PNAS
www.pnas.org

Supplementary Information for:

A majority of *R. sphaeroides* promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation

Kemardo K. Henry, Wilma Ross, Kevin S. Myers, Kimberly C. Lemmer, Jessica M. Vera, Robert Landick, Timothy J. Donohue, Richard L. Gourse

Corresponding authors: Richard L. Gourse, Wilma Ross

Email: rgourse@bact.wisc.edu, weross@wisc.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.2010087117

1. Expanded Materials and Methods

Bacterial strains and growth conditions. Bacterial strains are listed in Table S1. *E. coli* strains were grown at 37°C in either LB Lennox media or on LB agar plates supplemented with kanamycin (30 µg/mL) or chloramphenicol (20 µg/mL) when required. *R. sphaeroides* 2.4.1 strains were grown in succinate-based minimal medium (SIS) (1) at 30°C with kanamycin 25 µg/mL or sucrose (10%) when appropriate. For liquid growth conditions, *R. sphaeroides* 2.4.1 was grown in 500 mL cultures with constant sparging with 69% N_2, 30% O_2 and 1% CO_2 at 30°C for aerobic growth. For aerobic growth curves, samples were removed from the main culture under positive pressure, optical density was determined at OD_{600}, and viable cell counts were determined by plating on solid minimal medium.

Construction of chromosomally encoded *R. sphaeroides* His10-β'. For purification of RNAP by Ni-NTA chromatography, a thrombin-His10 tag was constructed in frame with the C-terminus of the RNAP β' subunit and introduced into wild-type *R. sphaeroides* 2.4.1 using the non-replicative plasmid pk18mobsacB (2) and a two-step recombination method, and the primers are listed in Table S4.

First, a 3 kb fragment containing 1.5 kb of sequence flanking either side of the stop codon of *rpoC* was PCR amplified from chromosomal DNA with primers RLG7576 and RLG7577. The pk18mobsacB vector was PCR linearized with oligonucleotides RLG7578 and RLG7579 that contained sequence overlapping the 3 kb fragment. The 3 kb fragment was then joined to linearized pk18mobsacB by HiFi Assembly (NEB) to

form pRLG14688. Insertion of the 3 kb fragment into pk18mobsacB was confirmed with primers RLG7584 and RLG7585.

Next, pRLG14688 was linearized by performing PCR with outward directed primers RLG7594 and RLG7589 lacking the *rpoC* stop codon. The thrombin-His10 tag was produced by annealing complementary oligonucleotides RLG7588 and RLG7589, which were then extended by PCR to create a 130 bp fragment with sequence overlapping the arms of linearized pRLG14688. This fragment was then joined to linearized RLG14688 by Hi-Fi Assembly to create pRLG14713. The resulting plasmid was verified by PCR with primers that annealed to the thrombin-His10 coding sequence (RLG7598) and within the *rpoC* gene (RLG7431) and then sequenced using the oligonucleotides listed in Table S4.

pRLG14713 was conjugated into *R. sphaeroides* 2.4.1 using *E. coli* S17-1 as a donor. Strains containing single crossovers were selected for kanamycin resistance, indicating integration of the pk18mobsacB into the *rpoC* locus. Double crossovers were then selected for the loss of sucrose sensitivity on agar plates containing sucrose (10%), and strains containing the thrombin-His10 coding sequence were identified by PCR using primers RLG7598 and RLG7431. The genomic region flanking the site of homologous recombination was PCR amplified using oligonucleotides RLG7511 and RLG7654, and the PCR fragment was sequenced. Glycerol stocks of the resulting strain RLG14765 were made from single colonies and stored at -80°C.

Purification of *R. sphaeroides* RNAP Holoenzyme. *R. sphaeroides* holoenzyme was purified using Ni-agarose affinity chromatography from aerobically grown cells containing the chromosomally encoded His10-β' (RLG14765). Cells were harvested at OD_{600} ~0.5-0.6 (log phase), and cell pellets were resuspended in Bugbuster (Novagen), 5 mL of Bugbuster per gram of wet cell, supplemented with 5 µL per gram Lysonase (Novagen), 1X HALT protease inhibitor (ThermoFisher Scientific), and 23 µg/mL phenylmethylsulfonyl fluoride (PMSF). After incubation at room temperature with gentle rocking for 1 hour, cell debris was removed by centrifugation at 14,000 RPM for 20 min at 4°C, and the cleared lysate was then added to a 1 mL Ni-NTA resin column (Qiagen) pre-equilibrated with RNAP wash buffer (50 mM NaH_2PO_4 [pH 8], 300 mM NaCl, 20 mM imidazole, 5% glycerol, 0.1% Tween 20, and 5 mM β-mercaptoethanol). After washing the column with RNAP wash buffer, RNAP was eluted with wash buffer containing 300 mM imidazole, dialyzed twice into 2X RNAP storage buffer (10 mM Tris-Cl [pH 8.0], 100 mM KCl, 10 mM MgCl2, 0.1 mM EDTA, 1 mM DTT, 20 µM ZnCl2) at 4°C, and concentrated using a 5 mL spin concentrator (Sartorius) with a 5 kDa molecular weight cutoff. An equal volume of 100% glycerol was added, and the protein concentration was assayed using the Bradford reagent (Bio-Rad) using BSA as a standard. Proteins were stored at -20°C. Native *E. coli* holo-RNAP was purified as described (3). *R. sphaeroides* σ^{93} was a gift from K. Wassarman (UW-Madison).

As is often observed with RNAP preparations made this way, a small number of other proteins co-purified with the *R. sphaeroides* RNAP. These proteins are present in very low, substoichiometric amounts compared to the other subunits (see Figure S1) and

could have co-purified with the RNAP due to weak interactions with RNAP itself (e.g. they could represent true RNAP-associated factors), or they could have co-purified because of affinity for the N-NTA agarose column matrix. Attempts to purify the *R. sphaeroides* RNAP further by heparin-sepharose chromatography were not successful. Although we cannot exclude the possibility that these proteins could influence transcription, they should not affect the interpretation of the results reported here, which in almost all cases compare the effects of the same RNAP preparation on different promoters or the effects of CarD on the same RNAP preparation.

Construction, overexpression, and purification of SUMO-tagged CarD wild-type and variant proteins. The *R. sphaeroides carD* gene was cloned into the pETSUMO vector for overexpression and purification from *E. coli* BL21 using its IPTG-inducible T7 promoter. The coding sequence of CarD (*rsp_2425*) was produced as a codon optimized gene block (Table S5) from GeneArt (Life Technologies). CarD was PCR amplified from the gene block with primers RLG7660 and RLG7661 (Table S6) containing sequence homology to pETSUMO that had been linearized with oligonucleotides RLG7369 and RLG7659. The two DNA fragments were joined by Hi-Fi Assembly (NEB), electroporated into *E. coli* XL1-blue cells, and grown on LB supplemented with kanamycin (30 μg/mL). Plasmids with the correct insert were identified by PCR with oligonucleotides RLG4583 and RLG7386 and then sequenced with the oligonucleotides listed in Table S6. The resulting plasmid (pRLG15433) created an Ulp1 cleavable His10-SUMO tag on the N-terminus of CarD. Substitutions were

introduced by Multi-site Lightning QuikChange (Agilent) using oligonucleotides listed in Table S6 and verified by sequencing.

His10-SUMO-tagged CarD was overexpressed from BL21 (DE3) pLysS cells transformed with pETSUMO coding for CarD or CarD variants. Transformed cells were grown in 500 mL of LB containing kanamycin and chloramphenicol inoculated from an overnight culture to a starting OD_{600} of 0.05. Cultures were grown at 37°C to OD_{600} of 0.5-0.6, induced with 1 mM IPTG for 3-4 hr, and purified by Ni-NTA chromatography as described (4). Cell pellets were resuspended in 25 mL of lysis buffer (20 mM Tris-Cl pH 8, 500 mM NaCl, 10% glycerol (vol/vol), 0.5 mM β-mercaptoethanol), PMSF (23 µg/ml), 0.08% sodium deoxycholate and 1X HALT protease inhibitor cocktail (ThermoFisher Scientific), and lysed by sonication. Cell debris was pelleted by centrifugation at 15,000 RPM for 30 min at 4°C, and the cleared lysate was filtered through a 0.45 µm filter added to a 0.5 mL Ni-NTA agarose resin (Qiagen) pre-equilibrated with lysis buffer. The protein-Ni-NTA column was washed with lysis buffer containing 25 mM imidazole, and the protein was eluted with lysis buffer containing 300 mM imidazole. The concentration of the eluted His10-SUMO-tagged CarD was determined using the Bradford assay reagent with bovine serum albumen (BSA) as a standard. The His10-SUMO tag was removed by incubating the eluted His10-SUMO-tagged CarD with His8-Ulp1 protease at a mass ratio of 400:1 (His10-SUMO-tagged CarD to Ulp1 protease) and dialyzed against Ulp1 cleavage buffer (20 mM Tris (pH 8), 150 mM NaCl, and 1 mM β-mercaptoethanol) at 4°C overnight. Free His10-SUMO and His8-Ulp1 protease were removed by passage through a ~300 µL Ni-NTA column. The flow-through containing

untagged CarD was dialyzed overnight against CarD storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 2 mM DTT, 50% glycerol). Protein concentration was assayed by the Bradford assay. Proteins were stored at -20°C. For in vitro transcription assays, CarD was diluted into CarD storage buffer to the desired concentration.

We cannot eliminate the possibility that the reduction in activity of the CarD mutant proteins in our studies (and in previous studies on CarD) resulted from general folding defects. We also note that we could not compare the effects of the triple substitution, $CarD_{Rsp}$ variant Q31A, I33A, R53A, corresponding to *T. thermophilus* $CarD_{Mtb}$ RID domain residues R25, I27, and R47, to those of an analogous triple RID substitution in *T. thermophilus* CarD, because a triple substitution variant was not constructed in previous studies.

