JOURNAL OF BACTERIOLOGY, Jan. 2007, p. 269–271 Vol. 189, No. 1
0021-9193/07/$08.00+0 doi:10.1128/JB.01456-06

Linker Regions of the RhaS and RhaR Proteins▽

Ana Kolin,[1] Visnja Jevtic,[1] Liskin Swint-Kruse,[2] and Susan M. Egan[1]*

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045,[1] and Department of Biochemistry and Molecular Biology, MS 3030, The University of Kansas Medical Center, Kansas City, Kansas 66160[2]

Received 14 September 2006/Accepted 13 October 2006

Substitutions within the interdomain linkers of the AraC/XylS family proteins RhaS and RhaR were tested to determine whether side chain identity or linker structure was required for function. Neither was found crucial, suggesting that the linkers do not play a direct role in activation, but rather simply connect the two domains.

In the presence of the sugar L-rhamnose, RhaS and RhaR activate transcription of *Escherichia coli* genes, whose products are required for the uptake and catabolism of L-rhamnose (6, 7, 29, 30). RhaS and RhaR comprise a regulatory cascade in which L-rhamnose stimulates RhaR to activate transcription of *rhaSR* (29) and then RhaS activates transcription of the L-rhamnose catabolic operon *rhaBAD* (7) and the L-rhamnose transport gene *rhaT* (30). Both RhaS and RhaR are members of the large AraC/XylS family of transcription regulatory proteins, which share sequence similarity in a 100-amino-acid DNA-binding domain (5, 10, 11, 21). AraC/XylS family proteins regulate the expression of genes whose functions include carbon metabolism (3, 12, 25, 28), stress responses (4, 15–18, 22, 32), and pathogenesis (9, 14, 23, 24).

RhaS and RhaR each consist of two independently functional domains: an N-terminal domain required for dimerization and response to L-rhamnose (A. Kolin, J. R. Wickstrum, and S. M. Egan, unpublished results) and a C-terminal domain involved in DNA binding and transcription activation (1, 2, 6, 31; J. R. Wickstrum and S. M. Egan, unpublished results). Based on alignment with the AraC dimerization domain and its structure (26) and the MarA protein and its structure (22), we can predict the approximate boundaries of the RhaS and RhaR domains. Such analysis, combined with knowledge of AraC (8), leads to the prediction that there is a flexible linker that connects the two domains of RhaS and RhaR. This linker spans approximately residues 166 to 172 in RhaS and 198 to 207 in RhaR (Fig. 1). Genetic and biochemical studies have shown that many different single and multiple substitutions could be made within the AraC linker without impacting activation (8), suggesting that the AraC arabinose response does not depend on the identity of its linker residues. However, our finding that the ligand responses of RhaS and RhaR differ from that of AraC (A. Kolin and S. M. Egan, unpublished results) left open the possibility that the linker might participate in transmission of the L-rhamnose signal.

In this study, we used a genetic approach to analyze the linker regions of RhaS and RhaR. Our previous results indicate that L-rhamnose binds to the N-terminal domains of RhaS and RhaR, while transcription activation involves their C-terminal domains (2, 31; Kolin and Egan, unpublished). Therefore, we sought to determine whether the linker regions of RhaS and/or RhaR were involved in transmitting the L-rhamnose status from the N- to the C-terminal domain or whether, similar to AraC, the linker was required only to flexibly connect the two domains.

Single alanine substitutions in the linker region of RhaS and RhaR have, at most, small effects. The ability of RhaS and RhaR to differentially activate transcription in the presence but not the absence of L-rhamnose could be due to either stimulation of protein activity in the presence of L-rhamnose or inhibition of activity in the absence of ligand (the latter is the case for AraC). If the linker were involved in stimulation of activity in the presence of ligand or were otherwise required for activation, we would expect important mutations to exhibit a significant reduction of the activation by RhaS or RhaR in the presence of L-rhamnose. If, on the other hand, the linker were involved in the inhibition of RhaS or RhaR activity in the absence of ligand, we would expect that mutations that reduced this inhibition would increase activation in the absence of L-rhamnose.

