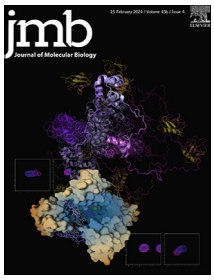
Research Article



SwrA-mediated Multimerization of DegU and an Upstream Activation Sequence Enhance Flagellar Gene Expression in *Bacillus subtilis*

Ayushi Mishra [1], Anna C. Hughes [1], Jeremy D. Amon [2], David Z. Rudner [2], Xindan Wang [1] and Daniel B. Kearns [1]*

*1 - **Department of Biology,** Indiana University, Bloomington, IN 47408, USA*
*2 - **Department of Microbiology,** Harvard Medical School, Boston, MA 02115, USA*

***Correspondence to Daniel B. Kearns:** dbkearns@indiana.edu (D.B. Kearns)*
https://doi.org/10.1016/j.jmb.2023.168419
Edited by Eric CASCALES

Abstract

The earliest genes in bacterial flagellar assembly are activated by narrowly-conserved proteins called master regulators that often act as heteromeric complexes. A complex of SwrA and the response-regulator transcription factor DegU is thought to form the master flagellar regulator in *Bacillus subtilis* but how the two proteins co-operate to activate gene expression is poorly-understood. Here we find using ChIP-Seq that SwrA interacts with a subset of DegU binding sites in the chromosome and does so in a DegU-dependent manner. Using this information, we identify a DegU-specific inverted repeat DNA sequence in the P_{flache} promoter region and show that SwrA synergizes with DegU phosphorylation to increase binding affinity. We further demonstrate that the SwrA/DegU footprint extends from the DegU binding site towards the promoter, likely through SwrA-induced DegU multimerization. The location of the DegU inverted repeat was critical and moving the binding site closer to the promoter impaired transcription by disrupting a previously-unrecognized upstream activation sequence (UAS). Thus, the SwrA-DegU heteromeric complex likely enables both remote binding and interaction between the activator and RNA polymerase. Small co-activator proteins like SwrA may allow selective activation of subsets of genes where activator multimerization is needed. Why some promoters require activator multimerization and some require UAS sequences is unknown.

Importance

In bacteria, promoter sequences are complex. The sigma subunit of RNA polymerase recognizes specific DNA sequences that determine where gene transcription begins, but some promoters also have upstream sequences to increase promoter activity. Some upstream sequences are recognized by the alpha subunit of RNA polymerase (e.g. UP elements) while others recruit transcription factors to increase promoter affinity or enhance open-complex formation. Here we show that the promoter for the 32 gene flagellar operon in *Bacillus subtilis* requires an upstream activation sequence (UAS), and still farther upstream, an inverted repeat sequence bound by the response regulator DegU. Our evidence suggests that DegU binds but does not activate flagellar gene expression unless another protein SwrA induces DegU oligomerization. Heteromeric activator complexes are known but poorly-understood in bacteria and we speculate they may be needed to activate gene expression in the context of intervening cis-elements.

Introduction

Bacterial flagella are complex, multi-subunit, trans-envelope machines that propel cells to swim in liquid or swarm over solid surfaces. Flagella are composed of dozens of different subunits, expressed in a series of one to four hierarchical tiers roughly corresponding to the order in which the subunits are assembled.[1–4] The top of the hierarchy is often a protein complex called a master regulator that directly enhances the expression of the earliest genes in flagellar synthesis that encode the flagellar basal body. Mutation of the master regulator tends to impair or abolish motility while overexpression of the master regulator can lead to hyper-flagellation and, in some systems, promote swarming motility atop solid surfaces.[5] Master regulators bind upstream of one or more promoters in the flagellar regulon and often occur in heterodimeric pairs.[6–8] While clearly important for motility gene expression, these master regulators are among the most species-specific and the least-studied components of the hierarchy.

In *Bacillus subtilis,* one part of the master flagellar regulator is SwrA, a small (117 amino acid), positively-charged protein with no predicted motifs that is narrowly-conserved in a closely-related subset of species within the genus *Bacillus.*[9,10] SwrA was first discovered as a protein that was required for swarming over surfaces but not swimming in liquid.[9] Suppressors that restored swarming in the absence of SwrA improved the consensus of the vegetative SigA-dependent P_{flache} promoter that directs transcription of the 32-gene long *fla/che* operon encoding flagellar basal body components.[11–13] Consistent with P_{flache} being the primary target in swarming motility, SwrA was shown to activate expression of the P_{flache} promoter and basal body number over a relatively narrow (~4-fold) range.[13,14] Finally, whereas moderate SwrA levels support a high frequency of swimming cells in liquid, SwrA levels increase on a surface and enhance P_{flache} expression necessary to elevate flagellar synthesis for swarming.[13,15–17] The mechanism by which SwrA enhances P_{flache} promoter activity is poorly-understood but SwrA has been shown interact with the DNA binding protein DegU.[14,18–20]

DegU is the other part of the master flagellar regulator in *B. subtilis* and is a transcription factor in the response regulator family of proteins. DegU confers pleiotropic phenotypes, either when mutated or hyperactivated, that include defects in genetic competence, exoprotease production, exopolymer production, biofilm formation, and motility.[21–30] A consensus binding sequence for DegU has not been determined and the mechanism by which DegU differentially regulates the wide variety of targets under its control is poorly-understood.[31,32] DegU is complicated in that it binds DNA in both phosphorylated and unphosporylated forms and may bind targets differently in depending on its phosphorylation state.[14,25,33–35] Regulation is also complex as DegU can be phosphorylated either by the soluble histidine kinase DegS[36,37] or by the small metabolite acetyl phosphate,[38] and two poorly-understood proteins DegQ and DegR promote the phosphorylated state.[25,39] Finally, SwrA binds to the DegU receiver domain and may alter DegU DNA binding activity.[19,20] Consistent with being a modulator of DegU function, cells lacking SwrA have been reported to have phenotypes beyond defects in motility, many of which overlap with cells lacking DegU.[40–42]

Here we characterized SwrA proximity to the chromosome using chromatin immunoprecipitation coupled to deep sequencing (ChIP-seq) and found that SwrA was enriched at a subset of DegU binding sites. Bioinformatic analysis predicted an asymmetric sequence common to SwrA and DegU ChIP targets, and asymmetry may have been due to a variable length AT-rich spacer between poorly-conserved half-sites. SwrA enhanced DegU binding at a variety of the target sites *in vitro* with the highest affinity displayed at P_{flache}, consistent with genetic results suggesting that it was the primary biological SwrA target.[13] Moreover, the P_{flache} promoter had a perfect 5-8-5 inverted repeat similar to and overlapping with the predicted consensus, and the sequence was shown to be required for both DegU-binding and promoter activation. SwrA synergized with phosphorylation to increase DegU's DNA binding affinity and expanded DegU binding towards the P_{flache} promoter likely by inducing DegU oligomerization. The location of the DegU binding site was critical and moving the repeat closer to the sigma binding site abolished promoter activity by disrupting a previously-unrecognized, cis-acting upstream activation sequence (UAS). We suggest that DegU oligomerization induced by SwrA is necessary to allow both the remote binding of DegU and interaction with RNA polymerase at P_{flache} and perhaps other promoters.

Results

SwrA interacts with DNA indirectly. SwrA is part of the master regulator of motility in *B. subtilis* as it activates the P_{flache} promoter that controls 32 genes involved in flagellar assembly and chemotaxis.[13,14] Other targets of SwrA have been reported but the extent of the SwrA regulon, and the mechanism of SwrA-mediated transcriptional activation are poorly-understood.[13,19,20,41,42] To investigate whether SwrA is found in proximity to DNA *in vivo*, we performed chromatin immunoprecipitation coupled to deep sequencing (ChIP-Seq) on wild type and *swrA* mutant cells. Mid-log phase cells were treated with formaldehyde, and after lysis and DNA fragmentation, SwrA was immunoprecipitated with antibodies raised against the full-length protein. After reversing the crosslinks, the DNA

associated with SwrA was subjected to next generation sequencing. Chromatin immuno-enrichment was calculated as the ratio of ChIP-Seq signal to genomic DNA plotted as peaks in 1 kb windows that spanned the entire genome. SwrA candidate binding sites were defined as peaks that were enriched in wild type replicates but not in the *swrA* mutant control (Figure 1A).

SwrA was enriched at thirty-four genomic locations and all but two were located in intergenic regions consistent with possible promoters (Table S1). To investigate whether SwrA affected

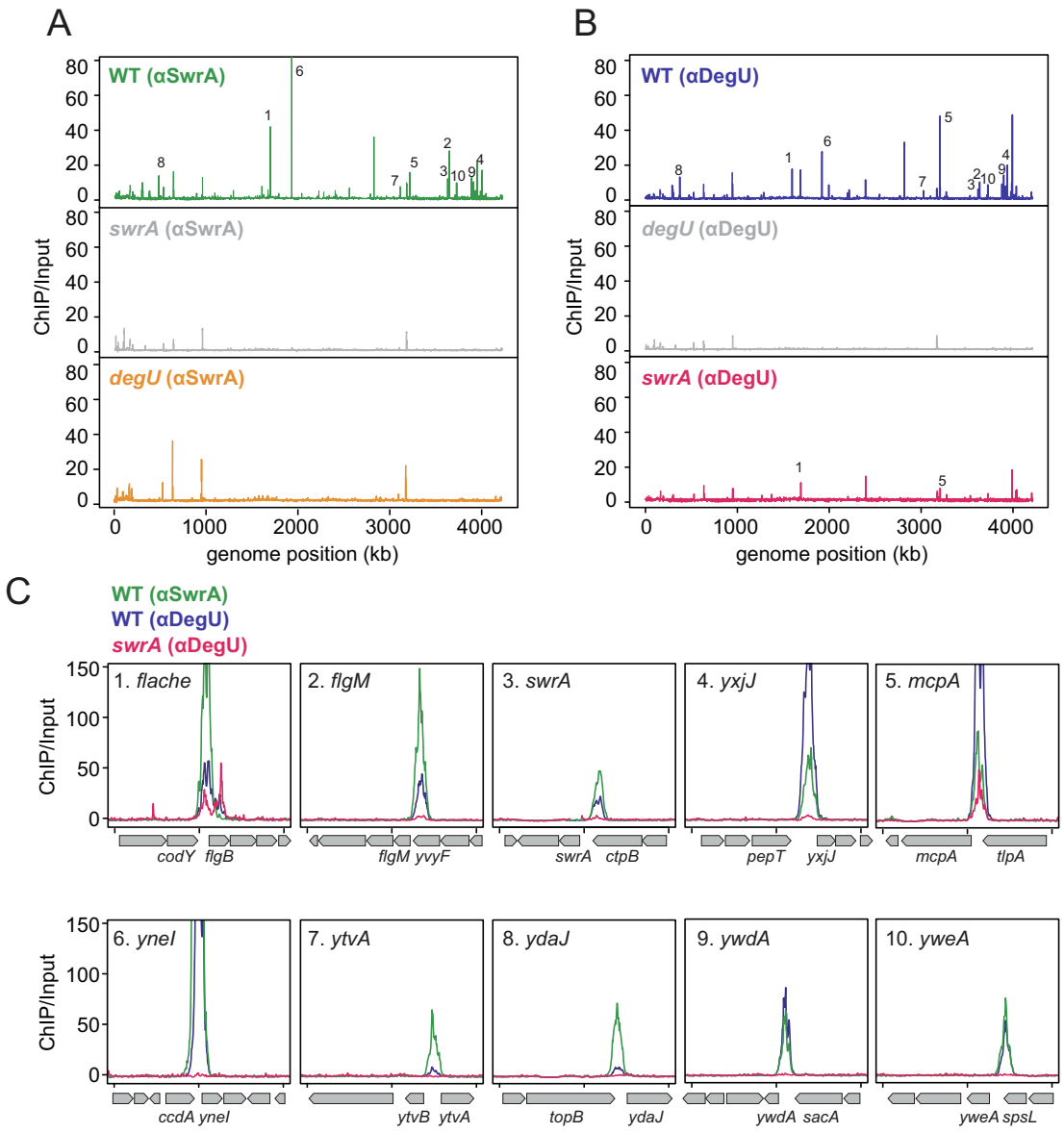


Figure 1. SwrA enriches a subset of DegU-enriched promoters. A) ChIP-Seq analysis using a primary antibody to SwrA (αSwrA). The number of sequencing reads were normalized by the total number of reads for each sample. The fold of enrichment (ChIP/Input) were calculated and plotted in 1 kb bins. The following strains were used to generate this panel: WT (3610), *swrA* (DS2415), and *degU* (DS3649). Peaks of particular interest were numbered 1–10. B) ChIP-Seq analysis using a primary antibody to DegU (αDegU). Data were processed in the same way as in panel A. The following strains were used to generate this panel: WT (3610), *degU* (DS3649) and *swrA* (DS2415). Peaks of particular interest were numbered 1–10 to match the same peaks in panel A. C) ChIP-Seq analysis of peaks 1–10 from panels A and B enlarged to show detail over a 4 kb range. Each panel is named according to the promoter region for the likely gene downstream. Gene size and identity is indicated below the X-axis. The X-axis is marked in 2 kb intervals. Green lines indicated WT ChIP-Seq data using αSwrA antibody, blue lines indicate WT ChIP-Seq data using αDegU antibody, and magenta lines indicate *swrA* mutant ChIP-seq data using αDegU antibody.

transcription of genes near to which it bound, 13 promoter regions containing strongly-enriched peaks (P_{flache}, P_{flgM}, P_{swrA}, P_{yxjJ}, P_{mcpA}, P_{ycdC}, P_{ynel}, P_{ytvA}, P_{ydaJ}, P_{ywdA}, P_{yweA}, P_{tlpA}, and P_{sacX}) were cloned upstream of the *lacZ* gene encoding β-galactosidase, and β-galactosidase activity was measured in various genetic backgrounds. As anticipated, expression of P_{flache}-*lacZ* decreased relative to wild type in cells lacking SwrA, and increased when SwrA was overexpressed, consistent with previous reports (Figure 2A).[13,14] The remaining reporters responded to the presence and absence of SwrA in a variety of ways. Expression from four promoters, P_{flgM}, P_{swrA}, P_{yxjJ}, and P_{mcpA} was reduced ~2-fold in a *swrA* deletion and increased ~2-fold when SwrA was artificially overexpressed in a manner that was similar to P_{flache} (Figure 2A). Five promoters, P_{ynel}, P_{ytvA}, P_{ydaJ}, P_{ywdA}, and P_{yweA} produced low but detectable levels of activity that was not altered by either mutation or overexpression of SwrA (Figure 2A). Finally, three promoters, P_{tlpA}, P_{ycdC} and P_{sacX}, produced no activity above background (<2 Miller units, MU) in any strain and were omitted from the study. We conclude that while ChIP-Seq indicated SwrA-enrichment of several promoter regions, enrichment did not necessarily reflect an effect of SwrA on reporter expression. We further conclude that SwrA either directly or indirectly activated the P_{flache}, P_{flgM}, P_{swrA}, P_{yxjJ}, and P_{mcpA} promoters.