Construction of in vitro transcription templates. *R. sphaeroides* promoter fragments were PCR amplified from *R. sphaeroides* 2.4.1 chromosomal DNA and inserted into pRLG770 (5) using the oligonucleotides listed in Table S2. Each oligonucleotide set created promoter fragments with the described endpoints relative to the predicted transcription start site based on the presence of predicted -10 and -35 elements. The resulting promoter fragments contained sequences for ligation to linearized pRLG770 (amplified using oligonucleotides RLG7057 and RLG7056) by Hi-Fi Assembly (NEB). Plasmids were electroporated into *E. coli* XL1-blue, and those confirmed to have the correct insert size on gels were sequenced with oligonucleotides RLG4252 and RLG4253. Promoter variants were created by Multi-site Lightning QuikChange (Agilent)

as described above using the mutagenic primers listed in Table S2. Promoter insertions in pRLG15507 and pRLG15543 (Table S5) were gene blocks from IDT that were ligated to linearized pRLG770. pRLG770 and its derivatives all contain the promoter and gene coding for the RNA I transcript, a naturally occurring product of ColE1 plasmids from which these plasmids derive. The RNA I promoter is located within the plasmid replication control region, and makes RNA I, an antisense regulatory RNA that controls the primer for DNA replication, RNA II (6). pRLG15474 containing the *T7A1lacO34* promoter was amplified from pIND4 (a replicable plasmid used for protein overexpression in *R. sphaeroides*) (7). *T7A1lacO34* is sometimes referred to as *T7A1lacO-1* (8). Plasmid DNA was purified from *E. coli* XL1-blue by miniprep (Qiagen), phenol extracted, diluted to 200 ng/µL in 10 mM Tris-HCl (pH 8.0), and stored at -20°C.

in vitro transcription. Multiple round transcription assays were carried out on supercoiled plasmid templates containing the promoters listed in Table S3 using *R. sphaeroides* RNAP (20-50 nM) or *E. coli* RNAP (10 nM) in 25 µL reactions. Template DNA (5 pmol or 400 ng) was incubated in transcription buffer (20 mM Tris-HCl [pH 7.9], 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA, and either 170 or 100 mM NaCl, as indicated, containing NTPs (200 µM ATP, GTP, CTP, 10 µM UTP, and 1.2 µCi α-^{32}P UTP [Perkin Elmer]) with CarD or CarD storage buffer as indicated at room temperature for 10 min. Transcription was initiated by addition of RNAP and reactions were incubated at 30°C for 20 min. The reactions were terminated by addition of 25 µL water-saturated phenol, aqueous and phenol phases were separated by brief centrifugation, and 16 µL of the aqueous phase containing the transcript was transferred to 16 µL 2X

loading solution (8 M urea, 10 mM EDTA, 2X TBE, 1% SDS and 0.02% bromophenol blue). Transcripts were denatured by heating the samples at 95°C for 1 min, resolved on 6% polyacrylamide 7M urea gels, imaged using a Typhoon phosphorimager, and quantified with ImageQuant 5.2 software.

Western blot analysis. *R. sphaeroides* cells were harvested from aerobically grown cultures at the OD_{600} indicated. Separate cultures were grown for each of the three growth phases, outgrowth, log, and stationary (Figs. 7 and S8). For outgrowth cultures, an aliquot of a stationary phase culture (OD_{600} ~3.5) was diluted 1:12.5 into fresh minimal medium to a starting OD_{600} of 0.25 at time 0, and samples were removed at the indicated times (from t=0 to 7 h). For the log phase culture, an aliquot of a stationary phase culture was diluted 1:50 to a starting OD_{600} of 0.02 into fresh medium at time 0, the culture was grown for 10 hr to an OD_{600} of ~0.2 before sampling, and samples were then collected at intervals from 10 to 18 hr, during which time the culture grew exponentially. For the stationary culture experiment, the culture was inoculated from a stationary phase culture to a starting OD_{600} of ~0.013 at t=0 and was sampled at times between 15 to 45 h, during which it entered and continued through stationary phase.

At the indicated OD_{600}, 2 OD equivalents of cells were pelleted, resuspended in 200 µL Laemmli loading buffer (75 mM Tris-HCl [pH 8], 2% SDS, 0.5 M β-mercaptoethanol, 10% glycerol, 0.2% Triton X100, 0.02% bromophenol blue), and heated at 95°C for 10 min. The whole cell lysates were cleared by brief centrifugation, and aliquots (10 µL/lane or 0.1 OD_{600} equivalent) and purified standards were separated on 12%

NuPAGE Bis-Tris gels (Invitrogen) in MES buffer. Proteins were then transferred onto a nitrocellulose membrane (GE Healthcare). The section of the membrane containing the CarD band was incubated for 2 hr with a primary antibody against CarD (rabbit polyclonal antibody raised against untagged CarD; Covance Research Products), and the section containing the α subunit of RNAP was incubated for 1 hr with a primary antibody against α at a 1:10,000 dilution at room temperature in blocking buffer (phosphate buffered saline [pH 7.4], 0.1% Tween20 (PBST) and 5% dried milk). In Fig. S1*B*, sections of the membrane were incubated with antibodies against the σ^{70} or the α subunits of RNAP to determine if the *R. sphaeroides* RNAP contained σ^{93} (with which the *E. coli* anti-σ^{70} cross reacts). The membranes were subsequently washed 3 times with PBST and then incubated with secondary antibody (goat anti-rabbit IgG horseradish peroxidase conjugate [Promega]; added at 1:10,000) for 1 hr in blocking buffer. The membrane was then washed 3 times with PBST again, incubated in SuperSignal West Dura Extended Duration Substrate (Thermo Scientific), and bands were detected with a LI-COR Odyssey Fc Imaging System. The intensity of each band was quantified using Image Studio Lite software (LI-COR). Polyclonal antibodies to the α subunit of *R. sphaeroides* RNAP (WI515) and σ^{70} (WI168) were gifts from Karen Wassarman (University of Wisconsin-Madison). CarD protein concentrations in the whole cell lysate were determined from a standard curve.

Viable cell counts. *R. sphaeroides* colony forming units (CFU) were determined by 10-fold serially diluting cultures into 1X M9 salts at the OD_{600} indicated and plating 10 uL (in triplicate) onto 1X SIS agar plate. Viable cells were scored after 48-96 hr at 30°C. The

number of viable cells loaded/lane (CFU/lane) was determined by dividing the CFU by the respective OD_{600} value multiplied by 0.1 (the OD_{600} equivalent loaded per lane).

Determination of CarD concentration in cells. The molar concentration of CarD in log phase cells was calculated from the number of moles of CarD in the samples, as determined from Western blots with purified CarD standards using a molecular weight of 19.045 g/mol for CarD, the number of cells per sample (CFU) determined from viable cell counts, and the volume of a log phase aerobically grown *R. sphaeroides* cell (0.87 μm^3 per CFU) (9).

Bioinformatic analysis of transcription start sites and promoter elements. Total *R. sphaeroides* RNA was extracted from separate cultures grown to mid-log phase aerobically or anaerobically, as described in the "Bacterial strains and growth conditions" section above and in reference (9). The RNA was then processed as described in reference (10), and transcription start sites (TSSs) were identified as described in reference (11). Both the anaerobic and aerobic sets of TSSs were then combined for promoter analysis. All TSSs were included for identification of promoter elements; the search did not exclude promoters used by holoenzymes containing sigma factors other than the major sigma, σ^{93} in *R. sphaeroides*. However, since most alternative sigmas are produced or activated only under specific stress conditions, most of promoters identified under the mid-log growth conditions we used would be expected to correspond to those utilizing σ^{93} holoenzyme.

For Fig. 6, we searched for probable -10 elements by analyzing the DNA sequences from -19 to -5 relative to each TSS (+1) using MEME version 5.1.0 (12) ("zoops" method, minimum width of 9 bp, maximum width of 10 bp, no palindromic sequence). TA motifs (corresponding to positions most frequently found at -12 and -11 at the upstream end of *E. coli* -10 elements) were identified in 763 putative promoters. The TA motifs were not required to be positioned at -12 and -11 relative to the TSS, since the spacing between the -10 hexamer and the TSS can vary. The base at position -7 in *R. sphaeroides* promoters was identified relative to the position of the TA motif. In cases where there were multiple closely-spaced TSSs associated with the same -10 element (~9% of genes), only one -10 element was used for quantitation of the percentage of -7T and extended -10 elements in the population. In cases where there was more than one TSS associated with the same gene, but they were not close together, the corresponding -10 elements for each putative promoter were included in the analysis. Since the genome of *R. sphaeroides* does not contain a gene coding for a σ^S homolog (13), and alternative holoenzymes (at least in *E. coli*) do not contain a TA motif at the upstream end of their -10 elements, the promoter collection analyzed in Fig. 6 for -7T and extended -10 elements likely consists primarily of σ^{93}-dependent promoters.

The -10 elements in our promoter library were separated into two groups based on the identity of the base at position -7, one group with -7T and the other with -7A/G/C combined. The percentage of promoters containing -7T, extended -10 elements, or both was determined for each species in Fig. 6 using published TSS information. The list of all *R. sphaeroides* TSSs, indicating those with probable -7T and extended -10 sequences, is provided as Dataset S1. We note that the number of *R. sphaeroides* promoters identified by MEME in our analysis is an underestimate of the true number of *R. sphaeroides* promoters, because we specified -10 elements with both a T at -12 and an A at -11. The same methods were used to calculate the percentages of promoters with -7T, extended -10 elements, or both for each bacterial species shown in Fig. 6. The sources of the TSS data used for this purpose were *R. sphaeroides* (11), *C. crescentus* (14), *Z. mobilis* (10), *E. coli* (15), *B. subtilis* (16)*, M. smegmatis* (17), *S. coelicolor* (18), and *B. cenocepacia* (19). Using our criteria, we estimate that ~65% of the TSSs were included in the analyses of -10 elements from each species.

