The *Escherichia coli* RhaS Transcriptional Activator: Transcriptional Activation by the DNA-Binding Domain, The Interdomain Effector Response, and Negative Autoregulation.

By

Jeffrey M. Skredenske

Submitted to the graduate degree program in Molecular Biosciences and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

_____

Dr. Susan M. Egan, Chairperson

_____

Dr. Stephen H. Benedict

_____

Dr. P. Scott Hefty

Date Defended: March 30, 2012

The Thesis Committee for Jeffrey M. Skredenske

certifies that this is the approved version of the following thesis:

The *Escherichia coli* RhaS Transcriptional Activator: Transcriptional Activation by the DNA-Binding Domain, The Interdomain Effector Response, and Negative Autoregulation.

_____

Dr. Susan M. Egan, Chairperson

Date approved: February 22, 2012

ABSTRACT

The chapters herein are the accepted manuscript versions of articles that were published independently in scholarly research journals. They have been combined and submitted in fulfillment of the thesis requirement for a Master of Arts degree in Microbiology from the University of Kansas Department of Molecular Biosciences.

In addition to the work presented here, my graduate work included two additional projects: a high-throughput screen to identify inhibitors of the *Escherichia coli* RhaS protein, and a site-directed mutagenesis screen to better understand the molecular mechanisms of the L-rhamnose response in RhaS. The high-throughput screen identified a compound that inhibits DNA binding by RhaS, the related *E. coli* RhaR protein and the virulence activators Rns from enterotoxigenic *E. coli* and VirF from *Shigella flexneri*, but by neither *E. coli* LacI nor CRP. It appears that this compound may have broad, specific inhibitory activity against AraC-family proteins, making it a candidate for development into an antimicrobial drug that functions by blocking the expression of certain bacterial virulence factors that require an AraC-family activator for expression. The compound likely binds in a pocket between the two helix-turn-helix motifs of the conserved AraC-family DNA-binding domain, thereby sterically prohibiting the protein from binding DNA.

In order to better understand the molecular mechanism by which the L-rhamnose signal is transmitted through RhaS from the N-terminal effector-binding domain to the C-terminal DNA-binding domain to regulate DNA binding in response to effector, I constructed a library of several dozen site-directed RhaS mutants. The goal of this work was to identify amino acids key to interdomain signaling by identifying point mutants with phenotypes consistent with defects in signaling. I focused my mutagenesis on regions of the protein predicted to be important in signaling, based on molecular modeling and similarities with related proteins. I isolated mutants

in the DNA-binding domain with nearly wild-type activity (-)L-rhamnose and reduced activity (+)L-rhamnose, consistent with a decreased ability to stimulate activity (+)L-rhamnose, at positions Asn174 and Leu175. We conclude that these two residues are likely important in the signal transduction pathway; future work will identify the region of the N-terminal domain involved in this interaction.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my graduate advisor, Dr. Susan Egan. From her I have learned not only literally everything I know about bacterial genetics, but also how to be a better writer, a better speaker, and a better student. I am incredibly grateful for the knowledge and experience I have gained by working in her laboratory. I also thank my advisory committee members, past and present: Dr. Steve Benedict, Dr. Scott Hefty, Dr. David Davido, Dr. Craig Martin, and Dr. William Picking. They have all given me advice and direction that have been indispensible throughout my time here.

My time and accomplishments here would not have been possible without the support of the great friends I have made over the years. My lab mates and neighbors have provided me with useful scientific advice and discussion, and also with countless memories from lunchtime, football games, lazy weekends and nights out. Ana, Vinitha, Bria, Swamy, Veer, Gurpreet, Jiaqin, Kelly, Ichie, Lindsay, Lindsey, Vinidhra, Jason… you will all be missed and will go on to do great things in life. I'll always treasure the good times I've spent with all of you.

I thank my family for the continuous love and support they have given me over the years. Mom, Dad, Kristin and Greg, you are all amazing people to me and I can't imagine life without you. It means so much to know that the people close to me in life are behind me no matter what I decide to do. And, finally, James, thank you for being the person that you are. The advice and support you have given me both in and out of the lab has helped me grow, develop myself and strive to be a better person. You are an incredible role model and a living testament that anyone can accomplish anything they want in life, no matter what challenges they face, as long as they commit themselves to the hard work and dedication it takes to achieve any given goal. You're my best friend, and I value that more than anything.

TABLE OF CONTENTS

CHAPTER 1

Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of its Effector-Binding Domain[†]

Jason R. Wickstrum[1#], Jeff M. Skredenske[1], Ana Kolin[1§], Ding J. Jin[2], Jianwen Fang[3] and Susan M. Egan[1]

[1]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA

[2]Transcription Control Section, Gene Regulation and Chromosome Biology Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, NIH, Bldg. 469, PO Box B, Frederick, MD 21702, USA

[3] Bioinformatics Core Facility, University of Kansas, Lawrence, KS 66045, USA

[#] Current Address: Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, KS 66160

[§]Current Address: Department of Microbiology and Immunology, 4301 Jones Bridge Road, Uniformed Services University of the Health Sciences, Bethesda, MD 20814

[†]This article is the accepted manuscript version of a subsequently published paper:

Wickstrum, J.R., J.M. Skredenske, A. Kolin, D.J. Jin, J. Fang and S.M. Egan. 2007. Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of its Effector-Binding Domain. *J. Bacteriol.* **189**:4984-4993.

The published version of this article is available online at:

http://dx.doi.org/10.1128/JB.00530-07.

ABSTRACT

The *Escherichia coli* L-rhamnose-responsive transcription activators RhaS and RhaR both consist of two domains, a C-terminal DNA-binding domain and an N-terminal dimerization domain. Both function as dimers and only activate transcription in the presence of L-rhamnose. Here, we examined the ability of the DNA binding domains of RhaS (RhaS-CTD) and RhaR (RhaR-CTD) to bind to DNA and activate transcription. RhaS-CTD and RhaR-CTD were both shown by DNase I footprinting to be capable of binding specifically to the appropriate DNA sites. *In vivo* as well as *in vitro* transcription assays showed that RhaS-CTD could activate transcription to high levels, whereas RhaR-CTD was capable of only very low levels of transcription activation. As expected, RhaS-CTD did not require the presence of L-rhamnose to activate transcription. The upstream half-site at rhaBAD and the downstream half-site at *rhaT* were found to be the strongest of the known RhaS half-sites, and a new putative RhaS half-site with comparable strength to known sites was identified. Given that CRP, the second activator required for full *rhaBAD* expression, cannot activate *rhaBAD* expression in a Δ*rhaS* strain, it was of interest to test whether CRP could activate transcription in combination with RhaS-CTD. We found that RhaS-CTD allowed significant activation by CRP, both *in vivo* and *in vitro*, although full-length RhaS allowed somewhat greater CRP activation. We conclude that RhaS-CTD contains all of the determinants necessary for transcription activation by RhaS.

INTRODUCTION

The RhaS protein functions to activate transcription of two of the operons in the *Escherichia coli* L-rhamnose regulon in response to the availability of L-rhamnose (11, 40). The two operons, *rhaBAD* and *rhaT*, encode the L-rhamnose catabolic enzymes (L-rhamnulokinase, L-rhamnose isomerase, and L-rhamnulose-1-phosphate aldolase) (2, 25) and an L-rhamnose-proton symporter that is responsible for transporting L-rhamnose into the cell (35), respectively. RhaS is encoded in an operon that also encodes a second L-rhamnose responsive transcription activator, RhaR (37). RhaR activates transcription of the operon that encodes the two activator proteins, *rhaSR* (37, 38). All three operons in the L-rhamnose regulon also require a second activator protein, CRP (cyclic AMP receptor protein), for full transcription activation (11, 16, 40).

RhaS and RhaR are both members of the AraC/XylS family of transcription activators. This very large family of transcription activators is defined by sequence similarity in a 100 amino acid region (10, 13). In all studied AraC/XylS family members, this 100 amino acid region functions as a DNA binding domain, and is referred to here as the AraC/XylS family domain. Most family members contain one or more domains in addition to the AraC/XylS family domain, but a few family members consist only of this single domain, such as MarA and SoxS (13). AraC and XylS, the namesakes of the family, are examples of two-domain family members in which the non-family domain functions in both effector binding and dimerization (7, 19, 33).

Detailed molecular structures have been determined for the DNA-binding domains of two AraC/XylS family members, MarA and Rob (20, 28). MarA and Rob share particularly high sequence similarity, 51%, and the structures of their DNA-binding domains (the only domain of MarA) are nearly identical (with a root mean square deviation of 0.9 Å) (20). The DNA-binding

domain of AraC/XylS family members contains two helix-turn-helix (HTH) DNA-binding motifs that contact consecutive major grooves of the DNA (20, 28). As a consequence, the binding site for each monomer (referred to as a half-site for dimers) is at least 17 bp long (13). In addition to DNA-binding, the AraC/XylS family domain of a number of family members has been shown to be involved in transcription activation, making contacts with the C-terminal domain of the alpha subunit of RNAP (α-CTD), the σ^{70} subunit of RNAP, or both (reviewed in (21)).

Although membership in the AraC family is defined by sequence similarity within a single domain, RhaS and RhaR share amino acid sequence identity with each other, as well as with AraC, over their entire lengths. All three proteins are therefore predicted to have similar three-dimensional structures for both of their domains. The N-terminal RhaS and RhaR domains function in both ligand binding and dimerization (Kolin and Egan, unpublished results), while the C-terminal domains are responsible for both DNA binding and direct contact with RNA polymerase (RNAP) to activate transcription (4, 5, 42). We have previously identified several amino acid-base pair contacts that are involved in DNA-binding by RhaS at *rhaBAD* (4). We have also identified two residues in RhaS and one in RhaR that are required to contact the σ^{70} subunit of RNAP to activate transcription, and further have identified the residues in σ^{70} that each of these activator residues contacts (5, 42). Interestingly, the RhaS and RhaR residues involved in these contacts with σ^{70} are all located in one of the H-T-H motifs of the proteins.

Among AraC/XylS family proteins that consist of more than one domain, it is interesting that there have been a variety of findings regarding whether the DNA-binding domain alone is sufficient to activate transcription (7, 18, 19, 24, 26, 36). One example is the *Pseudomonas putida* activator XylS (XylS-DN209). When overexpressed to sufficiently high levels, XylS-DN209 can activate transcription of the TOL plasmid *Pm* promoter to the same high level as full-

length XylS (19). Interestingly, at this high level of expression, full-length XylS becomes independent of its effector, activating to the same high levels in the absence and the presence of ligand. Another example is the DNA-binding domain of AraC. This domain alone could activate transcription of *araBAD* up to 15% as well as full-length AraC when alone, or to the same level as full-length AraC when fused to an unrelated dimerization domain (7, 36). In contrast, the DNA-binding domain of the MelR protein (MelR173) is unable to activate transcription either at the wild type target promoter (p*melAB*) or at promoters where the promoter-proximal MelR half-site is improved (18) (and S. Busby, personal communication).

In the current study, we tested the C-terminal AraC/XylS family domains of RhaS and RhaR for their ability to bind DNA and activate transcription in the absence of their N-terminal domains. We found that, while RhaS-CTD was able to activate transcription to high levels, RhaR-CTD could only activate to very low levels. DNase I footprinting indicated that both purified RhaS-CTD and RhaR-CTD were able to bind to DNA at their respective binding sites. Comparison of all of the RhaS half-sites showed that $rhaI_1$ was the strongest site, and that RhaS-CTD and full-length RhaS had similar profiles for binding to the different half-sites. Finally, we demonstrated the ability of RhaS-CTD to activate transcription *in vitro*, and further that CRP was capable of significant *in vitro* activation in combination with RhaS-CTD.

MATERIALS AND METHODS

Culture media and conditions. *E. coli* cultures for β-galactosidase assays were grown in MOPS buffered minimal medium using the protocol developed by Neidhardt *et al* (4, 23). Tryptone broth (0.8% tryptone; 0.5% NaCl; pH 7.0) was used to grow cultures in preparation for phage infection or transduction. $CaCl_2$ was added (final concentration 5 mM) to tryptone broth cultures used for bacteriophage P1 infection or transduction and maltose was added (final concentration 0.2%) to tryptone broth cultures used for bacteriophage λ infection or transduction. Tryptone-yeast extract (TY) liquid medium (0.8% tryptone; 0.5% yeast extract; and 0.5% NaCl; pH 7.0) was used to grow cells for most other experiments. All cultures were grown at 37°C. Antibiotics were used as indicated at the following concentrations: ampicillin (200 μg/ml), chloramphenicol (25 μg/ml), kanamycin (25 μg/ml), and tetracycline (20 μg/ml).

General Methods, Strains and plasmids. Standard methods were used for restriction endonuclease digestion and ligation. Oligonucleotides synthesized for this study were synthesized by MWG-Biotech (High Point, NC). A list of oligos used in this study is available at http://www.molecularbiosciences.ku.edu/faculty/egan.html. The Expand High Fidelity PCR System (Roche; Indianapolis, IN) was used to amplify DNA fragments for cloning as well as to generate template DNA for sequencing of genes that were recombined onto the chromosome. DNA sequencing was performed at the Molecular Research Core Facility at Idaho State University. The DNA sequence of both strands was determined for the entire cloned region of all cloned, mutagenized, and recombined DNA fragments.

Table 1 contains the list of strains and plasmids used in this study. All strains used in β-galactosidase assays were derived from ECL116 (1). All *lacZ* fusions used in this study were translational fusions with the exception of the Φ(*rhaT-lacZ*)Δ84 fusion, which was a transcriptional fusion. The *lacZ* fusions are named such that "Φ" stands for fusion and the

TABLE 1. Strains and plasmids used in this study.

Strain or plasmid	Genotype	Source or reference
E. coli strains		
BL21(DE3)	F$^-$ *ompT gal dcm lon hsdS$_B$* λDE3	Novagen
ECL116	F$^-$ Δ*lacU169 endA hsdR thi*	(1)
SME1048	ECL116 *recA::cat*	Laboratory collection
SME1051	ECL116 Δ*rhaSR::kan*	Laboratory collection
SME2986	ECL116 λΦ(*rhaB-lacZ*)Δ84 Δ*rhaSR::kan recA::cat*	This study
SME2999	ECL116 λΦ(*rhaS-lacZ*)Δ85 Δ*rhaSR::kan*	This study
SME3000	ECL116 λΦ(*rhaB-lacZ*)Δ84 Δ*rhaSR::kan*	This study
SME3066	ECL116 Δ*rhaRSBAD zih*-35::Tn*10*	This study
SME3072	ECL116 λΦ(*rhaB-lacZ*)Δ66 Δ*rhaSR::kan*	This study
SME3089	ECL116 λΦ(*rhaT-lacZ*)Δ84 Δ*rhaSR::kan*	This study
SME3114	ECL116 λΦ(*rhaB-lacZ*)Δ110 Δ*rhaSR::kan recA::cat*	This study
Plasmids		
pBluescript II SK	Ap^r *lacZ*α	Stratagene
pET15b	Ap^r *lacI* (ColE1 origin from pBR322)	Novagen
pHG165	*lacZα rop* Ap^r (ColE1 origin from pBR322)	(34)
pRS414	*lac'ZYA*	(32)
pSE101	pTZ 18R Ap^r *'rhaTrhaSRrhaBA'*	Laboratory collection
pSE227	pET15b *rhaR*196-312 (encodes N-terminal His6-tagged RhaR residues 196 through 312)	This study
pSE230	pET15b *rhaS*163-278 (encodes N-terminal His6-tagged RhaS residues 163 through 278)	This study

pSE250	pUC18 *rhaSRrhaT'* wild-type	This study
pSE262	pHG165 + pSR^{con} promoter	This study
pSE265	pSE262 *rhaS*	This study
pSE268	pSE262 *rhaS*163-278 (encodes RhaS-CTD)	This study
pSE271	pSU18 *rhaS*163-278 (encodes His6-tagged RhaS-CTD from *lac* promoter)	This study
pSE272	pSU18 *rhaR*196-312 (encodes His6-tagged RhaR-CTD from *lac* promoter)	This study
pSE273	pSU18 *rhaS*	This study
pSE274	pSU18 *rhaS*163-278 (encodes RhaR-CTD)	This study
pSE276	pRS414 Φ(*rhaB-lacZ*)Δ66	This study
pSE283	pTS134 with *rhaBAD* Δ110 promoter replacing *rhaSR* promoter	This study
pSU18	*lacZa* Cm^r (P15A *ori*)	(3)
pTS134	pBluescript II SK *rhaSR* promoter	(44)
pUC18	Ap^r *lacZ*α	(47)

upstream endpoint of each fusion relative to the transcription start site (for example –84, but without the negative sign) is given after the "Δ". The *lacZ* translational fusions were initially constructed on the plasmid pRS414, while the transcriptional fusion was constructed on pRS415 (32). The *lacZ* fusions used in all experiments except those in Figure 5 were then recombined onto the genome of bacteriophage λ, and integrated into the bacterial chromosome as lysogens (32). Single copy λ lysogens were identified using β-galactosidase assays and then confirmed using the Ter test (15). β-galactosidase assays were performed using the Miller method, as previously described (4, 22). Specific activities were averaged from at least three independent assays, with two replicates in each assay. In all β-galactosidase assays, error was less than 20% of the average values.

Construction of plasmids for overexpression and purification of His6-RhaS-CTD and His6-RhaR-CTD proteins. The N-terminal His6-tagged versions of RhaR-CTD and RhaS-CTD were expressed from pSE227 and pSE230, respectively. The RhaR and RhaS-coding regions of pSE227 and pSE230 were amplified by PCR using pSE101 as the template and the following oligonucleotides: 2345 and 2346 for *rhaR*; and 2349 and 2350 for *rhaS*. The PCR-amplified DNA was then ligated to pET15b using the *Nde*I and *Bam*HI restriction sites such that the vector encoded N-terminal His6-tag was fused in-frame with the RhaS and RhaR-coding regions.

Overexpression and purification of His6-RhaS-CTD and His6-RhaR-CTD. All protein overexpression was performed in strain BL21(DE3) (Novagen). The cells were grown to an A_{600} of approximately 1.0, induced with 1 mM IPTG and incubated for an additional 3 hrs. After harvesting the cells, the cell pellets were resuspended in chromatography binding buffer (5 mM imidazole, 0.5 M sodium chloride, 20 mM Tris-HCl, pH 7.9) and sonicated. After sonication, the lysate was centrifuged to separate soluble proteins from insoluble proteins. At

this level of overexpression, the vast majority of the proteins were present in the insoluble, pellet fraction. Therefore, the pellet fractions from the sonication lysates were resuspended in chromatography binding buffer containing 6 M urea and rocked overnight at 4°C to solubilize the His6-RhaS-CTD and His6-RhaR-CTD proteins. The next day, the urea-containing suspensions were centrifuged to remove the remaining insoluble protein, and the supernatant fractions were loaded onto immobilized metal affinity chromatography columns made with Ni^{+}-charged Chelex 20 resin (Sigma) that had been pre-equilibrated with binding buffer containing 6 M urea. The columns were washed with five volumes of binding buffer (containing 6 M urea), and then with 3 volumes of wash buffer (60 mM imidazole, 0.5 M sodium chloride, 20 mM Tris-HCl, pH 7.9) containing 6 M urea. In order to allow refolding of the protein on the column, the columns were washed with three volumes of wash buffer without urea and the His6-tagged proteins were then eluted with three volumes of elution buffer (0.5 M imidazole, 0.5 M sodium chloride, 20 mM Tris-HCl, pH 7.9). CRP protein with no added His6-tag was also purified using immobilized metal affinity chromatography, using the procedure previously described (43).

Construction of RhaS, RhaS-CTD, and RhaR-CTD expression plasmids for *in vivo* experiments. In order to test the ability of His6-RhaS-CTD and His6-RhaR-CTD to activate transcription *in vivo*, the respective genes were subcloned from pSE230 and pSE227 into pSU18 to make pSE271 and pSE272, respectively (using primers 2453 and 2454 in both cases). The subcloning involved digesting the PCR products with *Eco*RI and *Hin*dIII and then ligating to similarly digested pSU18. In the resulting constructs, the *lac* promoter of pSU18 drives transcription, and the ribosome binding site from pET15b (which was subcloned along with the open reading frame from pSE230 and pSE227) drives translation. A non-His6-tagged version of RhaS-CTD (pSE274) in pSU18 and an equivalent version of full-length RhaS (pSE273) were also constructed by PCR amplification from pSE101, and adding a primer-encoded Shine-

Dalgarno sequence that was equivalent to the Shine-Dalgarno sequence in pET15b. The upstream primer was 2571 for RhaS and 2574 for RhaS-CTD, and the downstream primer was 2542 in both cases.

Plasmid pSE262 was constructed by adding a constitutive promoter to pHG165. The inserted promoter in pSE262 (pSR^{con}) is the core *rhaSR* promoter, except the –35 hexamer was changed so that it matches the consensus –35 hexamer sequence (therefore the -35 sequence is TTGACA, and the –10 sequence is TACTAT). Also, pSE262 has a Shine-Dalgarno sequence (GAAGGA) followed immediately by a *Bam*HI site. Placement of a translational start codon immediately downstream of the *Bam*HI site provides the correct spacing relative to the Shine-Dalgarno sequence for ribosome recognition. The *rhaS* gene and the gene encoding RhaS-CTD were amplified by PCR from pSE250 with primers 2731 or 2732 and 2542 and ligated to pSE262 to make pSE265 and pSE268, respectively.

Western blots to compare *in vivo* expression of His6-RhaS-CTD and His6-RhaR-CTD. To compare His6-RhaS-CTD expression to His6-RhaR-CTD expression *in vivo*, cultures of the strains used in β-galactosidase assays (Table 2) were grown to mid-log phase in tryptone broth (TB) with chloramphenicol. Cells were sedimented and resuspended in TB to identical cell densities (using A_{600}). 500 μL of each sample was sonicated and separated into soluble and insoluble fractions by centrifugation. The insoluble pellets were resuspended in 500 μL TB. Total protein concentration in each sample was determined by Bradford (6) protein assay (Bio-Rad, Hercules, CA). Soluble fractions were standardized to identical protein concentrations, as were insoluble fractions, with insoluble fractions generally containing approximately 10-fold less protein than soluble fractions (the same ratio found before the minor adjustments to standardize protein concentrations). The samples were then analyzed using western blots. Equal amounts of protein from the standardized His6-RhaS-CTD- and His6-RhaR-CTD-containing cell fractions

were loaded onto two 15% SDS-polyacrylamide gels, electrophoresed, and blotted onto a nitrocellulose membrane using standard procedures. We also loaded known amounts of either purified His6-RhaS-CTD or His6-RhaR-CTD on the gels to allow for quantification of His6-RhaS-CTD or His6-RhaR-CTD present in the soluble and insoluble fractions of each sample. We added total lysate from the vector-only sample (collected before fractionation) to the samples with purified protein to prevent differences in antibody-antigen binding due to the lack of other proteins in the sample. The primary antibodies (anti-RhaS and anti-RhaR) were custom-made polyclonal rabbit antibodies from Cocalico Biologicals (Reamstown, PA). Anti-RhaS antibody was pre-adsorbed against a lysate of strain SME3066 (Δ*rhaRSBAD)* in order to remove rabbit antibodies to other *E. coli* proteins. This pre-adsorption step was not necessary for anti-RhaR antibody. The Alexa Fluor 680-labeled secondary antibody (anti-rabbit) was obtained from Molecular Probes (Eugene, OR). The blots were imaged using an Odyssey Infrared Imaging System (LI-COR, Lincoln, NE).

DNase I footprinting. The template DNAs for DNase I footprinting were generated by PCR using the following templates and primers: *rhaBAD* was amplified from pSE101 using primers 2371 and 2410; *rhaT* was amplified from ECL116 cells using primers 2096 (^{32}P-labeled) and 2097 for one strand, and primers 2655 (^{32}P-labeled) and 2656 for the other strand; and *rhaSR* was amplified from pSE101 using primers 2371 and 2409. DNase I footprinting was performed as previously described (44). Gels were imaged by autoradiography. In addition to the results shown, similar results were obtained when the other DNA strand was labeled. All DNase I footprinting experiments were carried out at least twice.