One way in which SwrA could enrich target promoter regions is by binding directly to DNA. To determine whether SwrA bound DNA directly, electrophoretic mobility shift assays (EMSA) were performed on seven different promoter regions. The P_{flache} promoter was chosen as a known SwrA target and P_{flgM}, P_{yxjJ}, P_{swrA}, and P_{ynel} were added as candidates from the present study. The promoter P_{comK} expressing the gene for the master activator of competence gene expression ComK, and the promoter P_{hag} expressing the gene for the flagellar filament protein Hag were included as controls that were not pulled down in the ChIP-Seq experiment. The seven different promoter fragments were PCR amplified, radiolabeled, incubated with purified SwrA protein at a range of protein concentrations, resolved by native gel electrophoresis and analyzed by phosphorimager. In each case, addition of SwrA failed to alter migration of the radiolabeled DNA fragment except at the highest concentration of SwrA added (1 μM) (Figure S1) and there appeared to be little to no difference between targets predicted by ChIP-Seq and the not-predicted controls. These data, along with previous reports showing that SwrA did not bind to the P_{flache} or P_{ycdA} promoters,[14,20] argue that SwrA is likely not sufficient for DNA binding *in vitro*.

SwrA-DNA interaction is DegU-dependent. Previous reports suggests that SwrA interacts with the P_{flache} promoter indirectly by interacting with the response regulator DegU.[19,20] Consistent with a requirement of both proteins, cells lacking either SwrA or DegU are defective in swarming motility even when the other protein was artificially overexpressed (Figure S2).[20,25,26,43,9] To test whether DegU was required for SwrA's association with DNA *in vivo*, we repeated the SwrA ChIP-Seq in a *degU* mutant strain. In the absence of DegU, all SwrA-dependent peaks were abolished (Figure 1A). Furthermore, the *degU* mutation reduced the expression of the SwrA target promoters P_{flache}, P_{flgM}, P_{swrA}, P_{yxjJ}, and P_{mcpA} to levels similar to that observed in a *swrA* mutant, and SwrA overexpression failed to activate these promoters (Figure 2B). Finally, cells doubly mutated for both SwrA and DegU produced expression levels from each reporter comparable to the DegU mutant alone (Figure 2C). We conclude that SwrA association with DNA and transcriptional activation of the SwrA-responsive promoters was entirely dependent on the presence of DegU.

To investigate whether SwrA alters DegU DNA binding specificity *in vivo*, we performed ChIP-seq using a DegU antibody. DegU was enriched at 46 locations, many of which were upstream of genes involved in motility, competence, biofilm formation as well as genes of unknown function (Figure 1B; Table S1).[44] All 34 SwrA-enriched regions were also enriched in the DegU ChIP-seq (Figure 1C, Figure S3A). Furthermore, in cells lacking SwrA, there was a general reduction in sites enriched by DegU (Figure 1B,1C). Peaks were sorted into three different DegU-ChIP classes depending on the effect of SwrA (Table S1). Class I targets (61%) were SwrA-dependent as they were abolished in the absence of SwrA. Class II targets (11%) were SwrA-enhanced as they were reduced but not abolished when SwrA was absent. Class III targets (28%) were SwrA-independent as they appeared unaffected by its absence. We conclude that three-quarters of the DegU-enriched promoters were either enhanced by, or fully dependent on, the presence of SwrA. Therefore, SwrA potentiates binding of DegU to a subset of promoters in its regulon.

In an effort to determine a consensus binding site for DegU, 200 base pair fragments surrounding each ChIP-seq peak center were compiled and subjected to MEME sequence pattern analysis.[45] Combining sequences from all DegU peaks indicated an enriched sequence that did not contain a repeat element, contrary to what one might expect for a response regulator DNA binding sequence (Figure S4). Separate analysis of the class I, class II and class III DegU ChIP peaks, provided similarly-enriched sequences (Figure S4). We wondered whether our data set might be incomplete as indicated by the absence of peaks located near the characterized DegU-regulated promoters: P_{aprE} and P_{pgsB} directing alkaline protease and poly-γ-glutamate synthesis, respectively (Fig-

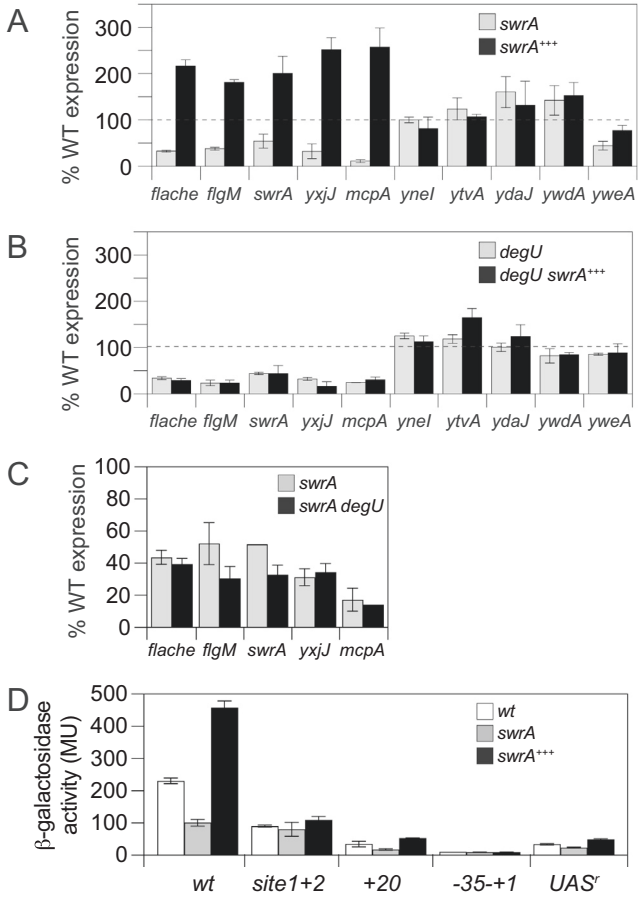


Figure 2. SwrA/DegU activate expression from a subset of enriched promoters. A) β-galactosidase activity from the indicated promoter region fused to the *lacZ* gene. Light gray bars indicate expression in a *swrA* mutant and black bars indicate expression when *swrA* was overexpressed from an IPTG-inducible promoter in meroploid. Activity was normalized to 100% wild type expression (dashed line). Error bars are the standard deviation of three replicates. The following strains were used to generate this panel: P_{flache}-*lacZ* (DK4730, DK4918), P_{flgM}-*lacZ* (DK4870, DK4919), P_{swrA}-*lacZ* (DK6624, DK6625), P_{yxjJ}-*lacZ* (DK4733, DK4929), P_{mcpA}-*lacZ* (DK4732, DK4928), P_{yneI}-*lacZ* (DK4731, DK4927), P_{ytvA}-*lacZ* (DK6500, DK6051), P_{ydaJ}-*lacZ* (DK6473, DK6477), P_{ywdA}-*lacZ* (DK6474, DK6478), and P_{yweA}-*lacZ* (DK6476, DK6480). B) β-galactosidase activity from the indicated promoter region fused to the *lacZ* gene. Light gray bars indicate expression in a *degU* mutant and black bars indicate expression when *swrA* was overexpressed from an IPTG-inducible promoter in meroploid. Activity was normalized to 100% wild type expression (dashed line). Error bars are the standard deviation of three replicates. The following strains were used to generate this panel: P_{flache}-*lacZ* (DK4734, DK4979), P_{flgM}-*lacZ* (DS3658, DK4983), P_{swrA}-*lacZ* (DK6626, DK6650), P_{yxjJ}-*lacZ* (DK4737, DK4982), P_{mcpA}-*lacZ* (DK4736, DK4981), P_{yneI}-*lacZ* (DK4735, DK4980), P_{ytvA}-*lacZ* (DK6502, DK6550), P_{ydaJ}-*lacZ* (DK6481, DK6509), P_{ywdA}-*lacZ* (DK6482, DK6524), and P_{yweA}-*lacZ* (DK6484, DK6510). C) β-galactosidase activity from the indicated promoter region fused to the *lacZ* gene. Light gray bars indicate expression in a *swrA* mutant and black bars indicate a *swrA degU* double mutant. Activity was normalized to 100% wild type expression. Error bars are the standard deviation of three replicates. The following strains were used to generate this panel: P_{flache}-*lacZ* (DK4730, DK7502), P_{flgM}-*lacZ* (DK4870, DB1032), P_{swrA}-*lacZ* (DK6624, DB1013), P_{yxjJ}-*lacZ* (DK4733, DB1015) and P_{mcpA}-*lacZ* (DK4732, DB1014). D) β-galactosidase activity from the indicated P_{flache} promoter mutants fused to the *lacZ* gene. White bars indicate expression in wildtype, light gray bars indicate expression in a *swrA* mutant and black bars indicate expression when *swrA* was overexpressed from an IPTG-inducible promoter in meroploid. Activity is represented in Miller units (MU). Error bars are the standard deviation of three replicates. The following strains were used to generate this panel: wt (DK5183, DK4730, DK4918), site1+2 (DB534, DB1074, DB573), +20 (DB556, DB570, DB578), – 35 +1(DK7928, DK7970, DK7976) and UAS^r (DB1012, DB1040, DB1053). Raw data included in Table S5.

ure S3B).[27,41,46–48] Accordingly, we performed a DegU ChIP-seq by immuno-precipitating DegU from a strain overexpressing the small phosphor-enhancer protein DegQ and a strain expressing the hyper-active DegU allele, DegU^{hy32}, in place of the wild type.[21,25,33] Both the DegQ overproduction

(Figure S5A, Table S2) and DegU^{hy32} (Figure S5B, Table S3) strains identified additional peaks that were enhanced by SwrA, but MEME analysis of each dataset still produced an asymmetric DegU target sequence similar to that of the wild type (Figure S4). Moreover, sequence analysis did not indicate how SwrA was differentiating a subset of the DegU peaks for enrichment, as there appeared to be no sequence in the DNA that was correlated with the presence of SwrA.

SwrA increases DegU affinity for DNA and expands the DegU binding site. To better understand the mechanism of SwrA-mediated DegU activation, DegU EMSAs were conducted on the same series of promoters previously used to test for direct interaction by SwrA. DegU bound poorly to each promoter but an electrophoretic mobility shift was observed with P_{flache} when DegU was phosphorylated by ATP and its cognate kinase, DegS (Figure S1).[14,25] Thus, the P_{flache} promoter seemed to contain the highest affinity binding site of those tested, and we note that previously studied promoters like P_{flgM} and P_{comK}, require higher concentrations of DegU than used here.[34,49] Next, the concentration of either DegU or DegU-P was held constant and increasing amounts of SwrA were added to the reaction. Consistent with previous reports, the presence of SwrA caused a supershift of the P_{flache} promoter and did so at lower concentrations when DegU was phosphorylated (Figure 3).[20] Moreover, the presence of SwrA induced a shift of all promoters except the non-predicted target P_{hag}, suggesting that SwrA enhanced the affinity of DegU-P for its targets in general.

To explore the effect of SwrA on DegU DNA binding activity, the concentration of SwrA was held constant, and EMSA experiments were conducted in triplicate on P_{flache} with increasing amounts of DegU and DegU-P. Addition of SwrA caused diffuse supershifted bands making it difficult to quantify the intensity of the bound state by densitometry. Therefore, we calculated the fraction of the unbound state and subtracted the value from 1. Phosphorylation of DegU increased DNA binding affinity by 10-fold (Figure 4A,B). SwrA increased the binding affinity of DegU (Figure 4A) and DegU-P (Figure 4B) by approximately 10-fold and 4-fold respectively, suggesting that SwrA affinity enhancement was phosphorylation-independent (Figure 4). High levels of SwrA caused smearing of the band signal even for non-target promoters like P_{hag} but did so with much lower affinity than observed for P_{flache} (Figure 4C). While ChIP-seq is non-quantitative, we note SwrA-enhancement of DegU DNA binding is consistent with the *in vivo* global reduction in DegU enrichment in the absence of SwrA (Figure 1B). We conclude that SwrA interacts with DegU and synergizes with phosphorylation to enhance DNA binding.

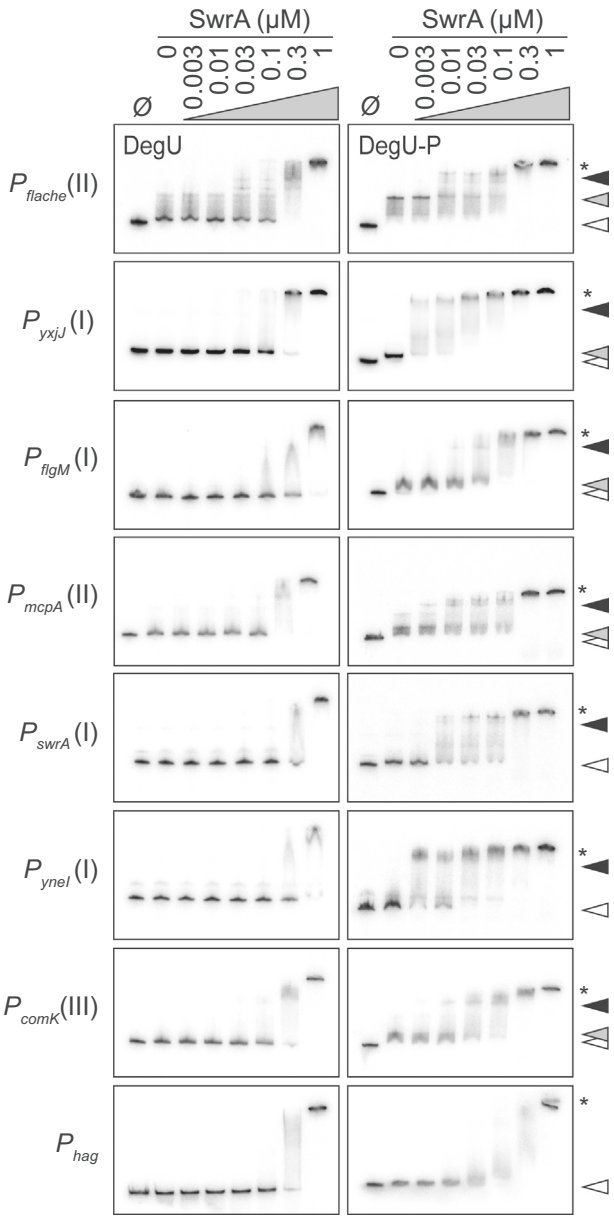


Figure 3. SwrA causes a supershift in DegU DNA binding. Electrophoretic mobility shift assays (EMSA) with the indicated radiolabeled promoter fragment of different classes (indicated in parentheses, Table S1) and protein. Ø indicates that no protein was added. Left panels, an increasing about of GST-SwrA was added to a constant 0.3 μM of DegU-His$_6$. Right panels, an increasing amount of GST-SwrA was added to a constant 0.3 μM of DegU-P-His$_6$ phosphorylated by DegS-His$_6$ and ATP. Gray carets indicate the position of DNA shifted by the presence of either DegU-His$_6$ or DegU-P-His$_6$ alone. Black carets indicated the position of DNA supershifted by the presence of either DegU-His$_6$ or DegU-P-His$_6$ and GST-SwrA. * indicates aggregation.