Table S1. Strains and plasmids used in this study

Strains	**Genotype**	**Source**
BL21 (DE3) pLysS	*E. coli* str. B F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S) pLysS[*T7p20 ori*$_{p15A}$](Cm^R)	Novagen
E. coli S17-1	TpR SmR recA, thi, pro, hsdR-M+RP4: 2-Tc:Mu: Km Tn7 λpir	(20)
E. coli XL1-Blue	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1* gyrA96 relA1 lac *[F'* proAB lacI^{q}ZΔM15 *Tn*10 *(Tet^r)]*	Stratagene
E. coli XL10-Gold	*endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]*	Stratagene
RLG14946	*Rhodobacter sphaeroides* 2.4.1 (Wild type lab strain)	
RLG14765	RLG14946 with a C-terminal thrombin-His_{10} chromosomally tagged β' (*rpoC*)	This study

Plasmids	**Description**	**Source**
pRLG770	Transcription vector	(21)
pETSUMO	pET derived plasmid that encodes a Ulp1 cleavable His_{10}SUMO tag. ORFs are cloned in frame with the coding sequence of the SUMO tag.	Invitrogen (4)
pK18mobsacB	Mobilization plasmid used for making chromosomal insertion into *R. sphaeroides* chromosome	(2)

Table S2. Oligonucleotides used to construct plasmid templates for in vitro transcription

Primers for linearizing RLG770

Oligonucleotide	Sequence
7057	5'-AAGCTTGGGTCCCACCTGACC-3'
7058	5'-GAATTCTTGAAGACGAAAGGGCC-3'

Primers for screening and sequencing promoters inserted into RLG770

Oligonucleotide	Sequence
4252	5'-CCCGAAAAGTGCCACCTGACG-3'
4253	5'-CAGTTCCCTACTCTCGCATGG-3'

Primers to construct *in vitro* transcription plasmids
(both forward and reverse primers)

Plasmid	Oligo	Sequence	Description
pRLG14562	7411	GGCCCTTTCGTCTTCAAGAATTC GATCCGGCGGATGGGAGCGGC	*R. sphaeroides rrnA* promoter, -320 to +50
	7412	GGTCAGGTGGGACCCAAGCTTG GCAGCACCGCCGCCTCTCAAC	
pRLG14564	7413	GGCCCTTTCGTCTTCAAGAATTC CCGCGAGCGGCTGACCGG	*R. sphaeroides rrnB* promoter, -230 to +50
	7415	GGTCAGGTGGGACCCAAGCTTCT GTCATCCGCTCCCTCGTCTCCGT CC	
pRLG14566	7416	GGCCCTTTCGTCTTCAAGAATTC GCCGTCTCGAGGGGGATCTGGA G	*R. sphaeroides rrnC* promoter, -270 to +50
	7412	GGTCAGGTGGGACCCAAGCTTG GCAGCACCGCCGCCTCTCAAC	
pRLG15054	8008	GGCCCTTTCGTCTTCAAGAATTC CATCCTGACGGAATGGGGCATC	*R. sphaeroides prrA* promoter, -379 to +51
	8009	GGTCAGGTGGGACCCAAGCTTA GTTCGAATACCAGATCCTCAGCC	
pRLG15224	8026	GGCCCTTTCGTCTTCAAGAATTC GGGCTGGCCCTTCTGGGATTTG	*R. sphaeroides rpsL* promoter, -132 to +65
	8136	GGTCAGGTGGGACCCAAGCTTG CCCGGATCATGACCACTACAG	
pRLG15228	8028	GGCCCTTTCGTCTTCAAGAATTCT CATCGGCTACTACTATGCCAAGG	*R. sphaeroides rpsM* promoter, -156 to +50
	8143	GGTCAGGTGGGACCCAAGCTTG ACTGTTCGCCGCAGGACAAAACC	
pRLG15227	8141	GGCCCTTTCGTCTTCAAGAATTC GTTGATGGCGGGCGATCTCGAG	*R. sphaeroides* $tRNA^{trp}$ RSP_4298 promoter, -106 to +56
	8142	GGTCAGGTGGGACCCAAGCTTCA CGACCCTCGGTTTTGGAGACC	

pRLG15229	8144	GGCCCTTTCGTCTTCAAGAATTC CCGCTCGAGGCTTGAGGATC	*R. sphaeroides rpoZ* promoter, -68 to +62
	8145	GGTCAGGTGGGACCCAAGCTTCA ACGCAGTCTTCAACCGTCACG	
pRLG15230	8146	GGCCCTTTCGTCTTCAAGAATTC GCGCCGGGGAAAGACAAGTGAT AC	*R. sphaeroides tufA* promoter, -60 to +25
	8147	GGTCAGGTGGGACCCAAGCTTC GACACCGGACAGCCCGTTC	
pRLG15472	8303	GGCCCTTTCGTCTTCAAGAATTCT CACCGCCCGATAGTGCTTC	*R. sphaeroides tRNA-val* (*RSP*_4309) promoter, -79 to +47
	8304	GGTCAGGTGGGACCCAAGCTTCA CACCTGTTCCGCGCGAG	
pRLG15474	8309	GGCCCTTTCGTCTTCAAGAATTC CCTATAAAAATAGGCGTATCACG AGG	Inducible pT7A1lacO34 promoter, -102 to +31 IPTG inducible promoter in pIND4 (5) (for overexpression in *R. sphaeroides*)
	8310	GGTCAGGTGGGACCCAAGCTTG ATGTGTGAAATTGTTATCCGCTC	
pRLG15601	8569	GGCCCTTTCGTCTTCAAGAATTC GATGCTGTGAAAAAGGCCGATGG	*R. sphaeroides acpP* promoter, -89 to +46
	8570	GGTCAGGTGGGACCCAAGCTTCA GGCAACTGGTGCAGTTCAG	
pRLG15602	8573	GGCCCTTTCGTCTTCAAGAATTC GGTCGGACGTCCCTGCACTG	*R. sphaeroides comL* promoter, -88 to +35
	8574	GGTCAGGTGGGACCCAAGCTTC CTCTGCCCCGTCAGACCATAG	
pRLG15604	8581	GGCCCTTTCGTCTTCAAGAATTC GGTGCTCCTTGCGACATGATCTT C	*R. sphaeroides ilvE1* promoter, -95 to +80
	8582	GGTCAGGTGGGACCCAAGCTTCA TGTCACCCTCCTGCGCATAG	
pRLG15606	8589	GGCCCTTTCGTCTTCAAGAATTC GTCCGGGGGCTCTACAGGATTC	*R. sphaeroides RSP*_1486 promoter, -78 to +30
	8590	GGTCAGGTGGGACCCAAGCTTGT CCGTCGCCTGTCGTCATTG	
pRLG15607	8591	GGCCCTTTCGTCTTCAAGAATTC CAGACAGCGGTCGCGGATGAG	*R. sphaeroides RSP*_3110 promoter, -90 to +72
	8592	GGTCAGGTGGGACCCAAGCTTC GAGCATGTCGGCGACCTTTCC	

pRLG15611	8603	GGCCCTTTCGTCTTCAAGAATTC GTCTTCTCCGCTGTGGCGAAATG	*R. sphaeroides rho* promoter, -94 to +110
	8604	GGTCAGGTGGGACCCAAGCTTG AGGTTGAGACGTTCGTTCATGGC	
pRLG15613	8571	GGCCCTTTCGTCTTCAAGAATTC GCGGAGGACCAAGGAAACAATG G	*R. sphaeroides atpH* promoter, -91 to +49
	8572	GGTCAGGTGGGACCCAAGCTTCA GAGGCATTGTATGGCGCGAG	
pRLG15619	8597	GGCCCTTTCGTCTTCAAGAATTC GATTGCGTGCGTCGACGGTTTTC	*R. sphaeroides putR* promoter, -100 to +40
	8598	GGTCAGGTGGGACCCAAGCTTG GTCGATGGGGTCCAGTTTCGATC	

Mutagenic primers for construction of transcription templates

Plasmid	**Oligo**	**Sequence** (mutations are in lowercase)	**Description**
pRLG14563	7364	TGCGGTTCCCTAGATtGC GCCTCACCGAAGC	*R. sphaeroides rrnA* A-7T promoter
pRLG14565	7365	CCGTCTGCTCCTAGAAtC CGCTTCACCGAGACG	*R. sphaeroides rrnB* A-7T promoter
pRLG14567	7364	TGCGGTTCCCTAGATtGC GCCTCACCGAAGC	*R. sphaeroides rrnC* A-7T promoter
pRLG15180	8148	CGCCCGGGGCCGTCTTg tgCTAGAAACCGCTTCAC	*R. sphaeroides rrnB* ext-10 promoter (to -17 to -14 TGTG)
pRLG15181	8149	CCAAAAAATCCGCTTGac aCCGGGGCCGTCTGCTC	*R. sphaeroides rrnB* -35 consensus promoter
pRLG15182	8150	GTCTGCTCCTAGAAACat aTTCACCGAGACGAAGA C	*R. sphaeroides rrnB* discriminator mutant promoter (-5 to -3, CCG to ATA)
pRLG15183	8151	GTCTGCTCCTAGAAAggG CTTCACCGAGACGAAG	*R. sphaeroides rrnB* discriminator mutant promoter (-6 to -5, CC to GG)
pRLG15185	8153	CTTGCGCCCGGGGCCG aTCTGCTCCTAGAAACC G	*R. sphaeroides rrnB* promoter with 18 bp spacer. insertion of A at -21
pRLG15186	8154	CCGCTTGCGCCCGGGG CGTCTGCTCCTAGAAAC C	*R. sphaeroides rrnB* promoter with 16 bp spacer, deletion of C at -21