Construction of *rhaI* half-site fusions on plasmids. The half-site fusions used to compare the strength of various RhaS DNA half-sites were constructed in the context of a Φ(*rhaB-lacZ*)Δ66 fusion (Fig. 1) in pRS414. At the wild-type *rhaBAD* promoter, the $rhaI_2$ half-

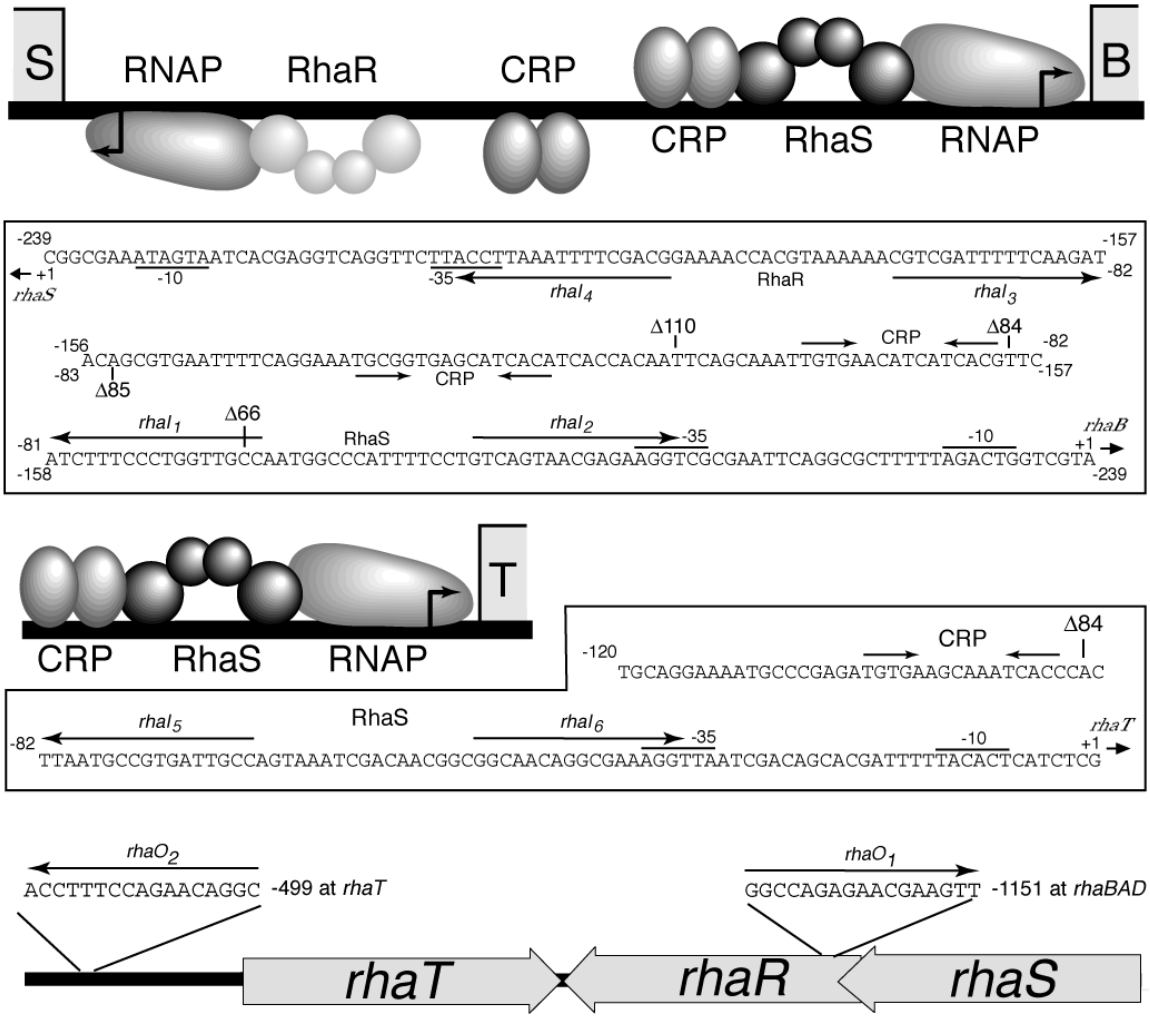

Figure 1. RhaS and RhaR binding sites within the L-rhamnose regulon. For each promoter (*rhaSR*, *rhaBAD*, and *rhaT*), there is a schematic showing the known proteins and binding sites and below each schematic is the DNA sequence of that promoter region. Within the *rhaSR-rhaBAD* intergenic region, labels for the *rhaSR* promoter are below the DNA sequence and labels for the *rhaBAD* promoter are above the DNA sequence. The upstream endpoints of the *lacZ* fusions used in this study (in bp upstream from the transcription start site) are indicated with a Δ symbol. The RhaS and RhaR binding sites are shown as half-sites and labeled with the half-site number ($rhaI_x$). The orientation of each half-site is indicated with an arrow. RhaS (dark gray) and RhaR (light gray) are depicted as dimers with each monomer consisting of two domains, a C-terminal DNA-binding domain and an N-terminal dimerization domain, depicted as spheres. The bottom figure shows the approximate positions of the *rhaO* half-sites (which are outside of the promoter regions) within the L-rhamnose region.

site overlaps the –35 hexamer of RNAP by 4 base pairs. The other RhaS half-sites were placed in the position of $rhaI_2$, however, the DNA sequence of the 4 base pair overlap with the –35 hexamer was not changed (Fig. 5C). (Constructs in which the –35 sequence was changed to match that of the $rhaI_1$, $rhaI_5$ and $rhaO_1$ half-site resulted in extremely low expression, unpublished results). The DNA sequence surrounding each half-site was identical in every case. The wild-type Φ(*rhaB-lacZ*)Δ66 fusion (pSE276) was created by PCR with oligonucleotides 2414 and 744, using pSE101 as the template. The other half-site fusions were constructed by PCR using an oligonucleotide-encoded half-site in the upstream primer (2413, 2441, 2442, 2445, and 2446) and 744 downstream, and the resulting plasmids were named pSE276 $rhaI_X$ (where "X" represents the half-site number).

***In vitro* transcription assays.** Single round *in vitro* transcription assays were carried out with core RNAP and σ^{70} purified as described (48, 49). Reconstitution of σ^{70} with core RNAP was carried out by mixing 4 µg core RNAP and 0.7 µg σ^{70} (1:1 molar ratio) in 100 µL of RNAP storage buffer (50 mM Tris-HCl, pH 8.0; 50% glycerol; 0.1 mM NaEDTA; 0.1 mM DTT; 50 mM NaCl) and incubating at 25°C for 10 minutes, then storing at –20°C. To prepare the transcription reaction, His6-RhaS-CTD and/or CRP was incubated with *rhaBAD* promoter template DNA (PCR amplified with oligonucleotides 744 and 2654) in IVT reaction buffer (final concentrations in reaction: 20mM Tris-HCl, pH 7.9; 50 mM KCl; 4 mM $MgCl_2$; 1 mM DTT; 0.1 mM KEDTA; 0.1 mg/mL BSA; 50 mM L-rhamnose; 0.2 mM cAMP) at 37°C for 10 minutes. RNAP was then added (10 nM final concentration) and the reaction was incubated for 5 minutes at 37°C. Initiation mix was added (final concentrations in reaction: 0.2 mM each of ATP, CTP, and GTP; 0.02 mM UTP; 100 mg/ml heparin; 0.2 µCi (α-^{32}P)UTP (3000 Ci/mmol). The reaction was next incubated at 37°C for 10 minutes, then stopped by addition of 0.25 volumes of stop solution (7 M urea; 0.1 M KEDTA; 0.4% SDS; 20 mM Tris-HCl, pH 7.9; 0.5% bromophenol

blue; 0.5% xylene cylanol). The reaction was then loaded directly onto a pre-heated 6% denaturing polyacrylamide gel for electrophoresis (0.3% *N,N*-methylenebisacrylamide, 8.9 mM Tris, 8.9 mM boric acid, 20 mM EDTA and 8 M urea). The gels were imaged and analyzed using a Cyclone Storage Phosphor System (PerkinElmer). The results shown are representative of three similar experiments.

RESULTS

***In vitro* DNA binding by His6-RhaS-CTD and His6-RhaR-CTD**. Specific residues in the C-terminal domains of RhaS and RhaR had previously been shown to contribute to DNA binding and transcription activation (4, 5, 38, 42). In order to test whether the C-terminal domain of each protein was sufficient for DNA binding and transcription activation we constructed plasmids expressing truncated versions of RhaS (encoding His6-RhaS-CTD, consisting of RhaS amino acids 163-278), and RhaR (encoding His6-RhaR-CTD, consisting of RhaR amino acids 196-312). We tested the *in vitro* DNA-binding activity of His6-RhaS-CTD and His6-RhaR-CTD by performing DNase I footprinting using the purified proteins. We found that His6-RhaS-CTD bound to the *rhaBAD* promoter region at two sites (Fig. 2A). The extent of the two footprinted regions corresponds very well with the two half-sites for RhaS binding previously predicted from the footprint of full-length RhaS and mutagenesis of the binding site region (12). There were two differences from the previously published footprints, however, both consistent with the prediction that His6-RhaS-CTD is monomeric, whereas full-length RhaS is dimeric. First, His6-RhaS-CTD did not protect the DNA between the two half-sites, while full-length RhaS did. Second, it was possible to observe differences in the binding strength to the two half-sites with His6-RhaS-CTD. We found that there were protein concentrations at which the footprint at $rhaI_1$ (the promoter distal half-site) was detectable while the footprint at $rhaI_2$ (the promoter proximal half-site) was no longer detectable, indicating that His6-RhaS-CTD bound more tightly to $rhaI_1$ than to $rhaI_2$ (Fig. 2A).

We also performed DNase I footprinting with His6-RhaS-CTD at the *rhaT* promoter. DNA sequence inspection and previous results indicating that RhaS was required for activation of *rhaT* expression (40) suggested that RhaS binds to DNA at the *rhaT* promoter, but direct evidence of RhaS protein binding to *rhaT* promoter DNA had not been obtained. Our DNase I

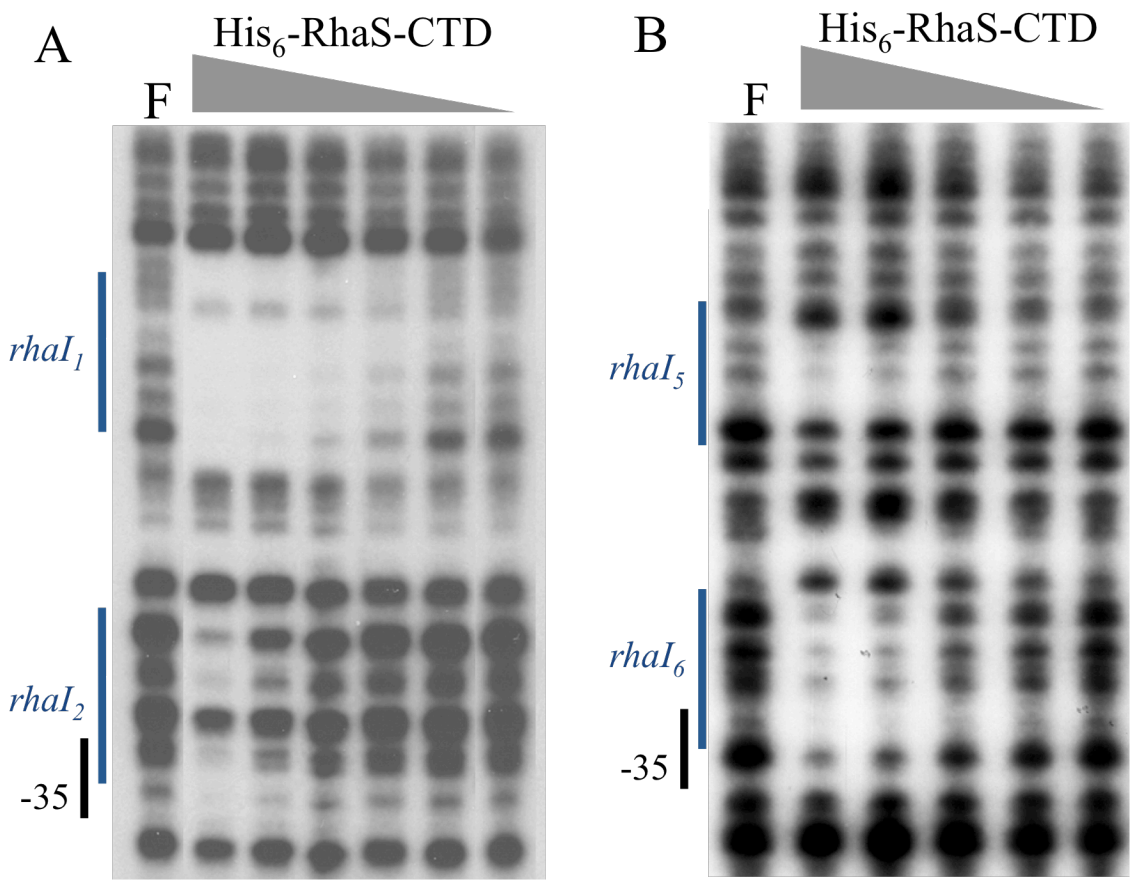

Figure 2. DNase I footprinting assay of His6-RhaS-CTD binding to the *rhaBAD* promoter (A) and the *rhaT* promoter (B). The DNA fragment used as the template for *rhaBAD* was generated by PCR with primers 2371 (^{32}P labeled) and 2410, while that for *rhaT* was generated with primers 2096 (^{32}P-labeled) and 2097. The positions of the RhaS half-sites were determined based on a DNA sequencing ladder (not shown). The highest His6-RhaS-CTD concentration was 6 μM and the dilution steps were 3-fold. F, free DNA.

footprinting results provide direct evidence of His6-RhaS-CTD binding to the predicted RhaS half-sites at the *rhaT* promoter (Fig. 2B). In this case, His6-RhaS-CTD appeared to have a slightly higher affinity for the $rhaI_6$ half-site (the promoter proximal half-site) than the $rhaI_5$ half-site (the promoter distal half-site). While the protection of $rhaI_5$ was weak, we were unable to use higher protein concentrations in this experiment due to the tendency of His6-RhaS-CTD to aggregate.

Finally, we tested *in vitro* DNA binding by His6-RhaR-CTD by DNase I footprinting. We found that His6-RhaR-CTD showed specific binding to two sites within *rhaSR* promoter DNA (Fig. 3). There was somewhat more ambiguity than usual in the exact extent of the protected regions in this case due to the relative lack of DNase I cleavage sites within the A-tracts, especially in the region of $rhaI_3$, however, the two protected regions appear to correspond well with the previously demonstrated RhaR half-sites (38, 44). There was not any substantial difference in the apparent strength of His6-RhaR-CTD binding to the two half-sites. These results indicate that our purified His6-RhaS-CTD and His6-RhaR-CTD protein preparations contained active proteins that were capable of specifically binding to DNA.

The C-terminal domain of RhaS but not RhaR is sufficient for transcription activation. His6-RhaS-CTD was tested for *in vivo* activation of *lacZ* fusions to the *rhaBAD* and *rhaT* promoters, while His6-RhaR-CTD was tested for *in vivo* activation of a *lacZ* fusion to the *rhaSR* promoter (Fig. 1). We found that plasmid-expressed His6-RhaS-CTD could activate transcription to high levels, with 1,000-fold-activation of the *rhaBAD* promoter and 180-fold activation of the *rhaT* promoter (Table 2). In contrast, plasmid-expressed His6-RhaR-CTD could only activate expression of the *rhaSR* promoter by two-fold (Table 2). Given that the fold-activation by full-length, chromosomally expressed RhaS at *rhaBAD* is approximately 33-fold higher than that of chromosomally expressed RhaR at *rhaSR*, comparable efficiencies of

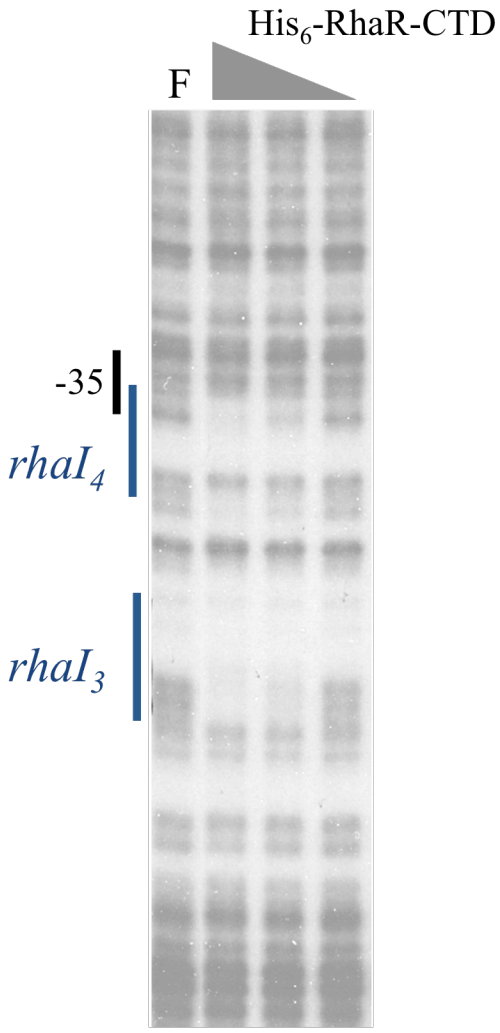

Figure 3. DNase I footprinting assay of His6-RhaR-CTD binding to the *rhaSR* promoter. The DNA fragment used as the template was generated by PCR with primers 2371 (^{32}P labeled) and 2409. The positions of the RhaR half-sites were determined based on a DNA sequencing ladder (not shown). The highest His6-RhaR-CTD concentration was 5 μM and the dilution steps were 3-fold. F, free DNA.

TABLE 2. Transcription activation by His6-RhaS-CTD and His6-RhaR-CTD

Promoter fusion[a]	β-galactosidase sp act[b]		His6-RhaS-CTD Fold Activation[d]
	Vector-only[c]	His6-RhaS-CTD[c]	
Φ(*rhaB-lacZ*)Δ84	0.021	21	1000
Φ(*rhaT-lacZ*)Δ84[e]	0.29	51	180
		His6-RhaR-CTD[c]	**His6-RhaR-CTD Fold Activation[d]**
Φ(*rhaS-lacZ*)Δ85	0.41	0.82	2.0

[a] Each strain carried a single copy *lacZ* fusion integrated into the chromosome as a λ lysogen, and also Δ(*rhaSR*).

[b] β-galactosidase activity was determined as described in Materials and Methods, and is expressed in Miller Units.

[c] Vector-only was pSU18, His6-RhaS-CTD was expressed from pSE271 and His6-RhaR-CTD was expressed from pSE272.

[d] Fold activation values were calculated by dividing the activity in the presence of His6-RhaS-CTD or His6-RhaR-CTD by the activity in the presence of vector only.

[e] Φ(*rhaT-lacZ*)Δ84 is a transcriptional fusion. All other *lacZ* fusions in this study were translational fusions.

activation by the C-terminal domains to their full-length counterparts would have resulted in His6-RhaR-CTD activating *rhaSR* by approximately 30-fold. This value is much higher than the two-fold value measured for His6-RhaR-CTD. To test whether the very poor activation might be due to an artifact of our His6-RhaR-CTD-expressing construct, we made a number of different RhaR-CTD constructs, however none were capable of significant transcription activation. Our results suggest that although His6-RhaR-CTD is capable of specific DNA binding, it is not capable of activating transcription well. Interestingly, Tobin and Schleif previously found that in the absence of L-rhamnose, full-length RhaR was able to bind to DNA, but not able to activate transcription (38, 39), suggesting that His6-RhaR-CTD (which lacks an L-rhamnose-binding domain) may be similar in its activity to full-length RhaR in the absence of L-rhamnose.

To test whether the low level of activation by His6-RhaR-CTD could be due to a low level of His6-RhaR-CTD protein expression or stability compared with the His6-RhaS-CTD protein, we performed western blots with samples of the same strains assayed in Table 2. We found that both soluble His6-RhaS-CTD (Fig. 4, top blot, third lane) and soluble His6-RhaR-CTD (Fig. 4, bottom blot, fifth lane) were present in substantial amounts, based on comparisons with known amounts of the respective purified proteins. However, soluble His6-RhaS-CTD (5.5 ng/μg of soluble protein) was present at a level approximately 5.5-fold higher than that of soluble His6-RhaR-CTD (1.0 ng/μg of soluble protein). This 5.5-fold difference in soluble protein levels could explain some of the decrease in activation by His6-RhaR-CTD at *rhaSR* relative to His6-RhaS-CTD at *rhaBAD*, however, the similarity in the activity of His6-RhaR-CTD to that of full-length RhaR in the absence of L-rhamnose suggests that His6-RhaR-CTD may simply be unable to activate transcription well. Our results also showed that His6-RhaS-CTD and His6-RhaR-CTD did not respond to L-rhamnose availability (data not shown), which

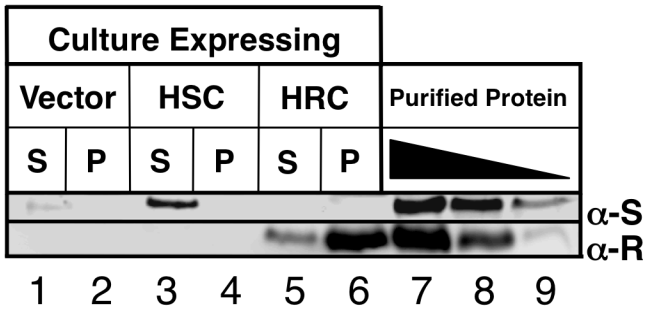

Figure 4. Western blots comparing *in vivo* expression of His6-RhaS-CTD and His6-RhaR-CTD. Soluble (S, supernatant) and insoluble (P, pellet) fractions after sonication were loaded as indicated. Vector only was pSU18, His6-RhaS-CTD (HSC) was expressed from pSE271, and His6-RhaR-CTD (HRC) was expressed from pSE272. Lanes 7-9 on each gel contained known amounts of purified His6-RhaS-CTD (top blot) or His6-RhaR-CTD (bottom blot). The His6-RhaS-CTD amounts were: 737 (lane 7), 368 (lane 8), 184 ng (lane 9). The His6-RhaR-CTD amounts were: 162 (lane 7), 54 (lane 8), 18 ng (lane 9). Two gels were prepared, with identical culture samples on each gel, and each blot was probed with the primary antibody corresponding to the purified protein samples loaded, as indicated to the right (α-S, anti-RhaS antibody; α-R, anti-RhaR antibody).

was expected since, based on sequence similarity with AraC, L-rhamnose binding is predicted to be a function of the RhaS and RhaR N-terminal domains.

Comparison of RhaS and RhaS-CTD activation of *rhaB-lacZ* fusions. The above results (Table 2) indicate that RhaS-CTD was capable of activating transcription from a *rhaB-lacZ* fusion that includes the full RhaS binding site. We next tested whether both or only one of the RhaS half-sites contributes to RhaS-CTD activation, and whether RhaS-CTD is sufficient to allow CRP protein to contribute to *rhaBAD* activation. Given that it lacks its dimerization domain, we predicted that RhaS-CTD would function as a monomer, similar to MarA (28). We expected that the RhaS-CTD monomer that bound to the half-site adjacent to RNAP would contribute to transcription activation based on our previous finding that RhaS contacts with σ^{70} contribute to transcription activation (5, 42). We also have some evidence that activation by RhaS may involve contacts with a-CTD (17), therefore it was possible that RhaS-CTD bound to the promoter distal RhaS half-site might further contribute to transcription activation. Previous *in vivo* experiments also indicate that CRP is not capable of activating *rhaBAD* expression in the absence of RhaS (11), perhaps suggesting that RhaS must bend the DNA to allow CRP to activate, although other possibilities exist as well. Therefore, we also tested whether RhaS-CTD was capable of fulfilling the function of RhaS that allows CRP activation.

To address these questions, we compared transcription activation by full-length RhaS and RhaS-CTD (no His6 tag) at three different truncations of the *rhaBAD* promoter (each fused to *lacZ* and carried as a single copy λ lysogen) (Fig. 1). At Φ(*rhaB-lacZ*)Δ66, which carries only one half-site of the full RhaS binding site, RhaS-CTD was capable of more than 2,000-fold activation (Table 3). Interestingly, this was a ten-fold higher level than similarly expressed full-length RhaS activated this same fusion. At the Φ(*rhaB-lacZ*)Δ84 fusion, which carries the full

TABLE 3. Transcription activation by RhaS-CTD compared to full-length RhaS.

Promoter Fusion[a]	β-galactosidase sp act[b]		Fold Activation with RhaS-CTD	β-galactosidase sp act[b]	Fold Activation with RhaS
	Vector Activity[c]	RhaS-CTD activity[c]		RhaS Activity[c]	
Φ(*rhaB-lacZ*)Δ66	0.015	34	2,300	3.5	230
Φ(*rhaB-lacZ*)Δ84	0.015	36	2,400	110	7,300
Φ(*rhaB-lacZ*)Δ110	0.018	91	5,100	700	39,000

[a] Each strain carried a single copy *lacZ* fusion integrated into the chromosome as a λ lysogen, and also Δ(*rhaSR*).

[b] β-Galactosidase activity was determined as described in Materials and Methods, and is expressed in Miller Units. All cultures were grown in the presence of L-rhamnose.

[c] Vector-only was pSE262, His6-RhaS-CTD was expressed from pSE268 and RhaS was expressed from pSE265.

[d] Fold activation values were calculated by dividing the activity in the presence of His6-RhaS-CTD or RhaS by the activity in the presence of vector only.

RhaS binding site, there was no increase in the activation by RhaS-CTD, indicating that RhaS-CTD activation occurs from the promoter proximal half-site, and consistent with the prediction that RhaS-CTD functions as a monomer. In contrast, full-length RhaS activated this fusion to a level more than 30-fold higher than its activation of the fusion containing only a single RhaS half-site, and three-fold higher than the activation by RhaS-CTD. Finally, at the Φ(*rhaB-lacZ*)Δ110 fusion, which contains the CRP site required for full *rhaBAD* activation as well as the full RhaS binding site, there was a 2-fold contribution to activation by CRP when in combination with RhaS-CTD, and a 5-fold contribution to activation by CRP when in combination with full-length RhaS. This suggests that RhaS-CTD can fulfill at least part of the function of RhaS that allows CRP activation at *rhaBAD*.

***In vivo* comparison of different RhaS half-sites**. The DNase I footprinting assays indicated that there were differences in the relative strength of the RhaS half-sites at the *rhaBAD* and *rhaT* promoters. In order to further compare RhaS binding to the RhaS half-sites, each half-site was placed at the same position in the context of the Φ(*rhaB-lacZ*)Δ66 promoter (referred to as "half-site fusions"). In these constructs, each of the half-sites replaces $rhaI_2$, the wild-type promoter proximal half-site at this promoter (Fig. 1). In addition to the four previously identified RhaS half-sites, we also tested two additional DNA sequences that were identified using a string-matching program (written in the computer language Perl) to identify potential RhaS half-sites within the entire *rha* region. The program identified only two DNA sequences with perfect matches in sequence and spacing to the six base pairs previously identified as most important for RhaS binding (12) (Fig. 5C). One of the potential half-sites ($rhaO_1$) is located within the *rhaR* gene, and is centered at -1153 relative to the *rhaBAD* transcription start site, or at +914 relative to the *rhaSR* transcription start site, while the other potential site ($rhaO_2$) is centered at -499 relative to the *rhaT* transcription start site.