To further explore the effect of SwrA on DegU-P, DNAse I protection assays were conducted on

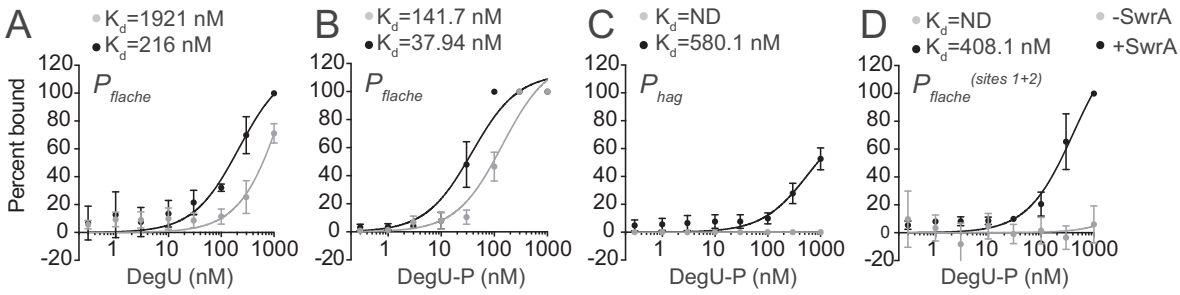


Figure 4. SwrA increases DegU DNA-binding affinity. Densitometry analysis of electrophoretic mobility shift assays (EMSA) (Figure S12) with the indicated radiolabeled promoter fragment and various concentrations of either DegU-His$_6$ or DegU-P-His$_6$ phosphorylated by ATP and DegS-His$_6$. Gray circles are EMSA densitometry in the presence of the indicated form of DegU-His$_6$. Black circles are EMSA densitometry in the presence of the indicated form of DegU-His$_6$ and 0.1 μM GST-SwrA. Each point is the average of three replicates and error bars are standard deviations. A) Densitometry values of radiolabeled P_{flache} promoter sequence shifted by DegU-His$_6$; B) densitometry values of radiolabeled P_{flache} promoter sequence shifted by DegU-P-His$_6$; C) densitometry values of radiolabeled P_{hag} promoter sequence shifted by DegU-P-His$_6$; D) densitometry values of radiolabeled P_{flache} promoter sequence doubly mutated at site 1 and site 2 shifted by DegU-P-His$_6$.

target promoters (Figure 5A,B). DegU-P protected a region of the P_{flache} promoter 40 base pairs (bp) upstream of the sigma -35 box, consistent with being a SwrA-enhanced class II promoter (Figure 5A, Table S1). Centered within the protected region was a perfect 5-8-5 inverted repeat of CTAGG separated by an intervening 8 base pairs (Figure 5C). Addition of SwrA to the reaction expanded the protection to both the left and right of the repeat thereby increasing the total protected area to approximately 70 bp. DegU-P alone did not provide DNAse I protection of P_{yxjJ} at the concentration used, consistent with a SwrA-dependent class I promoter, and addition of SwrA induced protection again with a width of approximately 70 bp (Figure 5B). While the footprinting assay was not optimized for other targets, large regions of protection of 70 bp or more in the presence of both DegU-P and SwrA was observed (Figure S6). We conclude that SwrA alters the way in which DegU-P binds DNA and creates a wide footprint of DNase I protection at multiple promoters.

The consensus sequence to which DegU-P binds is poorly-understood and we focused our attention on the putative 5-8-5 inverted repeat within the highest-affinity target, the P_{flache} promoter. To test whether the inverted repeat element was important for SwrA/DegU-dependent activation, two bases were changed in each half-site of the repeat separately and together at the native site in the chromosome (Figure 6A). Mutation of either the promoter-distal site (site 1) or the promoter-proximal site (site 2) had little effect on swarming motility but mutation of both sites simultaneously caused a severe defect (Figure 6B). Overexpression of SwrA shortened the lag period of the single mutants and restored partial swarming to the double mutant. Moreover, EMSA indicated that mutation of both sites simultaneously abolished and dramatically reduced binding of DegU-P in the absence and presence of SwrA respectively (Figure 4D). Finally, mutation of both sites in the P_{flache} promoter fused to the *lacZ* gene reduced β-galactosidase activity to levels comparable with a SwrA mutant and could not be increased by SwrA overexpression (Figure 2D). Consistent with an important regulatory element, the sequence upstream of the P_{flache} SigA binding boxes, including the putative 5-8-5 repeat, was highly-conserved in a wide variety of *B. subtilis* relatives that also encode SwrA and DegU (Figure 5C) and conservation degraded farther upstream (Figure S7). We conclude that the conserved inverted repeat protected in the DegU-P DNase I protection assay is required for SwrA/DegU-dependent activation of P_{flache}.

We noted that the 5-8-5 inverted repeat was part of the consensus sequence predicted by MEME analysis but the actual output emphasized the AT-rich spacer between the repeats rather than the repeats themselves (Figure S4). We wondered why the spacer sequence was so highly-conserved and homopolymer replacements were made to test its importance. When six of the spacer residues were changed to either all adenines (A) or all thymines (T), no change on swarming motility was observed and overexpression of SwrA shortened the lag period like wild type (Figure 6C). When the residues were changed to all guanines (G), swarming was severely impaired but in a way that could be rescued by SwrA overexpression. When the residues were changed to all cytosines (C) swarming motility was abolished and overexpression of SwrA had little effect. We conclude that the nature of the sequence between

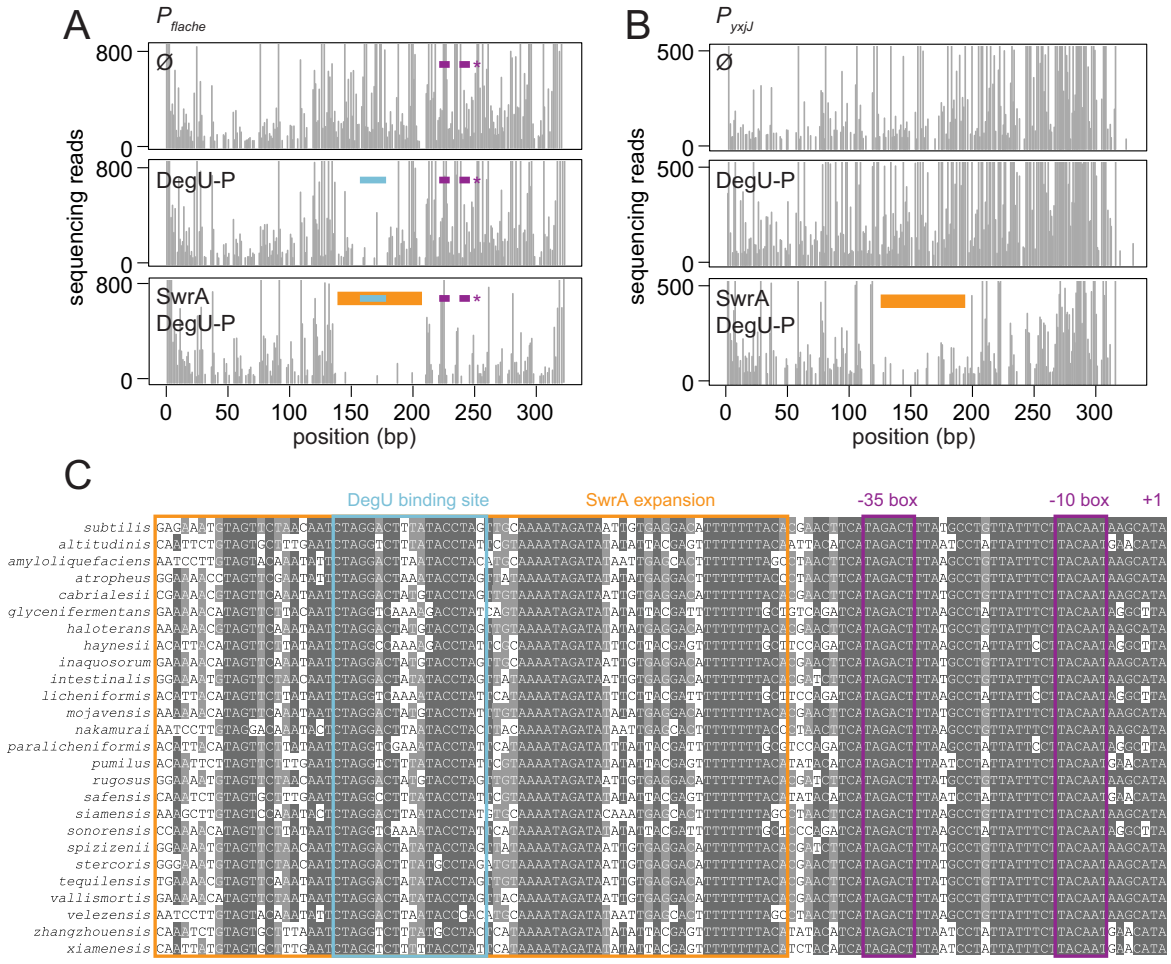


Figure 5. SwrA increases the length of DNA protected by DegU-P. A) DNaseI sequencing footprint analysis of the P_{flache} promoter. 300 bp fragments of dsDNA were fluorescently labeled on the forward strand, the indicated protein was added, followed by partial digestion with DNase I and sequencing of the digested fragments. Top panel, 1 μM BSA added; middle panel 3 μM of DegU-P-His$_6$ added; bottom panel 3 μM of DegU-P-His$_6$ and 1μ M of GST-SwrA added. Blue bars indicate the location of the 5-8-5 inverted repeat; orange bars indicate the expanded region of SwrA/DegU-P protection; purple bars indicate the -35 SigA box, the -10 SigA box, and the +1 transcriptional start sites of the P_{flache} promoter are marked with an asterisk. Peaks indicate the number of sequence reads terminating at that location on one strand. B) DNaseI sequencing footprint analysis of the P_{yxjJ} promoter. Top panel, 1 μM BSA added; middle panel 3 μM of DegU-P-His$_6$ added; bottom panel 3 μM of DegU-P-His$_6$ and 1 μM of GST-SwrA added. Peaks indicate the number of sequence reads terminating at that location. The location of the promoter and transcriptional start site for P_{yxjJ} is unknown. C) Alignment of the P_{flache} promoter fragment. Blue boxes highlight the location of the 5-8-5 inverted repeat (repeat sequences in bold); Orange box highlight the expanded region of SwrA/DegU-P protection; purple highlights the -35 and -10 SigA boxes of the P_{flache} promoter and +1 indicates the transcriptional start site.

the inverted repeats is relevant insofar as only an A-T rich sequence is tolerated for full functionality. We hypothesize that the variable length and composition of the homopolymer tract, the ability to mutate inverted repeat half sites while maintaining functionality, and SwrA-induced DegU affinity enhancement at weak sites, likely explains why the 5-8-5 repeat was difficult to detect both in footprint and MEME analysis of other promoters.

DegU-P bound to the P_{flache} 5-8-5 repeat but did not activate expression until the presence of SwrA expanded the region of protection closer to the promoter. We wondered whether the effectiveness of the 5-8-5 repeat was dependent on the distance from the sigma -35 box. To test position-dependence, the repeat was moved closer to (by deletion) and farther away from the promoter (by insertion of randomized sequence) by full helical turns of the DNA (10 bp) to maintain register with the SigA boxes (Figure 6A). Movement of the repeat 20 and 10 bases father away (-20 and -10) abolished and reduced swarming respectively, and SwrA overexpression enhanced swarming of the -10 reposition (Figure 6D). Moving the repeat

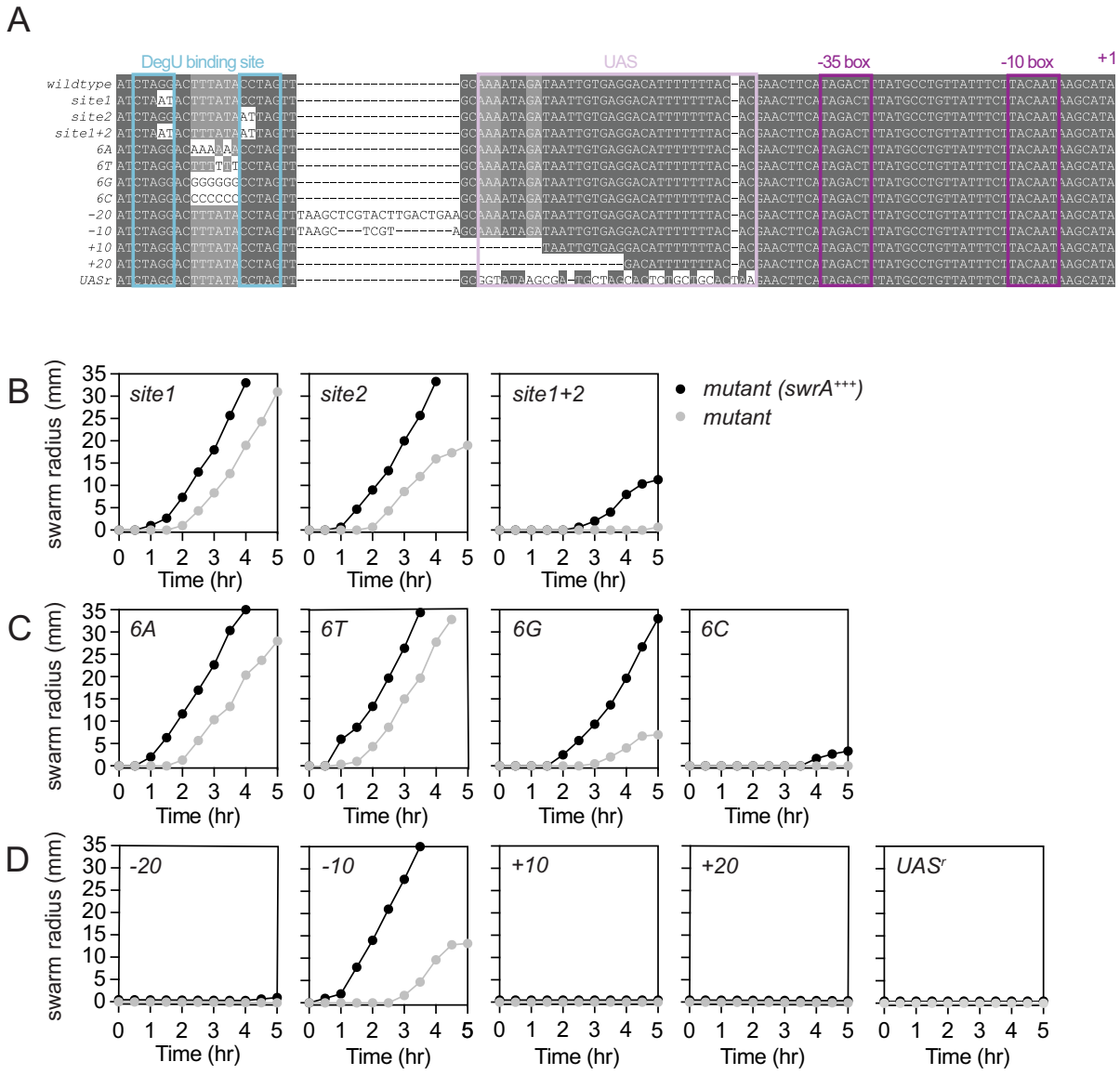


Figure 6. The distance between the DegU binding site and the P_{flache} promoter is important for activation. A) Alignment of P_{flache} promoter mutations used in the assay below. Blue boxes highlight the location of the 5-8-5 inverted repeat; Pink box highlights the upstream activating sequence; purple highlights that -35 and -10 SigA boxes of the P_{flache} promoter and +1 indicates the transcriptional start site. B) Quantitative swarm expansion assays of strains mutated for the DegU binding site inverted repeat sequences (gray circles), and strains containing the indicated mutations and SwrA overexpressed by IPTG induction (*swrA*$^{+++}$, black circles). The following strains were used to generate this panel: *site1* (DB164, DB198), *site 2* (DB102, DB197), and *site 1+2* double mutant (DB49, DB58). C) Quantitative swarm expansion assays of strains with the indicated sequence between the DegU binding site repeat elements (gray circles), and strains containing the indicated sequence and SwrA overexpressed by IPTG induction (*swrA*$^{+++}$, black circles). The following strains were used to generate this panel: *6A* (DB48, DB196), *6 T* (DB47, DB228), *6G* (DB165, DB199), and *6C* (DB166, DB200). D) Quantitative swarm expansion assays of strains in which the 5-8-5 DegU binding site motif was moved farther away (-20 or −10 bp, respectively) or towards (+10 or +20 bp, respectively) the P_{flache} promoter (gray circles), and strains containing the indicated relocation and SwrA overexpressed by IPTG induction (*swrA*$^{+++}$, black circles) The following strains were used to generate this panel: *−20* (DB488, DB524), *−10* (DB487, DB523), *+10* (DB195, DB220), *+20* (DB486, DB522) and UAS^{r} (DB1020, DB1071). Each data point is the average of three replicates.

