JB
Journals.ASM.org

The *syp* Enhancer Sequence Plays a Key Role in Transcriptional Activation by the σ^{54}-Dependent Response Regulator SypG and in Biofilm Formation and Host Colonization by *Vibrio fischeri*

Valerie A. Ray, Justin L. Eddy, Elizabeth A. Hussa,* Michael Misale, Karen L. Visick

Department of Microbiology and Immunology, Loyola University, Chicago, Illinois, USA

Biofilm formation by *Vibrio fischeri* is a complex process that requires multiple regulators. One such regulator, the NtrC-like response regulator SypG, controls biofilm formation and host colonization by *V. fischeri* via its impact on transcription of the symbiosis polysaccharide (*syp*) locus. SypG is predicted to activate *syp* transcription by binding to the *syp* enhancer (SE), a conserved sequence located upstream of four *syp* promoters. In this study, we performed an in-depth analysis of the sequences necessary for SypG to promote *syp* transcription and biofilm formation. We found that the SE sequence is necessary for SypG-mediated *syp* transcription, identified individual bases necessary for efficient activation, and determined that SypG is able to bind to *syp* promoter regions. We also identified SE sequences outside the *syp* locus and established that SypG recognizes these sequences as well. Finally, deletion of the SE sequence upstream of *sypA* led to defects in both biofilm formation and host colonization that could be restored by reintroducing the SE sequence into its native location in the chromosome. This work thus fills in critical gaps in knowledge of the Syp regulatory circuit by demonstrating a role for the SE sequence in SypG-dependent control of biofilm formation and host colonization and by identifying new putative regulon members. It may also provide useful insights into other bacteria, such as *Vibrio vulnificus* and *Vibrio parahaemolyticus*, that have *syp*-like loci and conserved SE sequences.

Biofilms represent the preferred lifestyle for many bacteria in nature (1–3). Bacteria often utilize biofilms to interact with their host organisms in either a pathogenic or mutualistic manner. One simple model used to study natural biofilm formation in the context of bacterium-host interactions is the mutualistic relationship between the marine bacterium *Vibrio fischeri* and its host, the Hawaiian bobtail squid *Euprymna scolopes* (reviewed in references 4 to 6). During the initial stages of this mutualism, *V. fischeri* forms a biofilm-like aggregate that promotes efficient colonization of a specific squid organ, the light organ (5, 7–9). Genetic manipulation of *V. fischeri* has revealed the requirement for a large polysaccharide locus, the symbiosis polysaccharide or *syp* locus, in both symbiotic aggregate formation and colonization, as well as in the production of biofilm phenotypes in culture (e.g., wrinkled colonies and pellicle formation) (8, 10, 11).

syp-dependent biofilm formation is regulated via a two-component signal transduction (2CST) system (reviewed in reference 12). These types of systems typically consist of a sensor kinase (SK), which senses a specific signal, and a response regulator (RR), which promotes the appropriate cellular response to that signal (e.g., transcription from a gene or set of genes) (13–15). The key 2CST system regulators that control biofilm formation by *V. fischeri* are the SK RscS and the predicted DNA binding RR SypG, which appear to exert their effects by inducing *syp* transcription (Fig. 1A) (8, 10, 16–18). The working model predicts that RscS senses an as-yet-unknown signal, autophosphorylates, and initiates a phosphorelay that ultimately leads to the phosphorylation of SypG. Phospho-SypG is predicted to promote transcription of the *syp* locus, resulting in the synthesis of proteins involved in the regulation, production, and transport of a polysaccharide necessary for biofilm formation and colonization (Fig. 1A) (10, 11).

Biofilm formation is also regulated at a level below *syp* transcription via the RR SypE, which controls the phosphorylation state of SypA (Fig. 1A) (9, 19). The specific role of SypA below *syp* transcription is unknown, but the phosphorylated form of this protein is unable to promote biofilm formation and colonization (9, 19). In culture, SypE functions primarily as a kinase to inhibit biofilm formation unless induced by specific conditions, such as overexpression of RscS, to serve instead as a phosphatase (9). Thus, biofilm formation depends both on activation of *syp* transcription (and the synthesis of Syp structural proteins necessary for polysaccharide production) and a switch of SypE's activity from kinase to phosphatase (Fig. 1A) (9, 20). SypG is unable to induce this switch and thus fails to promote biofilm formation unless SypE is absent or its kinase activity inactivated (19). Recently, we determined that the regulatory control by SypE can also be overcome by an excess of SypA target protein, presumably because some of it escapes phosphorylation (see Fig. S1 in the supplemental material) (19). We have used this phenomenon as a tool here to probe the role of SypG in controlling *sypA* transcription.

SypG is a multidomain protein that contains an N-terminal receiver (REC) domain with the predicted site of phosphorylation (D53), a centrally located AAA+ domain (for ATP hydrolysis), and a C-terminal DNA binding domain (10). Based on the pres-

Received 19 June 2013 Accepted 27 September 2013
Published ahead of print 4 October 2013
Address correspondence to Karen L. Visick, kvisick@lumc.edu.
* Present address: Elizabeth A. Hussa, Department of Biology, Illinois Wesleyan University, Bloomington, Illinois, USA.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.00689-13.

doi:10.1128/JB.00689-13

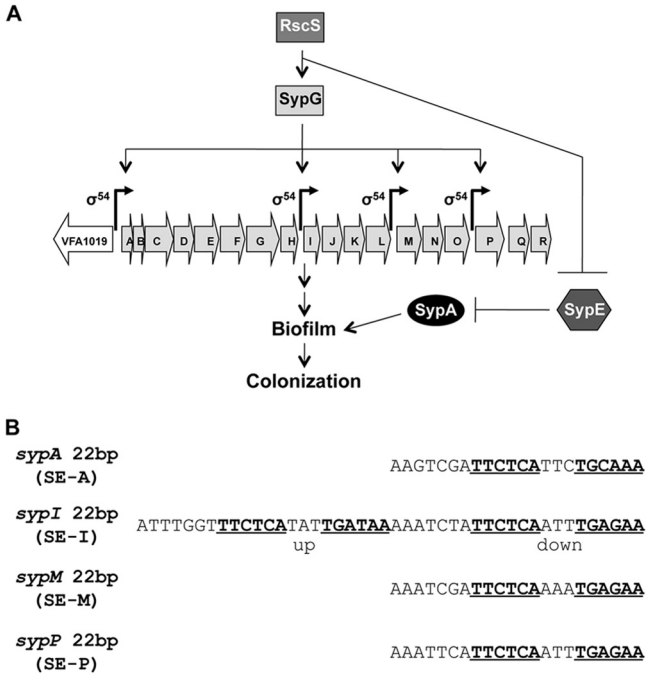


FIG 1 The Syp pathway and conserved *syp* enhancer sequences. (A) The *syp* locus contains 18 genes (gray arrows) involved in the regulation, production, and transport of a polysaccharide necessary for biofilm formation and host colonization. The *syp* locus contains four operons with σ^{54}-dependent promoters (bent arrows) located upstream of *sypA*, *sypI*, *sypM*, and *sypP*. The *syp* locus is transcriptionally controlled by the sensor kinase RscS (dark gray rectangle with white letters on the top of the diagram) and the response regulator (RR) SypG (gray rectangle, black letters). The RR SypE (gray hexagon) inhibits biofilm formation by inhibiting the activity of SypA (black oval), which functions downstream of *syp* transcription and is required for biofilm formation. Upon activation, RscS inactivates SypE's inhibitory activity, and SypA is able to promote biofilm formation. The gene directly upstream of the *syp* locus, *VF_A1019* (white block arrow), also has an SE sequence associated with it (not shown). (B) The nucleotide sequence of each *syp* enhancer is shown, with the inverted repeat underlined and in bold. Note that *sypI* contains two tandem SE sequences (SE-I-up and SE-I-down).

ence of these domains and the overall similarity of SypG to proteins such as NtrC and LuxO, SypG is predicted to function as a σ^{54}-dependent transcriptional activator. Transcription by the σ^{54}-containing holoenzyme differs from other holoenzymes in that this RNA polymerase recognizes sequences at −12 and −24 (rather than −10 and −35) and fails to initiate transcription without the help of a σ^{54}-dependent transcriptional activator (reviewed in references 21, 22, and 23). These activator proteins typically bind to an enhancer sequence exhibiting dyad symmetry (i.e., an inverted repeat [IR]) located 80 to 150 bp upstream of the promoter sequence. Upon activation (e.g., phosphorylation), the activator protein oligomerizes (typically into a hexamer), interacts with σ^{54} via DNA bending, and provides the energy (via ATP hydrolysis) necessary for RNA polymerase to promote open complex formation, leading to subsequent transcription of a particular gene(s).

The *syp* locus contains four promoters (upstream of *sypA*, *sypI*, *sypM*, and *sypP*), each with a predicted σ^{54} recognition sequence (Fig. 1A); of these, the promoters for *sypA*, *sypI*, and *sypM* have been confirmed via primer extension (10). In addition, we previously reported that multicopy expression of *sypG* induces expression of two *syp*::*lacZ* reporters (*sypD*::Tn*10lacZ* and *sypN*::Tn*10lacZ*) in a σ^{54}-dependent manner and that SypG is required for the RscS-induced transcription of *sypA* (10). From these studies, we concluded that SypG likely regulates transcription from at least two of the four putative σ^{54}-dependent *syp* promoters. Additionally, bioinformatic analyses of the *syp* promoters revealed a conserved 22-bp element present 50 to 90 bp upstream of each predicted σ^{54} recognition sequence (10), and in this work, we have identified a second conserved sequence upstream of *sypI* (Fig. 1B). This conserved region includes an IR consisting of two 6-bp half-sites separated by a 3-bp intervening sequence. Three of these sequences associated with the *syp* locus (*sypI*-down, *sypM*, and *sypP*) contain perfect IRs, while the upstream IR sequence of *sypI* and that associated with *sypA* diverge in the 3′ half-site (Fig. 1B; see also Fig. S2A in the supplemental material). The position and composition of this conserved element is consistent with those of other enhancer binding sequences bound by σ^{54}-dependent activators, such as NtrC (23). Thus, we hypothesized that this 22-bp sequence serves as a SypG binding site to facilitate transcriptional activation at each of the *syp* promoters (10). Although our experiments here did not address the ability of this sequence to serve as an enhancer *per se* (defined as activating transcription from a distance), it is, in fact, located at a distance from the promoter (see Fig. S3 in the supplemental material). Thus, for simplicity and consistency with the literature, we have designated these sequences *syp* enhancer (SE) sequences.

In this work, we assessed the role of the 22-bp SE sequence in *syp* transcription, biofilm formation, and host colonization. We demonstrated that the SE sequence is necessary for SypG-mediated *syp* transcription and biofilm formation and that SypG is able to recognize the SE sequence and bind to the *syp* promoter regions. In addition, SE sequences were found in other locations in the *V. fischeri* genome, and we demonstrated that a subset of these SE sequences was also recognized by SypG, suggesting that SypG regulates genes outside the *syp* locus. Finally, we found that the SE sequence was necessary for efficient host colonization. Overall, this work expands our understanding of the complex regulatory control necessary for *V. fischeri* to promote *syp* transcription and biofilm formation and underscores the importance of these processes for the biology of this organism.

MATERIALS AND METHODS

Strains and media. Strains used in this study are listed in Table 1. The parental *V. fischeri* strain used in this work was ES114, a strain isolated from *E. scolopes* (24). All derivatives were generated by conjugation, as previously described (25). *Escherichia coli* strains GT115 (Invivogen, San Diego, CA), TAM1 (Active Motif, Carlsbad, CA), TAM1 *pir* (Active Motif, Carlsbad, CA), CC118 *pir* (26), β3914 (27), and π3813 (27) were used for the purposes of cloning, plasmid maintenance, and conjugation. *V. fischeri* strains were grown in a complex medium (LBS [28] or seawater-tryptone [SWT] [10]), or in HEPES minimal medium (HMM) (29). *E. coli* strains were grown in LB (30) or brain heart infusion medium (BHI; Difco). As necessary, LB medium was supplemented to 0.3 mM with thymidine for *E. coli* strain π3813 or with diaminopimelic acid (DAP) for *E. coli* strain β3914. The following antibiotics were added to growth media as necessary, at the indicated final concentrations: chloramphenicol (Cm) at 2.5 μg/ml (*V. fischeri*) or 12.5 μg/ml (*E. coli*); erythromycin (Em) at 5 μg/ml (*V. fischeri*) or 150 μg/ml (*E. coli*); tetracycline (Tc) at 5 μg/ml in LBS or 30 μg/ml in HMM and SWT (*V. fischeri*) or 15 μg/ml (*E. coli*); ampicillin (Ap) at 100 μg/ml (*E. coli*).

Genetic and molecular techniques. Plasmids and primers used in this study are listed in Table S1 and Table S2, respectively, in the supplemental material. All plasmids were constructed using standard molecular biology

TABLE 1 Strains used in this study

Strain	Genotype or description	Derivation[a]	Reference
ES114	Wild-type *V. fischeri*	NA	24
KV1787	Δ*sypG*	NA	16
KV3246	attTn*7*::P_{sypA}-*lacZ* Em^r (FL + SE-A)	NA	19
KV3628	attTn*7*::P_{sypI}-*lacZ* Em^r (no SE-I-up)	pEAH120	This study
KV3629	attTn*7*::P_{sypI}-*lacZ* Em^r (FL + SE-I)	pEAH121	This study
KV3631	attTn*7*::P_{sypM}-*lacZ* Em^r (no SE-M)	pEAH123	This study
KV3632	attTn*7*::P_{sypM}-*lacZ* Em^r (FL + SE-M)	pEAH124	This study
KV3636	attTn*7*::P_{sypA}-*lacZ* Em^r (no SE-A)	pEAH128	This study
KV4522	attTn*7*::P_{sypA}-*lacZ* Em^r (IG + SE-A)	pKV318	This study
KV4523	attTn*7*::P_{sypI}-*lacZ* Em^r (IG + SE-I)	pKV312	This study
KV4524	attTn*7*::P_{sypM}-*lacZ* Em^r (IG + SE-M)	pKV316	This study
KV4525	attTn*7*::P_{sypP}-*lacZ* Em^r (IG + SE-P)	pKV311	This study
KV4526	attTn*7*::P_{sypP}-*lacZ* Em^r (FL + SE-P)	pKV313	This study
KV4527	attTn*7*::P_{sypP}-*lacZ* Em^r (no SE-P)	pKV314	This study
KV4970	ΔSE-A	pMSM25	This study
KV6716	ΔSE-A::SE-A	pKV461	This study

[a] The plasmids used to derive the strains generated in this study. NA, not applicable.

techniques, with restriction and modification enzymes obtained from New England BioLabs (Beverly, MA) or Thermo Fisher (Pittsburgh, PA). In some cases where PCR was used to generate DNA fragments, the PCR cloning vector pJET1.2 (Fisher Scientific, Pittsburgh, PA) was used as an intermediate vector prior to cloning into the final vector.