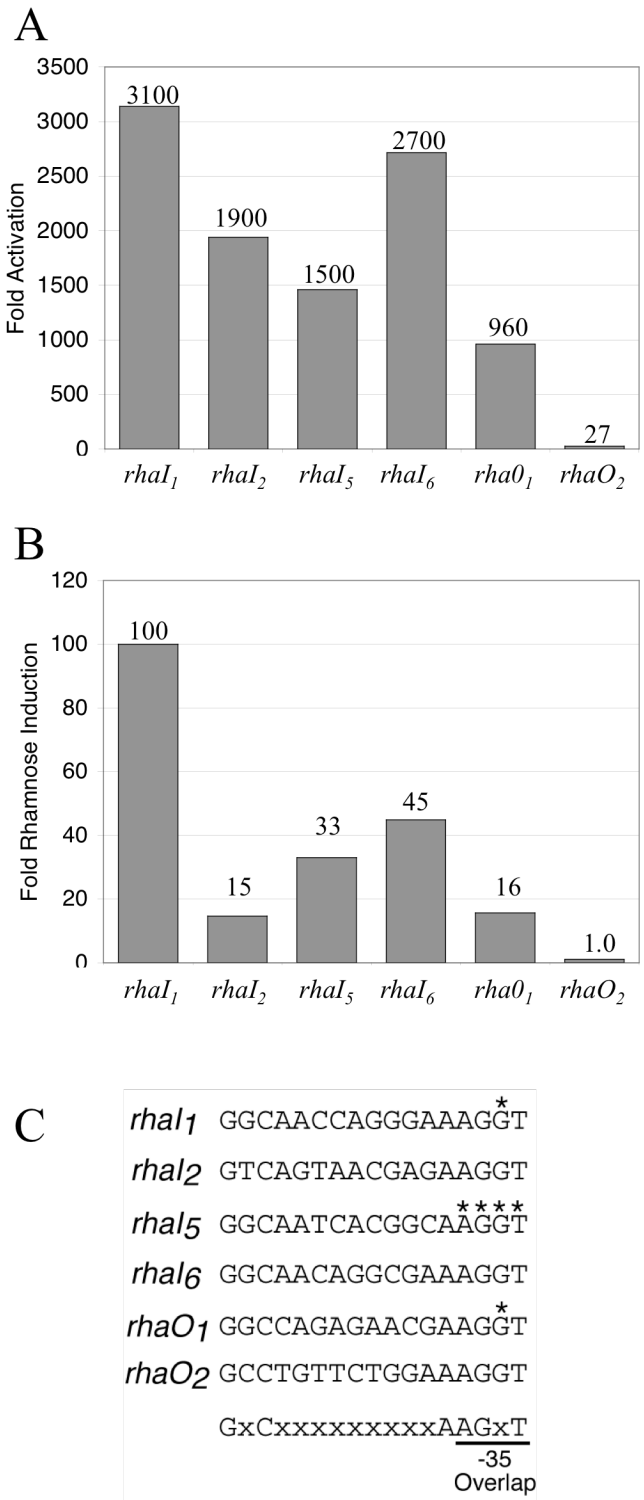

Figure 5. *In vivo* transcription activation by His6-RhaS-CTD or RhaS. The indicated RhaS half-sites, in the context of Φ(*rhaB-lacZ*)Δ66 on multicopy plasmids (pSE276 and derivatives), were assayed for β-galactosidase activity. Cells were grown in TY media. (A) Activation by His6-RhaS-CTD, expressed from pSE271 in SME1051 [Δ(*rhaSR*)]. Fold activation was determined by dividing the activity obtained with RhaS-CTD by the activity of pSU18 alone (vector only) at

(Figure 5, continued.) each promoter. The activity of pSU18 alone ranged from 0.39 to 0.99 Miller Units. (B) Transcription activation by RhaS expressed from the chromosome in SME1048 (wild-type *rhaSR*). Fold rhamnose induction was determined by dividing the activity of each fusion in the presence of rhamnose by the activity of the same fusion in the absence of rhamnose. The (-) rhamnose activities ranged from 0.21 to 0.29 Miller Units. (C) The DNA sequences of the half-sites used in these experiments are shown. The asterisks indicate positions at which base pairs were changed from the wild type half-site sequences (see Fig. 1) so that the DNA sequence of the overlapping –35 element was not changed. The bottom line indicates the sequence and position of the 6 important base pairs in the RhaS binding site that were used to identify $rhaO_1$ and $rhaO_2$.

We first tested the ability of each of the half-site fusions to be activated *in vivo* by His6-RhaS-CTD (Fig. 5A). Among the previously identified RhaS half-sites, we found the greatest fold-activation at $rhaI_1$, followed by $rhaI_6$, then $rhaI_2$ and finally $rhaI_5$. These results confirm the relative half-site strengths identified by DNase I footprinting, and further provide information about the relative strength of the *rhaBAD* vs. the *rhaT* half-sites. We also found that His6-RhaS-CTD could activate transcription from the fusion carrying the $rhaO_1$ half-site to an extent comparable with the previously identified half-sites, while there was only a very low level of activation from $rhaO_2$. The same order of relative half-site strength was also determined using electrophoretic mobility shift assays with purified His6-RhaS-CTD (data not shown). The ability of His6-RhaS-CTD to activate transcription to a high level from $rhaO_1$ confirms that the six base pairs used to identify this site are important for RhaS binding, however, the low level of activation with $rhaO_2$ suggests that the context of these six base pairs is also important. We also tested the same set of half-site fusions for activation by full-length RhaS expressed from the chromosome (Fig. 5B). We found a very similar order of apparent half-site strengths in this experiment, although the position of $rhaI_2$ in the order was different and the magnitude of activation by RhaS from these fusions was much lower than that by His6-RhaS-CTD. In this case, there was no activation from the $rhaO_2$ half-site, further supporting the idea that it is, at best, a marginal RhaS half-site.

Transcription activation by His6-RhaS-CTD *in vitro*. Given that RhaS-CTD (both with and without a His6-tag) was capable of activating transcription *in vivo*, and was capable of specifically binding to DNA at RhaS half-sites *in vitro*, we investigated its ability to activate transcription in a purified *in vitro* transcription assay. We have never been able to carry out i*n vitro* transcription with full-length RhaS due to its insolubility. In preliminary experiments we found that His6-RhaS-CTD activated transcription much more efficiently from linear DNA

templates than from supercoiled DNA templates (data not shown), in contrast to full-length RhaR, which required supercoiled DNA templates for efficient *in vitro* transcription activation (39, 44). We found that CRP protein alone was not capable of activating *rhaBAD* expression, similar to previous findings *in vivo*, and that His6-RhaS-CTD alone substantially activated *rhaBAD* expression (Fig. 6). We also found that CRP and His6-RhaS-CTD together could activate transcription to a level that was three-fold higher than that by His6-RhaS-CTD alone. The three-fold contribution to activation by CRP in this experiment is very similar to the two-fold contribution found in the above *in vivo* experiment (Table 3), confirming that RhaS-CTD is sufficient to allow at least partial CRP activation of *rhaBAD* expression. This *in vitro* transcription system mimics the *in vivo* activation of *rhaBAD* by RhaS-CTD, and also to a great extent, the *in vivo* activation of *rhaBAD* by full-length RhaS, each in the presence and absence of CRP. Consistent with the very low level of activation in our *in vivo* results, we were not able to detect transcription activation by His6-RhaR-CTD (data not shown).

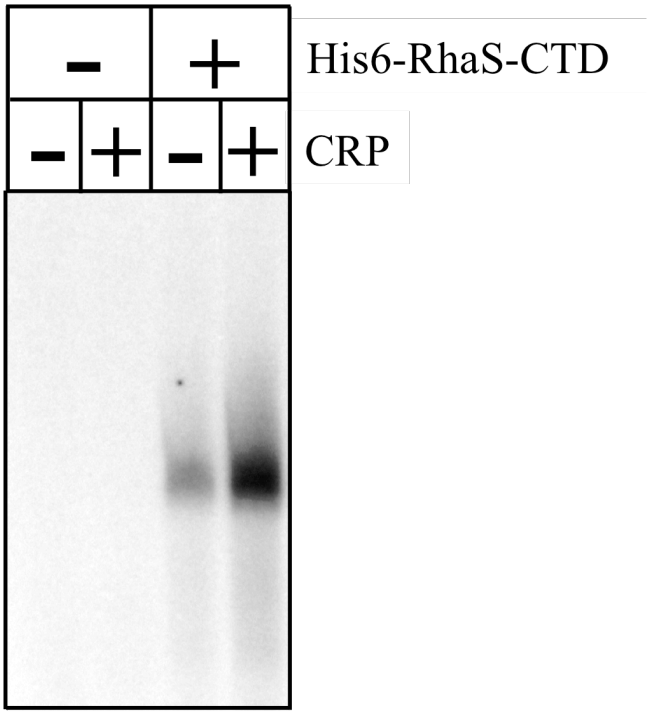

Figure 6. *In vitro* transcription activation by His6-RhaS-CTD. The linear template DNA containing the *rhaBAD* promoter was generated by PCR using primers 744 and 2654 with pSE283 as the template. When present (as indicated above the gel), the His6-RhaS-CTD concentration was 10 μM and the CRP concentration was 1μM. cAMP was present in all reactions.

DISCUSSION

***In vivo* transcription activation by RhaS-CTD**. Our results indicate that RhaS-CTD (with and without a His6-tag) could activate transcription very well, and that purified His6-RhaS-CTD protein was able to bind to DNA at the previously identified or predicted RhaS half-sites at *rhaBAD* and *rhaT*. Based on the amino acid sequence alignment of RhaS with AraC, as well as our studies of RhaS (Kolin and Egan, unpublished results), we predicted that the dimerization interface of RhaS would be located in its N-terminal domain. Several pieces of evidence in this study are consistent with the prediction that RhaS-CTD functions as a monomer. His6-RhaS-CTD did not footprint the DNA between the two RhaS half-sites, nor did binding by His6-RhaS-CTD exhibit much if any cooperative binding to the two half-sites at *rhaBAD* and *rhaT*, in both cases unlike full-length RhaS (12). The level of transcription activation by RhaS-CTD also did not increase with the addition of a second RhaS half-site, again unlike full-length RhaS.

We were also not surprised to find that RhaS-CTD was capable of equivalent transcription activation in the absence and the presence of L-rhamnose (data not shown), since amino acid sequence alignment with AraC suggests that the RhaS N-terminal domain likely binds L-rhamnose. The differential activation of transcription by full-length RhaS in the absence and presence of L-rhamnose could be due either to an inhibition of activity in the absence of ligand or a stimulation of activity in the presence of ligand. The finding that RhaS-CTD activates transcription very well in the absence of RhaS-NTD suggests that there may be an inhibition of full-length RhaS activity in the absence of ligand. The light-switch mechanism used by AraC to respond to its ligand arabinose also involves inhibition in the absence of ligand (14, 27, 29, 31, 45, 46). However, our more recent results suggest that the L-rhamnose response of

RhaS most likely involves an active stimulation of activity in the presence of L-rhamnose (Kolin and Egan, unpublished; see below).

***In vitro* DNA-binding and transcription activation by His6-RhaS-CTD.** We found that purified His6-RhaS-CTD was capable of specific DNA binding and also activation of transcription in purified *in vitro* reactions. Prior to this work, *in vitro* studies of the L-rhamnose regulon have been severely hampered by the strong tendency of full-length RhaS to aggregate. Our only previously published *in vitro* studies involving RhaS utilized partially purified protein that was refolded on a "per reaction" basis in the presence of specific DNA (12). We have never been able to obtain full-length RhaS that is both soluble and active by refolding of protein purified under denaturing conditions, nor by fusion with proteins that promote solubility. Therefore, our finding that His6-RhaS-CTD is capable of both specific DNA binding and activation of transcription *in vitro* represents a major breakthrough in our studies of the L-rhamnose regulon. While His6-RhaS-CTD is not entirely free of aggregation problems, its aggregation is substantially more manageable than that of full-length RhaS.

RhaS-CTD binding to *rhaI* half-sites. Our DNase I footprinting and *in vivo* transcription assays both indicated that His6-RhaS-CTD bound to the $rhaI_1$ half-site at *rhaBAD* significantly more strongly than to the $rhaI_2$ half-site. This finding is similar to previous findings with the AraC and MelR proteins in which the activator binds to its upstream half-site much more tightly than its downstream half-site (8, 9, 18, 31), and in the absence of ligand forms a DNA loop that represses transcription (30, 41). To begin to address the question of whether RhaS regulation might involve DNA-looping, we used a bioinformatics approach to look for other potential RhaS half-sites in the *rha* region. One of the two sites we identified, $rhaO_1$, is located within the *rhaR* gene at –1153 relative to the *rhaBAD* transcription start site or +914 relative to the *rhaSR* transcription start site. The finding that RhaS binds to $rhaO_1$ with an

apparent strength that is comparable to the known RhaS half-sites suggests the possibility that it has some *in vivo* function, although whether there is a role, and whether that role might involve DNA looping remains to be determined.

Differences between RhaS-CTD and RhaR-CTD. Given that RhaS-CTD and RhaR-CTD share 34% amino acid sequence identity, we were surprised to find that His6-RhaS-CTD efficiently activated transcription, while His6-RhaR-CTD only barely activated transcription. Although low protein levels may partially account for the lower activation by His6-RhaR-CTD, we would argue it is not the full explanation. Purified His6-RhaR-CTD protein was capable of binding to DNA, indicating that this fusion protein contained the necessary determinants for DNA binding and was capable of folding correctly. One hypothesis to explain the very low activation by His6-RhaR-CTD might be that its DNA-binding motifs are correctly folded, but that its transcription activation determinants are not properly folded. This hypothesis seems highly unlikely given that RhaR residue D276 is located within the stabilizing helix of one of the H-T-H DNA-binding motifs, and that this σ^{70}-contacting residue is responsible for approximately two-thirds of the transcription activation by RhaR (42). Also, as mentioned above, Tobin and Schleif (38, 39) previously found that, similar to His6-RhaR-CTD, full-length RhaR in the absence of L-rhamnose was capable of binding to DNA but not capable of activating transcription. It seems likely, therefore, that RhaR-CTD requires a signal from RhaR-NTD in the presence of L-rhamnose in order to activate transcription, and therefore is unable to activate in the absence of RhaR-NTD. RhaS, therefore, provides an additional example of an AraC/XylS family protein whose DNA-binding domain is capable of efficient transcription activation in the absence of a second domain. In contrast, RhaR may be an additional example of an AraC/XylS family protein whose DNA-binding domain alone is capable of little or no transcription activation.

ACKNOWLEDGMENTS

We would like to thank Peter Gegenheimer for use of his Cyclone Storage Phosphor System, and members of the lab of William Picking, especially Marianela Espina and Andrew Olive, for help with Western blots and image analysis. Western blot images were obtained on an Odyssey Infrared Imager through collaboration with LI-COR Inc.

This work was supported by NIH grant GM55099 from the National Institute of General Medical Sciences and NIH Grant P20 RR17708 from the Institutional Development Award (IDeA) Program of the National Center for Research Resources, both to S.M.E., NIH grant P20 RR016475 to J.F., and by the Intramural Research Program of NIH, National Cancer Institute, Center for Cancer Research to D.J.J.

REFERENCES

1. **Backman, K., Y.-M. Chen, and B. Magasanik.** 1981. Physical and genetic characterization of the *glnA-glnG* region of the *Escherichia coli* chromosome. Proc. Nat. Acad. Sci., U.S.A. **78:**3743-3747.
2. **Badia, J., L. Baldoma, J. Aguilar, and A. Boronat.** 1989. Identification of the *rhaA*, *rhaB* and *rhaD* gene products from *Escherichia coli* K-12. FEMS Microbiol. Lett. **65:**253-258.
3. **Bartolome, B., Y. M. Jubee, E. , and F. de la Cruz.** 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. Gene **102:**75-78.
4. **Bhende, P. M., and S. M. Egan.** 1999. Amino acid-DNA contacts by RhaS: an AraC family transcription activator. J. Bacteriol. **181:**5185-5192.
5. **Bhende, P. M., and S. M. Egan.** 2000. Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with sigma 70. J. Bacteriol. **182:**4959-4969.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248-254.
7. **Bustos, S. A., and R. F. Schleif.** 1993. Functional domains of the AraC protein. Proc. Natl. Acad. Sci. USA **90:**5638-5642.
8. **Carra, J. H., and R. F. Schleif.** 1993. Variation of half-site organization and DNA looping by AraC protein. EMBO J. **12:**35-44.
9. **Caswell, R., C. Webster, and S. Busby.** 1992. Studies on the binding of the Escherichia coli MelR transcription activator protein to operator sequences at the MelAB promoter. Biochem J **287:**501-8.
10. **Egan, S. M.** 2002. Growing repertoire of AraC/XylS activators. J Bacteriol **184:**5529-32.
11. **Egan, S. M., and R. F. Schleif.** 1993. A regulatory cascade in the induction of *rhaBAD*. J. Mol. Biol. **234:**87-98.
12. **Egan, S. M., and R. F. Schleif.** 1994. DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. J. Mol. Biol. **243:**821-829.
13. **Gallegos, M.-T., R. Schleif, A. Bairoch, K. Hofmann, and J. L. Ramos.** 1997. AraC/XylS family of transcriptional regulators. Microbiol. Mol. Biol. Rev. **61:**393-410.
14. **Ghosh, M., and R. F. Schleif.** 2001. Biophysical evidence of arm-domain interactions in AraC. Anal. Biochem. **295:**107-12.
15. **Gottesman, M. E., and M. B. Yarmolinsky.** 1968. The integration and excision of the bacteriophage lambda genome. Cold Spring Harbor Symp. Quant. Biol. **33:**735-747.
16. **Holcroft, C. C., and S. M. Egan.** 2000. Interdependence of activation at *rhaSR* by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain and RhaR. J. Bacteriol. **182:**6774-6782.
17. **Holcroft, C. C., and S. M. Egan.** 2000. Roles of cyclic AMP receptor protein and the carboxyl-terminal domain of the a subunit in transcription activation of the *Escherichia coli rhaBAD* operon. J. Bacteriol. **182:**3529-3535.
18. **Howard, V. J., T. A. Belyaeva, S. J. Busby, and E. I. Hyde.** 2002. DNA binding of the transcription activator protein MelR from Escherichia coli and its C-terminal domain. Nucleic Acids Res **30:**2692-700.
19. **Kaldalu, N., U. Toots, V. de Lorenzo, and M. Ustav.** 2000. Functional domains of the TOL plasmid transcription factor XylS. J. Bacteriol. **182:**1118-26.

20. **Kwon, H. J., M. H. J. Bennik, B. Demple, and T. Ellenberger.** 2000. Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. Nature Structural Biology **7:**424-430.
21. **Martin, R. G., and J. L. Rosner.** 2001. The AraC transcriptional activators. Curr. Opin. Microbiol. **4:**132-7.
22. **Miller, J. H.** 1972. Experiments in Molecular Genetics, vol. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736-747.
24. **Poore, C. A., C. Coker, J. D. Dattelbaum, and H. L. Mobley.** 2001. Identification of the domains of UreR, an AraC-like transcriptional regulator of the urease gene cluster in Proteus mirabilis. J Bacteriol **183:**4526-35.
25. **Power, J.** 1967. The L-rhamnose genetic system in Escherichia coli K-12. Genetics **55:**557-68.
26. **Prouty, M. G., C. R. Osorio, and K. E. Klose.** 2005. Characterization of functional domains of the Vibrio cholerae virulence regulator ToxT. Mol Microbiol **58:**1143-56.
27. **Reed, W. L., and R. F. Schleif.** 1999. Hemiplegic mutations in AraC protein. J Mol Biol **294:**417-25.
28. **Rhee, S., R. G. Martin, J. L. Rosner, and D. R. Davies.** 1998. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc. Natl. Acad. Sci. USA **95:**10413-10418.
29. **Saviola, B., R. Seabold, and R. F. Schleif.** 1998. Arm-domain interactions in AraC. J. Mol. Biol. **278:**539-548.
30. **Schleif, R.** 2000. Regulation of the L-arabinose operon of Escherichia coli. Trends Genet **16:**559-65.
31. **Seabold, R. R., and R. F. Schleif.** 1998. Apo-AraC actively seeks to loop. J Mol Biol **278:**529-38.
32. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85-96.
33. **Soisson, S. M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** 1997. Structural basis for ligand-regulated oligomerization of AraC. Science **276:**421-425.
34. **Stewart, G. S., S. Lubinsky-Mink, C. G. Jackson, A. Cassel, and J. Kuhn.** 1986. pHG165: a pBR322 copy number derivative of pUC8 for cloning and expression. Plasmid **15:**172-81.
35. **Tate, C. G., J. A. R. Muiry, and P. J. F. Henderson.** 1992. Mapping, cloning, expression, and sequencing of the *rhaT* gene which encodes a novel L-Rhamnose-H^+ transport protein in *Salmonella typhimurium* and *Escherichia coli.* J. Biol. Chem. **287:**6923-6932.
36. **Timmes, A., M. Rodgers, and R. Schleif.** 2004. Biochemical and physiological properties of the DNA binding domain of AraC protein. J Mol Biol **340:**731-8.
37. **Tobin, J. F., and R. F. Schleif.** 1987. Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. J. Mol. Biol. **196:**789-799.
38. **Tobin, J. F., and R. F. Schleif.** 1990. Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. J. Mol. Biol. **211:**75-89.
39. **Tobin, J. F., and R. F. Schleif.** 1990. Transcription from the *rha* operon p_{sr} promoter. J. Mol. Biol. **211:**1-4.

40. **Via, P., J. Badia, L. Baldoma, N. Obradors, and J. Aguilar.** 1996. Transcriptional regulation of the *Escherichia coli rhaT* gene. Microbiology **142:**1833-1840.
41. **Wade, J. T., T. A. Belyaeva, E. I. Hyde, and S. J. Busby.** 2000. Repression of the Escherichia coli *melR* promoter by MelR: evidence that efficient repression requires the formation of a repression loop. Mol. Microbiol. **36:**223-9.
42. **Wickstrum, J. R., and S. M. Egan.** 2004. Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. J Bacteriol **186:**6277-85.
43. **Wickstrum, J. R., and S. M. Egan.** 2002. Ni+-affinity purification of untagged cyclic AMP receptor protein. Biotechniques **33:**728-730.
44. **Wickstrum, J. R., T. J. Santangelo, and S. M. Egan.** 2005. Cyclic AMP receptor protein and RhaR synergistically activate transcription from the L-rhamnose-responsive rhaSR promoter in Escherichia coli. J Bacteriol **187:**6708-18.
45. **Wu, M., and R. Schleif.** 2001. Mapping arm-DNA-binding domain interactions in AraC. J Mol Biol **307:**1001-9.
46. **Wu, M., and R. Schleif.** 2001. Strengthened arm-dimerization domain interactions in AraC. J Biol Chem **276:**2562-4.
47. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.
48. **Zhi, H., and D. J. Jin.** 2003. Purification of highly-active and soluble *Escherichia coli* sigma 70 polypeptide overproduced at low temperature. Methods Enzymol **370:**174-80.
49. **Zhi, H., W. Yang, and D. J. Jin.** 2003. *Escherichia coli* proteins eluted from mono Q chromatography, a final step during RNA polymerase purification procedure. Methods Enzymol **370:**291-300.

CHAPTER 2

Differences in the Mechanism of the Allosteric L-Rhamnose Responses of the AraC/XylS Family Transcription Activators RhaS and RhaR[†]

Ana Kolin[#§], Vinitha Balasubramaniam[§],

Jeff Skredenske, Jason Wickstrum[^], and Susan M. Egan

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas

[§]These authors contributed equally to this work.

[#]Current address: Department of Microbiology and Immunology, Uniformed Services University of the Health Sciences, Bethesda, MD

[^]Current address: Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, KS

[†]This article is the accepted manuscript version of a subsequently published paper:

Kolin, A., V. Balasubramaniam, J.M. Skredenske, J.R. Wickstrum and S.M. Egan. 2008. Differences in the Mechanism of the Allosteric L-Rhamnose Responses of the AraC/XylS Family Transcription Activators RhaS and RhaR. *Mol. Microbiol.* **68**:448-461.

The published version of this article is available online at:

http://dx.doi.org/10.1111/j.1365-2958.2008.06164.x.

ABSTRACT

Proteins in the largest subset of AraC/XylS family transcription activators, including RhaS and RhaR, have C-terminal domains (CTDs) that mediate DNA-binding and transcription activation, and N-terminal domains (NTDs) that mediate dimerization and effector binding. The mechanism of the allosteric effector response in this family has been identified only for AraC. Here, we investigated the mechanism by which RhaS and RhaR respond to their effector, L-rhamnose. Unlike AraC, N-terminal truncations suggested that RhaS and RhaR don't use an N-terminal arm to inhibit activity in the absence of effector. We used random mutagenesis to isolate RhaS and RhaR variants with enhanced activation in the absence of L-rhamnose. NTD substitutions largely clustered around the predicted L-rhamnose-binding pockets, suggesting that they mimic the structural outcome of effector binding to the wild-type proteins. RhaS-CTD substitutions clustered in the first HTH motif, and suggested that L-rhamnose induces improved DNA binding. In contrast, RhaR-CTD substitutions clustered at a single residue in the second HTH motif, at a position consistent with improved RNAP contacts. We propose separate allosteric mechanisms for the two proteins: Without L-rhamnose, RhaS doesn't effectively bind DNA while RhaR doesn't effectively contact RNAP. Upon L-rhamnose binding, both proteins undergo structural changes that enable transcription activation.