To test the roles of individual linker residues, we constructed derivatives of RhaS and RhaR with single alanine substitutions in their linker regions by PCR (Expand High Fidelity PCR system; Roche, Indianapolis, IN). Both DNA strands of the entire cloned *rhaS* or *rhaR* regions were sequenced in all cases (Molecular Core Research Facility, Idaho State University, or Northwestern University Biotechnology Laboratory). Plasmid pHG165 (27) carrying the *rhaS* or *rhaR* derivative was transformed into a strain carrying a single-copy *rhaB-lacZ* fusion and a *rhaS* deletion (SME1088) (2) or a single-copy *rhaS-lacZ* fusion and a *rhaSR* deletion (SME1076) (13), respectively. Transcription activation was determined by β-galactosidase assays of cultures grown in morpholinepropanesulfonic acid (MOPS)-buffered minimal medium using the method described by Neidhardt et al. (1, 20) with glycerol as the carbon source and in the absence or presence of 0.2% L-rhamnose (1). Assays were performed by using the method described by Miller (1, 19).

We found that all of the RhaS derivatives were able to activate transcription to within approximately twofold of the activation by wild-type RhaS (Table 1), suggesting that none of

* Corresponding author. Mailing address: Department of Molecular Biosciences, 1200 Sunnyside Ave., University of Kansas, Lawrence, KS 66045. Phone: (785) 864-4294. Fax: (785) 864-5294. E-mail: sme@ku.edu.

▽ Published ahead of print on 27 October 2006.

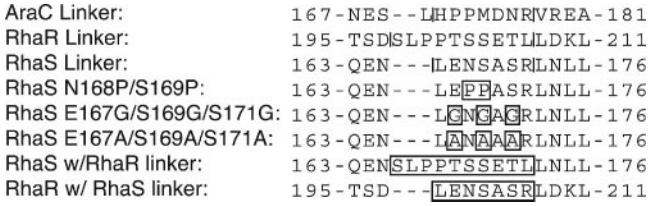

FIG. 1. Alignment of the RhaS, RhaR, and AraC linker regions. The approximate minimal linkers are enclosed by vertical lines in the wild-type sequences. The amino acid substitutions in the linker regions of the RhaS and RhaR derivatives are enclosed in boxes.

the individual linker residues was vital for activation by RhaS. In addition, none of the RhaS derivatives exhibited more than about a twofold increase in activation in the absence of L-rhamnose. RhaS R172A had a phenotype we did not predict, with a greater-than-fivefold defect in the absence of L-rhamnose but only about a twofold defect in the presence of L-rhamnose. Similar to RhaS, all of the RhaR derivatives activated transcription to within twofold of the activation by wild-type RhaR, and none of the derivatives resulted in a >1.5-fold increase in RhaR activation in the absence of L-rhamnose (Table 2). Interestingly, and again similar to results for RhaS, two of the RhaR derivatives (S204A and T206A) had fivefold or greater defects in the absence of L-rhamnose but no more than twofold defects in the presence of L-rhamnose, resulting in much higher activation values than for the wild type. These results suggest that none of the individual residues in the RhaS and RhaR linkers is required for the response to L-rhamnose or for the ability to activate transcription.

Structural requirements of the RhaS linker. Although the alanine substitution results argue that the identities of the individual RhaS linker residues are not important for its L-rhamnose response, they do not rule out the possibility that a linker region secondary structure, possibly an α-helix, is required for RhaS function. To test this hypothesis, we constructed RhaS derivatives with multiple linker substitutions that were predicted to either disrupt (double proline and triple glycine substitutions) or not disrupt (triple alanine substitutions) the formation of an α-helix (Fig. 1). We predict that disruption of the α-helix would reduce activity in the presence of L-rhamnose if the linker were required to stimulate activity with L-rhamnose or increase activity in the absence of L-rhamnose if the linker inhibited activity without L-rhamnose. In spite of the finding that the double-proline derivative activated to a lower level than the wild type, the finding that the triple-glycine derivative was not defective suggests that the RhaS linker is probably not required to form an α-helix (Table 3). The defect of the double-proline derivative suggests the possibility that the linker requires flexibility. The triple alanine substitution resulted in only a small decrease in RhaS activation.

