6726–6745 *Nucleic Acids Research, 2018, Vol. 46, No. 13* *Published online 9 June 2018*
doi: 10.1093/nar/gky475

Multifactorial control of the expression of a CRISPR-Cas system by an extracytoplasmic function σ/anti-σ pair and a global regulatory complex

Diego Bernal-Bernal[1,†], Javier Abellón-Ruiz[1,†], Antonio A. Iniesta[1], Elena Pajares-Martínez[1], Eva Bastida-Martínez[1], Marta Fontes[1], S. Padmanabhan[2,*] and Montserrat Elías-Arnanz[1,*]

[1]Departamento de Genética y Microbiología, Área de Genética (Unidad Asociada al IQFR-CSIC), Facultad de Biología, Universidad de Murcia, 30100 Murcia, Spain and [2]Instituto de Química Física 'Rocasolano', Consejo Superior de Investigaciones Científicas (IQFR-CSIC), Serrano 119, 28006 Madrid, Spain

Received April 5, 2018; Revised May 13, 2018; Editorial Decision May 14, 2018; Accepted May 15, 2018

ABSTRACT

Expression of CRISPR-Cas systems is a prerequisite for their defensive role against invading genetic elements. Yet, much remains unknown about how this crucial step is regulated. We describe a new mechanism controlling CRISPR-*cas* expression, which requires an extracytoplasmic function (ECF) σ factor (DdvS), its membrane-bound anti-σ (DdvA) and a global regulatory complex (CarD–CarG). Transcriptomic analyses revealed that the DdvS/CarD/CarG-dependent regulon comprises a type III-B CRISPR-Cas system in *Myxococcus xanthus*. We mapped four DdvS-driven CarD/CarG-dependent promoters, with one lying immediately upstream of the *cas* cluster. Consistent with direct action, DdvS and CarD–CarG localize at these promoters *in vivo*. The *cas* genes are transcribed as a polycistronic mRNA that reads through the leader into the CRISPR array, a putative σ^A-dependent promoter in the leader having negligible activity *in vivo*. Consequently, expression of the entire CRISPR-Cas system and mature CRISPR-RNA (crRNA) production is DdvS/CarD/CarG-dependent. DdvA likely uses its large C-terminal domain to sense and transduce the extracytoplasmic signal triggering CRISPR-*cas* expression, which we show is not starvation-induced multicellular development. An ECF-σ/anti-σ pair and a global regulatory complex provide an effective mechanism to coordinate signal-sensing with production of precursor crRNA, its processing Cas6 endoribonuclease and other Cas proteins for mature crRNA biogenesis and interference.

INTRODUCTION

CRISPR (clustered regularly interspaced short palindromic repeats) genetic loci and their associated (*cas*) genes are adaptive, small RNA-guided 'immune' systems for sequence-specific destruction of nucleic acids in archaea and bacteria (1–5). Their role in the defense against foreign genetic elements, such as plasmids and phages, has been well established (6), and they have been linked to other cellular functions like gene regulation, biofilm formation, multicellular development and virulence (4,7–11). They are also providing revolutionary tools and applications, from microbes to plants and animals, for gene and genome editing to probe gene function and regulation with unprecedented specificity (12).

CRISPR arrays consist of a variable number (usually <50) of short (20–50 bp) repeats separated by similarly sized spacers of diverse sequence acquired from foreign genetic material in a process called 'adaptation', mediated by Cas1 and Cas2 proteins (2). During CRISPR-Cas 'expression', the array is usually transcribed from a promoter in the leader region immediately upstream. This generates a long precursor RNA (pre-crRNA) that is processed by specific Cas endonucleases (generally expressed from a *cas* operon) to small CRISPR RNAs or crRNAs. Mature crRNAs assemble with defined Cas proteins as a ribonucleoprotein 'interference' complex to target nucleic acids with sequence complementarity for degradation (6,13–16). CRISPR-Cas systems are diverse and have been classified thus far into two classes, six types and over 20 subtypes based on locus arrangement and signature *cas* genes (1,17). Types I, III and

*To whom correspondence should be addressed. Tel: +34 868 887 134; Fax: +34 868 883 963; Email: melias@um.es
Correspondence may also be addressed to S. Padmanabhan. Tel: +34 915619400; Fax: +34 915642431; Email: padhu@iqfr.csic.es
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present address: Javier Abellón-Ruiz, Institute for Cell and Molecular Biosciences, The Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK.

IV, with multiprotein crRNA–effector complexes, are class 1 systems; types II, V and VI, with a single protein–crRNA effector complex, are class 2. All CRISPR-Cas systems require Cas proteins and crRNAs for function, and CRISPR-*cas* expression is a prerequisite to acquire new spacers, process pre-crRNA and assemble ribonucleoprotein crRNA interference complexes for target degradation (1–5,13). Yet, much remains to be discovered about the crucial step of how CRISPR-*cas* expression is triggered and controlled *in vivo,* which may occur at different stages of the invasion process (18).

Regulatable CRISPR-*cas* expression is desirable as it would keep the system silent until required (as upon phage infection), economize the cellular costs of expressing such large clusters and avoid unwanted build-up of products, like nucleases, that can be detrimental to the host. However, active CRISPR-*cas* expression in the absence of infection has been reported for *Thermus thermophilus*, *Clostridium difficile, Pectobacterium atrosepticum* and *Legionella pneumophila* under a variety of laboratory growth conditions (19–22). On the other hand, global regulatory factors such as H-NS (histone-like nucleoid-structuring protein) in *Escherichia coli* and some other species, and LRP (leucine-responsive regulatory protein) in *Salmonella enterica* have been implicated in CRISPR-*cas* repression, which must somehow be relieved upon infection (23–26). Changes in host metabolism in response to invasion by foreign elements, sensed via the global regulators CRP (cAMP receptor protein) or LeuO (a LysR-type transcriptional factor), have also been linked to CRISPR-*cas* changes in expression (23,27–30). Interestingly, few specific transcription factors have been implicated to date. These include an activator (Csa3a) and a repressor (Csa3b) in the archaeal *Sulfolobus* (31–33), which are associated with or proximal to the CRISPR-*cas* loci; and DevTRS in *Myxococcus xanthus,* whose genes form part of a *cas* operon and are involved in negative autoregulation linked to spore differentiation within the multicellular fruiting bodies that develop upon starvation (9,11,34). Likewise, quorum-sensing mechanisms were recently shown to regulate CRISPR-*cas* expression in *Pseudomonas aeruginosa* and *Serratia* (35,36). In many of the above bacterial systems, CRISPR-*cas* transcription relies on promoters recognized by RNA polymerase (RNAP) holoenzyme containing the primary σ^{70}/σ^{A} factor (RNAP-σ^{A}). Moreover, the aforementioned regulators generally target *cas* gene promoters with few, if any, CRISPR array-specific regulators known, at least in bacteria (18). The bacterial cell envelope is at the frontline for dealing with extracellular threats or stresses and would be the first to sense any intrusion that triggers CRISPR-*cas* activation. But this is also the least understood step, although membrane perturbation via the two-component systems BaeSR in *E. coli* (37) and VicRK in *Streptococcus mutans* (38) has been linked to CRISPR-*cas* activation.

In the present study, we report a previously undescribed mechanism for regulated expression of a CRISPR-Cas system. We have discovered that an extracytoplasmic function (ECF) σ factor, DdvS, drives the expression of all the *cas* genes as well as of the CRISPR array of a type III-B CRISPR-Cas system in *M. xanthus*, CRISPR4-Cas, one of three such systems found in this bacterium (Figure 1A and Supplementary Figures S1–3). ECF σ factors are alternative σ factors of the bacterial σ^{70} family and generally act in response to specific extracellular signals (39–41). Like typical ECF σ factors, DdvS is negatively regulated by its specific membrane-associated anti-σ factor DdvA (42), which would be expected to directly or indirectly sense and/or transduce the signal that triggers CRISPR4-*cas* expression. Our data indicate that this signal, whose identity remains elusive, is not related to starvation, which triggers development of fruiting bodies, since neither CRISPR4-*cas* expression nor mature crRNA formation was observed during multicellular development and normal fruiting bodies formed even when the CRISPR4-Cas system is artificially expressed. We demonstrate that expression of the DdvS-dependent CRISPR4-Cas system *in vivo* requires CarD and CarG. These two proteins always act in unison as a complex and bind to RNAP and to DNA via CarD, the founding member of a large and important family of RNAP-binding proteins with global regulatory roles, the CarD–CarG complex itself being implicated in the action of several ECF σ factors in *M. xanthus* (42–49). Whereas the only known interaction of CarG is that with the CarD N-terminal domain, which also binds to the RNAP β subunit, the CarD C-terminal domain, an intrinsically disordered, eukaryotic high-mobility group A-like domain with four AT-hook DNA-binding motifs, binds to the minor groove of AT-rich DNA tracts to maximize CarD activity. Our present study identified three promoters dependent on DdvS, CarD and CarG at the CRISPR4-*cas* locus, besides that of the *ddvS–ddvA* pair and, consistent with a direct regulatory action, we show that DdvS and CarD–CarG localize at these promoters *in vivo*. The data also revealed that the *cas* genes form an operon, which is transcribed as a single polycistronic mRNA that reads into the CRISPR4 array. Although the leader region contains a putative RNAP-σ^{A} promoter, its activity was negligible *in vivo*. Moreover, the CRISPR4-Cas system co-expresses with the pre-crRNA a Cas6 homolog that, we show, is required to generate mature crRNAs. Thus, an ECF-σ/anti-σ pair, together with a global regulatory complex, effectively regulate coordinated production of pre-crRNA, its processing Cas6 endoribonuclease and other Cas proteins for mature crRNA biogenesis and interference.

MATERIALS AND METHODS

Strains, plasmids, growth conditions and strain construction

Plasmids and *M. xanthus* strains used in this study are listed, respectively, in Supplementary Tables S1 and 2. *Myxococcus xanthus* vegetative growth was carried out at 33°C in the rich casitone-Tris (CTT) medium (1% Casitone, 10 mM Tris–HCl pH 8.0, 1 mM KH_2PO_4-K_2HPO_4, 8 mM $MgSO_4$; final pH 7.6), with antibiotic (kanamycin, Km, at 40 μg/ml or oxytetracycline, Tc, at 10 μg/ml) when necessary. CTT agar (1.5%) was used for growth on solid medium. For conditional gene expression, the medium was supplied with inducer (0.5 mM vanillate or 1 mM isopropyl-β-d-thiogalactopyranoside-IPTG) as required (50). Fruiting body development was induced on CF agar plates and examined using a Zeiss dissecting microscope. For this, cells grown in CTT to an OD_{550} of 0.6 (~10^8 cells ml^{-1}) were

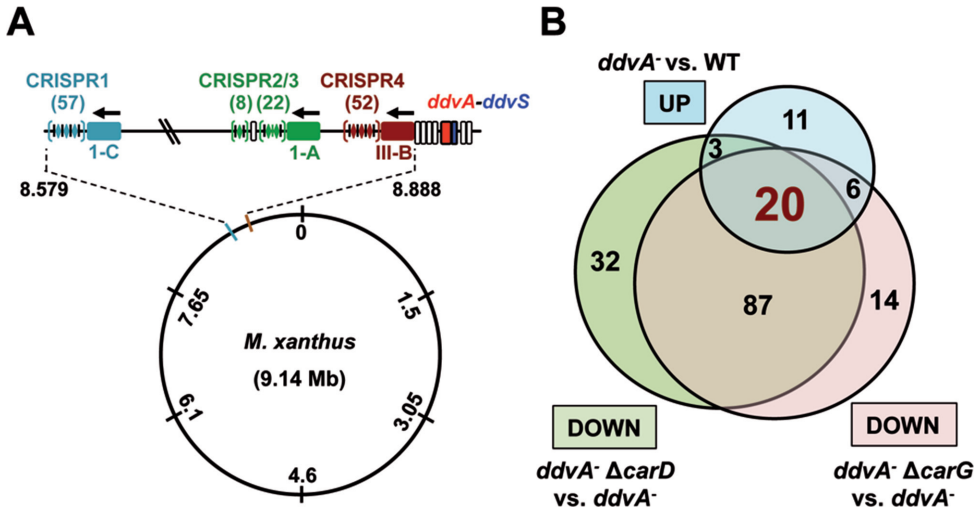

Figure 1. *Myxococcus xanthus* CRISPR-Cas systems and the DdvS regulon. (**A**) The three *M. xanthus* CRISPR-Cas systems and their location in the 9.14 Mb circular chromosome. Colored rectangles indicate the *cas* gene clusters of each system, whose type is labeled below. The four CRISPR arrays and the number of spacers in each are indicated. Vertical blue and red bars indicate genes *ddvS* and *ddvA*, respectively, and the unfilled bars are for genes encoding hypothetical proteins. The arrow points in the direction of transcription. (**B**) Venn diagram representation of microarray transcriptomic data for genes upregulated in the *ddvA*$^-$ strain (MR1543) relative to the WT strain (DK1622; blue), and downregulated in the *ddvA*$^-$ Δ*carD* strain (MR1544) or in the *ddvA*$^-$ Δ*carG* strain (MR1545) relative to MR1543 (green and pink, respectively). Genes and microarray data with differential expression are listed in Table 1 and Supplementary Table S4.

washed twice by centrifugation in TPM (10 mM Tris–HCl pH 8.0, 1 mM KH_2PO_4-K_2HPO_4, 8 mM $MgSO_4$; final pH 7.6) buffer, concentrated 10-fold in TPM buffer, spotted (10 μl) on CF plates and incubated at 33°C. *Escherichia coli* DH5α, used for plasmid constructions, was grown at 37°C in Luria broth supplemented with the appropriate antibiotic. Plasmids were constructed using standard protocols, verified by DNA sequencing, and introduced into *M. xanthus* by electroporation, where they integrate by homologous recombination. Plasmids used for IPTG or vanillate-inducible gene expression and the integrative plasmids with a 1.38 or 1.31-kb *M. xanthus* DNA for integration at a chromosomal site with no promoter activity have been described elsewhere (50). Strains with in-frame gene deletions were constructed following the two-step allele exchange protocol using Km^R negative selection/*galK* positive selection (Gal^S), as described previously (50). The polymerase chain reaction (PCR) overlap extension method was employed for site-directed mutagenesis.

Microarray analysis

Total RNA was isolated for microarray analysis as follows. The *M. xanthus* strain of interest was grown in 50 ml CTT to exponential phase (OD_{550} = 1), and 1 ml (for 5′ RACE) or 6 ml (for transcriptomic analysis) of the culture were then pelleted by centrifugation and stored at −70°C. Frozen cell pellets were resuspended in 300 μl of a 0.3 M sucrose/0.01 M sodium acetate pH 4.5 solution, transferred to an Eppendorf tube containing 300 μl of 2% sodium dodecyl sulfate, 0.01 M sodium acetate pH 4.5, incubated at 65°C for 1–2 min, mixed well with 400 μl of phenol and incubated at 65°C for 3 min after thorough mixing. After snap freezing for 1–2 min in liquid nitrogen, it was thawed, then centrifuged (16 000 *g*, 5 min) and the resulting aqueous layer was pipetted into a fresh tube with 600 μl of hot phenol. The mixing, 3 min-incubation at 65°C, snap-freezing in liquid nitrogen, and centrifugation were repeated as before. The aqueous phase obtained was mixed with 600 μl 1:1 phenol:chloroform in a new tube and centrifuged, and the process repeated using 600 μl pure chloroform. From this last step, the aqueous portion was mixed with 40 μl of 3 M sodium acetate (pH 4.5) and 900 μl of 96% ethanol, followed by 30-min incubation at −20°C and centrifuged (16 000 *g*, 20 min, 4°C). The supernatant was discarded, and the pellet was washed with 200 μl of ice-cold 70% ethanol, dried in a SpeedVac and suspended in 50 μl of 0.2% diethyl pyrocarbonate (DEPC)-treated water. It was then treated with 20 units of RNase-free recombinant DNase I in DNase I buffer and 20 units of Protector RNase Inhibitor (all from Roche) for 30 min and at 37°C. RNA was purified using QIAGEN RNeasy Mini Kit and its quality and quantity were assessed by gel electrophoresis and NanoDrop ND-1000 (Thermo Scientific) using an extinction coefficient at 260 nm of 40 ng-cm/μl.