pRLG15350	8287	GGCCGTCTGCTCCTtGAAACCGCTTCACC	*R. sphaeroides rrnB* promoter mutant, A-11T
pRLG15351	8288	GGCCGTCTGCTCCTgGAAACCGCTTCACC	*R. sphaeroides rrnB* promoter mutant, A-11G
pRLG15352	8289	GCCGTCTGCTCCTAtAAACCGCTTCACCG	*R. sphaeroides rrnB* promoter mutant, G-10T
pRLG15485	8383	TGCGGGTTTTTTTGCGGTgtgCTAGATAGCGCCTCACCG	*R. sphaeroides rrnA*, ext-10 promoter mutant GTG from -16 to -14
pRLG15488	8376	GGGCGGCCGTTACATtGGCCGCTC	*R. sphaeroides rpoZ* promoter mutant, A-7T
pRLG15489	8377	GAGGCGGGGCGGCtGcTACATAGGCCGCT	*R. sphaeroides rpoZ* ext-10 consensus TGC from -15 to -13
pRLG15490	8378	GGGGCAAAAGCCCATAGATtACGGCGTCCC	*R. sphaeroides rpsM* promoter mutant, C-7T
pRLG15491	8379	GGCCGGGGCAAAAGCtgcTAGATCACGGCGTCC	*R. sphaeroides rpsM* ext-10 consensus, TGC from -15 to -13
pRLG15505	8381	TCTCTTCGTCATTTTTCCTCTTGCGGccggggccGCGGTTCCCTAGATAGCGCCTCACCG	*R. sphaeroides rrnA* promoter, CCGGGGCC spacer substitution from -29 to -22

Table S3. Plasmids used for in vitro transcription

Plasmids	Promoter[a]	Endpoints[b]	Source
pRLG770	No promoter		(21)
pRLG1616	*E. coli rrnB* (*rrnB* P1)	-88 to +50	(21)
pRLG593	*lacUV5*	-59 to +38	(22)
pRLG3265	*RNA I*	-61 to +1	(22)
pRLG14562	*R. sphaeroides rrnA*	-320 to +50	This study
pRLG14563	*R. sphaeroides rrnA* A-7T	-320 to +50	This study
pRLG14564	*R. sphaeroides rrnB*	-230 to +50	This study
pRLG14565	*R. sphaeroides rrnB* A-7T	-230 to +50	This study
pRLG14566	*R. sphaeroides rrnC*	-270 to +50	This study
pRLG14567	*R. sphaeroides rrnC* A-7T	-270 to +50	This study
pRLG15054	*R. sphaeroides prrA*	-379 to +51	This study
pRLG15224	*R. sphaeroides rpsL*	-132 to +65	This study
pRLG15227	*R. sphaeroides tRNA^{trp}* (*RSP*_4298)	-106 to +56	This study
pRLG15228	*R. sphaeroides rpsM*	-156 to +50	This study
pRLG15229	*R. sphaeroides rpoZ*	-68 to +62	This study
pRLG15230	*R. sphaeroides tufA*	-60 to +25	This study
pRLG15350	*R. sphaeroides rrnB* A-11T	-230 to +50	This study
pRLG15351	*R. sphaeroides rrnB* A-11G	-230 to +50	This study
pRLG15352	*R. sphaeroides rrnB* G-10T	-230 to +50	This study
pRLG15472	*R. sphaeroides tRNA^{val}* (*RSP*_4309)	-79 to +47	This study
pRLG15474	*T7A1lacO34 =A1lacO-1*	-102 to +31	(8)
pRLG15485	*R. sphaeroides rrnA*, ext-10 GTG -16 to -14	-320 to +50	This study
pRLG15488	*R. sphaeroides rpoZ* A-7T	-68 to +62	This study
pRLG15489	*R. sphaeroides rpoZ* ext-10 TGC -15 to -13	-68 to +62	This study
pRLG15490	*R. sphaeroides rpsM* C-7T	-156 to +50	This study
pRLG15491	*R. sphaeroides rpsM* ext-10 TGC -15 to -13	-156 to +50	This study
pRLG15505	*R. sphaeroides rrnA* CCGGGGCC from -29 to -22	-320 to +50	This study
pRLG15507	*R. sphaeroides ssrA*	-71 to +55	This study
pRLG15543	*R. sphaeroides rrnB* full consensus UP element	-230 to +50	This study
pRLG15601	*R. sphaeroides acpP*	-89 to +46	This study
pRLG15602	*R. sphaeroides comL*	-88 to +35	This study
pRLG15604	*R. sphaeroides ilvE1*	-95 to +80	This study
pRLG15606	*R. sphaeroides RSP*_1486 (TetR transcription regulator family)	-78 to +30	This study
pRLG15607	*R. sphaeroides RSP*_3110 (putative Glutathione S-transferase)	-90 to +72	This study
pRLG15611	*R. sphaeroides rho*	-94 to +110	This study

pRLG15613	*R. sphaeroides atpH*	-91 to +49	This study
pRLG15619	*R. sphaeroides putR*	-100 to +40	This study

[a] Promoters are named for the first gene in the operon
[b] endpoints are relative to the predicted transcription start site (unless previously published otherwise in the indicated source)

Table S4. Primers used to construct chromosomally-encoded C-terminal thrombin-His10 tagged β'

1. Insertion of *rpoC* chromosomal fragment into pk18mobsacB (pRLG14688)

Primers to isolate *rpoC* chromosomal fragment

Oligo	Sequence
7576	GCCGCCCAGTCGATCGG
7577	AGGCGAACCTTGGCGACCTTAC

Primers to linearized pK18mobsacB

Oligo	Sequence
7578	CGCTATGCGTAAGGTCGCCAAGGTTCGCCTAAGCTTGG CACTGGCCGTC
7579	CGTACCCGGTTCGCCGATCGACTGGGCGGCGGATCCC CGGGTACCGAG

Primers to confirm 3 kb insert in pK18mobsacB

Oligo	Sequence
7584	CACTCATTAGGCACCCCAGGC
7585	CGCCAGGGTTTTCCCAGTCAC

2. Creation of C-terminally tagged *rpoC* with thrombin-His_{10} in pK18mobsacB (pRLG14713)

Primers to linearize pK18mobsacB with *rpoC* chromosomal fragment

Oligonucleotide	Sequence
7594	CCTGACCATGGGCCCGC
7595	CAGCCCCGAGTCCTCGG

Primers to introduce thrombin-His_{10} into *rpoC* C-terminus (Region of complementarity is underlined)

Oligonucleotide	Sequence
7588	CCGAGGACTCGGGGCTGGTGGAAACGGTGGAGAACCGCGAG GAGCTGGTGCCGCGCGGCAGCCATCACCACCATCATCAC
7589	TGCGCGGGCCCATGGTCAGGCGAGAGAGGGGAGGATTAATG GTGGTGATGGTGATGATGGTGGTGATGGC

Primers to confirm insertion of thrombin-His_{10} tag onto *rpoC* C-terminus

Oligo	Sequence
7598	ATGATGGTGGTGATGGCTGCC
7431	GCGTCAAGAAGATCGCGCAC

Primers to sequence RLG14713 and RLG14765

Oligo	Sequence
7431	GCGTCAAGAAGATCGCGCAC
7432	GTGATCCCGAGCAGGATCGGTTC
7512	TGCCGATCATCGCCGAGAAGG
7513	GAACCGATCCTGCTCGGGATC
7515	CGGGTTGCCGTCCATGATG
7516	GCCAGTCGGACACGATCTTCTG
7584	CACTCATTAGGCACCCCAGGC
7594	CCTGACCATGGGCCCGC
7602	CGAGAGTCGCACGAAATCCTCC
7612	GGGCCTCTTCGCTATTACGCC
7615	CGTGCTGAAGGAGCCACTGG
7616	CCCTCCGACTCCCCCTATACG
7617	GCGCCAGATCGTGAAGAGC
7618	CGTACGATGAGCAGGACGCC

Primers to amplify crossover region in RLG14765

Oligo		Sequence
7511		GCTCACCTGTGAAGCGGAAGAG
7654		GATCTTTGACACCCTGGGTATCGAGG

Table S5. Gene blocks used in this study

Plasmid	Sequence[a]	Description
pRLG15507	**CACGAGGCCCTTTCGTCTTCAAGAATTC**TG CCCCACGGATCCGCCCCGGCACGCCTGCG GCCGCTTGCGGAGGCCGCCCCTCCCGCCT ATATTCCGCCTGTCCGGGGCAACCCGGACT ATGGACATAAACGCGCTCGTCATAAGCGGA TCGGAC**AAGCTTGGGTCCCACCTGACCCC ATG**	*R. sphaeroides ssrA* promoter -71 to +55
pRLG15543	**GGCCCTTTCGTCTTCAAGAATTCG**gaaaatttttt tcaaaaagtaCTTGCGCCCGGGGCCGTCTGCT CCTAGAAACCGCTTCACCGAGACGAAGACC GGCAGCGCCGGACGGAGACGAGGGAGCG GATGACAG**AAGCTTGGGTCCCACCTGACC**	*R. sphaeroides rrnB* promoter with full consensus UP element (-36 to +50)

Gene block for CarD (*RSP*_2425)

Plasmid	Sequence	Description
pRLG15433	**ATG**ACCAAAACCAAAAAACCGGAATTTCGT CCGAATGAGTTTGTTGTTTATCCGGCACAT GGTGTGGGTCGTATTATTAGCATTGAAGAA CAAGAAATTGCCGGTATTCGTCTGGAACTG TTTGTGATCAGCTTCGAGAAAGATAAAATGA CCCTGCGTGTTCCGACCCATAAAGCAACCG AAGTTGGTATGCGTAGCCTGAGCACACCGG ATGTTGTTACCAAAGCACTGGATACCCTGA AAGGTAAAGCACGTGTGAAACGTGCAATGT GGTCACGTCGTGCACAAGAATATGAACAGA AAATTAACAGCGGTGATCTGATGAGCATTG CCGAAGTTGTTCGTGATCTGCATCGTACCG ATGATCAGCGTGAACAGAGCTATAGCGAAC GTCAGCTGTATGAAGCAGCACTGGAACGTC TGACCCGTGAAGTTGCAGCAGTTAGCGGTG TTGATGAAGCCGGTGCACAGAAAGCAGTTG ATGCAGTTCTGGTTAGCCGTGCAGCA**TAA**	Sequence of codon optimized (GeneArt, Life Technologies) *R. sphaeroides* CarD (*RSP*_2425)

[a] Promoter sequences in these plasmids were gene blocks (IDT) inserted into RLG770 by HiFi Assembly (NEB). Sequence homology to RLG770 is in boldface. -35 element to -10 element is underlined. Consensus UP element introduced into *R. sphaeroides rrnB* promoter is shown in lowercase. Start and stop codons of CarD are in boldface.

