Canadian Journal of Microbiology
Revue canadienne de microbiologie

**Functional reconstitution of an unusual Firmicutes σ factor
into a Gram negative heterologous host**

Journal:	<i>Canadian Journal of Microbiology</i>
Manuscript ID	cjm-2015-0408.R1
Manuscript Type:	Article
Date Submitted by the Author:	11-Aug-2015
Complete List of Authors:	Davis, Maria C.; University of New Brunswick, Department of Biology Kesthely, Christopher A.; University of New Brunswick, Department of Biology Smith, Logan K.; University of New Brunswick, Department of Biology Breen, Jillian; University of New Brunswick, Department of Biology MacLellan, Shawn R.; University of New Brunswick, Department of Biology
Keyword:	sigma factor, RNA polymerase, transcription, two-subunit

SCHOLARONE™
Manuscripts

**Functional reconstitution of an unusual Firmicutes σ factor into a Gram negative
heterologous host**

Maria C. Davis^a, Christopher A. Kesthely^b, Logan K. Smith^c, Jillian Breen^d, and Shawn R.
MacLellan^{e#}

^a maria.davis@unb.ca

^b chris.k@unb.ca

^c LoganK.Smith@unb.ca

^d breen1@ualberta.ca

^e srm@unb.ca

[#] corresponding author:

Shawn MacLellan

Department of Biology

University of New Brunswick

Fredericton, New Brunswick

E3B 5A3

srm@unb.ca

Abstract

σ factors are single subunit proteins that reversibly bind RNA polymerase and play an important role in the transcription initiation process. An unusual two-subunit σ factor, consisting of proteins SigO and RsoA, activates transcription from a group of related promoters in *Bacillus subtilis*. These two proteins specifically interact with each other, and with RNA polymerase subunits. This system is widespread amongst species in several *Bacillus*-related genera, but otherwise appears restricted to the Firmicutes. Here we reconstituted SigO-RsoA, and a cognate promoter, into the distantly related heterologous host *Escherichia coli* to examine whether this system can function in bacteria outside of the Firmicutes. We show that these proteins can productively associate with *E. coli* RNA polymerase and activate transcription, demonstrating that there are no structural barriers to function. In parallel, we tested a wide array of protein-protein interaction mutations and promoter mutations that impact SigO-RsoA function in both *B. subtilis* and *E. coli* and conclude that the SigO-RsoA system behaves, in most instances, similarly in both genetic backgrounds. These data raise the possibility of genetically isolating the system in this heterologous host and away from unknown *B. subtilis* factors that may also be playing a role in SigO-RsoA regulatory pathways, thus facilitating further study of the system. As a result of this work, we also provide a comprehensive mutational analysis of a SigO-RsoA promoter and report the preliminary identification of amino acids in SigO that play a role in mediating the SigO-RsoA protein-protein interaction.

Key Words: sigma factor, RNA polymerase, transcription, two-subunit

Introduction

RNA polymerase (RNAP) is the multi-subunit complex responsible for transcription in all cellular life forms. Although subunit complexity differs amongst the Bacteria, Archaea, and Eukarya, the core polymerase consisting of (in bacterial nomenclature) the β , β' , ω , α subunits are thought to have arisen from a common ancestor (Werner 2008, Lane and Darst 2010, Werner and Grohmann 2011). Core RNAP is capable of RNA synthesis when provided with single-stranded DNA template, but gene specific transcription requires one or more accessory subunits that target or recruit RNAP holoenzyme to gene promoters, and also play a role in the initiation of transcription.

In bacteria, the σ factor family of general transcription factors fulfills the aforementioned specificity function by binding reversibly to core RNAP to form holoenzyme and targeting the complex to gene promoters (Burgess et al. 1969, Burgess and Anthony 2001). Importantly, σ factors also play a critical role in transcription initiation by binding the non-template strand of the -10 promoter element DNA, thereby trapping the promoter in a single-stranded state that encompasses the -10 element and the transcription (+1) start site (Juang and Helmann 1994, deHaseth and Helmann 1995). This single-stranded state is called an open-complex, and its formation is required for initiation of RNA synthesis from the template strand at the RNAP active site.