Construction of *syp*-promoter *lacZ*-fusions and β-galactosidase assays. To evaluate promoter activity of the *sypA*, *sypI*, *sypM*, and *sypP* promoters, we fused the *syp* promoter regions (containing or lacking SE sequences) upstream of a promoterless *lacZ* gene. The reporter fusions were introduced into the Tn*7* delivery plasmid pEVS107 (31). Tetraparental matings with wild-type *V. fischeri*, *E. coli* carrying pEVS104 (32), *E. coli* carrying the pEVS107 derivatives, and *E. coli* carrying the Tn*7* transposase plasmid pUX-BF13 (33) were used to insert the DNA at the Tn*7* site in the chromosome of *V. fischeri*, as previously described (31, 32). This approach permitted us to assay the transcription from *syp* promoters present in single copy in the chromosome, in the presence and absence of a plasmid, pEAH73, that expressed *sypG* from the *lac* promoter. To assay β-galactosidase activity, cultures of the *lacZ* reporter strains (Table 1; see also Fig. S3 in the supplemental material) carrying the vector control plasmid pKV69 or the *sypG* plasmid pEAH73 were grown in HMM containing Tc at 28°C with shaking. Samples (50 μl) were collected at 24 h, and 50 μl of Pierce β-galactosidase assay reagent (Pierce Biotechnology, Rockford, IL) was added to each sample. Measurements were taken in a microtiter dish by using an ELx800 absorbance microplate reader (BioTek, Winooski, VT) with the appropriate settings. As a measure of *syp* transcription, β-galactosidase activity was determined as previously described (34).

Development of a biofilm-based assay of SypG activity. As an indirect measure of promoter activity, we developed an assay based on the observation that a strain that contained two plasmids, one that expressed *sypG* (from the *lac* promoter) and one that expressed *sypA*, but not one or the other plasmid, could produce a biofilm (19). When cultures were spotted onto LBS agar plates (containing Cm and Tc), the resulting colony developed a wrinkled morphology (indicative of biofilm formation) or remained smooth (indicative of the lack of biofilm formation). In that published study, *sypA* was under the control of both its native promoter and the vector-based *lac* promoter. Here, to understand the role of the SE sequences in promoting *sypA* transcription, we cloned the *sypA* gene under the control of its native promoter region, either lacking or containing the SE sequence upstream of *sypA* (SE-A), in an orientation opposite to that of the *lac* promoter contained within pKV69 (35) to generate pKV300 and pKV301, respectively. Derivatives of pKV300, which lacks SE-A but instead contains a nonnative ApaI restriction site (GGGCC/C, where the slash represents the cleavage site), were generated as follows: pKV300 was digested with ApaI and NcoI, for which a site is located within the Cm^r gene of the cloning vector. Sets of complementary primers were generated that contained SE sequences as described in the text and in Table S2 in the supplemental material and, on the ends, nonnative sequences complementary to the ApaI and NcoI restriction sites (see Fig. S5 in the supplemental material). These sets of primers were annealed and ligated into ApaI/NcoI-digested pKV300. Each resulting plasmid contained SE sequences and a 5-bp nonnative "scar" (GGCCC) as a result of ligation to the ApaI site (see Fig. S5). All derivatives were compared to pVAR58, which was similarly derived from pKV300 by the insertion of the wild-type SE-A at ApaI/NcoI and also contained the nonnative "scar" (see Fig. S5). The pKV300 derivatives all lost Cm^r but retained Tc^r. These Tc^r plasmids were fully compatible with pARM9 (19), the pVSV105-based Cm^r plasmid that carries the *sypG* gene. Both pARM9 and the pKV300 derivatives were introduced into wild-type strain ES114 via conjugation and selection with Cm and Tc. As a negative control, pKV300 was digested with ApaI and NcoI, the overhangs were filled in, and the plasmid was self-ligated to obtain pKV465. To assess the consequence of the insertion of the nonnative "scar" sequences, pVAR57 was generated from pKV301 by deletion of the sequences between ApaI and NcoI; the ApaI site in pKV301 is on the 5′ end of SE-A, and thus there was no "scar" between SE-A and the remainder of the promoter region in front of *sypA*. Additional constructs were generated in the same way, with the insertion of specific SE sequences into the ApaI/NcoI sites of pKV300.

Wrinkled colony assay. Wrinkled colony assays were performed as described previously (36), with the exception that the strains were cultured in and grown on LBS medium containing Tc and Cm. Wrinkled colony development was monitored over time, and representative images are shown below for the indicated time points. Although the specific endpoint of the assay on the third day varied between 68 and 72 h, there were no significant differences observed in wrinkled colony development within that time frame. We classified strains that carried *sypA* constructs with mutant SE-A sequences into three groups: similar to the positive control, similar to the negative control, or intermediate, based on the timing and extent of wrinkled colony development.

Bioinformatics. The IR for the *syp* enhancer sequence (TTCTCANN NTGMDWN, where N represents any nucleotide, M represents A or C, W represents A or T, and D represents A, T, or G) was found by searching the 500-bp regions upstream of promoter-proximal open reading frames in the *V. fischeri* genome with the genome-scale DNA pattern search in the Regulatory Sequence Analysis Tools (RSAT) package (http://rsat.ulb.ac.be) (37–39).

Chromatin immunoprecipitation (ChIP) analysis. Plasmids containing *sypG*-FLAG (pVAR45) alone or with either pVAR57 (*sypA* with the native SE-A sequence) or pKV465 (*sypA* without the SE-A sequence) were introduced into Δ*sypG V. fischeri* cells (KV1787). Bacterial strains were cultured in HMM containing Tc (for pVAR45) or Tc and Cm (for pVAR45 and pKV465 or pVAR57) at 28°C for 24 h with shaking. Cells (~0.05 g) were harvested by centrifugation (13,000 × *g* for 10 min) and washed twice in 1 ml of phosphate-buffered saline (PBS). Samples were subsequently prepared and immunoprecipitated using the Dynabeads co-immunoprecipitation kit (Invitrogen, Grand Island, NY). Cell samples were resuspended and lysed in 900 μl of extraction buffer (EB; 1× immunoprecipitation [IP] buffer, 100 mM NaCl, 1 mM dithiothreitol; Invitrogen, Grand Island, NY). The samples were sonicated using a Branson sonifier 250 (Branson Ultrasonics, Danbury, CT) on ice for a total of six 15-s cycles per sample, generating DNA fragments averaging 500 bp. Rabbit anti-FLAG antibodies (25 mg; Sigma-Aldrich, St. Louis, MO) were coupled to magnetic Dynabeads (5 mg; Invitrogen, Grand Island, NY) according the manufacturer's protocol. As a negative control, Dynabeads (5 mg) were coupled with nonspecific mouse anti-rabbit IgG antibodies (5 mg; Cell Signaling, Danvers, MA). For the immunoprecipitation, antibody-coupled beads were incubated with 450 μl of whole-cell extracts at 4°C with rocking for 30 min. Beads were collected and washed according to the manufacturer's protocol. Protein-DNA complexes were eluted in 60 μl elution buffer (Invitrogen, Grand Island, NY). DNA was recovered

using the MiniElute reaction cleanup kit (Qiagen, Valencia, CA) and resuspended in 50 μl of distilled water. Quantitative real-time PCR (qRT-PCR) was performed using an Opticon 2 real-time detector (Bio-Rad, Hercules, CA). PCR cycles entailed denaturation at 94°C (30 s) and annealing at 59°C (30 s); no extension phase was necessary. PCR data were analyzed using Opticon Monitor 3 software (Bio-Rad, Hercules, CA). Modification levels were calculated as the fold enrichment via the following equations (where C_T is the threshold cycle value and normalized input standardization [NIS] and mock represent the standardization of total DNA in cell culture used for immunoprecipitation and background negative controls, respectively): ΔC_T for the normalized ChIP value = [C_T for ChIP − C_T for input − $\log_2$(input dilution factor)]; the percent input = $2^{-\Delta CT}$ for the normalized ChIP value; $\Delta\Delta C_T$ for ChIP/NIS = ΔC_T for the normalized ChIP value − ΔC_T for NIS/mock; finally, the fold enrichment = $2^{-\Delta\Delta CT}$ for ChIP/NIS (ChIP/NIS is the standardization of immunoprecipitation to the total DNA in the cell lysate from each independent experiment, and NIS/mock is the standardization of all immunoprecipitations to the relative background across samples and independent experiments).

Generation of an SE-A deletion mutant. Deletion of the genomic SE-A sequence was accomplished by using a modified approach of that described by LeRoux et al. (27). Briefly, ~600 bp of sequences upstream and downstream of SE-A were amplified using PCR and fused as indicated in Fig. S6 in the supplemental material. For deletion of SE-A (ΔSE-A), the junction contained a nonnative ApaI site but no other extraneous sequences. This deletion sequence was cloned into pKV363 (11), a suicide plasmid that contains the arabinose-inducible *ccdB* toxin gene. The plasmid was introduced into *V. fischeri* cells, and cells that had undergone a single-crossover event were selected with Cm. Subsequently, cells that had undergone a second recombination event were selected by the addition of arabinose. Cells that contained the desired deletion were identified using PCR. The same approach was used to repair the ΔSE-A mutant, KV4970, by insertion of SE-A back into the ApaI site (see Fig. S6). The resulting strain, KV6716, contained a 5-bp "scar" sequence in addition to the native SE-A sequences.

Colonization assay. Experiments involving *E. scolopes* animals were carried out via approaches described in an Animal Component of Research Protocol (ACORP) approved by Loyola University's Institutional Animal Care and Use Committee (IACUC; LU 107314, 201297). To perform colonization assays, juvenile *E. scolopes* squid were placed in artificial seawater (Instant Ocean; Aquarium Systems, Mentor, OH) and incubated with 700 to 1,500 *V. fischeri* cells per ml of seawater for 3 to 4 h. The squid were then washed in artificial seawater and placed into fresh (*V. fischeri*-free) artificial seawater, and colonization was allowed to proceed for 13 to 16 h postwash (total experimental time, 16 to 20 h). The squid were then homogenized to release the contents of their light organs, and the homogenates were diluted and plated onto SWT to determine the number of CFU per squid.

RESULTS

SypG-mediated induction of *syp* promoters requires a conserved 22-bp sequence. Evidence gathered to date suggest that SypG, a putative σ^{54}-dependent transcriptional activator, serves as the direct transcriptional activator of the *syp* locus (Fig. 1) (10, 18, 40). However, whether SypG activates all four *syp* promoters (*sypA*, *sypI*, *sypM*, and *sypP*) and whether its activity depends on the conserved SE sequences located upstream of these promoters remain unknown. To answer these questions, we generated *lacZ* transcriptional reporter fusions to sequences upstream of the *sypA*, *sypI*, *sypM*, and *sypP* genes. These constructs included the entire intergenic (IG) sequence and, in some cases, sequences in the adjacent upstream gene (see Fig. S3 in the supplemental material). The fusions were placed in single copy in the chromosome of wild-type (ES114) *V. fischeri* at a benign site distal to the *syp* locus (the Tn*7* site). We then introduced either a multicopy SypG expression vector (p*sypG*) or the vector control (VC). SypG expressed from p*sypG* induced a substantial increase in β-galactosidase activity of each of the reporters (Fig. 2, IG set). These data thus established an important role for SypG in inducing *syp* transcription at each of the *syp* promoters. They also indicated that a SypG-responsive promoter exists upstream of *sypP*, which previously had been unconfirmed.

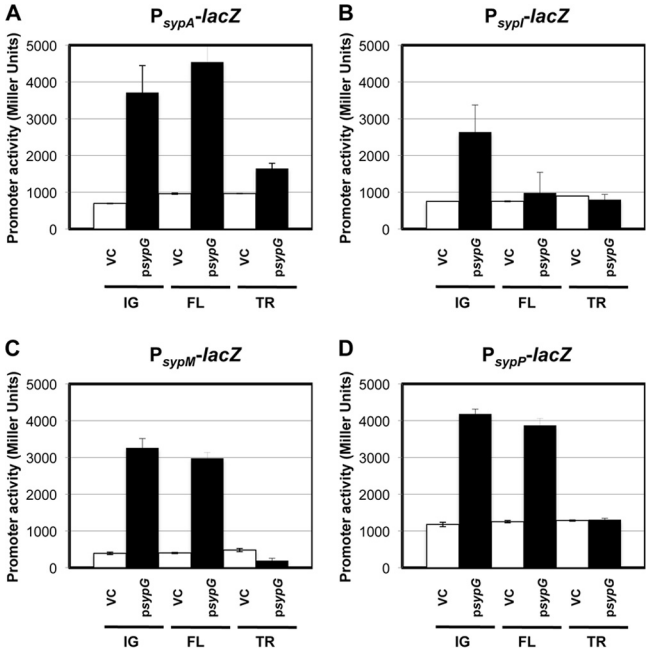


FIG 2 Activity of *sypA*, *sypI*, *sypM*, and *sypP* reporter fusions. pEAH73 (p*sypG*)-containing derivatives of the indicated strains were grown in HMM containing Tc for 24 h and then harvested for a β-galactosidase assay to measure promoter activity (reported in Miller units). IG, FL, and TR refer to intergenic, full-length, and truncated constructs, respectively, as defined in the text and also shown in Fig. S3 in the supplemental material. (A) P_{sypA}-*lacZ* fusions with IG (KV4522), FL (KV3246), or TR (KV3636); (B) P_{sypI}-*lacZ* fusions with IG (KV4523), FL (KV3629), or TR (KV3628); (C) P_{sypM}-*lacZ* fusions with IG (KV4524), FL (KV3632), or TR (KV3631); (D) P_{sypP}-*lacZ* fusions with IG (KV4525), FL (KV4526), or TR (KV4527). Error bars represent standard deviations. These graphs are representative of at least two independent experiments.

