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

Kemardo K. Henry[a], Wilma Ross[a,1], Kevin S. Myers[b], Kimberly C. Lemmer[b], Jessica M. Vera[b], Robert Landick[a,b,c], Timothy J. Donohue[a,b], and Richard L. Gourse[a,1]

[a]Department of Bacteriology, University of Wisconsin–Madison, Madison, WI 53706; [b]Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin–Madison, Madison, WI 53726; and [c]Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706

Edited by Lucia B. Rothman-Denes, The University of Chicago, Chicago, IL, and approved September 25, 2020 (received for review May 19, 2020)

Using an in vitro transcription system with purified RNA polymerase (RNAP) to investigate rRNA synthesis in the photoheterotrophic α-proteobacterium *Rhodobacter sphaeroides*, we identified a surprising feature of promoters recognized by the major holoenzyme. Transcription from *R. sphaeroides* rRNA promoters was unexpectedly weak, correlating with absence of −7T, the very highly conserved thymine found at the last position in −10 elements of promoters in most bacterial species. Thymine substitutions for adenine at position −7 in the three rRNA promoters strongly increased intrinsic promoter activity, indicating that *R. sphaeroides* RNAP can utilize −7T when present. rRNA promoters were activated by purified *R. sphaeroides* CarD, a transcription factor found in many bacterial species but not in β- and γ-proteobacteria. Overall, CarD increased the activity of 15 of 16 native *R. sphaeroides* promoters tested in vitro that lacked −7T, whereas it had no effect on three of the four native promoters that contained −7T. Genome-wide bioinformatic analysis of promoters from *R. sphaeroides* and two other α-proteobacterial species indicated that 30 to 43% contained −7T, whereas 90 to 99% of promoters from non–α-proteobacteria contained −7T. Thus, promoters lacking −7T appear to be widespread in α-proteobacteria and may have evolved away from consensus to enable their coordinated regulation by transcription factors like CarD. We observed a strong reduction in *R. sphaeroides* CarD levels when cells enter stationary phase, suggesting that reduced activation by CarD may contribute to inhibition of rRNA transcription when cells enter stationary phase, the stage of growth when bacterial ribosome synthesis declines.

CarD | RNA polymerase | −10 element | *Rhodobacter sphaeroides* | promoter

The α-proteobacteria are a Gram-negative, metabolically diverse, biotechnologically important class of bacteria. Although regulation of transcription has been well characterized in the γ-proteobacterium *Escherichia coli*, much less is known about transcription and its regulation in α-proteobacteria or in most other bacterial classes and phyla. *Rhodobacter sphaeroides* is a purple nonsulfur α-proteobacterium capable of aerobic and anaerobic respiration, photosynthesis, and CO_2 and N_2, as well as H_2 and polyhydroxybutyrate, production (1, 2). Aerobic growth is similar to that of other Gram-negative bacteria. However, when oxygen levels decline, *R. sphaeroides* switches to photosynthetic growth, completely remodeling its intracellular membrane and creating pigments necessary to capture light energy (3). In addition to its use as a photosynthetic α-proteobacterial model system, *R. sphaeroides* has also been studied extensively for its potential in bio-based production of fuels and chemicals (1, 4, 5). A better understanding of its basic transcription properties would improve the usefulness of *R. sphaeroides* for bioproduction as well as our understanding of α-proteobacterial gene regulation in general.

In *E. coli*, the synthesis of the translation machinery is tightly regulated at the level of rRNA transcription initiation in order to ensure a sufficient number of ribosomes to support the cellular growth rate. Study of rRNA transcription from the promoters for the seven *E. coli* rRNA operons has led to many general insights about transcription and its regulation, including the discovery that the nucleoid protein Fis is an important transcription factor (6), that the α-subunit of RNAP is a DNA-binding protein that contributes to specific promoter recognition (7), and that DksA is a transcription factor that functions in conjunction with the second messenger ppGpp to regulate large numbers of bacterial promoters (8, 9). Therefore, we focused on studying the properties of *R. sphaeroides* rRNA promoters to obtain insights about the mechanism of transcription initiation and its control in α-proteobacteria.

In contrast to *E. coli*, *R. sphaeroides* 2.4.1 has two chromosomes and only three rRNA operons, *rrnA* on chromosome 1 and *rrnB* and *rrnC* on chromosome 2. The transcription start sites for the *rrnA*, *rrnB*, and *rrnC* promoters were mapped previously by primer extension (10, 11). In several bacterial species, including *Mycobacterium tuberculosis*, *Thermus thermophilus*, and *Myxococcus xanthus*, rRNA promoters are activated in vitro by CarD, a small protein that binds to the lobe of the RNAP β-subunit and interacts with promoter DNA just upstream of the −10 hexamer (12–16). In the α-proteobacterium *Caulobacter crescentus*, the

Significance

Bacterial promoters are often predicted by similarity to the *Escherichia coli* −10 and −35 consensus elements. Although these elements are highly conserved in diverse bacterial phyla, only 30 to 43% of promoters we analyzed from *Rhodobacter sphaeroides* and two other α-proteobacteria contained −7T, a base in the −10 element present in 90 to 99% of promoters from non–α-proteobacteria. Expression from 15 of 16 *R. sphaeroides* promoters tested in vitro that lacked −7T was very weak, but we identified an essential transcription factor, CarD, that activated all 15 of these promoters. We suggest that promoters lacking a critical base in a consensus element can shape the transcriptome by coordinating expression of large numbers of genes by a single transcription factor.

Author contributions: K.K.H., W.R., K.S.M., R.L., T.J.D., and R.L.G. designed research; K.K.H., K.S.M., and J.M.V. performed research; K.C.L. contributed new reagents/analytic tools; K.K.H., W.R., K.S.M., T.J.D., and R.L.G. analyzed data; and K.K.H., W.R., K.S.M., R.L., T.J.D., and R.L.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: weross@wisc.edu or rgourse@bact.wisc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2010087117/-/DCSupplemental.

First published November 9, 2020.

CarD homolog localizes to rRNA promoters, and depletion decreases rRNA transcription (17). (The CarD homologs in *M. xanthus* and *C. crescentus* are sometimes referred to as CdnL.)

CarD family members are found in the *Actinomycetes, Firmicutes*, *Deinococcus/Thermus*, *Spirochaetes*, δ-proteobacteria, and most classes of α-proteobacteria, but are not found in β- and γ-proteobacteria (13, 14, 17). Although *carD* is essential in some species, and its depletion or deletion affects expression of many genes in *M. tuberculosis* and *C. crescentus* (12, 17–19), a direct, widespread role for CarD in transcription from non-rRNA promoters has not been demonstrated previously with purified components in vitro.

We report here that *R. sphaeroides* rRNA promoters are activated by $CarD_{Rsp}$ in vitro, at least in part because these promoters lack a critical thymine at the last position on the nontemplate strand of the −10 hexamer (−7T) that is present in almost all *E. coli* promoters recognized by the major holoenzyme. The T at −7 in *E. coli* promoters fits tightly in a pocket in σ^{70}, an interaction that is critical for transcription initiation (20). Not only do *R. sphaeroides* rRNA promoters utilize $CarD_{Rsp}$ to compensate, at least in part, for the absence of this −7T interaction with σ, but we also show here that many other *R. sphaeroides* promoters are activated by $CarD_{Rsp}$. In fact, bioinformatic analyses indicate that the majority of promoters in *R. sphaeroides*, as well as in two other α-proteobacterial species analyzed, lack a T at the last position in the −10 element, suggesting that many promoters in α-proteobacteria may also utilize CarD to increase transcription initiation.

Results

***R. sphaeroides* rRNA Promoters Initiate Poorly In Vitro with *R. sphaeroides* RNAP.** To purify *R. sphaeroides* RNAP for analysis of regulation of *R. sphaeroides* rRNA promoters in vitro, we inserted sequences coding for a C-terminal 10× histidine tag into the *rpoC* gene in the *R. sphaeroides* chromosome by homologous recombination, then purified *R. sphaeroides* RNAP ($RNAP_{Rsp}$) by nickel-affinity chromatography (*SI Appendix*, Fig. S1*A* and *Expanded Materials and Methods*). The resulting purified $RNAP_{Rsp}$ preparation contained proteins of the sizes expected for the β, β′, α, and ω subunits, as well as the major σ subunit, σ^{93}. The identity of σ^{93} was confirmed by Western blotting with a polyclonal anti-*E. coli* σ^{70} antibody that cross-reacted with *R. sphaeroides* σ^{93} (*SI Appendix*, Fig. S1*B*). Unless otherwise indicated, we use "$RNAP_{Rsp}$" to refer to the holoenzyme containing σ^{93}.

In vitro transcription with the three *R. sphaeroides* rRNA promoters was carried out with supercoiled plasmids containing rRNA promoter fragments (nontemplate strand sequences shown from −57 to +1 with respect to the transcription start sites in Fig. 1*A*) inserted into plasmid pRLG770 (6). Transcript lengths of ∼200 nt were predicted based on termination at the *E. coli rrnB* T1 terminator sequence downstream of the promoter fragment insertion site in the plasmid. Few transcripts of this size were detected from the *R. sphaeroides rrnA*, *rrnB*, *and rrnC* promoters with $RNAP_{Rsp}$ (Fig. 1*B*, lanes 1 to 6) or $RNAP_{Eco}$ (Fig. 1*C*, lanes 1 to 6) under conditions where transcription was detected from the RNA I promoter, part of the replication control system encoded by the plasmid templates (Fig. 1 *B* and *C*). In contrast, an abundant transcript of the expected size was detected in reactions containing the *E. coli rrnB* P1 promoter and either *R. sphaeroides or E. coli* RNAP (Fig. 1 *B* and *C*, lanes 7 and 8). $RNAP_{Rsp}$ also produced transcripts from two mutant *E. coli* or phage promoters that were tested, *lacUV5* and T7A1*lacO34* (*SI Appendix*, Fig. S2), consistent with the conclusion that the lack of transcription from the *R. sphaeroides* rRNA promoters was not a result of inactivity of the purified $RNAP_{Rsp}$ holoenzyme.