10 and 20 bases closer to the promoter (+10 and +20), abolished swarming in a manner that could not be rescued by SwrA overexpression (Figure 6D). We conclude that the position of the DegU binding site is important and moving it from its proper location impairs swarming motility.

The strain containing the DegU binding site shifted 20 base pairs closer to the promoter was unusual in that it exhibited mucoid colonies on plates and was difficult to pellet in liquid. Mucoid colonies and loose pellets are indicative of a defect in the alternative sigma factor SigD encoded at the 3′ end of the *fla/che* operon, perhaps suggesting a loss of P_{flache} activity beyond that observed in cells mutated for SwrA/DegU.[11,12,50–52] To measure SigD activity, a reporter in which the SigD-dependent P_{hag}-promoter was transcriptionally fused to the gene encoding green fluorescent protein (*gfp*) and introduced at an ectopic site in a variety of strains. Wild type exhibited a high frequency of brightly fluorescent P_{hag} ON cells with rare dark P_{hag} OFF cells (Figure 7). Mutation of SwrA, DegU, or both proteins simultaneously decreased the frequency of P_{hag} ON cells, with OFF cells growing as long chains (Figure 7). Strains in which the DegU binding site was moved 20 or 10 base pairs farther away (-20 and -10) or 10 base pairs closer (+10) to the promoter produced populations similar to that of either the SwrA or DegU mutant, consistent with the inability of the protein complex to activate the P_{flache} promoter properly (Figure 7). The strain that moved the DegU binding site 20 base pairs closer (+20) however, abolished ON cell production with ubiquitous cell chains, resembling cells that lack SigD (Figure 7). Moreover, DegU did not become inhibitory in the +20 relocation strain as mutation of either SwrA or DegU was insufficient to restore GFP expression (Figure S8). We infer that another cis element important for P_{flache} expression resides between the DegU binding site and the SigA -35 box.

To further explore the nature of the putative cis element, we cloned the promoter region of the +20 relocation upstream of the *lacZ* gene and inserted the construct at an ectopic site. The +20 relocation reporter reduced β-galactosidase activity lower than observed for a wild type promoter or a promoter in which the DegU binding site had been mutated (*site 1+2*) (Figure 2D). Consistent with requiring an additional cis element, cloning just the SigA binding site of -35

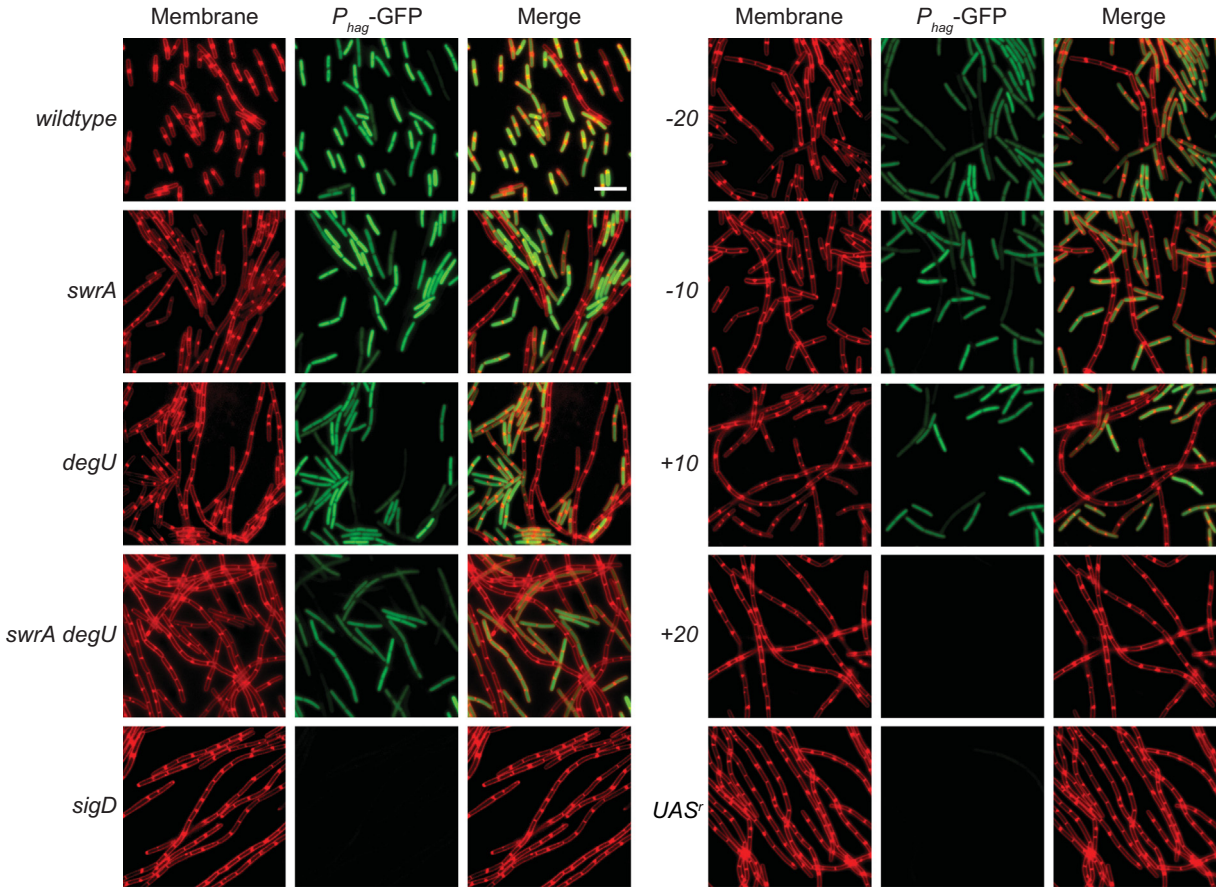


Figure 7. Moving the DegU binding site towards P_{flache} resembles a *sigD* mutant. Fluorescent micrographs of strains expressing P_{hag}-GFP in the indicated genetic backgrounds. Membrane was false colored red and P_{hag}-GFP reporter was false colored green. The following strains are used to generate this panel: *wt* (DK3858), *swrA* (DB549), *degU* (DB550), *swrA degU* (DB1031), *sigD* (DB612), *−20* (DB521), *−10* (DB520), *+10* (DB504), *+20* (DB519) and UAS^r (DB1046). Scale bar is 8 μm.

to +1 region upstream of *lacZ* abolished expression all-together (Figure 2C). Moreover, the region containing the putative cis element was replaced with a randomized sequence (UAS^r), and despite the presence and proper position of the DegU binding site, expression of the mutated P_{flache} reporter was severely impaired even when SwrA was overexpressed (Figure 2D). Finally, introduction of the randomized sequence at the native site also abolished swarming motility and P_{hag}-GFP expression similar to that observed for the +20 DegU binding site relocation (Figures 6D and 7). Consistent with a critical requirement, the cis element was highly conserved in a sequence alignment of P_{flache} promoter sequences from closely-related species (Figure 5C, Table S8). We conclude that the sequence between the DegU binding site and the SigA -35 box contains a cis-element that functions as an upstream activating sequence (UAS) that is essential for basal expression.

To further explore the relative functions of the DegU binding site and the UAS, suppressor mutations that restored swarming motility were selected in cells mutated for either cis element (Table 1, Figure S9). Consistent with both cis elements having an activator role, suppressor mutations arose in either background that deleted the terminator of the *codY* gene residing immediately upstream of the P_{flache} promoter. Previous studies have shown that the *codY* terminator deletions increase expression of the *flache* operon by read-through transcription originating from a strong promoter upstream of *codY*.[35] Thus, the *codY* terminator deletions increase *flache* expression bypassing the need for DegU, SwrA, and the UAS. Additional suppressor mutations that restored swarming motility to the DegU binding site mutants changed the −10 box closer to the SigA consensus whereas mutations that restored swarming motility to the randomized UAS changed the -35 box closer to consensus. Both likely increased expression from the P_{flache} promoter but may do so for different reasons as indicated by the correlation between each cis element and suppressors in particular SigA boxes. We conclude that both DegU and the UAS activate P_{flache} and may do so by different mechanisms. We further conclude that the position of the DegU binding site is constrained by the UAS and an additional factor SwrA is required to oligomerize DegU and enhance interaction with RNA polymerase (Figure 8A).

Discussion

SwrA is a poorly-understood activator of transcription, which while conserved in close relatives of *Bacillus subtilis*, is mutated in commonly-used domesticated strains.[9,40,53–55] How SwrA activates gene expression is unknown but recent work has indicated that SwrA modifies the function of the response regulator DegU.[19,20,42] Here we took a global ChIP-seq approach to show that SwrA interacts with DNA indirectly at a subset of DegU binding sites *in vivo*. The DNA binding consensus of DegU is unknown but by focusing on one of the strongest targets, the P_{flache} promoter, we found an inverted repeat that DegU bound with high affinity. Moreover, addition of SwrA synergized with DegU phosphorylation to enhance DNA binding and caused DegU to protect a region larger than the inverted repeat. We hypothesize that the inverted repeat in P_{flache} represents the DegU consensus binding site and that SwrA-induced DegU oligomerization may permit binding at other sites in the chromosome where consensus conservation is poor. Moreover, we suggest that SwrA functions as a co-activator to allow DegU to bind DNA far from the promoter while permitting contact with RNA polymerase. Thus, the SwrA/DegU complex constitutes the master activator of flagellar biosynthesis in *B. subtilis*.

Table 1 Suppressors of *site1+2* and *UAS^r* mutants.

Suppressors of *site$^+$*		
Strain	**Suppressor**	**Mutation**
DB382	1	P_{flache} TACAAT > TATAAT (-10 SigA box)
DB383	2	deletion in *codY* rho-independent terminator
DB384	3	P_{flache} TACAAT > TATAAT (-10 SigA box)
DB386	4	T > G at P_{flache} promoter position -16
DB387	5	T > G at P_{flache} promoter position -16
DB388	6	P_{flache} TACAAT > TATAAT (-10 SigA box)
Suppressors of *UAS^r*		
Strain	**Suppressor**	**Mutation**
DB1090	1	deletion in *codY* rho-independent terminator
DB1091	2	deletion in *codY* rho-independent terminator
DB1096	3	deletion in *codY* rho-independent terminator
DB1097	4	deletion in *codY* rho-independent terminator
DB1100	5	deletion in *codY* rho-independent terminator
DB1102	6	P_{flache} TAGACT > TTGACT (-35 SigA box)

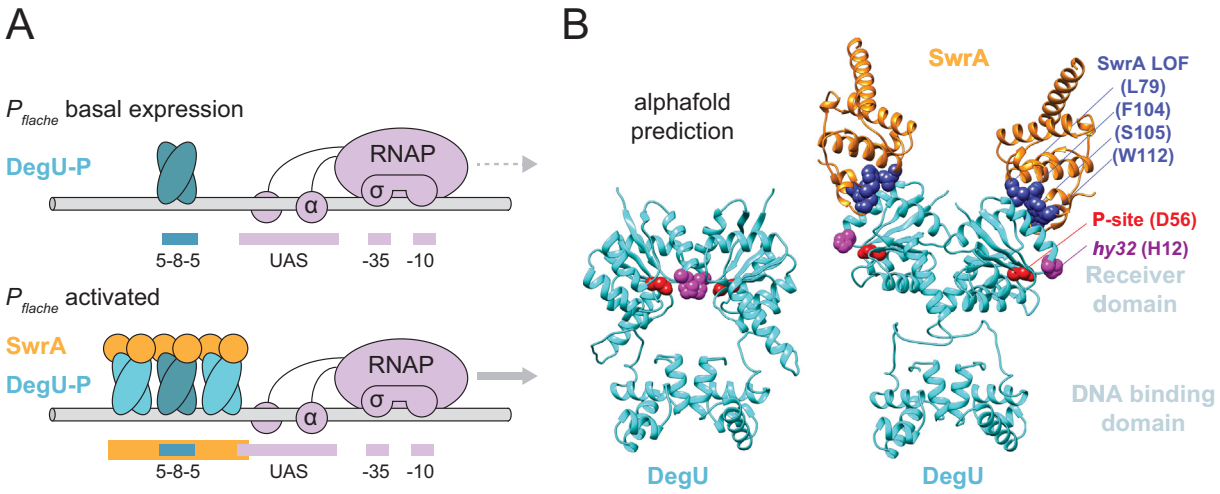


Figure 8. Flagellar gene expression requires a UAS for basal expression and SwrA-induced expansion of the DegU binding site for further activation. A) Model of SwrA-induced DegU oligomerization. Top, a dimer of DegU-P (blue) binds to the P_{flache} promoter region (gray cylinder) at the 5-8-5 inverted repeat site. Meanwhile, RNA polymerase (pink) α subunit and σ subunit bind at the UAS sequence and -35/-10 SigA boxes respectively to initiate a basal level of expression (dashed gray arrow). The regulatory state shown in this panel reflects what happens in laboratory strains lacking SwrA. Bottom, SwrA binds to DegU-P causing it to oligomerize and make putative contact with the α C-terminal domain and enhance expression (solid gray arrow). The regulatory state shown in this panel reflects what happens in swarming cells. B) AlphaFold-Multimer prediction of the DegU dimer (cyan) and DegU dimer bound by SwrA (orange). Red, residue phosphorylated by DegS. Purple, residue mutated by the hyperactive hy32 allele. Dark blue, residues identified in a forward genetic screen which when mutate abolished SwrA activity[16]. AlphaFold-Multimer prediction confidence graphs included in Figure S10.