σ factors are single subunit proteins ranging in size from about 70 kDa (in the case of primary Group I σ factors) to about 20 kDa (in the case of Group IV or ECF (Extracytoplasmic Function) σ factors). ECF σ factors are the smallest of the σ^{70} family of general transcription factors and possess what are called regions 2 and 4 (Missiakas and Raina 1998, Helmann 2002). The domain

67 encompassing region 4 interacts with the core RNAP β subunit and binds the -35 promoter
68 element via a helix-turn-helix motif (Geszvain et al. 2004, Lane and Darst 2006). The domain
69 encompassing region 2 interacts with the core RNAP β' subunit and, critically, also binds the -10
70 element non-template DNA strand to stabilize open-complex.

71 Previously, an unusual two-subunit σ factor was identified in the Gram positive bacterium
72 *Bacillus subtilis* (MacLellan et al. 2009a). These two subunits, SigO and RsoA, are homologous
73 to σ^{70} -family σ factors (MacLellan et al. 2008, 2009a). SigO is 23 kDa in size, carries a σ^{70} region
74 4 in its C-terminal domain, but has a highly divergent N-terminal domain with very limited
75 similarity to the canonical σ^{70} region 2. SigO appears to be a partially inactivated ECF σ factor
76 and it cannot stabilize open-complex *in vitro*, or activate transcription *in vivo*, on its own.

77 Instead, the presence of both SigO and RsoA are required for *in vitro* and *in vivo* activities (for a
78 current model, see Fig 1B) (MacLellan et al. 2009a). RsoA is a 10 kDa protein with sequence
79 similarity to sub-regions within σ^{70} domain 2 including sub-region 2.3 which plays a major role
80 in binding the -10 non-template DNA and stabilizing open complex. SigO and RsoA interact with
81 the β and β' subunits, respectively, and RsoA specifically interacts with the N-terminal domain
82 of SigO (MacLellan et al. 2009a). The mechanistic details of how SigO and RsoA cooperate to
83 activate transcription are not yet understood. The simplest hypothesis is that SigO contributes
84 region 4 activity and RsoA contributes region 2 function, including the critical role of open-
85 complex stabilization by region 2.3. However, we have recently mapped a 10-15 amino acid
86 motif in RsoA that mediates its interaction with SigO and this motif is the peptide with high
87 similarity to region 2.3 (manuscript in preparation) (also see Fig. S1A). It is not yet clear whether
88 region 2.3 of RsoA can function simultaneously to mediate interaction with SigO and stabilize

open-complex. Accordingly, a second hypothesis regarding SigO-RsoA mechanism is that RsoA performs a bridging role between core RNAP and SigO, and helps to orient the highly divergent, but perhaps still functional, N-terminus of SigO for proper interaction with the promoter -10 element and subsequent stabilization of open-complex.

The SigO-RsoA system is upregulated in the cell during growth under acidic conditions (MacLellan et al. 2009b) and after exposure to sub-lethal concentrations of certain cell wall-acting antibiotics (Hachmann et al. 2009). We have identified close to one hundred non-redundant orthologs in various species within the subtilis and cereus clades of bacilli (not shown), but the SigO-RsoA system so far seems restricted to a subset of genera with the Firmicutes. In this work, we examine whether there are structural reasons for a narrow host range, or whether this atypical σ factor could in principle function in a wider array of bacterial genera. To answer this question, we reconstituted SigO and RsoA into the very distantly related Gram negative bacterium *Escherichia coli*, and examined whether the two proteins can functionally associate with *E. coli* RNAP and initiate transcription from a cognate promoter in these cells. The *E. coli* and *B. subtilis* RNAP subunits are, in general, highly conserved. However, there are insertional variations between the catalytic β and β' subunits in these bacteria, and indications of phenotypic differences between the activities of RNAP from *E. coli* and *B. subtilis* (Zhou and Jin 1998, Artsimovitch et al. 2000). It was therefore possible that SigO-RsoA may not function in a Gram negative bacterium. An additional practical reason for conducting the present work was to explore the possibility of using *E. coli* as a heterologous host for SigO-RsoA, thereby simultaneously taking advantage of certain plasmid-based genetic techniques available

in *E. coli*, and also isolating the system from other *B. subtilis* proteins that may be playing an unknown role in SigO-RsoA regulation and function.

