Residues near the Amino Terminus of Rns Are Essential for Positive Autoregulation and DNA Binding

Georgeta N. Basturea, Maria D. Bodero, Mario E. Moreno, and George P. Munson*

Department of Microbiology and Immunology, University of Miami Miller School of Medicine, P.O. Box 016960 (R-138), Miami, Florida 33101

Received 23 October 2007/Accepted 13 January 2008

Most members of the AraC/XylS family contain a conserved carboxy-terminal DNA binding domain and a less conserved amino-terminal domain involved in binding small-molecule effectors and dimerization. However, there is no evidence that Rns, a regulator of enterotoxigenic Escherichia coli virulence genes, responds to an effector ligand, and in this study we found that the amino-terminal domain of Rns does not form homodimers in vivo. Exposure of Rns to the chemical cross-linker glutaraldehyde revealed that the full-length protein is also a monomer in vitro. Nevertheless, deletion analysis of Rns demonstrated that the first 60 amino acids of the protein are essential for the activation and repression of Rns-regulated promoters in vivo. Amino-terminal truncation of Rns abolished DNA binding in vitro, and two randomly generated mutations, H14T and N16D, that independently abolished Rns autoregulation were isolated. Further analysis of these mutations revealed that they have disparate effects at other Rns-regulated promoters and suggest that they may be involved in an interaction with the carboxy-terminal domain of Rns. Thus, evolution may have preserved the amino terminus of Rns because it is essential for the regulator's activity even though it apparently lacks the two functions, dimerization and ligand binding, usually associated with the amino-terminal domains of AraC/XylS family members.

The expression of several pilus serotypes in enterotoxigenic Escherichia coli (ETEC), an enteric pathogen that causes diarrheal disease in humans and livestock, is dependent upon the transcriptional regulator Rns (GenBank accession no. P16114). These pilus serotypes include the CS1, CS2, CS3, and CS4 pili (5, 6, 11). Rns has also been shown to repress the expression of an inner membrane lipoprotein involved in the biogenesis of outer membrane vesicles by preventing the formation of an RNA polymerase open complex at nlpAp, the lipoprotein’s promoter (2, 24, 42). Rns positively autoregulates its own expression (12) and is interchangeable with several other virulence regulators, including Cfd (accession no. P25393) and CsrV (accession no. CAA47200), which are carried by some ETEC strains (6). VirF (accession no. NP_085206) from Shigella flexneri, and AggR (accession no. P43464) from enterohaemorrhagic E. coli (EAEC) (26). In the case of VirF and Rns this is somewhat surprising because their regulons share no homologous genes. Rns may be considered the archetype for this group of conserved virulence regulators because it is the only member with a well-developed system for in vitro studies. This has facilitated characterization of its interactions with DNA and its effects upon RNA polymerase at various promoters and identification of additional genes within the Rns regulon (2, 26–28, 30).

Linker scanning mutagenesis of Rns has revealed a region within the protein that accepts insertion of 19 amino acids (M. D. Bodero and G. P. Munson, unpublished data). This region is comprised of residues 100 through 131 and probably functions as a flexible linker between two domains that are roughly the same size (see Fig. 1A). The carboxy-terminal domain (CTD) of Rns contains two putative helix-turn-helix (HTH) motifs connected by an α-helix, the signature feature of proteins belonging to the AraC/XylS superfamily of transcriptional regulators (13). Uracil interference studies suggested that Rns places a recognition helix from each HTH motif in the major groove of DNA (27, 28) in a manner similar to that seen in a DNA cocrystal of another AraC/XylS family member, MarA (PDB 1BL0) (34). Mutagenesis of either HTH motif has been shown to reduce or abolish the activity of VirF (32) and Rns in vivo (Bodero and Munson, unpublished data), presumably by disrupting DNA binding. Thus, it is likely that the CTD of Rns contains most or all of the residues that make direct contact with DNA.

Unlike the CTD, the function of the amino-terminal domain (NTD) of Rns has been obscure. The NTDs of many AraC/XylS family members are known or thought to bind effector ligands. They may also contain a dimerization interface for the formation of homodimers. However, there is no evidence that the activity of Rns is modulated by exogenous ligands, and in this paper we report that the NTD of Rns lacks the ability to form homodimers. Nevertheless, the amino terminus of Rns is essential for DNA binding, and we identified two residues in the Rns NTD that may interact with the CTD of Rns in a manner that facilitates DNA binding.