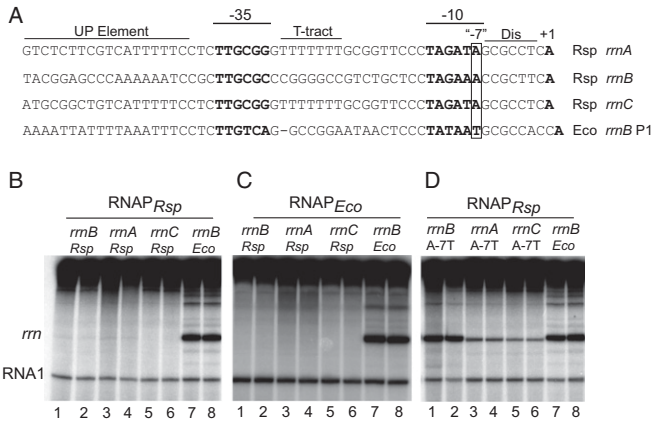

Fig. 1. Activity of *R. sphaeroides* rRNA promoters with either *R. sphaeroides* RNAP or *E. coli* RNAP. (*A*) Sequences of the three *R. sphaeroides* rRNA promoters, *rrnA*, *rrnB*, and *rrnC*, from −57 to the transcription start site, +1, and the *E. coli rrnB* P1 promoter. The discriminator (Dis) region, T-tract sequence, and UP element are indicated, in addition to the −10 and −35 RNAP recognition hexamers and the transcription start site, which are in bold. The last bp in the −10 element is referred to as the "−7" position (boxed), although it is 8 or 9 bp rather than 7 bp upstream from the transcription start site due to the larger-than-consensus number of bp between the −10 element and the TSS in these promoters. (*B*) In vitro transcription of the indicated rRNA promoters from plasmid templates with *R. sphaeroides* RNAP in buffer containing 170 mM NaCl (*SI Appendix, Expanded Materials and Methods*). Duplicate lanes are shown for each promoter. The RNA I promoter and transcript are part of the plasmid replication control system (*SI Appendix, Expanded Materials and Methods*). (*C*) In vitro transcription of the indicated promoters as in *B*, but with *E. coli* RNAP. (*D*) In vitro transcription with *R. sphaeroides* RNAP as in *B*, but with the A-7T promoter variants of the three *R. sphaeroides* rRNA promoters or with the wild type *E. coli rrnB* P1 promoter. A higher concentration of *R. sphaeroides* RNAP (50 nM) than *E. coli* RNAP (10 nM) was used to ensure that the absence of transcription from the *R. sphaeroides* rRNA promoters was not a result of limiting $RNAP_{Rsp}$. Robust, approximately equivalent transcription was observed from the *E. coli rrnB* promoter at 50 nM $RNAP_{Rsp}$ and 10 nM $RNAP_{Eco}$ (*B–D*).

The inactivity of the *R. sphaeroides* rRNA promoters was unexpected given that rRNA promoters are very active in moderate- to fast-growing bacterial cells (11, 21). In addition, the *R. sphaeroides* rRNA promoters contain several features characteristic of *E. coli* rRNA promoters (11), including a TTG sequence (nontemplate strand) at the upstream end of a putative −35 element, a TA sequence at the upstream end of a likely −10 element, an A+T-rich UP element-like sequence upstream of the −35 element (22), and a G+C-rich sequence immediately downstream from the −10 element, corresponding to the discriminator sequence in *E. coli* rRNA promoters (23) (Fig. 1*A*). The high A+T content of the UP element-like sequence and the 7-nt T-tract in the spacer region of *rrnA* and *rrnC* were especially striking given the ∼70% G+C content of the *R. sphaeroides* genome (1).

In contrast to *E. coli* rRNA promoters (and most other $E\sigma^{70}$-dependent *E. coli* promoters), each of the *R. sphaeroides* rRNA promoters contains an A rather than the highly conserved T at the last position in the −10 element (designated here as position −7; Fig. 1*A*). In fact, the predicted rRNA promoters from other sequenced *R. sphaeroides* strains also lack a T at −7 (*SI Appendix*, Fig. S3). The significance of the conserved T at −7 for $E\sigma^{70}$-dependent promoter activity in *E. coli* was demonstrated long ago by genetic analyses and by structural studies showing that only the thymine, and not other bases at position −7, is compatible with binding in a pocket in σ^{70} (20, 24–26). The σ-residues that contribute to the −7 pocket are conserved in the major $RNAP_{Rsp}$, consistent with the ability of $RNAP_{Rsp}$ to

recognize *E. coli* promoters (*SI Appendix*, Fig. S4*A*). Also consistent with the absence of −7T as a determinant of the low activity of the *R. sphaeroides* rRNA promoters in vitro, we found that a −7T substitution increased transcription from each of the rRNA promoters with $RNAP_{Rsp}$ (Fig. 1*D*), although the mutated *rrnA*(A–7T) and *rrnC*(A–7T) promoters were much weaker than *rrnB*(A–7T) or *E. coli rrnB* P1. Taken together, these results suggested that *R. sphaeroides* rRNA promoters might require a transcription activator for function.

***R. sphaeroides* CarD Activates rRNA Promoters.** Previous studies have reported that the CarD protein from *M. tuberculosis*, Mycobacterium *smegmatis*, *T. thermophilus*, and *M. xanthus* stimulates rRNA transcription in vitro (12, 13, 27, 28). Among the proteobacteria, CarD is found in the α and δ, but not in the β and γ classes (i.e., not in *E. coli*) (13, 14). The *R. sphaeroides* CarD homolog (*rsp_2425*) is similar in size (169 amino acids), is 25 to 60% identical to CarD proteins from the other bacteria (*SI Appendix*, Fig. S5), and is essential (5), consistent with a role in one or more critical functions including rRNA transcription.

To test this idea, we purified *R. sphaeroides* CarD after overexpression in *E. coli* and measured its effects on *R. sphaeroides* rRNA promoters in vitro (Fig. 2*A*). $CarD_{Rsp}$ activated transcription from the *R. sphaeroides rrnB* promoter at least 10-fold in vitro, with a half-maximal stimulatory effect (EC_{50}) at 85 nM (Fig. 2*B*). When added to $RNAP_{Rsp}$, $CarD_{Rsp}$ also activated the *R. sphaeroides rrnA* and *rrnC* promoters, although transcription from *rrnA* and *rrnC* was weaker than from *rrnB*, and the fold activation by $CarD_{Rsp}$ was smaller (Fig. 2 *C* and *D*). These results are consistent with a previous report that the *R. sphaeroides rrnB* promoter was stronger than the *rrnA* promoter in vivo based on measurements of promoter–*xylE* fusions (11).

To determine whether the 7-nt T-tract found in the spacer region of *R. sphaeroides rrnA* (and *rrnC*, which is identical to *rrnA*), but not in *rrnB*, could account for the promoter activity difference, we constructed a derivative of *rrnA* containing a replacement of the T-tract with the corresponding spacer sequence of *rrnB*. Replacement of the T-tract resulted in a threefold increase in basal *rrnA* promoter activity but did not alter the extent of activation by CarD (two- to threefold; Fig. 2 and *SI Appendix*, Fig. S6). Thus, the T-tract accounts, in part, for the reduced activities of *rrnA* and *rrnC* but does not account for the difference in activation by CarD (2- to 3-fold for *rrnA* vs. 10- to 12-fold for *rrnB*).

$CarD_{Rsp}$ did not increase transcription from the plasmid-encoded RNA I promoter when added to $RNAP_{Rsp}$ (Fig. 2 *A*–*C*) or from *R. sphaeroides rrnB* using *E. coli* RNAP (Fig. 2*C*). The failure of $CarD_{Rsp}$ to stimulate *E. coli* RNAP is consistent with previous reports that effects of CarD are species-specific because the interacting residues in the CarD RID domain and in the β-subunit of RNAP are not highly conserved (28, 29).

We note that a low level of *R. sphaeroides rrnB* transcript was detected in a previous study with *R. sphaeroides* RNAP (30). Based on the results reported here, we suggest that the observed activity may have resulted from the presence of small amounts of CarD in the RNAP used in the previous study. In fact, using an antibody raised against $CarD_{Rsp}$ (see below), we detected a very low level of CarD in our purified *R. sphaeroides* RNAP (~0.01 mol $CarD_{Rsp}$ per mol RNAP). In our in vitro transcription experiments, the concentration of CarD introduced from $RNAP_{Rsp}$ was well below the EC_{50} determined for CarD (85 nM) but perhaps sufficient to account for the previously reported transcription from *rrnB* (*SI Appendix*, Fig. S1*C*, lanes 7 and 8).