The paradigm of transcriptional activators was established with the pleiotropic CAP/CRP protein of *E. coli*.[56–58] Briefly, CAP binds as a dimer to an inverted repeat sequence[59–61] upstream of weak sigma binding boxes and increases RNA polymerase recruitment[62,63] by interaction with the C-terminal domain of α subunit (α-CTD).[64,65] The location of the CAP binding site is important, such that if the CAP site is relocated too far from, too close to, or on the opposite face of DNA from the sigma -35 box, activation is impaired.[66–69] Like CAP, DegU is a dimeric, pleiotropic, regulatory protein that binds to an inverted repeat sequence upstream of the P_{flache} promoter. DegU binding however was not sufficient and required an additional protein SwrA to increase P_{flache} expression. Moreover, moving the DegU binding site abolished activation of the P_{flache} promoter and moving it too far forward caused a near complete failure of expression, a defect greater than mere deactivation. Further analysis indicated that moving the DegU binding site too close to the -35 box interrupted an upstream activation sequence (UAS) that was required for promoter activity. Thus, SwrA and DegU not only activate expression of the P_{flache} promoter but must do so in a way that accommodates a UAS cis element that intercedes between the SigA-recognition elements and the DegU binding site.

The P_{flache} promoter-proximal UAS cis-element was required for high level promoter activity as randomizing the sequence dramatically reduced expression despite an intact and properly spaced DegU binding sequence. The UAS likely acts as an UP element. UP elements are sequences located upstream of the -35 box of certain promoters[70–72] and are bound by the alpha subunit of RNA polymerase to either increase RNA polymerase stability at the promoter or promote open complex formation to initiate transcription.[70,71,73,74] While we did not determine whether the RNAP alpha interacts with the UAS, we note that the sequence is very similar to previously-characterized UP elements that enhance the flagella-related P_{hag} and P_{fliD} promoters in *B. subtilis*.[75–78] Moving the DegU binding site 20 base pairs closer to the -35 box disrupted, and failed to substitute for, the UAS element suggesting that the two sequences serve different functions in promoting transcription. Suppressors that restored swarming to cells mutated for either the DegU binding site or the UAS element increased expression of the *flache* operon but altered different SigA box elements. Perhaps one cis-element acts to increase the affinity of RNA polymerase for the promoter, while the other may enhance open complex formation. Whatever the case, it appears that DegU evolved to bind as close to the promoter as possible while not interrupting the UAS, and SwrA extends the reach of DegU.

We infer that SwrA expands the region of protection on P_{flache} by changing the oligomeric state of the DegU response regulator (Figure 8A).

There is no evidence that SwrA interacts with DNA directly and a bacterial two-hybrid assay indicated that SwrA interacted with the N-terminal domain of DegU seemingly far from the DNA binding surface.[19] The N-terminal DegU-SwrA interaction is supported by AlphaFold-Multimer and SwrA residues previously reported as essential for activity are located at the putative interface between the two proteins (Figure 8B, Figure S10).[16] Moreover, the AlphaFold model suggests that SwrA interaction induces freedom of rotation in the DegU N-terminal dimerization domain, which could provide a mechanism for lateral oligomerization. SwrA-induced oligomerization might also explain the diffuse appearance of some EMSA supershifted bands and why high concentrations of the complex appeared to cause clogging of polyacrylamide wells. DegU oligomerization might also permit and/or enhance interaction with weak targets in the chromosome and explain why an inverted-repeat consensus binding site is difficult to detect at many sites. Finally, we note the residue altered by the enigmatic DegU^{hy32} hyper-active allele isolated in laboratory strains lacking SwrA sits at the putative DegU dimerization interface and might induce mobility of the DegU N-terminal domain to partially mimic the SwrA-bound state (Figure 8B).[21,22,33,35,42,43]

Understanding the role of DegU in *B. subtilis* physiology has been complicated, at least in part, by the use of hyper-active alleles such as DegU^{hy32}. DegU^{hy32} is thought to represent a hyper-phosphorylated form that causes a wide variety of different phenotypes including the inhibition of motility.[33,35,79] As phosphorylated DegU is normally required for the activation the motility, it was inferred that hyper-phosphorylated DegU might somehow convert DegU from being an activator to a repressor at the P_{flache} promoter 35). Here we show that DegU^{hy32} dramatically increases the number of sites at which DegU interacts *in vivo* and while it inhibits expression of the downstream SigD-dependent P_{hag} promoter (Figure S11), it does not do so by repressing P_{flache} (Figure S11). If and how DegU transitions from being an activator to a repressor at some sites, the target promoters at which this happens, and the mechanism of transcriptional repression are unknown. Alternatively, we speculate that DegU^{hy32} could appear to function as a repressor if it were to activate the expression of an inhibitor. Consistent with DegU primarily being an activator of motility, suppressors that restored motility to strain expressing DegU^{hy32} are the same as the suppressors that restored swarming to cells mutated for the UAS (Table 1, Figure S9).[35]

Promoters have transcriptional activators to increase expression in response to environmental input, but why some promoters have UAS/UP elements is unknown. Experimentally, promoter activity has traditionally been studied in the context of either an activator or an UP element, making promoters under the control of both seem either rare or simply unappreciated.[80,81] Here, the P_{flache} promoter requires an UP-element-like UAS for basal expression and a DegU binding site located further upstream to induce high frequency swimming in liquid and swarming motility on surfaces. The two elements seem non-redundant and, when bound by their cognate proteins, may enhance promoter activity by parallel mechanisms as one seems unable to substitute for the other. Also curious is the fact that the seemingly-conventional response regulator DegU requires a co-activator SwrA to expand the binding site and activate P_{flache} gene expression. We speculate that complex transcription factors may be needed at particular promoters, like P_{flache}, to accommodate the juxtaposition of two spatially-constrained cis-element sequences that cannot be superimposed. While DegU may be unique in requiring a promoter-specific co-activator, we note that a poorly-understood enterobacterial protein called Sxy/TfoX is required for CAP to activate a subset of its regulon dedicated to DNA uptake.[82–86] Sxy/TfoX, while dissimilar in sequence, resembles SwrA in that it interacts with a transcription factor,[87,88] increases expression of genes for a transenvelope nanomachine,[89–91] activates a promoter with a putative UP element,[92] and is proteolytically-regulated by Lon.[93] We infer that protein co-activators may be a generalizable paradigm to selectively activate a subset of genes in a transcription factor's regulon.

Materials and Methods

Strain and growth conditions: *B. subtilis* and *E. coli* 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. When appropriate, antibiotics were added at the following concentrations: 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). Isopropyl β-D-thiogalactopyranoside (IPTG, Sigma) was added to LB medium at 1 mM concentration when required.

Strain construction: All PCR products were amplified from *B. subtilis* chromosomal DNA, from the indicated strains. Constructs were transformed into the naturally competent strain DK1042 (3610 *comI*^{Q12L})[94] and transduced using SPP1-mediated generalized transduction to other genetic backgrounds.[95]

SPP1 phage transduction: To 0.2 ml of dense culture (OD600- 0.6–1.0), grown in TY broth (LB broth supplemented with 10 mM $MgSO_4$ and

100 μM $MnSO_4$), serial dilutions of SPP1 phage stock were added and statically incubated for 15 min at 37 °C. To each mixture, 3 ml TYSA (molten TY supplemented with 0.5% agar) was added, poured atop fresh TY plates, and incubated at 37 °C overnight. Top agar from the plate containing near confluent plaques was harvested by scraping into a 15 ml conical tube, vortexed, and centrifuged at 5,000 X g for 10 min. The supernatant was passed through a 0.45 μm syringe filter and stored at 4 °C. Recipient cells were grown to OD600- 0.6–1.0 in 3 ml TY broth at 37 °C. One ml cells were mixed with 25 μl of SPP1 donor phage stock. Nine mL of TY broth was added to the mixture and allowed to stand at 37 °C for 30 min. The transduction mixture was then centrifuged at 5,000 X g for 10 min, the supernatant was discarded, and the pellet was resuspended in the remaining volume. 100 μL of cell suspension was plated on TY fortified with 1.5% agar, the appropriate antibiotic, and 10 mM sodium citrate and incubated at 37 °C overnight. All strains used in this study are listed in Table 2. All primers used to build strains for this study are listed in Table S6 and all plasmids are listed in Table S7.

Transcriptional reporter constructs: The P_{swrA}, P_{yxjJ}, P_{mcpA}, P_{ynel}, P_{ydaJ}, P_{ywdA}, P_{sinI} and P_{tlpA} promoter regions were amplified from *B. subtilis* strain DK1042 as the template using the primer pairs (353/354), (5120/5121), (5118/5119), (5116/5117), (6481/6482), (6483/6484), (6187/6188) and (6189/6190), respectively. The amplicons were digested with EcoRI and BamHI and ligated into the EcoRI and BamHI sites of pDG268 containing the *lacZ* gene and the *cat* gene for chloramphenicol resistance between arms of *amyE*[96] to generate pDP144, pDP464, pDP463, pDP462, pAM08, pAM09, pDP521 and pDP522.

The P_{ytvA} and P_{yweA} promoter regions were amplified from *B. subtilis* strain DK1042 as the template using the primer pair (6487/6488) and (6489/6490) respectively. The amplicons were digested with MfeI and BamHI and ligated into the EcoRI and BamHI sites of pDG268 to generate pAM11 and pAM12.

The P_{sacX} promoter region was amplified from *B. subtilis* strain DK1042 as the template using the primer pair (6069/6064). The amplicon was digested with HinDIII and BamHI and ligated into the HinDIII and BamHI sites of pDG268 to generate pAM01.

P_{flache} region was amplified from chromosomal DNA from strains DB49(*site1+2*), DB486(*+20*) and DB1012(UAS^r) using primer (8008/8007), digested with EcoRI and BamHI and ligated into EcoRI and BamHI of pDG268 to generate pAM72, pAM77 and pDP617 respectively.

-35 to +1 of P_{flache} was amplified from DK1042 chromosomal DNA using primers (7227/7228), digested with EcoRI and BamHI and ligated into EcoRI and BamHI of pDG268 to generate pAM27.

Native site mutants: Native site mutants (Figure 6A) were created by allelic replacement. To generate the following mutations at the native site, the P_{flache} region was amplified from DK1042 chromosomal DNA using the primer pairs *6T* (7817/7820, 7819/7818), *6A* (7817/7822, 7821/7818), *site1* (7817/7824, 7823/7818), *site2* (7817/7826, 7825/7818), *site1+2* (7817/7828, 7827/7818), *6G* (7817/7870, 7869/7818), *6C* (7817/7872, 7871/7818), *+10* (7817/7887, 7886/7885), *+20* (7817/7980, 7979/7885), *−10* (7817/7982, 7981/7885) and *−20* (7817/7984, 7983/7885). The fragments were assembled by Gibson assembly[97] and the assembled product was digested with EcoRI and SalI. and ligated into the EcoRI and SalI site of pminiMAD carrying a temperature sensitive origin of replication in *B. subtilis* and the *erm* gene conferring mls resistance,[98] to generate pAM39, pAM40, pAM41, pAM42, pAM43, pAM48, pAM49, pAM56, pAM67, pAM68 and pAM69 respectively.

To generate UAS^r native site mutant, the P_{flache} region was amplified using primer pairs 8205/8206 and 8207/8208. The fragments were digested with EcoRI/NheI and NheI/BamHI respectively and ligated into the EcoRI and BamHI site of pminiMAD to generate pDP616.

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 mls 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 mls selection. Cells were serially diluted and plated on LB agar plates in the absence of mls. Individual colonies were replica patched on LB agar plates with and without mls to identify mls sensitive colonies that have successfully evicted the plasmid. Chromosomal DNA was isolated form the colonies that had excised the plasmid and allelic replacement was confirmed by PCR amplification of the P_{flache} region using primers 1921 and 3042 followed by sequencing using the same set of primers individually.

Swarm expansion assay: Cells were grown to mid-log phase (OD_{600} 0.3–1.0) at 37 °C in lysogeny broth (LB) and resuspended to and OD_{600} of 10 in pH 8.0 PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4) 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

Table 2 Strains.

Strain	Genotype[a]
E. coli	
DE776	*pT7 DegU-6His amp*
DE891	*pT7 DegS-6His kan*
DE2671	*pT7-GST-SwrA amp*
B. subtilis	
3610	*wildtype*
DB47	*comI^{Q12L} P$_{flache}$$^{DegUBS(6T)}$*
DB48	*comI^{Q12L} P$_{flache}$$^{DegUBS(6A)}$*
DB49	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$*
DB58	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ thrC::P$_{hyspank}$-swrA mls*
DB102	*comI^{Q12L} P$_{flache}$$^{DegUBS(site2)}$*
DB164	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1)}$*
DB165	*comI^{Q12L} P$_{flache}$$^{DegUBS(6G)}$*
DB166	*comI^{Q12L} P$_{flache}$$^{DegUBS(6C)}$*
DB195	*comI^{Q12L} P$_{flache}$$^{DegUBS(+10)}$*
DB196	*comI^{Q12L} P$_{flache}$$^{DegUBS(6A)}$ thrC::P$_{hyspank}$-swrA mls*
DB197	*comI^{Q12L} P$_{flache}$$^{DegUBS(site2)}$ thrC::P$_{hyspank}$-swrA mls*
DB198	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1)}$ thrC::P$_{hyspank}$-swrA mls*
DB199	*comI^{Q12L} P$_{flache}$$^{DegUBS(6G)}$ thrC::P$_{hyspank}$-swrA mls*
DB200	*comI^{Q12L} P$_{flache}$$^{DegUBS(6C)}$ thrC::P$_{hyspank}$-swrA mls*
DB220	*comI^{Q12L} P$_{flache}$$^{DegUBS(+10)}$ thrC::P$_{hyspank}$-swrA mls*
DB228	*comI^{Q12L} P$_{flcahe}$$^{DegUBS(6T)}$ thrC::P$_{hyspank}$-swrA mls*
DB382	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup1*
DB383	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup2*
DB384	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup3*
DB386	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup4*
DB387	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup5*
DB388	*comI^{Q12L} P$_{flache}$$^{DegUBS(site1+2)}$ sup6*
DB456	*comI^{Q12L}thrC::P$_{hag}$-GFP mls*
DB486	*comI^{Q12L}P$_{flache}$$^{DegUBS(+20)}$*
DB487	*comI^{Q12L}P$_{flache}$$^{DegUBS(-10)}$*
DB488	*comI^{Q12L}P$_{flache}$$^{DegUBS(-20)}$*
DB504	*comI^{Q12L} P$_{flache}$$^{DegUBS(+10)}$ amyE::P$_{hag}$-GFP cat*
DB519	*comI^{Q12L} P$_{flache}$$^{DegUBS(+20)}$ amyE::P$_{hag}$-GFP cat*
DB520	*comI^{Q12L} P$_{flache}$$^{DegUBS(-10)}$ amyE::P$_{hag}$-GFP cat*
DB521	*comI^{Q12L} P$_{flache}$$^{DegUBS(-20)}$ amyE::P$_{hag}$-GFP cat*
DB522	*comI^{Q12L} P$_{flache}$$^{DegUBS(+20)}$ thrC::P$_{hyspank}$-swrA mls*
DB523	*comI^{Q12L} P$_{flache}$$^{DegUBS(-10)}$ thrC::P$_{hyspank}$-swrA mls*
DB524	*comI^{Q12L} P$_{flache}$$^{DegUBS(-20)}$ thrC::P$_{hyspank}$-swrA mls*
DB534	*comI^{Q12L}amyE::P$_{flache}$$^{DegUBS\ (site1+2)}$-lacZ cat*
DB549	*comI^{Q12} swrA::kan amyE::P$_{hag}$-GFP cat*
DB550	*comI^{Q12} degU::TnYLB kan amyE::P$_{hag}$-GFP cat*
DB556	*comI^{Q12L}amyE::P$_{flache}$$^{DegUBS(+20)}$-lacZ cat*
DB570	*comI^{Q12L}swrA::kan amyE::P$_{flache}$$^{DegUBS(+20)}$-lacZ cat*
DB573	*comI^{Q12L}thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache}$$^{DegUBS(site1+2)}$-lacZ cat*
DB578	*comI^{Q12L}thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache}$$^{DegUBS(+20)}$-lacZ cat*
DB612	*comI^{Q12L}sigD::tet amyE::P$_{hag}$-GFP cat*
DB613	*comI^{Q12} P$_{flache}$$^{DegUBS(+20)}$ swrA::kan amyE::P$_{hag}$-GFP cat*
DB643	*comI^{Q12L}thrC::P$_{flache}$-lacZ mls*
DB795	*comI^{Q12L} P$_{flache}$$^{DegUBS(+20)}$ degU::TnYLB kan amyE::P$_{hag}$-GFP cat*
DB1012	*comI^{Q12L}amyE::P$_{flache}$^{UASr}-lacZ cat*
DB1013	*ΔswrA ΔdegU amyE::P$_{swrA}$-lacZ cat*
DB1014	*ΔswrA ΔdegU amyE::P$_{mcpA}$-lacZ cat*
DB1015	*ΔswrA ΔdegU amyE::P$_{yxjJ}$-lacZ cat*
DB1020	*comI^{Q12L}P$_{flache}$^{UASr}*
DB1031	*ΔswrA ΔdegU amyE::P$_{hag}$-GFP cat*
DB1032	*ΔswrA ΔdegU amyE::P$_{flgM}$-lacZ cat*
DB1041	*comI^{Q12L}swrA::kan amyE::P$_{flache}$^{UASr}-lacZ cat*
DB1046	*comI^{Q12L}P$_{flache}$^{UASr} amyE::P$_{hag}$-GFP cat*
DB1053	*comI^{Q12L}thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache}$^{UASr}-lacZ cat*
DB1071	*comI^{Q12L}P$_{flache}$^{UASr} thrC::P$_{hyspank}$-swrA mls*
DB1074	*comI^{Q12L}swrA::kan amyE::P$_{flache}$$^{DegUBS\ (site1+2)}$-lacZ cat*