INTRODUCTION

The RhaS and RhaR proteins are AraC/XylS family transcription activators of the *Escherichia coli* L-rhamnose catabolic regulon and are 30% identical to each other (Egan and Schleif, 1993; Egan and Schleif, 1994; Tate *et al.*, 1992; Tobin and Schleif, 1987). RhaS activates transcription of the *rhaBAD* and *rhaT* operons, which encode the L-rhamnose catabolic enzymes and the L-rhamnose transport protein, respectively (Power, 1967; Tate *et al.*, 1992). RhaR activates transcription of the *rhaSR* operon, which encodes RhaS and RhaR (Tobin and Schleif, 1987). The protein levels as well as the protein activities of RhaS and RhaR each increase in the presence of L-rhamnose, with the protein activities increasing on the order of 300-fold and 7-fold , respectively (Egan and Schleif, 1993; Tobin and Schleif, 1990b; Via *et al.*, 1996) (unpublished results). The DNA binding sites for RhaS and RhaR dimers consist of two 17 bp half-sites separated by 16 or 17 bp, respectively, and overlapping the -35 promoter hexamer by four bp (Egan and Schleif, 1994; Tobin and Schleif, 1990a). In addition to the transcription activators RhaS and RhaR, the cAMP receptor protein (CRP) is also required for the full activation of all three of the *rha* operons (Egan and Schleif, 1993; Holcroft and Egan, 2000; Via *et al.*, 1996).

The AraC/XylS family of transcription regulatory proteins are defined by a 100-amino-acid region of sequence similarity that comprises the DNA binding domain of the family members (Egan, 2002; Gallegos *et al.*, 1993; Gallegos *et al.*, 1997; Ramos, 1990; Tobin and Schleif, 1987). The majority of the proteins in the family, including RhaS and RhaR, consist of the conserved DNA binding domain, as well as at least one additional domain (Egan, 2002; Gallegos *et al.*, 1997). The crystal structures of the AraC/XylS family DNA binding domains of MarA and Rob have been solved (MarA consists of only the DNA binding domain while Rob also has a second domain) (Kwon *et al.*, 2000; Rhee *et al.*, 1998). The MarA-DNA complex

serves as our model for the RhaS- and RhaR-CTD structures given that its two helix-turn-helix (HTH) motifs each contact the DNA (Kwon *et al.*, 2000; Rhee *et al.*, 1998), similar to our findings for RhaS (Bhende and Egan, 1999). Our identification of several amino acid-base pair contacts in RhaS, allowed us to identify the orientation of the RhaS monomers at *rhaBAD* (Bhende and Egan, 1999). The conserved DNA binding domains of many AraC/XylS proteins also activate transcription by directly contacting the RNA polymerase (RNAP) α-subunit C-terminal domain and/or residues near the C-terminal end of σ^{70} (Dangi *et al.*, 2004; Grainger *et al.*, 2004a; Grainger *et al.*, 2004b; Jair *et al.*, 1995; Jair *et al.*, 1996a; Jair *et al.*, 1996b; Landini *et al.*, 1997; Lonetto *et al.*, 1998; Martin *et al.*, 2002; Ruiz *et al.*, 2001; Shah and Wolf, 2004). We've identified two residues in RhaS and one in RhaR and the residues in σ^{70} that they each contact to activate transcription (Bhende and Egan, 2000; Wickstrum and Egan, 2004).

The other domain of AraC/XylS proteins (usually the N-terminal domain, NTD) is not conserved throughout the family; however, proteins that also share sequence similarity with AraC in this domain make up the largest subset of the family. Among proteins in this subset, including RhaS and RhaR as well as XylS, MelR, and UreR, this domain is generally required for dimerization and/or effector binding. RhaS- and RhaR-NTDs have approximately 15% amino acid identity and 38% similarity with the AraC-NTD. The tertiary structure of the dimerization and effector-binding domain of AraC has been solved (Soisson *et al.*, 1997a, b; Weldon *et al.*, 2007), and serves as a model for the RhaS- and RhaR-NTD structures. Our results indicate that the RhaS and RhaR NTDs also function in both dimerization and effector (L-rhamnose) binding (Wickstrum *et al.*, 2007) (Kolin, Hunjan and Egan, unpublished results). Our work also indicates that, similar to AraC, RhaS and RhaR may have flexible linker regions that connect their two domains (Carra and Schleif, 1993; Kolin *et al.*, 2007). Given that the effector-binding site is physically separated from the DNA-binding and RNAP-contacting

domain in these proteins, an allosteric mechanism must communicate the effector-binding status of each NTD to the respective CTD.

The molecular mechanism of the effector response has been well defined for AraC and its effector L-arabinose, and is referred to as the "light-switch" mechanism (Harmer *et al.*, 2001; Reed and Schleif, 1999; Saviola *et al.*, 1998; Wu and Schleif, 2001a, b). The key to this mechanism is a small group of residues at the very N-terminal end of AraC known as the "N-terminal arm". The N-terminal arm binds over the L-arabinose binding pocket in the presence of L-arabinose (Soisson *et al.*, 1997b), leaving the two AraC domains in each monomer flexibly connected, and allowing transcription activation from the *araBAD* promoter-proximal half-sites (Harmer *et al.*, 2001; Seabold and Schleif, 1998; Wu and Schleif, 2001a). In the absence of arabinose, the arms instead bind to the C-terminal domain, rigidly connecting the domains so that a DNA loop forms that prevents transcription activation (Ghosh and Schleif, 2001; Ross *et al.*, 2003; Saviola *et al.*, 1998; Wu and Schleif, 2001a, b), (Lobell and Schleif, 1990; Seabold and Schleif, 1998).

In this study we sought to identify the mechanisms used by RhaS and RhaR to mediate their allosteric L-rhamnose responses. We tested N-terminal deletions of each protein and found that the L-rhamnose responses of RhaS and RhaR are most likely not mediated by N-terminal arms, and therefore may not be similar to the AraC mechanism. RhaS and RhaR variants were then identified that conferred increased activation in the absence of L-rhamnose compared to wild type. Our results suggest that the RhaS and RhaR NTDs likely undergo similar L-rhamnose-mediated structural changes. Substitutions conferring increased activation in the absence of L-rhamnose in the CTDs, however, were found in different regions of RhaS versus RhaR. We propose that in the absence of L-rhamnose, RhaS is limited in its ability to bind to DNA, whereas RhaR is limited in its ability to contact RNAP. Upon L-rhamnose binding, there

are allosteric structural changes transmitted from the NTDs to the CTDs in RhaS and RhaR that overcome their respective limitations and thereby allow them to activate transcription.

MATERIALS AND METHODS

Culture media and conditions. Cultures for β-galactosidase assays were grown in morpholinepropanesulfonic acid-buffered minimal medium (Neidhardt *et al.*, 1974), with 0.4% glycerol as the carbon source and in the absence or presence of 0.4% L-rhamnose and appropriate antibiotic (Bhende and Egan, 2000). All other cultures were grown in tryptone-yeast extract (TY) medium (0.8% tryptone, 0.8% yeast extract, 0.5% NaCl, pH 7.0) (Maloy *et al.*, 1996; Miller, 1972). For solid media, cells were grown on nutrient agar plates (2.3% Difco Nutrient agar, 0.5% NaCl) with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside, 40 μg/ml) for testing *lacZ* expression. Ampicillin (200 μg/ml), chloramphenicol (25 μg/ml) and L-rhamnose (0.2%) were added as indicated. All cultures were grown at 37°C.

General methods and strains. Oligonucleotide primers used in this study were synthesized by MWG-Biotech, Inc. (High Point, NC), and their DNA sequences are available upon request. Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Restriction endonuclease digestions, ligations and transformations were performed using standard procedures. All strains used in the study are derived from *E. coli* strain ECL116 (F^- Δ*lacU169 endA hsdR thi*) (Backman *et al.*, 1981), and genotypes list additional alleles in this genetic background.

Most β-galactosidase assays of activation by RhaS or its variants were performed in strain SME3000 [λΦ(*rhaB-lacZ*)Δ84 Δ(*rhaSR*)::Km]. The promoter driving *lacZ* expression in this strain extends to -84 (in all cases, relative to the transcription start site) and does not include the CRP binding site upstream of *rhaBAD*. The exception was the assays in Table 1 which were performed in SME1087 [λΦ(*rhaB-lacZ*)Δ226 Δ*rhaS*]. The promoter driving *lacZ* expression in SME1087 extends to -226 and includes all of the *rhaBAD* promoter elements. Most β-galactosidase assays of activation by RhaR or its variants were performed in strain SME3160

[λΦ(*rhaS-lacZ*)Δ85 Δ(*rhaSR*)::Km *recA::cat*]. The promoter driving *lacZ* expression in this strain extends to -85 and does not include the CRP binding site upstream of *rhaSR*. The exception was the assays in Table 1, which were performed in strain SME1076 [λΦ(*rhaS-lacZ*)Δ216 Δ(*rhaSR*)::Km]. The promoter driving *lacZ* expression in this strain extends to -216 and includes all of the *rhaSR* promoter elements.

Mutagenesis of *rhaS* and *rhaR*. All full-length wild type and RhaS and RhaR variants (including the N-terminal deletions) were expressed from the plasmid pHG165 (Ap^R, essentially pUC8 with the copy number of pBR322) (Stewart *et al.*, 1986). Long-way-around (inverse) PCR was used to construct 5' deletions of *rhaS* and *rhaR,* using the Expand High Fidelity PCR System (Roche, Indianapolis, IN), with the wild-type *rhaS* or *rhaR* in pHG165 (Stewart *et al.*, 1986) as the template. PCR products were designed to carry *Ear*I restriction sites at each end, thereby allowing seamless reconstruction of truncated *rhaS* or *rhaR* genes (LaPointe and Taylor, 2000; Wickstrum and Egan, 2004). Expression was under the control of the *lac* promoter located on the vector. These variants all retained the RhaS or RhaR start codon, but had progressively longer deletions ranging from codons 2 through 4 (in RhaS) up to deletion of codons 2 through 52 (in RhaR) (Table 1).

Random mutagenesis of *rhaS* or *rhaR* was performed by PCR amplification of the entire open reading frame using *Taq* DNA polymerase (Promega, Madison, WI) under standard reaction conditions (Zhou *et al.*, 1991). PCR products of randomly mutagenized *rhaS* or *rhaR* were ligated into the *Eco*RI and *Hind*III restriction sites of pGH165 (Stewart *et al.*, 1986) such that expression was driven by a vector-encoded *lac* promoter. Separation of the *rhaS* or *rhaR* L-rhamnose-independent double mutants into single mutants was performed by PCR amplification using the Expand High Fidelity PCR system (Roche; Indianapolis, IN). Fragments of *rhaS* or *rhaR* carrying a single mutation were seamlessly pieced together with fragments

lacking mutations (LaPointe and Taylor, 2000; Wickstrum and Egan, 2004) upon cloning into pHG165.

To test the phenotypes of the L-rhamnose-independent *rhaS* mutants when expressed in RhaS-CTD alone (residues 163-278, referred to as RhaS(163-278)), the portion of *rhaS* that encodes RhaS(163-278) was PCR amplified using the Expand High Fidelity PCR system (Roche; Indianapolis, IN) and ligated into pSE257 using *Bam*HI and *Hin*dIII restriction sites. pSE257 was derived from pSU18 (Cm^R, essentially the pUC18 multiple cloning site/*lacZ*α region on a plasmid with the P15A *ori*) (Bartolome *et al.*, 1991) by the addition of a promoter (driving downstream gene expression) that combines a near-consensus σ^{70} -35 element (TTGACT) with the -10 element from the *rhaSR* promoter (TACTAT), to construct a moderately strong, constitutive promoter (Wickstrum and Egan, unpublished).

The genes encoding the wild type and the L-rhamnose-independent variant T279A of RhaR-CTD alone (residues 196-311, referred to as RhaR(196-311)) were cloned into pSE262. pSE262 was derived from pHG165 (in a manner analogous to pSE257) by the addition of a promoter to drive downstream gene expression that combines a consensus σ^{70} -35 element (TTGACA) with the -10 element from the *rhaSR* promoter (TACTAT), to construct a relatively strong, constitutive promoter (Wickstrum and Egan, unpublished). PCR amplification was performed using the Expand High Fidelity PCR system (Roche; Indianapolis, IN). For these constructs, the insert contained the translation start codon, whereas the Shine-Dalgarno sequence was plasmid-encoded, with a *Bam*HI site between them to provide proper spacing (8 nucleotides) for ribosome recognition. Plasmids expressing either the wild-type RhaR(196-311) or RhaR(196-311) T279A were transformed into a strain that carried a *rhaSR* deletion and a single-copy fusion of *lacZ* with a *rhaSR* promoter that did not include the CRP binding site.

DNA sequencing. DNA sequencing was performed at the Molecular Research Core Facility of Idaho State University (Pocatello, ID) using an ABI3100 automated sequencer or the Northwestern University Biotechnology Laboratory (Chicago, IL) using an ABI3100 or an ABI3730 automated sequencer. The DNA sequences of both strands were determined for the entire cloned region of all cloned and mutagenized DNA fragments.

β-Galactosidase assays. β-Galactosidase assays were performed as previously described (Bhende and Egan, 1999) using the method of Miller (Miller, 1972). Briefly, all strains for β-galactosidase assays were grown in three serial steps: tryptone-yeast extract culture with ampicillin; overnight culture (MOPS-buffered minimal medium containing 0.04% glycerol as limiting carbon source and ampicillin); and growth culture (MOPS-buffered mimimal medium containing 0.4% glycerol as the carbon source, with or without 0.4% L-rhamnose and with ampicillin), based on the method of Neidhardt (Neidhardt *et al.*, 1974). Specific activities were averaged from at least two, usually three, independent assays with two replicates in each assay. β-Galactosidase activity is expressed in Miller units (Miller, 1972).

RESULTS

Our first hypothesis was that the mechanisms of the RhaS and RhaR allosteric effector responses might be similar to that of AraC. In the AraC mechanism, the very N-terminal residues (the arm) of AraC bind to the CTD in the absence of arabinose to prevent transcription activation and cause formation of a DNA loop. Consistent with the arm inhibiting activity in the absence of effector, N-terminal deletions of seven to 20 residues (within the arm) led to increased basal activation by AraC (were constitutive) (Saviola *et al.*, 1998; Soisson *et al.*, 1997b). In spite of the lack of evidence for DNA looping by RhaS or RhaR, binding of an N-terminal arm to their CTDs could, in principle, prevent binding to the adjacent DNA half-sites from which they activate transcription without requiring binding to alternative distant sites. If residues at the very N-terminus of RhaS or RhaR performed functions similar to the AraC N-terminal arm, then one or more N-terminal truncations would be predicted to result in activation of transcription in the absence of L-rhamnose to higher basal levels than wild-type RhaS or RhaR.

N-terminal truncations of RhaS and RhaR. Truncated variants of RhaS were assayed for *in vivo* transcription activation of a *rhaBAD* promoter that included the RhaS binding site but not the CRP binding site. In the absence of L-rhamnose, we found that the variants activated transcription to levels comparable to or lower than wild-type RhaS, indicating that none of the truncations conferred a constitutive phenotype (Table 1). In the presence of L-rhamnose, all of the truncations resulted in large defects in the ability of RhaS to activate transcription. Western blots (data not shown) as well as the ability of the variants to activate to nearly wild-type levels in the absence of L-rhamnose support the conclusion that the altered activity was not the result of low protein levels. We also constructed three additional RhaS deletions with intermediate endpoints and found that none of these conferred constitutive phenotypes (data not shown).

TABLE 1. Transcription activation by RhaS and RhaR N-terminal deletion variants

RhaS or RhaR Variant	β-Galactosidase Activity* (-) L-rhamnose	(+) L-rhamnose
A. RhaS variants		
Wild type	20	1070
Δ2-7	14	107
Δ2-13	23	35
Δ2-19	12	16
B. RhaR variants		
Wild type	220	528
Δ2-6	188	514
Δ2-12	177	427
Δ2-18	190	494
Δ2-24	144	277
Δ2-29	210	484
Δ2-34	108	394
Δ2-40	33	64
Δ2-46	2.6	10
Δ2-52	0.4	0.5

* β-Galactosidase activity (in Miller Units) was measured from a single-copy *lacZ* fusions in strain SME1087 [λΦ(*rhaB-lacZ*)Δ226 Δ*rhaS*] (A) or strain SME1076 [λΦ(*rhaS-lacZ*)Δ216 Δ(*rhaSR*)::Km] (B). Strains were transformed with plasmids (in vector pHG165) encoding either wild type or N-terminal deletion variants of RhaS or RhaR, as indicated. Cultures were grown in MOPS growth medium containing glycerol and ampicillin, with or without L-rhamnose as indicated. Standard errors were less than 30% of the average units. The activity from each fusion in the absence of plasmid-encoded RhaS or RhaR was less than one Miller Unit.

Overall, these results suggest that the mechanism of the RhaS L-rhamnose response does not involve an N-terminal arm inhibiting activity in the absence of L-rhamnose, and thereby differs from the AraC L-arabinose response.

RhaR has 33 additional N-terminal residues compared with RhaS, so we constructed a larger number of N-terminal deletions and tested their ability to activate transcription from the *rhaSR* promoter. None resulted in increased activation in the absence of L-rhamnose (Table 1). In addition, deletion of the entire 33-residue N-terminal RhaR extension relative to RhaS had very little affect on RhaR activity. The apparent dispensability of the RhaR N-terminal extension is consistent with the observation that most RhaR orthologs outside of *E. coli* and *Shigella* spp. do not have this extension. Deletion of the first 40 or more RhaR residues resulted in increasingly large losses of activity both in the presence and absence of L-rhamnose, and western blots indicated that these same truncations were detected at substantially lower levels than wild-type RhaR (data not shown). Therefore, while the instability of the longest truncations introduces some uncertainty, we conclude that RhaR most likely does not use an N-terminal arm to mediate its response to L-rhamnose.

Isolation of partially constitutive RhaS and RhaR variants. The AraC allosteric response to effector is the only mechanism that has been well characterized in the AraC/XylS family. In addition, our previous results suggested that the RhaS and RhaR linkers are not central to the signal transmission from the NTDs to the CTDs (Kolin *et al.*, 2007). We were therefore unable to hypothesize which residues might be important for the RhaS and RhaR allosteric effector responses, and turned to random mutagenesis. To avoid the problem of distinguishing interesting mutations from the large background of those that reduced activity for reasons unrelated to the effector response, we chose to screen for RhaS and RhaR variants with elevated transcription activation in the absence of L-rhamnose. We used PCR amplification to

introduce random mutations, and then transformed the cloned genes into strains carrying single-copy translational fusions to *lacZ* of either a *rhaBAD* promoter (for RhaS) or a *rhaSR* promoter (for RhaR). For simplicity, all promoters lacked the upstream CRP binding site. The level of *lacZ* expression from these fusions was such that, when plated on media containing X-gal, wild-type transformants yielded white colonies in the absence of L-rhamnose and blue colonies in the presence of L-rhamnose. We therefore screened for blue colonies in the absence of L-rhamnose. We screened approximately 240,000 clones obtained from 90 independent PCR amplifications of *rhaS*, and identified blue colonies in 69 of the amplifications. Early in the screening process we began to isolate independent mutants (from separate PCR amplifications) with identical substitutions. Subsequently, we used β-galactosidase assays to identify candidates whose activation levels appeared to differ from those already isolated, in an attempt to isolate as many unique substitutions as possible. Based on the assay results, we sequenced a total of 32 candidates, each from an independent PCR pool. We isolated nine unique RhaS variants, five of which were isolated at least twice (accounting for 19 of the 32 candidates) (Tables 2 and 3). The 13 remaining isolates each had one substitution that was identical to one of the nine unique substitutions as well as a second substitution. In each case, the β-galactosidase activity indicated that the second substitution did not contribute to the phenotype, so we did not further analyze these isolates. The nine unique variants had substitutions that were approximately evenly divided between RhaS-NTD and RhaS-CTD (Tables 2 and 3). These results indicate that, at least within the mutation spectrum of the PCR mutagenesis, a limited number of substitutions could confer a phenotype of increased activity in the absence of L-rhamnose upon RhaS.

To identify RhaR variants, we screened approximately 170,000 clones derived from 70 independent *rhaR* PCR amplifications. Of the twenty-nine colonies that were more blue than wild type in the absence of L-rhamnose (each from an independent PCR amplification), 15 had

TABLE 2. Transcription activation by partially constitutive variants of RhaS with substitutions that map to RhaS-NTD

			β-Galactosidase Activity[c]	
Expt	RhaS variant[a]	No. of isolates[b]	(-) L-rhamnose	(+) L-rhamnose
A.	Wild type		0.55	218
	F28L/F50L	1	57	61
	E37K	2	5.6	3.9
	T56I	2	6.1	360
	N83H	1	2.9	2.7
	P141L	3	12	304
B.	Wild type		0.34	206
	F28L		0.30	1.3
	F50L		0.38	3.6

[a] Wild type and RhaS variants were encoded on pHG165 transformed into SME3000 [λΦ(*rhaB-lacZ*)Δ84 Δ(*rhaSR*)::Km].

[b] Number of times this substitution was independently isolated in random mutagenesis screen.

[c] Cultures were grown in MOPS growth medium containing glycerol plus ampicillin with or without L-rhamnose as indicated. β-Galactosidase activity (in Miller Units) was assayed. Standard errors were less than 20% of the average units, except for a few of the samples with activities below 3 Miller units. These had errors up to 31% of the average units.

TABLE 3. Transcription activation by partially constitutive variants of RhaS with substitutions that map to RhaS-CTD

			β-Galactosidase Activity[c]	
Expt	RhaS variant[a]	No. of isolates[b]	(-) L-rhamnose	(+) L-rhamnose
A.	Wild type		0.55	218
	F184Y/Q207R	1	8.6	8.9
	L201R	6	54	329
	L208F	2	5.6	285
	Q210R	1	2.6	326
B.	Wild type		0.34	206
	F184Y		0.69	218
	Q207R		0.73	243

[a] Wild type and RhaS variants were encoded on pHG165 transformed into SME3000]λΦ(*rhaB-lacZ*)Δ84 Δ(*rhaSR*)::Km].

[b] Number of times this substitution was independently isolated in random mutagenesis screen.

[c] Cultures were grown in MOPS growth medium containing glycerol plus ampicillin with or without L-rhamnose as indicated. β-Galactosidase activity (in Miller Units) was assayed. Standard errors were less than 20% of the average units, except for a few of the samples with activities below 3 Miller units. These had errors up to 31% of the average units.

two-fold or higher activity than wild-type RhaR in the absence of L-rhamnose (data not shown). Full-gene sequencing showed that 13 of the isolates encoded a substitution at residue T279, in RhaR-CTD, changing the wild-type Thr to either Ala (10 times) or Ser (3 times). Some of these had one or more additional substitutions; however, all of them activated transcription to levels that were the same (within error) as RhaR T279A and T279S alone, so those with multiple substitutions were not further analyzed (data not shown). We also found that RhaR T279A and T279S activated to the same levels as each other (data not shown), therefore we proceeded only with RhaR T279A. Only two of the 15 independent RhaR variants did not have a substitution at RhaR T279. Both of these had two substitutions, with all four located outside of RhaR-CTD (in either RhaR-NTD or the unstructured linker between the two domains). Therefore, among the partially constitutive RhaR variants there were only four unique variants, and we further analyzed three of them: F61L/W75R; H165R/D197G; and T279A.

Partially constitutive RhaS variants with substitutions in RhaS-NTD. Among the nine RhaS variants with increased activity in the absence of L-rhamnose, five had substitutions within the L-rhamnose-binding and dimerization domain, RhaS-NTD. Table 2 shows the ability of these RhaS variants to activate transcription. In the absence of L-rhamnose, they activated transcription to levels from six- to more than 100-fold higher than wild-type RhaS. Two of the variants, RhaS T56I and P141L, were able to respond to L-rhamnose, and activated transcription in the presence of L-rhamnose to levels that were somewhat higher than wild-type RhaS. In contrast, the other three variants, RhaS E37K, N83H and the variant with two substitutions, F28L/F50L, were unable to activate transcription to higher levels in the presence of L-rhamnose compared with their respective activities in the absence of L-rhamnose, and therefore we refer to them as uninducible.

We used the structure of AraC-NTD in the presence of L-arabinose (Soisson *et al.*, 1997a) as a model to identify the positions of these RhaS residues (Fig. 1A). All of the substitutions were located within the predicted RhaS β-barrel, relatively near the sugar-binding pocket, with the exception of RhaS P141L. The RhaS substitutions do not appear to form a surface-exposed cluster of residues that might define a site of contact between RhaS-NTD and RhaS-CTD. However, note that the RhaS-CTD would connect to this model of the RhaS-NTD at the upper right corner (as drawn). Given that a second monomer of RhaS is expected to dimerize along the α-helical face of the structure, the RhaS-CTD might be expected to sit above the RhaS-NTD in this orientation, and the RhaS substitutions might provide clues to structural changes transmitted from the L-rhamnose binding pocket toward RhaS-CTD. Western blot analysis (data not shown) indicated that the protein levels of RhaS P141L were at least several fold higher than wild-type RhaS, perhaps explaining the increased activity of this variant, whereas none of the remaining variants were detected at higher levels than wild-type RhaS. We separated the two substitutions in RhaS F28L/F50L and found that neither alone conferred increased activation in the absence of L-rhamnose (Table 2B), indicating that both of these substitutions were necessary for the constitutive phenotype.