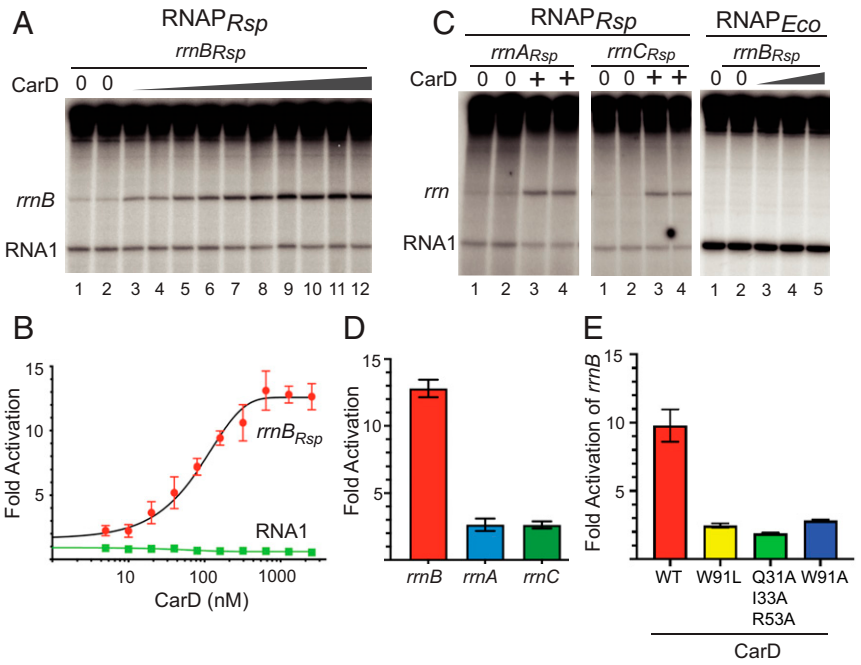

Fig. 2. Activities of the *R. sphaeroides* rRNA promoters in vitro with or without purified *R. sphaeroides* CarD. (*A*) In vitro transcription of the *R. sphaeroides rrnB* promoter with 20 nM *R. sphaeroides* RNAP with or without $CarD_{Rsp}$ (wedge indicates $CarD_{Rsp}$ range of 5 to 2,560 nM) in buffer with 170 mM NaCl. (*B*) Fold activation of the *R. sphaeroides rrnB* promoter by $CarD_{Rsp}$ (with/without the indicated concentration of $CarD_{Rsp}$) from experiments like that in *A*. The RNA I transcript is from a plasmid-encoded promoter. Error bars indicate SD from $n = 3$ separate assays. (*C*) In vitro transcription of the *R. sphaeroides rrnA* or *rrnC* promoters with *R. sphaeroides* RNAP as in *A* with or without 1,280 nM CarD or of the *R. sphaeroides rrnB* promoter with 10 nM *E. coli* RNAP with or without 20, 40, or 80 nM CarD. (*D*) Fold activation of the *R. sphaeroides rrnA, rrnB,* and *rrnC* promoters by 1,280 nM CarD with SD from $n = 3$ assays for each promoter. (*E*) Fold activation of the *R. sphaeroides rrnB* promoter by $RNAP_{Rsp}$ and 1,280 nM wild-type CarD or W91A, W91L, or Q31A/I33A/R53A triple-mutant CarD (SD from $n = 3$ assays). Wild-type CarD used in *B* and *E* was from different protein preparations, so differences in specific activities could account for the slight differences in fold activation.

To address whether the mechanism of activation by $CarD_{Rsp}$ is similar to that by $CarD_{Mtb}$, we created substitutions in $CarD_{Rsp}$ at positions analogous to those previously shown to affect the DNA interaction or RNAP interaction domain (RID) activities of $CarD_{Mtb}$ (13, 14). $CarD_{Mtb}$ interacts with DNA at the upstream edge of the −10 hexamer using residue W85 to help wedge open the −10 hexamer. Alignment of the amino acid sequences from *Rhodobacter*, *Mycobacterium*, *Thermus*, *Myxococcus*, and *Caulobacter* CarD homologs indicated that $CarD_{Rsp}$-W91 corresponds to $CarD_{Mtb}$-W85 (14, 31) (*SI Appendix*, Fig. S5). Substitutions for $CarD_{Rsp}$-W91 ($CarD_{Rsp}$-W91L or -W91A) greatly reduced activation of *R. sphaeroides rrnB* transcription, decreasing the 10-fold effect observed with wild-type $CarD_{Rsp}$ to only ∼2.5-fold (Fig. 2*E* and *SI Appendix*, Fig. S5), a partial defect similar to that observed for $CarD_{Mtb}$ W85A and $CarD_{Tth}$ W86A (14). Conserved basic residues in CarD also interact with the DNA near W86 in the *T. thermophilus* CarD-RNAP-promoter DNA complex structure (13), suggesting that the partial activity of the W91A or W91L $CarD_{Rsp}$ proteins can be attributed to DNA interactions that are weakened but not eliminated, rather than to general folding defects. A $CarD_{Rsp}$ variant with three substitutions, Q31A, I33A, R53A, corresponding to *T. thermophilus* CarD RID domain residues R25, I27, and R47, activated $rrnB_{Rsp}$ transcription only ∼twofold (Fig. 2*E* and *SI Appendix*, Fig. S5). We tentatively conclude that $CarD_{Rsp}$ uses a mechanism similar to that of previously characterized CarD proteins.

CarD Directly Activates Other *R. sphaeroides* Promoters. Depletion of CarD from *M. smegmatis* and *M. tuberculosis* in vivo (12, 18), or deletion of the *C. crescentus* homolog (19), suggested that CarD affects gene expression broadly, although the approaches did not distinguish direct from indirect effects. CarD colocalized with *M. smegmatis* RNAP in ChIP-seq experiments at all promoters recognized by the primary σ-factor in vivo (14, 32). Two *M. tuberculosis* promoters, the AP3 rRNA promoter and the VapB promoter, were also shown to be activated by CarD in vitro (16).

To evaluate whether CarD has a more general role in *R. sphaeroides* gene expression, we tested its direct effects on transcription in vitro from a set of previously uncharacterized *R. sphaeroides* promoters. Sixteen additional *R. sphaeroides* promoters were chosen based on several criteria: genome-wide ChIP signals for σ^{93} from an *R. sphaeroides* ChIP-seq dataset (3), proximity to mapped *R. sphaeroides* transcription start sites (TSSs) (33), and sequence similarity to elements of an *E. coli* $E\sigma^{70}$ promoter consensus (25). We note that 13 of the 16 non-rRNA promoters lacked −7T (*SI Appendix*, Fig. S7), even though they were not chosen based on that criterion. PCR-derived fragments containing sequences from ∼150 bp upstream to 30 to 50 bp downstream of the proposed TSSs were cloned into pRLG770, the same plasmid described above for measuring transcription from rRNA promoters, and promoter activities were determined with and without 1,280 nM purified $CarD_{Rsp}$.

The sizes of the transcripts produced in these assays relative to standards of known length confirmed the predicted locations of the promoters (Fig. 3 and *SI Appendix*, Fig. S7). As with the rRNA promoters, $CarD_{Rsp}$ increased the activities of 12 of the 16 non-rRNA promoters. The 12 non-rRNA promoters that were activated by $CarD_{Rsp}$ included two for ribosomal protein operons (*rpsM* and *rpsL*), two for tRNAs ($tRNA^{trp}$ and $tRNA^{val}$), *rpoZ* (ω-subunit of RNAP), *tufA* (translation factor EF-Tu), *prrA* (transcription factor PrrA), *rho* (transcription termination factor Rho), *ilvE1* (a putative aminotransferase), *acpP* (acyl carrier protein), *comL* (branched-chain amino acid BamD subunit), and *atpH* (ATP synthase F1 δ-subunit; Fig. 3*A*). For the promoters whose basal activities were high enough to quantify, activation by $CarD_{Rsp}$ ranged from two- to almost sixfold (Fig. 3*B*).

Four of the 16 non-rRNA promoters, including *ssrA* (tmRNA), *putR* (a putative transcription factor), *rsp_3110* (a putative GST), and *rsp_1486* (tetR family homolog), were unaffected by $CarD_{Rsp}$ (Fig. 3*C*). $CarD_{Rsp}$ increased transcription less than 1.4-fold from three promoters whose basal activities were high enough to quantify the fold effect (Fig. 3*D*). In addition, three of the four *E. coli* promoters tested with $RNAP_{Rsp}$ were unaffected by $CarD_{Rsp}$ (*lacUV5, T7A1lacO34*, RNA I; *SI Appendix*, Fig. S2). Thus, positive regulation by $CarD_{Rsp}$ is promoter-specific and not limited to rRNA promoters, and not all promoters are dependent on $CarD_{Rsp}$ for activity.

Promoter Features that Correlate with Increased Transcription by $CarD_{Rsp}$. An image depicting the degree of sequence conservation at each promoter position in the 15 *R. sphaeroides* promoters activated by $CarD_{Rsp}$ (the 12 above plus the 3 rRNA promoters) is shown in Fig. 3*E*, and promoter sequences are shown in *SI Appendix*, Fig. S7. None of the 15 promoters activated by $CarD_{Rsp}$ contained a T at −7. Each of the other three bases at this position (A, G, and C) was represented in the activated promoter population, suggesting that the absence of −7T rather than the presence of another specific base was the relevant determinant for activation by $CarD_{Rsp}$. In contrast, three of the four *R. sphaeroides* promoters unaffected by $CarD_{Rsp}$ contained a T at −7 (Fig. 3*C* and *SI Appendix*, Fig. S7).

To address whether sequence features in addition to the absence of −7T might also contribute to activation by $CarD_{Rsp}$, we constructed *R. sphaeroides rrnB* promoter variants with increased or decreased similarity to the *E. coli* consensus for the UP element, −35 hexamer, extended −10 element, −10 hexamer, discriminator region, and −10/−35 spacer length (35) (Fig. 4*A*). Transcription activities of the variant promoters were compared to the wild-type promoter in the absence and presence of 1,280 nM $CarD_{Rsp}$ in vitro (Fig. 4*B*).