Table S6. Primers used to construct pETSUMO-CarD

Primers to linearize pETSUMO (pRLG14641)

Oligo	Sequence
7369	TCGTAACAAAGCCCGAAAGGAAGCTG
7659	TCCACCAATCTGTTCTCTGTGAGCC

Primers to PCR amplify CarD with sequence overlapping pETSUMO

Plasmid	Oligo	Sequence	Description
pRLG15433	7660	ACAGAGAACAGATTGGTGGAATGACCAAAACCAAAAAACCGGAATTTCG	Primers to amplify codon optimized CarD, which generates overlapping sequence to pETSUMO (pRLG14641)
	7661	CCTTTCGGGCTTTGTTACGATTATGCTGCACGGCTAACCAGAAC	

Primers to sequence pETSUMO-CarD (pRLG15433)

Oligo	Sequence
7386	ACGCTGCGCGTAACCACC
7837	GACAGGGTAAGGAAATGGACTCC
7839	GCGTGAACAGAGCTATAGCGAACG
4583	GGTGATGTCGGCGATATAGGC
7840	CATCCGGTGTGCTCAGGCTAC
7838	GGTGTCCGATGGATCTTCAGAGATC

Mutagenic primers to construct CarD variants

Plasmid	Oligo	Sequence	Description
pRLG15053	8058	GTGTGAAACGTGCAATGGCGTCACGTCGTGCACAAG	CarD W91A substitution
pRLG15434	8195	TGTGAAACGTGCAATGTTGTCACGTCGTGCACAAG	CarD W91L substitution
pRLG15435	8190	TGGTGTGGGTCGTATTATTAGCATTGAAGAAGCAGAAGCTGCCGGTATTCGTCTG	CarD Q31A/I33A/ R53A substitution
	8191	GAAAGATAAAATGACCCTGGCTGTTCCGACCCATAAAGCA	

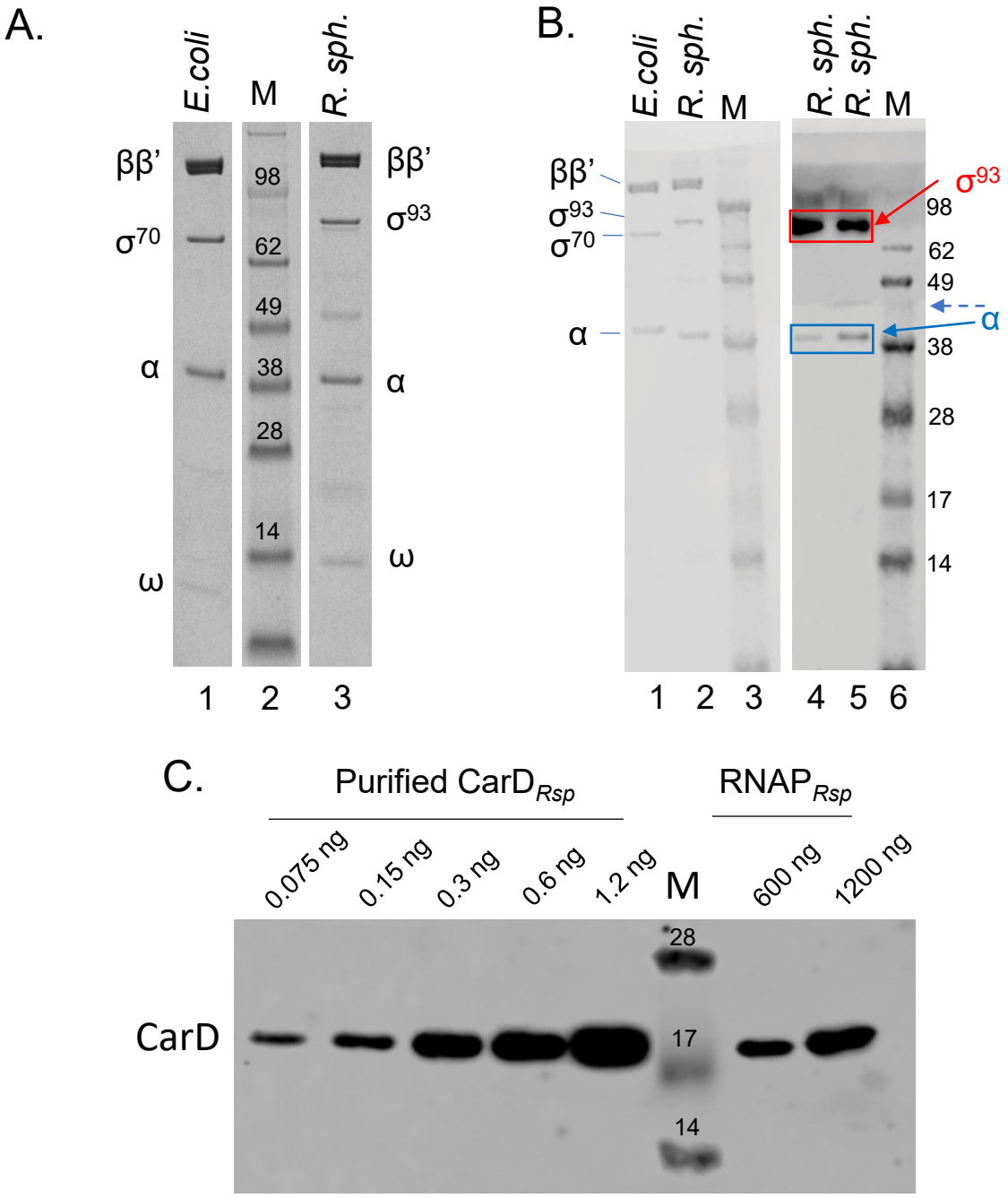


Fig. S1. RNAP purification. (*A*) 1.0 µg purified *E. coli* RNAP (lane 1) or *R. sphaeroid*es RNAP (lane 3) run on a 4-12% Invitrogen NuPAGE Bis-Tris gel with MES buffer. MW (kDa) of See Blue 2 size markers (Invitrogen, lane 2) are indicated above bands. Intervening lanes were removed for clarity. (*B*) All lanes are from the same gel (12% NuPAGE Bis-Tris/MES). Lanes 1- 3 were Coomassie stained. Lanes 1 and 2, 1 µg of the indicated RNAP. Lane 3, Size marker as in (*A*). Lanes 4-6, Western blot of purified $RNAP_{Rsp}$ to detect σ and α subunits. Lane 4, 10 ng $RNAP_{Rsp}$. Lane 5, 20 ng $RNAP_{Rsp}$, Lane 6, marker as in (*A*). Top portion of blot (above dotted arrow) was probed with anti-σ^{70} antibody WI168. Lower portion of blot was probed with anti-α antibody (WI515). All lanes (1-6) were visualized by Li-COR imaging. (*C*) Western blot with anti-$CarD_{Rsp}$ antibody, using indicated amounts of purified $CarD_{Rsp}$ or purified $RNAP_{Rsp}$. The RNAP:CarD molar ratio was 126:1. Marker (M) is the same as in (*A*).