MATERIALS AND METHODS

Rns and cl expression plasmids. Plasmid pGPMRns is a derivative of pNEB193 (New England Biolabs) that expresses Rns from lac (2). Amino-terminal truncations of Rns were constructed and epitope tagged by performing inverse PCR with pGPMRns using primer NEB-hisTag (AATGCATGCCGTG GTGGTGTTGTTGCTAGTGCATATGCTGTGGTCTTCT) and primer 057-F7 (CCAAGCATGCGAAGCGAATTTTATCAG), 057-F8 (GAGGCCA...
TGCTAATACGCTATGTAAGAAGACATGACCAGAA, or 5'-FGGAGCAT GTAACCTCGGTTGAGAATTATTTAAAATATA (underlining in the primer sequences indicates primer-template mismatches). The PCR products were then digested with SpH1 and circularized with T4 DNA ligase to produce plasmids pGPM1025, pGPM1036, and pGPM1037, which express His6-Rns, His6-Rns-100, and His6-Rns-110, respectively. Transposon mutagenesis of pGPMRs produced pGPMRe-Gatc-2, which carries rnc-lan. Plasmid pEU2030 (12) expresses Rns from a promoter within cloning vector pUC18 (accession no. L08752). Random mutagenesis of pEU2030 (see below) produced plasmid pRs100, which carries the rns-100 allele (codon 14, ATT [Ile] changed to ACT [Thr]), and plasmid pRs101, which carries rns-101 (codon 16, AAT [Asn] changed to GAT [Asp]).

Plasmids pMBRns1, pMBRns64, and pMBRns58 are derivatives of the protein expression vector pMal-c2 (New England Biolabs) that expresses maltose binding protein (MBP)-Rns, MBP-Rns (80-265), and MBP-Rns (128-265), respectively, from the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter tacp. Construction of pMBRns1 has been previously reported (2). Plasmids pMBRns564 and pMBRns58 were constructed by performing inverse PCR with plasmid pEU7500 (29) using primers 057R1 (AAAGTGGATCCGTCCTTCCTCC TCGATGCGAAGGTGTGTTT) and 057F1 (GGAGGATCCGAGAGCTATCTTTATATATAC) and RsflFlagBamv3 (ACGGATCTACTTCTAGGGCTCCTTGGTGATCAGTCAAGAA), or 057-F9 (GGAGGCAT TGCCTCAAGCGGCATTCG) for 20 to 30 min at 37°C. The PCR products were digested with SpH1 and then circularized with T4 DNA ligase. The plasmids used for expression of cl included pFG157, which expresses full-length cl; pKH101, which expresses the DNA binding domain (DBD) of cl; and pHJJ70 (30) containing the leucine zipper of GCN4 (22, 23). Plasmid pBGP17 expresses clRns100-154 (31) and was constructed by digesting a fragment of plasmid pBGP17Rns (32) with Rsfl (33) and EcoRI and then ligating it into pGPM1 (34) digested with Rsfl (35) and EcoRI. The E. coli strain EU2103, MC4100/ relA1 flhD5301 deoC1 ptsF25 rbsR leu was transformed with pGPM1025, pGPM1036, and pGPM1037, which express His6-Rns(61-265), His6-Rns(100-265), and His6-Rns(110-265), respectively, from the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter tacp. Expression of cl was checked by SDS-PAGE analysis.

**RESULTS**

**Rns is not a homodimer.** Like several other AraC family members that have been characterized, Rns contains two do-
mains that are joined by a flexible linker (Fig. 1A). If Rns is
like AraC and other dimeric family members, its NTD should
contain all of the residues necessary for the formation of
homodimers. However, we have also identified a potential α-helix
in the CTD of Rns (Fig. 1A) that has homology, albeit limited,
to the dimerization helices in the NTDs of AraC, XylS, UreR,
and ToxT (8, 21, 31, 35). Therefore, in order to determine if
Rns forms homodimers in vivo, we replaced the carboxy-
terminal dimerization domain of the dimerization helix was
included as a monomeric negative control. Each protein construct
was also expressed in AG1688 and assayed to
determine its ability to confer immunity to the lytic phage
λKH54 (Δcl). I, immune to infection and lysis by λKH54; S, sensitive to
infection and lysis by λKH54. (C) Fusion protein cIDBD-Rns(1-154)FLAG has an expected molecular mass of 34 kDa, and its expression
from plasmid pGB17 was confirmed by Western blotting using an
anti-FLAG polyclonal antibody against the FLAG epitope tag. Lane Std. contained
protein standards.