Partially constitutive RhaR variants with substitutions in or near RhaR-NTD. The RhaR variants F61L/W75R and H165R/D197G activated transcription of *rhaS-lacZ* to levels that were approximately 4- and 6-fold higher, respectively, than wild-type RhaR in the absence of L-rhamnose (Fig. 2). In the presence of L-rhamnose, RhaR F61L/W75R activated to approximately the same level as wild-type RhaR, while H165R/D197G activated to a level 7-fold higher than wild-type RhaR. Using the structure of AraC-NTD in the presence of L-arabinose (Soisson *et al.*, 1997a) as a model we found that RhaR F61L and W75R both mapped within the β-barrel sugar-binding subdomain of AraC-NTD (Fig. 1B). The H165R substitution of RhaR

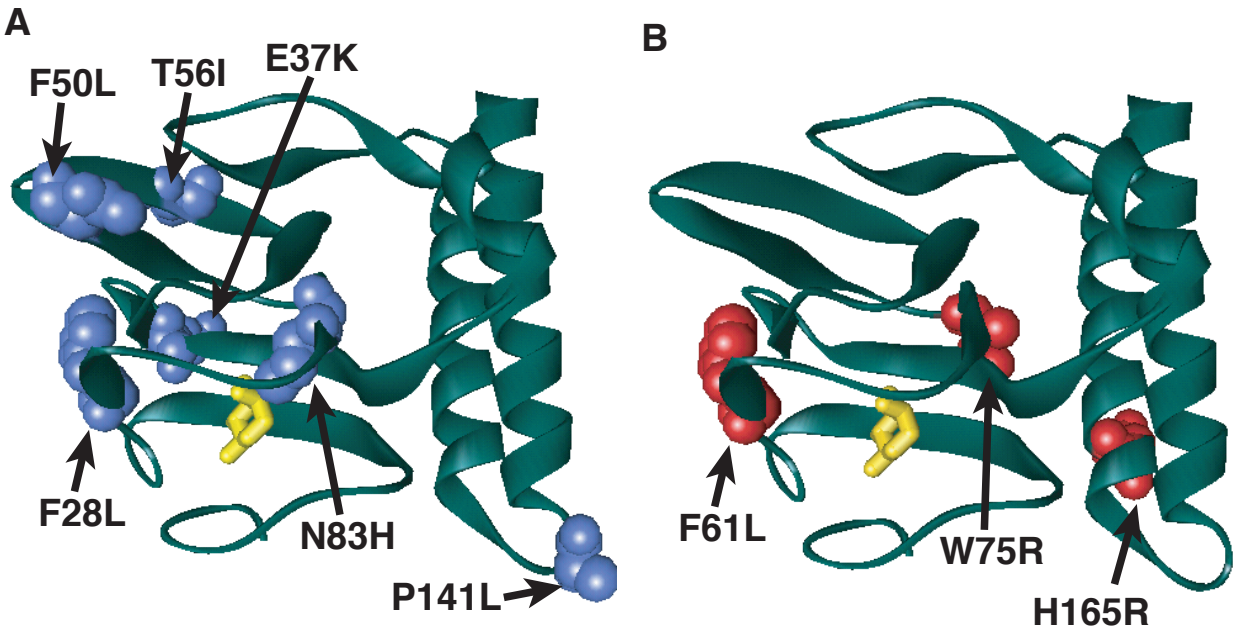

Figure 1. Model of substitutions leading to a partially constitutive phenotype within RhaS- or RhaR-NTD. The model for the structure of RhaS- or RhaR-NTD is the structure of one monomer of AraC-NTD in a ribbon representation (green) in the presence of L-arabinose (yellow stick model) (Soisson *et al.*, 1997b). (A) The positions of RhaS-NTD substitutions are shown in blue as space filling representations. (B) The positions of RhaR-NTD substitutions are shown in red as space filling representations. RhaS F28L/F50L, RhaR F61L/W75R, and RhaR H165R/D197G substitutions were isolated as double mutants. RhaR D197G is not shown, since it is located just beyond the structured residues.

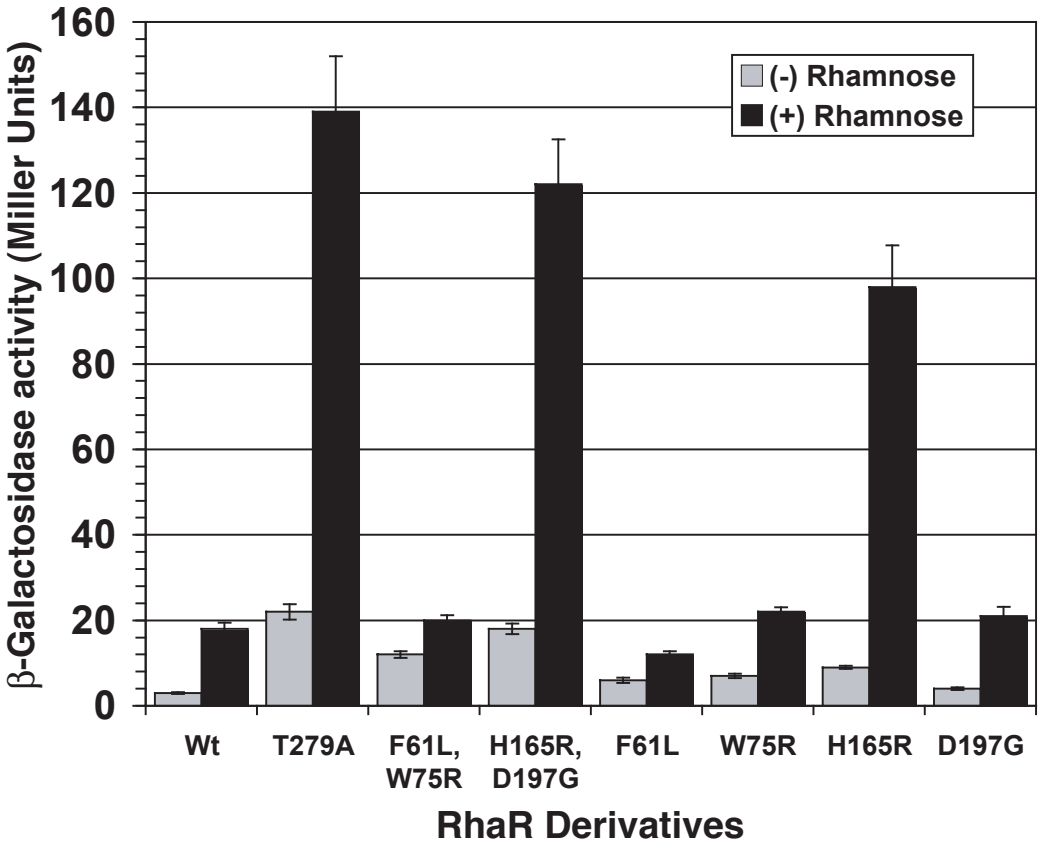

Figure 2. Transcription activation by partially constitutive RhaR variants. Plasmids expressing wild-type RhaR or variants were transformed into strain SME3160 [Φ(*rhaS-lacZ*)Δ85 Δ(*rhaSR*)::Km], the cells were grown with or without L-rhamnose, as indicated, and β-galactosidase activity was measured.

H165R/D197G mapped to the first of the two predicted α-helices of RhaR-NTD (Fig. 1B). While this residue does not map to the β-barrel, it is in position to potentially interact with the β-barrel. The other substitution in this variant, D197G, was located three residues beyond the structured residues of the AraC-NTD model (Soisson *et al.*, 1997b), but within the region defined biochemically as the minimal AraC-NTD (Eustance *et al.*, 1994). This residue is therefore predicted to lie either at the very end of RhaR-NTD or at the beginning of the linker connecting the RhaR domains. Western blots showed that both of the RhaR variants were expressed at levels similar to or slightly lower than wild-type RhaR (data not shown); therefore, increased protein levels did not explain their increased basal activity relative to wild-type RhaR.

We separated the two double mutants into single mutants in order to determine whether both substitutions in the RhaR variants were required for their increased activity in the absence of L-rhamnose. RhaR F61L and W75R each had somewhat higher activation than wild-type RhaR in the absence of L-rhamnose, but both substitutions were required for the full activity of RhaR F61L/W75R in the absence of L-rhamnose (Fig. 2). RhaR H165R activated to higher levels than wild-type RhaR in the absence of L-rhamnose, whereas RhaR D197G did not have significantly higher activation than wild-type RhaR (Fig. 2); but may have a subtle effect on transcription activation in combination with RhaR H165R. RhaR F61L was not fully inducible to the wild-type level in the presence of L-rhamnose, while all of the other single variants were L-rhamnose inducible to at least wild-type levels (Fig. 2).

Partially constitutive RhaS variants with substitutions in RhaS-CTD. Four RhaS variants with increased activity in the absence of L-rhamnose had substitutions in the DNA-binding and transcription activation domain (RhaS-CTD). Table 3 shows the amino acid substitutions found in each of these RhaS variants as well as the level of transcription activation by each. These variants activated transcription in the absence of L-rhamnose to levels ranging

from five- to nearly 100-fold higher than wild-type RhaS. Of these variants, only RhaS F184Y/Q207R was uninducible by L-rhamnose. Each of the other variants was able to respond to L-rhamnose, and all activated to levels that were somewhat higher than that of wild-type RhaS. We separated the two substitutions in RhaS F184Y/Q270R, and found that neither alone conferred the phenotype of the parent (Table 3B). Each of them conferred only a very small increase in activity in the absence of L-rhamnose, and the uninducible phenotype of the parent was also lost, indicating that both substitutions were required for the phenotype.

Substitutions in RhaS-CTD could, in principal, result in increased activity in the absence of L-rhamnose due to loss-of-function changes, such as the loss of inhibitory domain-domain contacts. Alternatively, they could result in increased activity due to gain-of-function changes, such as improved DNA binding, interactions with RNAP, or protein folding/stability. Improved DNA binding appeared likely given that all of the RhaS-CTD variants include a substitution in the recognition helix of the first HTH motif (Fig. 3A); three out of four of these residues were changed to positively charged Arg residues; and we have previously shown that nearby residues R202 and R206 make base-specific contacts with the DNA (Bhende and Egan, 1999). The gain-of-function possibilities in RhaS-CTD are predicted to exhibit an altered phenotype (relative to the appropriate wild type) both in full-length RhaS and in RhaS(163-278) (which is RhaS-CTD alone, residues 163-278). In contrast, the loss-of-function possibilities are predicted to exhibit an altered phenotype (relative to the appropriate wild type) in full-length RhaS, but not in RhaS(163-278). Therefore, we could distinguish these possibilities by subcloning the substitutions into truncated *rhaS* genes that encode RhaS(163-278).

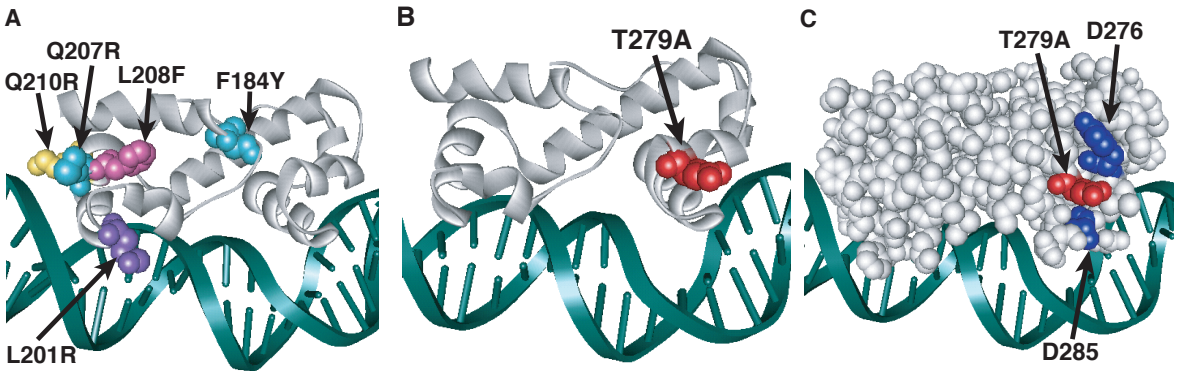

Figure 3. Model of substitutions leading to a constitutive phenotype within RhaS-CTD and RhaR-CTD. The structure of MarA in complex with DNA is used as a model for RhaS-CTD and RhaR-CTD (Rhee *et al.*, 1998), with the DNA in green and MarA in light gray. The overlap with the -35 element is on the right in all cases. (A) The positions of the RhaS-CTD substitutions are shown in various colors as space filling representations on a ribbon model of MarA. The F184Y and Q207R substitutions were isolated as a double mutant. (B) The position of the RhaR-CTD substitution is shown in red as a space-filling representation on a ribbon model of MarA. (C) The position of residue RhaR T279 is shown relative to the positions of residues shown to contact RNAP in RhaS and RhaR. RhaR D276/RhaS D241 and RhaS D250 (that aligns with RhaR D285) have been shown to contact RNAP and are shown in blue (Bhende and Egan, 2000; Wickstrum and Egan, 2004). RhaR T279 is shown in red, all on a space-filling representation of MarA.

Figure 4 shows the levels of transcription activation by RhaS(163-278) variants carrying each of the substitutions. Three of the RhaS(163-278) variants, L201R, Q210R and F184Y/Q207R, showed significantly higher (1.7- to 4-fold) levels of transcription activation than wild-type RhaS(163-278). The fourth variant (L208F) activated transcription to a slightly higher level than wild-type RhaS(163-278), but this difference was not statistically significant. These results suggest that the substitutions within RhaS-CTD most likely all represent domain autonomous gain-of-function changes in RhaS, and based on their positions, likely confer increased DNA-binding upon RhaS. These results suggest a model in which RhaS is unable to effectively bind DNA in the absence of L-rhamnose, which is entirely consistent with the previous finding that RhaS DNA binding was only detectable in the presence of L-rhamnose (Egan and Schleif, 1994).

Partially constitutive RhaR variants with substitutions within RhaR-CTD. All 13 RhaR-CTD variants with increased activity in the absence of L-rhamnose had substitutions at residue T279, and activated transcription of *rhaS-lacZ* to a level approximately 7-fold higher than wild-type RhaR in the absence of L-rhamnose (Fig. 2). This residue mapped to the stabilizing helix of the second RhaR HTH motif (Fig. 3B). The alignment of this residue with AraC and its position on the MarA structure suggest two different hypotheses to explain the phenotype of RhaR T279A. RhaR T279 aligns with AraC R251, which has been proposed to interact with residues in the N-terminal arm of AraC (Wu and Schleif, 2001b). Although our results suggest that the very N-terminal residues of RhaR are not involved in its L-rhamnose response (Table 1), one hypothesis is that RhaR T279 could interact with a different region of RhaR-NTD as part of its L-rhamnose response. A second hypothesis was based on the location of RhaR T279 on the surface of RhaR-CTD between residues D276 and D285, which we've previously identified as sites of contact between RhaS and RhaR and the σ^{70} subunit of RNAP (Bhende and Egan, 2000; Wickstrum and Egan, 2004) (Fig. 3C). Therefore, our second

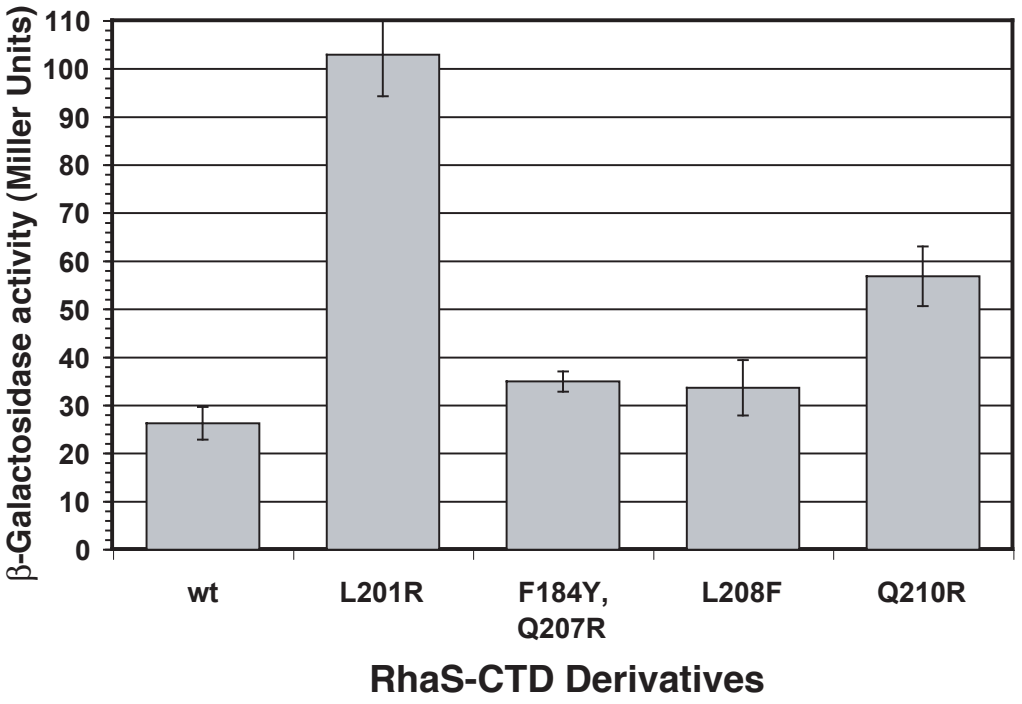

Figure 4. Transcription activation by variants of the RhaS-CTD expressed alone [RhaS(163-278)]. Wild-type RhaS(163-278) or variants (plasmid-encoded) in strain SME 3000 [Φ(*rhaB-lacZ*)Δ84 Δ(*rhaSR*)::Km] were assayed for transcription activation of the single copy *lacZ* fusion. Cultures were grown in minimal glycerol medium with ampicillin and without L-rhamnose.

hypothesis was that the L-rhamnose-independent phenotype of RhaR T279A might be due to increased interaction with the σ^{70} subunit of RNAP, perhaps by decreasing steric inhibition imposed by T279 in the absence of L-rhamnose.

In order to distinguish these two hypotheses for the phenotype of RhaR T279, we subcloned the genes encoding wild-type RhaR and RhaR T279A such that only the CTD of RhaR was expressed [RhaR(196-311), residues 196-311]. Similar to the RhaS case above, gain-of-function substitutions that resulted in improved DNA binding, interactions with RNAP, or protein folding/stability are predicted to have an altered phenotype relative to the appropriate wild type in both full-length RhaR and RhaR(196-311). On the other hand, loss-of-function changes that result in the loss of inhibitory domain-domain contacts that altered inter-domain contacts are predicted to have an altered phenotype relative to the appropriate wild type in full-length RhaR but not in RhaR(196-311). Figure 5 shows the ability of RhaR(196-311) T279A to activate *rhaS-lacZ* expression compared with wild-type RhaR(196-311). As expected, neither wild-type RhaR(196-311) nor RhaR(196-311) T279A was able to respond to L-rhamnose. We found that RhaR(196-311) T279A activated transcription to significantly higher levels (3-fold) than wild-type RhaR(196-311) in the absence of L-rhamnose. This suggests that the increased activity of RhaR T279A (in full-length RhaR) is unlikely to be due to inter-domain interactions. Our RhaR results, therefore, suggest a model in which RhaR is unable to effectively contact RNAP in the absence of L-rhamnose. This model is consistent with the previous finding that RhaR is able to bind to DNA but is not able to activate transcription in the absence of L-rhamnose (Tobin and Schleif, 1990a, b).

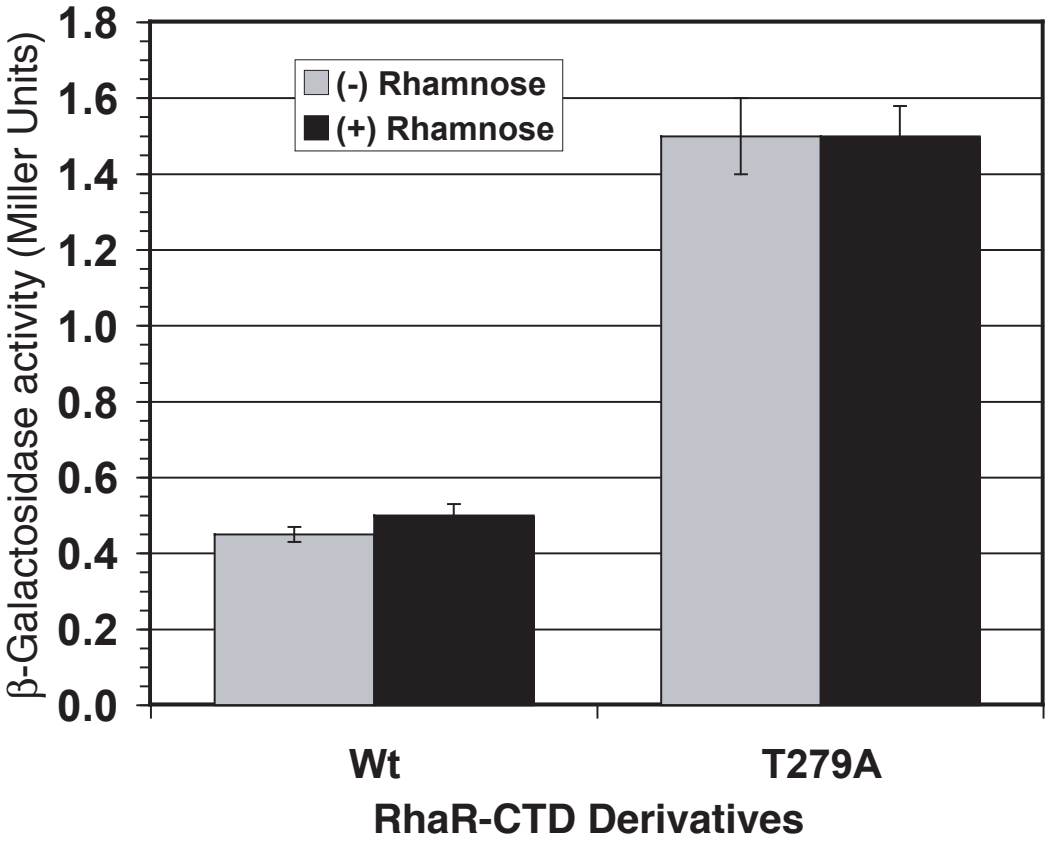

Figure 5. Transcription activation by RhaR T279A in RhaR-CTD expressed alone [RhaR(196-311)]. Wild-type RhaR(196-311) or RhaR(196-311) T279A (plasmid-encoded) in strain SME3160 [Φ(*rhaS-lacZ*)Δ85 Δ(*rhaSR*)::Km] were assayed for transcription activation of the single copy *lacZ* fusion. Cultures were grown in minimal glycerol medium with ampicillin and with or without L-rhamnose, as indicated.

DISCUSSION

RhaS and RhaR do not use "light-switch" allosteric effector mechanisms. The goal of this study was to identify the allosteric mechanisms used by the RhaS and RhaR proteins to respond to their common effector, L-rhamnose. Our first hypothesis was that the RhaS and RhaR mechanism might involve inhibition of activity by residues at their very N-termini in the absence of effector. This would be similar to the N-terminal arm that mediates the L-arabinose "light-switch" mechanism in AraC (Harmer *et al.*, 2001; Reed and Schleif, 1999; Saviola *et al.*, 1998). However, unlike AraC, deletion of N-terminal residues of RhaS and RhaR did not result in effector-independent phenotypes (Table 1), suggesting that the mechanism of their effector responses differ from that of AraC.

Alternative allosteric effector hypotheses. Given that AraC's is the only well-characterized allosteric effector mechanism among AraC/XylS family proteins, we considered other mechanisms that could explain the allosteric effect of L-rhamnose on the activity of RhaS and RhaR. One hypothesis was that the linker could transmit a signal from the RhaS- and RhaR-NTD to their respective CTDs; however, our previous results suggested that the linker was not central to the L-rhamnose signal transmission (Kolin *et al.*, 2007). We therefore hypothesized that the L-rhamnose-binding (NTD) and DNA-binding (CTD) domains might make effector-dependent contacts that could occur, in principle, in the absence of L-rhamnose to inhibit activity, or in the presence of L-rhamnose to stimulate activity. Our N-terminal deletion results indicated that these residues did not inhibit activity in the absence of L-rhamnose. However, these results did not rule-out models in which the very N-terminal residues made stimulatory contacts in the presence of L-rhamnose, or those in which other residues in the NTDs made inhibitory (or stimulatory) contacts in the absence (or presence) of L-rhamnose.

Somewhat surprisingly, we had previous evidence that the mechanism of allosteric activation might differ at some level between RhaS and RhaR. We've found that the RhaS C-terminal domain expressed alone, RhaS(163-278), was similar to L-rhamnose-activated full-length RhaS in that both could bind specifically to DNA and activate transcription (2,000- and 7,000-fold compared with Δ*rhaS*, respectively); whereas L-rhamnose-free full-length RhaS was unable to bind to DNA (Egan and Schleif, 1994; Wickstrum *et al.*, 2007). In contrast, the activity of the RhaR C-terminal domain expressed alone, RhaR(196-311), more closely mimicked L-rhamnose-free full-length RhaR in that both were capable of binding to DNA, but barely if at all (two-fold or less compared with Δ*rhaR*), able to activate transcription. Only L-rhamnose-activated full-length RhaR was capable of robust transcription activation (Tobin and Schleif, 1990a, b; Wickstrum *et al.*, 2007).

One model to explain the above findings is that the RhaS allosteric effector mechanism involves inhibition in the absence of L-rhamnose and the RhaR mechanism involves stimulation in the presence of L-rhamnose. However, we conclude that a model in which RhaS requires L-rhamnose to effectively bind DNA while RhaR requires L-rhamnose to effectively contact RNAP better fits the available data (further explained below). We also previously noticed that RhaS and RhaR were not alone in this difference in the activities of their DNA-binding domains. Expression of the DNA-binding domains alone of several AraC/XylS family proteins resulted in high levels of transcription activation (Gendlina *et al.*, 2002; Kaldalu *et al.*, 2000; Miyazaki *et al.*, 2003; Poore *et al.*, 2001; Wickstrum *et al.*, 2007), while expression of the DNA-binding domains of others resulted in little or no transcription activation (Howard *et al.*, 2002; Prouty *et al.*, 2005), suggesting that the RhaS and RhaR mechanisms may apply to other AraC/XylS family proteins as well.