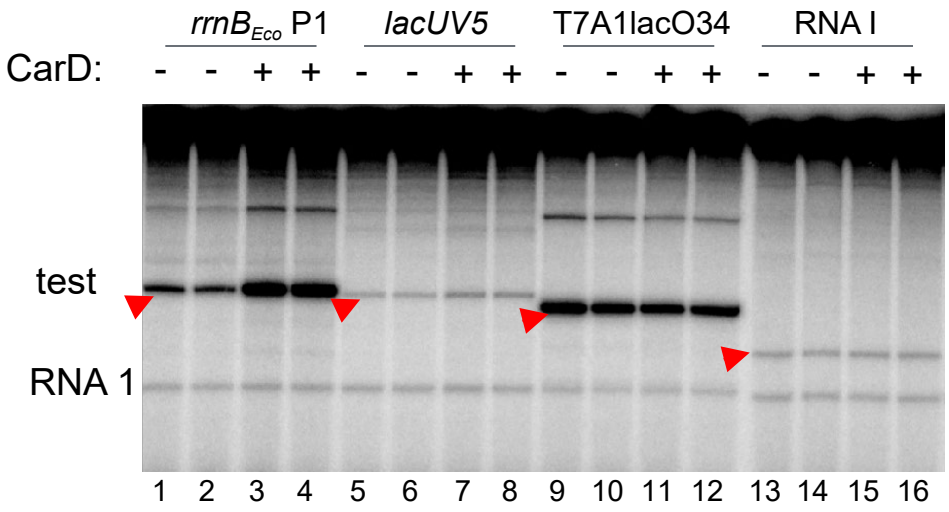


Fig. S2. Transcription of non-*R. sphaeroides* promoters with 20 nM RNAP$_{Rsp}$. Buffer contained 170 mM NaCl. Only *E. coli rrnB* P1 was activated by CarD$_{Rsp}$. Promoter fragments were inserted upstream of the *rrnB* P1 terminator to produce the transcripts indicated by the red arrowheads. RNA 1 is encode Where indicated the CarD concentration was 1280 nM. d by the plasmid template (see *Expanded Materials and Methods*). The *lacUV5* promoter is a variant of the *E. coli lac* promoter. The *T7A1lacO34* promoter is a variant of the phage T7A1 promoter (see *Expanded Materials and Methods*). In lanes 13-16, the template contains the RNA I promoter inserted upstream of the *rrnB* T1 terminator, so that RNA I is transcribed from both its native location in the plasmid and also from the site of fragment insertion.

```
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_2.4.1_rrnA
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_2.4.1_rrnB
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_2.4.1_rrnC

CTTGCGGTGCCCCGAGAGAGCGCCTAGATAGCCGCTCCACG          Rsph_ATCC_17025_rrnA
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCTTCACCG          Rsph_ATCC_17025_rrnB
CTTGCAGACCGGGGCGACTGCCCCTAAAAACCGCTTCACCG          Rsph_ATCC_17025_rrnC
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17025_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnA
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnB
CTTGCGCGGGCCGTCTGCTCCCCGTAGAAACCGCTTCACCG          Rsph_ATCC_17029_rrnC
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnA
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_KD131_rrnC
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_WS8N_rrnA
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_WS8N_rrnB
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_WS8N_rrnC

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC           Rsph_MBTJL-8_rrnA
CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-8_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACC           Rsph_MBTJL-8_rrnC

CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-13_rrnA
CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-13_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACC           Rsph_MBTJL-13_rrnC
```

Fig. S3: *R. sphaeroides* rRNA promoter sequences. An A at position -7 (6th position of the -10 hexamer; highlighted in red) is present in each of the rRNA promoters (*rrnA, rrnB, rrnC*, and where present, *rrnD*) in six different *R. sphaeroide*s strains. Genome sequences were obtained from NCBI. Strain accession numbers for Rsph 2.4.1: GCA_003324715.1; Rsph_ATCC_17025: GCA_000016405.1; Rsph_ATCC_17029: GCA_000015985.1; Rsph_KD131: GCA_000021005.1; Rsph_WS8N: GCA_000212605.1; Rsph_MBTJL-8: GCA_001576595.1; Rsph_MBTJL-13: GCA_001685625.1.

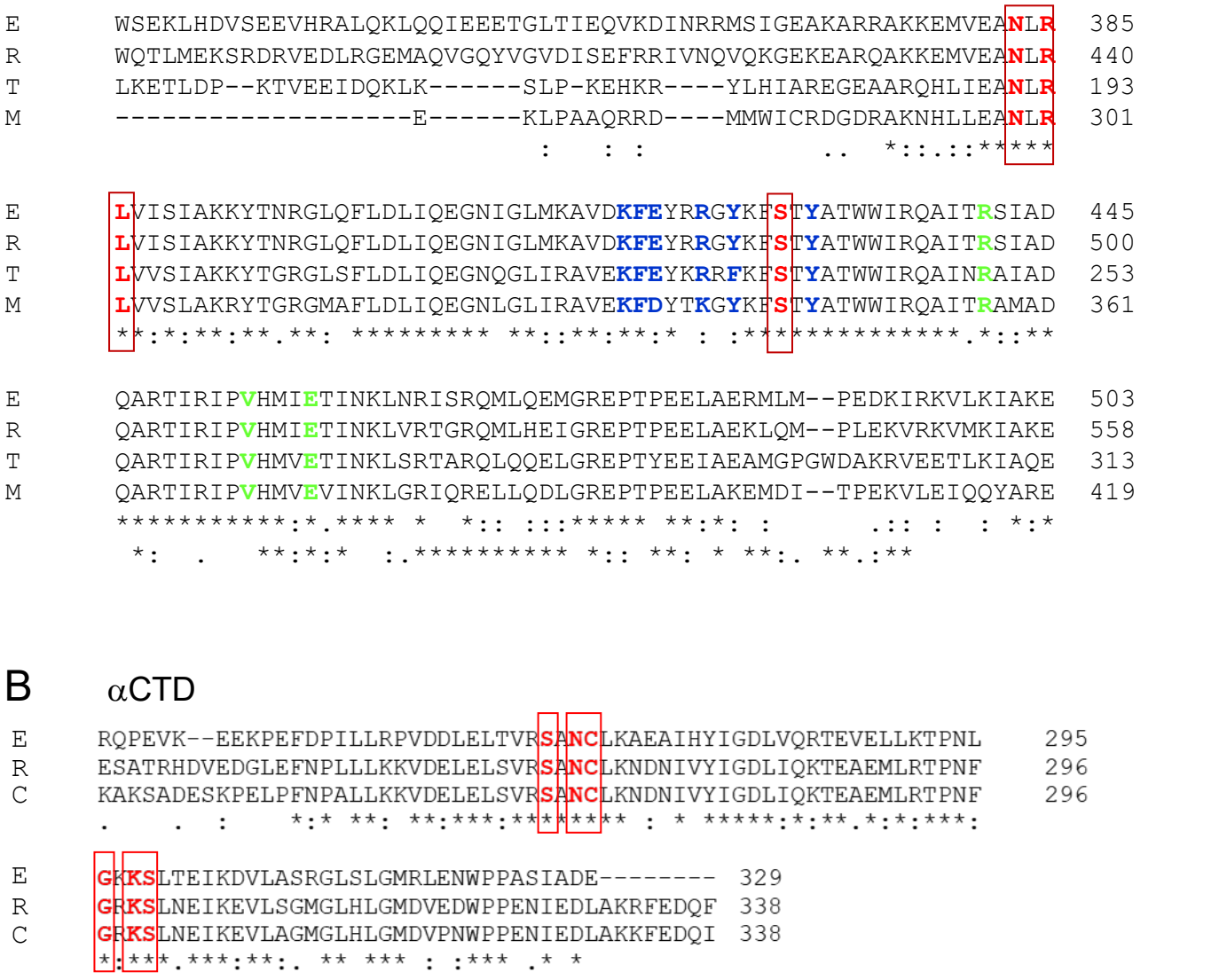


Fig. S4. **Major σ factor and αCTD alignments showing in (*A*) conservation of residues involved in promoter -10 recognition or in (*B*) residues involved in recognition of UP elements**. (*A*) Alignment of part of major σ factor amino acid sequences from *E. coli* (E), *R. sphaeroides* (R), *Thermus thermophilus* (T), and *Mycobacterium tuberculosis* (M). Conserved residues that contribute directly to the -7 pocket are in red and are boxed, the -11 pocket residues are in blue, and residues that interact with the extended -10 element are in green. Residue numbers are indicated at the end of each line. Alignments were made using Clustal Omega. The residues comprising the -7T binding pocket in *E. coli* σ^{70} include N383, R385, L386, and S428, which are conserved in *R. sphaeroides* σ^{93} (N499, R501, L502, S543) and in σ factors from species as evolutionarily diverse as *E. coli* (23, 24), *T. thermophilus* / *T.aquaticus* (25-27), and *M. smegmatis* / *M.tuberculosis* (28). (*B*) Alignment of the alpha C-terminal domains from *E. coli* (E), *R. sphaeroides* (R), and *C. crescentus* (C) RNAP, showing DNA binding residues for UP element recognition (red and boxed) (29).

A

```
Tth      -----MKEFRPGDKVVLPPYGVGVVAGIAQRSVSGVSRAYYQVDFPGSRSKAYVPVEAPH      55
Myx      ----MQTSFKTGDKAVYPGQGVGEVMGIEHTEVAGQRQSFYVLRILENGMRIMIPINKVG      56
Mtb      ------MIFKVGDTVVYPHHGAALVEAIETRTIKGEQKEYLVLKVAQGDLTVRVPAENAE      54
Rsp      MTKTKKPEFRPNEFVVYPAHGVGRIISIEEQEIAGIRLELFVISFEKDKMTLRVPTHKAT      60
Ccr      -MSKTGLEFSVGDHVVYPAHGVGSIQAIETQEVAGMSLEVYVITFDHEKMTLRVPTKKAK      59
                *  .: .* *  *.. : .*   : *       : .        :* .

Tth      SVGLRKALAPEEVPVILDLLKNGR-MPLPKQWAARHRKTSEILADGNPYRIAQMAGQLRA      114
Myx      SVGLREIISEEDVKQVYSILKEKDISVDSTTWNRRYREYMEKIKTGSVFEIAEVLRDLYL      116
Mtb      YVGVRDVVGQEGLDKVFQVLRAPH-TEEPTNWSRRYKANLEKLASGDVNKVAEVVRDLWR      113
Rsp      EVGMRSLSTPDVVTKALDTLKGKA-RVKRAMWSRRAQEYEQKINSGDLMSIAEVVRDLHR      119
Ccr      TAGLRPLAEGGTVSQALTTLKGRA-RVKRTMWSRRAQEYEAKINSGDLISIAEVVRDLHR      118
          .*:*       :      *:          *  * :     :  *.   :*::  :*

Tth      WEVERGLPDLDRQALRRAIHLLAEEVAQSLEITVQEAKRLFEEAWGEELN  164
Myx      LKGDKDLSFGERKMLDTARSLLIKELSLAKDCSEDEIESDLKKIFNLA--  164
Mtb      RDQERGLSAGEKRMLAKARQILVGELALAESTDDAKAETILDEVLAAAS-  162
Rsp      TDDQREQSYSERQLYEAALERLTREVAAVSGVDEAGAQKAVDAVLVSRAA  169
Ccr      AENQPEQSYSERQLYESALDRMAREVAAIERIDREAAIGILTKSLVKAA-  167
```