We report that SigO and RsoA do function in *E. coli* and based on a variety of comparative analyses, appear to function similarly in both bacteria. As a consequence of this work, we also report an analysis of SigO-RsoA promoter structure and sequence, and report the first identification of amino acid residues in the N-terminus of SigO that mediate its interaction with RsoA.

Materials and methods

Bacterial strains and growth conditions

Bacterial strains used in this study are listed in Table S1. Plasmids used are listed in Table S2. *B. subtilis* strains were usually cultured in LB broth (Fisher Scientific product number BP1426) or this medium solidified with 1.6% agar. For experiments involving σ factor expression in *E. coli*, M9 minimal media (Miller 1987) was used. Antibiotic concentrations used were: Ampicillin (100 $\mu\text{g/mL}$), Kanamycin (50 $\mu\text{g/mL}$), Streptomycin (100 $\mu\text{g/mL}$), Spectinomycin (100 $\mu\text{g/mL}$), Lincomycin (25 $\mu\text{g/mL}$), Erythromycin (1 $\mu\text{g/mL}$), Chloramphenicol (10 $\mu\text{g/mL}$). Indicator media included either 40 $\mu\text{g/mL}$ (*E. coli*) or 80 $\mu\text{g/mL}$ (*B. subtilis*) 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal). Isopropyl- β -D-1-thiogalactopyranoside (IPTG) was used at a final concentration of 1 mM.

Bacterial Two-Hybrid Assay

Characterizations of protein-protein interactions were conducted using an adenylate cyclase-based BACTH assay (Euromedex Inc) (Karimova et al. 1998). *sigO* and *rsoA* gene sequences were

amplified using PCR and ligated into pUT18C and pKT25 (respectively) to generate C-terminal fusions to the CyaA T18 and T25 catalytic sub-domains. All possible combinations of fusions using the available BACTH plasmids and the *sigO* and *rsoA* genes have been tested and all yield comparable results in terms of apparent interactions (data not shown). Constructed bait and prey plasmids were co-transformed into the test strain *E. coli* BTH101. In this system, positive interactions are indicated by the expression of a *lacZ* reporter gene. Quantitative measurements were conducted by growing transformants at 37°C overnight, subculturing a 2.5 % volume into fresh LB broth and growing to an O.D. of 0.3-0.4 at 30°C before subjecting cells to a β -galactosidase assay using ONPG (ortho-Nitrophenyl - β -galactoside) as substrate.

Promoter activity assays

In these experiments, a promoter from upstream of the *oxdC* gene (P_{oxdC}) (see Fig. 1A) was used to monitor SigO-RsoA-dependent transcription activity in cells. In *E. coli* cells, a P_{oxdC} -*lacZ* fusion was carried on plasmid pSAD00 or pMCD6 (or their mutant derivatives) and these plasmids were co-transformed into cells along with a SigO-RsoA expression plasmid (pSAD51). For *B. subtilis*, a P_{oxdC} -*lacZ* fusion was generated by ligating the promoter into plasmid pDG1663 (Guérout-Fleury et al. 1996), which was subsequently integrated ectopically into the *thrC* site of the *B. subtilis* genome. For initial experiments to determine whether SigO-RsoA could activate transcription in *E. coli* (Fig. 2), a P_{oxdC} -*gfp* fusion was generated by cloning the promoter into pSRM1, a derivative of pOT1 (Allaway et al. 2001).

Genetic techniques

Standard techniques were used for PCR amplification and cloning. For generating site-directed mutations, we utilized a technique based upon the Stratagene Quikchange method as previously described (MacLellan et al. 2008) or using overlap extension PCR.

Computational analyses

Clustal Omega was used for sequence alignments and WebLogo was used to generate sequence logo representations of conserved nucleotides in SigO-RsoA promoters. Spectrophotometric measurements were acquired using a Molecular Dynamics Paradigm multimodal plate reader. Gfp activity was visualized over a UV box and photographed. Photoshop was used to prepare the composite image shown in Fig. 2.

Results and Discussion

The SigO-RsoA system regulates the expression of five genes from three related promoters in *B. subtilis* (Figure 1). Three of these genes encode regulatory proteins including the autoregulated *sigO-rsoA* operon, and a gene downstream of *oxdC* called *rsiO* (formerly *yvrL*). We have recently determined that *rsiO* encodes a membrane-spanning anti- σ factor that is cognate for the SigO-RsoA system (manuscript in preparation). These regulatory proteins (SigO, RsoA, and RsiO) seem focused on regulating the synthesis of OxdC, an oxalate decarboxylase that may play a role in pH homeostasis in *B. subtilis*. The fifth gene regulated by SigO-RsoA is *yvrJ*, which is currently of unknown function.