(*continued on next page*)

Table 2 (*continued*)

Strain	Genotype[a]
DB1090	*comI^{Q12L}P$_{flache}$^{UASr} sup1*
DB1091	*comI^{Q12L} P$_{flache}$^{UASr} sup2*
DB1096	*comI^{Q12L}P$_{flache}$^{UASr} sup3*
DB1097	*comI^{Q12L}P$_{flache}$^{UASr} sup4*
DB1100	*comI^{Q12L}P$_{flache}$^{UASr} sup5*
DB1102	*comI^{Q12L}P$_{flache}$^{UASr} sup6*
DB1159	*ΔpgsB degU::TnYLB kan amyE::P$_{hyspank}$-degUhy(32) lacI spec thrC::P$_{flache}$-lacZ mls*
DB1160	*ΔpgsB degU::TnYLB kan swrA::tet amyE::P$_{hyspank}$-degUhy(32) lacI spec thrC::P$_{flache}$-lacZ mls*
DB1161	*ΔpgsB degU::TnYLB kan amyE::P$_{hyspank}$-degUhy(32) lacI spec thrC::P$_{hag}$-GFP mls*
DB1162	*ΔpgsB degU::TnYLB kan swrA::tet amyE::P$_{hyspank}$-degUhy(32) lacI spec thrC::P$_{hag}$-GFP mls*
DK1042	*comI^{Q12L}* [98]
DK3858	*comI^{Q12L}amyE::P$_{hag}$-GFP cat*
DK4148	*comI^{Q12L}amyE::P$_{yneI}$-lacZ cat*
DK4149	*comI^{Q12L}amyE::P$_{mcpA}$-lacZ cat*
DK4150	*comI^{Q12L}amyE::P$_{yxjJ}$-lacZ cat*
DK4730	*ΔswrA amyE::P$_{flache}$-lacZ cat*
DK4731	*ΔswrA amyE::P$_{yneI}$-lacZ cat*
DK4732	*ΔswrA amyE::P$_{mcpA}$-lacZ cat*
DK4733	*ΔswrA amyE::P$_{yxjJ}$-lacZ cat*
DK4734	*ΔdegU amyE::P$_{flache}$-lacZ cat*
DK4735	*ΔdegU amyE::P$_{yneI}$-lacZ cat*
DK4736	*ΔdegU amyE::P$_{mcpA}$-lacZ cat*
DK4737	*ΔdegU amyE::P$_{yxjJ}$-lacZ cat*
DK4813	*comI^{Q12L}amyE::P$_{sacX}$-lacZ cat*
DK4870	*ΔswrA amyE::P$_{flgM}$-lacZ cat*
DK4918	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache}$-lacZ cat*
DK4919	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{flgM}$-lacZ cat*
DK4927	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{yneI}$-lacZ cat*
DK4928	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{mcpA}$-lacZ cat*
DK4929	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{yxjJ}$-lacZ cat*
DK4979	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache}$-lacZ cat degU::TnYLB kan*
DK4980	*thrC::P$_{hyspank}$-swrA mls amyE:: P$_{yneI}$-lacZ cat degU::TnYLB kan*
DK4981	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{mcpA}$-lacZ cat degU::TnYLB kan*
DK4982	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{yxjJ}$-lacZ cat degU::TnYLB kan*
DK4983	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{flgM}$-lacZ cat degU::TnYLB kan*
DK5183	*comI^{Q12L}amyE::P$_{flache}$-lacZ cat*
DK5666	*comI^{Q12L}amyE::P$_{sinI}$-lacZ cat*
DK5667	*comI^{Q12L}amyE::P$_{tlpA}$-lacZ cat*
DK6387	*comI^{Q12L}amyE::P$_{ydaJ}$-lacZ cat*
DK6388	*comI^{Q12L}amyE::P$_{ywdA}$-lacZ cat*
DK6390	*comI^{Q12L}amyE::P$_{yweA}$-lacZ cat*
DK6471	*comI^{Q12L}amyE::P$_{ytvA}$-lacZ cat*
DK6473	*ΔswrA amyE::P$_{ydaJ}$-lacZ cat*
DK6474	*ΔswrA amyE::P$_{ywdA}$-lacZ cat*
DK6476	*ΔswrA amyE::P$_{yweA}$-lacZ cat*
DK6477	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{ydaJ}$-lacZ cat*
DK6478	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{ywdA}$-lacZ cat*
DK6480	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{yweA}$-lacZ cat*
DK6481	*ΔdegU amyE::P$_{ydaJ}$-lacZ cat*
DK6482	*ΔdegU amyE::P$_{ywdA}$-lacZ cat*
DK6484	*ΔdegU amyE::P$_{yweA}$-lacZ cat*
DK6500	*ΔswrA amyE::P$_{ytvA}$-lacZ cat*
DK6501	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{ytvA}$-lacZ cat*
DK6502	*ΔdegU amyE::P$_{ytvA}$-lacZ cat*
DK6509	*ΔdegU thrC::P$_{hyspank}$-swrA mls amyE::P$_{ydaJ}$-lacZ cat*
DK6510	*ΔdegU thrC::P$_{hyspank}$-swrA mls amyE::P$_{yweA}$-lacZ cat*
DK6524	*ΔdegU thrC::P$_{hyspank}$-swrA mls amyE::P$_{ywdA}$-lacZ cat*
DK6550	*ΔdegU thrC::P$_{hyspank}$-swrA mls amyE::P$_{ytvA}$-lacZ cat*
DK6613	*amyE::P$_{swrA}$-lacZ cat*
DK6624	*ΔswrA amyE::P$_{swrA}$-lacZ cat*
DK6625	*thrC::P$_{hyspank}$-swrA mls amyE::P$_{swrA}$-lacZ cat*
DK6626	*ΔdegU amyE::P$_{swrA}$-lacZ cat*
DK6650	*ΔdegU thrC::P$_{hyspank}$-swrA mls amyE::P$_{swrA}$-lacZ cat*

Table 2 (*continued*)

Strain	Genotype[a]
DK7502	*ΔswrA ΔdegU amyE::P$_{flache}$-lacZ cat*
DK7631	*comI^{Q12L}amyE::P$_{ycdC}$-lacZ cat*
DK7928	*comI^{Q12L}amyE::P$_{flache(-35-+1)}$-lacZ cat*
DK7970	*ΔswrA amyE::P$_{flache(-35-+1)}$-lacZ cat*
DK7976	*comI^{Q12L}thrC::P$_{hyspank}$-swrA mls amyE::P$_{flache(-35-+1)}$-lacZ cat*
DK9504	*ΔpgsB amyE::P$_{hyspank}$-degQ spec*
DK9523	*ΔpgsB degU::TnYLB kan amyE::P$_{hyspank}$-degUhy(32) lacI spec*
DK9524	*ΔpgsB degU::TnYLB kan swrA::tet amyE::P$_{hyspank}$-degUhy(32) lacI spec*
DK9525	*ΔpgsB swrA::tet amyE::P$_{hyspank}$-degQ spec*
DS811	*amyE::P$_{flgM}$-lacZ cat*[13]
DS1868	*amyE::P$_{flache}$-lacZ cat*
DS2415	*ΔswrA*[107]
DS3534	*degU::TnYLB kan amyE::P$_{hyspank}$-degU lacI spec*
DS3649	*ΔdegU*[107]
DS3658	*ΔdegU amyE::P$_{flgM}$-lacZ cat*
DS6262	*ΔswrA ΔdegU*
DS8094	*ΔswrA amyE::P$_{hyspank}$-swrA spec*
DS8111	*ΔswrA amyE::P$_{hyspank}$-degU spec*
DS8259	*ΔdegU amyE::P$_{hyspank}$-swrA spec*

[a] All *B. subtilis* strains are in either 3610 or DK1042 genetic backgrounds and the *E. coli* strains are in BL21 background.

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. IPTG was added to the medium at final concentration of 1 mM to the LB broth and plates when appropriate.

β-galactosidase assay: *B. subtilis* strains were grown in LB broth at 37 °C with constant rotation to OD (0.7–1.0). One ml of cells were harvested and assayed for β-galactosidase activity. The OD_{600} of each sample was measured and the cells were centrifuged and resuspended in 1 ml of Z-buffer (40 mM NaH_2PO_4, 60 mM Na_2HPO_4, 1 mM $MgSO_4$, 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 Na_2CO_3. The OD_{420} of the reaction mixtures were recorded and the β-galactosidase specific activity was calculated according to the equation: (OD_{420}/time x OD_{600})] x dilution factor x 1000. IPTG was added to the medium at final concentration of 1 mM to the LB broth and plates when appropriate. Average β-galactosidase activities and standard deviations in Figure 2 are presented in Table S5.

DegU-His$_6$ protein purification: The DegU-His$_6$ fusion protein expression vector pNW43 was transformed into BL21(DE3) *E. coli*, and the resulting strain, DE776, was grown to an OD_{600} of 0.7 at 37 °C with constant shaking in 1 liter of Terrific Broth (12 g tryptone, 25 g yeast extract, 4 ml glycerol and 100 ml sterile potassium phosphate solution (2.31 g KH_2PO_4, 12.54 g K_2HPO_4) supplemented with 100 μg/ml ampicillin. Protein expression was induced by adding 1 mM IPTG, and growth was continued overnight at 16 °C. Cells were pelleted, resuspended in lysis buffer (50 mM Tris HCl pH7.6 and 150 mM NaCl), treated with lysozyme and lysed by sonication. Lysed cells were centrifuged at 8000*g* for 30 min at 4 °C. Supernatant was isolated and combined with 2 ml of nickel-nitrilotriacetic acid (Ni-NTA) His Bind resin (Novagen), equilibrated in lysis buffer and incubated for 1 h at 4 °C with gentle rotation. The resin-lysate mixture was poured into a 1-cm separation column (Bio-Rad), the resin was allowed to pack, and the lysate was allowed to flow through the column. The resin was washed with wash buffer (50 mM Tris HCl pH7.6, 150 mM NaCl and 5 mM Imidazole). DegU-His$_6$ protein bound to the resin was eluted in a stepwise manner using lysis buffer supplemented with 10, 30, 100, 250 and 500 mM Imidazole. Elution products were separated by 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie stained to verify purification of the DegU-His$_6$ protein and fractions containing clean DegU-His$_6$ were pooled and concentrated to 2 ml. Finally, the concentrated protein was purified via size exclusion

chromatography using a Superdex 75 16/60 column (GE Heathcare). DegU-His_6 was stored in storage buffer (20 mM Tris pH7.6, 10% glycerol, 200 mM NaCl) at −80C. Concentration of protein was determined by Bradford assay (Bio-rad).

DegS-His_6 protein purification: The DegS-His_6 fusion protein expression vector pYH8 was transformed into BL21(DE3) *E. coli*, and the resulting strain, DE891, was grown to an OD_{600} of 0.5 at 37 °C with constant shaking in 500 ml of LB Broth supplemented with 25 μg/ml kanamycin. Protein expression was induced by adding 1 mM IPTG, and growth was continued for an additional 4 h at 30 °C. Cells were pelleted, resuspended in lysis buffer (50 mM Tris HCl pH7.6 and 150 mM NaCl), treated with lysozyme and lysed by sonication. Lysed cells were centrifuged at 8000*g* for 30 min at 4 °C. Insoluable pellets were resuspended in 4 ml of resuspension buffer (6 M guanidine HCl, 50 mM Tris HCl pH7.6 and 150 mM NaCl), stored on ice for 30 min followed by centrifugation at 30,000*g* for 15 min. The supernatant was isolated and combined with 2 ml of nickel-nitrilotriacetic acid (Ni-NTA) His Bind resin (Novagen), incubated at room temperature for 20 min and poured into a 1-cm separation column (Bio-Rad). To renature the lysate-resin mixture, the column was serially washed with 10 ml of lysis buffer supplemented with 8, 6, 4 and 2 M Urea respectively, followed by a final wash of 10 ml lysis buffer. DegS-His6 was serially eluted with lysis buffer supplemented with 2 ml 100 mM, 250 mM and 500 mM imidazole respectively. Elution products were separated by 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie stained to verify purification of the DegS-His_6 protein and fractions containing clean DegS-His_6 were pooled and concentrated to 2 mL. Finally, the concentrated protein was dialyzed against storage buffer (50 mM Tris-HCl pH7.6, 200 mM KCl, 10 mM MgCl2, 1 mM dithiothreitol (DTT), 20% (v/v) glycerol) and stored at −80 °C. Concentration of protein was determined by Bradford assay (Bio-rad).