We did not detect changes in basal promoter activity resulting from replacement of the *rrnB* UP element region with either a non-UP element-like sequence or with the *E. coli* UP element consensus because transcription was too weak for quantitation (Fig. 4 *B* and *C*). However, in the presence of $CarD_{Rsp}$, it was apparent that either the native UP element sequence or the consensus UP element increased transcription (Fig. 4 *B* and *D*), consistent with conservation of the DNA binding residues in the *R. sphaeroides* and *E. coli* α-subunit C-terminal domains (22) (*SI Appendix*, Fig. S4*B*).

Substitutions toward consensus for recognition by RNAP in the −35 hexamer, −10 hexamer, and discriminator region increased basal *R. sphaeroides rrnB* promoter activity (Fig. 4 *B* and *C*, compare lane 1 with lanes 7, 13, 16, 17, 19, and 21). Not surprisingly, since the wild-type *R. sphaeroides rrnB* promoter already has consensus −10/−35 spacing (17 bp), changing the spacing to either 16 bp or 18 bp did not increase basal promoter activity (Fig. 4 *B* and *C*, compare lane 2 with lane 10 or 12). Each of these *R. sphaeroides rrnB* promoter variants was activated by $CarD_{Rsp}$ (Fig. 4 *B* and *D*), suggesting that they altered transcription without fully bypassing the step(s) affected by $CarD_{Rsp}$. In contrast, the improved extended −10 element and the A–7T mutations increased basal promoter activity to a much greater extent than any of the other substitutions (Fig. 4 *B* and *C*, compare lane 1 with lanes 13 and 17), and these promoter variants were no longer activated at all by $CarD_{Rsp}$ (Fig. 4 *B* and *C*). We conclude that the absence of −7T or an extended −10 element can create a barrier to transcription initiation, and that $CarD_{Rsp}$ or mutations to consensus or both can alleviate this barrier.

Context Dependence of Promoter Sequences that Bypass the Effect of CarD on *rrnB*. We next addressed whether the −7T or extended −10 substitutions would bypass the requirement for CarD in three other contexts, the *rpsM*, *rpoZ*, and *rrnA* promoters

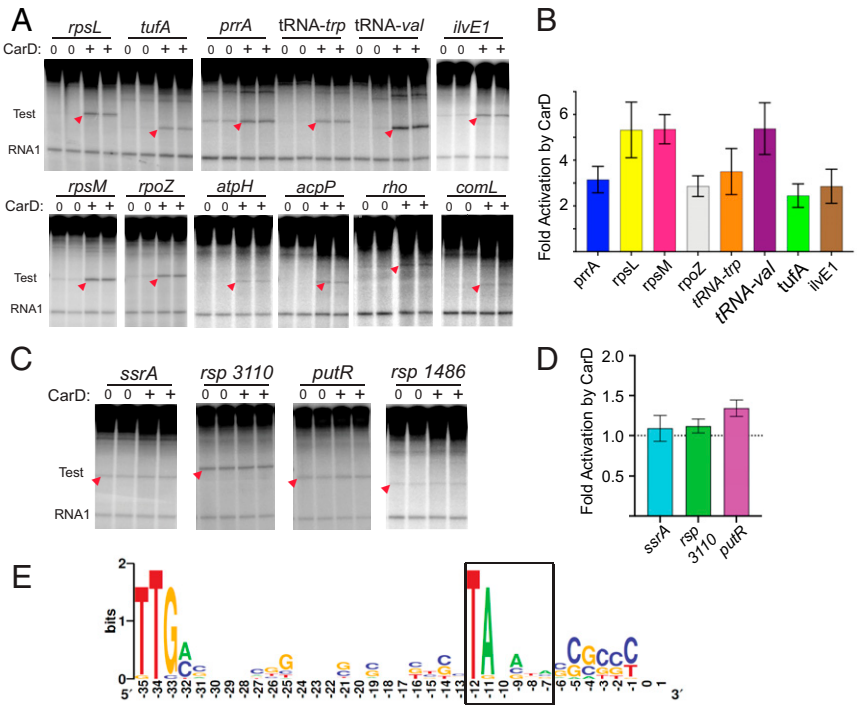

Fig. 3. Activation of *R. sphaeroides* non-rRNA promoters in vitro with $CarD_{Rsp}$. (*A*) Transcription in vitro with $RNAP_{Rsp}$ (20 nM) and $CarD_{Rsp}$ (1,280 nM) in buffer with 100 mM NaCl (*SI Appendix, Expanded Materials and Methods*). Duplicate lanes with or without CarD are shown for each promoter. Transcripts derived from test promoters are indicated with red arrowheads. Gel images also show position of RNA 1 transcript from the plasmid promoter. *SI Appendix*, Fig. S7, shows promoter sequences. (*B*) Fold activation by CarD (+CarD/no CarD) for promoters shown in *A* whose activities were high enough to quantify accurately. SDs are shown from $n = 3$ to 6 assays except for *ilvE1* (range from two assays). (*C*) Transcription in vitro of four *R. sphaeroides* promoters whose activities were unaffected by CarD under the same conditions as in *B*. *SI Appendix*, Fig. S7, shows promoter sequences. (*D*) Fold activation of promoters in *C* with or without CarD. The SDs shown are from $n = 5$ assays for *ssrA*, and the ranges are shown from $n = 2$ assays for *rsp_3110* and *putR*. (*E*) WebLogo (34) representation of consensus sequence for 15 promoters whose transcription was activated by CarD in vitro (3 rRNA promoters and the 12 promoters shown in *A*). The −10 hexamer (−12 to −7) is boxed.

(Fig. 5*A*). Both substitutions increased the basal activity of the *rpsM* promoter substantially, and although the substitutions reduced the extent of activation by $CarD_{Rsp}$ from ~fivefold to ~twofold (Fig. 5 *B* and *C*), they did not completely eliminate the effect of $CarD_{Rsp}$. As with the *rpsM* and *rrnB* promoters, both substitutions increased the basal activities of the *rrnA* and *rpoZ* promoters. However, unlike *rpsM* and *rrnB*, the *rpoZ* and *rrnA* mutant promoters containing the extended −10 element or −7T substitutions were still activated by $CarD_{Rsp}$ almost to the same extent as with the wild-type promoters (Fig. 4 *B*–*D*).

Thus, there are promoters lacking −7T (i.e., *rsp_3110*, above) that are not activated by $CarD_{Rsp}$, and there are mutant promoters containing −7T that are still activated, at least to some extent, by $CarD_{Rsp}$. These "context effects" indicate that $CarD_{Rsp}$ can compensate for rate-determining steps in promoter complex formation that derive from other promoter sequence features in addition to or instead of −7T. Such context effects have long been recognized in studies of effects of promoter substitutions on transcription (24, 36). The mechanistic explanations for the context effects reported here are described more extensively in the *Discussion*.

The Majority of *R. sphaeroides* Promoters Lack the T at the Last Position in the −10 Element. To identify likely −10 elements used by *R. sphaeroides* RNAP genome-wide, we analyzed available *R. sphaeroides* transcription start site (TSS) data obtained from cells in exponential growth where most transcription is likely by $E\sigma^{93}$, the major holoenzyme (33). This analysis was followed by a bioinformatic search for a TA sequence corresponding to the highly conserved TA at positions −12 and −11 in *E. coli* −10 hexamers, located at an appropriate distance upstream of each TSS (Dataset S1). The A at position −11 is almost universally conserved in *E. coli* promoters utilizing the major σ-factor (25) and is essential for promoter activity with $RNAP_{Eco}$ (24, 37). Residues involved in recognition of A-11 are conserved in the *R. sphaeroides* primary sigma factor (σ^{93}; *SI Appendix*, Fig. S4*A*). We confirmed that −11A is essential for activity of the *R. sphaeroides rrnB* promoter with $RNAP_{Rsp}$ by showing that substitutions for −11A eliminated transcription (*SI Appendix*, Fig. S5*C*).

In Fig. 6, we analyze the promoters in *R. sphaeroides* that contain −7T genome-wide, based on the presence of the TA motif appropriately positioned upstream of the experimentally defined TSSs, as described above. In support of the interpretation that this promoter collection consists primarily of σ^{93}-dependent promoters, the genome of *R. sphaeroides* does not contain a gene coding for a σ^{S} homolog (11), and other holoenzymes containing alternative σ-factors (at least in *E. coli*) do not contain a TA motif at the upstream end of their −10 elements. The *Bioinformatic Analysis* section in *SI Appendix*, *Expanded Materials and Methods*, contains further details about the promoters included in the genome-wide promoter analysis.

In contrast to *E. coli* promoters, where 95% contain a T at −7, only 43% of *R. sphaeroides* promoters contained −7T (Fig. 6). As with the *R. sphaeroides* promoters lacking −7T that we analyzed in vitro, each of the other three bases at position −7 (A, G, and C) was represented in the genome-wide collection of *R. sphaeroides* promoters lacking −7T, indicating the importance of the absence of −7T rather than the presence of another specific base (Dataset S1 and *Bioinformatic Analysis* section in *SI Appendix*, *Expanded Materials and Methods*).