SigO and RsoA activate transcription in a heterologous host

To test whether SigO and RsoA could functionally associate with *E. coli* RNAP, we co-transformed *E. coli* cells with a plasmid expressing the genes from separate IPTG inducible

promoters. Activation of transcription was monitored by using a plasmid carrying a target promoter (the *oxdC* operon promoter, *PoxdC*) transcriptionally fused to *gfp*. Cells were patched onto two different base media and incubated overnight at three different temperatures. Gfp was maximally synthesized during growth on M9 minimal medium in the presence of IPTG with incubation at 22°C (Fig. 2) indicating that SigO and RsoA can productively associate with *E. coli* RNAP. Gfp activity was low on LB medium at all temperatures tested, and on M9 medium at 30°C and 37°C. This is likely due to protein solubility effects. SigO, like all σ factors, accumulates in *E. coli* predominantly as inclusion bodies (MacLellan et al. 2009a) when overexpressed. The slower growth of cells at low temperatures maximizes soluble protein synthesis and the attenuated growth of *E. coli* on minimal medium and at 22°C likely results in a larger fraction of soluble SigO synthesis.

The effects of promoter mutations on patterns of expression

SigO-RsoA target promoters possess unusual features as deduced from a sequence logo depiction from an alignment of 168 known or predicted *sigO-rsoA*, *yvrJ*, and *oxdC-rsiO* promoters (Fig 3ABCD). Although SigO has the dimensions of a Type IV (ECF) σ factor and both proteins have greater primary amino acids sequence similarity to alternative σ factors, rather than the primary σ factors, the -35 element of target promoters (T-T-G/T-A-C-T) is very similar to the *B. subtilis* and *E.coli* primary σ factor consensus (T-T-G-A-C-A), a consensus widely shared amongst most bacterial primary σ factors. In both *B. subtilis* and *E.coli* on the other hand, many ECF promoter -35 elements resemble the general consensus sequence G-G/A-A-A-C-T-T (Lane and Darst 2006). The SigO-RsoA -10 element (C-A-C-T-T-A) consensus is distinct from that of primary σ factors (T-A-T-A-A-T). The spacer regions between the -35 and -10 elements are also

unusually T-rich in SigO-RsoA promoters (Fig. 3ABCD), a feature that may be topologically important for recognition by this two-subunit σ factor.

To compare the activity of SigO and RsoA in *B. subtilis* and *E. coli* transcription complexes, we tested a comprehensive array of *PoxdC* -10 element (Fig. 4A) mutations. These 18 mutations have similar effects when tested in the native *B. subtilis* host and the heterologous *E. coli* host (Fig. 4B). Most of the mutations, particularly those in the 5' half of the element, substantially impair transcription, and transversion mutations tend to have a greater negative effect than transition mutations. These data demonstrate a close concordance between SigO-RsoA activity in *E. coli* and *B. subtilis* transcription complexes, amongst a large number of promoter mutants. In Fig. 4B we relate reporter gene expression as a percentage of wt promoter activity in each genetic background. In absolute terms, the activity in the *E. coli* cells is much higher than in *B. subtilis* cells (see legend). This reflects the fact that both the expression plasmid and reporter plasmid are single-copy genomic integrants in *B. subtilis*, while in *E. coli* they are multi-copy replicating plasmids.

The effects of RsoA protein-protein interaction motif mutations on patterns of expression

SigO and RsoA associate with RNAP subunits, and with each other in a transcription complex. RsoA specifically interacts with the N-terminal half of SigO (MacLellan et al. 2009a) and we have recently mapped the interaction motif of RsoA to a highly hydrophobic segment of amino acids in its C-terminus (manuscript in preparation) (also see FigS1A). Two C-terminal residues in the 79 amino acid RsoA protein (F67 and F70) are conserved amongst all RsoA homologs, and most canonical σ factors. Their substitution with alanine greatly impairs the SigO-RsoA interaction and essentially abolishes transcription activity in *B. subtilis*. Mutation E69V that was isolated in