To begin to delimit the regulatory region important for SypG-mediated *syp* induction, we made additional reporter derivatives with 5′ truncations. Specifically, we generated a set of constructs, designated FL (full length), that began at the previously documented SE sequence (10), and a second set, designed TR (truncation), that began immediately downstream of that sequence (see Fig. S3 in the supplemental material). For strains containing the FL constructs, with one exception, we saw similar levels of β-galactosidase activity, indicating that the sequences included in these shorter constructs were sufficient for the observed SypG-mediated induction (Fig. 2, FL set). The exception was *sypI*, and upon closer inspection we noticed that the *sypI* promoter region contained two sets of the conserved SE sequence (Fig. 1B); the upstream SE sequence was truncated in the *sypI* FL construct (see Fig. S3). These data suggest either that both sets of the sequence are required for optimal transcriptional activation from the *sypI* promoter or that the upstream sequence is the more important of the two. For strains containing the TR constructs, we found that the

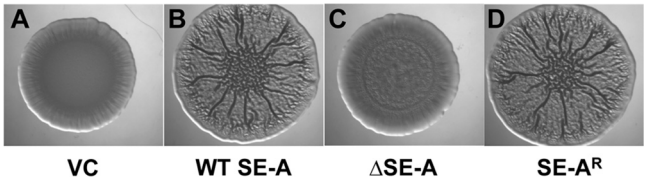


FIG 3 Biofilm formation induced by cooverexpression of *sypG* and *sypA*. Images show the wrinkled colony morphologies of spotted cultures at 68 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and either vector control (VC) or a *sypA* plasmid derivative that contained or lacked the *sypA* enhancer (SE-A), as indicated. (A) VC (pKV282); (B) WT SE-A (pVAR57); (C) ΔSE-A (pKV465); (D) SE-A^R (pVAR58). Data are representative of at least three independent experiments.

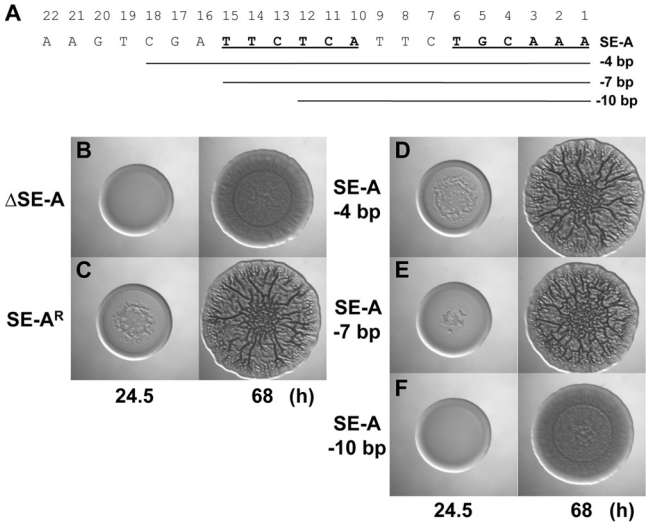


FIG 4 Effects of truncating SE-A in the *sypA*-dependent biofilm assay. (A) The full-length and 5′ truncation derivatives of SE-A generated in this study. (B to E) To assess the importance of the SE-A sequence, wrinkled colony formation was monitored up to 68 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and a *sypA* plasmid derivative containing the indicated full-length or truncated SE-A sequence, as follows: (B) ΔSE-A (pKV465); (C) SE-A^R (pVAR58); (D) SE-A −4 bp (pSLN4); (E) SE-A −7 bp (pSLN5); (F) SE-A −10 bp (pSLN7). Data are representative of at least three independent experiments.

ability of SypG to induce transcription was severely diminished or abolished (Fig. 2, TR sets). Thus, SypG requires the SE sequences to induce *syp* transcription.

Development of a biofilm assay for *sypA* expression. The β-galactosidase reporter experiments confirmed our prediction that the SE sequences located upstream of the four *syp* promoters were necessary for SypG-mediated transcription, but they did not address the requirement for specific nucleotides. To assess which nucleotides are critical for *syp* activation, we developed a biofilm-based assay of SypG activity. We used this assay instead of the β-galactosidase assay due to technical difficulties associated with generating numerous strains with single-copy point mutation derivatives of the *lacZ* fusion constructs. The biofilm assay provides an indirect measure of SypG activity based on its ability to promote biofilm formation. Previous studies demonstrated that expression of SypG from a multicopy plasmid fails to promote biofilm formation, even though it induces *syp* transcription (18). However, biofilm formation is induced when the *sypG* plasmid is introduced along with a second plasmid that contains the SypG-controlled gene *sypA* (along with its upstream regulatory region) (see Fig. S1 in the supplemental material) (19). We hypothesized that, under these conditions, SypG induces expression of high levels of SypA such that some SypA escapes the inhibitory phosphorylation mediated by SypE, thereby permitting biofilm formation to proceed.

Here, we tested one part of this hypothesis, that biofilms form due to the ability of SypG to activate the *sypA* promoter present on the multicopy plasmid. First, we asked whether SypG expressed from a multicopy plasmid increased the levels of SypA protein, and we found that it did (see Fig. S4 in the supplemental material). As predicted from the β-galactosidase assay results, this increase in SypA protein levels depended on the presence of an intact *sypA* enhancer (SE-A) sequence (see Fig. S4).

Next, we asked whether biofilm formation induced upon introduction of the *sypG* and *sypA* plasmids similarly depended on the presence of the SE-A sequence. To evaluate biofilm formation, we spotted a culture of cells onto a plate and monitored the development of the spot from a smooth to wrinkled morphology over time. We anticipated that the decreased amount of SypA due to the loss of SE-A would result in diminished or delayed wrinkled colony development. As previously demonstrated (18, 19), cells that contained either the *sypG* plasmid or the *sypA* plasmid alone formed smooth colonies (Fig. 3A and data not shown), while cells that contained both plasmids formed wrinkled colonies (Fig. 3B). In contrast, cells that contained the *sypG* plasmid along with a *sypA* plasmid lacking SE-A (ΔSE-A) were unable to form wrinkled colonies within the same time frame as the positive control (compare Fig. 3B and C). Although colonies formed by this strain did eventually wrinkle, the extent of wrinkling was minimal even at the late time point of this data set (Fig. 3). This strain thus served as the negative control throughout the remainder of these experiments.

To verify that wrinkled colony formation depended on SE-A, we restored SE-A to the ΔSE-A construct at an engineered ApaI restriction site (SE-A-restored, or SE-A^R) (see Fig. S5 in the supplemental material). Cells containing this modified plasmid formed wrinkled colonies with the same timing and intensity as the native SE-A-containing control strain (Fig. 3B and D and data not shown). Although restoration of SE-A resulted in the additional insertion of 5 bp of nonnative sequence between the enhancer and the promoter, we could detect no consistent differences in biofilm formation by cells carrying the restored or native SE-A constructs. Because the SE-A^R construct most closely matched other constructs that we subsequently generated, we used it as a positive control throughout the remainder of these experiments.

Together, our data indicated that SE-A plays an important role in biofilm formation. Because SE-A was also necessary for *sypA-lacZ* reporter expression and induction of SypA protein production, we conclude that biofilm formation in this assay depends on the ability of SypG to induce *sypA* transcription and, as a result, promote SypA protein production. This assay thus provides a means to perform an in-depth analysis of the sequences necessary for SypG-mediated *syp* transcription and biofilm formation.

Identification of nucleotides necessary for SypG-mediated biofilm induction. We next constructed *sypA* plasmids with modifications to the SE-A sequence. First, we generated nested deletion derivatives (Fig. 4A). Loss of the four 5′-most nucleotides (−4 bp) did not affect the timing of wrinkled colony formation,

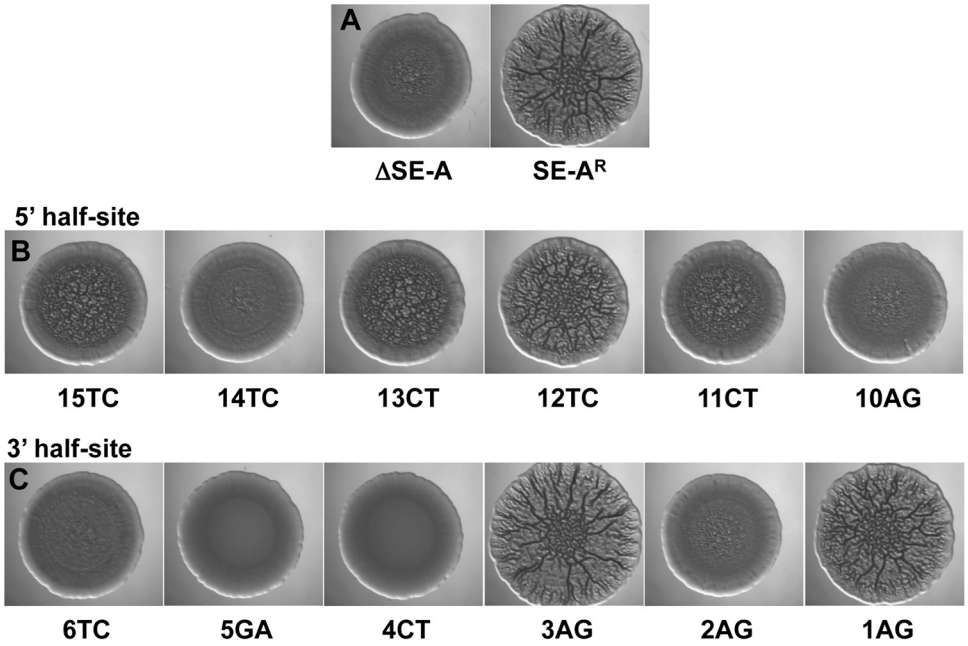


FIG 5 Effects of SE-A transition mutations in the *sypA*-dependent biofilm assay. The impacts of transition mutations within the SE-A sequence were assessed in a *sypA*-dependent biofilm assay. Images show the wrinkled colony morphologies of spotted cultures at 71.5 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and a *sypA* plasmid derivative containing the indicated point mutation. (A) controls ΔSE-A (pKV465) and SE-A^R (pVAR58); (B) 5′ half-site with 15TC (pKV444), 14TC (pKV445), 13CT (pKV446), 12TC (pKV447), 11CT (pKV448), or 10AG (pKV449); (C) 3′ half-site with 6TC (pKV450), 5GA (pKV451), 4CT (pKV452), 3AG (pKV453), 2AG (pKV454), or 1AG (pKV455). Data are representative of at least three independent experiments. Refer to Fig. 4A for the numbering scheme for the individual nucleotides within SE-A.

suggesting that these sequences are not important for *sypA* induction (Fig. 4, compare panels C and D). However, loss of the first seven nucleotides (−7 bp) caused a minor delay in biofilm formation, resulting in a slightly smaller wrinkled colony at later time points, indicating that *sypA* induction is likely diminished (Fig. 4, compare panels C and E). When SE-A was truncated even further to remove a portion of the 5′ half-site of the IR (−10 bp), biofilm formation was reduced to that of the negative control (Fig. 4, compare panels B and F). Thus, the presence of an intact IR is sufficient to promote biofilm formation, but the presence of upstream sequences permits the best induction of biofilm formation.

To further understand the requirements of the SE-A sequence for biofilm formation, we generated transition mutations (exchanging A for G and T for C, and vice versa) at each of the nucleotides within the 5′ and 3′ half-sites of the IR region (Fig. 4A, underlined regions). Only two substitutions resulted in wrinkled colony formation that was comparable to or better than that of the control strain: 3AG and 1AG (Fig. 5, compare panels A and C). In particular, the strain carrying the 3AG substitution in SE-A consistently exhibited wrinkled colony formation that was faster than the positive control (data not shown). Of note, 3AG is a change that makes SE-A more closely match a perfect IR sequence; in contrast to the other SE sequences (SE-I-up, SE-M, and SE-P), which contain perfect IR sequences (TTCTCA-N3-TG**AG**AA), SE-A contains an imperfect IR (TTCTCA-N3-TG**CA**AA) (Fig. 1).

We categorized the biofilm phenotypes of the remaining strains into two classes: (i) those with diminished biofilm formation relative to the positive control (intermediate phenotype) and (ii) those that were biofilm defective (indistinguishable from the negative control) (Fig. 5). The first category included those with substitutions within the highly conserved 5′ half-site of SE-A (15TC, 13CT, 12TC, and 11CT). The second category included strains with substitutions within both the 3′ half-site (6TC, 5GA, 4CT, and 2AG) and the 5′ half-site (14TC and 10AG). We interpreted the reduced and defective biofilm formation of these strains to mean that SypG has reduced or defective binding to the SE sequences and thus fails to promote the same levels of *sypA* transcription as the positive control. Thus, we concluded that 10 of the 12 nucleotides of the IR play important roles in promoting SypG-mediated biofilm formation in this assay.

Because our conclusions were based on a single base change, we expanded our screen for a subset of these nucleotides (positions 2, 4, 5, and 14) by generating the other two possible base changes (transversions). For position 2A, a change to T was permissive to promote biofilm formation, but a change to C was not (Fig. 6B). Similarly, for 4C, a change to A, which brings the 3′ half-site closer to perfect, was permissive, but a change to G was not (Fig. 6C). In contrast, for 5G and 14T, all substitutions resulted in biofilm formation that was indistinguishable from that of the negative control (Fig. 6, compare panel A with D and E), indicating that the original bases at these positions are indeed critical determinants for *sypA* expression. These data thus indicate that while some changes are more permissive than others, most of the nucleotides within the IR play key roles in SypG-induced biofilm formation.