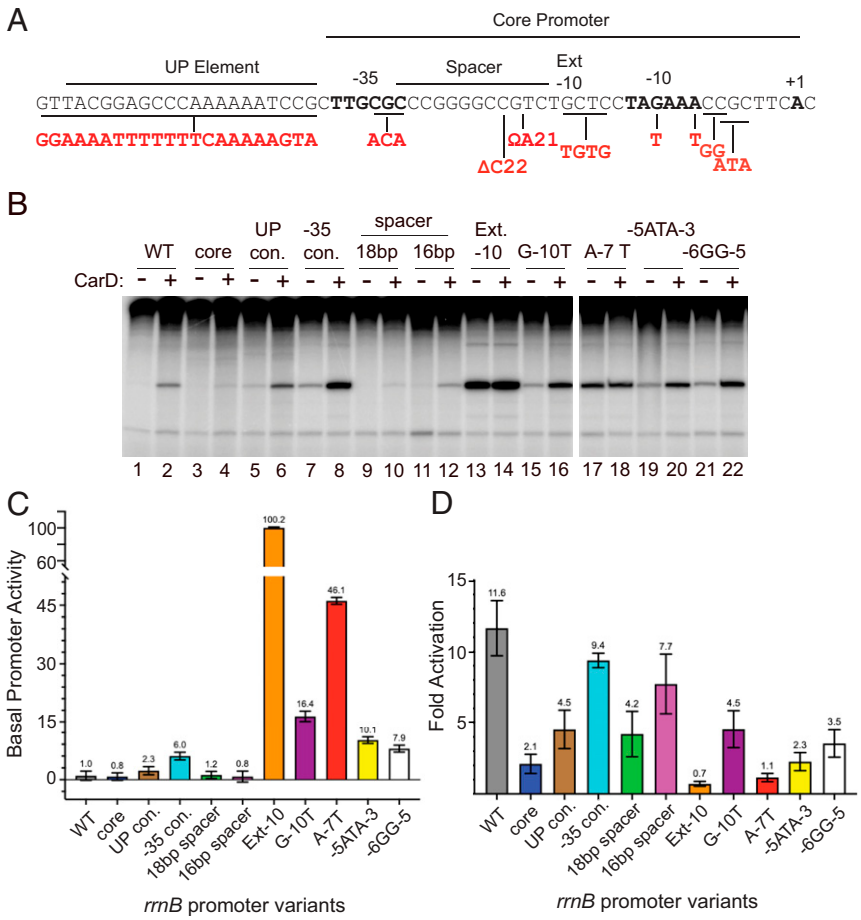

Fig. 4. Effects of individual *R. sphaeroides rrnB* promoter mutations on activation by CarD. (*A*) Promoter recognition elements (UP element, −35 hexamer, spacer, extended −10, −10 hexamer, and core promoter) are indicated above the *R. sphaeroides rrnB* promoter sequence. Mutations to create promoter variants are indicated in red (*SI Appendix*, Tables S2 and S3). Each mutation creates the *E. coli* σ^{70} consensus element sequence at the indicated position except the spacer insertion or deletion for which the wild-type *R. sphaeroides* promoter has the *E. coli* 17-bp consensus spacing. Triple-mutation −5 ATA −3 alters the base composition in the discriminator region. (*B*) In vitro transcription of each promoter mutant with $RNAP_{Rsp}$ (20 nM) with or without CarD (1,280 nM) in buffer with 170 mM NaCl. (*C*) Basal activities of mutant *rrnB* promoters relative to transcription from the wild-type promoter, with SDs from $n = 3$ to 6 assays for each promoter. (*D*). Fold activation of each promoter by CarD (+CarD/no CarD). SDs are from $n = 3$ to 6 assays for each promoter.

We also used available TSS data to analyze the predicted −10 elements from a selection of other bacterial species, including the α-proteobacteria *C. crescentus* (38) and *Zymomonas mobilis* (39), the γ-protobacterium *E. coli* (40), the firmicute *Bacillus subtilis* (41), the actinobacteria *Streptomyces coelicolor* and *M. smegmatis* (42, 43), and the β-proteobacterium *Burkholderia cenocepacia* (44). This analysis showed that ∼39% of the predicted promoters in *C. crescentus* and 30% in *Z. mobilis* contain −7T, whereas 95%, 99%, 90%, 98%, and 98% of the predicted promoters from *E. coli, B. subtilis*, *M. smegmatis*, *S. coelicolor*, and *B. cenocepacia*, respectively, contain −7T (Fig. 6). Thus, although the number of species that we analyzed is limited, our results strongly suggest that the −7T is much less conserved in α-proteobacterial promoters than in promoters from other bacterial phyla or classes.

Because the presence of an extended −10 element eliminated activation of the *R. sphaeroides rrnB* promoter by $CarD_{Rsp}$ (Fig. 4), we also analyzed the percentage of promoters with this element in the genomes of *R. sphaeroides* and the other bacterial species analyzed above (Fig. 6). Unlike −7T, the extended −10 motif was found in about the same percentage of promoters from both α-proteobacteria and non–α-proteobacteria. The number of promoters with an extended −10 element was much smaller than the percentage with a −7T in most of the species analyzed (Fig. 6).

Although an extended −10 element compensated for the absence of −7T in an *rrnB* mutant promoter variant, and this promoter was no longer activated by $CarD_{Rsp}$ (Fig. 4), an extended −10 element did not correlate well with activation by $CarD_{Rsp}$ in the 19 native promoters tested in vitro, in stark contrast to −7T (Fig. 3 and *SI Appendix*, Fig. S7). Together with our results that the majority of putative promoters in α-proteobacteria lack an extended −10 element (Fig. 6), that the presence of an extended −10 element in six other native *R. sphaeroides* promoters tested in vitro did not correlate with activation by $CarD_{Rsp}$ (*SI Appendix*, Fig. S7), and that 85% of *R. sphaeroides* promoters lack an extended −10 element, we conclude that the absence of an extended −10 element is not predictive of activation by $CarD_{Rsp}$.

Levels of $CarD_{Rsp}$ Change In Vivo. One prediction of our data is that promoters lacking −7T may have evolved in α-proteobacteria to be regulated by changing concentrations of CarD. To test whether CarD levels change in vivo, we used purified $CarD_{Rsp}$ to develop a CarD-specific antibody (*SI Appendix, Expanded Materials and Methods*), and $CarD_{Rsp}$ levels were analyzed by

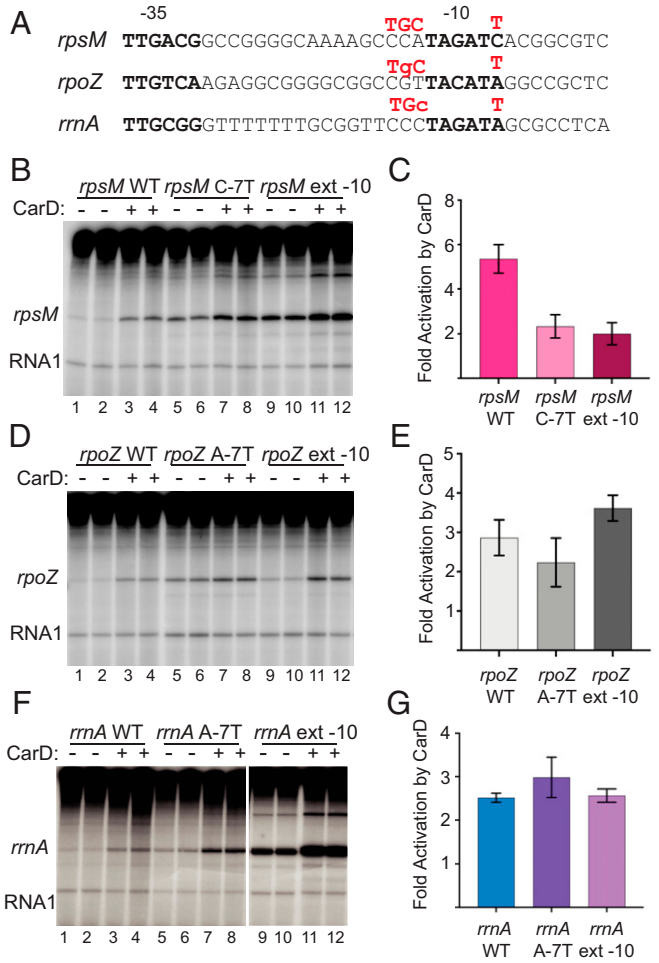

Fig. 5. Context dependence of effects of −7T or extended −10 mutants. (*A*) Positions in three CarD-activated *R. sphaeroides* promoters where mutations were made are indicated in red above each promoter sequence, either a T at −7 or a TGC at −15, −14, −13. (*B*) In vitro transcription with *rpsM*, *rpsM* C-7T, and *rpsM* extended −10 with or without 1,280 nM CarD with 20 nM *R. sphaeroides* RNAP in buffer with 100 mM NaCl. Duplicate samples are shown for each. (*C*) Fold activation (+CarD/no CarD) with SD from three assays like that shown in *B*. (*D*) Same as *B* except the *rpoZ*, *rpoZ* A-7T, and *rpoZ* extended −10 promoter plasmids were used as templates. (*E*) Fold activation (+CarD/no CarD) with SD from three assays like that shown in *D*. (*F*) Same as *B* and *D* except the *rrnA*, *rrnA* A-7T, and *rrnA* extended −10 promoter plasmids were used as templates and the buffer contained 170 mM NaCl. All lanes are from the same gel, with intervening lanes removed for clarity. (*G*) Fold activation (+CarD/no CarD) with range from two assays like that shown in *F*.