GST-SwrA protein purification: The GST-SwrA fusion protein expression vector pSM94 was transformed into BL21(DE3) *E. coli*, and the resulting strain, DE2671, was grown to an OD_{600} of 0.5 at 37 °C with constant shaking in 500 ml of LB broth supplemented with 100 μg/ml ampicillin. Protein expression was induced by adding 1 mM IPTG, and growth was continued for an additional 4 h at 30C. Cells were pelleted, resuspended in lysis buffer (25 mM Tris-HCl pH 8.0, 1 mM DTT, 1 mM EDTA, 0.1% Triton X-100, 150 mM NaCl, and protease inhibitor cocktail (Roche)), treated with lysozyme and lysed by sonication. Lysed cells were centrifuged at 8000*g* for 30 min at 4 °C. The supernatant was isolated and combined with 2 ml of Glutathione-Sepharose resin (GE Healthcare) and incubated for 1 h at 4 °C. The mixture was poured into a a 1-cm separation column (Bio-Rad) and the column was washed with wash buffer (25 mM Tris-HCl pH 8.0, 1 mM DTT, 0.1% NP-40, 250 mM NaCl, 10% glycerol, and protease inhibitor cocktail (Roche)). GST-SwrA was eluted form the resin using elution buffer 25 mM Tris-HCl pH 8.5, 20 mM glutathione, 1 mM DTT, 250 mM NaCl, 10% glycerol, and protease inhibitor cocktail (Roche)). Elution products were separated by 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie stained to verify purification of the GST-SwrA protein and fractions containing clean GST-SwrA were pooled and concentrated to 2 ml. Finally, the concentrated protein was dialyzed against storage buffer (25 mM Tris-HCl [pH 8.0], 1 mM DTT, 250 mM NaCl, 10% glycerol, 20% sucrose) and stored at −80 °C. Concentration of protein was determined by Bradford assay (Bio-rad).

Antibody generation: 1 mg His_6-DegU protein was sent to Cocalico Biologicals for serial injection into a rabbit host for antibody generation. The antibody was purified from serum by mixing it with His_6-DegU-conjugated Affigel-10 and incubating overnight at 4 °C. The slurry was loaded onto a 1 cm column (BioRad) and eluted with 100 mM glycine (pH 2.5) dropwise and neutralized with 2 M unbuffered Tris base. Elutions were separated by SDS-PAGE and Coomassie stained. Peak elutions were pooled, dialyzed into 1 X PBS, 50% glycerol, and BSA was added to a final concentration of 1 mg/ml prior to storage at −20 °C.

Chromatin Immunoprecipitation Sequencing (ChIP-Seq): *Bacillus subtilis* cultures were grown to an OD_{600} 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 ∼200 bp using Qsonica sonicator (Q8000R), and then incubated overnight at 4 °C with α-SwrA or α-DegU as indicated. 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 UltraII DNA library prep kit (NEB #E7645L). Paired end sequencing of the libraries was performed using Illumina NextSeq 550 platform and atleast 3 million paired end reads were obtained for each sample. Two or three biological replicates were sequenced for each sample.

Whole genome sequencing (WGS): *Bacillus subtilis* cultures were grown to an OD_{600} 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). Genomic DNA was sonicated using Qsonica sonicator (Q8000R) and the sonicated DNA was used to prepare libraries using the NEBNext UltraII DNA library prep kit (NEB #E7645L). Paired end sequencing of the libraries was performed using Illumina NextSeq 550 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 (CP020102)[99] 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 was 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: 200 bp sequence surrounding the DegU peak center in *WT, degQ^{+++}* and *degU degUhy32^{+++}* (Table S1, S2 and S3) 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 6 -maxw 30 -objfun classic -revcomp -markov_order 0)[45] (Table S2). 30 bp highly enriched motif sequences were extracted and sequence logo presented in Figure S4 was created by WebLogo using default parameters.[100]

Electromobility shift assay: DNA probes of 150–200 bp regions surrounding DegU ChIP-Seq peak centers were generated by PCR amplification using DK1042 chromosomal DNA and the following primer pairs: P_{flache} (7231/1782), P_{yxjJ} (7463/7464), P_{flgM} (7459/7460), P_{mcpA} (7554/7555), P_{swrA} (7461/7462), P_{ynel} (7465/7466), P_{comK} (7229/7230) and P_{hag} (7597, 7598). When required the P_{flache} *(site1+2)* region was amplified using primer pairs (7231, 1782) using DB49 chromosomal DNA as the template. 25 nM of DNA was radiolabeled using T4-Polynucleotide kinase (NEB M0201S) and 0.5 μl of ATP(γ-^{32}P) (Perkin, 3000 Ci/mmol) in 20 μl reactions. Excess unincorporated ATP(γ-^{32}P) was removed by passing the reaction through as G-50 Micro Columns (Cytiva) and the radiolabeled DNA was stored at 4 °C until further use. When appropriate, DegU phosphorylation reactions were performed by adding DegU-His_6 and DegS-His_6 at a ratio of 1:5 and 1 mM cold ATP to kinase buffer (50 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 10 mM $MgCl_2$, 1 mM DTT, 50 mM KCl, 0.1 mg/ml BSA, 10% glycerol) and the reaction was incubated at room temperature for 20 min. 20 μl binding reactions were prepared with various concentrations of DegU, DegU-P and SwrA as indicated, 1ul of radiolabeled probe and 5 ng/ul polydI-dC (Roche) in binding buffer (50 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 10 mM $MgCl_2$, 1 mM DTT, 50 mM KCl, 0.1 mg/ml BSA, 10% glycerol, 0.1 mM ATP) and incubated at 30 °C for 30 min. 6.5% native polyacrylamide gel was prepared using 19:1 acrylamide/bisacrylamide (Biorad), 1X Tris-Glycine-EDTA buffer (25 mM Tris base, 250 mM Glycine and 1 mM EDTA pH 8.0) and 5% glycerol. Glycerol was added to the binding reaction at a final concentration of 10% to facilitate loading and 12ul of the reaction mixture was fractionated on a 6.5% native gel in 1X TGE running buffer at room temperature for 1 h at 100 V (constant). Gels were dried, exposed to a phosphorimager screen overnight and radioactive signal was detected using Typhoon FLA 9500. Fiji v 2.1.0[96] was used to quantify the fraction of unbound DNA the fraction of bound DNA was calculated by subtracting the fraction of unbound DNA from 1. DNA binding curves were generated in GraphPad Prism 9 using Non-linear regression and one site specific binding parameters. Gels used for Kd analysis are included in Figure S12.

DnaseI-footprinting assay: 300–350 bp fluorescently labelled DNA probes surrounding the DegU ChIP peak was generated by PCR amplification using a forward primer with a 5′-FAM fluorescent tag (IDT) and a reverse primer with a 5′-HEX fluorescent tag (IDT). Promoter regions used in the assay were amplified using DK1042 chromosomal DNA as template and primer pairs P_{flache}(7548/7549), P_{yxjJ}(7562/7563), P_{flgM}(7550/7551), P_{mcpA}(7566/7567), P_{swrA}(7564/7565) and P_{ynel}(7568/7569) respectively. The optimal concentration of DnaseI (NEB #M0570) at which uniform cleavage was observed across the probe was assessed over a range of concentrations (1X – 1024X) prepared in the presence of 1X DNAseI buffer. DegU was phosphorylated as described earlier. 20ul binding reactions were set up in 1X binding buffer to which 5 ng/μl polydIdC, 20 nM DNA probe, 1X DnaseI buffer and either no protein or indicated amounts of DegU-P and SwrA were added and incubated at 30 °C for 30 min. 5 μl of optimized DNAseI dilution was added to the reaction and incubated at room temperature for 15 min. Reaction was quenched by addition of 25 μl of 0.5 M EDTA pH 8.0. DNA fragments were cleaned using Qaigen MinElute PCR purification kit (#28004). Fragment analysis was performed by

Genewiz, Azenta Life Sciences using a 3730 DNA analyzer and fragment size was determined using a GeneScan 500LIZ DNA size standard. Data was analyzed using Peak Scanner software v1.0 and the peak height corresponding each fragment size was exported into a text file. The values were plotted using a custom R-script and DNaseI protection was determined by an absence of peaks across a range of consecutive fragment sizes.

Alignment of P_{flache} region: All bacteria that contain both SwrA and DegU were identified by BLAST+ v 2.12.0.[101] P_{flache} sequence ranging from +1 to −120 (Figure 5) and −121 to −250 (Figure S7) were extracted and aligned by Clustal Omega v 1.2.4 using default parameters.[102] Alignment was shaded using Jalview v 2.11.2.7[103] using a 60% identity threshold and false colored using Adobe illustrator.

Microscopy: For microscopy, 3 ml of LB broth was inoculated with a single colony and grown at 37C to OD600 05–0.8. 1 ml of culture was pelleted and resuspended in 30 μl 1X PBS buffer containing 5 μg/ml FM 4–64 (Invitrogen #T13320) and incubated for 2 min at room temperature in the dark. Excess dye was washed with 1 ml of PBS, cells were spun down and resuspended in a final volume of 30 μl of PBS. Flat agarose pads (1% agarose in PBS) were created by on a slide, 5 μl of sample was spotted on the Agarose pads 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.[104]

Structure prediction: Structure prediction was performed using Alphafold.[105] The *B. subtlis* 3610 DegU and SwrA sequences were separated by a colon (:) whenever necessary and prediction was performed using parameters colabfold_batch –num-recycle 20 –amber –templates –model-type alphafold2_multimer_v2. The structures were visualized using UCSF Chimera v 1.15.[106]

Funding

Support for this work comes from National Institutes of Health R35GM145299 to DZR, 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.

CRediT authorship contribution statement

Ayushi Mishra: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anna C. Hughes:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation. **Jeremy D. Amon:** Formal analysis, Methodology, Writing – original draft. **David Z. Rudner:** Funding acquisition, Writing – original draft. **Xindan Wang:** Data curation, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – original draft. **Daniel B. Kearns:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

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, members of the Van Kessel lab for biochemical support and we thank Zhongqing Ren for assistance with ChIP-seq. We thank the Indiana University Center for Genomics and Bioinformatics for high throughput sequencing and AlphaFold-Multimer was performed using IU Carbonate supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2023.168419.

Received 30 August 2023;
Accepted 19 December 2023;
Available online 21 December 2023

Keywords:
UP element;
SwrA;
transcription activation;
two-component system;
flagella

References

1. Komeda, Y., (1982). Fusions of flagellar operons to lactose genes on a Mu *lac* bacteriophage. *J. Bacteriol.* **150**, 16–26.
2. Kutsukake, K., Ohya, Y., Iino, T., (1990). Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J. Bacteriol.* **172**, 741–747.
3. Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M.G., Alon, U., (2001). Ordering genes in a flagellar pathway by analysis of expression kinetics from living bacteria. *Science* **292**, 2080–2083.
4. Dasgupta, N., Wolfgang, M.C., Goodman, A.L., Arora, S. K., Jyot, J., Lory, S., Ramphal, R., (2003). A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **50**, 809–824.
5. Patrick, J.E., Kearns, D.B., (2012). Swarming motility and the control of master regulators of flagellar biosynthesis. *Mol. Microbiol.* **83**, 14–23.
6. Liu, X., Matsumura, P., (1994). The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *J. Bacteriol.* **176**, 7345–7351.
7. Claret, L., Hughes, C., (2000). Functions of the subunits in the $FlhD_2C_2$ transcriptional master regulator of bacterial flagellum biogenesis and swarming. *J. Mol. Biol.* **303**, 467–478.
8. Sourjik, V., Muschler, P., Scharf, B., Schmitt, R., (2000). VisN and VisR are global regulators of chemotaxis, flagellar, and motility genes in *Sinorhizobium (Rhizobium) meliloti*. *J. Bacteriol.* **182**, 782–788.
9. Kearns, D.B., Chu, F., Rudner, R., Losick, R., (2004). Genes governing swarming in *Bacillus subtilis* and evidence of a phase variation mechanism controlling surface motility. *Mol. Microbiol.* **52**, 357–369.
10. Calvio, C., Celandroni, F., Ghelardi, E., Amati, G., Salvetti, S., Ceciliani, 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.
11. Márquez-Magaña, L.M., Chamberlin, M.J., (1994). Characterization of the *sigD* transcriptional unit of *Bacillus subtilis*. *J. Bacteriol.* **176**, 2427–2434.
12. West, J.T., Estacio, W., Márquez-Magaña, L., (2000). Relative roles of the *fla/che* $_{P}$A, $_{PD-}$3, and $_{Psig}$D promoters in regulating motility and *sig*D expression in *Bacillus subtilis*. *J. Bacteriol.* **182**, 4841–4848.
13. Kearns, D.B., Losick, R., (2005). Cell population heterogeneity during growth of *Bacillus subtilis*. *Gene Dev* **19**, 3083–3094.
14. 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.
15. Mukherjee, S., Bree, A.C., Liu, J., Patrick, J.E., Chien, P., Kearns, D.B., (2015). Adaptor-mediated Lon proteolysis restricts *Bacillus subtilis* hyperflagellation. *PNAS* **112**, 250–255.
16. Hughes, A.C., Subramanian, S., Dann III, C.E., Kearns, D. B., (2018). The C-terminal region of *Bacillus subtilis* SwrA is required for activity and adaptor-dependent LonA proteolysis. *J. Bacteriol.* **200**, e00659–e00717.
17. Olney, S.G., Chien, P., Kearns, D.B., (2022). SmiA is a hybrid priming/scaffolding adaptor for the LonA protease in *Bacillus subtilis*. *J. Biol. Chem.* **298**, 102045
18. Calvio, C., Osera, C., Amati, G., Galizzi, A., (2008). Autoregulation of *swrAA* and motility in *Bacillus subtilis*. *J. Bacteriol.* **190**, 5720–5728.
19. 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.
20. 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 *Bacillus subtilis*. *PLoS One* **8**, e85065.
21. Kunst, F., Pascal, M., Lepesant-Kejzlarova, J., Lepesant, J.A., Billault, A., Dedonder, R., (1974). Pleiotropic mutations affecting sporulation conditions and the synthesis of extracellular enzymes in *Bacillus subtilis*. *Biochimie* **56**, 1481–1489.
22. Ayusawa, D., Yoneda, Y., Yamane, K., Maruo, B., (1975). Pleiotropic phenomena in autolytic enzyme(s) content, flagellation, and simultaneous hyperproduction of extracellular α-amylase and protease in a *Bacillus subtilis* mutant. *J. Bacteriol.* **124**, 459–469.
23. Kunst, F., Debarbouille, M., Msadek, T., Young, M., Mauel, C., Karamata, D., Klier, D., Rapoport, G., Dedonder, R., (1988). Deduced polypeptides encoded by the *Bacillus subtilis sacU* locus share homology with two-component sensor-regulator systems. *J. Bacteriol.* **170**, 5093–5101.
24. Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G., Dedonder, R., (1990). Signal transduction pathway 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.
25. Kobayashi, K., (2007). Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **66**, 395–409.
26. Verhamme, D.T., Kiley, T.B., Stanley-Wall, N.R., (2007). DegU co-ordinates multicellular behavior exhibited by *Bacillus subtilis*. *Mol. Microbiol.* **65**, 554–568.
27. Ohsawa, T., Tsukahara, K., Ogura, M., (2009). *Bacillus subtilis* response regulator DegU is a direct activator of *pgsB* transcription involved in γ-poly-glutamic acid synthesis. *Biosci. Biotech. Bioch.* **73**, 2096–2102.
28. Kimura, K., Tran, L.S.P., Do, T.H., Itoh, Y., (2009). Expression of the *pgsB* encoding the poly-gamma-DL-