219 a random screen for interaction defects also abolishes activity in *B. subtilis*, while mutation
220 M74A has little effect on transcription. Minor topological differences between *B. subtilis* and *E.*
221 *coli* transcription complexes could lead to differential effects from the consequences of SigO-
222 RsoA interaction mutations. We therefore compared the transcriptional effects of RsoA
223 interaction mutations known to display varying degrees of interaction impairment (F67A, E69V,
224 F70A, and M74A) in both genetic backgrounds (Figure 5). Three of the mutations (F67A, E69V,
225 and F70A) that impair transcription in *B. subtilis* similarly impaired transcription in the *E. coli*
226 background. Interestingly, mutation M74A has little effect on transcription in *B. subtilis*, but has
227 a dramatic effect on transcription in *E. coli*. This construct was re-built and re-tested to confirm
228 this result. The concordance between the effects of mutations in both genetic backgrounds
229 suggests that *E. coli* can be used as a heterologous host for the preliminary identification of
230 important SigO-RsoA mutations. The differential effect from mutation M74A however, may
231 reflect that mutations having little impact in the context of host RNAP could perturb function
232 with a non-host RNAP. This observation underscores the need for thorough characterization of
233 mutations that have been presumptively identified in *E. coli*, in the native context. Differential
234 phenotypes could also emerge from the differential expression, proper folding, or accumulation
235 of certain mutant proteins in a non-native host. We did not examine these parameters in this
236 survey, but controlling for expression levels would be a required element of any comprehensive
237 mutational analysis.

238 **The effects of SigO protein-protein interaction motif mutations on patterns of expression**

239 While RsoA residues important for the SigO-RsoA interaction have been identified, thus far
240 amino acids in SigO important for the interaction have not been reported. Here we

communicate the preliminary identification of three contiguous residues in the N-terminus of SigO that abolish the SigO-RsoA interaction. Using a sequence alignment-guided approach, we used site-directed mutagenesis to impart alanine substitutions in 19 residues that are highly conserved amongst the N-termini of SigO homologues. An adenylate cyclase-based bacterial two-hybrid assay (BACTH) assay was used to test the effect of these alanine substitutions in SigO on its interaction with RsoA. A number of mutations had mild effects on the interaction between SigO and RsoA, but three near-contiguous substitutions (L48A, D49A, and F52A) in SigO (see Fig. S1B) led to pronounced impairment of the interaction (Figure 5). Understanding the topology of the SigO-RsoA transcription complex is an important component of understanding the mechanism by which these proteins activate transcription. The data in Figure 5 are an incremental step (along with our previous characterization of the RsoA interaction motif) toward defining the SigO-RsoA interaction interface.

As we did for RsoA interaction mutations (Figure 5), we compared the *in vivo* transcription effects of the three aforementioned SigO mutations (L48A, D49A, and F52A), and a fourth mutation (F55A) that has a moderate effect on interaction, in the *B. subtilis* and *E. coli* backgrounds (Figure 6). Mutation of SigO residues L48, D49, and F52 reduced transcription to between 10-20% of wt SigO in both genetic backgrounds. Mutation F55A, however, had no detectable effect on reporter gene expression in *B. subtilis* but reduced transcription in *E. coli* to approximately 30% of wt activity, again illustrating the need for testing all mutations isolated in an *E. coli* background in *B. subtilis*.

The rationale for testing protein-protein interaction defects in both genetic backgrounds is that topological differences amongst *B. subtilis* and *E. coli* transcription complexes could either

accentuate or suppress the effects of these mutations. From a gross phenotypic perspective, however, most mutations have essentially indistinguishable effects in the divergent genetic backgrounds. As noted for both RsoA and SigO however, two mutations (M74A in RsoA and F55A in SigO) have differential effects that are more pronounced in *E. coli*. This could reflect a sub-optimal association of SigO and RsoA with *E. coli* RNAP that results in an exaggerated mutant phenotype in this species.