Western blot during aerobic growth in minimal medium as cells emerged from stationary phase (outgrowth), grew throughout log phase (log; Fig. 7 *A*–*C*), and went into and remained in stationary phase (stationary; Fig. 7 *A*–*C*). During outgrowth, the $CarD_{Rsp}$ concentration rose gradually to the level present in log phase, whereas the level of the α-subunit of RNAP changed little. The levels of CarD and α remained constant per OD_{600} (an estimate of total protein concentration) during log phase. However, the CarD concentration dropped more than 10-fold during stationary phase, whereas the RNAP α-subunit concentration decreased only slightly. We analyzed the cellular CarD concentration relative to colony-forming units and found that the decrease in CarD concentration did not reduce cell viability (*SI Appendix*, Fig. S8). Based on the purified $CarD_{Rsp}$ standards run in parallel (Fig. 7*A*) and a volume of 0.87 μm^3 for aerobically grown log-phase cells (45), we estimate that the concentrations of CarD are ~1.1 μM in log phase and <100 nM in stationary phase (calculation provided in *SI Appendix, Expanded Materials and Methods*), in the range of the concentrations of CarD that affected transcription in vitro. Taken together, our data indicate that CarD concentrations change in vivo and thus have the potential to regulate transcription in cells. Future studies will investigate the mechanisms responsible for regulating the concentrations (or activities) of CarD under different conditions.

Discussion

The Absence of −7T Is a Major Determinant of the *R. sphaeroides* Transcriptome. We discovered an unexpected feature of *R. sphaeroides* promoters and the promoters of two other α-proteobacterial species, *C. crescentus* (38) and *Z. mobilis* (39). Only 30 to 43% of the promoters from these three α-proteobacterial species contain a thymine base at promoter position −7, the most downstream position in the −10 element, whereas this base is 90 to 99% conserved in promoters in other bacterial phyla (Fig. 6). Most of the 19 *R. sphaeroides* promoters tested in vitro were very poorly transcribed by *R. sphaeroides* RNAP and lacked −7T, suggesting an activator might be required to compensate for the absence of the −7T interaction with σ. We suggest that activation of the large number of promoters lacking −7T by transcription factor(s) makes a major contribution to shaping the transcriptome of α-proteobacteria.

Role of −7T in Promoter Function. Promoters recognized by the major holoenzyme, $E\sigma^{70}$ in *E. coli* or $E\sigma^{93}$ in *R. sphaeroides*, have sequences that vary except at six very highly conserved positions, three each in the −35 and −10 hexamers (35), one of which is the thymine at −7. Transcription initiation is characterized by a series of conformational changes in the promoter and in RNAP, driven by binding free energy, in which RNAP first binds to the promoter to form a closed complex, followed by a series of steps in which RNAP melts the DNA strands to form an open complex (35, 46–48). The −7T plays an important role in this multistep process, fitting base-specifically into a conserved pocket formed by the major σ-factor in promoter complexes in very diverse bacterial phyla (e.g., *E. coli*, *T. thermophilus*, and *M. smegmatis*)

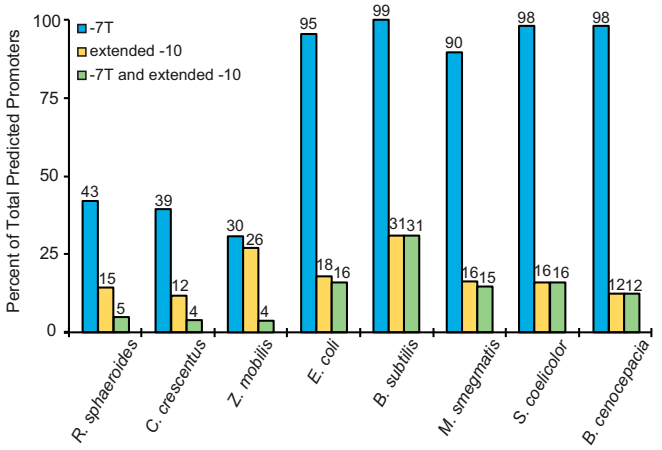

Fig. 6. Percentage of promoters in indicated bacterial species containing a thymine at the last position of the −10 hexamer (−7) or an extended −10 consensus (TG) at −14, −15, or both. Percentages of promoters with identifiable −10 elements are shown on the *y* axis. Percentages of those promoters with −7T or extended −10 elements or both are indicated at the top of each bar and are relative to all promoters with identified −10 elements for that species. Promoters were identified based on published TSS data (cited in the main text) and sequence similarity to consensus *E. coli* −10 element as described in *SI Appendix, Expanded Materials and Methods*.

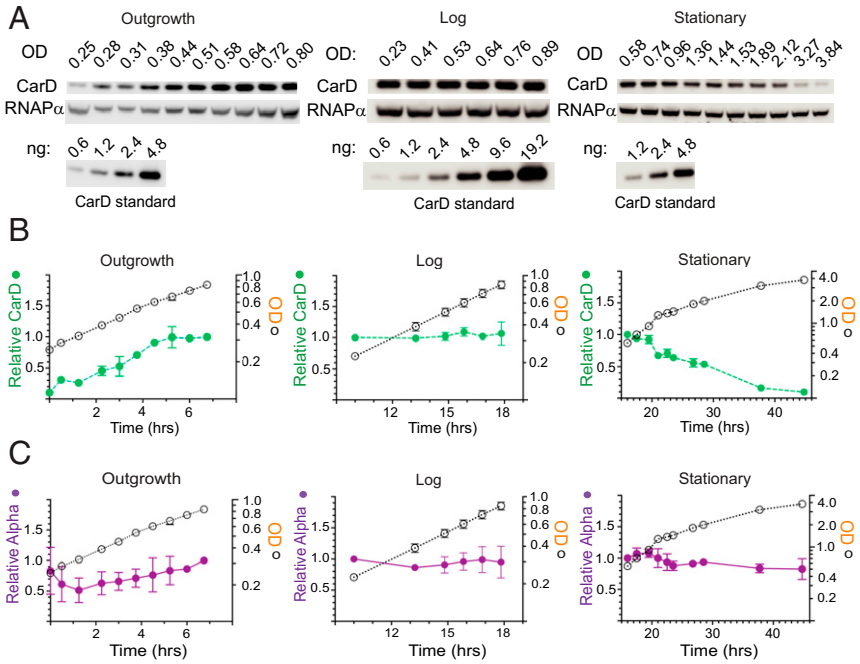

Fig. 7. CarD$_{Rsp}$ levels at different phases of growth. (*A*) Western blots were performed from *R. sphaeroides* aerobic cultures grown at 30 °C, diluted from stationary phase into fresh minimal medium at time 0, and sampled at the indicated times. Separate experiments were conducted for each growth phase (outgrowth, log, and stationary; detailed in *SI Appendix, Expanded Materials and Methods*). For each sample, the OD_{600} was measured, CarD levels were determined by Western blotting with anti-CarD$_{Rsp}$ antibody, and RNAP$_{Rsp}$ levels were determined with an anti–α-subunit antibody. Purified CarD standards were analyzed in parallel for quantitation of CarD levels. (*B*) Relative CarD levels and OD_{600} values for cultures sampled during outgrowth, log, and stationary phase. (*C*) Relative RNAP$_{Rsp}$ α-subunit levels in the same samples analyzed for CarD levels. Further details are provided in *SI Appendix, Expanded Materials and Methods*. Values for CarD$_{Rsp}$, RNAP$_{Rsp}$ α-subunit, and OD_{600} were plotted on the same scale by normalizing to the CarD level at an overlapping OD_{600}. Error ranges were determined from two experiments for each.

(20, 26, 49). This interaction with σ occurs subsequent to the formation of an initial partially opened complex containing a 5-nt transcription bubble (47). The −7T interaction with σ facilitates further DNA strand separation, displacement of the N-terminal domain of σ ($\sigma_{1.1}$) from the main DNA channel, and stabilization of the open complex to yield a 13-nt bubble (47, 48). The absence of the −7T interaction in the majority of *R. sphaeroides* promoters would be expected to reduce the rate of formation and stability of the open complex, creating the need for a transcription factor to enhance the rate-limiting kinetic step(s) at these promoters and thereby enhance promoter activity.

CarD Facilitates a Step in the Mechanism of Promoter Opening that Is Affected by the RNAP Interaction with Promoter Position −7T. We identified a strong correlation between the absence of −7T and activation of native *R. sphaeroides* promoters by the transcription factor CarD. CarD increased the activity of 15 of the 16 promoters that lacked −7T, whereas it had no effect on 3 of the 4 promoters that contained −7T (Fig. 3 and *SI Appendix*, Fig. S7). *R. sphaeroides* CarD shares significant similarity with other CarD homologs from different phyla (e.g., *Mycobacteria* and *Thermus*). Transcription activation by CarD$_{Rsp}$ required a highly conserved tryptophan residue (W91) in the predicted DNA-binding domain, corresponding to W85 in *Mycobacteria* and W86 in *Thermus*, suggesting that CarD$_{Rsp}$ has mechanistic features in common with these previously characterized CarD proteins (Fig. 2) (13, 14, 27). Previous structural studies indicated that CarD$_{Tth}$ interacts sequence-nonspecifically with promoter DNA at the junction of double-stranded and single-stranded DNA at the upstream end of the −10 hexamer (12–16, 27), adjacent to position −11, where strand separation initiates. Kinetic studies with the *Mycobacterial* and *Thermus* CarDs indicated that they affect isomerization step(s) during open complex formation, stimulating formation of a partially melted intermediate and stabilizing open complexes (13, 16, 27, 50).

Thus, CarD and −7T both act on isomerization steps involving extension of a partially open to a fully open transcription bubble and stabilization of the open complex, consistent with the ability of CarD$_{Rsp}$ to compensate, at least in part, for the absence of −7T in many promoters.

