using NEBNext Ultrall DNA library prep kit (NEB #E7645L). Paired end sequencing of the libraries was performed using NextSeq 500 platform and at least 3 million paired-end reads were obtained for each sample. Two or three biological replicates were sequenced for each sample.

Whole genome sequencing (WGS)

B. subtilis cultures were grown to an OD₆₀₀ of 1.0 at 37 °C with constant rotation and 5 ml of cells were collected, pelleted and DNA was extracted using Qiagen DNeasy kit (#69504). Sonication of genomic DNA was performed using Qsonica sonicator (Q8000R) and the sonicated DNA was used to prepare libraries using the NEBNext Ultrall DNA library prep kit (NEB #E7645L). Paired end sequencing of the libraries was performed using NextSeq 500 platform and at least 3 million paired end reads were obtained for each sample. Data from WGS was used as input for the ChIP.

Analysis of ChIP-Seq and WGS data

Sequencing reads for both ChIP and WGS were mapped individually to *B. subtilis* 3610 genome (NZ_CP020102.1)⁵⁵ using CLC Genomics Workbench software (Qiagen). The enrichment at ribosomal RNA locations were eliminated and the number of reads mapped to each base pair in the genome was exported into a .csv file. Data were normalized to the total number of reads and fold enrichment was calculated as the ratio of number of reads at each genome location in ChIP-Seq and WGS (ChIP/input). Analysis was performed and graphs were plotted in 1-kb bins to show enrichment across the entire genome using custom R-scripts. When required, individual peaks were plotted in 10-bp bins across a 4-kb range centered around the peak summit. Detailed protocols and scripts are available upon request.

MEME analysis

A 200-bp sequence surrounding each peak center was extracted using a custom perl script and a fasta file was created. Sequences were subjected to Multiple Em for Motif Elicitation (MEME) v 5.5.2 using parameters (meme sequences.fa -dna -oc. -nostatus -time 14,400 -mod anr -nmotifs 3 -minw 21 -maxw 21 -objfun classic -revcomp -markov_order 0). 21 bp highly enriched motif sequences were

extracted and sequence logo generated by MEME is presented in [Figure 2B](#).

Microscopy

For microscopy, 3 ml of LB broth was inoculated with a single colony and grown at 37 °C. 1 ml of culture at OD₆₀₀ 0.5–0.8 was pelleted and resuspended in 30 µl 1X PBS buffer supplemented with 5 µg/ml FM 4–64 (Invitrogen #T13320) and incubated at room temperature for 2 min in the dark. The cells were washed once with 1 mL of PBS, spun down and resuspended in 30 µl of PBS. 5 µl of sample was spotted onto flat agarose pads (1% agarose in PBS) on slides and covered with a glass coverslip. Fluorescence microscopy was performed with a Nikon 80i microscope with a phase contrast objective Nikon Plan Apo 100X and an Excite 120 metal halide lamp. FM4-64 was visualized with a C-FL HYQ Texas Red Filter Cube (excitation filter 532–587 nm, barrier filter >590 nm). GFP was visualized using a C-FL HYQ FITC Filter Cube (FITC, excitation filter 460–500 nm, barrier filter 515–550 nm). Images were captured with a Photometrics Coolsnap HQ2 camera in black and white using NIS elements software and subsequently false colored and superimposed using Fiji v 2.1.0.⁵⁶

Structure prediction

Multimer structure prediction of SinR·SinR, SlrR·SlrR, SinR·SlrR and YgzD·YgzD was performed using Alphafold2.⁵⁷ For multimer prediction, amino acid sequence for each protein from *Bacillus subtilis* 3610 genome (NZ_CP020102.1) was separated by a colon (:) and prediction was performed using parameters colabfold_batch -num-recycle 20 -amber -templates -model-type alphafold2_multimer_v2. Structures were visualized and shaded using UCSF Chimera v 1.15.⁵⁸

Sequence alignment

Amino acid sequence of YgzD, SinR and SlrR protein from *Bacillus subtilis* 3610 genome (NZ_CP020102.1) were aligned by Clustal Omega v 1.2.4 using default parameters.⁵⁹ Alignment was shaded using Jalview v 2.11.2.7 using a 50% identity threshold.⁶⁰