Conclusions

We have shown that the *B. subtilis* proteins SigO and RsoA are able to functionally associate with *E. coli* RNAP and activate transcription from a target promoter. A number of promoter and protein-protein interaction motif mutations have essentially indistinguishable effects when compared in *B. subtilis* and *E. coli* genetic backgrounds. The mutations tested are part of ongoing investigations into the mechanism by which SigO-RsoA activates transcription. Mapping interaction motifs has helped establish the topological organization of a transcription complex involving these atypical σ factor-related proteins. Furthermore, the comprehensive set of *PoxdC* -10 element promoter mutations discussed here are being used to seek suppressor mutations in either SigO or RsoA, in an attempt to determine which of these two proteins interact with the -10 element to stabilize the pre-transcription initiation open complex. These various mutations also afforded the ability to examine two other broad questions: namely, can this atypical firmicutes σ factor system also function in an evolutionarily distant context, and can this system be studied using *E. coli* as a heterologous host to facilitate some mechanistic investigations.

From an evolutionary perspective, it is not surprising that σ factors are cross-functional between *B. subtilis* and *E. coli*, given the conservation amongst RNAP subunit proteins and amongst σ factor proteins within most groups of bacteria. However, structural and functional differences between Gram positive and Gram negative RNAP have been documented (Zhou and Jin 1998, Artsimovitch et al. 2000). For example, the *E. coli* and *B. subtilis* catalytic β and β' share only about 50% amino acid identity and both *B. subtilis* catalytic subunits are substantially smaller than those in *E. coli*. These enzyme complexes also differ in their recognition of some regulatory sequences including certain promoters (Zhou and Jin 1998, Artsimovitch et al. 2000). It was possible that these differences might preclude functioning of the atypical two-subunit SigO-RsoA system in an *E. coli* background. We demonstrate that this non-canonical system, which appears confined to the Firmicutes group, could in principle arise outside of this family. The study of SigO-RsoA also has broader evolutionary implications: the evolution of a two-subunit σ factor represents an increase in RNAP holoenzyme complexity. How multi-subunit complexes evolve increasing complexity is of considerable general interest and both neutral and adaptive mechanisms have been proposed (Stoltzfus 1999, Lukeš et al. 2011). It is also worthwhile to note that core RNAP in the bacteria, archaea, and eukaryotes are believed to have a common ancestral origin; yet in the latter two domains, substantial increases in subunit complexity have occurred. It is possible that the two-subunit system in *B. subtilis* has arisen in a way that parallels the subunit amplification occurring in archaea and eukaryotes and could, in principal, arise in any bacteria.

We established that the heterologous host *E. coli* provides a reasonable genetic background in which to study SigO-RsoA function, even though all mutant phenotypes identified in *E. coli* must subsequently be tested in the native host, *B. subtilis*.

The SigO-RsoA system is activated in *B. subtilis* during acid stress. The ability to reconstitute these proteins and a recently characterized cognate anti- σ factor (RsiO, manuscript in preparation) in this heterologous host will help establish whether other currently unknown *B. subtilis* factors are important for this pH-related activation. In addition, although *B. subtilis* is genetically versatile, some genetic manipulations and the search for important mutations can be facilitated by employing *E. coli* as a genetic host. Indeed, detecting protein-protein interaction defects using a BACTH assay, for example, must be conducted in an *E. coli* background, as no comparable Gram positive system is currently available. The results of these present experiments indicate that, in general, *E. coli* provides a cogent model system for facilitating investigation of this unusual two-subunit σ factor.

Acknowledgments

This work was funded by an NSERC grant to S.R.M. M.C.D. and C.A.K. contributed equally to this work.

References

- Allaway, D., Schofield, N. A., Leonard, M.E., Gilardoni, L., Finan, T.M., and Poole, P.S. 2001. Use of differential fluorescence induction and optical trapping to isolate environmentally induced genes. *Environ. Microbiol.* **3**(6): 397–406. doi: 10.1046/j.1462-2920.2001.00205.x.
- Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R.R., and Landick, R. 2000. RNA Polymerases from *Bacillus subtilis* and *Escherichia coli* Differ in Recognition of Regulatory Signals In Vitro. *J. Bacteriol.* **182**(21): 6027–6035. doi: 10.1128/JB.182.21.6027-6035.2000.