glutamate synthase of *Bacillus subtilis (natto*). *Biosci. Biotech. Bioch.* **73**, 1149–1155.
29. Verhamme, D.T., Murray, E.J., Stanley-Wall, N.R., (2009). DegU and Spo0A jointly control transcription of two loci required for complex colony development by *Bacillus subtilis*. *J. Bacteriol.* **191**, 100–108.
30. Baptista, C., Barreto, H.C., São-José, C., (2013). High levels of DegU-P activate an Esat-6-like secretion system in *Bacillus subtilis*. *PLoS One* **8**, e67840.
31. Ogura, M., Yamaguchi, H., Yoshida, K., Fujita, Y., Tanaka, T., (2001). DNA microarray analysis of Bacillus subtilis DegU, ComA, and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. *Nucl. Acids Res.* **29**, 3804–3813.
32. Mäder, U., Antelmann, H., Buder, T., Dahl, M.K., Hecker, M., Homuth, G., (2002). *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics. *Mol. Genet. Genomics* **268**, 455–467.
33. Dahl, M.K., Tsadek, T., Kunst, F., Rapoport, G., (1992). The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of gene competence in *Bacillus subtilis*. *J. Biol. Chem.* **267**, 14509–14514.
34. Hamoen, L.W., Werkhoven, A.F.V., Venema, G., Dubnau, D., (2000). The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*. *PNAS* **97**, 9246–9251.
35. Amati, G., Bisicchia, P., Galizzi, A., (2004). DegU-P represses expression of the motility *fla-che* operon in *Bacillus subtilis*. *J. Bacteriol.* **186**, 6003–6014.
36. Mukai, K., Kawata, M., Tanaka, T., (1990). Isolation and phosphorylation of the *Bacillus subtilis degS* and *degU* gene products. *J. Biol. Chem.* **265**, 20000–20006.
37. Dahl, M.K., Tsadek, 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.
38. Cairns, L.S., Martyn, J.E., Bromley, K., Stanley-Wall, N. R., (2015). An alternate route to phosphorylating DegU of *Bacillus subtilis* using acetyl phosphate. *BMC Microbiol.* **15**, 78.
39. Mukai, K., Kawata-Mukai, M., Tanaka, T., (1992). Stabilization of phosphorylated *Bacillus subtilis* DegU by DegR. *J. Bacteriol.* **174**, 7954–7962.
40. Stanley, N.R., Lazazzera, B.A., (2005). Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-γ-DL-glutamic acid production and biofilm formation. *Mol. Microbiol.* **57**, 1143–1158.
41. Osera, C., Amati, G., Calvio, C., Galizzi, A., (2009). SwrAA activates poly-γ-glutamate synthesis in addition to swarming in *Bacillus subtilis*. *Microbiol* **155**, 2282–2287.
42. Ermoli, F., Bontà, V., Vitali, G., Calivo, C., (2021). SwrA as global modulator of the two-component system DegSU in *Bacillus subtilis*. *Res. Microbiol.* **172**, 103877
43. Sanchez, S., Snider, E.V., Wang, X., Kearns, D.B., (2021). Identification of genes required for swarming motility in *Bacillus subtilis* using transposon mutagenesis and high-throughput sequencing (TnSeq). *J. Bacteriol.* **204**, e0008922.
44. Murray, E.J., Kiley, T.B., Stanley-Wall, N.R., (2009). A pivotal role for the response regulator DegU in controlling multicellular behavior. *Microbiol* **155**, 1–8.
45. Bailey, T.L., Johnson, J., Grant, C.E., Noble, W.S., (2015). The MEME suite. *Nucl. Acids Res.* **43**, W39–W49.
46. Henner, D.J., Ferrari, E., Perego, M., Hoch, J.A., (1988). Location of the targets of the *hpr*-97, *sacU32*(Hy), and *sacQ36*(Hy) mutations in upstream regions of the subtilisin promoter. *J. Bacteriol.* **170**, 296–300.
47. Ogura, M., Shimane, K., Asai, K., Ogasawara, N., Tanaka, T., (2003). Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol. Microbiol.* **49**, 1685–1697.
48. Shimane, K., Ogura, M., (2004). Mutational analysis of the helix-turn-helix region of Bacillus subtilis response regulator DegU, and indentification of *cis*-acting sequences for DegU in the *aprE* and *comK* promoters. *J. Biochem.* **136**, 387–394.
49. Hsueh, Y.H., Cozy, L.M., Sham, L.T., Calvo, R.A., Gutu, A.D., Winkler, M.E., Kearns, D.B., (2011). DegU-phosphate activates expression of the anti-sigma factor FlgM in *Bacillus subtilis*. *Mol. Microbiol.* **81**, 1092–1108.
50. Helmann, J.D., Marquez, L.M., Chamberlin, M.J., (1988). Cloning, sequencing, and disruption of the *Bacillus subtilis* $\sigma^2$8 gene. *J. Bacteriol.* **170**, 1568–1574.
51. Chan, J.M., Guttenplan, S.B., Kearns, D.B., (2014). Defects in the flagellar motor increase synthesis of poly-γ-glutamate in *Bacillus subtilis*. *J. Bacteriol.* **195**, 740–753.
52. Mukherjee, S., Kearns, D.B., (2014). The structure and regulation of flagella in *Bacillus subtilis*. *Annu. Rev. Genet.* **48**, 319–340.
53. 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.
54. Senesi, S., Ghelardi, E., Celandroni, F., Salvetti, S., Parisio, E., Galizzi, A., (2004). Surface-associated flagellum formation and swarming differentiation in *Bacillus subtilis* are controlled by the *ifm* locus. *J. Bacteriol.* **186**, 1158–1164.
55. Zeigler, D.R., Prágai, Z., Rodriguez, S., Chevreux, B., Muffler, A., Albert, T., Bai, R., Wyss, M., Perkins, J.B., (2008). The origins of 168, W23, and other *Bacillus subtilis* legacy strains. *J. Bacteriol.* **190**, 6983–6995.
56. Busby, S., Ebright, R.H., (1999). Transcriptional activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293**, 199–213.
57. Lawson, C.L., Swigon, D., Murakami, K.S., Darst, S.A., Berman, H.M., Ebright, R.H., (2004). Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* **14**, 10–20.
58. Lee, D.J., Minchin, S.D., Busby, S.J.W., (2012). Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152.
59. Simpson, R.B., (1980). Interaction of the cAMP receptor protein with the *lac* promoter. *Nucl. Acids Res.* **8**, 759–766.
60. Schmitz, A., (1981). Cyclic AMP receptor protein interacts with lactose operator DNA. *Nucl. Acids Res.* **9**, 277–292.
61. Schultz, S.C., Shields, G.C., Steitz, T.A., (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. *Science* **253**, 1001–1007.
62. Malan, T.P., Kolb, A., Buc, H., McClure, W.R., (1984). Mechanism of CRP-cAMP activation of *lac* operon transcription initiation activation of the P1 promoter. *J. Mol. Biol.* **180**, 881–909.

63. Zhou, Y., Merkel, T.J., Ebright, R., (1994). Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP) II. Role at class I and class II CAP-dependent promoters. *J. Mol. Biol.* **243**, 603–610.
64. Igarashi, K., Ishihama, A., (1991). Bipartite functional map of the *E. coli* RNA polymerase α subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65**, 1015–1022.
65. Chen, Y., Ebright, Y.W., Ebright, R.H., (1994). Identification of the target of a transcription activator protein by protein-protein photocrosslinking. *Science* **265**, 90–92.
66. Gaston, K., Bell, A., Kolb, A., Buc, H., Busby, S., (1990). Stringent spacing requirements for transcription activation by CRP. *Cell* **62**, 733–743.
67. Ushida, C., Aiba, H., (1990). Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucl. Acids Res.* **18**, 6325–6330.
68. Wing, H.J., Williams, S.M., Busby, S.J.W., (1995). Spacing requirements for transcription activation by *Escherichia coli* FNR protein. *J. Bacteriol.* **177**, 6704–6710.
69. Egland, K.A., Greenberg, E.P., (1999). Quorum sensing in *Vibrio fischeri*: elements of the *luxI* promoter. *Mol. Microbiol.* **31**, 1197–1204.
70. Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., Gourse, R.L., (1993). A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262**, 1407–1413.
71. Rao, L., Ross, W., Appleman, J.A., Gaal, T., Leirmo, S., Schlax, P.J., Record, M.T., Gourse, R.L., (1994). Factor independent activation of *rrnB* P1: an "extended" promoter with an upstream element that dramatically increases promoter strength. *J. Mol. Biol.* **235**, 1421–1435.
72. Estrem, S.T., Gaal, T., Ross, W., Gourse, R.L., (1998). Identification of an UP element consensus sequence for bacterial promoters. *PNAS* **95**, 9761–9766.
73. Strainic Jr, M.G., Sullivan, J.J., Velevis, A., deHaseth, P. L., (1998). Promoter recognition by *Escherichia coli* RNA polymerase: effects of the UP element on open complex formation and promoter clearance. *Biochem* **37**, 18074–18080.
74. Tagami, H., Aiba, H., (1999). An inactive open complex mediated by an UP element at *Escherichia coli* promoters. *PNAS* **96**, 7202–7207.
75. Fredrick, K., Caramori, T., Chen, Y.-F., Galizzi, A., Helmann, J.D., (1995). Promoter architecture in the flagellar regulon of *Bacillus subtilis*: high-level expression of flagellin by the σ^D RNA polymerase requires an upstream promoter element. *PNAS* **92**, 2582–2586.
76. Helmann, J.D., (1995). Compilation and analysis of *Bacillus subtilis* $^{\sigma}$A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucl. Acids Res.* **23**, 2351–2360.
77. Caramori, T., Galizzi, A., (1998). The UP element of the promoter for the flagellin gene, *hag*, stimulates transcription from both SigD- and SigA-dependent promoters in *Bacillus subtilis*. *Mol Gen Genet* **258**, 385–388.
78. Meijer, W.J.J., Salas, M., (2004). Relevance of UP elements for three strong *Bacillus subtilis* phage Φ29 promoters. *Nucl. Acids Res.* **32**, 1166–1176.
79. Tokunaga, T., Rashid, M.H., Kuroda, A., Sekiguchi, J., (1994). Effect of *degS-degU* mutations on the expression of *sigD*, encoding an alternative sigma factor, and autolysin operon of *Bacillus subtilis*. *J. Bacteriol.* **176**, 5177–5180.
80. Ross, W., Schneider, D.A., Paul, B.J., Mertens, A., Gourse, R.L., (2003). An intersubunit contact stimulating transcription initiation by E. coli RNA polymerase: interaction of the α C-terminal domain and σ region 4. *Genes Dev.* **17**, 1293–1307.
81. Chen, H., Tang, H., Ebright, R.H., (2003). Functional interaction between RNA polymerase α subunit C-terminal domain and σ^{70} in UP-element- and activator-dependent transcription. *Mol. Cell* **11**, 1621–1633.
82. Redfield, R.J., (1991). *sxy-1*, a Haemophilus influenza mutation causing greatly enhanced spontaneous competence. *J. Bacteriol.* **173**, 5612–5618.
83. Williams, P.M., Bannister, L.A., Redfield, R.J., (1994). The *Haemophilus influenze sxy-1* mutation is in a newly identified gene essential for competence. *J. Bacteriol.* **176**, 6789–6794.
84. Zulty, J.J., Barcak, G.J., (1995). Identification of a DNA transformation gene required for *com101A+* expression and supertransformer phenotype in *Haemophilus influenzae*. *PNAS* **92**, 3616–3620.
85. Karudapuram, S., Barcak, G.J., (1997). The *Haemophilus influenza dprABC* genes constitute a competence-inducible operon that requires the product of the *tfoX (sxy)* gene for transcriptional activation. *J. Bacteriol.* **179**, 4815–4820.
86. Pollack-Berti, A., Wollenberg, M.S., Ruby, E.G., (2010). Natural transformation of *Vibrio fischeri* requires *tfoX* and *tfoY*. *Environ. Microbiol.* **12**, 2302–2311.
87. Redfield, R.J., Cameron, A.D.S., Qian, Q., Hinds, J., Ali, T. R., Kroll, J.S., Langford, P.R., (2005). A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenza*. *J. Mol. Biol.* **347**, 735–747.
88. Søndberg, E., Sinha, A.K., Gerdes, K., Semsey, S., (2019). CRP interacts specifically with Sxy to activate transcription in *Escherichia coli*. *Front. Microbiol.* **10**, 2053.
89. Sinha, S., Cameron, A.D.S., Redfield, R.J., (2009). Sxy induces a CRP-S regulon in *Escherichia coli*. *J. Bacteriol.* **191**, 5180–5195.
90. Sinha, S., Mell, J.C., Redfield, R.J., (2012). Seventeen Sxy-dependent cyclic AMP receptor protein site-regulated genes are needed for natural transformation in *Haemophilus influenzae*. *J. Bacteriol.* **194**, 5245–5254.
91. Scrudato, M.L., Blokesch, M., (2012). The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.* **8**, e1002778.

92. Cameron, A.D.S., Redfield, R.J., (2008). CRP binding and transcriptional activation at CRP-S sites. *J. Mol. Biol.* **383**, 313–323.
93. Jaskólska, M., Gerdes, K., (2015). CRP-dependent positive autoregulation and proteolytic degradation regulate competence activator Sxy of *Escherichia coli*. *Mol. Microbiol.* **95**, 833–845.
94. Konkol, M.A., Blair, K.M., Kearns, D.B., (2013). Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.* **195**, 4085–4093.
95. Yasbin, R.E., Young, F.E., (1974). Transduction in *Bacillus subtilis* by bacteriophage SPP1. *J. Virol.* **14**, 1343–1348.
96. Antoniewski, C., Savelli, B., Stragier, P., (1990). The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. *J. Bacteriol.* **172**, 86–93.
97. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345.
98. Patrick, J.E., Kearns, D.B., (2008). MinJ (YvjD) is a topological determinant of cell division in *Bacillus subtilis*. *Mol. Microbiol.* **70**, 1166–1179.
99. 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–e00417.
100. Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., (2004). WebLogo: a sequence logo generator. *Gen Research* **14**, 1188–1190.
101. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopouos, J., Bealer, K., et al., (2009). BLAST+: architecture and applications. *BMC Bioinf.* **10**, 421.
102. 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.
103. 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.
104. 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.
105. Jumper, J., Evans, R., Pritzel, A., et al., (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
106. 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.
107. Patrick, J.E., Kearns, D.B., (2009). Laboratory strains of *Bacillus subtilis* do not exhibit swarming motility. *J. Bacteriol.* **191**, 7129–7133.