RNA extraction and analysis

RNA was extracted from *B. subtilis* as described earlier⁸ with slight modifications. *B. subtilis* strains were grown in LB overnight, diluted the next day, and grown to an OD₆₀₀ of ~1.0. 5 ml of culture was flash frozen by adding an equal volume of cold methanol that was pre-chilled at -80 °C. The mixture was centrifuged at 5000g for 10 min at 4 °C, supernatant was discarded, and the pellets were stored at -80 °C. Pellets were resuspended in

800 μ l of hot LETS buffer (10 mM Tris–HCl pH 7.4, 50 mM LiCl, 10 mM EDTA pH 8.0, 1% SDS) pre-incubated at 75 °C. The resuspension was added to a mixture of 650 mg acid-washed glass beads and 600 μ l hot acid-saturated phenol pH 4.6 pre-incubated at 75 °C. The mixture was vortexed for 3 min and 600 μ l of chloroform was added. This mixture was vortexed for 30 sec and centrifuged at 3200g for 10 mins at 4 °C. 600 μ l of the top aqueous layer was added to 800 μ l of hot phenol–chloroform (pH 4.3, 5:1, pre-incubated at 75 °C), vortexed for 3 min, and centrifuged at 3200g for 10 mins at 4 °C. The aqueous phase was collected and added to an equal volume of isopropanol, mixed by inversion, and left at room temperature for 10 mins. The mixture was centrifuged at 4 °C for 25 min at maximum speed. Supernatant was discarded and the pellet was resuspended in 1 ml ice-cold 75% ethanol to wash. The resuspension was centrifuged at maximum speed for 5 min at 4 °C, after which the supernatant was removed and the pellet was dried at room temperature for 10 min. The pellet was resuspended in 20 μ l nuclease-free water at 55 °C for 5 min. 300 μ l of TRIzol was added, and the mixture was vortexed for 15 sec and incubated at room temperature for 5 min. 60 μ l of chloroform was added, and the mixture was inverted for 15 sec and incubated at room temperature for 2 mins. The mixture was centrifuged at 4 °C at maximum speed for 15 min. The above steps, starting with the collection of the aqueous layer and ending with the incubation at 55 °C for 5 min, were repeated. RNA was treated with RNase-free DNase I at 37 °C for 30 min (Invitrogen AM2222) according to manufacturer's instructions. RNA was re-extracted the same way as before starting at the TRIzol step. rRNA depletion and library preparation was performed by Indiana University Center for Genomics and Bioinformatics. Paired end sequencing of the libraries was performed using NextSeq 550 platform and at least 5 million paired end reads were obtained for each sample. Reads were mapped against NCBI 3610 genome (NZ_CP020102.1)⁵⁴ and TPM (Transcript per kilobase million) were calculated using CLC genome browser. TPM values are presented in Table S10.

CRedit authorship contribution statement

Ayushi Mishra: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Abigail E. Jackson:** Writing – review & editing, Methodology, Investigation. **Xindan Wang:** Writing – review & editing, Software, Resources, Methodology, Funding acquisition, Formal analysis. **Daniel B. Kearns:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project

administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This work is supported by National Institutes of Health R01GM141242, R01GM143182, R01AI172822 to XW, and R35GM131783 to DBK. This research is a contribution of the GEMS Biology Integration Institute, funded by the National Science Foundation DBI Biology Integration Institutes Program, Award #2022049 to XW.

Data availability

ChIP-Seq and RNA-Seq data have been submitted to the Gene Expression Omnibus and are available under accession numbers **GSE285064** and **GSE285065** respectively. Protocols and scripts used in this study are available upon request.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank members of the Kearns lab for helpful discussions and members of the Wang lab for assistance with ChIP-seq. AlphaFold was performed using IU carbonate, a part of Indiana University Pervasive Technology Institute that is supported by Lilly Endowment, Inc. We thank Indiana University Center for Genomics and Bioinformatics for their help with next-generation sequencing.

Appendix A. Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.jmb.2025.169123>.