Evaluation of other *syp* enhancers. We next wondered whether replacing the SE-A sequence with other SE sequences (i.e., those associated with *sypI*, *sypM*, and *sypP*) would similarly promote biofilm formation in our assay. Indeed, insertion of either SE-M or SE-P promoted biofilm formation in the context of

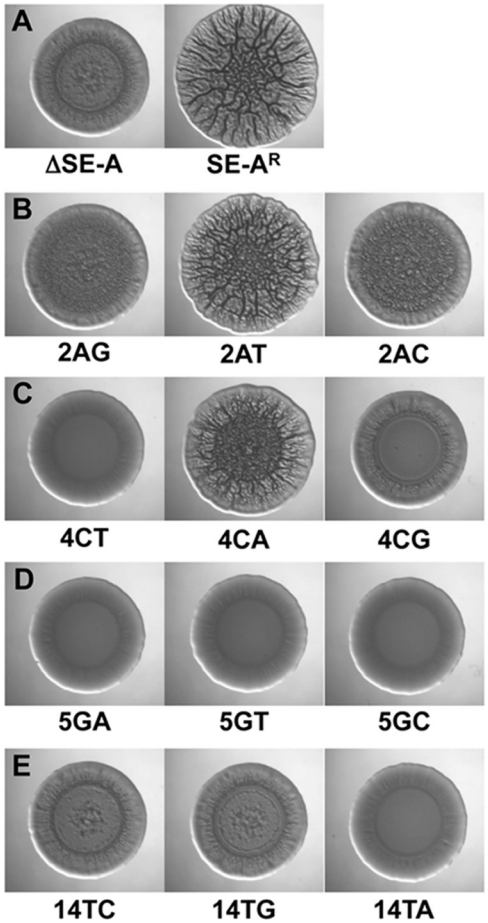


FIG 6 Effects of transversion point mutations in the *sypA*-dependent biofilm assay. The impacts of SE-A transversion mutations were assessed in a *sypA*-dependent biofilm assay. Images show the wrinkled colony morphologies of spotted cultures at 72 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and a *sypA* plasmid derivative containing the indicated point mutation: (A) controls ΔSE-A (pKV465) and SE-A^R (pVAR58); (B) 2AG (pKV454), 2AT (pVAR84), or 2AC (pVAR83); (C) 4CT (pKV452), 4CA (pKJW11), or 4CG (pKJW10); (D) 5GA (pKV451), 5GT (pKJW8), or 5GC (pKJW9); (E) 14TC (pKJW4), 14TG (pKJW5), or 14TA (pKV452). The original transition mutation constructs are included here for reference. Data are representative of at least three independent experiments. Refer to Fig. 4A for the numbering scheme for the individual nucleotides within SE-A.

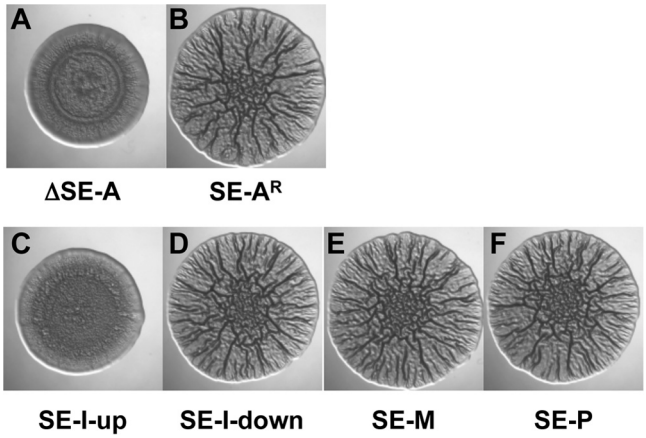


FIG 7 Impacts of substitution of SE-I-up, SE-I-down, SE-M, and SE-P for SE-A in the *sypA*-dependent biofilm assay. The ability of SypG to recognize the other SE sequences was assessed in the *sypA*-dependent biofilm assay. Images show the wrinkled colony morphologies of spotted cultures at 72 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and a *sypA* plasmid derivative with the indicated SE sequence in place of SE-A: (A) ΔSE-A (pKV465); (B) SE-A^R (pVAR58); (C) SE-I-up (pKV437); (D) SE-I-down (pKV438); (E) SE-M (pKV439); (F) SE-P (pKV440). Data are representative of at least three independent experiments.

this assay (Fig. 7). Interestingly, when we evaluated the two *sypI* enhancers (SE-I-up and SE-I-down [Fig. 1B]), we found that SE-I-down promoted biofilm formation, whereas SE-I-up did not (Fig. 7C and D). Thus, in the context of the *sypA* promoter region, the upstream *sypI* enhancer (SE-I-up), which is further removed from the IR consensus, is insufficient to promote biofilm formation. These results stand somewhat in contrast to what we observed for transcription from the *sypI* promoter, in which loss of the upstream, less-conserved SE sequence led to a loss of transcription. We address this apparent contradiction in the Discussion section. Overall, the SE sequences associated with the *syp* promoters were able to promote biofilm formation in our assay, indicating that SypG is able to recognize these sequences.

VF_A1019, VF_A0120, and VF_A0550 contain *syp* enhancer sequences. Equipped with the information about the specific nucleotides critical for SypG-mediated activation, we next examined the *V. fischeri* genome for potential SypG binding sites. With the Regulatory Sequence Analysis Tools (RSAT) program (37–39), we identified a number of potential binding sites (see Table S3 in the supplemental material). We predicted that a member of the SypG regulon would also require a σ^{54} binding site, and we found that the promoter regions of only three genes (other than *syp*) contained both a putative SypG binding site and a putative downstream σ^{54} binding site. One potential SypG binding sequence was located immediately adjacent to the *syp* locus in front of the divergently transcribed gene *VF_A1019* (Fig. 1A and 8A). The other two were located upstream of *VF_A0120* and *VF_A0550* (Fig. 8A). Upon closer inspection of the predicted SE sequences associated with these genes, we noted that the 5′ half-site was an exact match to the SE sequences, but the 3′ half-site deviated from the IR sequence (Fig. 8A; see also Fig. S2B in the supplemental material). Compared to the 3′ half-site of SE-A, the 3′ half-site associated with *VF_A1019* differed at a single position, position 1 (change from A to T) (Fig. 8A), a position that does not appear important for SypG's recognition of SE-A (Fig. 5C). Similarly, the 3′ half-site associated with *VF_A0120* differed at nucleotide 4 (C to A), a change that was permissive in the context of SE-A (Fig. 6C). Finally, the 3′ half-site associated with *VF_A0550* had nucleotide changes at positions 1 to 4; some substitutions at these positions were permissive for biofilm formation in the context of SE-A (Fig. 5C and 6B and C). Therefore, we determined whether SypG could recognize these SE sequences by replacing SE-A upstream of *sypA* with the SE sequence from *VF_A1019* (SE-1019), *VF_A0120* (SE-0120), or *VF_A0550* (SE-0550) in our plasmid-based biofilm assay. We found that introduction of any of these SE sequences in place of SE-A induced biofilm formation to a level above that of the negative control (Fig. 8A), and indeed, to a degree indistinguishable from the positive control (Fig. 8B), suggesting that SypG can recognize these SE sequences despite their differences. Taken together, these data suggest that *VF_A1019*, *VF_A0120*, and *VF_A0550* comprise part of the SypG regulon.

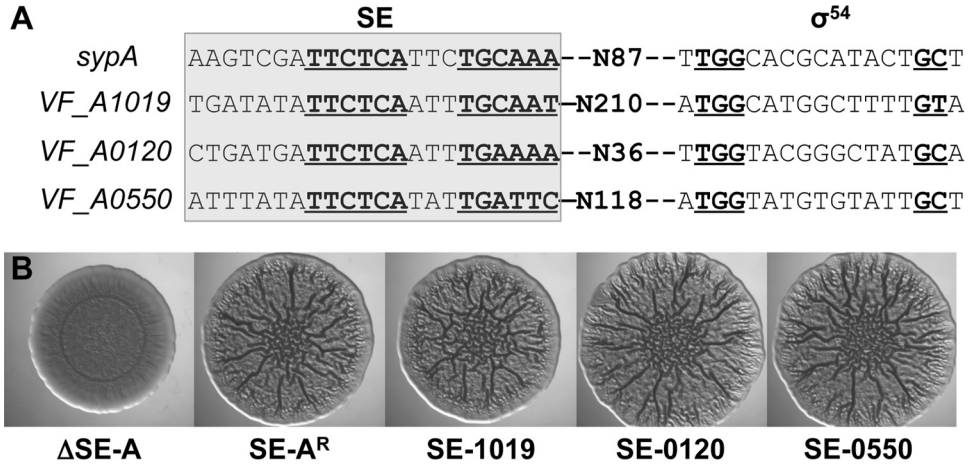


FIG 8 Effects of SE-1019, SE-0120, and SE-0550 in the *sypA*-dependent biofilm assay. (A) The SE sequences (shaded) and predicted σ^{54} binding sites of *sypA* (SE-A), *VF_A1019*, *VF_A0120*, and *VF_A0550* are shown. (B) The ability of SypG to recognize the putative SE sequences associated with *VF_A1019*, *VF_A0120*, and *VF_A0550* was assessed in a *sypA*-dependent biofilm assay. Images show the wrinkled colony morphologies of spotted cultures at 68 h postspotting for ES114 containing the *sypG* overexpression plasmid pARM9 and a *sypA* plasmid derivative containing the indicated SE sequence in place of SE-A: ΔSE-A (pKV465), SE-A^R (pVAR58), SE-1019 (pKV463), SE-0120 (pKV462), or SE-0550 (pKV464). Data are representative of at least three independent experiments.

SypG binds to *syp* regulatory regions. Our data to date are consistent with the hypothesis that SypG binds to SE sequences to activate *syp* transcription. However, we have not directly demonstrated that SypG can bind to the *syp* promoter regions. Attempts to purify SypG for use in electrophoretic mobility shift assays were unsuccessful, as purified SypG precipitated out of solution (data not shown). Therefore, a ChIP approach commonly utilized for eukaryotic studies was used to assess SypG binding to SE regions *in vivo*.

We assessed binding of SypG by using a Δ*sypG* strain that expressed a FLAG epitope-tagged SypG from a plasmid (pVAR45). We then used anti-FLAG antibodies to immunoprecipitate SypG-FLAG and associated DNA from cell lysates. DNA isolated from the immunoprecipitation was then subjected to qRT-PCR with primer sets specific to the regions of interest, including the *sypA*, *sypI*, *sypM*, and *sypP* promoter regions, with an internal *syp* control (internal to *sypF*) (Fig. 9A). The qRT-PCR results demonstrated that SypG was enriched at SE-A, 2.7-fold ± 0.8-fold (mean ± standard deviation [SD]) above the background negative control IgG immunoprecipitation, which was set to 1 (Fig. 9B). Similar results were observed for the promoters of *sypI*, *sypM*, and *sypP*, where SypG was enriched at these regions 3.5- to 4.9-fold above background levels (Fig. 9B). As an internal *syp* control, we quantified the enrichment of SypG within the *sypF* region, which was not predicted to bind SypG. The level of internal *sypF* that coimmunoprecipitated with SypG-FLAG was not enriched significantly over the IgG negative control ChIP and was significantly lower than that of any of the *syp* promoters that contained putative SE sequences (Fig. 9B). We next tested whether SypG was associated with the promoter of *VF_A0120*, one of the other genes with a putative SE sequence. We found that SypG was indeed enriched 3.2-fold ± 0.4-fold over the IgG control at the *VF_A0120* promoter (Fig. 9B). These data support the hypothesis that SypG binds to *syp* promoter regions to activate *syp* transcription.

We then expanded this approach to confirm our findings from our biofilm assay. First, we coexpressed *sypG*-FLAG and a plasmid that contained the *sypA* promoter with the wild-type SE-A sequence in the Δ*sypG* strain. The ChIP results confirmed that SypG-FLAG bound equally to the native chromosomal SE-A and the *sypA* promoter containing SE-A on the plasmid, as indicated by equal enrichment at both sites (Fig. 9C, SE-A). We next evaluated a strain containing a plasmid with the *sypA* promoter that lacked the SE-A sequence (ΔSE-A). SypG was enriched at the native chromosomal *sypA* promoter to the same degree as the other

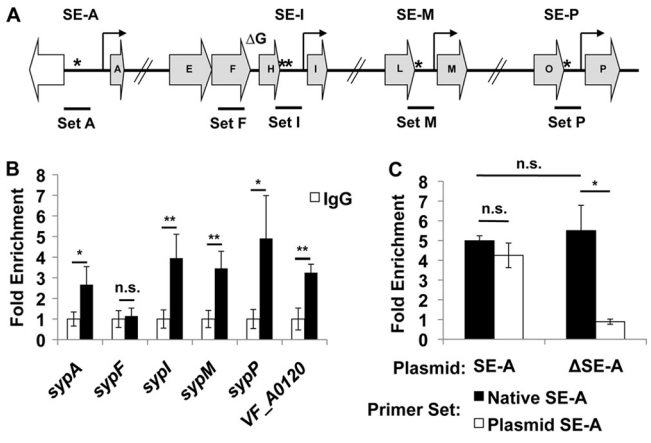


FIG 9 ChIP analysis of SypG binding at the *syp* promoter regions. (A) The *syp* locus (gray arrows), with SE sequences (asterisks) shown upstream of promoters (bent arrows). The regions used for ChIP analyses are indicated below the promoter. (B) ChIP analysis of a Δ*sypG* strain containing pVAR45 (SypG-FLAG). qRT-PCR was used to quantify the amount of DNA that coimmunoprecipitated with SypG-FLAG. SypG binding to *syp* promoter regions or a site predicted to not bind SypG (*sypF*) is presented as the fold enrichment. The graph shows the averages of 4 independent experiments ± the SD; the IgG background was set to 1. (C) ChIP analysis with a Δ*sypG* strain cooverexpressing pVAR45 (SypG-FLAG) and either pVAR57 (+SE-A) or pKV465 (ΔSE-A). qRT-PCR was used to quantify the amount of DNA that coimmunoprecipitated with SypG-FLAG. Binding of SypG to native SE-A on the chromosome (black bars) or on the plasmid (+SE-A or ΔSE-A; white bars) is presented as the fold enrichment. A representative graph from 1 of 3 independent experiments is shown (with SD). *, $P < 0.05$; **, $P < 0.01$.

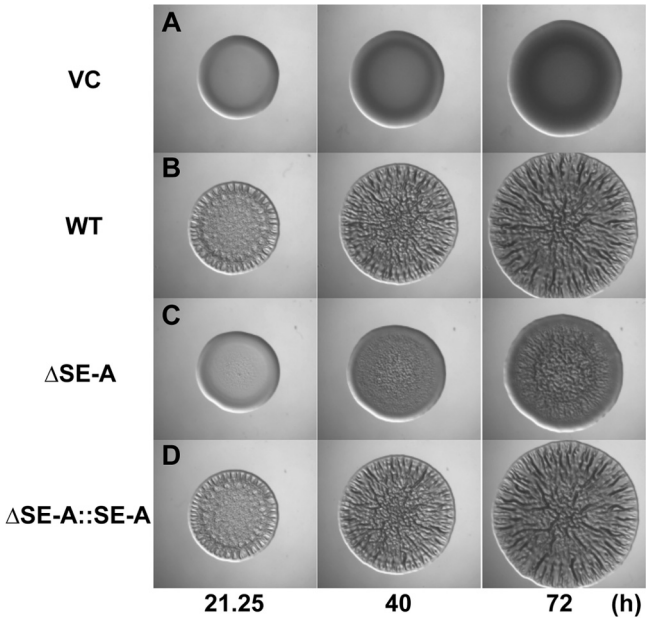


FIG 10 Wrinkled colony development by ΔSE-A and restored strains. The impact of deleting SE-A from its native position in the chromosome was assessed by monitoring wrinkled colony formation for 72 h postspotting for the following strains: (A) Wild type (WT; ES114) carrying pKV69 (vector control); (B) WT (ES114) carrying the *rscS* overexpression vector pKG11; (C) ΔSE-A (KV4970) carrying pKG11; (D) ΔSE-A::SE-A (KV6716) carrying pKG11. Data are representative of at least three independent experiments.