- 332 Burgess, R.R., and Anthony, L. 2001. How sigma docks to RNA polymerase and what sigma does.
333 Curr. Opin. Microbiol. **4**(2): 126–31.
- 334 Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. 1969. Factor stimulating transcription by
335 RNA polymerase. Nature **221**(5175): 43–6.
- 336 deHaseth, P.L., and Helmann, J.D. 1995. Open complex formation by *Escherichia coli* RNA
337 polymerase: the mechanism of polymerase-induced strand separation of double helical
338 DNA. Mol. Microbiol. **16**(5): 817–24.
- 339 Geszvain, K., Gruber, T.M., Mooney, R.A., Gross, C.A., and Landick, R. 2004. A hydrophobic
340 patch on the flap-tip helix of *E.coli* RNA polymerase mediates sigma(70) region 4 function.
341 J. Mol. Biol. **343**(3): 569–87. doi: 10.1016/j.jmb.2004.08.063.
- 342 Guérout-Fleury, A.M., Frandsen, N., and Stragier, P. 1996. Plasmids for ectopic integration in
343 *Bacillus subtilis*. Gene **180**(1-2): 57–61. doi: 10.1016/S0378-1119(96)00404-0.
- 344 Hachmann, A.-B., Angert, E.R., and Helmann, J.D. 2009. Genetic analysis of factors affecting
345 susceptibility of *Bacillus subtilis* to daptomycin. Antimicrob. Agents Chemother. **53**(4):
346 1598–609. doi: 10.1128/AAC.01329-08.
- 347 Helmann, J.D. 2002. The extracytoplasmic function (ECF) sigma factors. Adv. Microb. Physiol.
348 **46**: 47–110.
- 349 Juang, Y.L., and Helmann, J.D. 1994. A promoter melting region in the primary sigma factor of
350 *Bacillus subtilis*. Identification of functionally important aromatic amino acids. J. Mol. Biol.
351 **235**: 1470–1488. doi: S0022-2836(84)71102-8 [pii]\n10.1006/jmbi.1994.1102 [doi].
- 352 Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. 1998. A bacterial two-hybrid system based
353 on a reconstituted signal transduction pathway. Proc. Natl. Acad. Sci. U. S. A. **95**(10): 5752–
354 6.
- 355 Lane, W.J., and Darst, S.A. 2006. The structural basis for promoter -35 element recognition by
356 the group IV σ factors. PLoS Biol. **4**: 1491–1500. doi: 10.1371/journal.pbio.0040269.
- 357 Lane, W.J., and Darst, S.A. 2010. Molecular Evolution of Multisubunit RNA Polymerases:
358 Sequence Analysis. J. Mol. Biol. **395**: 671–685. doi: 10.1016/j.jmb.2009.10.062.
- 359 Lukeš, J., Archibald, J.M., Keeling, P.J., Doolittle, W.F., and Gray, M.W. 2011. How a neutral
360 evolutionary ratchet can build cellular complexity. IUBMB Life **63**(7): 528–37. doi:
361 10.1002/iub.489.

- 362 MacLellan, S.R., Guariglia-Oropeza, V., Gaballa, A., and Helmann, J.D. 2009a. A two-subunit
363 bacterial sigma-factor activates transcription in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S.
364 A. **106**: 21323–21328. doi: 10.1073/pnas.0910006106.
- 365 MacLellan, S.R., Helmann, J.D., and Antelmann, H. 2009b. The Yvrl alternative σ factor is
366 essential for acid stress induction of oxalate decarboxylase in *Bacillus subtilis*. J. Bacteriol.
367 **191**: 931–939. doi: 10.1128/JB.01435-08.
- 368 MacLellan, S.R., Wecke, T., and Helmann, J.D. 2008. A previously unidentified σ factor and two
369 accessory proteins regulate oxalate decarboxylase expression in *Bacillus subtilis*. Mol.
370 Microbiol. **69**: 954–967. doi: 10.1111/j.1365-2958.2008.06331.x.
- 371 Miller, H. 1987. Practical aspects of preparing phage and plasmid DNA: growth, maintenance,
372 and storage of bacteria and bacteriophage. Methods Enzymol. **152**: 145–70.
- 373 Missiakas, D., and Raina, S. 1998. The extracytoplasmic function sigma factors: role and
374 regulation. Mol. Microbiol. **28**(6): 1059–66.
- 375 Stoltzfus, A. 1999. On the possibility of constructive neutral evolution. J. Mol. Evol. **49**(2): 169–
376 181. doi: 10.1007/PL00006540.
- 377 Werner, F. 2008. Structural evolution of multisubunit RNA polymerases. doi:
378 10.1016/j.tim.2008.03.008.
- 379 Werner, F., and Grohmann, D. 2011. Evolution of multisubunit RNA polymerases in the three
380 domains of life. Nat. Rev. Microbiol. **9**(2): 85–98. doi: 10.1038/nrmicro2507.
- 381 Zhou, Y.N., and Jin, D.J. 1998. The *rpoB* mutants destabilizing initiation complexes at stringently
382 controlled promoters behave like “stringent” RNA polymerases in *Escherichia coli*. Proc.
383 Natl. Acad. Sci. U. S. A. **95**(6): 2908–2913. doi: 10.1073/pnas.95.6.2908.