Received 28 January 2025;
Accepted 29 March 2025;
Available online 3 April 2025

Keywords:

motility;
biofilm;
bistability;
heterodimer;
transcription

References

- Grant, G.F., Simon, M.I., (1969). Synthesis of bacterial flagella II. PBS1 transduction of flagella specific markers in *Bacillus subtilis*. *J. Bacteriol.* **99**, 116–124.
- Nishihara, T., Freese, E., (1975). Motility of *Bacillus subtilis* during growth and sporulation. *J. Bacteriol.* **123**, 366–371.
- Kearns, D.B., Losick, R., (2005). Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev.* **19**, 3083–3094.
- Márquez, L.M., Helmann, J.D., Ferrari, E., Parker, H.M., Ordal, G.W., Chamberlin, M.J., (1990). Studies of σ^D -dependent functions in *Bacillus subtilis*. *J. Bacteriol.* **172**, 3435–3443.
- Chen, R., Guttenplan, S.B., Blair, K.M., Kearns, D.B., (2009). Role of the sigmaD-dependent autolysins in *Bacillus subtilis* population heterogeneity. *J. Bacteriol.* **191**, 5775–5784.
- Cozy, L.M., Kearns, D.B., (2010). Gene position in a long operon governs motility development in *Bacillus subtilis*. *Mol. Microbiol.* **76**, 273–285.
- Serizawa, M., Yamamoto, H., Yamaguchi, H., Fujita, Y., Kobayashi, K., Ogasawara, N., Sekiguchi, J., (2004). Systematic analysis of SigD-regulated genes in *Bacillus subtilis* by DNA microarray and Northern blotting analyses. *Gene* **329**, 125–136.
- Cozy, L.M., Phillips, A.M., Calvo, R.A., Bate, A.R., Hsueh, Y.-H., Bonneau, R., Eichenberger, P., Kearns, D.B., (2012). SlrA/SinR/SlrR inhibits motility gene expression upstream of a hypersensitive and hysteretic switch at the level of σ^D in *Bacillus subtilis*. *Mol. Microbiol.* **83**, 1210–1228.
- Norman, T.M., Lord, N.D., Paulsson, J., Losick, R., (2013). Memory and modularity in cell-fate decision making. *Nature* **503**, 481–486.
- Baker, J.D., Kysela, D.T., Zhou, J., Madren, S.M., Wilkens, A.S., Brun, Y.V., Jacobson, S.C., (2016). Programmable, pneumatically actuated microfluidic device with an integrated nanochannel array to track development of individual bacteria. *Anal. Chem.* **88**, 8476–8483.
- Guttenplan, S.B., Shaw, S., Kearns, D.B., (2013). The cell biology of peritrichous flagella in *Bacillus subtilis*. *Mol. Microbiol.* **87**, 211–229.
- Helmann, J.D., Márquez, L.M., Chamberlin, M.J., (1988). Cloning, sequencing, and disruption of the *Bacillus subtilis* σ^{28} gene. *J. Bacteriol.* **170**, 1568–1574.
- Márquez-Magaña, L.M., Chamberlin, M.J., (1994). Characterization of the *sigD* transcriptional unit of *Bacillus subtilis*. *J. Bacteriol.* **176**, 2427–2434.
- West, J.T., Estacio, W., Márquez-Magaña, L., (2000). Relative roles of the *fla/che* P_A, P_{D-3}, and P_{sigD} promoters in regulating motility and *sigD* expression in *Bacillus subtilis*. *J. Bacteriol.* **182**, 4841–4848.
- Mirel, D.B., Lauer, P., Chamberlin, M.J., (1994). Identification of flagellar synthesis regulatory and structural genes in a σ^D -dependent operon of *Bacillus subtilis*. *J. Bacteriol.* **176**, 4492–4500.
- Caramori, T., Barillà, D., Nessi, C., Sacchi, L., Galizzi, A., (1996). Role of FlgM in σ^D -dependent gene expression in *Bacillus subtilis*. *J. Bacteriol.* **178**, 3113–3118.
- Frederick, K., Helmann, J.D., (1996). FlgM is a primary regulator of σ^D activity, and its absence restores motility to a *sinR* mutant. *J. Bacteriol.* **178**, 7010–7013.