L-Rhamnose likely induces allosteric changes in RhaS- and RhaR-NTD. Our screen for RhaS and RhaR substitutions that conferred increased activity in the absence of L-rhamnose resulted in the isolation of a number of substitutions within the NTDs. The positions of these residues within the β-barrel of the AraC-NTD model of RhaS- and RhaR-NTD along with their phenotypes suggests that they may mimic structural changes that normally occur upon binding L-rhamnose, and that may be transmitted from the sugar-binding pocket to the CTDs (Fig. 1). RhaR H165R was instead located in the α-helical region of RhaR-NTD and, unlike RhaS P141L, high protein levels did not explain its elevated activity in the absence of L-rhamnose. The positively charged H165R substitution is positioned, however, across from a negatively charged aspartic acid residue in the RhaR β-barrel (data not shown), suggesting a potential influence of this substitution on the β-barrel structure.

L-Rhamnose induces an allosteric change in RhaS-CTD HTH#1. Each of the four RhaS-CTD substitutions that conferred increased activity in the absence of L-rhamnose map to the predicted recognition helix of the first HTH motif (Fig. 3A). Three of these residues were changed to Arg (L201R, Q207R and Q210R) which – combined with our previous identification of two nearby Arg residues in RhaS HTH#1 (R202 and R206) that make base-specific DNA contacts (Bhende and Egan, 1999) – strongly supports the hypothesis that these RhaS variants bind to DNA more tightly than wild-type RhaS. These substitutions also conferred increased activity upon RhaS(163-278), again supporting the hypothesis that DNA-binding is increased (Fig. 4). Although purified RhaS(163-278) is significantly less prone to aggregation than full-length RhaS (Wickstrum *et al.*, 2007), quantitative assays to directly compare the relative DNA binding strength of these variants have not been possible due to residual aggregation. The previous finding that DNA binding by RhaS has not been detected in the absence of L-rhamnose

(Egan and Schleif, 1994) supports the model that DNA binding is central to the RhaS L-rhamnose response.

L-Rhamnose induces an allosteric change in RhaR-CTD HTH#2. All of the RhaR variants within RhaR-CTD that conferred increased activity in the absence of L-rhamnose substituted small residues (Ala or Ser) for the Thr at position 279. This residue is predicted to lie within the stabilizing helix of the second HTH motif of RhaR-CTD (Fig. 3B), and aligns with AraC R251. Although AraC R251 (in the DNA binding domain of AraC) has been proposed to interact with residues in the N-terminal arm of AraC in the absence of L-arabinose (Wu and Schleif, 2001b), suggesting that RhaR might use a similar mechanism, our results suggest that the RhaR L-rhamnose response is not similar to the AraC light-switch mechanism (Saviola *et al.*, 1998; Wu and Schleif, 2001b). The position of residue T279 on the structural model of RhaR-CTD between two residues that participate in protein-protein interactions with residues in the σ^{70} subunit of RNAP (Bhende and Egan, 1999; Wickstrum and Egan, 2004) (Fig. 3C), suggests that the substitutions to the smaller Ala or Ser residues may allow RhaR to more effectively contact RNAP in the absence of L-rhamnose. The previous finding that RhaR is capable of binding to DNA both in the absence and presence of L-rhamnose, but cannot activate transcription without L-rhamnose (Tobin and Schleif, 1990a, b) supports the hypothesis that RhaR contacts with σ^{70} are central to the RhaR L-rhamnose response.

Model for the RhaS and RhaR allosteric effector responses. Overall, we propose that both the RhaS and the RhaR L-rhamnose responses involve the stimulation of the activity of the CTDs by the NTDs in the presence of L-rhamnose. We further hypothesize that in the absence of L-rhamnose, RhaS is limited in its ability to bind to DNA, and that L-rhamnose binding induces a structural change that alters HTH#1 and increases DNA binding (Fig. 6B). In contrast, we propose that in the absence of L-rhamnose, RhaR is limited in its ability to contact RNAP, and

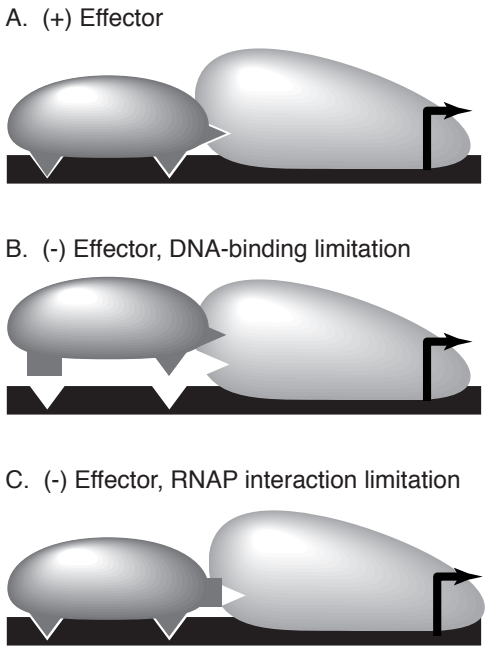

Figure 6. Model of the effector responses of RhaS and RhaR. (A) In the presence of effector, both RhaS and RhaR can efficiently bind to DNA and contact RNAP to activate transcription. (B) In the absence of an effector, RhaS is limited in its ability to bind to DNA. (C) In the absence of effector, RhaR is limited in its ability to effectively contact RNAP. Black line, DNA; dark gray, the DNA binding domain of the promoter-proximal monomer of either RhaS or RhaR; light gray, RNAP. Triangular protrusions or indentations represent sites of contact between the DNA-binding domain of RhaS or RhaR and DNA or RNAP**,** whereas rectangular protrusions represent sites that are not in the correct conformation to make contact.

that a structural change is required in HTH#2 upon L-rhamnose binding to allow RhaR to effectively contact RNAP (Fig. 6C). The previous results that RhaS is unable to bind to DNA in the absence of L-rhamnose, and that RhaR is capable of binding to DNA in the absence of L-rhamnose, but is not capable of activating transcription (Egan and Schleif, 1994; Tobin and Schleif, 1990a, b), are entirely consistent with our proposed mechanisms. The finding that MarA consists of two structurally similar subdomains (Rhee *et al.*, 1998), one including HTH#1 (involved in the RhaS mechanism), and one including HTH#2 (involved in the RhaR mechanism), suggests the possibility that the underlying mechanisms of the RhaS and RhaR allosteric responses may be more similar than they appear. This proposal is also supported by the similarities in our findings in the RhaS- and RhaR-NTDs.

The model that the RhaS- and RhaR-NTD's must stimulate their respective CTDs in the presence of L-rhamnose in order to activate transcription easily explains why RhaR(196-311) alone could not activate transcription (Wickstrum *et al.*, 2007). However, reconciling the finding that RhaS(163-278) alone activated transcription to high levels (Wickstrum *et al.*, 2007) with this model is more difficult. This apparent contradiction might be explained by the specific nature of the step limiting activation by RhaS-CTD. We propose that the DNA-binding defect of RhaS in the absence of L-rhamnose [and similarly of RhaS(163-278)] may be relatively small, but nonetheless sufficient to prevent the low concentrations of RhaS expressed in the absence of L-rhamnose from binding to DNA. We further propose that the expression of RhaS(163-278) from even relatively low copy number plasmids is sufficiently high (compared with chromosomally expressed full-length RhaS) to overcome this DNA binding defect and to drive RhaS(163-278) onto the DNA. In contrast, similar overexpression of RhaR(196-311) is not expected to overcome limitations in the ability to contact RNAP, since the mass action advantage of overexpression no longer has an influence once a single protein binds to the DNA. Although

this aspect of our RhaS model does not incorporate the simplest hypothesis to explain the activity of RhaS(163-278), we find it to be the best model to explain the available data.

Implications for other AraC/XylS family proteins. As mentioned above, the DNA-binding domains of many multi-domain AraC/XylS family proteins appear to fall into two classes: (1) Those that are capable of high levels of transcription activation when expressed alone, including RhaS; and (2) Those that show little or no activation when expressed alone, including RhaR. Our models suggest a possible explanation for these two classes.

The AraC/XylS family proteins whose DNA binding domains activate transcription well –including XylS, UreR and XynR (Gendlina *et al.*, 2002; Kaldalu *et al.*, 2000; Miyazaki *et al.*, 2003; Poore *et al.*, 2001; Wickstrum *et al.*, 2007) – may be similar to RhaS and unable to bind DNA effectively in the absence of effector (Fig. 6B). The DNA-binding domains in each of these studies were overexpressed to some extent, in line with the expected mass action effect on DNA binding. Also consistent with this hypothesis are the findings that: two UreR mutations that confer a constitutive phenotype have been shown to increase DNA binding by the full-length protein (Gendlina *et al.*, 2002); and that XylS expression from the stronger of its two natural promoters causes it to become effector-independent (Gallegos *et al.*, 1996; Gonzalez-Perez *et al.*, 2004; Ramos *et al.*, 1987).

The AraC/XylS family proteins whose DNA binding domains do not activate transcription well – including MelR and ToxT (Howard *et al.*, 2002; Prouty *et al.*, 2005) – may be similar to RhaR and unable to effectively contact RNAP (or otherwise activate transcription at a step after DNA binding) in the absence of effector (Fig. 6C). MelR-CTD binds to DNA but does not activate transcription (Howard *et al.*, 2002). MelR is somewhat of a special case, however, given that in the absence of melibiose, MelR has been proposed to form a DNA loop that actively represses *melR* transcription (Wade *et al.*, 2000). It remains to be seen whether the

MelR mechanism involves the transmission of structural changes from its NTD to its CTD in the presence of melibiose in addition to releasing its DNA loop. Similarities in the mechanisms are suggested by the observation that several melibiose-independent variants of MelR have substitutions that align or are adjacent to substitutions we identified in RhaS and RhaR (Kahramanoglou *et al.*, 2006).

Although the available evidence is consistent with additional AraC/XylS family employing allosteric effector mechanisms similar to those we've proposed for RhaS and RhaR, there is no evidence that these mechanisms apply to AraC itself. AraC-NTD does not undergo conformational changes in its β-barrel upon arabinose binding (Soisson *et al.*, 1997b; Weldon *et al.*, 2007), arguing against the transmission of any structural changes from the NTD to the CTD in the presence of L-arabinose. The importance of aligned residues AraC R251 and RhaR T279 in their respective mechanisms, however, suggests that the AraC and RhaR mechanisms may share the common principle of preventing contacts with σ^{70} in the absence of effector.

ACKNOWLEDGEMENTS

We would like to thank members of the lab of William Picking, especially Marianela Espina and Andrew Olive, for help with Western blots and image analysis. Western blot images were obtained on an Odyssey Infrared Imager through collaboration with LI-COR Inc. We also thank Liskin Swint-Kruse (KU Med Cntr) and Shivakumara Siddaramappa for comments on the manuscript, Matthew Buechner for use of a gradient thermocycler, and Brian Easley, Remealle How and Christian Stopp for technical assistance.

This work was supported by NIH grant GM55099 from the National Institute of General Medical Sciences, NIH Grant P20 RR17708 from the Institutional Development Award (IDeA) Program of the National Center for Research Resources, and the Department of Molecular Biosciences, University of Kansas, all to S.M.E.

REFERENCES

Backman, K., Chen, Y.-M., and Magasanik, B. (1981) Physical and genetic characterization of the *gln A-glnG* region of the *Escherichia coli* chromosome. *Proc. Nat. Acad. Sci., U.S.A.* **78**: 3743-3747.

Bartolome, B., Jubee, Y.M., E. , and de la Cruz, F. (1991) Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**: 75-78.

Bhende, P.M., and Egan, S.M. (1999) Amino acid-DNA contacts by RhaS: an AraC family transcription activator. *J. Bacteriol.* **181**: 5185-5192.

Bhende, P.M., and Egan, S.M. (2000) Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with sigma 70. *J. Bacteriol.* **182**: 4959-4969.

Carra, J.H., and Schleif, R.F. (1993) Variation of half-site organization and DNA looping by AraC protein. *EMBO J.* **12**: 35-44.

Dangi, B., Gronenborn, A.M., Rosner, J.L., and Martin, R.G. (2004) Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA. *Mol Microbiol* **54**: 45-59.

Egan, S.M., and Schleif, R.F. (1993) A regulatory cascade in the induction of *rhaBAD*. *J. Mol. Biol.* **234**: 87-98.

Egan, S.M., and Schleif, R.F. (1994) DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. *J. Mol. Biol.* **243**: 821-829.

Egan, S.M. (2002) Growing repertoire of AraC/XylS activators. *J Bacteriol* **184**: 5529-5532.

Eustance, R.J., Bustos, S.A., and Schleif, R.E. (1994) Reaching out: Locating and lengthening the interdomain linker in AraC protein. *J. Mol. Biol.* **242**: 330-338.

Gallegos, M.-T., Michan, C., and Ramos, J.L. (1993) The XylS/AraC family of regulators. *Nucl. Acids Res.* **21**: 807-810.

Gallegos, M.-T., Marques, S., and Ramos, J.L. (1996) Expression of the TOL plasmid *xylS* gene in *Pseudomonas putida* occurs fom a s^{70}-dependent promoter or from s^{70}- and s^{54}-dependent tandem promoters according to the compound used for growth. *J. Bacteriol.* **178**: 2356-2361.

Gallegos, M.-T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J.L. (1997) AraC/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61**: 393-410.

Gendlina, I., Gutman, D.M., Thomas, V., and Collins, C.M. (2002) Urea-dependent signal transduction by the virulence regulator UreR. *J Biol Chem* **277**: 37349-37358.

Ghosh, M., and Schleif, R.F. (2001) Biophysical evidence of arm-domain interactions in AraC. *Anal. Biochem.* **295**: 107-112.

Gonzalez-Perez, M.M., Ramos, J.L., and Marques, S. (2004) Cellular XylS levels are a function of transcription of xylS from two independent promoters and the differential efficiency of translation of the two mRNAs. *J Bacteriol* **186**: 1898-1901.

Grainger, D.C., Belyaeva, T.A., Lee, D.J., Hyde, E.I., and Busby, S.J. (2004a) Transcription activation at the Escherichia coli melAB promoter: interactions of MelR with the C-terminal domain of the RNA polymerase alpha subunit. *Mol Microbiol* **51**: 1311-1320.

Grainger, D.C., Webster, C.L., Belyaeva, T.A., Hyde, E.I., and Busby, S.J. (2004b) Transcription activation at the Escherichia coli melAB promoter: interactions of MelR with its DNA target site and with domain 4 of the RNA polymerase sigma subunit. *Mol Microbiol* **51**: 1297-1309.

Harmer, T., Wu, M., and Schleif, R. (2001) The role of rigidity in DNA looping-unlooping by AraC. *Proc Natl Acad Sci U S A* **98**: 427-431.

Holcroft, C.C., and Egan, S.M. (2000) Interdependence of activation at *rhaSR* by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain and RhaR. *J. Bacteriol.* **182**: 6774-6782.

Howard, V.J., Belyaeva, T.A., Busby, S.J., and Hyde, E.I. (2002) DNA binding of the transcription activator protein MelR from Escherichia coli and its C-terminal domain. *Nucleic Acids Res* **30**: 2692-2700.

Jair, K., Martin, R.G., Rosner, J.L., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1995) Purification and regulatory properties of MarA protein, a transcriptional activator of *Escherichia coli* multiple antibiotic and superoxide resistance promoters. *J. Bacteriol.* **177**: 7100-7104.

Jair, K.-M., Yu, X., Skarstad, K., Thony, B., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1996a) Transcriptional activation of promoters of the superoxide and multiple antibiotic resistance regulons by Rob, a binding protein of the *Escherichia coli* origin of chromosomal replication. *J. Bacteriol.* **178**: 2507-2513.

Jair, K.-W., Fawcett, W.P., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1996b) Ambidextrous transcriptional activation by SoxS: requirement for the C-terminal domain of the RNA polymerase alpha subunit in a subset of *Escherichia coli* superoxide-inducible genes. *Molec. Microbiol.* **19**: 307-317.

Kahramanoglou, C., Webster, C.L., El-Robh, M.S., Belyaeva, T.A., and Busby, S.J. (2006) Mutational analysis of the Escherichia coli melR gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. *J Bacteriol* **188**: 3199-3207.

Kaldalu, N., Toots, U., de Lorenzo, V., and Ustav, M. (2000) Functional domains of the TOL plasmid transcription factor XylS. *J. Bacteriol.* **182**: 1118-1126.

Kolin, A., Jevtic, V., Swint-Kruse, L., and Egan, S.M. (2007) Linker regions of the RhaS and RhaR proteins. *J Bacteriol* **189**: 269-271.

Kwon, H.J., Bennik, M.H.J., Demple, B., and Ellenberger, T. (2000) Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nature Structural Biology* **7**: 424-430.

Landini, P., Gaal, T., Ross, W., and Volkert, M.R. (1997) The RNA polymerase a subunit carboxyl-terminal domain is required for both basal and activated transcription from the *alkA* promoter. *J. Biol. Chem.* **272**: 15914-15919.

LaPointe, C.F., and Taylor, R.K. (2000) The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J Biol Chem* **275**: 1502-1510.

Lobell, R.B., and Schleif, R.F. (1990) DNA looping and unlooping by AraC protein. *Science* **250**: 528.

Lonetto, M.A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase s^{70} subunit. *J. Mol. Biol.* **284**: 1353-1365.

Maloy, S.R., Stewart, V.J., and Taylor, R.K. (1996) *Genetic analysis of pathogenic bacteria.* Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.

Martin, R.G., Gillette, W.K., Martin, N.I., and Rosner, J.L. (2002) Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in Escherichia coli. *Mol Microbiol* **43**: 355-370.

Miller, J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Miyazaki, K., Miyamoto, H., Mercer, D.K., Hirase, T., Martin, J.C., Kojima, Y., and Flint, H.J. (2003) Involvement of the multidomain regulatory protein XynR in positive control of xylanase gene expression in the ruminal anaerobe Prevotella bryantii B(1)4. *J Bacteriol* **185**: 2219-2226.

Neidhardt, F.C., Bloch, P.L., and Smith, D.F. (1974) Culture medium for enterobacteria. *J. Bacteriol.* **119**: 736-747.

Poore, C.A., Coker, C., Dattelbaum, J.D., and Mobley, H.L. (2001) Identification of the domains of UreR, an AraC-like transcriptional regulator of the urease gene cluster in Proteus mirabilis. *J Bacteriol* **183**: 4526-4535.

Power, J. (1967) The L-rhamnose genetic system in Escherichia coli K-12. *Genetics* **55**: 557-568.

Prouty, M.G., Osorio, C.R., and Klose, K.E. (2005) Characterization of functional domains of the Vibrio cholerae virulence regulator ToxT. *Mol Microbiol* **58**: 1143-1156.

Ramos, J.L., Mermod, N., and Timmis, K.N. (1987) Regulatory circuits controlling transcription of TOL plasmid operon encoding meta-cleavage pathway for degradation of alkylbenzoates by Pseudomonas. *Mol Microbiol* **1**: 293-300.

Ramos, J.L., F. Rojo, L. Zhou, K. N. Timmis (1990) A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucl. Acids Res.* **18**: 2149-2152.

Reed, W.L., and Schleif, R.F. (1999) Hemiplegic mutations in AraC protein. *J Mol Biol* **294**: 417-425.

Rhee, S., Martin, R.G., Rosner, J.L., and Davies, D.R. (1998) A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc. Natl. Acad. Sci. USA* **95**: 10413-10418.

Ross, J.J., Gryczynski, U., and Schleif, R. (2003) Mutational analysis of residue roles in AraC function. *J Mol Biol* **328**: 85-93.

Ruiz, R., Ramos, J.L., and Egan, S.M. (2001) Interactions of the XylS regulators with the C-terminal domain of the RNA polymerase a subunit influence the expression level from the cognate Pm promoter. *FEBS Letters* **491**: 207-211.

Saviola, B., Seabold, R., and Schleif, R.F. (1998) Arm-domain interactions in AraC. *J. Mol. Biol.* **278**: 539-548.

Seabold, R.R., and Schleif, R.F. (1998) Apo-AraC actively seeks to loop. *J Mol Biol* **278**: 529-538.

Shah, I.M., and Wolf, R.E., Jr. (2004) Novel Protein-Protein Interaction Between Escherichia coli SoxS and the DNA Binding Determinant of the RNA Polymerase alpha Subunit: SoxS Functions as a Co-sigma Factor and Redeploys RNA Polymerase from UP-element-containing Promoters to SoxS-dependent Promoters during Oxidative Stress. *J Mol Biol* **343**: 513-532.

Soisson, S.M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997a) The 1.6 A crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. *J Mol Biol* **273**: 226-237.

Soisson, S.M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997b) Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**: 421-425.

Stewart, G.S.A.B., Lubinsky-Mink, S., Jackson, C.G., Cassel, A., and Kuhn, J. (1986) pHG165: a pBR322 copy number derivative of pUC8 for cloning and expression. *Plasmid* **15**: 172-181.

Tate, C.G., Muiry, J.A.R., and Henderson, P.J.F. (1992) Mapping, cloning, expression, and sequencing of the *rhaT* gene which encodes a novel L-Rhamnose-H^+ transport protein in *Salmonella typhimurium* and *Escherichia coli*. *J. Biol. Chem.* **287**: 6923-6932.

Tobin, J.F., and Schleif, R.F. (1987) Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. *J. Mol. Biol.* **196**: 789-799.

Tobin, J.F., and Schleif, R.F. (1990a) Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. *J. Mol. Biol.* **211**: 75-89.

Tobin, J.F., and Schleif, R.F. (1990b) Transcription from the *rha* operon p_{sr} promoter. *J. Mol. Biol.* **211**: 1-4.

Via, P., Badia, J., Baldoma, L., Obradors, N., and Aguilar, J. (1996) Transcriptional regulation of the *Escherichia coli rhaT* gene. *Microbiology* **142**: 1833-1840.

Wade, J.T., Belyaeva, T.A., Hyde, E.I., and Busby, S.J. (2000) Repression of the Escherichia coli *melR* promoter by MelR: evidence that efficient repression requires the formation of a repression loop. *Mol. Microbiol.* **36**: 223-229.

Weldon, J.E., Rodgers, M.E., Larkin, C., and Schleif, R.F. (2007) Structure and properties of a truely apo form of AraC dimerization domain. *Proteins* **66**: 646-654.

Wickstrum, J.R., and Egan, S.M. (2004) Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. *J Bacteriol* **186**: 6277-6285.

Wickstrum, J.R., Skredenske, J.M., Kolin, A., Jin, D.J., Fang, J., and Egan, S.M. (2007) Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of Its Effector-Binding Domain. *J Bacteriol* **189**: 4984-4993.

Wu, M., and Schleif, R. (2001a) Strengthened arm-dimerization domain interactions in AraC. *J Biol Chem* **276**: 2562-2564.

Wu, M., and Schleif, R. (2001b) Mapping arm-DNA-binding domain interactions in AraC. *J Mol Biol* **307**: 1001-1009.

Zhou, Y.H., Zhang, X.P., and Ebright, R.H. (1991) Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucleic Acids Res* **19**: 6052.

CHAPTER 3

The AraC/XylS Family Activator RhaS Negatively Autoregulates *rhaSR* Expression by Preventing Cyclic AMP Receptor Protein Activation[†]

Jason R. Wickstrum^, Jeff M. Skredenske,

Vinitha Balasubramaniam, Kyle Jones and Susan M. Egan

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas

^Current address: Kansas Health and Environmental Laboratories, Topeka, KS

[†]This article is the accepted manuscript version of a subsequently published paper:

Wickstrum, J.R., J.M. Skredenske, V. Balasubramaniam, K. Jones and S.M. Egan. 2010. The AraC/XylS Family Activator RhaS Negatively Autoregulates *rhaSR* Expression by Preventing Cyclic AMP Receptor Protein Activation. *J. Bacteriol.* **192**:225-232.

The published version of this article is available online at:

http://dx.doi.org/10.1128/JB.00829-08.

ABSTRACT

The *Escherichia coli* RhaR protein activates expression of the *rhaSR* operon in the presence of its effector, L-rhamnose. The resulting RhaS protein (plus L-rhamnose) activates expression of the L-rhamnose catabolic and transport operons, *rhaBAD* and *rhaT*, respectively. Here, we further investigated our previous finding that *rhaS* deletion resulted in a three-fold increase in *rhaSR* promoter activity, suggesting RhaS negative autoregulation of *rhaSR*. We found that RhaS autoregulation required the CRP binding site at *rhaSR*, and that RhaS was able to bind to the RhaR binding site at *rhaSR*. In contrast to the expected repression, we found that in the absence of both RhaR and the CRP binding site at the *rhaSR* promoter, RhaS activated expression to a level comparable with RhaR activation of the same promoter. However, when the promoter included the RhaR and CRP binding sites, activation by RhaS and CRP was to a much lower level than activation by RhaR and CRP, suggesting that CRP could not fully co-activate with RhaS. Taken together, our results indicate that RhaS negative autoregulation involves RhaS competition with RhaR for binding to the RhaR binding site at *rhaSR*. Although RhaS and RhaR activate *rhaSR* transcription to similar levels, CRP cannot effectively co-activate with RhaS. Therefore, once RhaS reaches a relatively high protein concentration, presumably sufficient to saturate the RhaS-activated promoters, there will be a decrease in *rhaSR* transcription. We propose a model in which differential DNA bending by RhaS and RhaR may be the basis for the difference in CRP co-activation.