Linker swapping between the RhaS and RhaR proteins. The RhaR linker is three residues longer than the RhaS linker, which is noteworthy given that the only other internal insertion/deletion in the alignment of RhaS and RhaR is a single-residue difference. To determine whether this linker length difference has functional consequences, we swapped the RhaS and RhaR linker regions (Table 3). Transcription activation by the RhaS derivative carrying the RhaR linker was ninefold lower than the wild type in the presence of L-rhamnose. However, this derivative maintained a 15-fold response to L-rhamnose, indicating that its L-rhamnose response was at least partially intact. The RhaR derivative carrying the RhaS linker region had small defects in both the presence and absence of L-rhamnose; however, its activation in response to L-rhamnose was somewhat greater than that of wild-type RhaR. The defect of the RhaS derivative carrying the RhaR linker may again

TABLE 1. Single alanine substitutions within the linker region of RhaS

RhaS derivative	β-Galactosidase activity[a]	
	−L-Rhamnose	+L-Rhamnose
Wild type	0.28	456
L166A	0.39	309
E167A	0.6	495
N168A	0.2	240
S169A	0.18	404
S171A	0.16	191
R172A	0.05	191

[a] β-Galactosidase activities are expressed in Miller units (19), with standard errors of less than 25%.

TABLE 2. Alanine substitutions of the linker region of RhaR

RhaR derivative	β-Galactosidase activity[a]	
	−L-Rhamnose	+L-Rhamnose
Wild type	117	592
S198A	90	562
L199A	123	535
P200A	92	575
P201A	125	680
T202A	178	576
S203A	101	506
S204A	19	359
E205A	57	583
T206A	3.4	304
L207A	143	392

[a] β-Galactosidase activities are expressed in Miller units (19), with standard errors of less than 34%.

TABLE 3. Multiple mutations in the RhaS and RhaR linker regions

RhaS or RhaR derivative	β-Galactosidase activity[a]	
	−L-Rhamnose	+L-Rhamnose
RhaS derivatives		
Wild type	0.32	179
N168P/S169P	0.47	55
E167G/S169G/S171G	0.24	206
E167A/S169A/S171A	0.08	120
RhaS with RhaR linker	0.31	25
RhaR derivatives		
Wild type	3.3	16
RhaR with RhaS linker	1	8.4

[a] β-Galactosidase activities are expressed in Miller units (19), with standard errors of less than 18% for RhaS derivatives and less than 14% for RhaR derivatives. The LacZ fusion did not contain the CRP binding sites, resulting in lower values than those in Tables 1 and 2.

reflect a requirement by RhaS for linker flexibility, given that the RhaR linker contains two adjacent proline residues.

Summary. Our results indicate that the individual side chains of the RhaS and RhaR linker residues are unlikely to play any crucial roles in activation by RhaS and RhaR or in the responses of these proteins to L-rhamnose. The two RhaS derivatives with the greatest defects both contained two consecutive prolines, suggesting that the prolines, and possibly the rigidity they impart, were responsible for the loss of RhaS activity, although we cannot rule out other possibilities. Our results indicate that the linker regions of RhaS and RhaR function only to flexibly connect the two domains of the proteins.

This work was supported by NIH grant GM55099 from the National Institute of General Medical Sciences (to S.M.E.) and NIH grant RR-P20 RR17708 from the Institutional Development Award (IdeA) Program of the National Center for Research Resources (to S.M.E. and L.S.-K.).