Context Affects Promoter Activity. Our in vitro experiments established a strong correlation between promoters lacking −7T and activation by CarD$_{Rsp}$. The high percentage of *R. sphaeroides* promoters genome-wide lacking −7T is consistent with the relative ease with which we identified specific promoters activated directly by CarD$_{Rsp}$, even though the promoters were not chosen for study because they lacked −7T. However, as described below, the absence of −7T is not a perfect predictor of activation by CarD in all promoter contexts.

As a result of the multistep nature of promoter complex formation and the influence of multiple RNAP interactions with other sequences in the promoter, substitutions at the same position in different promoters have long been known to exhibit different effects on transcription output (context effects; e.g., refs. 24, 36). Promoter interactions with RNAP containing the primary σ-factor involve not only the specific base interactions at the critical conserved positions in the consensus elements, but also phosphate backbone interactions within and outside of the consensus hexamers and base-specific interactions in regions of the promoter that may only be present in a subset of promoters. These interactions nevertheless collectively contribute to the rate of open complex formation and can lead to differences in the rate-determining steps that limit promoter activity at different promoters (48).

Transcription can only be activated by a factor that acts on a rate-determining kinetic step. In some contexts, an interaction involving a particular promoter position may be the major determinant of the step affected by the activator, but in other contexts, another promoter interaction, or more than one interaction, may limit promoter activity, and the activator may have little or no effect on transcriptional output (35, 46). The one promoter in our cohort of 16 native *R. sphaeroides* promoters examined in vitro that lacked −7T but was not activated by $CarD_{Rsp}$ is an example of such a context effect. This promoter, *rsp_3110*, contains an extended −10 element that we suggest bypasses the requirement for $CarD_{Rsp}$ by facilitating strand opening, just as a mutation to −7T in some promoter contexts (e.g., *rpoZ* and *rrnA*) does not fully eliminate activation by $CarD_{Rsp}$.

Taken together, our observations suggest that the absence of −7T is a major contributor to the dependence of many *R. sphaeroides* promoters on an activator for transcription. $CarD_{Rsp}$ may not be the transcription factor responsible for regulating every promoter lacking −7T that requires an activator, nor is the presence/absence of −7T likely to be the only determinant of a response to the regulator(s) in all promoter contexts. Thus, prediction of the promoters activated by $CarD_{Rsp}$ genome-wide, as well as prediction of the magnitude of the effect of $CarD_{Rsp}$ on a specific promoter, will be more complex than simply identifying promoters lacking −7T.

The Absence of a Conserved RNAP Recognition Feature in a Large Set of Promoters Creates a Regulon Controlled by a Transcription Factor. The unexpectedly low percentage of *R. sphaeroides* promoters containing −7T (43%) stands in stark contrast to the very high percentage (90 to 99%) containing −7T in other bacterial phyla (Fig. 6). Together, the strong correlation between the absence of −7T and activation by $CarD_{Rsp}$ in our in vitro experiments (Figs. 2 and 3 and *SI Appendix*, Fig. S7), along with the observations that CarD is essential in *R. sphaeroides* (5) and that CarD concentrations vary with growth phase (Fig. 7), suggest that CarD plays a role in activating many promoters lacking −7T. Most bacterial phyla (e.g., β- and γ-Proteobacteria, Firmicutes, Actinobacteria, Thermus-Deinococcus) have promoters containing −7T. How such a large subset of promoters lacking −7T evolved in α-proteobacteria whereas diverse bacterial phyla contain promoters with −7T remains a question for future exploration. We suggest that regulons in other bacterial phyla that derive from utilization of a transcription factor that compensates for the absence of a crucial base in the core promoter, not necessarily −7T, could be a common mechanism contributing to regulation.

The Role of CarD and −7T in Other α-Proteobacteria and Other Bacterial Phyla. Our bioinformatic analysis of published genome-scale TSS data indicated that the absence of −7T is a feature of a majority of promoters in *R. sphaeroides* and the two other α-proteobacterial species for which we were able to predict −10 element sequences genome-wide. Consensus sequences for promoters in the α-proteobacteria *C. crescentus* and *Sinorhizobium meliloti* have been proposed (38, 51–53). These sequence predictions included a −10 element consensus element consisting only of positions matching bases near the upstream end of the *E. coli* −10 element. Furthermore, the *rrnA* promoter from the δ-proteobacterium *M. xanthus* lacks −7T and is activated by CarD (14). Together, these studies are consistent with our proposal that CarD activates transcription in many other species. However, the previous studies did not correlate the absence of −7T with activation by CarD. Deletion of *C. crescentus* CarD (CdnL) was reported to affect transcription from many promoters in vivo, but it was proposed that the number of promoters regulated directly by CarD was small and that most effects of CarD on transcription were likely to be indirect (19). In contrast, our data suggest that α-proteobacterial CarD is likely to affect transcription initiation from many promoters directly.

In Mycobacterial systems, where CarD has been most thoroughly investigated in vitro, most promoters including the rRNA promoter AP3 contain −7T, and thus a role for CarD in compensating for the −7T interaction with σ was not proposed. Other interactions must be rate-limiting in Mycobacterial promoters as a result of DNA sequence differences in the promoter and/or amino acid sequence variation in RNAP among *Mycobacteria*, *Thermus*, and proteobacteria (13, 14, 27, 54). The effects of CarD on the *M. tuberculosis* rRNA promoter are also amplified by a second transcription factor, RbpA, not found in *R. sphaeroides* (16, 50).

Implications for Regulation of Synthesis of the *R. sphaeroides* Translation Apparatus. The *R. sphaeroides* promoters activated by CarD include a variety of housekeeping genes, including several involved in synthesis and assembly of the translation apparatus (rRNAs, ribosomal proteins, tRNAs, transcription factors; Fig. 3 and *SI Appendix*, Fig. S7). Consistent with its critical role in translation, *carD* is an essential gene in *R. sphaeroides* (5). The decrease in CarD levels in stationary phase may contribute to the decrease in expression of rRNA and other translation-related gene products at this stage in growth.

Additional transcription factors are also likely to contribute to rRNA transcription regulation in *R. sphaeroides*. In *E. coli*, the stringent response factors DksA and ppGpp, like CarD in *R. sphaeroides*, interact directly with RNAP and affect specific promoters because of their specific kinetic properties, in contrast to classical transcription factors whose promoter specificity results from DNA binding sites adjacent to individual promoters (8, 35, 46, 50). *R. sphaeroides* DksA was previously shown to regulate transcription by *E. coli* RNAP in vitro (55), suggesting that *R. sphaeroides* uses ppGpp/DksA as well as CarD to regulate rRNA transcription. In *E. coli*, the Fis protein activates rRNA promoters by binding to sites upstream of their −35 elements, contributing to their regulation (6). *R. sphaeroides* rRNA promoters could also respond to transcription factor(s) yet to be identified that bind upstream of the core promoter, since deletions upstream of the −35 element reduce rRNA promoter activity in vivo (11). The integration of effects of CarD and other regulators of *R. sphaeroides* rRNA transcription will be subjects of a separate study.

Materials and Methods

Further details for each section are provided in *SI Appendix, Expanded Materials and Methods*.

Bacterial Strains. *E. coli* and *R. sphaeroides* strains are listed in *SI Appendix*, Table S1.

Bacterial Growth. *R. sphaeroides* was grown aerobically in a succinate-based minimal medium (56).

Purification of *R. sphaeroides* RNAP. An *R. sphaeroides* 2.4.1 derivative for purification of RNAP was constructed by creating a His10 tag fused to the C terminus of the β′ subunit and introducing it into the *R. sphaeroides* chromosome using the nonreplicative plasmid pk18mobsacB (57) and a two-step recombination method. RNAP holoenzyme containing σ^{93} was purified by Ni-affinity chromatography from aerobically grown cells.

Construction and Purification of *R. sphaeroides* Sumo-Tagged CarD. A codon-optimized *carD* gene was inserted into pETSUMO for purification of an N-terminally tagged His10-SUMO–tagged CarD. Wild type and variant CarD proteins were overexpressed and purified from *E. coli* BL21 (DE3) pLysS cells (14).

In Vitro Transcription. Test promoter fragments were PCR-amplified from *R. sphaeroides* 2.4.1 chromosomal DNA and inserted into pRLG770 (6), and promoter variants were constructed using Multi Site Lightning Quick Change

Mutagenesis (Agilent) using primers listed in *SI Appendix*, Table S2. Transcription was performed with either *R. sphaeroides* RNAP holoenzyme containing the major σ-factor σ^{93} or *E. coli* $E\sigma^{70}$ as indicated.

Western Blot Analysis of CarD. CarD concentrations in aerobically grown *R. sphaeroides* cells were determined throughout a growth curve using a polyclonal antibody raised against CarD without a SUMO tag. Polyclonal antibody raised against the α-subunit of *R. sphaeroides* RNAP was used for comparison.

Bioinformatic Analysis of Promoter Elements. We used available transcription start site (TSS) information in the literature to determine the likely −10 elements of promoters from the bacterial species indicated in Fig. 6, first identifying the bases most likely corresponding to −11A and −12T and then the DNA sequences most likely corresponding to −7T and the extended −10 element, as described in detail in *SI Appendix, Expanded Materials and Methods*. The TSSs for *R. sphaeroides* were determined as previously described (33).

Data Availability. All study data are included in the article and supporting information.