An 8 × 15k microarray platform format (8 microarrays per slide each with 15 000 probes) was designed using eArray (https://earray.chem.agilent.com/earray/; Agilent Technologies, Santa Clara, USA), and in each microarray spot 7316 out of the 7441 annotated genes in the *M. xanthus* genome (excluding 79 RNA genes, 43 pseudogenes and genes with duplicate annotations) were represented by 60-bp amplicon probes (in duplicate and distributed randomly per microarray, with empty array features filled with randomly chosen amplicon probes). Total RNA isolated as described above was checked for quality and quantity using the Agilent Bioanalyzer 2100 and the Prokaryote Total RNA Nano assay, with those having an RNA integrity number ≥ 7 chosen for further analysis. This RNA was reverse-transcribed to yield cDNA that, after labeling with the cyanine Cy3 (green) or Cy5 (red) fluorescent dyes (GE

Healthcare), was hybridized to the microarray spot. Relative mRNA levels were determined using the Agilent two-color microarray-based prokaryote analysis for gene expression protocols, version 1.3. For each test and reference condition two or more independent microarray experiments were performed and analyzed separately. Microarray hybridization and data analyses were performed at Bioarray SL (Spain). Hybridized microarrays were scanned and the image analyzed to transform color intensity to numeric data using the Agilent Feature Extractor Software v.10.7. Raw intensity data were read and checked for quality using the Limma package for R (51), corrected for background using the *normexp+offset* method (52), and normalized with loess (intra-array) and quantile (inter-array) procedures (53). A simple linear model was fit to the data and variances were corrected using an empirical Bayes approach implemented in the Bioconductor limma package (54). Genes were considered as significantly affected if they displayed a mean $\log_2$ ratio < -1 or $> +1$ (corresponding to 2-fold change up or down) relative to the control and with adjusted *P*-values ≤ 0.05 (95% confidence limit) from two independent array runs. Microarray data have been deposited at NCBI Gene Expression Omnibus (GEO) with accession number GSE112385.

Real-time quantitative reverse transcription PCR (qRT-PCR)

Total RNA (2 μg) isolated as above was reverse transcribed to cDNA using random hexamer primers (Promega) and Transcriptor Reverse Transcriptase (Roche) in a 20 μl reaction mix as per instructions. The cDNA sample (2 μl) was added to a PCR mixture prepared using Power SYBR® Green PCR Master Mix (Applied Biosystems) that contained the respective primers at 100 nM. Reactions were performed in triplicate for cDNA obtained from each biological replicate; a control with an equivalent volume of RNA, but not reverse transcribed, was tested to rule out the presence of contaminant DNA. Reported values are the average and the standard error for three biological replicates. Primers (Supplementary Table S3) to amplify ∼50–150 bp region within each transcript were designed (16S rRNA was the internal standard) using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). RNA quantitation was carried out in an Applied Biosystems StepOne equipment using the 1-Step RT-PCR program cycle without the reverse transcription step. Melting and dissociation curves were determined from 60–95°C, 30 s and 95°C, 15 s. Each primer pair was tested for RT-PCR analysis on a standard curve generated from five 10-fold serial dilutions of cDNA. Only primer pairs with efficiency close to 100% were used and data were analyzed using the Applied Biosystems system software.

5′ RACE and RT-PCR analysis

Total RNA was isolated from the Δ*ddvA* or wild-type (WT) strain using High Pure RNA kit (Roche) or RNeasy Mini kit (Qiagen) as described above, and RNA recovered after two rounds of treatment with recombinant DNase I (Roche) to fully eliminate genomic DNA was suspended in 50 μl RNase-free water (pre-treated with 0.2% DEPC). For 5′ RACE (rapid amplification of cDNA ends), 2 μg of this purified RNA was reverse-transcribed using an oligonucleotide primer that hybridizes to the 3′ end of the gene of interest and the 5′/3′ RACE 2nd Generation kit in the presence of RNase inhibitor Protector (Roche). The cDNA product was purified, tagged with a 5′ polyA adaptor and used as template in a PCR reaction with, as primers, dT-Anchor and a cDNA-specific oligonucleotide that hybridizes to the 3′ end of the gene of interest, followed by a second PCR reaction with PCR Anchor instead of the dT-Anchor. The product was purified using High Pure PCR Product Purification Kit (Roche) and sequenced. For RT-PCR, 500 ng total RNA was reverse-transcribed using random hexamers primers (Promega) and Transcriptor Reverse Transcriptase (Roche). A total of 2 μl of the cDNA product was used as template for subsequent PCR using primer pairs that generate fragments of ≤500 bp spanning across the junctions of the annotated *cas* genes or across the leader to CRISPR4 spacers. As the negative (−RT) control, RNA that was not subjected to reverse transcription, but otherwise treated identically, was used as template in the PCR reaction.

β-galactosidase activity

Specific β-galactosidase activity (in nanomoles of o-nitrophenyl β-D-galactoside hydrolyzed/min/mg protein and at least three independent measurements) was performed on cells grown in CTT to early exponential phase or on cells undergoing development using a SpectraMax 340 microtitre plate reader (Molecular Devices), as described elsewhere (45). X-Gal (5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside)-agarose overlay assay for qualitative analysis of reporter *lacZ* expression was carried out as reported previously (50).

Chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR)

Cells grown in 50 ml CTT (with or without vanillate as required) at 33°C to mid-late exponential phase (OD_{550} = 0.70) were treated with rifampicin (25 μg/ml) for 30 min and then cross-linked with formaldehyde (1% v/v; Sigma-Aldrich) in the presence of 10 mM sodium phosphate (pH 7.6) for 10 min at room temperature with shaking (100 rpm). After cooling in ice for 30 min, cells were pelleted, washed twice with phosphate-buffered saline solution and stored at −80°C until further use. For chromatin immunoprecipitation (ChIP), the frozen pellet was thawed, resuspended in lysis buffer (470 μl, 10 mM Tris pH 7.5, 1 mM ethylenediaminetetraacetic acid (EDTA), 100 mM NaCl, 2.2 mg/ml lysozyme) and incubated at room temperature with shaking (100 rpm) for 30 min. It was mixed with ChIP buffer (550 μl, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8.1, 167 mM NaCl, 1X Roche complete protease inhibitor cocktail), incubated for 10 min at 37°C and sonicated (sixty 30 s on/30 s off cycles) in a Bioruptor (Diagenode) to generate fragments of ∼0.5 kb. The sample was clarified by centrifugation and an aliquot of 20 μl supernatant was kept aside

as the input sample. The rest was mixed with 550 μl ChIP buffer with 0.01% sodium dodecyl sulphate (SDS) and 5 μl of monoclonal anti-FLAG (Sigma-Aldrich) or anti-CarD antibodies (46), incubated at 4°C overnight with rotation and immunoprecipitated (5–6 h, 4°C) with rotation with 30 μl of protein A magnetic Dynabeads (Life Technologies) previously washed with PBS plus 1 mg/ml bovine serum albumin. The beads were then washed once each with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.1, 150 mM NaCl), the same buffer at 0.5 M NaCl, and with LiCl buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.1), and twice with TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). The protein–DNA complex was recovered in two 100 μl fractions of elution buffer (1% SDS, 0.1 M $NaHCO_3$) and incubated at 65°C for 10 min to disrupt the cross-links, treated with 5 μl proteinase K (20 μg/μl) at 42°C for 2 h and at 65°C for 6 h, followed by DNA isolation using High Pure PCR product purification kit (Roche). The input sample was also subjected to this cross-link reversal and DNA extraction protocol. qRT-PCR was then carried out using the SYBR Green reaction mix (BioRad) and specific primers (Supplementary Table S3) in 0.1 ml MicroAMP FAST optical 48-well reaction plates in a StepOne qPCR apparatus (Applied Biosystems). Standard curves were obtained for each DNA region of interest and its corresponding primer pair with serial dilutions of the input DNA sample. Signal enrichment at each promoter was estimated as: (i) percent of input and normalized relative to the sample obtained with cells lacking the immunoprecipitated protein; or (ii) ratio of the promoter-specific to intragenic signals of ChIP reactions relative to the values for the input sample; the intragenic region used as the non-promoter control spans nucleotides 402–460 of the *fruA* gene (55). Reported values are the mean and standard error from three independent experiments.

Electrophoretic mobility shift assays (EMSA)

Purification of CarD, CarG and *M. xanthus* RNAP-σ^A and electrophoretic mobility shift assays (EMSA) were carried out as described previously (45,56). The ^{32}P-5′-end radiolabeled double-stranded DNA probes used for EMSA were obtained by PCR-amplification from plasmid constructs bearing the required regions as DNA template. The DNA probes thus obtained were: (i) 304-bp segments corresponding to the entire CRISPR4 leader with the WT or mutant (Mut) putative −35 promoter element at the leader, to test for RNAP-σ^A binding; (ii) 170-bp segments covering the DdvS-dependent promoter regions, to test for CarD and CarD–CarG binding. Samples (20 μl) in EMSA buffer (80 mM KCl, 25 mM Tris pH 8.0, 5 mM $MgCl_2$, 1 mM dithiothreitol, 10% glycerol, 200 ng/μl bovine serum albumin) with 1 μg non-specific competitor poly[dG-dC] and 1 nM of a given radiolabeled DNA probe (∼13 000 cpm) were incubated for 30 min with RNAP-σ^A (100 nM) at 37°C, or CarD (440 or 520 nM) with or without CarG (12 μM) at 4°C. Then, to samples containing RNAP-σ^A, heparin (1 μg) was added and incubated for an additional 5 min to allow formation of only heparin-resistant open promoter complexes, followed by electrophoresis in 4% non-denaturing polyacrylamide gels at 200 V for 1.5 h in TBE buffer (45 mM Tris and boric acid, 1 mM EDTA) at 10°C. Gels were vacuum-dried and analyzed by autoradiography.

Northern blot analysis of pre-crRNA and mature crRNAs

Total RNA from 5 ml cell cultures in CTT (OD_{550} = 0.6) was isolated with 1 ml of Trizol reagent (Invitrogen) and 0.2 ml of chloroform (Sigma-Aldrich) and precipitated with 0.5 ml isopropanol (100%). The pellet was washed with 1 ml of pre-cooled 75% ethanol, resuspended in 50 μl of DEPC-treated water and RNA concentration determined for pre-crRNA analysis. To isolate RNA from fruiting bodies at each time point analyzed, cells from 10 ml CTT cultures (OD_{550} = 0.6) were washed and suspended in 2 ml TPM, then distributed on four CF plates as 20 μl spots and incubated at 33°C. At given times, developing cells were collected in tubes containing 1 ml TPM and a 0.5 ml equivalent of glass beads (0.1 mm diameter) for lysis using a Mini-beadbeater (BioSpec), and RNA was recovered from the supernatant as before. For mature crRNA detection, the total RNA was resuspended in 300 μl DEPC-treated water and large RNA molecules were selectively precipitated out by incubating with 5% polyethylene glycol 8000 and 0.5 M NaCl for 30 min in ice. The supernatant, recovered by centrifugation, was mixed with 100% ethanol (three volumes) and 3 M sodium acetate (pH 5; 0.1 volume), incubated overnight at −20°C to precipitate small RNAs, washed twice with 80% ethanol, dried and resuspended in DEPC-treated water. The RNA concentration was determined and northern blot analysis of mature crRNA was carried out as described below.

To detect pre-crRNA, ∼20 μg of total RNA obtained above was electrophoresed in a 1.2% agarose gel with 20 mM MOPS/5 mM sodium acetate/1 mM EDTA/7% formaldehyde buffer. The gel was washed with water (six changes, 90 min), followed by capillary transfer to a Hybond N+ membrane (GE Lifesciences) overnight and cross-linked by UV-irradiation (1.2×10^5 μJ/cm^2; Hoefer UVC 500 apparatus). Pre-crRNA was detected using a radiolabeled double-stranded 1044-bp DNA probe (spanning the segment from CRISPR4 spacer 37 to spacer 52) generated using DNA polymerase I Klenow (Takara) and α-[^{32}P]-dCTP (specific activity 3000 Ci/mmol) following standard protocols. 23S RNA, probed using a specific 1078-bp double-stranded DNA probe, was used as the loading control. Probes were incubated at 95°C for 5 min and rapidly cooled in ice prior to use. The pre-crRNA blot was incubated for 2 h at 65°C in 0.9 M NaCl, 1% SDS, 100 μg/ml sheared, denatured salmon sperm DNA and hybridized overnight at 65°C with probe in fresh buffer containing 10% dextran sulfate sodium salt and 50 μg/ml of the non-specific DNA. After hybridization, the blots were washed (3–5 times over 20 min) with 0.3 M NaCl, 0.03 M sodium citrate (SSC) buffer containing 0.1% SDS at 65°C, and analyzed by autoradiography (Kodak X-OMAT film; intensifying screen at −70°C). For northern blots of mature crRNAs, RNA (10 μg) free of large RNAs obtained above was electrophoresed alongside 5′-end [^{32}P]-radiolabeled RNA markers (Decade, Ambion) in a 7 M urea-15% polyacrylamide gel (19:1 acrylamide:bisacrylamide) with TBE buffer, electrotrans-

ferred to a Hybond N+ membrane (250 mA for 45 min in TBE) and UV cross-linked as before. Single-stranded synthetic DNA probes (20 pmol) to detect mature crRNA corresponding to the CRISPR4 leader-proximal spacer 52 (5′-TGCATGAGTCGAAAGAGATTTTGAAGCGC CGGAC-3′) and leader-distal spacer 1 (5′-GTCTGGCAA CACCAGAATGGCGGAACCGACTA-3′) were 5′-end-labeled with γ-[^{32}P]-ATP (6000 Ci/mmol specific activity; Perkin Elmer) and T4 polynucleotide kinase (Takara) using standard protocols. The loading control was 5S RNA probed with 5′-CTTAACTTCCGTGTTCGGGAT GGGAACGGGTGGGAC-3′. The mature crRNA blot was treated for 2 h at 30°C with 40% formamide, 7% SDS, 0.3 M NaCl, 0.05 M sodium phosphate pH 7 and Denhardt's solution (57) with 100 μg/ml sheared, denatured salmon sperm DNA (58) and hybridized with specific probes in the same solution for 16 h at 30°C. The blots were washed with SSC buffer containing 0.2% SDS solution at 50°C, and analyzed by autoradiography as before.

Bioinformatic analysis

Database searches and analysis were carried out with BLASTN (for nucleotide sequences) and DELTA-BLAST (for proteins) at http://blast.ncbi.nlm.nih.gov/Blast.cgi. CRISPR-Cas data were analyzed using CRISPI (http://crispi.genouest.org/; (59)) and CRISPRTarget (http://bioanalysis.otago.ac.nz/CRISPRTarget/crispr_analysis.html (60)). Genomic context and putative operons in the *M. xanthus* genome were analyzed using Artemis software (http://www.sanger.ac.uk/resources/software/artemis (61)). Putative DdvS promoters within a maximum distance to gene of 200 bp were identified in the *M. xanthus* genomic region from positions 8909344 to 8883115 (~26 kb at the *ddvS–ddvA* CRISPR4-Cas locus) with as query the sequence GTAAn$_{16}$CGT (n is any deoxyribonucleotide) in Virtual Footprint Regulon Prediction (http://prodoric.tu-bs.de/vfp; (62)). RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi; (63)) and mfold (http://unafold.rna.albany.edu/?q=mfold; (64) servers were used for secondary structure predictions.