- Bertero, M.G., Gonzales, B., Tarricone, C., Cecilian, F., Galizzi, A., (1999). Overproduction and characterization of the *Bacillus subtilis* anti-sigma factor FlgM. *J. Biol. Chem.* **274**, 12103–12107.
- Calvo, R.A., Kearns, D.B., (2015). FlgM is secreted by the flagellar export apparatus of *Bacillus subtilis*. *J. Bacteriol.* **197**, 81–91.
- Dahl, M.K., Msadek, T., Kunst, F., Rapoport, G., (1991). Mutational analysis of the *Bacillus subtilis* DegU regulator and its phosphorylation by the DegS protein kinase. *J. Bacteriol.* **173**, 2539–2547.
- Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G., Dedonder, R., (1990). Signal transduction pathways controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in *degS* and *degU*. *J. Bacteriol.* **172**, 824–834.
- Calvio, C., Celandroni, F., Ghelardi, E., Amati, G., Salvetti, S., Cecilian, F., Galizzi, A., Senesi, S., (2005). Swarming differentiation and swimming motility in *Bacillus subtilis* are controlled by *swrA*, a newly identified dicistronic operon. *J. Bacteriol.* **187**, 5356–5366.
- Tsukahara, K., Ogura, M., (2008). Promoter selectivity of the *Bacillus subtilis* response regulator DegU, a positive regulator of the *fla/che* operon and *sacB*. *BMC Microbiol.* **8**, 8.
- Ogura, M., Tsukahara, K., (2012). SwrA regulates assembly of *Bacillus subtilis* DegU via its interaction with N-terminal domain of DegU. *J. Biochem.* **151**, 643–655.
- Mordini, S., Osera, C., Marini, S., Scavone, F., Bellazzi, R., Galizzi, A., Calvio, C., (2013). The role of SwrA DegU, and P_{D3} in *fla/che* expression in *B. subtilis*. *PLoS One* **12**, e85065.
- Mishra, A., Hughes, A.C., Amon, J.D., Rudner, D.Z., Wang, X., Kearns, D.B., (2024). SwrA-mediated multimerization of DegU and an upstream activation sequence enhance flagellar gene expression in *Bacillus subtilis*. *J. Mol. Biol.* **436**, 168419.
- Lewis, R.J., Brannigan, J.A., Offen, W.A., Smith, I., Wilkinson, A.J., (1998). An evolutionary link between sporulation and prophage induction in the structure of a repressor:anti-repressor complex. *J. Mol. Biol.* **283**, 907–912.
- Scott, D.J., Leejeerajumnean, S., Brannigan, J.A., Lewis, R.J., Wilkinson, A.J., Hoggett, J.G., (1999). Quaternary rearrangement analysed by spectral enhancement: the interaction of a sporulation repressor with its antagonist. *J. Mol. Biol.* **293**, 997–1004.
- Colledge, V.L., Fogg, M.J., Levdikov, V.M., Leech, A., Dodson, E.J., Wilkinson, A.J., (2011). Structure and organization of SinR, the master regulator of biofilm formation in *Bacillus subtilis*. *J. Mol. Biol.* **411**, 597–613.
- Milton, M.E., Draughn, G.L., Bobay, B.G., Stowe, S.D., Olson, A.L., Feldmann, E.A., Thompson, R.J., Myers, K.H., Santoro, M.T., Kearns, D.B., Cavanagh, J., (2020). The solution structures and interaction of SinR and SinI: elucidating the mechanism of action of the master regulator switch for biofilm formation in *Bacillus subtilis*. *J. Mol. Biol.* **432**, 343–357.
- Kearns, D.B., Chu, F., Branda, S.S., Kolter, R., Losick, R., (2005). A master regulator for biofilm formation by *Bacillus subtilis*. *Mol. Microbiol.* **55**, 739–749.
- Blair, K.M., Turner, L., Winkelman, J.T., Berg, H.C., Kearns, D.B., (2008). A molecular clutch disables flagella in the *Bacillus subtilis* biofilm. *Science* **320**, 1636–1638.