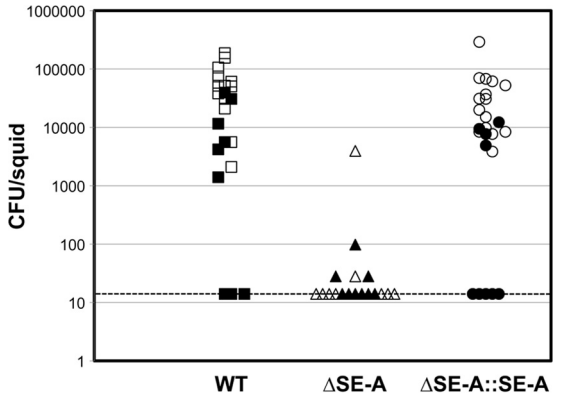


FIG 11 Colonization by wild-type (WT), ΔSE-A, and restored strains. The role of SE-A in colonization was assessed by inoculating newly hatched juvenile *E. scolopes* squid with WT (ES114 [squares]), the ΔSE-A mutant (KV4970 [triangles]), or the SE-A-restored strain (ΔSE-A::SE-A; KV6716 [circles]) for 3 h, followed by transfer to fresh artificial seawater and incubation for another 13 to 16 h. The number of CFU per squid was then assessed. Each marker represents an individual squid. The dashed line represents the limit of detection (14 CFU). The data shown are combined from two individual experiments (black and white symbols).

strain but demonstrated significantly reduced enrichment on the plasmid containing the *sypA* promoter lacking SE-A (Fig. 9C). These data suggest that SypG does indeed recognize and bind to the SE sequence, at least at the *sypA* promoter, and corroborate the findings we obtained from our biofilm assays.

The *sypA* enhancer is important for biofilm formation and colonization. To further investigate the importance of the SE-A sequence, we deleted it (SE-A plus 5 bp upstream) (see Fig. S6 in the supplemental material) from its native location in the chromosome and evaluated wrinkled colony formation. Because this strain expresses the SypE biofilm inhibitor, we induced *syp* by using the sensor kinase RscS, as this regulator functions both to activate *syp* transcription via SypG (18) and to inactivate SypE (Fig. 1A) (20). Whereas the wild-type strain (SE-A$^+$) formed wrinkled colonies by about 21 h, biofilm formation by the ΔSE-A strain was both delayed and diminished (Fig. 10A to C). To verify that the defect of the ΔSE-A strain was due to this deletion, we restored the SE-A sequence to its native location in the chromosome (ΔSE-A::SE-A) (see Fig. S6). This strain exhibited wrinkled colony formation similar to that of the positive control (Fig. 10, compare panels B and D). Thus, these data indicate that loss of SE-A from its native location in the chromosome indeed disrupts biofilm formation.

We next asked whether SE-A is similarly important for host colonization by *V. fischeri*. Previous studies had revealed a clear correlation between the ability to form a biofilm and colonization competence: strains unable to promote biofilm formation *in vitro* are also unable to proficiently colonize the host *in vivo* (8, 9, 11, 19). Furthermore, these experiments do not depend on overexpression of any regulatory protein, such as RscS or SypG. We thus inoculated juvenile *E. scolopes* squid with wild-type *V. fischeri*, the ΔSE-A mutant, or the ΔSE-A::SE-A strain and evaluated colonization at an early time point (see Materials and Methods). At this time point, wild-type-inoculated animals contained on average 4×10^4 bacteria, while squid inoculated with the ΔSE-A strain were uncolonized or contained, on average, fewer than 150 bacteria (Fig. 11). Finally, squid inoculated with the SE-A-restored strain (ΔSE-A::SE-A) contained wild-type levels of bacteria (Fig. 11). These data indicate that the SE-A sequence is indeed necessary for efficient initiation of host colonization, and they underscore the importance of this enhancer sequence to the lifestyle of *V. fischeri*.

DISCUSSION

In this study, we sought to determine whether a conserved sequence within the *syp* locus, the *syp* enhancer sequence, was necessary for the σ^{54}-dependent transcriptional activator SypG to be able to promote *syp* transcription, biofilm formation, and host colonization by *V. fischeri*. We have (i) demonstrated that the SE sequence is necessary for SypG-mediated *syp* transcription from the four *syp* promoters (*sypA*, *sypI*, *sypM*, and *sypP*), (ii) identified critical nucleotides within the 5′ and 3′ half-sites of the *sypA* enhancer (SE-A) necessary for SypG-mediated activation of biofilm formation, (iii) demonstrated that SypG is able to recognize the SE sequences associated with the *syp* locus, as well as those associated with genes outside the *syp* locus, (iv) determined that SypG binds to the *syp* promoter regions, and (v) found that loss of SE-A leads to defective biofilm formation and host colonization. Overall, this work filled in a critical gap in the Syp regulatory circuit by demonstrating that SypG binds to *syp* promoter regions in a manner that depends on SE sequences and allowed identification of three new potential SypG targets.

Through truncation analyses and the use of *lacZ*-reporter fusions, we demonstrated that *syp* reporter fusions that contain the SE sequence are activated in a SypG-dependent manner, while reporter fusions that lack the SE sequence are not (Fig. 2). The only exception was the *sypI* promoter, which contains two tandem

SE sequences (Fig. 1B). We found that truncation of the upstream SE sequence results in a loss of *syp* transcription (Fig. 2B), suggesting that either the upstream sequence or the presence of both sequences is necessary for *syp* transcription. Surprisingly, in our biofilm-based assay, we found that only the downstream SE-I sequence (SE-I-down) was sufficient to promote biofilm formation when used to control *sypA* expression (Fig. 7C and D). Although the IR of SE-I-up deviates from the consensus, the individual changes are permissive with respect to SE-A. Perhaps some combination of nucleotide substitutions prevents recognition by SypG in the context of the biofilm assay. In two other vibrios that contain a *syp*-like locus (i.e., *V. parahaemolyticus* and *V. vulnificus*) (10, 41, 42), the SE sequence is present and well conserved, but the promoter regions upstream of the *sypI* equivalent only contain one SE sequence (reference 42 and unpublished data). Thus, the presence of two SE sequences at the *sypI* promoter of *V. fischeri* merits further investigation.

In this study, we performed an in-depth analysis of the specific nucleotides necessary for SypG to recognize the SE sequence. This work was facilitated by the development of a novel biofilm-based assay that depends on the ability of SypG to activate transcription of *sypA*; increased SypA levels can overcome SypE-mediated inhibition and thus promote biofilm formation (see Fig. S1 in the supplemental material) (19). We categorized the phenotypes of our different strains into three simple categories: similar to the positive control, similar to the negative control, or an intermediate phenotype. Although there was a range of intermediate phenotypes, the assay was not robust enough to permit categorization of the intermediate phenotypes into subcategories. Nevertheless, we submit that the strains with diminished or disrupted biofilm formation represent those in which SypG-mediated transcription of *sypA* is diminished or defective. This conclusion is supported by results from both β-galactosidase and ChIP experiments.

In this assay, we found that loss of the first four bases of the conserved 22-bp sequence exerted no impact on biofilm formation, indicating that these bases are not required for SypG binding and activation (Fig. 4D). However, at least one and potentially up to three bases immediately upstream of the IR sequence contribute to SypG-mediated activation, as a truncation derivative lacking these bases exhibited a somewhat-reduced ability to promote biofilm formation (Fig. 4E). Perhaps not surprisingly, the IR sequences themselves were the critical determinants in this assay. Our ability to replace the SE-A sequence with specific point mutation derivatives permitted us to understand which positions are necessary for SypG recognition and which sequences are "flexible." Indeed, some flexibility was already suggested by the fact that the SE-A sequence differs from the other SE sequences in that it contains an imperfect IR, whereas SE-I-down, SE-M, and SE-P each contain a perfect IR (Fig. 1). We determined that mutations at position 1 or 3 of the IR did not cause a defect or delay in biofilm formation (Fig. 5). This result explains why SypG recognizes the SE sequence associated with *sypA*, as SE-A has an A at position 3, while the other *syp* SE sequences have a G (Fig. 1B; see also Fig. S2A in the supplemental material). These data also explain why SypG is able to recognize SE sequences associated with genes outside of the *syp* locus (i.e., SE-1019 and SE-0550) that have changes at position 1 and/or 3 (Fig. 8A; see also Fig. S2B).

Most of the other bases were required for SypG-mediated activation, as changes at nucleotide positions other than 1 and 3 resulted in reduced or no biofilm formation (Fig. 5). However, some specific changes in a required nucleotide were permissive. For example, at positions 2 and 4, an A-to-T change and a C-to-A change, respectively, were permissive, but other changes were not (Fig. 6B and C). Consistent with the A-to-T change at position 2 being permissive, this sequence is naturally present in SE-0550, which also promotes biofilm formation. As for the C-to-A change at position 4, this specific change brings SE-A closer to the perfect IR consensus present in the other SE sequences.

Combining our knowledge of the critical SE sequences necessary for SypG-mediated activation with bioinformatic analyses, we have identified a number of genes that have a putative SE sequence, some of which also have a putative σ^{54} recognition sequence associated with them (see Table S3 in the supplemental material). Of the non-*syp* SE sequences, we chose to characterize three, SE-1019, SE-0120, and SE-A0550, since these sequences also appeared to be associated with a putative downstream σ^{54} binding site. Of note, the three genes (*VF_A1019*, *VF_A0120*, and *VF_A0550*) encode proteins that are similar to each other (~40% identical and ~50% similar [43, 44]). We predict that, since SypG recognizes the SE sequences associated with them, these genes constitute part of the SypG regulon and might also be involved in biofilm formation. We are currently exploring the roles of these genes in biofilm formation and colonization.

The *V. fischeri* SE sequence is necessary not only for *syp* transcription and biofilm formation but also for host colonization. It is striking that deletion of only 27 bp from the *V. fischeri* genome (ΔSE-A) resulted in the inability to efficiently colonize the squid host (Fig. 11). This effect can be attributed to the severe decrease in transcription of the *sypA* operon (Fig. 2A), as *sypA* is known to be a critical colonization determinant (9), while the divergently transcribed *VF_A1019* is not required for colonization (10). Furthermore, reintroduction of SE-A into its native location on the chromosome restored colonization competence (Fig. 11), suggesting that the colonization defect was indeed due to the loss of this small sequence. Because the colonization experiments did not rely on the overexpression of regulatory proteins, this work implicates the SE-A sequence (and, likely, SE-I, SE-M, and SE-P) as a critical regulatory sequence in *V. fischeri* and gives us a better understanding of the biology of this organism.

Importantly, the work performed here to understand the role of a conserved sequence in biofilm formation and colonization has the potential for impacting more than our understanding of the marine bacterium *V. fischeri*. Numerous *Vibrio* species contain a similar locus, including the human pathogens *V. vulnificus* and *V. parahaemolyticus* (10, 41). The SE sequence we identified in *V. fischeri* (10) is nearly identical to a conserved sequence found in *V. vulnificus* within the *rbd* locus, a *syp*-like locus similarly involved in biofilm formation (42). Similar to SypG, the SypG-like regulator RbdG is capable of inducing expression of the *rbd* genes, although the role of the conserved IR sequences in transcriptional control has yet to be investigated (42). Of note, the SE-like sequences associated with *rbdA*, *rbdM*, and *rbdP* contain perfect IR sequences, while the SE-like sequence associated with *rbdI* is not a perfect IR. Furthermore, in contrast to *sypI*, *rbdI* appears to have only a single SE-like sequence. The similarities between these two studied loci are striking, but the differences will be informative. Undoubtedly, work on one locus will provide insights into the other as well as into numerous other *Vibrio* species that contain this locus.

ACKNOWLEDGMENTS

We thank Shannon Newman, Krystal Thomas-White, and Jakob Ondrey for their contributions to this project. Additionally, we are grateful to Alan Wolfe for helpful discussions and to members of the Visick lab for discussions and reading the manuscript.

This work was funded by NIH grant GM59690, awarded to K.L.V.