RESULTS

The ECF σ factor DdvS and the CarD–CarG global regulatory complex upregulate expression of the *cas* genes of a type III-B *M. xanthus* CRISPR-Cas system

The *ddvS–ddvA* (genome locus tag MXAN_7289-MXAN_7288) pair was first identified in an early Tn5 *lac* insertion analysis to screen for CarD-dependent genes in *M. xanthus* and then shown to depend on CarG as well (47,65). We demonstrated that DdvS is an ECF σ factor that drives its own expression and that of its translationally coupled cognate anti-σ, DdvA, in a manner that requires CarD and CarG; and that DdvA is a single-pass transmembrane protein, whose ~70-residue cytoplasmic N-terminal ZAS (zinc anti-sigma) domain physically interacts with DdvS to sequester it and maintain it inactive ((42); Supplementary Figure S3). DdvS and DdvA thus constitute an ECF σ/anti-σ pair (42). Like typical anti-σ of ECF σ factors, inactivation of DdvA to release DdvS would be expected to occur on sensing a specific external signal (40), presumably via its large (~900-residue) extracytoplasmic domain (Supplementary Figure S3). Since the exact identity of such a signal is unknown, our strategy to identify the DdvS regulon was to disrupt *ddvA*. Transcriptomic analyses using WT *M. xanthus* and its derivative MR1543 with a Tn5 *lac* disruption of *ddvA* (*ddvA*$^-$ allele), and MR1544 and MR1545 with deletions of *carD* (Δ*carD*) or *carG* (Δ*carG*), respectively, were performed to assign differentially expressed genes between these strains ('Materials and Methods' section). Relative to the WT, about 40 and 63 genes were upregulated and downregulated, respectively, in MR1543; and relative to the latter, 48 and 50 genes were upregulated in MR1544 and MR1545, respectively, while 142 and 127 genes were downregulated (Figure 1B and Supplementary Table S4). Of the 40 genes that were upregulated in MR1543, which could be directly or indirectly activated by DdvS, 20 (including *ddvS–ddvA*) were also downregulated in MR1544 and MR1545, indicating that their expression not only depends on DdvS but also on CarD and CarG (Figure 1B and Table 1). The majority of these 20 genes cluster at the *ddvS–ddvA* locus and are the ones with the highest fold-change in MR1543 relative to the WT (6- to 600-fold change; Table 1 and Supplementary Table S4). Two, MXAN_7291 and MXAN_7290, lie immediately upstream of *ddvS*, are translationally coupled and encode hypothetical proteins. The other 12, downstream of *ddvA*, include four contiguous genes for hypothetical proteins (MXAN_7287 to MXAN_7284) and further downstream, interestingly, the *cas* (*cmr*) genes (MXAN_7283 to MXAN_7276) of the CRISPR4-Cas type III-B system, one of the three CRISPR-Cas systems in *M. xanthus* (Figures 1A and 2A; Supplementary Figure S1). An enhanced expression in MR1543 relative to the WT, albeit lower, is also observed for five other genes at two different loci (MXAN_1892: a putative serine-threonine kinase; MXAN_1894: DNA-binding protein; MXAN_1810, MXAN_1891 and MXAN_1893: hypothetical proteins). Thus, microarray data suggest an unprecedented regulation of a CRISPR-Cas system by an ECF σ/anti-σ pair and a global regulatory complex, and we focused on this for further analysis below.

Four DdvS-driven promoters dependent on CarD–CarG occur at the CRISPR4-*cas* locus

To understand how DdvS controls expression of the genes in its vicinity, we first checked for the presence of DdvS-dependent promoter(s) in the corresponding genomic region. We have previously mapped the promoter of the *ddvS–ddvA* operon, P_{ddvSA} and have shown by mutational analysis that the AA and GT at the −35 and −10 promoter regions, respectively, are crucial elements ((42); Figure 2A and B). These bases form part of the AAc and cGT motifs (less conserved bases in lowercase) in the −35 and −10 regions, respectively, shared by several promoters dependent on other ECF σ factors in *M. xanthus* (42), as well as in other bacteria (66,67), with additional determinants around these conserved motifs generally providing promoter selectivity. An *in silico* scan of the *M. xanthus* genomic region spanning from ~1 kb upstream of MXAN_7291 to

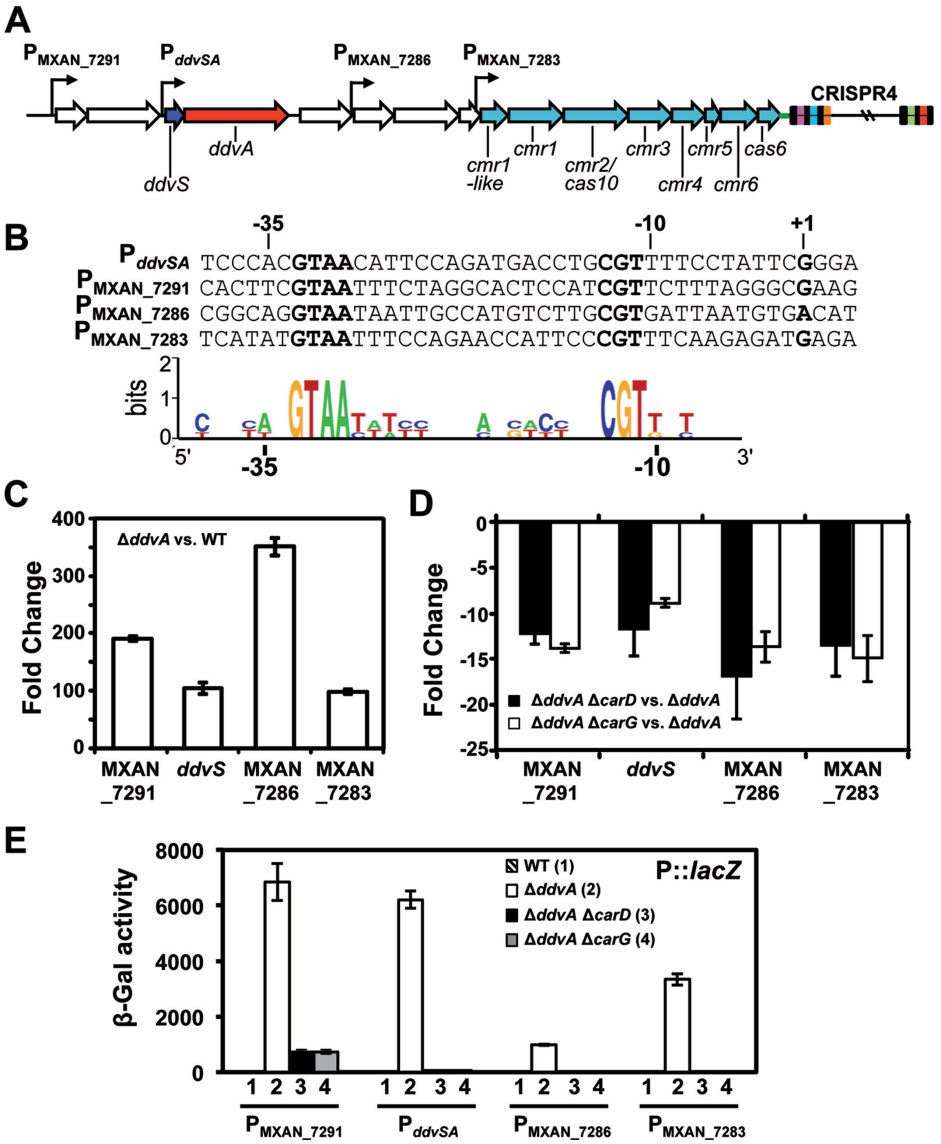

Figure 2. Analysis of DdvS-dependent promoters. (**A**) Schematic of the *Myxococcus xanthus* CRISPR4-Cas system and the DdvS-dependent promoters (right-angled arrows), P_{ddvSA}, and the three (P_{MXAN_7291}, P_{MXAN_7286} and P_{MXAN_7283}) identified in this study. Thick arrows (pointing in the direction of transcription) in cyan are *cas*/*cmr* genes, as labeled, and thick unfilled arrows are genes encoding hypothetical proteins. CRISPR4 array repeats are shown as black vertical bars and their spacers as colored bars. The leader (between *cas6* and the array) is shown as a green line. (**B**) Sequences of P_{ddvSA}, P_{MXAN_7291}, P_{MXAN_7286} and P_{MXAN_7283} with the GTAA and CGT of the $GTAAn_{16}CGT$ pattern used in the *in silico* search in boldface, with a Weblogo representation based on these below. The putative −10 and −35 regions were assigned by 5′-RACE mapping of the transcriptional start sites (+1 in bold; see text and Supplementary Figure S4). For P_{MXAN_7283}, the G at +1 and that of the annotated MXAN_7283 initiator ATG codon coincide, suggesting that the true initiator codon is probably an in-frame TTG eight codons downstream, which has a consensus ribosomal binding site. (**C**) Validation of transcript levels using qRT-PCR of MXAN_7291, *ddvS*, MXAN_7286 and MXAN_7283 in the Δ*ddvA* strain (MR1316) versus the WT (DK1050). (**D**) Same as in (C) for Δ*ddvA* Δ*carD* (MR1317) (black bar) or Δ*ddvA* Δ*carG* (MR2384) (unfilled bar) versus Δ*ddvA*. Fold change is the mean and standard error of three independent experiments. (**E**) Reporter *lacZ* expression (specific β-galactosidase activity, 'β-Gal activity', in nanomoles of o-nitrophenyl β-D-galactoside hydrolyzed/min/mg protein) of *lacZ* fusions to a fragment spanning positions (relative to the initiation codon of the corresponding gene): −400 to +63 at P_{ddvSA}, −400 to +130 at P_{MXAN_7283} and at P_{MXAN_7286}, and −292 to +197 at P_{MXAN_7291}. Fusions are integrated at a heterologous site (see text) in the WT (1), Δ*ddvA* (2), Δ*ddvA* Δ*carD* (3) and Δ*ddvA* Δ*carG* (4) strains. Activities correspond to the mean and standard error of three independent measurements.

∼1 kb downstream of the CRISPR4 array using as query the sequence $GTAAn_{16}CGT$, which matches that at P_{ddvSA}, yielded three new hits. These were within 50 bp upstream of the initiator codon of MXAN_7291, MXAN_7286, and MXAN_7283 (Figure 2A and B), whose expression was up-regulated by DdvS, CarD and CarG according to the microarray analysis (Table 1) and its further validation by qRT-PCR (Figure 2C and D). The use of strains with a Δ*ddvA* allele (in-frame deletion of *ddvA*) in the qRT-PCR analysis, rather than the Tn*5 lac* disrupted *ddvA*$^-$ allele used in the microarray analysis, may explain the different fold changes. However, it is less clear why fold changes from qRT-PCR for MXAN_7286 are larger than for the other genes tested (Figure 2C and D). Nonetheless, both microarray and qRT-PCR indicate that expression of all four genes

Table 1. Genes activated by DdvS and downregulated in the absence of CarD or CarG

Locus_tag	Annotation	FC (MR1543 versus DK1622)	FC (MR1544 versus MR1543)	FC (MR1545 versus MR1543)
MXAN_7291	Hypothetical protein	572.45	−4.66	−7.32
MXAN_7290	Hypothetical protein	285.32	−5.39	−4.15
MXAN_7289	DdvS	204.52	−5.75	−4.67
MXAN_7286	Hypothetical protein	122.96	−7.31	−8.73
MXAN_7283	Putative CRISPR-associated protein	66.61	−4.46	−4.25
MXAN_7282	CRISPR-associated RAMP Cmr1 family protein	65.46	−4.38	−4.09
MXAN_7285	Hypothetical protein/ThiF family protein	58.51	−7.15	−6.56
MXAN_7280	Putative CRISPR-associated Crm3 family protein	33.94	−3.40	−3.51
MXAN_7279	CRISPR-associated RAMP Cmr4 family protein	31.99	−3.11	–
MXAN_7277	CRISPR-associated RAMP Cmr6 family protein	25.10	−3.43	−3.09
MXAN_7284	Hypothetical protein	21.04	−5.27	−5.36
MXAN_7281	CRISPR-associated Cmr2/Cas10 family protein	18.32	−3.42	−3.62
MXAN_7276	CRISPR-associated Cas6 family protein	17.53	−3.25	−2.82
MXAN_7287	Hypothetical protein	8.61	−2.44	−2.32
MXAN_7288	DdvA	8.35	−2.92	−2.06
MXAN_7278	CRISPR-associated Cmr5 family protein	6.66	−2.88	−2.64
MXAN_1810	RDD family protein	3.50	−10.54	−3.61
MXAN_1893	Hypothetical protein	3.25	−7.52	−3.34
MXAN_1892	Putative serine/threonine protein kinase	3.16	−9. 49	−3.33
MXAN_1891	Hypothetical protein	2.69	−5.43	−2.96
MXAN_1894	DNA-binding protein	2.21	−3.05	−2.22
MXAN_5622	CarD	–	−55.39	–
MXAN_5623	CarG	–	–	−20.37

FC: Fold-change in expression in the strain indicated first relative to the second strain listed. Note that the negative sign in the entries of the two 'FC' columns on the right is used to only indicate downregulation.

is DdvS/CarD/CarG-dependent, and this was further supported by the data below.

To corroborate that the new hits correspond to DdvS-dependent promoters whose expression also requires CarD and CarG, we fused DNA segments (∼500 bp) including the putative promoters to a reporter *lacZ* gene in plasmid pMR3183 (50), which also contains a 1.38 kb *M. xanthus* DNA fragment with no promoter activity for integration into the chromosome at a heterologous site. Each plasmid construct was introduced into the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* and Δ*ddvA* Δ*carG* strains. All three fusions (P_{MXAN_7291}::*lacZ*, P_{MXAN_7286}::*lacZ* and P_{MXAN_7283}::*lacZ*) exhibited significant reporter *lacZ* activity in the Δ*ddvA* strain, like the positive control P_{ddvSA}::*lacZ* (42), but little or no activity in the WT, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains (Figure 2E). Moreover, mapping of the transcription start site in each of these constructs using 5′ RACE ('Materials and Methods' section) confirmed that the motifs identified from the sequence analysis corresponded well with the −35 and −10 promoter regions (Figure 2B and Supplementary Figure S4).

Direct action of DdvS and the CarD–CarG complex at the four DdvS-dependent promoters would require them to localize at these promoters *in vivo*, and we examined this by ChIP-qPCR. For DdvS, we used a FLAG-tagged version (shown to be functional *in vivo*; Supplementary Figure S5A) to enable immunoprecipitation with anti-FLAG antibodies. Furthermore, to achieve stimulus-independent activation of FLAG-DdvS and allow its conditional expression, the gene was placed under the control of a vanillate-inducible promoter and integrated into the chromosome at the heterologous site mentioned earlier (50). The induced expression of DdvS alone overcomes the stoichiometric sequestration by DdvA to provide free and active DdvS (42). ChIP-qPCR analysis using anti-FLAG antibodies of cells grown to mid-exponential phase in the presence of vanillate, and treated with rifampicin to trap RNAP complexes at promoters (68), revealed a six to seven-fold enrichment of DdvS at all four DdvS-dependent promoters relative to cells grown without vanillate (Figure 3A).