33. Guttenplan, S.B., Blair, K.M., Kearns, D.B., (2010). The EpsE flagellar clutch is bifunctional and synergizes with EPS biosynthesis to promote *Bacillus subtilis* biofilm formation. *PLoS Genet.* **6**, e1001243.
34. Chu, F., Kearns, D.B., McLoon, A., Chai, Y., Kolter, R., Losick, R., (2008). A novel regulatory protein governing biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **68**, 1117–1127.
35. Kobayashi, K., (2008). SlrR/SlrA controls the initiation of biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **69**, 1399–1410.
36. Chai, Y., Kolter, R., Losick, R., (2009). Paralogous antirepressors acting on the master regulator for biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **74**, 876–887.
37. Chai, Y., Norman, T., Kolter, R., Losick, R., (2010). An epigenetic switch governing daughter cell separation in *Bacillus subtilis*. *Genes Dev.* **24**, 754–765.
38. Newman, J.A., Rodrigues, C., Lewis, R.J., (2013). Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in *Bacillus subtilis*. *J. Biol. Chem.* **288**, 10766–10778.
39. Chu, F., Kearns, D.B., Branda, S.S., Kolter, R., Losick, R., (2006). Targets of the master regulator of biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **59**, 1216–1228.
40. Swint-Kruse, L., Matthews, K.S., (2009). Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* **12**, 129–137.
41. Lewis, D.E.A., Adhya, S., (2015). Molecular mechanisms of transcriptional initiation at *gal* promoters and the multi-level regulation by GalR, CRP and DNA loop. *Biomolecules* **5**, 2782–2807.
42. Browning, D.F., Butala, M., Busby, S.J.W., (2019). Bacterial transcription factors: regulation by pick “n” mix. *J. Mol. Biol.* **431**, 4067–4077.
43. Sonenshein, A.L., (2007). Control of key metabolic intersections in *Bacillus subtilis*. *Nature Rev. Microbiol.* **5**, 917–927.
44. Ratnayake-Lecamwasam, M., Serror, P., Wong, K.-W., Sonenshein, A.L., (2001). *Bacillus subtilis* CodY represses early-stationary-phage genes by sensing GTP levels. *Genes Dev.* **15**, 1093–1103.
45. Brinsmade, S.R., Kleijn, R.J., Sauer, U., Sonenshein, A.L., (2010). Regulation of CodY activity through modulation of intracellular branched-chain amino acid pools. *J. Bacteriol.* **192**, 6357–6368.
46. Brinsmade, S.R., Alexander, E.L., Livny, J., Stettner, A.I., Segrè, D., Rhee, K.Y., Sonenshein, A.L., (2014). Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *PNAS* **111**, 8227–8232.
47. Belitsky, B.R., Sonenshein, A.L., (2011). Roadblock repression of transcription by *Bacillus subtilis* CodY. *J. Mol. Biol.* **411**, 729–743.
48. Ababneh, Q.O., Herman, J.K., (2015). CodY regulates SigD levels and activity by binding to three sites in the *fla/che* operon. *J. Bacteriol.* **197**, 2999–3006.
49. Belugurov, G.A., Artsimovitch, I., (2015). Regulation of transcript elongation. *Annu. Rev. Microbiol.* **69**, 49–69.
50. Branda, S.S., Gonzalez-Pastor, J.E., Ben-Yehuda, S., Losick, R., Kolter, R., (2001). Fruiting body formation by *Bacillus subtilis*. *PNAS* **98**, 11621–11626.
51. Vlamakis, H., Aguilar, C., Losick, R., Kolter, R., (2008). Control of cell fate by the formation of an architecturally complex bacterial community. *Genes Dev.* **22**, 945–953.
52. López, D., Vlamakis, H., Losick, R., Kolter, R., (2009). Paracrine signaling in a bacterium. *Genes Dev.* **23**, 1631–1638.
53. Subramanian, S., Kearns, D.B., (2019). Functional regulators of bacterial flagella. *Annu. Rev. Microbiol.* **73**, 225–246.
54. Nye, T.M., Schroeder, J.W., Kearns, D.B., Simmons, L.A., (2017). Complete genome sequence of undomesticated *Bacillus subtilis* strain NCIB 3610. *Genome Announc.* **5**, e00364-17.
55. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Cardona, A., (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682.
56. Jumper, J., Evans, R., Pritzel, A., et al., (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
57. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
58. Sievers, F., Wilm, A., Dineen, D.G., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D., (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539.
59. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., (2009). Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191.
60. Bai, U., Mandic-Mulec, I., Smith, I., (1993). SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction. *Genes Dev.* **7**, 139–148.