INTRODUCTION

The *Escherichia coli rhaSR* operon encodes two L-rhamnose-responsive members of the AraC/XylS family of transcription activator proteins, RhaS and RhaR (28). In the presence of L-rhamnose, RhaS and RhaR are required to activate transcription of the operons encoding the L-rhamnose catabolic enzymes (*rhaBAD*) (Fig. 1) and the L-rhamnose uptake protein (*rhaT*). The sole identified function of RhaR is to activate transcription of *rhaSR*, and thereby to increase RhaS protein concentration in the presence of L-rhamnose, while RhaS directly activates transcription of the *rhaBAD* and *rhaT* operons (11, 12, 27-30). CRP is required for full expression of all three of the *rha* operons, but functions as a co-activator that only substantially activates transcription when RhaS or RhaR also bind to the promoter regions (11, 15, 31). CRP activation at the *rhaSR* operon was shown to require the RNAP α-subunit C-terminal domain (α-CTD) (33). It is likely that CRP co-activation also involves contacts with α-CTD at the *rhaBAD* and *rhaT* promoters. RhaS or RhaR may be required to bend the DNA to allow CRP to co-activate by contacting α-CTD, similar to other promoters (8).

RhaS and RhaR are 30% identical to each other and likely arose by gene duplication. Both proteins function as homodimers and comprise two domains, an N-terminal dimerization and L-rhamnose binding domain, and a C-terminal DNA binding domain [(34) and Hunjan, Kolin and Egan, unpublished]. Flexible linkers connect the RhaS and RhaR domains; however the sequences of the linkers do not appear to be critical for function (18). The RhaS and RhaR DNA binding sites consist of two imperfect 17-bp inverted repeat half-sites that are separated by 16 or 17 bp (12, 29). The downstream half-sites overlap the -35 hexamer by 4 bp, placing RhaS and RhaR in position to make protein-protein contacts with RNAP σ^{70} to activate transcription (6, 32). The RhaR binding site upstream of *rhaSR* contains four phased A-tracts (14), and is bent to an estimated angle of 70° in the absence of RhaR, and 160° upon RhaR binding (29). A single

four-bp A-tract is present in the RhaS binding site, suggesting that this DNA sequence is likely less dramatically bent.

Our previous results suggested that in addition to activation of *rhaBAD* and *rhaT*, RhaS likely also negatively autoregulates its own expression. Deletion of *rhaS* resulted in a three-fold increase in expression of a *rhaS-lacZ* translational fusion (extending from 312 bp upstream of the transcriptional start point through codon 20 of *rhaS*), while overexpression of *rhaS* resulted in a decrease in expression of the same fusion (11). It was proposed that this RhaS autoregulation could, in principle, occur by: formation of inactive RhaS-RhaR heterodimers; a DNA-looping mechanism in which RhaS bound to its site at the *rhaBAD* promoter would inhibit activation by RhaR bound to its site at the *rhaSR* promoter; or RhaS competing for binding to the RhaR binding site at the *rhaSR* promoter (11). Here, we have further investigated the hypothesis that RhaS negatively autoregulates its own expression as well as the mechanism of this *rhaSR* autoregulation. Our results suggest a somewhat novel mechanism in which RhaS itself is capable of activating *rhaSR* expression from the RhaR binding site to nearly as high a level as RhaR. However, the CRP contribution to *rhaSR* activation is substantially lower when RhaS is the primary activator than when RhaR is the primary activator, resulting in a relative decrease in *rhaSR* expression. Differences in DNA bending by RhaS and RhaR may play a role in the differential ability of CRP to co-activate with RhaS versus RhaR. RhaS negative autoregulation likely functions to limit positive autoregulation of *rhaSR* by RhaR.

MATERIALS AND METHODS

Culture media and conditions. *Escherichia coli* cultures for β-galactosidase assays were grown in MOPS (3-[*N*-morpholino]propanesulfonic acid)-buffered medium supplemented with 0.4% glycerol, 0.2% casamino acids, and 0.002% thiamine, (5) using the protocol of Neidhardt *et al*. (20). Tryptone broth (TB; 0.8% tryptone, 0.05% NaCl, pH 7.0) supplemented with 0.2% maltose was used to grow cultures for bacteriophage λ infection. Tryptone-yeast extract (TY) liquid medium (0.8% tryptone, 0.05% yeast extract, and 0.05% NaCl, pH 7.0) was used to grow cells for most other experiments, and was supplemented with 5 mM $CaCl_2$ to grow cultures for bacteriophage P1 infection. Antibiotics were used as indicated at 200 μg/ml for ampicillin and 25 μg/ml for chloramphenicol. All cultures were grown at 37°C with aeration.

General Methods. Standard methods were used for restriction endonuclease digestion and ligation. Transformation was carried out using chemically induced competent cells of *E. coli* and plasmid DNA was purified by alkaline lysis (7) or using an EZNA Plasmid Mini Kit I (Omega Bio-Tek, Inc.; Norcross, GA). Oligonucleotides were synthesized by MWG-Biotech (High Point, NC). The Northwestern University Genomics Core (Chicago, IL) performed DNA sequencing reactions. The DNA sequence of both strands was determined for the entire cloned region of all cloned DNA fragments. The Expand High Fidelity PCR System (Roche; Indianapolis, IN) was used to amplify non-mutagenized DNA fragments for cloning. The QIAquick PCR Purification Kit (Qiagen; Chatsworth, CA) was used to clean up PCR products. β-Galactosidase assays were performed using the method of Miller, as previously described (5, 19). Specific activities were averaged from at least two independent assays, with two replicates per assay. RhaS fold-repression values were calculated by dividing vector-only values by $rhaS^+$ values in $rhaR^+$ strains, while RhaS fold-activation values were calculated by dividing $rhaS^+$ values by vector-only values in Δ*rhaR* strains.

Strains, phage and plasmids. All strains in this study were derived from ECL116 (1). Plasmid-borne RhaS expression in all experiments was from promoters that were independent of RhaS, RhaR and L-rhamnose. The promoter was either the *lac* promoter (pHG165 and pSU18), or a constitutive promoter with the *rhaSR* -10 element, a near perfect -35 element (TTGACT), and no known upstream binding sites.

The *lacZ* fusions are named such that "Φ" stands for "fusion" and the upstream endpoint of each fusion relative to the transcription start site (for example, -84, but without the minus sign) is given after the "Δ." The downstream endpoint of all fusions, unless otherwise indicated, was within codon 20 of *rhaS*, at +84 relative to the transcription start site. The exceptions had downstream endpoints within codon 20 of *rhaR*, at +904 relative to the *rhaSR* transcription start site. Most *lacZ* fusions were translational fusions, unless otherwise noted. These were first constructed in plasmid pRS414, while transcriptional fusions were first constructed in pRS415 (25). The desired insert was generated by high-fidelity PCR of the promoter region of interest. Oligonucleotide 896 (11) was the downstream primer used to amplify *rhaSR* promoter regions with downstream endpoints at +84. Oligonucleotide 2832 (5'-GCGGGATCCTTATTCGCAATATGGCGTAC-3') was the downstream primer for the fusions with downstream endpoints at +904. The *lacZ* fusions in pRS414 or pRS415 were recombined onto the genome of bacteriophage λ and integrated into the *E. coli* chromosome as single copy lysogens (25). Several single copy lysogen candidates were tested using β-galactosidase assays to distinguish likely single copy lysogens from multiple lysogens based on activity levels. Single copy lysogens were confirmed using a PCR test as well as the Ter test (13, 21). Phage P1 transduction was used to introduce the Δ(*rhaSR*)::*kan* or Δ*rhaS rhaR*$^+$ (linked to *zih-35*::Tn*10*) alleles into strains as required.

Three new *rhaSR* promoter regions fused to *lacZ*, each with a different upstream endpoint, were constructed to identify the DNA elements required for RhaS autoregulation. These *rhaS-lacZ* fusions were constructed as described above using oligonucleotides 2727 (5'-GTCGAATTCTTTCCTGAAAATTCACGCTG-3'), 2726 (5'-GTCGAATTCTGCTCACCGCATTTCCTG-3'), and 2725 (5'-GGCGAATTCTGATGTGATGCTCACCGC-3') in combination with 896 to make Φ(*rhaS-lacZ*)Δ103, Φ(*rhaS-lacZ*)Δ114, and Φ(*rhaS-lacZ*)Δ122, respectively.

Two variant *rhaSR* promoters positioned the native *rhaSR* CRP binding site at -92.5 and -93.5, similar to the CRP site at the *rhaBAD* promoter (11). Oligonucleotides 2867 (5'-CACGAATTCTGTGATGCTCACCGCAGTATCTTGAAAAATCGACG-3') and 2868 (5'-CACGAATTCTGTGATGCTCACCGCATGTATCTTGAAAAATCGACG-3') were used with oligonucleotide 896, and placed the CRP binding site (underlined) 2 bp and 3 bp upstream of the RhaR binding site, respectively. The RhaS binding site at *rhaBAD* was also replaced with a RhaR binding site (with or without an upstream CRP site) by using a natural *Eco*RI site just downstream of the RhaS binding site.

Random mutagenesis of *rhaS* was performed as previously described (17). Briefly, the *rhaS* gene was PCR amplified with *Taq* polymerase under standard reaction conditions (36) and cloned into pHG165 (26).

Electrophoretic mobility shift assays. Electrophoretic mobility shift assays were performed as previously described (15), with the following modifications. The DNA fragments were generated by PCR with one primer labeled with ^{32}P at the 5' end (using T4 polynucleotide kinase) and one unlabeled primer. The 49 bp DNA fragments each contained one 17 bp RhaS or RhaR binding half-site (*rhaI*$_1$, *rhaI*$_2$, *rhaI*$_3$ or *rhaI*$_4$) flanked by the same 16 bp sequences for each half site. The flanking sequences were the same as the previously published half-site

fusions with *lacZ* (34). The mobility shift assay buffer did not contain Nonidet P-40 or cAMP, and contained 0.5 mM Tris[2-carboxyethyl]-phosphine (TCEP) instead of 1 mM dithiothreitol. His6-RhaS(163-278) was purified as previously described (34). After electrophoresis, the gels were dried, exposed to a phosphor screen, and analyzed using the Cyclone Storage Phosphor System (Packard).

DNase I footprinting. The DNA substrate for footprinting reactions was prepared by PCR amplification of the *rhaSR-rhaBAD* intergenic region using primers 2371 and 2409 (33) with one or the other of the primers ^{32}P-labeled in separate footprinting reactions. DNase I footprinting was performed as previously described (12).

RESULTS AND DISCUSSION

RhaS autoregulation of *rhaSR* requires the CRP binding site but not the RhaS binding site at *rhaBAD*. Our first goal was to identify the DNA element(s) in the *rhaSR-rhaBAD* intergenic region required for RhaS autoregulation (Fig. 1). To do this, we assayed the ability of RhaS to repress *rhaSR* expression from a number of single copy *lacZ* fusions with different upstream endpoints, in a strain carrying an in-frame deletion of the majority of the *rhaS* gene, Δ*rhaS rhaR*$^+$ (11). Here and throughout this work, we expressed RhaS from a moderate copy number plasmid to increase the repression by RhaS, and thereby enable detection of otherwise relatively small effects. The strains in which these assays were performed are represented schematically in Fig. 2, lines 2 and 3. Given the precedence of AraC protein forming a repressing DNA loop [reviewed in (22)], one question this experiment addressed was whether RhaS autoregulation might involve a DNA looping mechanism in which RhaS, bound to its site at the *rhaBAD* promoter, interacts with and inhibits activation by RhaR bound at the *rhaSR* promoter. The longest promoter fusion in these experiments was the same one that initially provided evidence of RhaS autoregulation (11). It extended upstream to -312 (relative to the *rhaSR* transcription start site) and included the RhaS binding site at the divergent *rhaBAD* promoter (Fig. 1).

We found that expression of RhaS from a plasmid resulted in 30-fold or greater decreases in *rhaSR* promoter expression (Table 1) with fusions that had upstream endpoints in the range -122 to -312 (Fig. 1). The RhaS binding site at *rhaBAD* is located between the -216 and -128 deletion endpoints, however, its deletion did not result in a significant change in RhaS autoregulation. This indicates that the RhaS binding site at *rhaBAD* is not required for RhaS negative autoregulation of *rhaSR*, and rules out a mechanism in which RhaS represses *rhaSR* expression by forming a DNA loop from its site at *rhaBAD*.

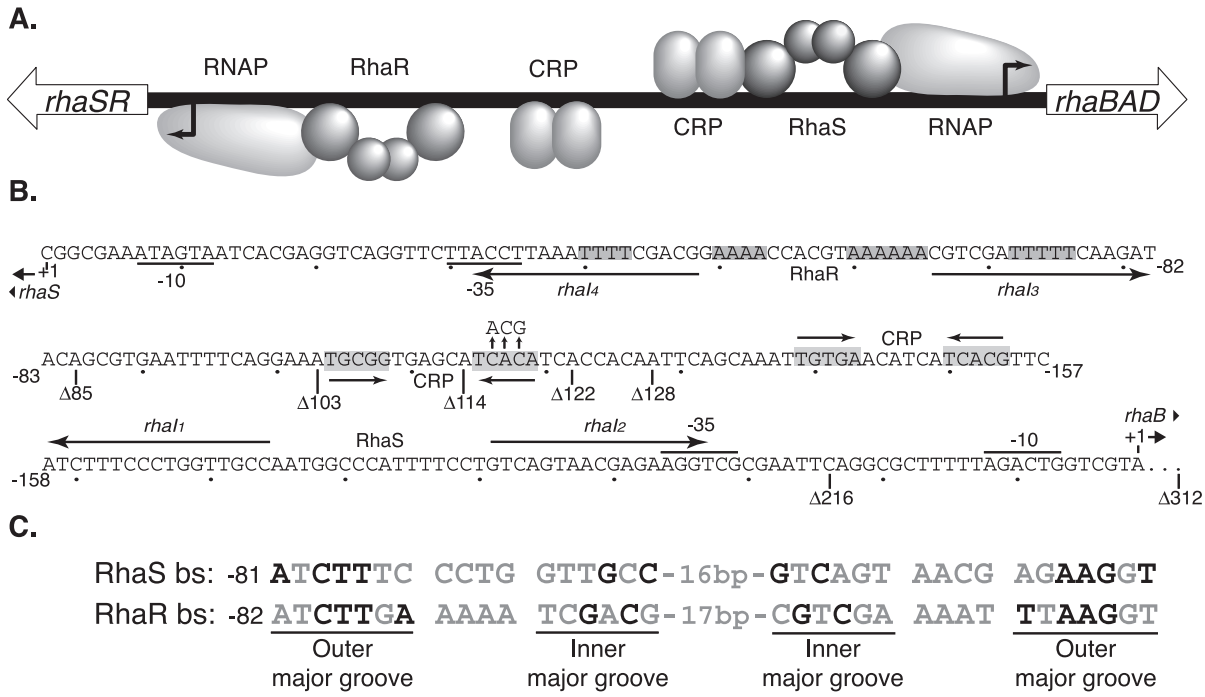

Figure 1. *rhaSR-rhaBAD* intergenic region. A) A schematic representation of the *rhaSR-rhaBAD* intergenic region. The relative positions of the RNA polymerases (RNAP), and the activator proteins RhaS, RhaR and CRP are shown. The activators and sites shown above the line all influence *rhaBAD* expression, while the activators and sites shown below the line influence *rhaSR* expression. B) The DNA sequence between the *rhaBAD* and *rhaSR* transcription start sites. The positions of the RhaS and RhaR binding sites are shown by arrows labeled with the half-site names (*rhaI*$_1$ through *rhaI*$_4$). The CRP binding site positions are shown as inverted arrows with shading, and the substitutions in the CRP binding site consensus positions are shown with vertical arrows above the CRP site. The -10 and -35 regions of the promoters are marked. Binding sites important at the *rhaBAD* promoter are shown above the line, while deletion endpoints (marked Δ), binding sites and distances relative to the *rhaSR* promoter are shown below the line. The 70 bp sequence between the *rhaBAD* transcription start site and the *rhaSR* -312 position is not shown. C) The sequences of the RhaS and RhaR binding sites are aligned. The black letters indicate those base pairs found important for sequence recognition by the activators, with all other letters in gray. The outer and inner major grooves of the inverted repeat half-sites are labeled.

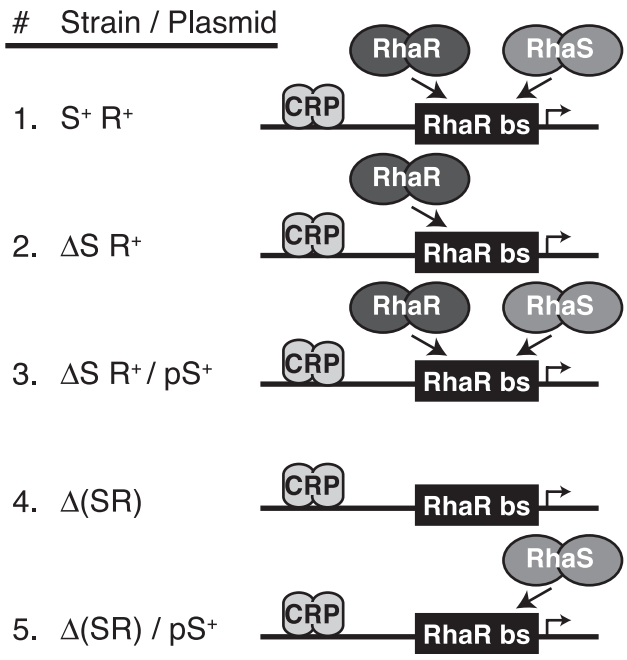

Figure 2. Schematic representation of strains used in this study. The relevant strain backgrounds used in many of these studies are represented schematically. In the strain designations: *rhaS* and *rhaR* are S and R, respectively; deletions of *rhaS* or both *rhaS* and *rhaR* are written as ΔS and Δ(SR), respectively; plasmid-borne *rhaS* is pS^+. DNA is a horizontal line; the RhaR binding site is a black box labeled "RhaR bs"; the RhaS, RhaR and CRP proteins are pairs of gray ovals, as labeled. The direction of transcription is indicated by bent arrows. RhaS or RhaR binding to the RhaR binding site is represented by arrows, with competition represented by two proteins attempting to bind to the same site.

Table 1. RhaS repression of various *rhaSR* promoter fusions

Φ(*rhaS-lacZ*)

Promoter	β-Galactosidase activity[a]		Fold RhaS
Truncation	Vector-only[b]	RhaS[b]	Repression
D85	8.9	7.0	1.3
D103	10	2.7	3.7
D114	12	2.9	4.1
D122	155	5.3	29
D128	204	6.9	30
D216	234	7.1	33
D312	203	4.9	41
D312 CRP^-	22	3.5	6.3

[a] β-Galactosidase activity was assayed from single-copy *rhaS-lacZ* fusions with the upstream promoter endpoints indicated. The Δ312 CRP^- promoter has point mutations in three consensus positions of the *rhaSR* promoter CRP binding site. Cultures were grown in minimal MOPS growth medium containing ampicillin, and L-rhamnose. The strain background was Δ*rhaS rhaR*$^+$ *zih-35*::Tn*10 recA*::*cat*. Standard error was less than 12% of the average units.

[b] The vector was pSE262 (34), which is pHG165 (26) with a stronger promoter driving expression. RhaS was expressed from plasmid pSE265 (34), which is pSE262 *rhaS*$^+$.

RhaS expression reduced expression from *rhaSR* promoter fusions with endpoints between -114 and -85 by only 1.3- to 4.1-fold (Table 1). The upstream endpoint of the CRP binding site required for full *rhaSR* activation is at -119, therefore the Δ114 fusion did not include the entire CRP binding site, whereas the Δ122 fusion did (Fig. 1). This suggested that the CRP binding site is required for maximal repression by RhaS. To further test this hypothesis, we assayed a fusion with an upstream endpoint at -312 and carrying point mutations at three consensus positions in the CRP binding site (Fig. 1) that we previously found greatly reduced CRP co-activation of *rhaSR* (34). We found that the CRP binding site point mutations reduced RhaS repression of the Δ312 fusion by more than six-fold, to a level similar to the Δ114 fusion (Table 1, Δ312 CRP^-). Taken together, these results support the hypothesis that the CRP binding site at *rhaSR*, but not the RhaS binding site at *rhaBAD*, is required for the RhaS-dependent decrease in *rhaSR* expression, or negative autoregulation.

RhaS binds to the RhaR binding site at *rhaSR*. We next considered the hypothesis that RhaS might compete with RhaR for binding to the RhaR binding site at the *rhaSR* promoter. As previously noted (5, 12), the DNA sequences of the outer major grooves of the RhaS and RhaR binding sites at the *rhaBAD* and *rhaSR* promoters, respectively, are nearly identical (Fig. 1C), raising the possibility that RhaS might be capable of binding to the RhaR binding site. We tested this using *in vitro* electrophoretic mobility shift assays with 49 bp DNA fragments, each containing one of the 17 bp half-sites for RhaR or RhaS binding, and identical flanking DNA sequences. Given that full-length RhaS protein severely aggregates when overexpressed, we used the purified C-terminal DNA binding domain of RhaS [His6-RhaS(163-278), previously called His6-RhaS-CTD, (34)] for these assays. His6-RhaS(163-278) was previously found capable of *in vitro* DNA binding and transcription activation (34). Using electrophoretic mobility shift assays, we found that His6-RhaS(163-278) was able to bind *in vitro* to DNA

fragments carrying each of the RhaR half-sites at *rhaSR*, although three- to ten-fold less tightly than to the RhaS half-sites at *rhaBAD* (Fig. 3). This apparent weaker binding is consistent with the expectation that an autoregulation mechanism would involve RhaS binding to the RhaR binding site only at RhaS protein concentrations above those necessary to saturate the RhaS binding sites at the *rhaBAD* and *rhaT* promoters. DNase I footprinting confirmed that His6-RhaS(163-278) bound to the expected RhaR half-sites (data not shown).

RhaS activates *rhaSR* expression, but CRP co-activation is reduced. The finding that RhaS can bind to the RhaR binding site at *rhaSR* suggested a model in which, upon reaching saturating levels in the cell, RhaS might compete for binding to the RhaR binding site and repress expression of *rhaSR*. However, the promoter-proximal RhaR half-site at *rhaSR* is identically positioned relative to the promoter as the similar RhaS half-site at *rhaBAD*, suggesting that RhaS might activate rather than repress transcription from this site. To test whether RhaS was able to activate *rhaSR* expression, we assayed expression from single copy *rhaS-lacZ* fusions in a strain lacking RhaR [Δ(*rhaSR*) strain background], with RhaS expressed from the vector pSU18 (2) [plasmid pSE273 (34)]. It was not possible to use a *rhaS*$^+$ Δ*rhaR* strain since *rhaS* expression requires RhaR. For comparison, we also assayed the same fusions in a strain expressing RhaR from the chromosome [Δ*rhaS rhaR*$^+$ strain background], again with RhaS expressed from plasmid pSE273 (34). In the Δ*rhaS rhaR*$^+$ strain, the vector-only samples represent the level of activation by RhaR, while the addition of the RhaS-encoding plasmid shows the RhaS-mediated reduction of the RhaR activation (Fig. 2, lines 2 and 3). In the Δ(*rhaSR*) strain, the vector-only samples represent the basal level of expression from the fusions, and the addition of the RhaS-encoding plasmid shows any ability of RhaS to activate expression of the fusions (Fig. 2, lines 4 and 5).

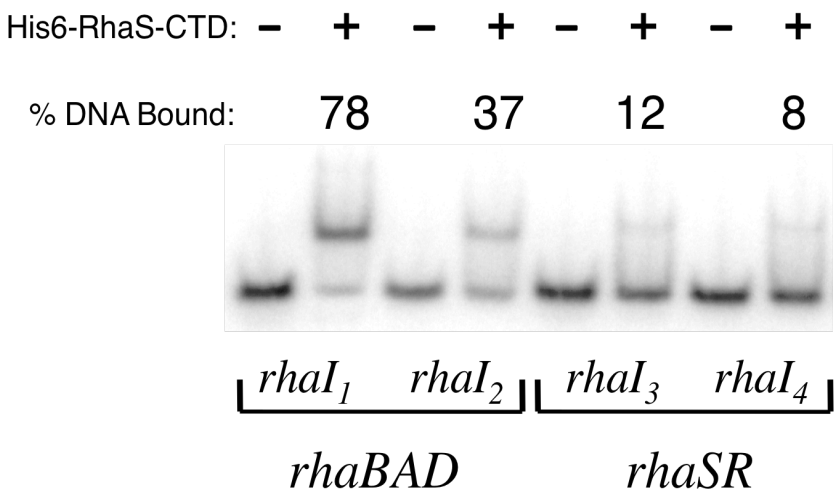

Figure 3. Electrophoretic mobility shift assays of His6-RhaS(163-278) binding to RhaS and RhaR half-sites. Electrophoretic mobility shift reactions were carried out in the absence (-) or presence (+) of 6 μM purified His6-RhaS(163-278) with 49 bp ^{32}P-labeled DNA fragments including each of the RhaS half-sites at *rhaBAD*, and the RhaR half-sites at *rhaSR*. The direction of electrophoresis was from the top down, as shown. The % DNA bound was averaged from four independent assays.