Since CarD and CarG always act in concert as a stable complex (45,47), we tested whether the complex localizes *in vivo* at these promoters by immunoprecipitating its CarD component from the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* and Δ*ddvA* Δ*carG* strains using anti-CarD monoclonal antibodies (46). This ChIP-qPCR analysis revealed about a 5-fold enrichment of CarD, relative to the Δ*ddvA* Δ*carD* strain (the negative control), at all four promoters in the Δ*ddvA* but not in the WT or Δ*ddvA* Δ*carG* strains (Figure 3B); comparable data were obtained when tested using an intragenic region, where CarD is not expected to be enriched (Supplementary Figure S5B). The data thus indicate that *in vivo* CarD preferentially localizes at these DdvS-driven promoters when actively transcribed and if CarG is also present, suggesting a direct positive role of the CarD–CarG complex on the expression from these promoters. Localization of the CarD–CarG complex at these promoters can be mediated by interactions, as noted in the Introduction, of CarD with RNAP as well as with DNA. Since CarD binds to the minor groove of AT-rich DNA and such AT-rich regions occur in the vicinity of the DdvS-driven promoters, we tested their binding *in vitro* to CarD alone or in the presence of CarG. EMSA using DNA probes covering these promoter regions confirmed the binding of CarD and

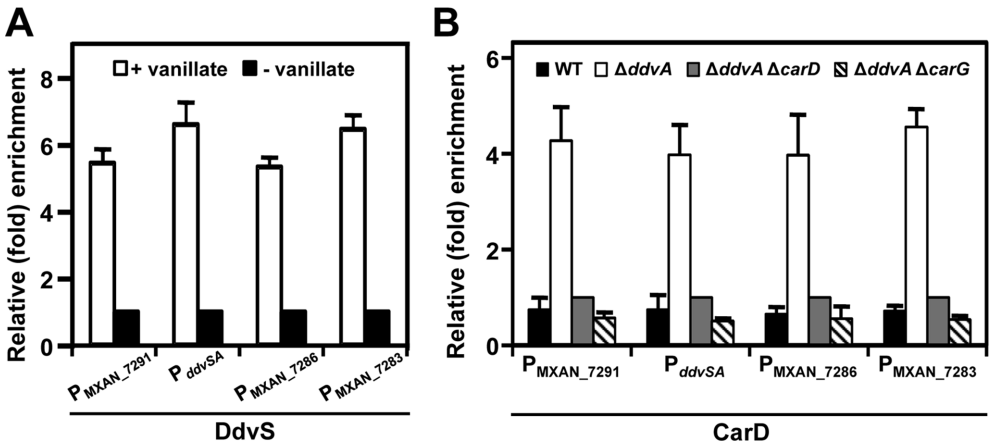

Figure 3. DdvS and the CarD–CarG complex localize at four DdvS/CarD/CarG-dependent promoters *in vivo*. (**A**) Occupancy of DdvS at the four promoters. ChIP of FLAG-DdvS was performed using M2 anti-FLAG antibodies, and enrichment at each promoter is the percent input for cells with P_{van}::*flag-ddvS* at a heterologous site (strain MR2971) grown in the presence of vanillate (unfilled bars) relative to that in its absence (black bars). (**B**) ChIP-qPCR data for enrichment of CarD at the four promoters. CarD was immunoprecipitated from cell cultures of strains DK1622, MR2590 (Δ*ddvA*), MR2766 (Δ*ddvA* Δ*carD*) and MR2767 (Δ*ddvA* Δ*carG*) using monoclonal anti-CarD antibodies and enrichment at each promoter is the percent input relative to the negative control (Δ*ddvA* Δ*carD* strain). Data are the mean and standard error from three independent experiments.

of the CarD–CarG complex (Supplementary Figure S5C and D). Although CarD alone binds to DNA *in vitro*, its enrichment at DdvS-dependent promoters *in vivo* probably owes to additional interactions with CarG and DdvS-associated RNAP. Altogether, the above data establish the presence of four promoters that depend directly on DdvS, CarD and CarG in the vicinity of the CRISPR4-*cas* locus, one of which (P_{MXAN_7283}) is directly upstream of the *cas*/*cmr* cluster.

The CRISPR4 array leader has negligible promoter activity

The data above indicated that expression of the *cas* genes of the CRISPR4-Cas system depends on DdvS and the CarD–CarG complex but how the CRISPR4 array is transcribed remained to be addressed. This has been described to occur in some bacteria from primary σ^{70}/σ^{A}-dependent promoters located in the CRISPR leader (25,69), as noted in the 'Introduction' section. Our inspection of the 304-bp CRISPR4 leader sequence did not reveal any promoter resembling P_{ddvSA}. However, we did notice a segment with a TTGACA sequence spaced 17 bp from a TTTATG downstream (Figure 4A) that would conform well to the −35 and −10 elements, respectively, of known *M. xanthus* σ^{A}-dependent promoters (56,70). We therefore first tested if this putative promoter is recognized *in vitro* using purified *M. xanthus* RNAP-σ^{A} holoenzyme in EMSA with ^{32}P 5′-labeled DNA probes corresponding to the 304-bp WT CRISPR4 leader segment or a variant (Mut) in which the TTG of the putative −35 region was mutated to CCA. In the assay, the labeled DNA probe and RNAP-σ^{A} holoenzyme were first incubated and heparin was then added to eliminate non-specific and closed promoter complexes, leaving only stable, transcriptionally competent open promoter–RNAP-σ^{A} (RP_o) complexes (56). This revealed a retarded band with the WT probe that was barely detected with the Mut probe (Figure 4B), suggesting that the proposed elements are recognized by RNAP-σ^{A} *in vitro*.

Next, we tested if this putative promoter in the CRISPR4 leader was active *in vivo* by generating a plasmid construct with a reporter *lacZ* gene fusion to the entire leader segment (Leader::*lacZ*) or with the leader bearing the mutation at the −35 region of its putative σ^{A} promoter (Leader(Mut)::*lacZ*). The two plasmids were then independently electroporated into the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* and Δ*ddvA* Δ*carG* strains, where they would integrate at the heterologous site mentioned earlier (Figure 4C). Both fusions exhibited negligible reporter *lacZ* expression, and hence promoter activity, in all four strains (Figure 4D), which was surprising given that RNAP-σ^{A} appeared to recognize *in vitro* the putative σ^{A} promoter in the leader. Since this lack of reporter activity might be due to absence of some critical element(s) at the heterologous site, we generated plasmid constructs that allow integration of the reporter Leader::*lacZ* or Leader(Mut)::*lacZ* at the endogenous site by homologous recombination, and introduced them into the WT, Δ*ddvA*, Δ*ddvA* Δ*carD*, and Δ*ddvA* Δ*carG* strains (Figure 4C). Significant Leader::*lacZ* reporter activity was now observed in the Δ*ddvA* strain but not in the WT, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains, and this remained unchanged with the Leader(Mut)::*lacZ* reporter (Figure 4D). Thus, the putative σ^{A} promoter in the leader has a marginal contribution, if any, to the expression of the CRISPR4 array *in vivo*.

The CRISPR4 array is expressed by transcriptional readthrough from the *cas* operon

Lack of promoter activity in the leader raises the possibility that transcription of the *cas* genes reads through the leader into the CRISPR4 array. If so, DdvS, CarD and CarG would coordinate transcription of the CRISPR4 array with that of the *cas* genes. To test this, we first generated strains with the appropriate genetic background in which a promoter-less *lacZ* gene was fused downstream of *cas6*, the gene at the 3′-end of the *cas* cluster and immediately upstream of the 5′-end of the leader. Reporter *lacZ* was

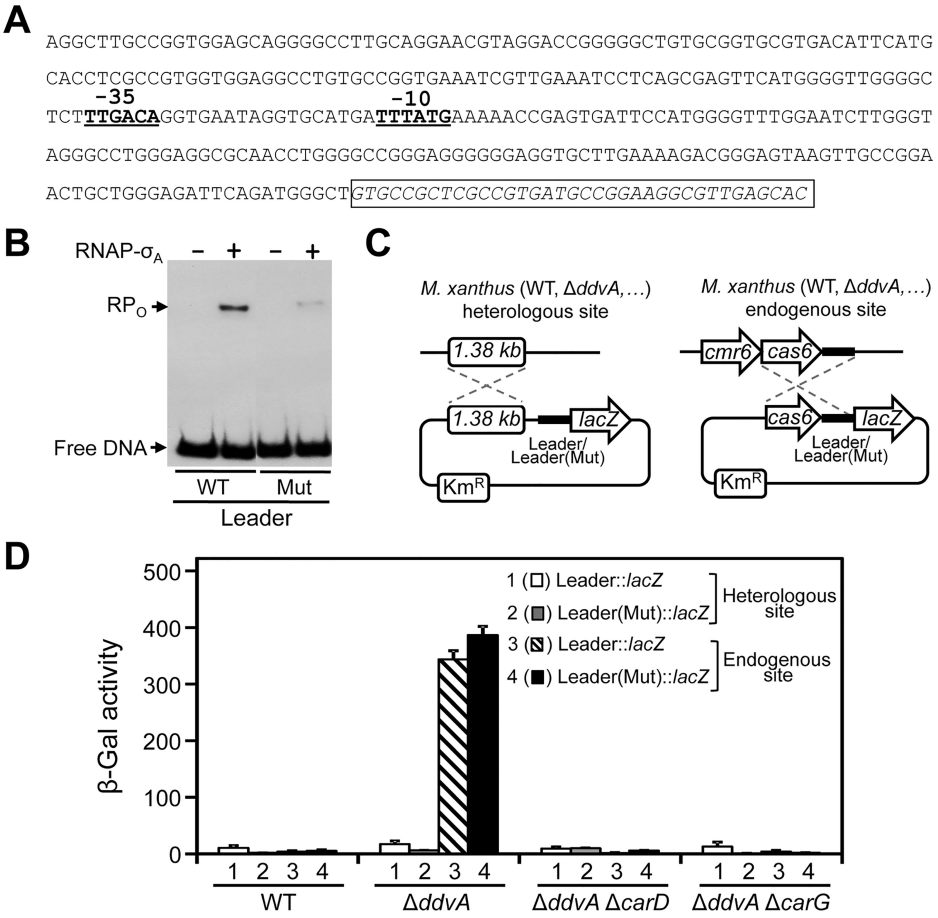

Figure 4. A putative σ^A-dependent promoter in the CRISPR4 leader has negligible activity *in vivo*. (**A**) Sequence of the 304-bp leader segment extending from the *cas6* 3′-end to the first CRISPR4 repeat (italics and boxed). The putative σ^A-promoter −35 and −10 elements are indicated. (**B**) Representative EMSA gel for the binding of RNAP-σ^A (100 nM) to the 5′-end ^{32}P-labeled 304-bp WT leader probe and its variant (Mut) with TTG of the −35 element of the putative σ^A-promoter mutated to CCA. Samples were incubated at 37°C in the presence of poly[dG-dC] as non-specific competitor for 30 min and then, after adding 1 μg heparin, for a further 5 min, followed by electrophoresis. (**C**) Schemes for the strategy used to introduce in the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains a Leader::*lacZ* or a Leader(Mut)::*lacZ* reporter fusion (Mut is the same as in B) at a heterologous (left) or at the endogenous site (right). The plasmid construct bearing Leader::*lacZ* (pMR4334) or Leader(Mut)::*lacZ* (pMR4385) was electroporated into the desired strain and colonies resulting from plasmid integration by recombination (dashed gray lines) at the heterologous 1.38 kb site were selected using the Km^R marker. Alternatively, a plasmid construct bearing *cas6*-Leader::*lacZ* (pMR4351) or *cas6*-Leader(Mut)::*lacZ* (pMR4402) was electroporated and colonies resulting from plasmid integration at the endogenous locus by recombination (dashed gray lines) were selected using the Km^R marker. (**D**) Expression of the Leader::*lacZ* or Leader(Mut)::*lacZ* reporter fusions (β-Gal activity, as the mean and standard error of at least three independent measurements) introduced into the *M. xanthus* genetic backgrounds indicated at the heterologous (strains MR2777-MR2780; MR2797-MR2799 and MR2851) or the endogenous site (strains MR2781-MR2784; MR2852-MR2855).

highly expressed in the Δ*ddvA* strain but not in the WT, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains (Figure 5A), suggesting that the DdvS-dependent transcription from *cas6* reads through downstream. Interestingly, this *lacZ* activity in the Δ*ddvA* strain was ∼3-fold higher than for the Leader::*lacZ* fusion at the endogenous site (see Figure 4D), suggesting that the leader reduces its expression. This may be because proteins that bind to the leader (Cas1/Cas2, H-NS, IHF have been shown to do so in other bacteria; (1,25)) impede transcription, and/or because the predicted leader RNA secondary structure (Supplementary Figure S6A; (71)) hinders translation.

The relative positioning of the *cas* genes, with predicted start and stop codons overlapping (*cmr1*-*cmr2*, *cmr3*-*cmr4*, *cmr4*-*cmr5*, *cmr5*-*cmr6* and *cmr6*-*cas6* pairs), or apart by only 2 bp (*cmr2*-*cmr3*) or 4 bp (MXAN_7283-*cmr1*), suggests that they form an operon, and hence would be expected to be transcribed as a polycistronic mRNA from P_{MXAN_7283} (immediately upstream of the *cas*/*cmr* cluster). To confirm this, we isolated total RNA from the Δ*ddvA* and WT strains and analyzed it by RT-PCR using appropriate primers to amplify regions across the junctions between different *cas* genes and, to check for transcriptional readthrough into the CRISPR4 array, from *cas6* to the first or second leader-proximal spacers (Figure 5B). Specific RT-PCR products were detected with RNA from the Δ*ddvA* strain but not in controls with the RT step omitted or with RNA isolated from the WT strain (Figure 5B). This indicates that the *cas* genes are transcribed as a polycistronic mRNA that reads through the leader into the CRISPR4 array, and that this occurs once DdvS is active.

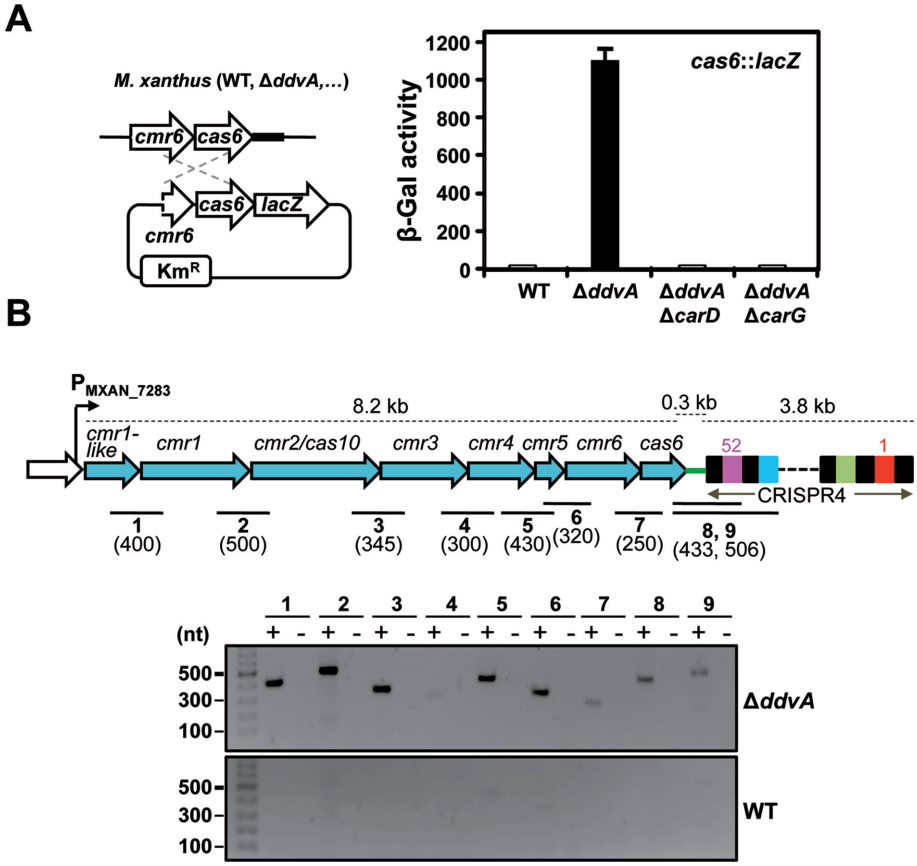

Figure 5. Transcriptional readthrough of the DdvS-dependent *cas* operon into the CRISPR4 array. (**A**) Scheme of the strategy used for introducing a CRISPR4 *cas6*-reporter *lacZ* fusion at the endogenous locus into the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains. A plasmid construct bearing the *cas6*::*lacZ* fusion (pMR4384, which also includes 440 bp of *cmr6* to facilitate integration at the endogenous site) was electroporated into the given strain and colonies resulting from plasmid integration at the endogenous locus by homologous recombination (dashed gray lines) were selected using the Km^R marker (strains MR2793-MR2796). Reporter *lacZ* expression (β-Gal activity) is the mean and standard error of at least three independent measurements. (**B**) RT-PCR analysis of RNA isolated from the Δ*ddvA* and WT strains. The scheme of CRISPR4-*cas* system on top indicates the span of this locus and of the regions (numbered lines with size, in bp, in parentheses) analyzed by RT-PCR. Spacers are numbered in descending order in the direction of transcription from P_{MXAN_7283}; thus, the first spacer from the 3′-end of the leader is numbered 52 and the last (farthest) spacer from the leader is numbered 1. RNA isolated from the indicated strain and reverse transcribed (+) or not (−; the negative control) served as the template to amplify the indicated regions. PCR products were analyzed in agarose gels alongside size markers.