REFERENCES

1. **Bhende, P. M., and S. M. Egan.** 1999. Amino acid-DNA contacts by RhaS: an AraC family transcription activator. J. Bacteriol. **181:**5185–5192.
2. **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.
3. **Bourgerie, S. J., C. M. Michan, M. S. Thomas, S. J. W. Busby, and E. I. Hyde.** 1997. DNA binding and DNA bending by the MelR transcription activator protein from *Escherichia coli*. Nucleic Acids Res. **25:**1685–1693.
4. **Cohen, S. P., H. Hachler, and S. B. Levy.** 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. J. Bacteriol. **175:**1484–1492.
5. **Egan, S. M.** 2002. Growing repertoire of AraC/XylS activators. J. Bacteriol. **184:**5529–5532.
6. **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.
7. **Egan, S. M., and R. F. Schleif.** 1993. A regulatory cascade in the induction of *rhaBAD*. J. Mol. Biol. **234:**87–98.
8. **Eustance, R. J., and R. F. Schleif.** 1996. The linker region of AraC protein. J. Bacteriol. **178:**7025–7030.
9. **Falconi, M., B. Colonna, G. Prosseda, G. Micheli, and C. O. Gualerzi.** 1998. Thermoregulation of *Shigella* and *Eshcherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. EMBO J. **17:**7033–7043.
10. **Gallegos, M.-T., C. Michan, and J. L. Ramos.** 1993. The XylS/AraC family of regulators. Nucleic Acids Res. **21:**807–810.
11. **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.
12. **Gonzalez-Perez, M. M., J. L. Ramos, M. T. Gallegos, and S. Marques.** 1999. Critical nucleotides in the upstream region of the XylS-dependent TOL meta-cleavage pathway operon promoter as deduced from analysis of mutants. J. Biol. Chem. **274:**2286–2290.
13. **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.
14. **Jordi, B. J. A. M., B. A. M. van der Zeijst, and W. Gaastra.** 1994. Regions of the CFA/I promoter involved in the activation by the transcriptional activator CfaD and repression by the histone-like protein H-NS. Biochimie **76:**1052–1054.
15. **Landini, P., and S. J. Busby.** 1999. The *Escherichia coli* Ada protein can interact with two distinct determinants in the σ^{70} subunit of RNA polymerase according to promoter architecture: identification of the target of Ada activation at the *alkA* promoter. J. Bacteriol. **181:**1524–1529.
16. **Lemotte, P. K., and G. C. Walker.** 1985. Induction and autoregulation of *ada*, a positively acting element regulating the response of *Escherichia coli* K-12 to methylating agents. J. Bacteriol. **161:**888–895.
17. **Li, Z., and B. Demple.** 1996. Sequence specificity for DNA binding by *Escherichia coli* SoxS and Rob proteins. Mol. Microbiol. **20:**937–945.
18. **Li, Z., and B. Demple.** 1994. SoxS, an activator of superoxide stress genes in *Escherichia coli*: purification and interaction with DNA. J. Biol. Chem. **269:**18371–18377.
19. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
20. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736–747.
21. **Ramos, J. L., F. Rojo, L. Zhou, and K. N. Timmis.** 1990. A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. Nucleic Acids Res. **18:**2149–2152.
22. **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.
23. **Sakai, T., C. Sasakawa, and M. Yoshikawa.** 1988. Expression of four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kilodalton *virF* protein. Mol. Microbiol. **2:**589–597.
24. **Savelkoul, P. H. M., G. A. Willshaw, M. M. McConnell, H. R. Smith, A. M. Hamers, B. A. M. van der Zeijst, and W. Gaastra.** 1990. Expression of CFA/I fimbriae is positively regulated. Microb. Pathog. **8:**91–99.
25. **Schleif, R.** 1996. Two positively regulated systems, *ara* and *mal*, p. 1300–1309. *In* F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, DC.
26. **Soisson, S. M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** 1997. Structural basis for ligand-regulated oligomerization of AraC. Science **276:**421–425.
27. **Stewart, G. S., S. Lubinsky-Mink, C. G. Jackson, A. Cassel, and J. Kuhn.** 1986. pHG165: a pBR322 copy number derivative of pUC18 for cloning and expression. Plasmid **15:**172–181.
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. Transcription from the *rha* operon p_{sr} promoter. J. Mol. Biol. **211:**1–4.
30. **Via, P., J. Badia, L. Baldoma, N. Obradors, and J. Aguilar.** 1996. Transcriptional regulation of the *Escherichia coli rhaT* gene. Microbiology **142:**1833–1840.
31. **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:**6277–6285.
32. **Wu, J., and B. Weiss.** 1991. Two divergently transcribed genes, *soxR* and *soxS*, control a superoxide response regulon of *Escherichia coli*. J. Bacteriol. **173:**2864–2871.