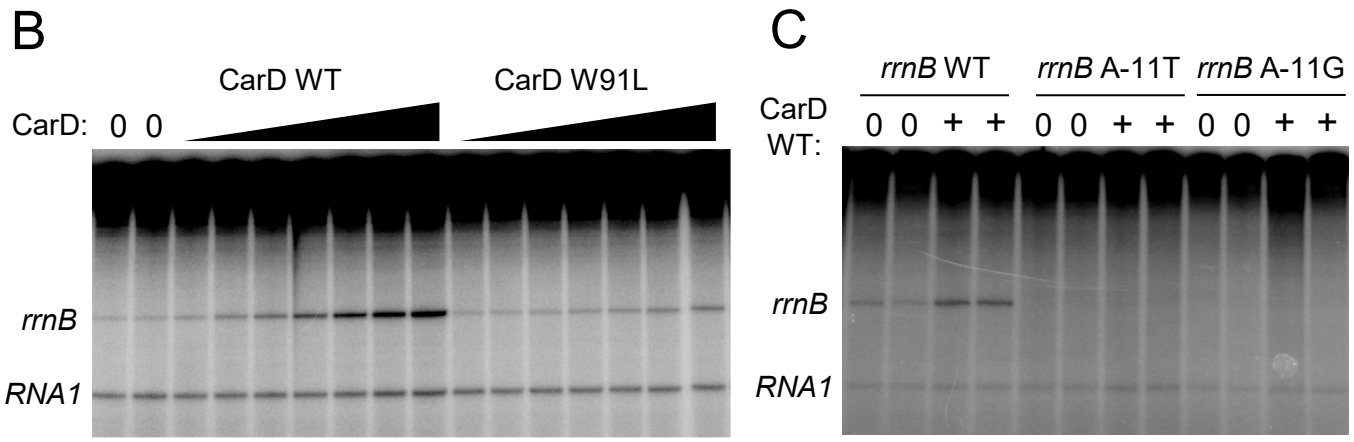


Fig. S5. Alignments of CarD, activation by wild-type or W91L mutant $CarD_{Rsp}$, and transcription from *R. sphaeroides rrnB* -11 mutant promoters. (A) Alignments of CarD from *Thermus thermophilus* HB8, *Myxococcus xanthus* DK1622 (also referred to as CdnL), *Mycobacterium tuberculosis* H37Rv, Rhodobacter sphaeroides 2.4.1 and *Caulobacter crescentus* NA1000 (CdnL). Substitutions introduced into $CarD_{Rsp}$ are boxed (Q31A/I33A/R53A; W91A; W91L). W91 is conserved among the CarD homologs (4). (B) Activation of $rrnB_{Rsp}$ transcription by 20 nM $RNAP_{Rsp}$ with 20 nM to 1280 nM $CarD_{Rsp}$ WT or $CarD_{Rsp}$ W91L. (C) In vitro transcription of the *rrnB* wild type, A-11T or A-11G mutant promoters by 20 nM $RNAP_{Rsp}$ ± 1280 nM $CarD_{Rsp}$. Buffer in (*B*) and (*C*) contained 170 mM NaCl.

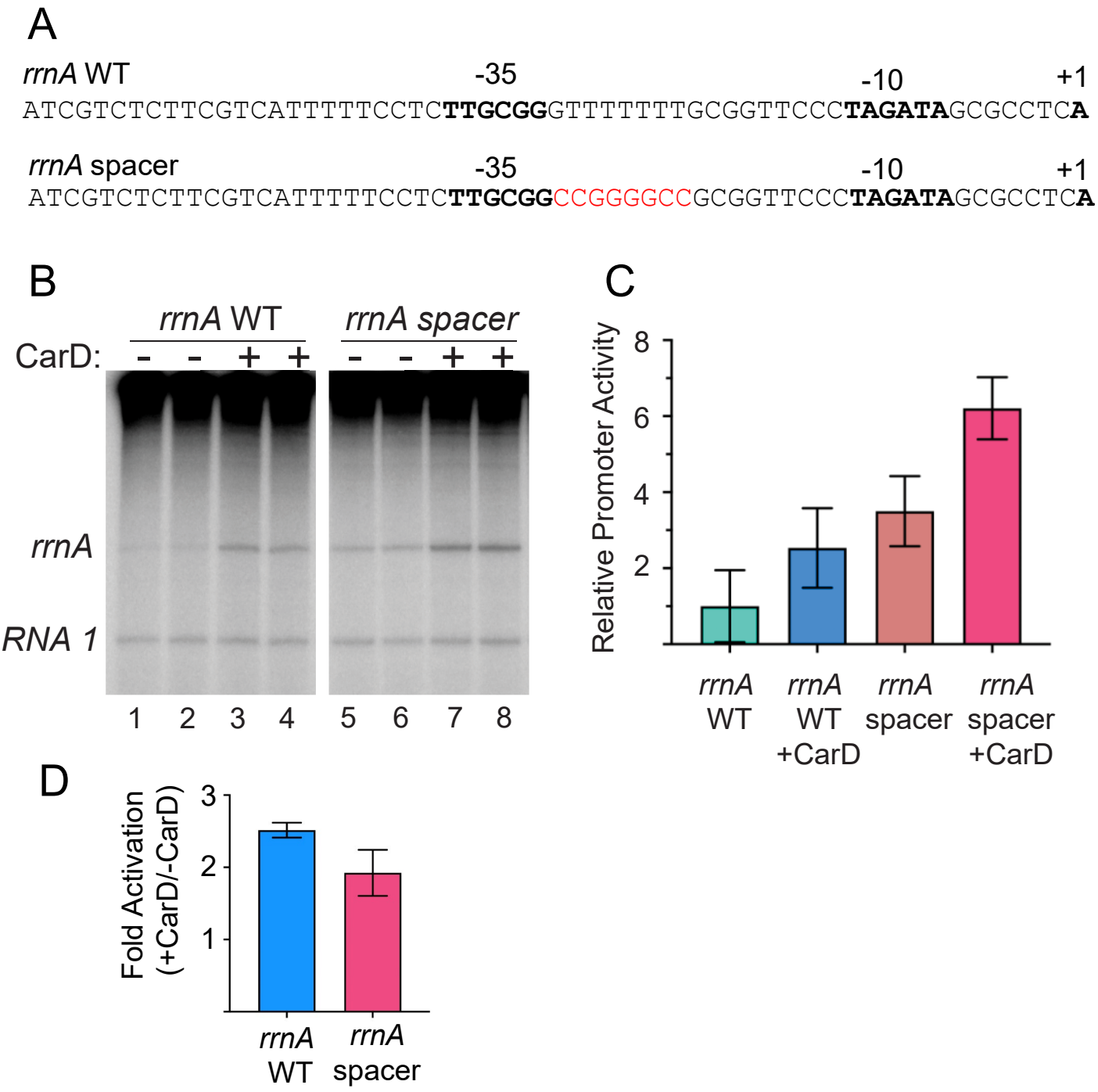
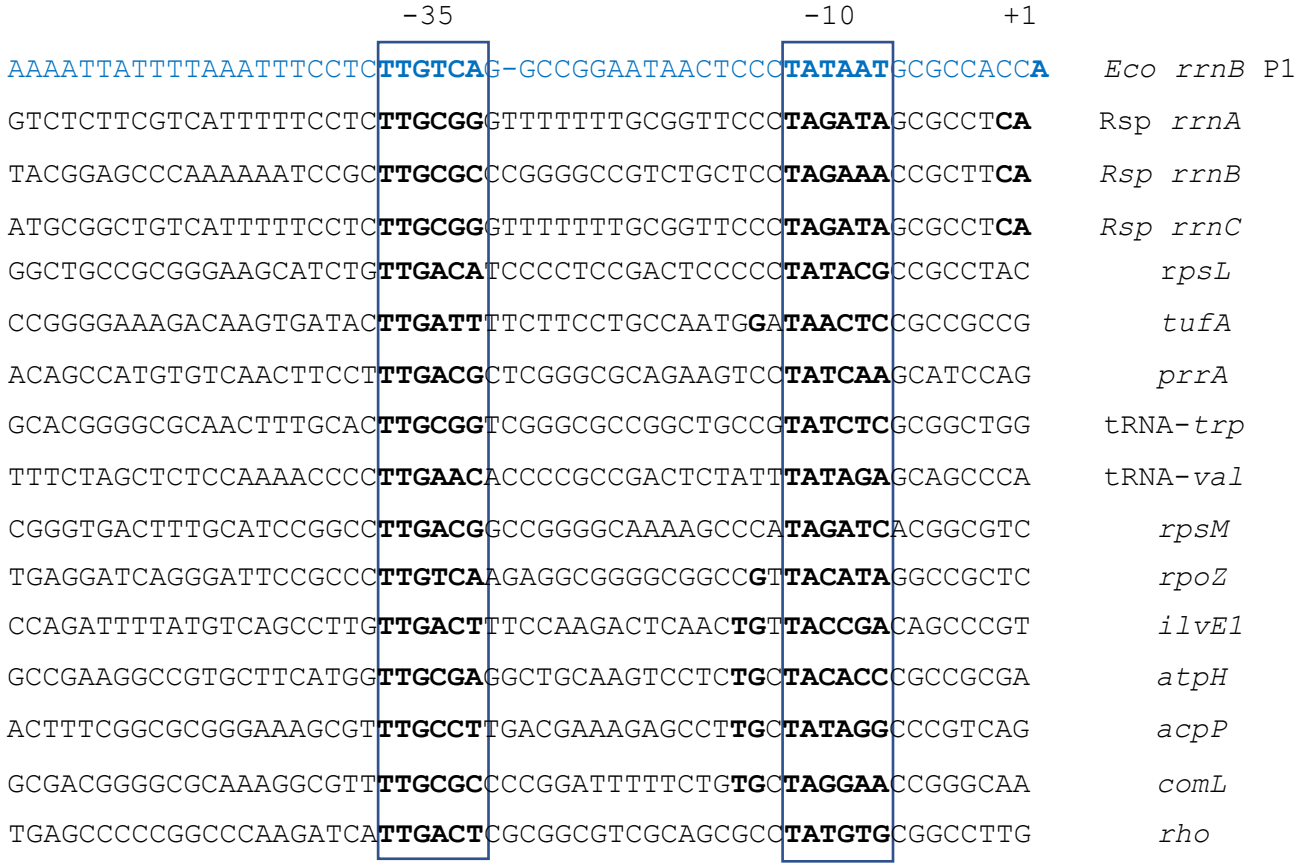