Mature CRISPR4 crRNA biogenesis is controlled by DdvS, CarD and CarG

CRISPR arrays transcribed from a promoter in its leader region generate pre-crRNA, which is then processed to smaller, mature crRNAs. In the case of the CRISPR4 array, transcriptional readthrough from the upstream *cas* genes would be expected to result in a large transcript (of at least ∼12 kb) encompassing the *cas* operon transcript and the ∼4 kb pre-crRNA. To examine if this transcript can be detected, we isolated total RNA from the WT, Δ*ddvA*, Δ*ddvA* Δ*carD* and Δ*ddvA* Δ*carG* strains, and carried out a Northern blot analysis using an ∼1-kb probe spanning CRISPR4 spacers 37–52. We also performed the analysis with a Δ*ddvA* strain bearing a copy of *ddvA* under the control of an IPTG-inducible promoter (P_{IPTG}::*ddvA*) at the heterologous site mentioned before, which allows *ddvA* expression (and hence presence) to be turned on or off as required, and testing if DdvA can function when supplied in *trans* (Figure 6A). The Northern blot revealed a smear of progressively increasing signal intensity extending from >4 to <0.2 kb (where the signal was most intense) for the Δ*ddvA* strain and its derivative with P_{IPTG}::*ddvA* grown in the absence of the inducer IPTG (Figure 6A). By contrast, this was not observed for the WT, Δ*ddvA* Δ*carD*, Δ*ddvA* Δ*carG* or the Δ*ddvA* P_{IPTG}::*ddvA* strain grown in the presence of IPTG, although in each case a weak band appeared between 3 and 4 kb. To check if this weak signal corresponds to pre-crRNA, perhaps due to marginal expression from the putative σ^A-dependent promoter in the leader, we generated strains with the CRISPR4 array deleted in the WT and Δ*ddvA* genetic backgrounds. The presence of the band even when CRISPR4 is deleted indicates that it must be a non-specific background signal (Figure 6A). The smeared signal, with RNA molecules <0.2 kb contributing overwhelmingly, observed for strains devoid of DdvA (which can function in *trans*) but not of CarD or CarG, suggests that the large initial transcript is rapidly processed to mature crRNAs, whose production was examined next.

To detect mature crRNAs generated due to transcriptional readthrough from the upstream *cas* genes into the

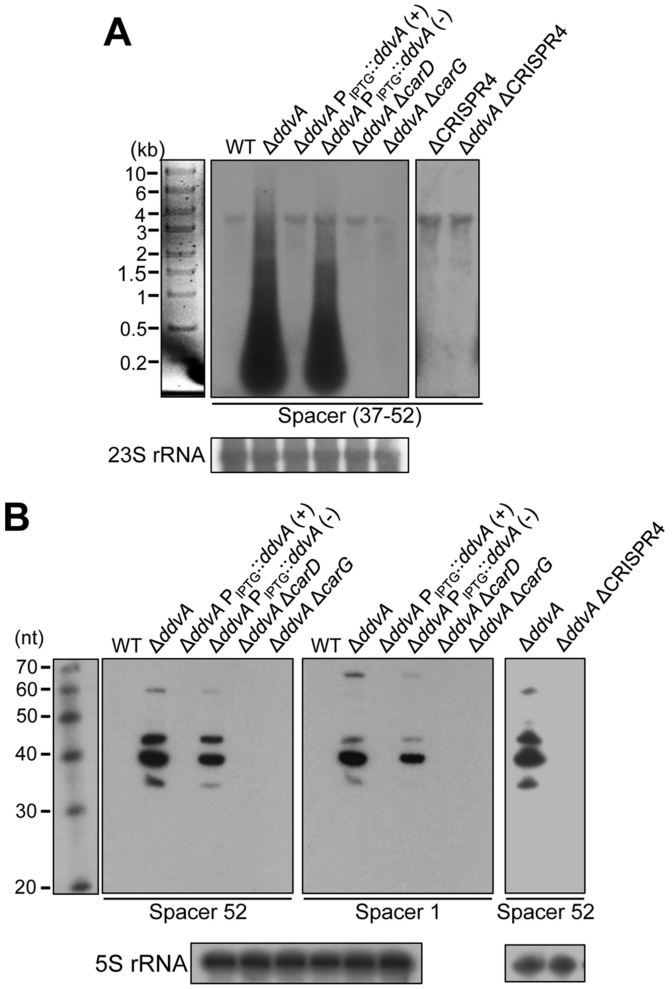

Figure 6. CRISPR4 crRNA analysis and its dependence on DdvS, CarD and CarG. (**A**) Northern blot analysis of total RNA isolated from the strains indicated on top using a probe corresponding to the segment spanning from spacers 37 to 52 of the CRISPR4 array. For the Δ*ddvA* P_{IPTG}::*ddvA* strain (MR2959), RNA was isolated after growth in the presence (+) or absence (−) of IPTG. 23S rRNA (2973 nt) was probed (bottom panel) as the internal loading control for each sample above. Size markers are shown on the left. (**B**) Representative Northern blot to detect mature CRISPR4 crRNAs by analyzing small RNAs isolated from the indicated strains using probes corresponding to CRISPR4 spacers 52 or 1. 5S rRNA (116 nt) probed as the internal loading control for each sample is shown below.

CRISPR4 array, small RNAs isolated from each strain above were analyzed in northern blots using short probes corresponding to the CRISPR4 spacer 52 (the first spacer from the leader 3′-end and the one most recently incorporated) or spacer 1 (the spacer farthest from the leader; Figure 5B and Supplementary Figure S2). With both probes, an intense band at ∼40 nt and weaker ones at ∼45 nt and at ∼35 nt, corresponding to mature CRISPR4 crRNAs, were detected in the Δ*ddvA* strain or its derivative with the P_{IPTG}::*ddvA* copy grown without IPTG (Figure 6B). That signals were detected using the spacer 1 probe indicates that transcription from the *cas* operon reads through the entire CRISPR4 array. However, no crRNAs were observed in the Δ*ddvA* strain with the P_{IPTG}::*ddvA* copy when grown in the presence of IPTG, nor in the WT, Δ*ddvA* Δ*carD* or Δ*ddvA* Δ*carG* strains (Figure 6B). Failure to detect any signal in the Δ*ddvA* strain with CRISPR4 deleted confirms that these crRNAs arise solely from this array (Figure 6B), while analysis using short probes complementary to those above indicated that the CRISPR4 array is not transcribed in the opposite direction (Supplementary Figure S6B). Thus, mature CRISPR4 crRNA biogenesis occurs only if DdvA is absent and CarD and CarG are available, which is also when the associated *cmr*/*cas* genes are expressed.

Biogenesis of mature CRISPR4 crRNAs depends on the Cas6 homolog of this type III-B system

The CRISPR4-*cas* and *ddvS-ddvA* loci have four genes (MXAN_7287 to MXAN_7284) in between that encode proteins of unknown function, and whose expression also depends on DdvS, CarD and CarG (Figure 2A and Table 1). The concurrent expression of MXAN_7287-7284 with that of CRISPR4-*cas* could imply a functional relationship between them that may be manifested in mature CRISPR4 crRNA formation. To test this we deleted the MXAN_7287-7284 segment in the Δ*ddvA* strain with a P_{IPTG}::*ddvA* copy at the heterologous site, leaving the *cas* genes and its promoter P_{MXAN_7283} intact (Figure 7A). Mature CRISPR4 crRNA was detected at comparable levels in the resulting strain and in that with intact MXAN_7287-7284 in the absence of DdvA (no IPTG), and was absent in both strains upon IPTG-induced *ddvA* expression (Figure 7A). This indicates that any functional link between MXAN_7287-7284 and the CRISPR4-Cas system is not at the level of mature crRNA production.

As mentioned earlier, the most downstream gene in the CRISPR4-*cas* cluster encodes a Cas6 homolog, the endonuclease expected to cleave pre-crRNA within each repeat and thus required to produce mature crRNAs. We tested this and whether it can do so in *trans*, since Cas6 of type III-B CRISPR systems reportedly acts in a standalone manner (13,14). For this, we generated a strain with an in-frame deletion of the CRISPR4 *cas6* gene in a Δ*ddvA* genetic background and its derivative bearing a vanillate-inducible copy of *cas6* (P_{van}::*cas6*) for conditional expression. Northern blots revealed that the mature CRISPR4 crRNAs observed in the Δ*ddvA* background were not detected when *cas6* was also deleted (Δ*ddvA* Δ*cas6*), and reappeared on inducing P_{van}::*cas6* expression (Figure 7B). Northern blots using the longer 1-kb probe yielded consistent data (Figure 7C). Absence of *cas6*, as in the Δ*ddvA* Δ*cas6* strain or the Δ*ddvA* Δ*cas6* P_{van}::*cas6* strain grown without vanillate, yielded a detectable smear unlike the WT strain, but without the highly intense signal at <0.2 kb that characterized the Δ*ddvA* strain or the Δ*ddvA* Δ*cas6* P_{van}::*cas6* strain grown in the presence of vanillate. This suggests that CRISPR4 transcripts produced in the Δ*ddvA* background are rapidly processed if Cas6 of this system is available, even when Cas6 is supplied in *trans*, whereas these transcripts are produced but not fully processed in the Δ*ddvA* Δ*cas6* genetic background. Interestingly, in the latter, the decrease in signal at <0.2 kb was not accompanied by a concomitant increase in RNA species >0.2 kb, and the relatively lower overall signal may be due to a somehow reduced expres-

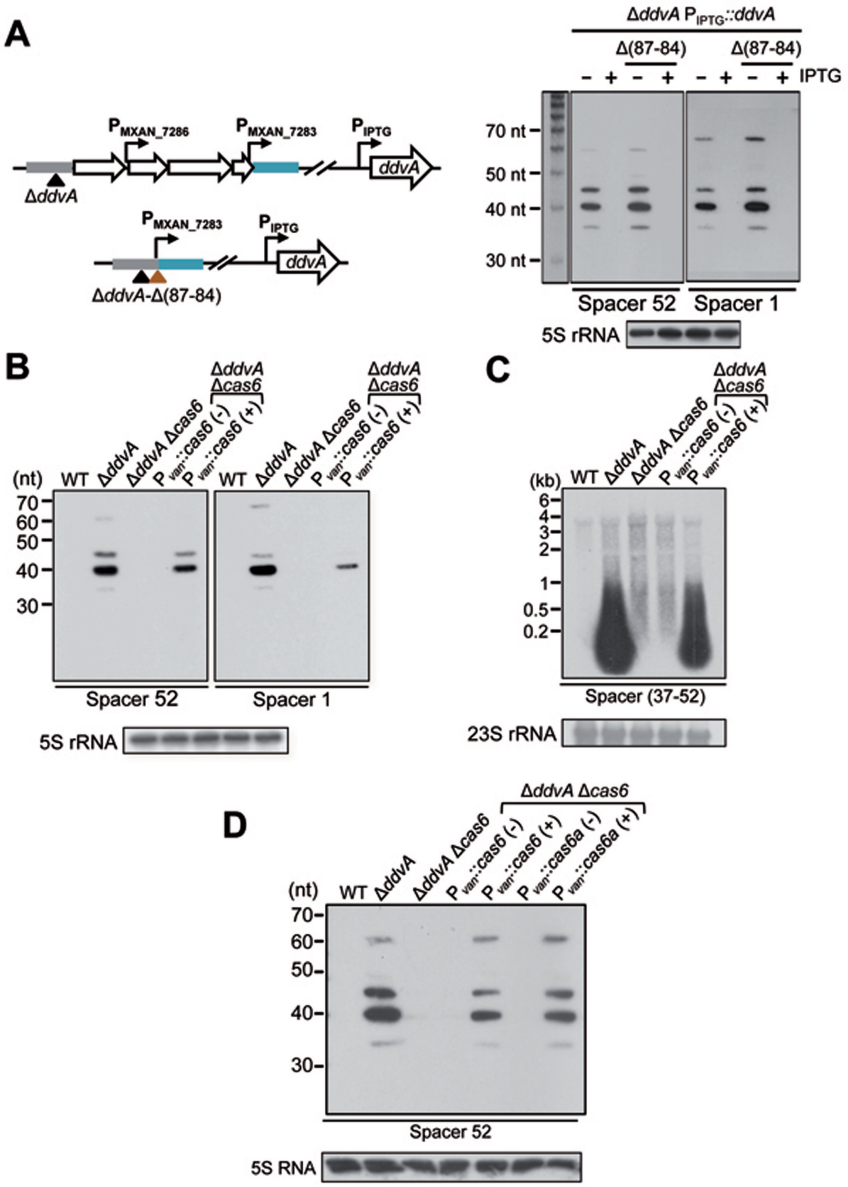

Figure 7. Cas6 dependence of mature CRISPR4 crRNA generation. (**A**) Scheme showing the Δ*ddvA* P_{IPTG}::*ddvA* strain (MR2959) and its derivative (MR3177) with Δ(87–84), where the segment encompassing MXAN_7287 to MXAN_7284 is deleted, leaving the *cas* genes and P_{MXAN_7283} intact. Northern blot using probes corresponding to CRISPR4 spacers 1 and 52 comparing mature crRNAs produced in the Δ*ddvA* P_{IPTG}::*ddvA* strain and its Δ(87–84) derivative grown with and without IPTG. Below is the blot probing 5S rRNA as the internal loading control for each sample. (**B**) Northern blot analysis of mature crRNAs using probes corresponding to CRISPR4 spacers 1 and 52 in the WT (DK1050), Δ*ddvA* (MR1316), Δ*ddvA* Δ*cas6* (MR2565) and Δ*ddvA* Δ*cas6* P_{van}::*cas6* (MR2862) strains; 5S rRNA was probed as the internal loading control for each sample. RNA was extracted for strain MR2862 grown in the absence (−) or presence (+) of vanillate. (**C**) Northern blot analysis of CRISPR4 pre-crRNA, for strains as in (B), using a DNA probe spanning the CRISPR4 segment from spacer 37–52. 23S rRNA was probed as the internal loading control for each sample. (**D**) Northern blot analysis of mature crRNAs using a probe corresponding to CRISPR4 spacer 52 in strains DK1622, MR2959, MR2962, MR2967 and MR3062, with relevant genotype indicated. The Δ*ddvA* strains used in this analysis have a P_{IPTG}::*ddvA* copy at the 1.38 kb heterologous site but are denoted Δ*ddvA*, for simplicity, as they were grown in the absence of IPTG. The 5S rRNA was probed as the internal loading control for each sample. Strains with a vanillate-inducible copy of CRISPR4 *cas6* (MR2967) or CRISPR2/3 *cas6a* (MR3062) were grown in the absence (−) or presence (+) of vanillate.

sion of pre-crRNA, or to the latter being more susceptible to non-specific degradation in the absence of Cas6. Altogether, these data indicate that Cas6 of the CRISPR4-Cas system is required to process pre-crRNA for mature crRNA biogenesis.