REFERENCES

1. **Stoodley P, Sauer K, Davies DG, Costerton JW.** 2002. Biofilms as complex differentiated communities. Annu. Rev. Microbiol. **56:**187–209.
2. **Branda SS, Vik S, Friedman L, Kolter R.** 2005. Biofilms: the matrix revisited. Trends Microbiol. **13:**20–26.
3. **Flemming HC, Wingender J.** 2010. The biofilm matrix. Nat. Rev. Microbiol. **8:**623–633.
4. **Ruby EG.** 1996. Lessons from a cooperative, bacterial-animal association: the *Vibrio fischeri-Euprymna scolopes* light organ symbiosis. Annu. Rev. Microbiol. **50:**591–624.
5. **Nyholm SV, McFall-Ngai MJ.** 2004. The winnowing: establishing the squid-*Vibrio* symbiosis. Nat. Rev. Microbiol. **2:**632–642.
6. **Stabb E, Visick K.** 2013. *Vibrio fischeri*: a bioluminescent light organ symbiont of the bobtail squid *Euprymna scolopes*, p 497–532. *In* Rosenberg E (ed), The prokaryotes. Springer-Verlag, Berlin, Germany.
7. **Nyholm SV, Stabb EV, Ruby EG, McFall-Ngai MJ.** 2000. Establishment of an animal-bacterial association: recruiting symbiotic vibrios from the environment. Proc. Natl. Acad. Sci. U. S. A. **97:**10231–10235.
8. **Yip ES, Geszvain K, DeLoney-Marino CR, Visick KL.** 2006. The symbiosis regulator *rscS* controls the *syp* gene locus, biofilm formation and symbiotic aggregation by *Vibrio fischeri*. Mol. Microbiol. **62:**1586–1600.
9. **Morris AR, Visick KL.** 2013. The response regulator SypE controls biofilm formation and colonization through phosphorylation of the *syp*-encoded regulator SypA in *Vibrio fischeri*. Mol. Microbiol. **87:**509–525.
10. **Yip ES, Grublesky BT, Hussa EA, Visick KL.** 2005. A novel, conserved cluster of genes promotes symbiotic colonization and σ^{54}-dependent biofilm formation by *Vibrio fischeri*. Mol. Microbiol. **57:**1485–1498.
11. **Shibata S, Yip ES, Quirke KP, Ondrey JM, Visick KL.** 2012. Roles of the structural symbiosis polysaccharide (*syp*) genes in host colonization, biofilm formation, and polysaccharide biosynthesis in *Vibrio fischeri*. J. Bacteriol. **194:**6736–6747.
12. **Ray VA, Visick KL.** 2012. Two-component regulators in the *Vibrio fischeri-Euprymna scolopes* symbiosis, p 371–404. *In* Gross R, Beier D (ed), Two-component systems in bacteria. Caister Academic Press, Norfolk, United Kingdom.
13. **Bourret RB, Hess JF, Simon MI.** 1990. Conserved aspartate residues and phosphorylation in signal transduction by the chemotaxis protein CheY. Proc. Natl. Acad. Sci. U. S. A. **87:**41–45.
14. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annu. Rev. Biochem. **69:**183–215.
15. **West AH, Stock AM.** 2001. Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem. Sci. **26:**369–376.
16. **Hussa EA, O'Shea TM, Darnell CL, Ruby EG, Visick KL.** 2007. Two-component response regulators of *Vibrio fischeri*: identification, mutagenesis, and characterization. J. Bacteriol. **189:**5825–5838.
17. **Geszvain K, Visick KL.** 2008. The hybrid sensor kinase RscS integrates positive and negative signals to modulate biofilm formation in *Vibrio fischeri*. J. Bacteriol. **190:**4437–4446.
18. **Hussa EA, Darnell CL, Visick KL.** 2008. RscS functions upstream of SypG to control the *syp* locus and biofilm formation in *Vibrio fischeri*. J. Bacteriol. **190:**4576–4583.
19. **Morris AR, Visick KL.** 2013. Inhibition of SypG-induced biofilms and host colonization by the negative regulator SypE in *Vibrio fischeri*. PLoS One **8**(3):e60076. doi:10.1371/journal.pone.0060076.
20. **Morris AR, Darnell CL, Visick KL.** 2011. Inactivation of a novel response regulator is necessary for biofilm formation and host colonization by *Vibrio fischeri*. Mol. Microbiol. **82:**114–130.
21. **Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD.** 2000. The bacterial enhancer-dependent σ^{54} (σ^{N}) transcription factor. J. Bacteriol. **182:**4129–4136.
22. **Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, Buck M.** 2008. *Modus operandi* of the bacterial RNA polymerase containing the σ54 promoter-specificity factor. Mol. Microbiol. **68:**538–546.
23. **Bush M, Dixon R.** 2012. The role of bacterial enhancer binding proteins as specialized activators of σ^{54}-dependent transcription. Microbiol. Mol. Biol. Rev. **76:**497–529.
24. **Boettcher KJ, Ruby EG.** 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. J. Bacteriol. **172:**3701–3706.
25. **DeLoney CR, Bartley TM, Visick KL.** 2002. Role for phosphoglucomutase in *Vibrio fischeri-Euprymna scolopes* symbiosis. J. Bacteriol. **184:**5121–5129.
26. **Herrero M, de Lorenzo V, Timmis KN.** 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J. Bacteriol. **172:**6557–6567.
27. **Le Roux F, Binesse J, Saulnier D, Mazel D.** 2007. Construction of a *Vibrio splendidus* mutant lacking the metalloprotease gene *vsm* by use of a novel counterselectable suicide vector. Appl. Environ. Microbiol. **73:**777–784.
28. **Graf J, Dunlap PV, Ruby EG.** 1994. Effect of transposon-induced motility mutations on colonization of the host light organ by *Vibrio fischeri*. J. Bacteriol. **176:**6986–6991.
29. **Ruby EG, Nealson KH.** 1977. A luminous bacterium that emits yellow light. Science **196:**432–434.
30. **Davis RW, Botstein D, Roth JR.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
31. **McCann J, Stabb EV, Millikan DS, Ruby EG.** 2003. Population dynamics of *Vibrio fischeri* during infection of *Euprymna scolopes*. Appl. Environ. Microbiol. **69:**5928–5934.
32. **Stabb EV, Ruby EG.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods Enzymol. **358:**413–426.
33. **Bao Y, Lies DP, Fu H, Roberts GP.** 1991. An improved Tn*7*-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. Gene **109:**167–168.
34. **Miller JH.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
35. **Visick KL, Skoufos LM.** 2001. Two-component sensor required for normal symbiotic colonization of *Euprymna scolopes* by *Vibrio fischeri*. J. Bacteriol. **183:**835–842.
36. **Ray VA, Visick KL.** 2012. LuxU connects quorum sensing to biofilm formation in *Vibrio fischeri*. Mol. Microbiol. **86:**954–970.
37. **van Helden J.** 2003. Regulatory sequence analysis tools. Nucleic Acids Res. **31:**3593–3596.
38. **Thomas-Chollier M, Sand O, Turatsinze JV, Janky R, Defrance M, Vervisch E, Brohee S, van Helden J.** 2008. RSAT: regulatory sequence analysis tools. Nucleic Acids Res. **36:**W119–W127.
39. **Thomas-Chollier M, Defrance M, Medina-Rivera A, Sand O, Herrmann C, Thieffry D, van Helden J.** 2011. RSAT 2011: regulatory sequence analysis tools. Nucleic Acids Res. **39:**W86–W91.
40. **Darnell CL, Hussa EA, Visick KL.** 2008. The putative hybrid sensor kinase SypF coordinates biofilm formation in *Vibrio fischeri* by acting upstream of two response regulators, SypG and VpsR. J. Bacteriol. **190:**4941–4950.
41. **Yildiz FH, Visick KL.** 2009. *Vibrio* biofilms: so much the same yet so different. Trends Microbiol. **17:**109–118.
42. **Guo Y, Rowe-Magnus DA.** 2011. Overlapping and unique contributions of two conserved polysaccharide loci in governing distinct survival phenotypes in *Vibrio vulnificus*. Environ. Microbiol. **13:**2888–2990.
43. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
44. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.

Supplementary Material

The *syp* enhancer sequence plays a key role in transcriptional activation by the σ^{54}-dependent response regulator SypG and in biofilm formation and host colonization by *Vibrio fischeri*

Valerie A. Ray, Justin L. Eddy, Elizabeth A. Hussa*, Michael Misale, and Karen L. Visick#
Department of Microbiology and Immunology, Loyola University Chicago

#Corresponding author
Department of Microbiology and Immunology
Health Sciences Division
Loyola University Chicago
2160 S. 1st Ave., Bldg. 105 Rm. 3933
Maywood, IL 60153
kvisick@lumc.edu
(708) 216-0869
(708) 216-9574 fax
kvisick@lumc.edu

***Current address**
Department of Bacteriology
University of Wisconsin-Madison
Microbial Science Building, Rm. 4545
1550 Linden Dr.
Madison, WI 53706
(608) 265-4307

Supplemental Table S1. Plasmids used in this study

Plasmid	Description	Relevant Primers[1]	Reference
pARM9	pVSV105 + *sypG*	N/A	(1)
pEAH90	pEVS107 + P_{sypA} - *lacZ* Em^R (FL + SE-A)	N/A	(2)
pEAH120	pEVS107 + P_{sypI} - *lacZ* Em^R (No SE-I)	783 and 786	This study
pEAH121	pEVS107 + P_{sypI} - *lacZ* Em^R (FL + SE-I)	784 and 786	This study
pEAH123	pEVS107 + P_{sypM} - *lacZ* Em^R (No SE-M)	787 and 718	This study
pEAH124	pEVS107 + P_{sypM} - *lacZ* Em^R (FL + SE-M)	788 and 718	This study
pEAH128	pEVS107 + P_{sypA} - *lacZ* Em^R (No SE-A)	675 and 713	This study
pEVS104	Conjugal helper plasmid (*tra trb*); Kn^R	N/A	(3)
pEVS107	Mini-Tn7 delivery plasmid; *oriR6K*, *mob*; Kn^R, Em^R	N/A	(4)
pKJW4	pKV300 ΔCm^R + SE-A - 14TG	1547, 1548	This study
pKJW5	pKV300 ΔCm^R + SE-A - 14TA	1549, 1550	This study
pKJW8	pKV300 ΔCm^R + SE-A - 5GT	1555, 1556	This study
pKJW9	pKV300 ΔCm^R + SE-A - 5GC	1557, 1558	This study
pKJW10	pKV300 ΔCm^R + SE-A - 4CG	1559, 1560	This study
pKJW11	pKV300 ΔCm^R + SE-A - 4CA	1561, 1562	This study
pKV69	Cm^R, Tc^R, *mob*, *oriT*	N/A	(5)
pKV300	pKV69 + P_{sypA} *sypA* (No SE-A); Cm^R, Tc^R	675, 807	This study
pKV301	pKV69 + P_{sypA} *sypA* (+ SE-A); Cm^R, Tc^R	673, 807	This study
pKV311	pEVS107+ P_{sypP} - *lacZ* Em^R (IG + SE-P)	719, 720	This study
pKV312	pEVS107+ P_{sypI} - *lacZ* Em^R (IG + SE-I)	937, 786	This study
pKV313	pEVS107 + P_{sypP} - lacZ Em^R (FL+ SE-P)	791, 720	This study
pKV314	pEVS107 + P_{sypP} - lacZ Em^R(No SE-P)	790, 720	This study
pKV316	pEVS107 + P_{sypM} - *lacZ* Em^R (IG + SE-M)	938, 718	This study
pKV318	pEVS107 + P_{sypA} - *lacZ* Em^R (IG + SE-A)	931, 714	This study
pKV363	Cm^R, *oriT*, *oriR6K*, *ccdB*	N/A	(6)
pKV437	pKV300 ΔCm^R + SE-I-down	1384, 1385	This study
pKV438	pKV300 ΔCm^R + SE-I -up	1386, 1387	This study
pKV439	pKV300 ΔCm^R + SE-M	1388, 1389	This study
pKV440	pKV300 ΔCm^R + SE-P	1390, 1391	This study
pKV444	pKV300 ΔCm^R + SE-A -15TC	1398, 1399	This study
pKV445	pKV300 ΔCm^R + SE-A - 14TC	1400, 1401	This study
pKV446	pKV300 ΔCm^R + SE-A - 13CT	1402, 1403	This study
pKV447	pKV300 ΔCm^R + SE-A - 12TC	1404, 1405	This study
pKV448	pKV300 ΔCm^R + SE-A - 11CT	1406, 1407	This study
pKV449	pKV300 ΔCm^R + SE-A - 10AG	1408, 1409	This study
pKV450	pKV300 ΔCm^R + SE-A - 6TC	1410, 1411	This study
pKV451	pKV300 ΔCm^R + SE-A - 5AG	1412, 1413	This study
pKV452	pKV300 ΔCm^R + SE-A - 4CT	1414, 1415	This study
pKV453	pKV300 ΔCm^R + SE-A - 3AG	1416, 1417	This study
pKV454	pKV300 ΔCm^R + SE-A - 2AG	1418, 1419	This study
pKV455	pKV300 ΔCm^R + SE-A - 1AG	1420, 1421	This study

pKV461	pMSM25 with SE-A inserted at non-native ApaI site; restores SE-A	1001, 1002	This study
pKV465	pKV300 ΔCm^R (No SE-A)	N/A	This study
pMSM25	pKV363 containing ~500 bp flanking SE-A on either site (deletes SE-A)	935, 936, 675, 807	This study
pSLN4	pKV300 ΔCm^R + SE-A Δ4 bp	1028, 1029	This study
pSLN5	pKV300 ΔCm^R + SE-A Δ7 bp	1030, 1031	This study
pSLN7	pKV300 ΔCm^R + SE-A Δ10 bp	1032, 1033	This study
pVAR45	pVSV105 + *sypG*-FLAG	1249, 1438	This study
pVAR57[2]	pKV301 ΔCm^R (+ SE-A)	N/A	This study
pVAR58[2]	pKV300 ΔCm^R + SE-A	1036, 1037	This study
pVAR83	pKV300 ΔCm^R + SE-A – 2AC	1628, 1629	This study
pVAR84	pKV300 ΔCm^R + SE-A – 2AT	1630, 1631	This study
pVAR91	pKV300 ΔCm^R + *sypA*-HA (+SE-A)	1798, 1800	This study
pVAR92	pKV300 ΔCm^R + *sypA*-HA (ΔSE-A)	1799, 1800	This study
pVSV105	Mobilizable vector, Cm^R	N/A	(7)

[1]Relevant primers for plasmids generated in this study; N/A, not applicable.
[2]pVAR57 and pVAR58 differ in the presence of 5 bp non-native ("scar") sequence on the 3' side of SE-A present in pVAR58 but not pVAR57.