384

385

386

387

388

389

390

391

392

393 **Figure Legends**

394

395 Figure 1. The *B. subtilis* SigO-RsoA two-subunit σ factor. A) Organization of the SigO-RsoA
396 regulon. SigO-RsoA autoregulate their own transcription, the expression of *yvrJ* (unknown
397 function), *oxdC* (oxalate decarboxylase) and *rsiO* (the putative cognate anti- σ factor) from three
398 related promoters. B) Current model of SigO-RsoA association with core RNA polymerase.

399

400 Figure 2. SigO and RsoA co-activate transcription in *E. coli*. Induction of co-expression of SigO
401 and RsoA from pET-DUET, but not each protein in isolation, activates transcription from a P_{oxdC} -
402 *gfp* transcriptional fusion when cells are grown on M9 minimal media at 22°C. This is a
403 composite colour image of GFP fluorescence under UV light.

404

405 Figure 3. Sequence conservation amongst SigO-RsoA target promoters. For each sequence logo
406 representation, *oxdC-rsiO* (P_{oxdC}), *sigO-rsoA* (P_{sigO}), and *yvrJ* (P_{yvrJ}) known or predicted promoters
407 were identified from non-redundant genome sequences and manually aligned. 168 promoters
408 were aligned to generate a consensus (Cons) SigO-RsoA target promoter sequence. The -10 and
409 -35 nucleotides are identified relative to the experimentally deduced transcription +1 site
410 (MacLellan et al. (2008)). Bacterial names and genome accession numbers used as source of
411 promoter sequences are listed as supplemental information (Figure S2).

412

413 Figure 4. *B. subtilis* *oxdC* promoter analysis. A) Sequence of *B. subtilis* SigO-RsoA target -10
414 elements. B) Activity of *B. subtilis* P_{oxdC} -10 element promoter mutants in *B. subtilis*

(open bars) and *E. coli* (grey bars) genetic backgrounds. Activity reported relative to respective wt promoter activities (*B. subtilis*, 117.6 +/- 15.1 Miller units; *E. coli*, 406.1 +/- 66.5 Miller units; mean +/- sd, n=5).

Figure 5. Comparison of the effects of RsoA mutations that impair SigO-RsoA interaction on transcription activity in *B. subtilis* and *E. coli*. wt RsoA and RsoA point mutants that impair the SigO-RsoA protein-protein interaction (manuscript in preparation) assessed for ability to activate transcription from a P_{oxdC} -*lacZ* transcriptional fusion. Activity reported relative to respective wt promoter activities (*B. subtilis*, 733.0 +/- 69.4 Miller units; *E. coli*, 2402.8 +/- 223.5 Miller units; mean +/- sd, n=3).

Figure 6. Identification of mutations in SigO that impair its interaction with RsoA using an adenylate cyclase-based bacterial two hybrid assay (Karimova et al. 1998). An alignment-guided approach was used to identify conserved amino acid residues in the N-terminus of SigO for targeted site-directed mutagenesis. Activity reported relative to wt SigO-RsoA interaction (4414.2 +/- 246.2 Miller units; mean +/- sd, n=3).

Figure 7. Comparison of the effects of SigO mutations that impair SigO-RsoA interaction on transcription activity in *B. subtilis* and *E. coli*. wt SigO and SigO point mutants shown (Fig. 6) to impair the SigO-RsoA protein-protein interaction assessed for ability to activate transcription from a P_{oxdC} -*lacZ* transcriptional fusion. Activity reported relative to respective wt promoter

436 activities (*B. subtilis*, 768.6 +/- 19.6 Miller units; *E. coli*, 1406.5 +/- 341.9 Miller units; mean +/-
437 sd, n=3).

438

Draft

Fig.1

Fig.2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

Fig. 7