Myxococcus xanthus contains a second Cas6 in the distinct type I CRISPR2/3-Cas system mentioned earlier that houses the *devTRS* genes, which WT cells reportedly express under starvation, the stress that induces fruiting body development (9,11,34). Interestingly, this Cas6 (hereafter denoted Cas6a to distinguish it from the other) shares high (66%) sequence identity to that of the type III-B system, and has an additional 27-residue N-terminal segment (Supplementary Figure S7A). Moreover, the 36/37-bp CRISPR

array repeats of the two systems differ at only two positions, and the CRISPR3 and CRISPR4 leaders are 74% identical (Supplementary Figures S1B and 7B). We therefore tested if Cas6a can also process CRISPR4 pre-crRNA. Northern blots revealed mature CRISPR4 crRNA in a Δ*ddvA* Δ*cas6* strain supplied with a vanillate-inducible copy of *cas6a* when grown in the presence of inducer (Figure 7D). These results thus indicate that Cas6a can also target CRISPR4 pre-crRNA, consistent with the high sequence similarity between the two *M. xanthus* Cas6 proteins and with Cas6a functioning in a standalone manner, as has been reported for type I-A and type III-B systems (13,14). Our data, together with the other similarities noted above and a previous observation that a 79-bp stretch in the antisense strand between the two CRISPR2/3 arrays is similar (with 14 mismatches) to a segment in MXAN_7283 of the CRISPR4-Cas system (9), thus raise the possibility of crosstalk between the two *M. xanthus* CRISPR-Cas systems. Such crosstalk has been observed for some, though not all, coexisting type III-B and type I systems, despite their distinct modes of target interference: type I systems recognize and degrade DNA targets, whereas type III systems recognize nascent RNA transcripts that include a reverse complement of the crRNA sequence to degrade both the transcript and its template DNA, a process described as transcription-dependent DNA interference (72). Crosstalk between the above two CRISPR-Cas systems in *M. xanthus* would require both to be simultaneously active. The CRISPR2/3-Cas system is reportedly activated during fruiting body development (9). Moreover, a gain-of-function transposon insertion at the CRISPR4 array was reported to severely affect the cell aggregation that precedes fruiting body formation, and it was speculated that this may be due to some CRISPR4 transcripts targeting genes critical for development (73). We therefore examined CRISPR4-*cas* expression and mature crRNA biogenesis during fruiting body development.

CRISPR4-*cas* expression is not activated by starvation

WT cells undergo a complex multicellular developmental program on starvation to form mature, dark spore-filled fruiting bodies, a process that is impaired in the Δ*carD* and Δ*carG* strains (45,47). On starvation plates, the Δ*ddvA* strain produced normal fruiting bodies at similar times as the WT but typically about half in number, in contrast to the defective fruiting bodies of the Δ*carD* strain used as negative control (Figure 8A). Thus, fruiting bodies develop even upon constitutive expression of DdvS, whose competition with other σ factor(s) that usually act during development may indirectly reduce the observed number of fruiting bodies. To test whether starvation affects CRISPR4-*cas* expression, we first analyzed expression during development of a *lacZ* reporter gene fused to the CRISPR4 leader at the endogenous site (described earlier). This was done qualitatively by monitoring the appearance of a blue color in an X-Gal overlay assay on starvation plates and quantitatively by measuring β-galactosidase activity. The Δ*ddvA* strain, in which the reporter *lacZ* is expressed during vegetative growth, exhibited a pronounced blue color at time zero, as expected, which persisted throughout development on starvation plates (Figure 8B); accordingly, the corresponding β-galactosidase activity was high and steady over the 0–24 h period examined (Figure 8C). For comparison, the late developmentally regulated marker (Ω4435; Supplementary Table S2) increased sharply after 12 h post-starvation from the very low levels at earlier times (Figure 8C). By contrast, the blue color as well as the β-galactosidase activity remained marginal throughout development for the WT, just as for the Δ*ddvA* Δ*carD* and Δ*ddvA* Δ*carG* strains used as negative controls for CRISPR4-*cas* expression (Figure 8B and C). To independently support these observations, we performed Northern blot analysis of mature CRISPR4 crRNAs of the WT and Δ*ddvA* strains. Mature CRISPR4 crRNAs were observed at even 48 h into development for the Δ*ddvA* strain, albeit at lower levels than at time zero, but not for the WT at any of the times examined (Figure 8D). Despite these crRNAs being present, the Δ*ddvA* strain forms normal fruiting bodies, suggesting that the developmental defect caused by the transposon insertion at CRISPR4 mentioned above is unlikely to stem from expression of the corresponding crRNAs, as was speculated. Altogether, our data suggest that starvation, which induces multicellular development, is not the trigger that activates CRISPR4-*cas* expression and mature CRISPR4 crRNA biogenesis in *M. xanthus*.

As noted earlier, the CRISPR2/3-Cas system, which includes the *cas* orthologs *devTRS*, has been reported to be developmentally induced at low levels and enhanced by disrupting the negative autoregulation by DevTRS (9,11). The corresponding mature crRNAs would thus be expected to appear in a developmental/DevTRS-dependent manner, but this has not been previously examined. We therefore analyzed this in Northern blots using a probe corresponding to the leader-proximal spacer 22 of the CRISPR2/3-Cas system (Figure 8D). Mature CRISPR2/3 crRNAs were not detected in the WT, likely due to its reported low induction during development, since in the Δ*devT* strain we could observe discrete signals (although of low intensity) at ∼90 nt, ∼150 nt and somewhat higher. These sizes are notably larger than the 35–45 nt mature CRISPR4 crRNAs, despite the comparable sizes of the repeats as well as of the spacers, and the high sequence similarity between the repeats of the two CRISPR systems and between their respective Cas6 homologs, both of which can act on CRISPR4 pre-crRNA in *trans* to generate mature crRNAs of similar size (Figure 7D). The larger CRISPR2/3 crRNAs observed in the Δ*devT* strain may stem from differences in the final trimming of these type I system RNAs relative to those of the type III-B CRISPR4-Cas system (13,14), from incomplete processing, or from lack of stabilization of the smaller mature crRNAs by an interference complex (Cascade) devoid of DevT, a Cas8a1 homolog and likely component (11). Altogether, these data indicate that the CRISPR4-Cas and the CRISPR2/3-Cas systems are not simultaneously expressed pre- or post-starvation in the WT (where neither system is expressed) or under conditions favoring expression of CRISPR4-*cas* (Δ*ddvA*) or CRISPR2/3-*cas* (Δ*devT*). Consequently, there is little possibility of crosstalk between the two CRISPR-Cas systems, at least under the conditions tested. However, given the similarities between the two systems noted above, crosstalk could nonetheless occur when

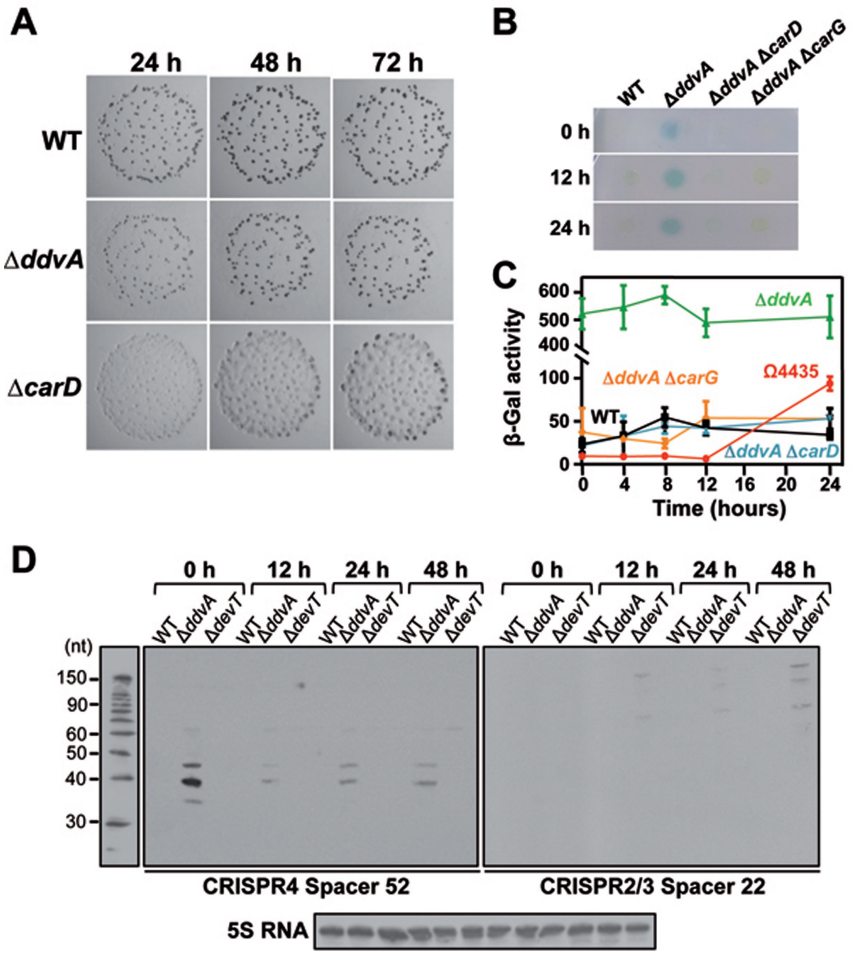

Figure 8. CRISPR-*cas* expression and starvation-induced fruiting body development. (**A**) Fruiting bodies formed by the WT (DK1622), Δ*ddvA* (MR2590) and Δ*carD* (MR1900) strains after incubation at 33°C for 24, 48 and 72 h on CF starvation plates. (**B**) An X-Gal overlay assay of WT (MR2781), Δ*ddvA* (MR2782), Δ*ddvA* Δ*carD* (MR2783) and Δ*ddvA* Δ*carG* (MR2784) strains, with a Leader::*lacZ* reporter at the endogenous CRISPR4-*cas* locus. Cells were spotted on CF starvation plates and incubated at 33°C for 12, 24 and 48 h, followed by an X-Gal overlay and examination after 2 h. (**C**) Activity of a Leader::*lacZ* reporter at the endogenous CRISPR4-*cas* locus, in strains as in (**B**), from 0 to 24 h upon starvation. The developmentally-induced Ω4435::Tn5-*lac* (Ω4435) insertion was used as a control. (**D**) Northern blot analysis to detect mature crRNAs isolated from the WT (DK1622), the Δ*ddvA* (MR2590) and the Δ*devT* (MR3078) strains incubated on CF starvation plates for the indicated times. The 5S rRNA, probed as the internal loading control for each sample, is shown below and size markers are indicated on the left.

both are active, although natural conditions for this, if any, remain to be determined.

DISCUSSION

CRISPR arrays and associated *cas* genes must be first expressed for function. Controlling and coordinating this essential step ensures that the machinery is active only when needed, and limits the metabolic costs and unwanted accumulation of nucleases. The present study has uncovered an unprecedented mechanism for regulated expression of a CRISPR-*cas* system, which depends on an alternative ECF σ factor (DdvS), its cognate anti-σ factor (DdvA) and a global regulatory complex (CarD–CarG). This tight multifactorial control of a type III-B CRISPR-*cas* system in *M. xanthus* enables coordinated supply of all the CRISPR-Cas components for mature crRNA biogenesis and its assembly as an interference complex *in vivo* (Figure 9).

ECF σ factors, small divergent members of the σ^{70} superfamily in bacteria, associate with RNAP to recognize specific promoters for the expression of defined genes in adaptive responses to various cell envelope and environmental stress signals (39–41). We had previously established that DdvS is kept inactive via its association with the cytoplasmic N-terminal ZAS domain of DdvA, a single-pass membrane-associated protein, such that disruption of *ddvA* or overexpression of *ddvS* provides free active DdvS for its own CarD/CarG-dependent expression *in vivo* (42). The specific natural signal that triggers DdvA inactivation to liberate DdvS, and the set of genes (beyond *ddvS* and *ddvA*) whose expression requires DdvS (its regulon) were, however, unknown. The current study has revealed that the CarD/CarG-dependent DdvS regulon comprises primarily a type III-B CRISPR-Cas system located in the proximity of the *ddvS-ddvA* locus. The few other genes included in the regulon encode products with unknown functions, unrelated to Cas proteins and whose role, if any, in CRISPR4-