Similar to the previous experiment (Table 1), in the *rhaR*$^+$ strain there was a 45 fold-repression by RhaS from the fusion that included the CRP binding site [(*rhaS-lacZ*)Δ128], but only a 1.5 fold-repression from the fusion that did not include the CRP binding site [(*rhaS-lacZ*)Δ85] (Table 2). We found that RhaS activated transcription of the (*rhaS-lacZ*)Δ85 fusion to a level similar to RhaR (4.5 and 6.4 Miller units, respectively), but activated the (*rhaS-lacZ*)Δ128 fusion to a much lower level than did RhaR (17 and 630 Miller units, respectively). The increased expression from the longer of these fusions is due to the contribution of CRP to the activation (15, 34), which in this case was nearly 100-fold in combination with RhaR, but less than four-fold in combination with RhaS. These results indicate that when RhaS binds to the RhaR binding site, RhaS itself is able to activate transcription nearly as well as RhaR. However, the contribution to activation by CRP at the promoter activated by RhaS was greatly reduced compared with the CRP-RhaR combination, thereby resulting in a relative repression of *rhaSR* expression upon activation by RhaS.

Upon addition of the RhaS-expressing plasmid, the expression levels from each fusion were nearly the same regardless of whether RhaR was present (4.4 vs. 4.5 and 14 vs. 17 Units) (Table 2). This suggests that under these conditions RhaR does not significantly contribute to the activation and that RhaS may fully out-compete RhaR for binding to the RhaR binding site. We don't expect RhaS will fully out-compete RhaR under physiological conditions (without RhaS overexpression). However, given the values in Table 2 and the three to four-fold increase in *rhaSR* expression upon deletion of chromosomal *rhaS* (11), we calculate that RhaS represents approximately 67% of the protein bound at the *rhaSR* promoter in wild-type cells in the presence of L-rhamnose.

RhaS is more highly expressed than RhaR. Given our finding that RhaS binds relatively weakly to the RhaR binding site (Fig. 3), we were surprised by the estimate that RhaS

Table 2. Repression and activation of *rhaSR* expression by RhaS

	β-Galactosidase activity (Miller Units)[a]					
	Δ*rhaS rhaR*$^{+}$ [b]			Δ(*rhaSR*)[b]		
	Vector[c] only	*rhaS*$^{+}$ [c]	Fold RhaS Repression	Vector[c] only	*rhaS*$^{+}$ [c]	Fold RhaS Activation
Φ(*rhaS-lacZ*)Δ85	6.4	4.4	1.5	0.17	4.5	26
Φ(*rhaS-lacZ*)Δ128	630	14	45	0.84	17	20

[a] β-Galactosidase activity was measured from single-copy *lacZ* fusion strains grown in minimal MOPS growth medium containing chloramphenicol and L-rhamnose. Standard errors were less than 30% of the average units.

[b] The strain background was either Δ*rhaS rhaR*$^{+}$ *zih-35*::Tn*10* or Δ(*rhaSR*)::Km.

[c] The vector was pSU18 (2). RhaS was expressed from plasmid pSE273 (34), which is pSU18 *rhaS*$^{+}$.

outcompetes RhaR for binding at *rhaSR* two-thirds of the time (at steady state levels in the presence of L-rhamnose). One potential explanation for this finding is that, although *rhaS* and *rhaR* are transcribed together, the level of protein synthesized from the two genes might differ. Inspection of the Shine-Dalgarno (SD) sequences upstream *rhaS* and *rhaR* suggests this might be the case, as the *rhaS* SD sequence is predicted to be strong (5'-AGGAGGC-3'), while the *rhaR* gene has a SD sequence predicted to be a much weaker (5'-GCCAGGG-3') relative to the consensus sequence (5'-AGGAGGT-3') (24).

To test the hypothesis that *rhaS* and *rhaR* are translated at different levels, we constructed three new *lacZ* fusions driven by the *rhaSR* promoter in addition to the (*rhaS-lacZ*)Δ128 translational fusion used in previous experiments. The first new *lacZ* fusion had the same promoter region endpoints at -128 and +84 as the previous fusion, but was a transcriptional fusion. We refer here to the fusions with endpoints at -128 and +84 as "short" fusions. The other two were a transcriptional and a translational fusion, both with endpoints at -128 and +904 ("long" fusions). The +904 downstream endpoint is within codon 20 of *rhaR*, therefore, the long fusions measure the relative transcription and translation of *rhaR*. The expression levels from the short and long transcriptional fusions were essentially the same as each other, as expected (Table 3). However, the short translational fusion (which was a measure of *rhaS* translation) was 28-fold more highly expressed than the long translational fusion (which was a measure of *rhaR* translation). These results suggest that RhaS is expressed to a significantly higher level than RhaR, and may explain how RhaS is able to dominate in the competition for binding to the RhaR binding site at steady state levels in the presence of L-rhamnose.

RhaS and RhaR differ in their optimal CRP binding site positions. It is likely relevant that although the RhaS and RhaR binding sites at *rhaBAD* and *rhaSR* are identically positioned (Fig. 1), the CRP [TGTGA motifs (4)] and RhaS sites at *rhaBAD* are separated by 3

Table 3. Comparison of transcription and translation of the *rhaS* and *rhaR* genes

		β-Galactosidase activity (Miller Units)[a]	
	Fusion type	(-) rha	(+) rha
Φ(*rhaS-lacZ*)(-128 → +84) [b]	transcriptional	20	1330
Φ(*rhaSR-lacZ*)(-128 → +904) [b]	transcriptional	26	1540
Φ(*rhaS-lacZ*)(-128 → +84) [b]	translational	0.04	170
Φ(*rhaSR-lacZ*)(-128 → +904) [b]	translational	0.02	6.0

[a] β-Galactosidase activity was measured from single-copy *lacZ* fusion strains grown in minimal MOPS growth medium with or without L-rhamnose, as indicated. Standard errors were less than 8% of the average units, except for values less than 1.0, where standard error was less than 25%.

[b] The strain background was ECL116 (1).

bp, while the CRP and RhaR sites at *rhaSR* are separated by 21 bp (12, 16). In both cases, there is little or no activation by CRP in the absence of RhaS or RhaR [Table 2 and (16)]; presumably CRP requires DNA bending by the primary activator to enable contacts with α-CTD (8). We've shown that CRP co-activation requires α-CTD at the *rhaSR* promoter (34). Tobin and Schleif (29) found that the RhaR binding site DNA at *rhaSR* was bent to approximately 70 degrees when alone (presumably due to the four phased A-tracts within the RhaR binding site, see Fig. 1), but that this bend was increased to approximately 160 degrees upon RhaR binding. There is at most one A-tract in and around the RhaS binding site (Fig. 1), suggesting that the *rhaBAD* promoter DNA is likely to be significantly less bent. Taken together, these results suggest that, in addition to its own DNA bending (23), CRP may require DNA bending by a second activator protein in order to co-activate transcription at *rhaSR* and *rhaBAD*, but that there might be differences in the extent of the DNA bending required at the two promoters.

We hypothesized that RhaS and RhaR differ in the position of the upstream CRP binding site that is optimal for co-activation. To test this hypothesis, we wished to move the CRP binding site at *rhaSR* to the position of the CRP binding site at *rhaBAD*. However, two possible lines of reasoning could identify the optimal position for the CRP site at *rhaSR*. The first was that the position of the CRP binding site relative to the transcription start site was the important parameter. We therefore constructed a *rhaS-lacZ* fusion in which the CRP site was positioned at -92.5 relative to the transcription start site, the same position as the site at *rhaBAD* (11). This placed the CRP site 2 bp upstream of the RhaR binding site [(*rhaS-lacZ*)CRP^{-2}]. The second possibility was that the distance upstream of the RhaR binding site was the key parameter. Since the CRP site at *rhaBAD* is 3 bp upstream of the RhaS binding site, we placed the CRP site in the second *rhaS-lacZ* fusion 3 bp upstream of the RhaR binding site [at -93.5, (*rhaS-lacZ*)CRP^{-3}].

We compared the expression from the *rhaS-lacZ* fusions with CRP sites 2 and 3 bp upstream of the RhaR binding site to expression from a fusion with no CRP binding site [(*rhaS-lacZ*)Δ85], and one with a CRP site at the wild-type position [(*rhaS-lacZ*)Δ128] (Table 4). The expression levels in the Δ(*rhaSR*) strain indicated that the basal expression levels were similar in all cases. In the Δ*rhaS rhaR*$^{+}$ strain, where RhaR but not RhaS can activate, we found that the wild-type CRP binding site [(*rhaS-lacZ*)Δ128] enabled a 30-fold increase in activation compared with no CRP binding site [(*rhaS-lacZ*)Δ85] (216 vs. 6.9 Units). In contrast, the addition of the CRP sites 2 or 3 bp upstream of the RhaR binding site did not result in increased expression compared with the (*rhaS-lacZ*)Δ85, indicating that there was no contribution of CRP to activation in combination with RhaR from these CRP site positions. This finding indicates that CRP can function as a co-activator from a site 21 bp, but not 2 or 3 bp, upstream of RhaR, and is consistent with the hypothesis that RhaS and RhaR differ in the position of the CRP binding site that is optimal for co-activation.

We also tested the ability of CRP to co-activate with RhaS at the CRP^{-2} and CRP^{-3} constructs by using a Δ(*rhaSR*) strain expressing RhaS from a plasmid (Table 4). When RhaS was the primary activator, a CRP site 2 bp upstream of the RhaR binding site resulted in a seven-fold reduction in expression [compared with no CRP binding site, (*rhaS-lacZ*)Δ85]. This suggests that CRP binding may interfere with RhaS binding at this promoter (a small reduction in expression was also seen with RhaR as the primary activator at this construct). However, the CRP site 3 bp upstream of RhaS (bound at the RhaR binding site) enabled a 3.5-fold increase in expression relative to no CRP binding site [(*rhaS-lacZ*)Δ85]. This was very similar to the CRP contribution to expression with the CRP site at the wild-type position [(*rhaS-lacZ*)Δ128] and RhaS as the primary activator. Therefore, although we expected the CRP site in the CRP^{-3} construct to be optimally positioned for CRP co-activation with RhaS, there was no increase in

Table 4. Influence of CRP binding site position on CRP co-activation of *rhaSR*

	β-Galactosidase activity (Miller Units)[a]		
	Δ(*rhaSR*) [b]	Δ*rhaS rhaR*$^{+}$ [b]	Δ(*rhaSR*) p*rhaS*$^{+}$ [b]
Φ(*rhaS-lacZ*)Δ85	0.20	6.9	7.4
Φ(*rhaS-lacZ*)Δ128	0.38	216	31
Φ(*rhaS-lacZ*)CRP^{-2}	0.40	4.3	1.1
Φ(*rhaS-lacZ*)CRP^{-3}	0.15	7.1	26

[a] β-Galactosidase activity was measured from single-copy *lacZ* fusion strains grown in minimal MOPS growth medium containing ampicillin and L-rhamnose. The CRP^{-2} and CRP^{-3} constructs replaced the native CRP site (20 bp upstream of RhaR binding site) with the same sequence 2 or 3 bp upstream, as described in text. Standard errors were less than 23% of the average units.

[b] The strain background was either Δ*rhaS rhaR*$^{+}$ *zih-35*::Tn*10* or Δ(*rhaSR*)::*kan*. Each strain was transformed with empty vector, pHG165 (26), or plasmid pSE289, which is pHG165 *rhaS*$^{+}$ (17).

CRP co-activation compared with the wild-type CRP position at *rhaSR*. One explanation for this finding might be that the four phased A-tracts within the RhaR binding site (Fig. 1) increase the extent of DNA bending by RhaS at *rhaSR* compared with RhaS binding at *rhaBAD*. This increase in DNA bending could change the position/geometry of the CRP protein and potentially decrease the ability of CRP to contact α-CTD.

These results support the hypothesis that the optimal CRP binding site position differs for RhaS versus RhaR. The position of the CRP site at *rhaBAD* (-92.5) is a fairly common class III activator position, and is similar to the position of CRP at *araBAD* as well as the optimal upstream CRP site position in studies with tandem CRP sites at class I and class II positions (3, 35). The position of the CRP site at *rhaSR* (-111.5) is less typical, and in the tandem CRP site studies, CRP sites near this position made little if any contribution to activation (3). As such, we hypothesize that the four phased A-tracts may contribute to the ability of CRP to activate well from such a distant site at *rhaSR*. Although the A-tracts may explain why CRP did not co-activate well with RhaS at (*rhaS-lacZ*)CRP^{-3}, they do not explain why CRP also did not co-activate well with RhaS at (*rhaS-lacZ*)Δ128, in spite of co-activating well with RhaR at this promoter. This finding indicates that there must be a difference between RhaS and RhaR that contributes to their respective optimal CRP site positions, perhaps a difference in the extent of DNA bending by the two proteins.

RhaS variants with increased activation at *rhaSR*. Given the 30% amino acid sequence identity and 62% similarity between RhaS and RhaR, we attempted to identify the difference between RhaS and RhaR that dictates their difference in optimal CRP binding site position. We tested whether it would be possible to obtain increased CRP co-activation at *rhaSR* in combination with RhaS by screening for RhaS variants with increased activation at (*rhaS-lacZ*)Δ128, but without any increase in activation at (*rhaS-lacZ*)Δ85. In this way, we

expected to eliminate from consideration substitutions that simply increased RhaS binding to the RhaR binding site, for example. We screened 40 independent pools of *rhaS* genes that had been PCR mutagenized and cloned into plasmid pHG165 as previously described (17). Although we previously isolated apparent gain-of-function mutants from these pools (17), here we were unable to isolate any RhaS variants that met both of these criteria among the approximately 38,000 clones screened. However, we characterized three (of many) RhaS variants with increased activation at both fusions. These RhaS variants, RhaS H205R (isolated twice) and H253Y, activated transcription at each of the (*rhaS-lacZ*) fusions to 20- to 60-fold higher levels than wild-type RhaS (Table 5). RhaS H205R and H253Y are located at positions five and three of the recognition helices of the first and second HTH motifs of RhaS, respectively (5). In each case the substitutions replaced the RhaS residue with the wild-type RhaR residue at the aligned position. This suggests that the substitutions increase RhaS binding to the RhaR binding site.

The inability to isolate RhaS variants that enabled increased CRP co-activation at *rhaSR* suggests that the difference that allows RhaR to enable CRP co-activation at *rhaSR* may involve more than one or two simple amino acid substitutions relative to RhaS. On the other hand, we were able to easily isolate RhaS variants with increased activity at both promoter fusions. This, and the finding that these variants had substitutions that made the RhaS amino acid sequence more like that of RhaR, further support our conclusion that RhaS autoregulates *rhaSR* expression by binding to the RhaR binding site.

Summary of RhaS autoregulation model. Our overall model for RhaS autoregulation is as follows (Fig. 4): Upon encountering L-rhamnose, RhaR binds upstream of *rhaSR* and activates transcription. The resulting RhaS protein then binds to its sites at the *rhaBAD* and *rhaT* promoters and activates transcription. As the RhaS protein concentration increases, RhaS also competes with RhaR for binding at the *rhaSR* promoter. CRP is unable to efficiently co-activate

Table 5. RhaS variants with increased activation at *rhaSR*

	β-Galactosidase activity[a]	
Activator	Φ(*rhaS-lacZ*)Δ85	Φ(*rhaS-lacZ*)Δ128
RhaR	13	1013
RhaS	7	22
RhaS H205R	160	440
RhaS H253Y	438	679

[a] β-Galactosidase activity was measured from single-copy *lacZ* fusion strains grown in minimal MOPS growth medium containing ampicillin and L-rhamnose. Standard errors were less than 10% of the average units. The strain background was Δ(*rhaSR*)::*kan recA*::*cat*. Wild-type RhaS, wild-type RhaR or the RhaS variants were expressed from plasmid pHG165 (26), as previously described (17).

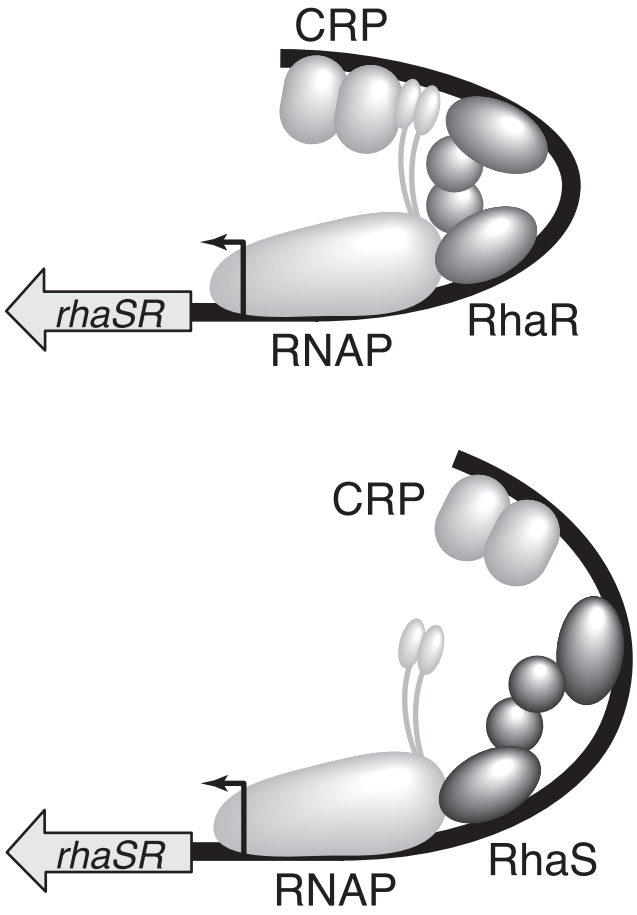

Figure 4. Model for RhaS Negative Autoregulation. Expression of the *rhaSR* operon in the presence of L-rhamnose: Top: At relatively low RhaS protein concentrations, RhaR and CRP both contribute to *rhaSR* activation, therefore expression is at its maximal level. CRP contacts α-CTD (shown as two small ovals connected by a flexible linker to RNAP). Bottom: At relatively high RhaS protein concentrations, RhaS binds to the RhaR binding site, thereby replacing RhaR. RhaS contributes to *rhaSR* activation, but CRP does not effectively co-activate, therefore *rhaSR* expression is reduced. CRP does not effectively contact α-CTD.

rhaSR transcription in combination with RhaS, likely as a result of inefficient contacts with α-CTD in this context. The end result is a decrease in *rhaSR* expression by three to four-fold. As RhaS protein levels fluctuate, this autoregulation would compensate by increasing or decreasing expression of *rhaSR* to return the RhaS protein concentration to its optimal level. We hypothesize that RhaS may not bend the DNA appropriately for CRP to co-activate *rhaSR* transcription, as described above, and illustrated in Fig 4. We have not, however, ruled out the possibility that RhaS somehow interferes with CRP binding to its DNA site at *rhaSR*, but this seems unlikely given the 21 bp spacing between the sites.

The RhaS autoregulation mechanism is unusual in that a transcriptional activator protein functions to decrease expression of an operon by activating transcription, but not allowing a second activator protein – in this case CRP – to co-activate transcription. The *rhaSR* mechanism is not unique among characterized regulatory schemes, however, with one similar system being regulation of the *E. coli napF* promoter (9, 10). At *napF*, the NarL protein competes with the 44% identical NarP protein for binding to a common DNA site. NarL binding to the site results in reduced *napF* expression relative to NarP binding, including a reduced contribution to activation by the Fnr protein (related to CRP) from its DNA site upstream of the NarP/NarL site (9, 10).

ACKNOWLEDGMENTS

We thank Gurpreet Kaur Hunjan and Bria Wilkins for helpful comments on the manuscript. This work was supported by NIH grant GM55099 from the National Institute of General Medical Sciences, NIH Grant P20 RR17708 from the Institutional Development Award (IDeA) Program of the National Center for Research Resources, and the Department of Molecular Biosciences, University of Kansas, all to S.M.E.

REFERENCES

1. **Backman, K., Y.-M. Chen, and B. Magasanik.** 1981. Physical and genetic characterization of the *gln A-glnG* region of the *Escherichia coli* chromosome. Proc. Nat. Acad. Sci., U.S.A. **78:**3743-3747.
2. **Bartolome, B., Y. M. Jubee, E. , and F. de la Cruz.** 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. Gene **102:**75-78.
3. **Belyaeva, T. A., V. A. Rhodius, C. L. Webster, and S. J. W. Busby.** 1998. Transcription activation at promoters carrying tandem DNA sites for the *Escherichia coli* cyclic AMP receptor protein: organisation of the RNA polymerase a subunits. J. Mol. Biol. **277:**789-804.
4. **Berg, O. G., and P. H. von Hippel.** 1988. Selection of DNA binding sites by regulatory proteins. Trends Biochem Sci **13:**207-11.
5. **Bhende, P. M., and S. M. Egan.** 1999. Amino acid-DNA contacts by RhaS: an AraC family transcription activator. J. Bacteriol. **181:**5185-5192.
6. **Bhende, P. M., and S. M. Egan.** 2000. Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with sigma 70. J. Bacteriol. **182:**4959-4969.
7. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res **7:**1513-23.
8. **Browning, D. F., and S. J. Busby.** 2004. The regulation of bacterial transcription initiation. Nat Rev Microbiol **2:**57-65.
9. **Darwin, A. J., and V. Stewart.** 1995. Nitrate and nitrite regulation of the Fnr-dependent *aeg-46.5* promoter of *Escherichia coli* K-12 is mediated by competition between homologous response regulators (NarL and NarP) for a common DNA-binding site. J. Mol. Biol. **251:**15-29.
10. **Darwin, A. J., E. C. Ziegelhoffer, P. J. Kiley, and V. Stewart.** 1998. Fnr, NarP, and NarL regulation of Escherichia coli K-12 napF (periplasmic nitrate reductase) operon transcription in vitro. J Bacteriol **180:**4192-8.
11. **Egan, S. M., and R. F. Schleif.** 1993. A regulatory cascade in the induction of *rhaBAD*. J. Mol. Biol. **234:**87-98.
12. **Egan, S. M., and R. F. Schleif.** 1994. DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. J. Mol. Biol. **243:**821-829.
13. **Gottesman, M. E., and M. B. Yarmolinsky.** 1968. The integration and excision of the bacteriophage lambda genome. Cold Spring Harbor Symp. Quant. Biol. **33:**735-747.
14. **Haran, T. E., and U. Mohanty.** 2009. The unique structure of A-tracts and intrinsic DNA bending. Q Rev Biophys **42:**41-81.
15. **Holcroft, C. C., and S. M. Egan.** 2000. Interdependence of activation at *rhaSR* by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain and RhaR. J. Bacteriol. **182:**6774-6782.
16. **Holcroft, C. C., and S. M. Egan.** 2000. Roles of cyclic AMP receptor protein and the carboxyl-terminal domain of the a subunit in transcription activation of the *Escherichia coli rhaBAD* operon. J. Bacteriol. **182:**3529-3535.
17. **Kolin, A., V. Balasubramaniam, J. M. Skredenske, J. R. Wickstrum, and S. M. Egan.** 2008. Differences in the mechanism of the allosteric L-rhamnose responses of the AraC/XylS family transcription activators RhaS and RhaR. Mol Microbiol **68:**448-61.

18. **Kolin, A., V. Jevtic, L. Swint-Kruse, and S. M. Egan.** 2007. Linker regions of the RhaS and RhaR proteins. J Bacteriol **189:**269-71.
19. **Miller, J. H.** 1972. Experiments in Molecular Genetics, vol. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
20. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736-747.
21. **Powell, B. S., M. P. Rivas, D. L. Court, Y. Nakamura, and C. L. Turnbough, Jr.** 1994. Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic Acids Res **22:**5765-6.
22. **Schleif, R.** 2003. AraC protein: a love-hate relationship. Bioessays **25:**274-82.
23. **Schultz, S. C., G. C. Shields, and T. A. Steitz.** 1991. Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. Science **253:**1001-7.
24. **Shine, J., and L. Dalgarno.** 1974. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A **71:**1342-6.
25. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85-96.
26. **Stewart, G. S. A. B., S. Lubinsky-Mink, C. G. Jackson, A. Cassel, and J. Kuhn.** 1986. pHG165: a pBR322 copy number derivative of pUC8 for cloning and expression. Plasmid **15:**172-181.
27. **Tate, C. G., J. A. R. Muiry, and P. J. F. Henderson.** 1992. Mapping, cloning, expression, and sequencing of the *rhaT* gene which encodes a novel L-Rhamnose-H^+ transport protein in *Salmonella typhimurium* and *Escherichia coli.* J. Biol. Chem. **287:**6923-6932.
28. **Tobin, J. F., and R. F. Schleif.** 1987. Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. J. Mol. Biol. **196:**789-799.
29. **Tobin, J. F., and R. F. Schleif.** 1990. Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. J. Mol. Biol. **211:**75-89.
30. **Tobin, J. F., and R. F. Schleif.** 1990. Transcription from the *rha* operon p_{sr} promoter. J. Mol. Biol. **211:**1-4.
31. **Via, P., J. Badia, L. Baldoma, N. Obradors, and J. Aguilar.** 1996. Transcriptional regulation of the *Escherichia coli rhaT* gene. Microbiology **142:**1833-1840.
32. **Wickstrum, J. R., and S. M. Egan.** 2004. Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. J Bacteriol **186:**6277-85.
33. **Wickstrum, J. R., T. J. Santangelo, and S. M. Egan.** 2005. Cyclic AMP receptor protein and RhaR synergistically activate transcription from the L-rhamnose-responsive *rhaSR* promoter in *Escherichia coli*. J Bacteriol **187:**6708-18.
34. **Wickstrum, J. R., J. M. Skredenske, A. Kolin, D. J. Jin, J. Fang, and S. M. Egan.** 2007. Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of Its Effector-Binding Domain. J Bacteriol **189:**4984-93.
35. **Zhang, X., and R. Schleif.** 1998. Catabolite gene activator protein mutations affecting activity of the *araBAD* promoter. J. Bacteriol. **180:**195-200.
36. **Zhou, Y. H., X. P. Zhang, and R. H. Ebright.** 1991. Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. Nucleic Acids Res **19:**6052.