ACKNOWLEDGMENTS. We thank A. Mehle for access to the Li-Cor Fc instrument used for imaging Western blots and D. Grainger for providing genome-wide *B. subtilis* transcription start site information prior to publication. We also thank F. Alberge for measurements of *R. sphaeroides* cell volume. Work in the R.L.G. laboratory is supported by NIH R01 GM37048. K.K.H. was also supported by fellowships from the Jack Kent Cooke Foundation, an NIH Molecular Biosciences Training Grant (T32GM007215), and the Department of Bacteriology. Work in the T.J.D. and R.L. labs is supported by US Department of Energy, Office of Biological and Environmental Research, under Award DE-SC0018409.

1. C. Mackenzie *et al*., Postgenomic adventures with *Rhodobacter sphaeroides*. *Annu. Rev. Microbiol.* **61**, 283–307 (2007).
2. S. Imam, D. R. Noguera, T. J. Donohue, Global insights into energetic and metabolic networks in Rhodobacter sphaeroides. *BMC Syst. Biol.* **7**, 89 (2013).
3. S. Imam, D. R. Noguera, T. J. Donohue, Global analysis of photosynthesis transcriptional regulatory networks. *PLoS Genet.* **10**, e1004837 (2014).
4. J. H. Zeilstra-Ryalls, M. Gomelsky, A. A. Yeliseev, J. M. Eraso, S. Kaplan, Transcriptional regulation of photosynthesis operons in *Rhodobacter sphaeroides* 2.4.1. *Methods Enzymol.* **297**, 151–166 (1998).
5. B. T. Burger, S. Imam, M. J. Scarborough, D. R. Noguera, T. J. Donohue, Combining genome-scale experimental and computational methods to identify essential genes in *Rhodobacter sphaeroides*. *mSystems* **2**, e00015–e00017 (2017).
6. 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).
7. W. Ross *et al*., A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407–1413 (1993).
8. R. L. Gourse *et al*., Transcriptional responses to ppGpp and DksA. *Annu. Rev. Microbiol.* **72**, 163–184 (2018).
9. P. Sanchez-Vazquez, C. N. Dewey, N. Kitten, W. Ross, R. L. Gourse, Genome-wide effects on *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8310–8319 (2019).
10. S. C. Dryden, S. Kaplan, Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter sphaeroides*. *Nucleic Acids Res.* **18**, 7267–7277 (1990).
11. S. C. Dryden, S. Kaplan, Identification of cis-acting regulatory regions upstream of the rRNA operons of *Rhodobacter sphaeroides*. *J. Bacteriol.* **175**, 6392–6402 (1993).
12. C. L. Stallings *et al*., CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. *Cell* **138**, 146–159 (2009).
13. B. Bae *et al*., CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife* **4**, 08505 (2015).
14. D. B. Srivastava *et al*., Structure and function of CarD, an essential mycobacterial transcription factor. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12619–12624 (2013).
15. A. Gallego-García *et al*., Structural insights into RNA polymerase recognition and essential function of *Myxococcus xanthus* CdnL. *PLoS One* **9**, e108946 (2014).
16. E. A. Hubin *et al*., Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *eLife* **6**, e22520 (2017).
17. A. Gallego-García *et al*., *Caulobacter crescentus* CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. *Sci. Rep.* **7**, 43240 (2017).
18. D. X. Zhu, A. L. Garner, E. A. Galburt, C. L. Stallings, CarD contributes to diverse gene expression outcomes throughout the genome of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13573–13581 (2019).
19. S. A. Woldemeskel *et al*., The conserved transcriptional regulator CdnL is required for metabolic homeostasis and morphogenesis in ***Caulobacter***. *PLoS Genet.* **16**, e1008591 (2020).
20. A. Feklistov, S. A. Darst, Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**, 1257–1269 (2011).
21. J. Ryals, R. Little, H. Bremer, Control of rRNA and tRNA syntheses in *Escherichia coli* by guanosine tetraphosphate. *J. Bacteriol.* **151**, 1261–1268 (1982).
22. S. T. Estrem, T. Gaal, W. Ross, R. L. Gourse, Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9761–9766 (1998).
23. S. P. Haugen *et al*., rRNA promoter regulation by nonoptimal binding of sigma region 1.2: An additional recognition element for RNA polymerase. *Cell* **125**, 1069–1082 (2006).
24. H. Moyle, C. Waldburger, M. M. Susskind, Hierarchies of base pair preferences in the P22 ant promoter. *J. Bacteriol.* **173**, 1944–1950 (1991).
25. R. K. Shultzaberger, Z. Chen, K. A. Lewis, T. D. Schneider, Anatomy of *Escherichia coli* sigma70 promoters. *Nucleic Acids Res.* **35**, 771–788 (2007).
26. Y. Zhang *et al*., Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012).
27. E. Davis, J. Chen, K. Leon, S. A. Darst, E. A. Campbell, Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. *Nucleic Acids Res.* **43**, 433–445 (2015).
28. D. García-Moreno *et al*., 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 (2010).
29. L. F. Westblade *et al*., Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. *Nucleic Acids Res.* **38**, 8357–8369 (2010).
30. R. K. Karls, D. J. Jin, T. J. Donohue, Transcription properties of RNA polymerase holoenzymes isolated from the purple nonsulfur bacterium *Rhodobacter sphaeroides*. *J. Bacteriol.* **175**, 7629–7638 (1993).
31. L. A. Weiss *et al*., Interaction of CarD with RNA polymerase mediates *Mycobacterium tuberculosis* viability, rifampin resistance, and pathogenesis. *J. Bacteriol.* **194**, 5621–5631 (2012).
32. R. Landick, A. Krek, M. S. Glickman, N. D. Socci, C. L. Stallings, Genome-wide mapping of the distribution of CarD, RNAP σ^A, and RNAP β on the *Mycobacterium smegmatis* chromosome using chromatin immunoprecipitation sequencing. *Genom. Data* **2**, 110–113 (2014).
33. K. S. Myers *et al*., Genome-wide identification of transcription start sites in two alphaproteobacteria, rhodobacter sphaeroides 2.4.1 and novosphingobium aromaticivorans DSM 12444. *Microbiol. Resour. Announc.* **9**, e00880-20 (2020).
34. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
35. S. P. Haugen, W. Ross, R. L. Gourse, Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6**, 507–519 (2008).
36. D. Graña, T. Gardella, M. M. Susskind, The effects of mutations in the ant promoter of phage P22 depend on context. *Genetics* **120**, 319–327 (1988).
37. E. Heyduk, K. Kuznedelov, K. Severinov, T. Heyduk, A consensus adenine at position -11 of the nontemplate strand of bacterial promoter is important for nucleation of promoter melting. *J. Biol. Chem.* **281**, 12362–12369 (2006).
38. B. Zhou *et al*., The global regulatory architecture of transcription during the *Caulobacter* cell cycle. *PLoS Genet.* **11**, e1004831 (2015).
39. J. M. Vera *et al*., Genome-scale transcription–translation mapping reveals features of *Zymomonas mobilis* transcription units and promoters. *mSystems* **5**, e00250-20 (2020).
40. 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).
41. E. Warman, D. Forrest, J.T. Wade, D.C. Grainger, Widespread divergent transcription from prokaryotic promoters. bioRxiv:10.1101/2020.01.31.928960 (2 Feb 2020).
42. Y. Jeong *et al*., The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat. Commun.* **7**, 11605 (2016).
43. X. Li *et al*., Transcriptome landscape of *Mycobacterium smegmatis*. *Front. Microbiol.* **8**, 2505 (2017).
44. A. M. Sass *et al*., Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315. *BMC Genomics* **16**, 775 (2015).
45. K. C. Lemmer *et al*., The NtrYX two-component system regulates the gram-negative cell envelope. *mBio* **11**, e00957-20 (2020).

46. E. A. Galburt, The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11604–E11613 (2018).
47. J. Chen *et al*., Stepwise promoter melting by bacterial RNA polymerase. *Mol. Cell* **78**, 275–288.e6 (2020).
48. E. F. Ruff, M. T. Record, Jr, I. Artsimovitch, Initial events in bacterial transcription initiation. *Biomolecules* **5**, 1035–1062 (2015).
49. H. Boyaci, J. Chen, R. Jansen, S. A. Darst, E. A. Campbell, Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* **565**, 382–385 (2019).
50. J. Rammohan, A. Ruiz Manzano, A. L. Garner, C. L. Stallings, E. A. Galburt, CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. *Nucleic Acids Res.* **43**, 3272–3285 (2015).
51. J. Malakooti, S. P. Wang, B. Ely, A consensus promoter sequence for Caulobacter crescentus genes involved in biosynthetic and housekeeping functions. *J. Bacteriol.* **177**, 4372–4376 (1995).
52. S. R. MacLellan, A. M. MacLean, T. M. Finan, Promoter prediction in the rhizobia. *Microbiology (Reading)* **152**, 1751–1763 (2006).
53. J. P. Schlüter *et al*., Global mapping of transcription start sites and promoter motifs in the symbiotic α-proteobacterium *Sinorhizobium meliloti* 1021. *BMC Genomics* **14**, 156 (2013).
54. T. Kouba *et al*., The core and holoenzyme forms of RNA polymerase from *Mycobacterium smegmatis*. *J. Bacteriol.* **201**, e00583-18 (2019).
55. C. W. Lennon *et al*., A Rhodobacter sphaeroides protein mechanistically similar to *Escherichia coli* DksA regulates photosynthetic growth. *mBio* **5**, e01105–e01114 (2014).
56. W. R. Sistrom, A requirement for sodium in the growth of *Rhodopseudomonas spheroides*. *J. Gen. Microbiol.* **22**, 778–785 (1960).
57. 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).