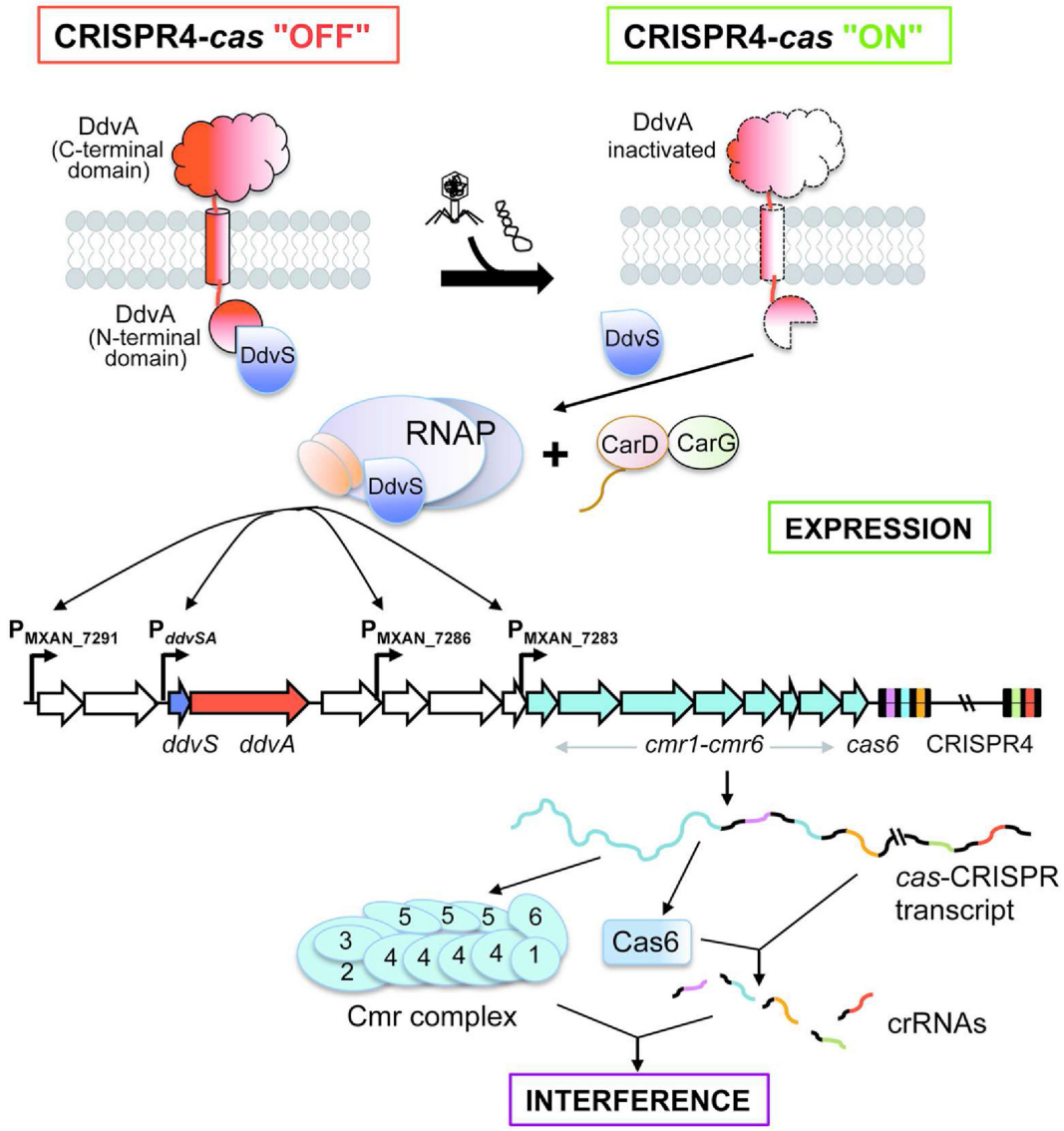

Figure 9. Model for regulated CRISPR4-*cas* expression by the DdvS–DdvA ECF-σ/anti-σ pair and the CarD–CarG regulatory complex. In the CRISPR4-*cas* 'OFF' state, the ∼70-residue cytoplasmic, N-terminal ZAS domain of DdvA physically sequesters its cognate ECF-σ DdvS and impedes DdvS-driven expression of the CRISPR4-Cas system. In the CRISPR4-*cas* 'ON' state, DdvA is inactivated by an external signal (that may be phage or foreign DNA) freeing DdvS to associate with RNAP and activate, together with the CarD–CarG complex (the wavy line in CarD represents its DNA-binding domain), expression from P_{MXAN_7283} (as well as from the other three DdvS/CarD/CarG-dependent promoters at the locus). This leads to simultaneous expression of the *cas* genes and, by transcriptional readthrough, of the CRISPR4 array to produce a polycistronic mRNA encoding the Cas proteins comprising the multiprotein Cmr complex, the endoribonuclease Cas6 and the pre-crRNA. Processing of pre-crRNA, dependent on Cas6, produces mature crRNAs, which associate with the Cmr complex (represented based on available descriptions of other Cmr complexes) to form the ribonucleoprotein complex that mediates interference.

Cas function is unclear. At least four of these, located in the segment between *ddvS-ddvA* and the downstream CRISPR4-*cas* operon, could be deleted with no apparent loss of mature CRISPR4 crRNA production. We mapped three new DdvS-driven CarD/CarG-dependent promoters in this region, besides that previously identified for *ddvS-ddvA*. One of these occurs immediately upstream of the first gene of the *cas* cluster, and the other two are further upstream and drive the expression of genes encoding proteins of unknown functions. These promoters are all silent unless DdvA is inactivated and, consistent with a direct action, DdvS and the CarD–CarG complex localize *in vivo* at these promoters when the latter are turned on. DdvS specificity for these promoters may stem from their conserved $GTAAn_{16}CGT$ motif and spacer length and the dependence on CarD and CarG.

The exact molecular details on how CarD and CarG exert their positive role on DdvS activity (or on the activity of a number of other ECF σ factors in *M. xanthus*; (42)) remain enigmatic. *Myxococcus xanthus* has a large number of ECF σ factors (∼45) and the activity of several of these is modulated by the CarD–CarG complex, which is thus an additional regulatory layer to the negative control exerted by the cognate anti-σ factor (42). Since promoter

recognition by ECF σ factors usually adheres to consensus −35 and −10 elements, variations from the consensus and in the length (14–17 bp) and sequence of the spacer between them determine ECF σ promoter selectivity (39–41,66,67). Regulation by factors like the CarD–CarG complex, which could also contribute to promoter selectivity (42), may enable DdvS to target its promoters. The CarD–CarG complex binds to the minor groove of AT-rich DNA sites using the intrinsically disordered CarD C-terminal domain, and to the RNAP β-subunit through interactions with the CarD N-terminal domain (44–48). Consequently, the CarD–CarG complex may facilitate recruitment and/or positioning of the RNAP-DdvS holoenzyme, once available, at target promoters through interactions with the promoter proximal AT-rich DNA sites and with RNAP. Interestingly, the CarD N-terminal domain is the founding member of a large family of bacterial RNAP-binding proteins denoted CdnL, which play a crucial role in stabilizing transcriptionally competent open promoter–RNAP complexes formed at σ^A-dependent promoters (56,74–76). That CarD shares the RNAP recognition mode and some other crucial functional determinants with CdnL (48) suggests that its complex with CarG could have a similar stabilizing role in transcription initiation at DdvS-dependent promoters.

The CRISPR4-Cas system, we show, is an operon with the *cas* genes transcribed as a polycistronic mRNA that reads through the leader segment into the CRISPR4 array in a DdvS/CarD/CarG-dependent manner. Curiously, the leader has a putative σ^A-dependent promoter but whose activity appears to be negligible *in vivo*. As a consequence, expression of the CRISPR4 array, whose leader lacks a DdvS-dependent promoter, nonetheless requires DdvS, CarD and CarG. This contrasts with most other CRISPR-*cas* systems that have been studied, where transcription of the CRISPR array, as with the *cas* genes, relies on promoters recognized by RNAP holoenzyme containing the primary σ^A factor. Dependence on an ECF σ factor (together with CarD and CarG) to express *cas* genes as well as the CRISPR array thus ensures simultaneous production of all the essential components of the CRISPR4-Cas system upon receiving and transducing the signal that leads to its activation.

How CRISPR-*cas* activation is triggered is clearly a critical step in its action, but the underlying mechanisms are among those least characterized. In line with its role as the anti-σ of an ECF σ factor, DdvA would be at the forefront in sensing and/or transmitting a putative extracytoplasmic signal that would culminate in its inactivation to liberate DdvS. Being membrane-associated, DdvA inactivation is very likely accompanied by a perturbation of its membrane environment. CRISPR-*cas* expression has been previously linked to membrane stress, but the nature of the actual cues involved remain speculative and the underlying mechanisms complex and poorly understood (37,38). Our data rule out one potential source of membrane stress as the trigger for CRISPR4-*cas* expression and mature crRNA generation, that stemming from the cellular differentiation and morphological changes that occur on starvation-induced fruiting body development of *M. xanthus*. Conceivably, the large (∼900-residue) extracytoplasmic DdvA C-terminal domain plays a crucial role in sensing the external signal and in the subsequent inactivation of DdvA. The sequence of this domain predicts an N-terminal region with three tandem tetratricopeptide (TPR) repeat modules that typically mediate protein–protein interactions (77), and a C-terminal CHAT (Caspase HetF Associated with Tprs) family cysteine protease module (Supplementary Figure S3). The CHAT domain is widespread in bacteria (78) but remains poorly characterized, although a protease activity has been demonstrated for two such proteins, including HetF, in cyanobacteria (79,80). An external signal inducing protease action, with associated protein–protein interactions, is a frequently employed mechanism for anti-σ inactivation (40). For the DdvS–DdvA pair, this could be orchestrated by the putative CHAT and TPR modules in the DdvA C-terminal domain. A precedent for autoproteolysis of an anti-σ factor (unrelated to DdvA) has been reported (81). Identifying the molecular mechanism involved will require pinning down the signal triggering CRISPR4-*cas* expression. One attractive possibility to pursue in future studies is that the DdvA C-terminal domain is directly, or indirectly, involved in phage recognition.

Deleting the *cas6* homolog encoded by the CRISPR4-Cas system abrogated mature crRNA formation, which was restored on supplying this Cas6 in *trans*, indicating its requirement for processing of the pre-crRNA. Intriguingly, the Cas6 protein, the array repeats and the leader segment of the CRISPR4-Cas system closely resemble those of the type I-A CRISPR2/3-Cas system in *M. xanthus* (but share no similarity with the third, uncharacterized CRISPR1-Cas type I-C system). Also, the first *cas* gene of the CRISPR4-Cas system has a 79-nt segment that coincides (with 14 mismatches) with a stretch in the region between the CRISPR2 and CRISPR3 arrays (9,34), and in several natural *M. xanthus* isolates the type III-B CRISPR4-Cas system co-occurs with the CRISPR2/3-Cas system (82). These similarities suggest that the two systems likely share a common evolutionary origin, and also raise the possibility of crosstalk between them. Consistent with this, we found that Cas6a of CRISPR2/3 supplied in *trans* could complement the lack of CRISPR4-associated Cas6, suggesting that the two Cas6 homologs are functionally interchangeable. Any crosstalk between the two CRISPR-Cas systems would require that both are simultaneously active. However, CRISPR4-*cas* expression is controlled tightly by DdvS, DdvA, CarD and CarG, and is not induced by starvation, which triggers multicellular development, whereas CRISPR2/3-*cas* expression is regulated differently. It is silent during vegetative growth but induced upon starvation, although the levels are low due to negative autoregulation of its promoter P_{dev} (which is very different from P_{ddvSA}) by the products of three of its own *cas* genes (*devTRS*) (9,11). Consistent with this, we found that mature CRISPR2/3 crRNAs were undetected in the WT strain but appeared in a Δ*devT* strain during development, albeit at low levels and, possibly, incompletely processed. The *devTRS* products have been shown to regulate the timing of sporulation during development by preventing overproduction of the sporulation inhibitor DevI (encoded by the first gene of the *dev* operon) through negative autoregulation of *dev* transcription, presumably by a DevTRS Cascade-like subcomplex (11). This was observed in the standard laboratory strain. However, many natural *M. xanthus* isolates undergo normal devel-

opment and sporulation despite lacking a functional P_{dev} and *devI*, and DevTRS orthologs appear to be absent in most other myxobacteria with proficient fruiting body development (34,82), leading to the hypothesis that the relationship between a functional *dev* CRISPR2/3 system and development may have evolved recently (11). The link, if any, between CRISPR2/3 crRNA production and development, however, remains an open issue. Nevertheless, at least during development under the conditions tested with the standard laboratory strain, our data do not indicate any crosstalk between CRISPR4-*cas* and CRISPR2/3-*cas* expression,

An interesting alternative is that the two CRISPR-Cas systems are linked via the more common role in phage immunity. It was recently reported that a type III-B CRISPR-Cas interference complex could use not only its own mature crRNAs but also naturally co-opt those from a type I locus, despite differences in their repeat sequences, to counter phage infection (72). The study proposed that plasticity in crRNA utilization by type III-B systems may be a general feature to supplement a co-existing type I system as an additional line of defense (72). It was also noted that type III systems often co-exist with type I systems in prokaryotic genomes, and their greater tolerance to mutations in target protospacer and flanking sequences can limit phage escape effectively thus providing a back-up system against phage invasion (72). It is thus noteworthy that the type III-B CRISPR4-Cas system co-occurs with the type 1-A CRISPR-Cas system in several natural *M. xanthus* isolates (82). In the *M. xanthus* CRISPR2/3-Cas system, the first unique spacer downstream of the leader following the *cas2* gene matches perfectly with a sequence in the integrase gene (*intP*) of bacteriophage Mx8 (Supplementary Figure S7C; (9)), strongly hinting at a role of this CRISPR-Cas system in defense against Mx8 or related viruses. Neither this nor any of the other spacers in the CRISPR2/3 array occurs in CRISPR4, whose spacers show no significant matches to any known bacteriophage or plasmid DNA. It must be noted that the genome sequence is available for Mx8 but not for any of the other *M. xanthus* phages that have been isolated. It would therefore be interesting to examine if these or other to-be-identified phages elicit a direct response from the CRISPR4-Cas system, or indirectly through crosstalk with the other two type I systems in *M. xanthus.* We are therefore currently exploring these aspects and the possible interplay with the DdvS–DdvA pair in these intriguing immune systems.

In conclusion, our study has uncovered a novel regulatory mechanism for CRISPR-*cas* expression and crRNA biogenesis that is based on the use of a specific ECF-σ and its cognate anti-σ, together with a global regulatory complex. This provides an elegant mechanism to detect an external signal such as phage or plasmid invasion, directly or indirectly, by the anti-σ factor and couple it to anti-σ inactivation and release of the ECF-σ factor for a coordinated expression of a CRISPR-*cas* cluster and mature crRNA generation.

DATA AVAILABILITY

Microarray data are accessible through GEO accession number GSE112385, deposited at the NCBI′s GEO (https://www.ncbi.nlm.nih.gov/geo/).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Cesar Flores-Flores (University of Murcia) for DNA sequencing, Dr Luis Alcaraz (Bioarray SL, Spain) for the microarray analysis, Prof. Lee Kroos (Michigan State University) for plasmid pRR023 and J.A. Madrid and Victoria López Egea (both from the University of Murcia) for assistance.

FUNDING

Ministerio de Economía y Competitividad-Spain/European Regional Development Fund (MINECO/FEDER) [BFU2012-40184-C02-01, BFU2015-67968-C2-1P to M.E.-A, BFU2012-40184-C02-02, BFU2015-67968-C2-2-P to S.P]; FPI Ph.D. Fellowships/contracts (to J.A.-R., D.B.-B); Fundación Séneca (Murcia)-Spain [19429/PI/14 to M.E.-A.]. Funding for open access charge: Ministerio de Economía y Competitividad-Spain [BFU2015-67968-C2-1P].
Conflict of interest statement. None declared.

REFERENCES

1. Hille,F., Richter,H., Wong,S.P., Bratovic,M., Ressel,S. and Charpentier,E. (2018) The biology of CRISPR-Cas: backward and forward. *Cell*, **172**, 1239–1259.
2. Marraffini,L.A. (2015) CRISPR-Cas immunity in prokaryotes. *Nature*, **526**, 55–61.
3. van der Oost,J., Westra,E.R., Jackson,R.N. and Wiedenheft,B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **12**, 479–492.
4. Westra,E.R., Buckling,A. and Fineran,P.C. (2014) CRISPR-Cas systems: beyond adaptive immunity. *Nat. Rev. Microbiol.*, **12**, 317–326.
5. Wright,A.V., Nunez,J.K. and Doudna,J.A. (2016) Biology and applications of CRISPR aystems: harnessing Nature's toolbox for genome engineering. *Cell*, **164**, 29–44.
6. Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S., Romero,D.A. and Horvath,P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
7. Babu,M., Beloglazova,N., Flick,R., Graham,C., Skarina,T., Nocek,B., Gagarinova,A., Pogoutse,O., Brown,G., Binkowski,A. *et al.* (2011) A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.*, **79**, 484–502.
8. Sampson,T.R., Saroj,S.D., Llewellyn,A.C., Tzeng,Y.L. and Weiss,D.S. (2013) A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature*, **497**, 254–257.
9. Viswanathan,P., Murphy,K., Julien,B., Garza,A.G. and Kroos,L. (2007) Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.*, **189**, 3738–3750.
10. Louwen,R., Staals,R.H., Endtz,H.P., van Baarlen,P. and van der Oost,J. (2014) The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev.*, **78**, 74–88.
11. Rajagopalan,R. and Kroos,L. (2017) The *dev* operon regulates the timing of sporulation during *Myxococcus xanthus* development. *J. Bacteriol.*, **199**, e00788–00716.