Supplemental Table S2. Primers used in this study

Primer	Sequence[1]
393	GCTACACTTTCACTAGACGC
559	ggtaccGGTACCTCATTCCGATTCTTCATAG
633	TTTGCAGAATGAGAATCGACTT
673	aagggcccCTCTTAAGTCGATTCTCATTCTGCAAACTGCA
675	aagggcccCTGCATTGCAAATTGAGAATATATC
684	ACTCGCATCGGTGTCAGC
713	aaactagtcattAACCGATGGCGTCCATATCAC
714	aaactagtcattAACCGATGGCGTCCATATCACCTTGAA
718	aaactagtcattaaAGGCGAGGAGTTGGCAACTC
719	aaagggcccGCACATCTCCTGTATTAAGCCG
720	aaactagtcattACCCGCCAGGTGAAAGGTG
783	aagggcccAAATAAAACAGCATATAAATTAATCATTAAAA
784	aagggcccGATAAAAATCTATTCTCAATTTGAGAAAAATA
786	aaactagtcattaaCGGTATTTCTGATGGTTGATGCTTTTA
787	aagggcccATTTAGAATAACCCTAATCAAAATAACATATC
788	aagggcccCTTTCAAATCGATTCTCAAAATGAGAAATTTA
790	aagggcccAACAAAAACAACGATAACTAACACATTG
791	aagggcccTTACTAAATTCATTCTCAATTTGAGAAAACAA
807	ATGTGTCATACAGTTAAAATGGTG
931	aaagggcccAGCTTCTTCCTTATAGTTATGATG
935	ctgcagTGTTTTATCCGAAGGTAACCC
936	gggcccTTGATAGGTAAATAGATTAAAAATG
937	aagggcccGTCTGCTTATCAAATTTAATTTGG
938	aagggcccAAAGACCGAATTGATCCCGC
1001	AAGTCGATTCTCATTCTGCAAAggcc
1002	TTTGCAGAATGAGAATCGACTTggcc
1028	catgCGATTCTCATTCTGCAAAggcc
1029	TTTGCAGAATGAGAATCG
1030	catgTTCTCATTCTGCAAAggcc
1031	TTTGCAGAATGAGAA
1032	catgTCATTCTGCAAAggcc
1033	TTTGCAGAATGA
1036	catgAAGTCGATTCTCATTCTGCAAAggcc
1037	TTTGCAGAATGAGAATCGACTT
1223	GAATGTCTTGCTAAGTACCTG
1249	catactaagtgcggccgcctaAAACAAGGTTTCTCAAAATAAAAG
1384	catgAAATCTATTCTCAATTTGAGAAggcc
1385	TTCTCAAATTGAGAATAGATTT
1386	catgATTTGGTTTCTCATATTGATAAggcc

1387	TTATCAATATGAGAAACCAAAT
1388	catgAAATCGATTCTCAAAATGAGAAggcc
1389	TTCTCATTTTGAGAATCGATTT
1390	catgAAATTCATTCTCAATTTGAGAAggcc
1391	TTCTCAAATTGAGAATGAATTT
1398	catgAAGTCGA**C**TCTCATTCTGCAAAggcc
1399	TTTGCAGAATGAGA**G**TCGACTT
1400	catgAAGTCGAT**C**CTCATTCTGCAAAggcc
1401	TTTGCAGAATGAG**G**ATCGACTT
1402	catgAAGTCGATT**T**TCATTCTGCAAAggcc
1403	TTTGCAGAATGAA**A**AATCGACTT
1404	catgAAGTCGATTC**C**CATTCTGCAAAggcc
1405	TTTGCAGAATG**G**GAATCGACTT
1406	catgAAGTCGATTCT**T**ATTCTGCAAAggcc
1407	TTTGCAGAAT**A**AGAATCGACTT
1408	catgAAGTCGATTCTC**G**TTCTGCAAAggcc
1409	TTTGCAGAA**C**GAGAATCGACTT
1410	catgAAGTCGATTCTCATTC**C**GCAAAggcc
1411	TTTGC**G**GAATGAGAATCGACTT
1412	catgAAGTCGATTCTCATTCT**A**CAAAggcc
1413	TTTG**T**AGAATGAGAATCGACTT
1414	catgAAGTCGATTCTCATTCTG**T**AAAggcc
1415	TTT**A**CAGAATGAGAATCGACTT
1416	catgAAGTCGATTCTCATTCTGC**G**AAggcc
1417	TT**C**GCAGAATGAGAATCGACTT
1418	catgAAGTCGATTCTCATTCTGCA**G**Aggcc
1419	T**C**TGCAGAATGAGAATCGACTT
1420	catgAAGTCGATTCTCATTCTGCAA**G**ggcc
1421	**C**TTGCAGAATGAGAATCGACTT
1438	aaaaaggtaccttatttatcatcatcatctttataatcTTCCGATTCTTCATAGGCTTCCCA
1453	CAGATTGCACGCAATTAATTTCAT
1455	TTGTCTGCTTATCAAATTTAATTT
1457	CAAACAGGAAAAACACTATGATGG
1458	TGCAAGAGCTGTAACCCATTGTCG
1459	AGTATGTCAAAAGCCAATTAAAAC
1464	GGCTCACCATCCTTGCAGTTTCTT
1468	TGAATTGCAGAGTTTGTGCCA
1471	GGAGCGGAAACGCTAAGTGGTTGA
1472	TTCCGCTCGACTTGCATGTGTTAG
1547	catgAAGTCGAT**G**CTCATTCTGCAAAggcc
1548	TTTGCAGAATGAG**C**ATCGACTT

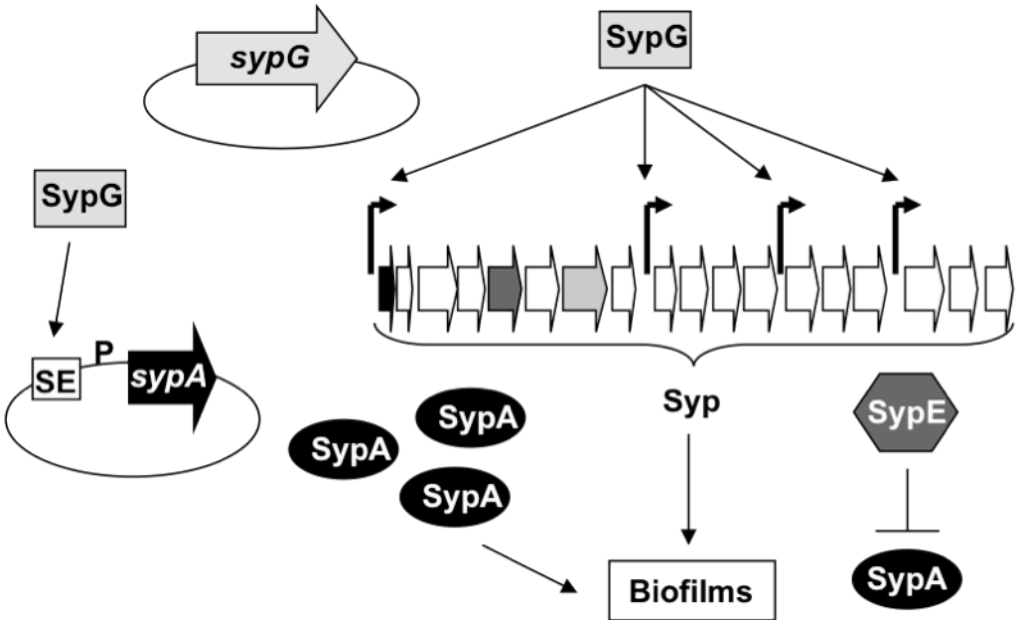
1549	catgAAGTCGAT**A**CTCATTCTGCAAAggcc
1550	TTTGCAGAATGAG**T**ATCGACTT
1555	catgAAGTCGATTCTCATTCT**T**CAAAggcc
1556	TTTG**A**AGAATGAGAATCGACTT
1557	catgAAGTCGATTCTCATTCT**C**CAAAggcc
1558	TTTG**G**AGAATGAGAATCGACTT
1559	catgAAGTCGATTCTCATTCTG**G**AAAggcc
1560	TTT**C**CAGAATGAGAATCGACTT
1561	catgAAGTCGATTCTCATTCTG**A**AAAggcc
1562	TTT**T**CAGAATGAGAATCGACTT
1628	catgAAGTCGATTCTCATTCTGCA**C**Aggcc
1629	T**G**TGCAGAATGAGAATCGACTT
1630	catgAAGTCGATTCTCATTCTGCA**T**Aggcc
1631	T**A**TGCAGAATGAGAATCGACTT
1752	CGCAGTACTGTTGTAATTCATTAAG
1753	GATATATTCTCAATTTGCAATGCAGG
1798	gccttgcgtataatatttgcccatggAAGTCGATTCTCATTCTGCAAA
1799	gccttgcgtataatatttgcccatggCTGCATTGCAAATTGAGAATATA
1800	gattacgccaagcttgcatgcctgcaggaattcgagctcggtacc

[1]Non-native sequences are in lower case letters; point mutations are shown in bold.

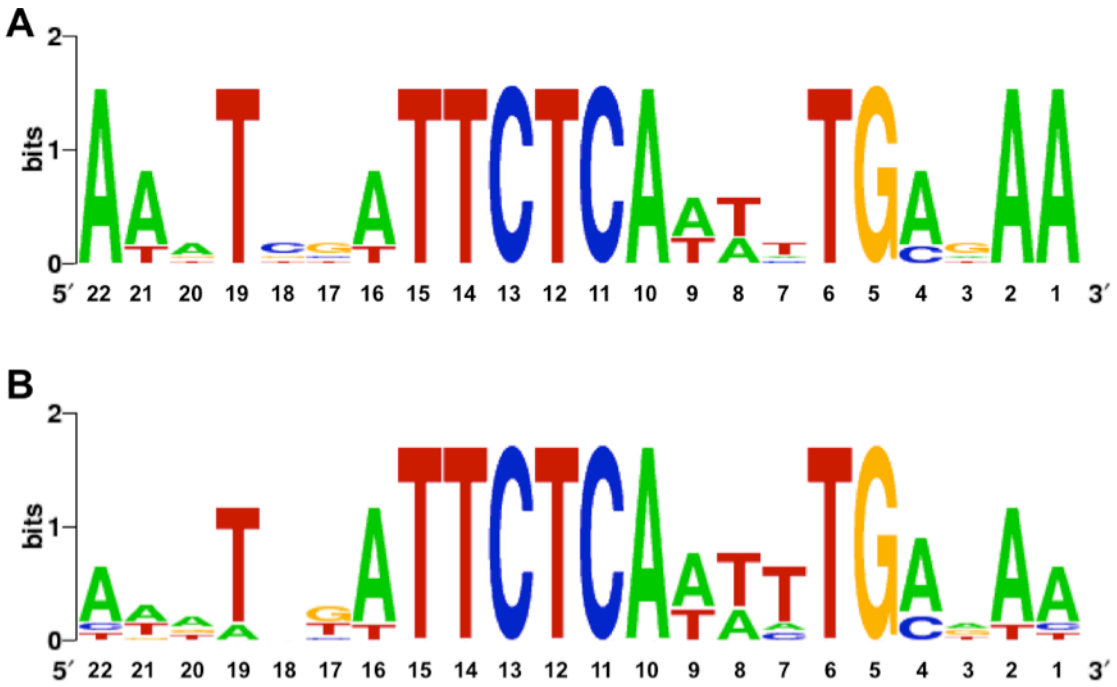
Supplementary Table S3. Bioinformatics search for the SE sequence in *V. fischeri*.

Gene	σ^{54}	Name	Predicted protein
VF_0270		*hepA*	ATP-dependent helicase
VF_0703		*ribH*	6,7-dimethyl-8-ribityllumazine synthase
VF_0885			ABC transporter permease
VF_0952		*ruvB*	Holliday junction DNA helicase
VF_1010			Hypothetical protein
VF_1062			Adenylosuccinate synthase
VF_1063			Transcriptional regulator
VF_1326			Acetyltransferase
VF_2641			Hypothetical protein
VF_1423			Hypothetical protein
VF_1537[1]			Hypothetical protein
VF_1538[1]			Hypothetical protein
VF_1849		*fliI*	Flagellum-specific ATP synthase
VF_2089		*lepA*	GTP binding protein
VF_A0119			Lactoylglutathione lyase
VF_A0120	YES		Conserved hypothetical protein
VF_A0159		*fuhD*	Iron-hydroxamate transporter subunit
VF_A0222			Chromosome partitioning ATPase
VF_A0253		*pepT*	Peptidase T
VF_A0550	YES		Conserved hypothetical protein
VF_A0559			Hypothetical protein
VF_A0640[1]			Glyoxylase
VF_A0642[1]			Hypothetical protein
VF_A0683			Hypothetical protein
VF_A0734			Hypothetical protein
VF_A1019	YES		Conserved hypothetical protein
VF_A1020	YES	*sypA*	Sulphate transporter, anti-sigma factor antagonist
VF_A1028	YES	*sypI*	Group 1 glycosyltransferase
VF_A1028	YES	*sypI*	Group 1 glycosyltransferase
VF_A1032	YES	*sypM*	Acetyltransferase
VF_A1035	YES	*sypP*	Glycosyltransferase
VF_A1070			Hypothetical protein
VF_A1083			Methyl-accepting chemotaxis protein

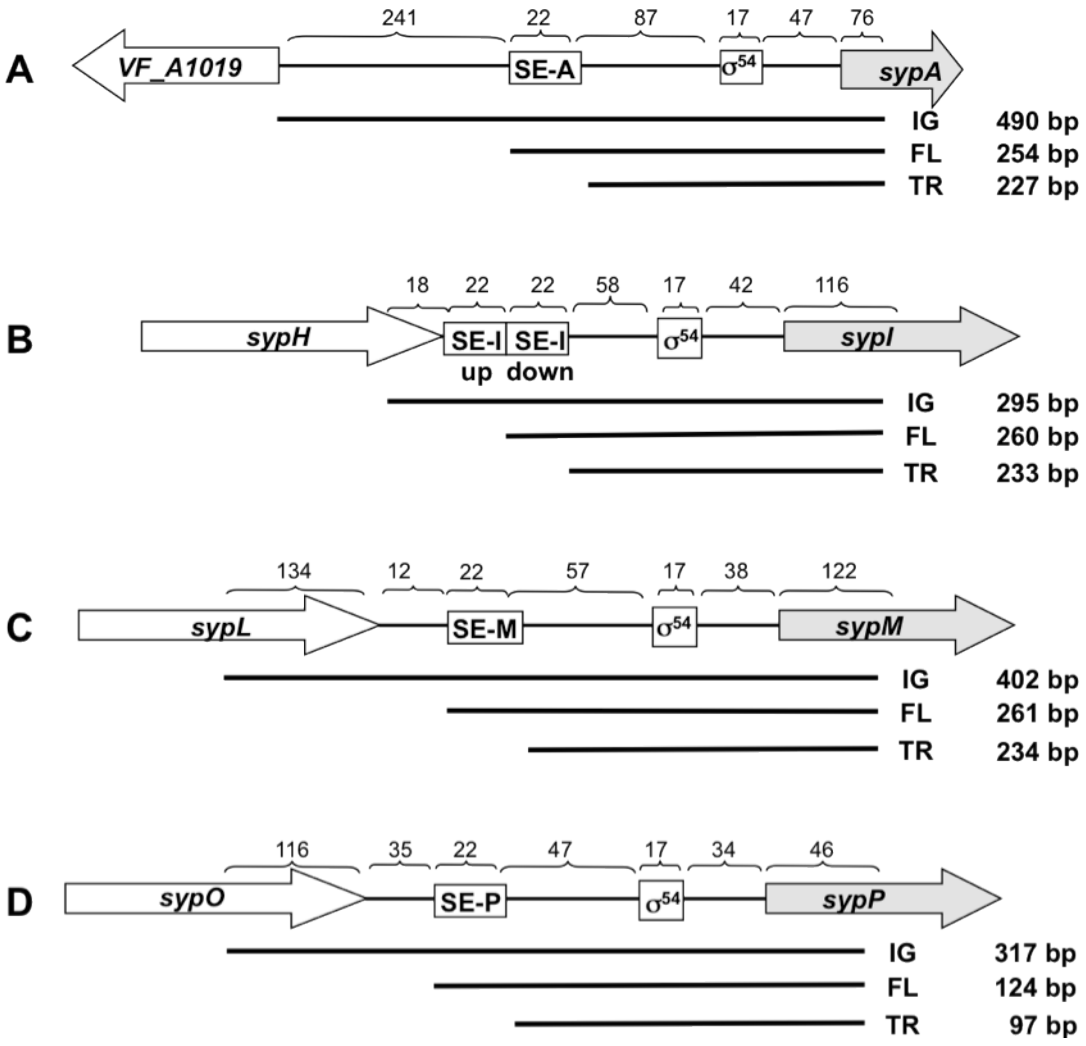
[1]It is unclear which gene is associated with the predicted SE sequence.