Fig. S6. In vitro transcription from *R. sphaeroides rrnA* wild-type and spacer mutant promoters. (*A*) Sequence of the WT *rrnA* promoter (-59 to +1), or spacer mutant with the sequence in red (CCGGGGCC) derived from the corresponding region of the *R. sphaeroides rrnB* promoter replacing the sequence GTTTTTTT. (*B*) Representative gel showing transcripts produced with 20 nM $RNAP_{Rsp}$ and1280 nM CarD where indicated in buffer with 170mM NaCl. The RNA I transcript is from a plasmid-encoded promoter (See Expanded Materials and Methods). (*C*) Quantitation of *rrnA* transcripts from gels like that in (B). Values were normalized to that of the wild type promoter in the absence of CarD. Error bars indicate range from two assays. (*D*) Quantitation of fold activation by CarD for either wild-type *rrnA* (lanes 3,4 vs lanes 1,2) or the *rrnA* spacer mutant (lanes 7,8 vs lanes 5,6) from two assays like that in (*B*).

A. Promoters Activated by CarD



B. Promoters Not Activated by CarD

Fig. S7. *R. sphaeroides* promoter sequences. (*A*) Promoter sequences (~-60 to +1) from 15 CarD-activated *R. sphaeroides* promoters and *E. coli rrnB* P1, which is activated by CarD when transcribed by *R. sphaeroides* RNAP but not by *E. coli* RNAP (see also Fig. S2). (*B*) Promoter sequences of 4 *R. sphaeroides* promoters and 3 *E. coli* promoters (in blue) whose activities are not altered by CarD. Positions corresponding to the -35 and -10 recognition motifs are in bold and boxed. For promoters with 18 bp spacing between the -35 and -10, the sequences are aligned with the -10 hexamer resulting in a shift of the -35 hexamer 1 bp further upstream.

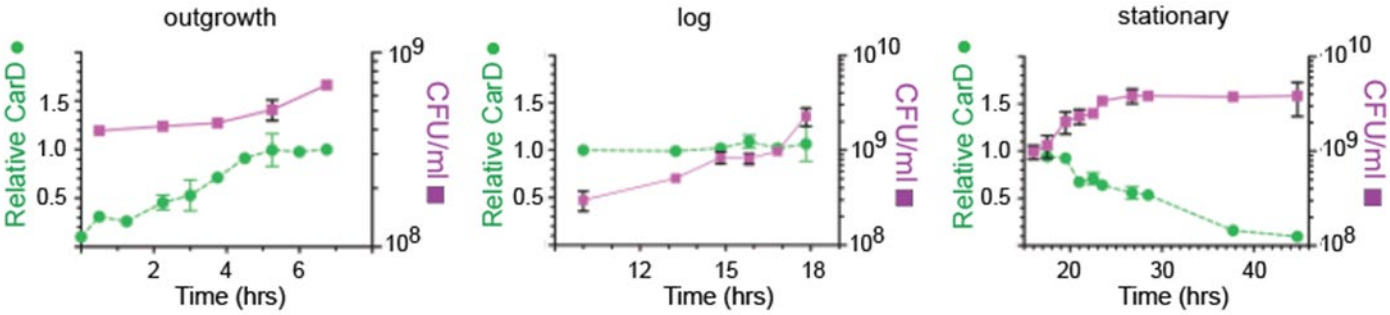


Fig. S8. Relative levels of $CarD_{Rsp}$ and CFU/ml in *R. sphaeroides* cultures during aerobic growth. CarD levels and CFU/ml were determined in samples from outgrowth, log phase, and stationary phase cultures and plotted as a function of time after inoculation of the culture. Data in each panel represent average and range from 2 experiments. Outgrowth, log, and stationary phase samples were obtained in separate experiments. Representative Western blots and OD_{600} values for these experiments are shown in Fig. 7*A*.

4. Legend for Dataset 1

The spreadsheet contains the chromosome location (A), locus tag (B), gene name if known (C), proposed product (D), position of the transcription start site (E), whether the start site is on the + or – strand (F), promoter sequence from -20 to +1 corresponding to the TSS (G), the sequence of the predicted extended -10 and -10 hexamer (H), and the identity of the base at position -7 (I). Detailed information about the bioinformatic analysis is in Expanded Materials and Methods.

5. SI References

1. W.R. Sistrom, A requirement for sodium in the growth of *Rhodopseudomonas spheroides*, *J Gen Microbiol* **22**:778–785 (1960).

2. A. Schäfer *et al*, Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum, *Gene* **145**:69–73 (1994).

3. R.R. Burgess, J.J. Jendrisak, Procedure for the rapid, large-scale purification of Escherichia coli DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**:4634–4638 (1975).

4. D.B. Srivastava *et al*, Structure and function of CarD, an essential mycobacterial transcription factor, *Proc Natl Acad Sci USA* **110**:12619-12624 (2013).

5. W. Ross, J.F. Thompson, J.T. Newlands, R.L. Gourse, *E.coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo, *EMBO J* **9**:3733–3742 (1990).

6. Davison J., Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids--a review. *Gene* **28**:1-15 (1984).

7. A.C. Ind *et al*, Inducible-Expression Plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*, *Appl Environ Microbiol* **75**:6613–6615 (2009).

8. R. Lutz H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**:1203-1210 (1997).

9. K. Lemmer *et al*, The NtrYX two-component system regulates the gram-negative cell envelope, *mBio* **11**:e00957-20.

10. J.M. Vera *et al*, Genome-scale transcription–translation mapping reveals features of *Zymomonas mobilis* transcription units and promoters, *mSystems* **5**:e00250-20 (2020).

11. K.S. Myers *et al*, Genome-wide Identification of Transcription Start Sites for *Rhodobacter sphaeroides* Grown Aerobically and Photosynthetically, *ASM Microbiology Resource Announcements* (2020), in press.

12. T.L. Bailey *et al*, MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**:W202-8 (2009).

13. C. Mackenzie *et al*, The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. *Photosynth Res* **70**:19-41 (2001).

14. B. Zhou *et al*, The global regulatory architecture of transcription during the *Caulobacter* cell cycle, *PLoS Genet* **11**:e1004831 (2015).

15. M.K. Thomason *et al*, Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*, *J Bacteriol* **197**:18–28 (2015).

16. E. Warman, D. Forrest, J.T. Wade, D.C. Grainger, Widespread divergent transcription from prokaryotic promoters, doi.org/10.1101/2020.01.31.928960doi: bioRxiv preprint first posted online Feb. 2, 2020.

17. X. Li *et al*, Transcriptome landscape of *Mycobacterium smegmatis*, *Front Microbiol* **8**:2505 (2017).

18. Y. Jeong *et al*, The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2), *Nat Commun* **7**:11605 (2016).

19. A.M. Sass *et al*, Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315, *BMC Genomics* **16**:775 (2015).

20. R. Simon, U. Priefer, A. Pühler, A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria, *Nat Biotech* **1**:784–791 (1983).

21. W. Ross, S.E. Aiyar, J. Salomon, R.L. Gourse, *Escherichia coli* Promoters with UP elements of different strengths: modular structure of bacterial promoters, *J Bacteriol* **180**:5375–5383 (1998).

22. T. Gaal, M.S. Bartlett, W. Ross, C.L. Turnbough, R.L. Gourse, Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria, *Science* **278**: 2092–2097 (1997).

23. Y. Zuo, T.A. Steitz, Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Mol Cell* **58**:534-540 (2015).

24. A. Narayanan *et al*, Cryo-EM structure of *Escherichia coli* σ^{70} RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation, *J Biol Chem* **293**:7367-7375 (2018).

25. A. Feklistov, S.A. Darst, Structural basis for promoter -10 element recognition by the bacterial RNA polymerase σ subunit, *Cell* **147**:1257-1269 (2011).

26. Y. Zhang *et al*, Structural basis of transcription initiation, *Science* **338**:1076-1080 (2012).

27. B. Bae, A. Feklistov, A. Lass-Napiorkowska, R. Landick, S.A. Darst, Structure of a bacterial RNA polymerase holoenzyme open promoter complex, *eLife* **4**:e08504 (2015).

28. E.A. Hubin, M. Lilic, S.A. Darst, E.A. Campbell, Structural insights into the mycobacteria transcription initiation complex from analysis of X-ray crystal structures, *Nat Commun* **8**:16072 (2017).

29. B. Benoff *et al*, H. Yang, C.L. Lawson, G. Parkinson, J. Liu, E. Blatter, Y.W. Ebright, H.M. Berman, R.H. Ebright, Structural basis of transcription activation: the CAP-alpha CTD-DNA complex, *Science* **297**:1562-1566 (2002).