12. Ledford,H. (2016) CRISPR: gene editing is just the beginning. *Nature*, **531**, 156–159.
13. Charpentier,E., Richter,H., van der Oost,J. and White,M.F. (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.*, **39**, 428–441.
14. Hochstrasser,M.L. and Doudna,J.A. (2015) Cutting it close: CRISPR-associated endoribonuclease structure and function. *Trends Biochem. Sci.*, **40**, 58–66.
15. Brouns,S.J., Jore,M.M., Lundgren,M., Westra,E.R., Slijkhuis,R.J., Snijders,A.P., Dickman,M.J., Makarova,K.S., Koonin,E.V. and van der Oost,J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**, 960–964.
16. Carte,J., Wang,R., Li,H., Terns,R.M. and Terns,M.P. (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.*, **22**, 3489–3496.
17. Koonin,E.V., Makarova,K.S. and Zhang,F. (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.*, **37**, 67–78.
18. Patterson,A.G., Yevstigneyeva,M.S. and Fineran,P.C. (2017) Regulation of CRISPR-Cas adaptive immune systems. *Curr. Opin. Microbiol.*, **37**, 1–7.
19. Soutourina,O.A., Monot,M., Boudry,P., Saujet,L., Pichon,C., Sismeiro,O., Semenova,E., Severinov,K., Le Bouguenec,C., Coppee,J.Y. *et al.* (2013) Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet.*, **9**, e1003493.
20. Gunderson,F.F. and Cianciotto,N.P. (2013) The CRISPR-associated gene *cas2* of *Legionella pneumophila* is required for intracellular infection of amoebae. *Mbio*, **4**, e00074-00013.
21. Juranek,S., Eban,T., Altuvia,Y., Brown,M., Morozov,P., Tuschl,T. and Margalit,H. (2012) A genome-wide view of the expression and processing patterns of *Thermus thermophilus* HB8 CRISPR RNAs. *RNA*, **18**, 783–794.
22. Przybilski,R., Richter,C., Gristwood,T., Clulow,J.S., Vercoe,R.B. and Fineran,P.C. (2011) Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol.*, **8**, 517–528.
23. Medina-Aparicio,L., Rebollar-Flores,J.E., Gallego-Hernandez,A.L., Vazquez,A., Olvera,L., Gutierrez-Rios,R.M., Calva,E. and Hernandez-Lucas,I. (2011) The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.*, **193**, 2396–2407.
24. Pougach,K., Semenova,E., Bogdanova,E., Datsenko,K.A., Djordjevic,M., Wanner,B.L. and Severinov,K. (2010) Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.*, **77**, 1367–1379.
25. Pul,U., Wurm,R., Arslan,Z., Geissen,R., Hofmann,N. and Wagner,R. (2010) Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Mol. Microbiol.*, **75**, 1495–1512.
26. Westra,E.R., Pul,U., Heidrich,N., Jore,M.M., Lundgren,M., Stratmann,T., Wurm,R., Raine,A., Mescher,M., Van Heereveld,L. *et al.* (2010) H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.*, **77**, 1380–1393.
27. Agari,Y., Sakamoto,K., Tamakoshi,M., Oshima,T., Kuramitsu,S. and Shinkai,A. (2010) Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.*, **395**, 270–281.
28. Patterson,A.G., Chang,J.T., Taylor,C. and Fineran,P.C. (2015) Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res.*, **43**, 6038–6048.
29. Yang,C.D., Chen,Y.H., Huang,H.Y., Huang,H.D. and Tseng,C.P. (2014) CRP represses the CRISPR/Cas system in *Escherichia coli*: evidence that endogenous CRISPR spacers impede phage P1 replication. *Mol. Microbiol.*, **92**, 1072–1091.
30. Arslan,Z., Stratmann,T., Wurm,R., Wagner,R., Schnetz,K. and Pul,U. (2013) RcsB-BglJ-mediated activation of Cascade operon does not induce the maturation of CRISPR RNAs in *E. coli* K12. *RNA Biol.*, **10**, 708–715.
31. He,F., Vestergaard,G., Peng,W., She,Q. and Peng,X. (2017) CRISPR-Cas type I-A Cascade complex couples viral infection surveillance to host transcriptional regulation in the dependence of Csa3b. *Nucleic Acids Res.*, **45**, 1902–1913.
32. Lintner,N.G., Frankel,K.A., Tsutakawa,S.E., Alsbury,D.L., Copie,V., Young,M.J., Tainer,J.A. and Lawrence,C.M. (2011) The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. *J. Mol. Biol.*, **405**, 939–955.
33. Liu,T., Li,Y., Wang,X., Ye,Q., Li,H., Liang,Y., She,Q. and Peng,N. (2015) Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo spacer acquisition. *Nucleic Acids Res.*, **43**, 1044–1055.
34. Rajagopalan,R., Wielgoss,S., Lippert,G., Velicer,G.J. and Kroos,L. (2015) *devI* is an evolutionarily young negative regulator of *Myxococcus xanthus* development. *J. Bacteriol.*, **197**, 1249–1262.
35. Hoyland-Kroghsbo,N.M., Paczkowski,J., Mukherjee,S., Broniewski,J., Westra,E., Bondy-Denomy,J. and Bassler,B.L. (2017) Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 131–135.
36. Patterson,A.G., Jackson,S.A., Taylor,C., Evans,G.B., Salmond,G.P.C., Przybilski,R., Staals,R.H.J. and Fineran,P.C. (2016) Quorum sensing controls adaptive immunity through the regulation of multiple CRISPR-Cas systems. *Mol. Cell*, **64**, 1102–1108.
37. Perez-Rodriguez,R., Haitjema,C., Huang,Q., Nam,K.H., Bernardis,S., Ke,A. and DeLisa,M.P. (2011) Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.*, **79**, 584–599.
38. Serbanescu,M.A., Cordova,M., Krastel,K., Flick,R., Beloglazova,N., Latos,A., Yakunin,A.F., Senadheera,D.B. and Cvitkovitch,D.G. (2015) Role of the *Streptococcus mutans* CRISPR-Cas systems in immunity and cell physiology. *J. Bacteriol.*, **197**, 749–761.
39. Feklistov,A., Sharon,B.D., Darst,S.A. and Gross,C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.*, **68**, 357–376.
40. Sineva,E., Savkina,M. and Ades,S.E. (2017) Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Curr. Opin. Microbiol.*, **36**, 128–137.
41. Staron,A., Sofia,H.J., Dietrich,S., Ulrich,L.E., Liesegang,H. and Mascher,T. (2009) The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol. Microbiol.*, **74**, 557–581.
42. Abellón-Ruiz,J., Bernal-Bernal,D., Abellán,M., Fontes,M., Padmanabhan,S., Murillo,F.J. and Elías-Arnanz,M. (2014) The CarD/CarG regulatory complex is required for the action of several members of the large set of *Myxococcus xanthus* extracytoplasmic function sigma factors. *Environ. Microbiol.*, **16**, 2475–2490.
43. Elías-Arnanz,M., Padmanabhan,S. and Murillo,F.J. (2010) The regulatory action of the myxobacterial CarD/CarG complex: a bacterial enhanceosome? *FEMS Microbiol. Rev.*, **34**, 764–778.
44. García-Heras,F., Abellón-Ruiz,J., Murillo,F.J., Padmanabhan,S. and Elías-Arnanz,M. (2013) High-mobility-group a-like CarD binds to a DNA site optimized for affinity and position and to RNA polymerase to regulate a light-inducible promoter in *Myxococcus xanthus*. *J. Bacteriol.*, **195**, 378–388.
45. García-Heras,F., Padmanabhan,S., Murillo,F.J. and Elías-Arnanz,M. (2009) Functional equivalence of HMGA- and histone H1-like domains in a bacterial transcriptional factor. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 13546–13551.
46. Padmanabhan,S., Elías-Arnanz,M., Carpio,E., Aparicio,P. and Murillo,F.J. (2001) Domain architecture of a high mobility group A-type bacterial transcriptional factor. *J. Biol. Chem.*, **276**, 41566–41575.
47. Peñalver-Mellado,M., García-Heras,F., Padmanabhan,S., García-Moreno,D., Murillo,F.J. and Elías-Arnanz,M. (2006) Recruitment of a novel zinc-bound transcriptional factor by a bacterial HMGA-type protein is required for regulating multiple processes in *Myxococcus xanthus*. *Mol. Microbiol.*, **61**, 910–926.
48. Bernal-Bernal,D., Gallego-García,A., García-Martínez,G., García-Heras,F., Jiménez,M.A., Padmanabhan,S. and Elías-Arnanz,M. (2015) Structure-function dissection of *Myxococcus xanthus* CarD N-Terminal domain, a defining member of the CarD_CdnL_TRCF family of RNA polymerase interacting proteins. *PLoS One*, **10**, e0121322.
49. Cayuela,M.L., Elías-Arnanz,M., Peñalver-Mellado,M., Padmanabhan,S. and Murillo,F.J. (2003) The *Stigmatella aurantiaca* homolog of *Myxococcus xanthus* high-mobility-group A-type

transcription factor CarD: insights into the functional modules of CarD and their distribution in bacteria. *J. Bacteriol.*, **185**, 3527–3537.
50. Iniesta,A.A., García-Heras,F., Abellón-Ruiz,J., Gallego-García,A. and Elías-Arnanz,M. (2012) Two systems for conditional gene expression in *Myxococcus xanthus* inducible by isopropyl-β-D-thiogalactopyranoside or vanillate. *J. Bacteriol.*, **194**, 5875–5885.
51. Smyth,G.K., Michaud,J. and Scott,H.S. (2005) Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics*, **21**, 2067–2075.
52. Ritchie,M.E., Silver,J., Oshlack,A., Holmes,M., Diyagama,D., Holloway,A. and Smyth,G.K. (2007) A comparison of background correction methods for two-colour microarrays. *Bioinformatics*, **23**, 2700–2707.
53. Bolstad,B.M., Irizarry,R.A., Astrand,M. and Speed,T.P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**, 185–193.
54. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, 1–25.
55. Ogawa,M., Fujitani,S., Mao,X., Inouye,S. and Komano,T. (1996) FruA, a putative transcription factor essential for the development of *Myxococcus xanthus*. *Mol. Microbiol.*, **22**, 757–767.
56. Gallego-García,A., Mirassou,Y., García-Moreno,D., Elías-Arnanz,M., Jiménez,M.A. and Padmanabhan,S. (2014) Structural insights into RNA polymerase recognition and essential function of *Myxococcus xanthus* CdnL. *PLoS One*, **9**, e108946.
57. Denhardt,D.T. (1966) A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.*, **23**, 641–646.
58. Han,Y. and Grierson,D. (2002) The influence of inverted repeats on the production of small antisense RNAs involved in gene silencing. *Mol. Genet. Genomics*, **267**, 629–635.
59. Rousseau,C., Gonnet,M., Le Romancer,M. and Nicolas,J. (2009) CRISPI: a CRISPR interactive database. *Bioinformatics*, **25**, 3317–3318.
60. Biswas,A., Gagnon,J.N., Brouns,S.J., Fineran,P.C. and Brown,C.M. (2013) CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol.*, **10**, 817–827.
61. Rutherford,K., Parkhill,J., Crook,J., Horsnell,T., Rice,P., Rajandream,M.A. and Barrell,B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics*, **16**, 944–945.
62. Munch,R., Hiller,K., Grote,A., Scheer,M., Klein,J., Schobert,M. and Jahn,D. (2005) Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*, **21**, 4187–4189.
63. Hofacker,I.L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.*, **31**, 3429–3431.
64. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
65. Galbis-Martínez,M., Fontes,M. and Murillo,F.J. (2004) The high-mobility group A-type protein CarD of the bacterium *Myxococcus xanthus* as a transcription factor for several distinct vegetative genes. *Genetics*, **167**, 1585–1595.
66. Campagne,S., Allain,F.H. and Vorholt,J.A. (2015) Extra Cytoplasmic Function sigma factors, recent structural insights into promoter recognition and regulation. *Curr. Opin. Struct. Biol.*, **30**, 71–78.
67. Gaballa,A., Guariglia-Oropeza,V., Durr,F., Butcher,B.G., Chen,A.Y., Chandrangsu,P. and Helmann,J.D. (2018) Modulation of extracytoplasmic function (ECF) sigma factor promoter selectivity by spacer region sequence. *Nucleic Acids Res.*, **46**, 134–145.
68. Herring,C.D., Raffaelle,M., Allen,T.E., Kanin,E.I., Landick,R., Ansari,A.Z. and Palsson,B.O. (2005) Immobilization of *Escherichia coli* RNA polymerase and location of binding sites by use of chromatin immunoprecipitation and microarrays. *J. Bacteriol.*, **187**, 6166–6174.
69. Carte,J., Christopher,R.T., Smith,J.T., Olson,S., Barrangou,R., Moineau,S., Glover,C.V. 3rd, Graveley,B.R., Terns,R.M. and Terns,M.P. (2014) The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.*, **93**, 98–112.
70. Biran,D. and Kroos,L. (1997) *In vitro* transcription of *Myxococcus xanthus* genes with RNA polymerase containing *sigmaA*, the major sigma factor in growing cells. *Mol. Microbiol.*, **25**, 463–472.
71. Chen,C., Zhang,H., Broitman,S.L., Reiche,M., Farrell,I., Cooperman,B.S. and Goldman,Y.E. (2013) Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat. Struct. Mol. Biol.*, **20**, 582–588.
72. Silas,S., Lucas-Elio,P., Jackson,S.A., Aroca-Crevillen,A., Hansen,L.L., Fineran,P.C., Fire,A.Z. and Sanchez-Amat,A. (2017) Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. *Elife*, **6**, e27601.
73. Wallace,R.A., Black,W.P., Yang,X. and Yang,Z. (2014) A CRISPR with roles in *Myxococcus xanthus* development and exopolysaccharide production. *J. Bacteriol.*, **196**, 4036–4043.
74. Bae,B., Chen,J., Davis,E., Leon,K., Darst,S.A. and Campbell,E.A. (2015) CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *Elife*, **4**, e08505.
75. Gallego-García,A., Iniesta,A.A., Gonzalez,D., Collier,J., Padmanabhan,S. and Elías-Arnanz,M. (2017) *Caulobacter crescentus* CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. *Sci. Rep.*, **7**, 43240.
76. García-Moreno,D., Abellón-Ruiz,J., García-Heras,F., Murillo,F.J., Padmanabhan,S. and Elías-Arnanz,M. (2010) CdnL, a member of the large CarD-like family of bacterial proteins, is vital for *Myxococcus xanthus* and differs functionally from the global transcriptional regulator CarD. *Nucleic Acids Res.*, **38**, 4586–4598.
77. Blatch,G.L. and Lassle,M. (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays*, **21**, 932–939.
78. Aravind,L. and Koonin,E.V. (2002) Classification of the caspase-hemoglobinase fold: detection of new families and implications for the origin of the eukaryotic separins. *Proteins*, **46**, 355–367.
79. Klemencic,M., Novinec,M. and Dolinar,M. (2015) Orthocaspases are proteolytically active prokaryotic caspase homologues: the case of *Microcystis aeruginosa*. *Mol. Microbiol.*, **98**, 142–150.
80. Risser,D.D. and Callahan,S.M. (2008) HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. *J. Bacteriol.*, **190**, 7645–7654.
81. Bastiaansen,K.C., Otero-Asman,J.R., Luirink,J., Bitter,W. and Llamas,M.A. (2015) Processing of cell-surface signalling anti-sigma factors prior to signal recognition is a conserved autoproteolytic mechanism that produces two functional domains. *Environ. Microbiol.*, **17**, 3263–3277.
82. Wielgoss,S., Didelot,X., Chaudhuri,R.R., Liu,X., Weedall,G.D., Velicer,G.J. and Vos,M. (2016) A barrier to homologous recombination between sympatric strains of the cooperative soil bacterium *Myxococcus xanthus*. *ISME J.*, **10**, 2468–2477.