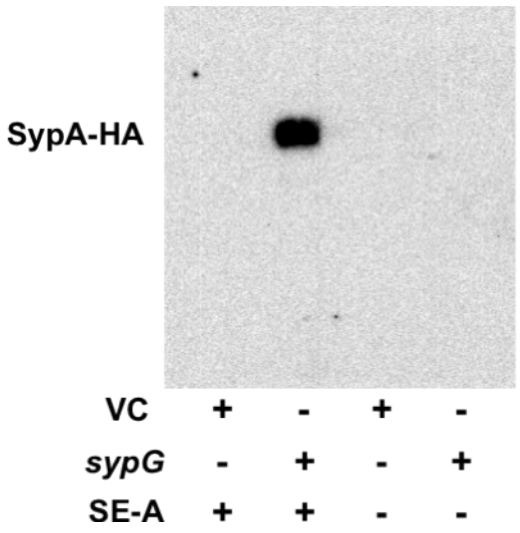
Supplementary Figure S1. Model for the *sypA*-based biofilm assay. Overexpression of *sypG* alone (from a plasmid, grey block arrow in open circle) activates transcription of the *syp* locus, but is unable to promote biofilm formation due to the inhibitory activity of SypE (grey hexagon) on SypA (black oval). However, co-overexpression of *sypG* and *sypA* [black block arrow in open circle with promoter (P) and *syp* enhancer (SE) sequence (white box)] promotes biofilm formation, likely by SypG-mediated *sypA* expression, leading to excess SypA, which can escape the inhibitory activity of SypE and promote biofilm formation.



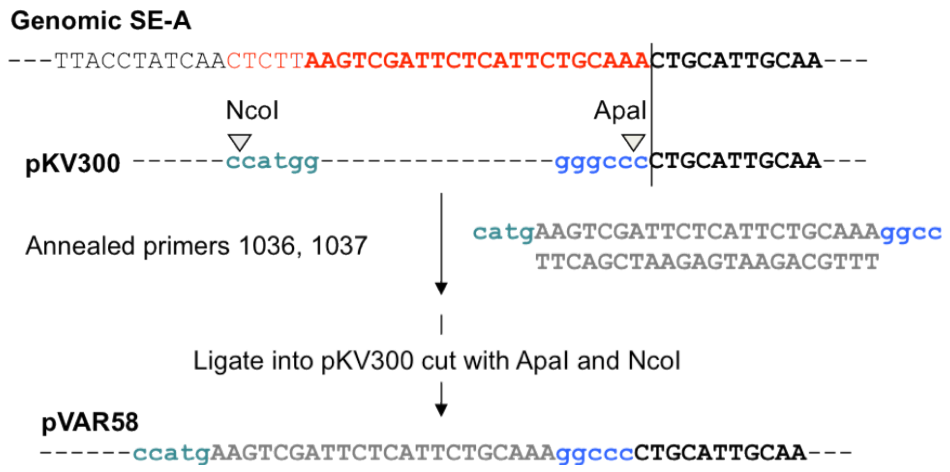
Supplementary Figure S2. Sequence analysis of SE sequences. Logo sequence analysis [weblogo.berkely.edu (8, 9)] comparing (A) SE-A, SE-I-up, SE-I-down, SE-M, and SE-P and (B) SE-A, SE-1019, SE-0120, and SE-0550.



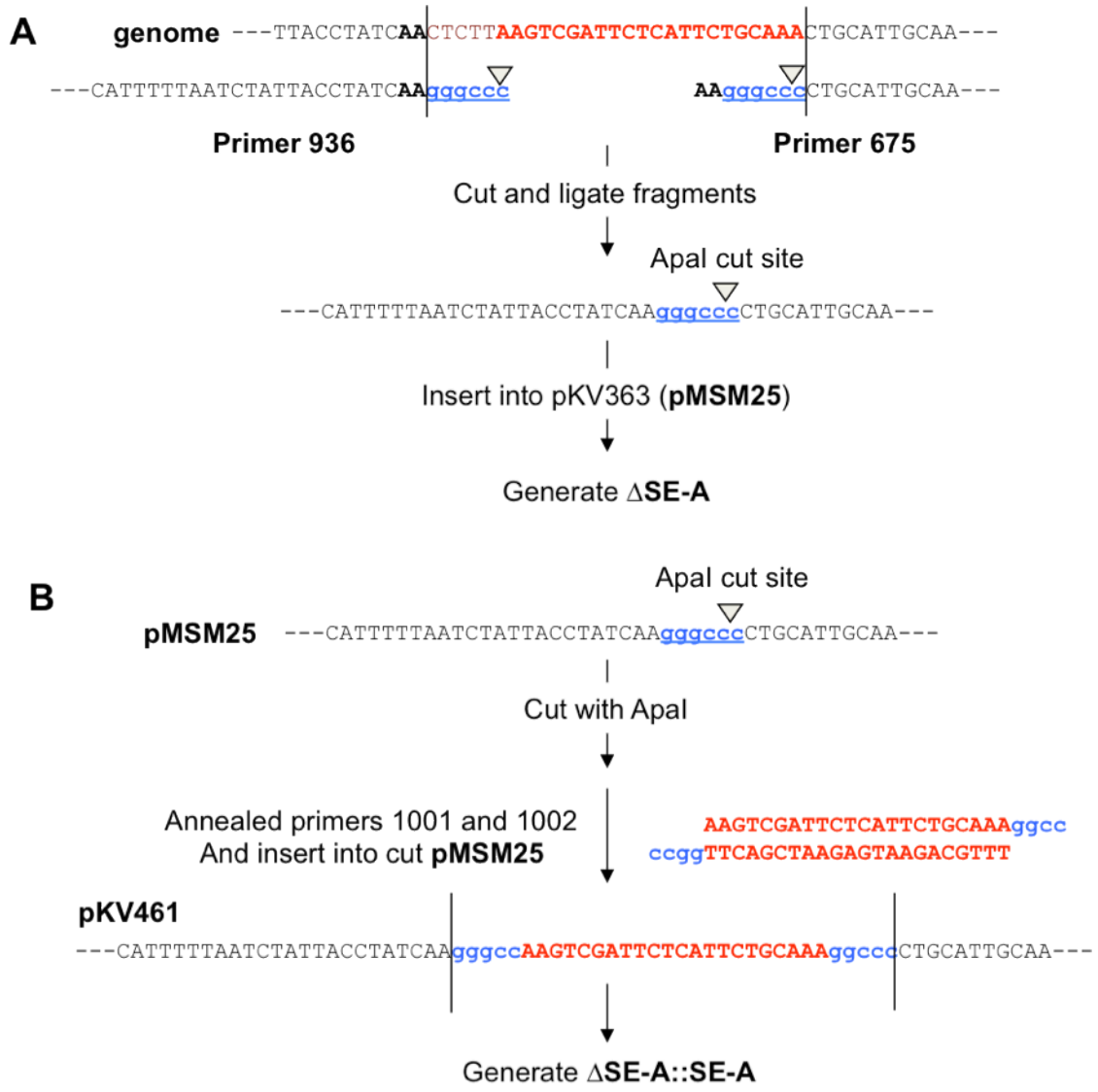
Supplementary Figure S3. Representation of the SE regions of the *syp* locus and regions used for *lacZ*-fusion derivatives. The promoter regions for (A) *sypA*, (B) *sypI*, (C) *sypM*, and (D) *sypP* are shown. The *syp* genes are indicated with block arrows, with the relevant *syp* gene shaded in gray. The predicted σ^{54} and SE sequences are indicated with boxes. Numbers and brackets represent the distance within the indicated region. The promoter derivative used in the *lacZ*-fusion construct is indicated beneath each promoter region. Intergenic (IG) and full-length (FL) constructs each contains the associated SE sequence, with the exception of FL *sypI*, which lacks SE-I-up. The truncation (TR) constructs lack SE sequences.



Supplementary Figure S4. Assessment of SypA protein levels in strains with *sypG* and *sypA* plasmids. Western blot analysis was performed on the following strains to determine the level of SypA in strains that contained vector control (VC – pVSV105) or the *sypG* plasmid (pARM9) and a plasmid containing an HA-tagged allele of *sypA* with or without the SE-A sequence [pVAR91 (+SE-A) and pVAR92 (-SE-A), respectively]: VC and pVAR91 (lane 1), pARM9 and pVAR91 (lane 2), VC and pVAR92 (lane 3), and pARM9 and pVAR92 (lane 4).



Supplementary Figure S5. Construction of pVAR58 and other derivatives of pKV300. The genomic context of SE-A (red bold letters) is shown and compared to pKV300, which contains only the sequence downstream of SE-A (bold black letters). pKV300 was cut with NcoI and ApaI as indicated and the annealed primer set 1036/1037 was ligated into the linearized vector. The resulting plasmid, pVAR58, contains an insertion of the SE-A sequence flanked on both sides by 5 non-native nucleotides. These additional nucleotides had no impact on the ability of SypG to recognize the SE-A sequence in our biofilm-based assay (see Figure 3). Other pKV300 derivatives were generated in a similar manner.



Supplementary Figure S6. Construction of SE-A deletion and restoration plasmids. (A) DNA fragments (~500 bp) flanking SE-A (bold red letters) were generated via PCR with primer sets 935/936 and 675/807 (not depicted). Primers 936 and 675 added an exogenous ApaI restriction site (blue letters), which was used to digest and ligate the fragments together, resulting in a fused DNA product that lacks SE-A (as well as 5 nucleotides upstream of SE-A (red letters)), but contains the ApaI site. This fused DNA segment was inserted into pKV363 (generating plasmid pMSM25) and used to delete SE-A from the genome of ES114 as described in Materials and Methods. (B) SE-A was restored to its native location in the chromosome as described below. pMSM25 was cut with ApaI and primers 1001/1002 were annealed and ligated into the cut plasmid; these primers have sites compatible with the ApaI cut site as indicated in blue. The resulting plasmid, pKV461, was used to restore SE-A (but not the upstream 5 nucleotides) to its native location in the chromosome. Of note, non-native sequences (blue) are present in the final construct, pKV461, and thus in the final SE-A-restored strain KV6716, but they did not affect the ability of this strain to promote biofilm formation.

Materials and Methods

Western blot analysis. Plasmids that contained epitope (HA) tagged *sypA* that either contained or lacked the SE-A sequence upstream of the *sypA* gene were generated via PCR using pARM36 (10) as a template and the indicated primers (Table S1 and S2) using Gibson Assembly (11) (master mix obtained from New England Biolabs, Beverly, MA, USA). The resulting plasmids were then introduced into *V. fischeri*. Western blot analysis was used to evaluate the levels of SypA protein produced when either pVAR91 (+SE-A) or pVAR92 (ΔSE-A) were expressed with either vector control (VC – pVSV105) or a *sypG* plasmid (pARM9) in the WT (ES114) strain. Briefly, cultures were grown overnight at 28˚C with shaking in LBS containing Tc and Cm. Samples were collected and standardized to an OD=2 in 500 µl, then resuspended in 500µl 2X SDS loading buffer (4% SDS, 10% 2-mercaptoethanol, 0.005% bromophenol blue, 20% glycerol, 0.1M Tris pH 7), boiled for 5 min, and loaded (10 µl) onto a 15% SDS polyacrylamide gel. After electrophoresis, proteins were transferred to a polyvinylidene fluoride membrane (PVDF) and probed with anti-HA antibody (Sigma-Aldrich, St. Louis, MO). Protein bands were visualized using a horseradish peroxidase-conjugated secondary antibody and ECL reagents (SuperSignal West Pico Chemiluminescent Substrate, Pierce Biotechnology, Rockford, IL).

References

1. Morris, A. R., and K. L. Visick. 2013. Inhibition of SypG-induced biofilms and host colonization by the negative regulator SypE in *Vibrio fischeri*. PLoS One 8:e60076.
2. Ray, V. A., and K. L. Visick. 2012. LuxU connects quorum sensing to biofilm formation in *Vibrio fischeri*. Mol Microbiol 86:954-970.
3. Stabb, E. V., and E. G. Ruby. 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods Enzymol 358:413-26.
4. McCann, J., E. V. Stabb, D. S. Millikan, and E. G. Ruby. 2003. Population dynamics of *Vibrio fischeri* during infection of *Euprymna scolopes*. Appl Environ Microbiol 69:5928-34.
5. Visick, K. L., and L. M. Skoufos. 2001. Two-component sensor required for normal symbiotic colonization of *Euprymna scolopes* by *Vibrio fischeri*. J Bacteriol 183:835-42.
6. Shibata, S., E. S. Yip, K. P. Quirke, J. M. Ondrey, and K. L. Visick. 2012. Roles of the structural symbiosis polysaccharide (*syp*) genes in host colonization, biofilm formation, and polysaccharide biosynthesis in *Vibrio fischeri*. J Bacteriol 194:6736-47.
7. Dunn, A. K., D. S. Millikan, D. M. Adin, J. L. Bose, and E. V. Stabb. 2006. New *rfp*- and pES213-derived tools for analyzing symbiotic Vibrio fischeri reveal patterns of infection and lux expression in situ. Appl Environ Microbiol 72:802-10.
8. Schneider, T. D., and R. M. Stephens. 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18:6097-100.
9. Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner. 2004. WebLogo: a sequence logo generator. Genome Res 14:1188-90.
10. Morris, A. R., and K. L. Visick. 2013. The response regulator SypE controls biofilm formation and colonization through phosphorylation of the *syp*-encoded regulator SypA in *Vibrio fischeri*. Mol Microbiol 87:509-25.
11. Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd, and H. O. Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343-5.