FIG. 1. Rns(1-154) does not contain a homodimerization domain.
(A) Diagram of Rns and its domain organization. The linker region
between the NTD and CTD of Rns was identified by linker scanning
mutagenesis (Bodero and Munson, unpublished data). Putative HTH
motifs were identified by secondary and tertiary modeling. A putative
dimerization helix was identified by homology to the dimerization
helices of AraC, XylS, UreR, and ToxT. (B) β-Galactosidase activity (n > 3)
from the cl-repressed promoter λP_{O_R}^{2-2} in K-12 hyogen AG1688/
λ202. cl and cIDBD-GCN4 are homodimers and were included as
positive controls. cIDBD was included as a monomeric negative control.
In addition to immunity assays, we tested the ability of cIDBD-
Rns(1-154)-FLAG to regulate the expression of β-galactosidase
from the cl-repressed λP_{O_R}^{2-2} promoter in reporter
strain AG1688/λ202. λ202 is a reporter phage containing lacZ
driven by the λP_{O_R}^{2-2} promoter. This promoter carries a
mutation in one of the three cl operator sites that allows
individual cl dimers to repress the promoter in the absence of
cooperative binding (23). As expected, both cl and cIDBD-
GCN4 repressed expression of β-galactosidase (Fig. 1B). In
contrast, both cIDBD-Rns(1-154)-FLAG and cIDBD failed to
repress expression of β-galactosidase (Fig. 1B). Because cl
dimerize must function as a repressor, the inability of Rns(1-
154) to restore repressor function to cIDBD indicates that the
fusion protein is not a dimer. These results are consistent with
those obtained in the immunity assays and demonstrate that
unlike other AraC/XylS family members that have been shown
to form homodimers, the first 154 amino acids of Rns are not
sufficient for dimerization.

We also tested the possibility that full-length Rns maybe a
dimer or other multimer by exposing MBP-Rns to the chemical
cross-linker glutaraldehyde. MBP is a monomeric protein that
has been previously shown to increase the solubility of Rns
without interfering with its activity in vivo or in vitro (27, 28).
This solubility tag is necessary for in vitro studies because Rns,
for all practical purposes, is insoluble in vitro, as are most
AraC/XylS family members. It has also been previously shown
that MBP does not interfere with glutaraldehyde cross-linking
of XylS dimers in vitro (35). Although we did observe that the
mobility of MBP-Rns was decreased as a result of various
amounts of glutaraldehyde bound to monomers of MBP-Rns,
dimers of the 74-kDa protein were not detected (Fig. 2A). At
the highest concentration of glutaraldehyde used, only large
nonspecific complexes trapped near the top of the gel were
observed. In contrast, dimers and trimers of SpaT (accession
no. NP_902290), a multimeric chaperone from the purple-
pigmented water and soil bacterium Chromobacterium viola-
were readily detected under the same cross-linking conditions (Fig. 2A). From these initial studies with SpaT we determined that 0.005% (vol/vol) glutaraldehyde was the optimal concentration for cross-linking protein multimers. We then treated a range of protein concentrations with this optimal concentration of glutaraldehyde to determine if higher concentrations of MBP-Rns facilitated the formation of homodimers (Fig. 2B). However, we found no evidence of MBP-Rns dimers, even when the MBP-Rns concentration was raised to 12 μM. In contrast, SpaT dimers were detected with 6 μM (Fig. 2B). These results indicate that full-length Rns protein is a monomer under these conditions and are consistent with the results for cIDBD heterologous fusions, which demonstrated that the NTD of Rns does not dimerize in vivo.

Amino terminus of Rns is required for DNA binding. Since the NTD of Rns does not contain a homodimerization interface, we designed a series of plasmids with progressive deletions of the Rns NTD to determine if the NTD is necessary for the regulator’s activity. However, only three plasmids expressed stable proteins, as determined by Western blotting using a polyclonal antibody to the hexahistidyl epitope tag (Fig. 3A). These plasmids, along with a plasmid expressing Rns or a plasmid carrying rns::kan, were then transformed into lysogens harboring msp-lacZ and cfaAp-lacZ reporter prophage to determine their ability to activate Rns-dependent promoters. In contrast to Rns, none of the deletion constructs were able to activate either promoter because the β-galactosidase expression was essentially the same as that of the rns::kan negative controls (Fig. 3B). We also found that none of the truncated proteins were able to efficiently repress nlpAp. With His6-Rns(100-265) we did observe that the expression of β-galactosidase from nlpAp was slightly less than that of the negative control strain carrying rns::kan; however, the difference between the two strains is within the uncertainty of the measurements and is therefore probably not significant (Fig. 3B). These results suggest that the truncated proteins cannot activate msp and cfaAp because they cannot bind DNA, and they eliminate the possibility that the truncated proteins bind DNA but fail to activate msp and cfaAp because they lack an activation domain.

We also determined that the amino terminus of Rns is required for DNA binding in vitro by DNase I footprinting. At the CFA/I pilus promoter 250 nM MBP-Rns is sufficient to saturate both of the previously reported (30) Rns binding sites, cfaAo1, and cfaAo2, as well as an additional binding site, cfaAo3, which were readily detected under the same cross-linking conditions (Fig. 2A). From these initial studies with SpaT we determined that 0.005% (vol/vol) glutaraldehyde was the optimal concentration for cross-linking protein multimers. We then treated a range of protein concentrations with this optimal concentration of glutaraldehyde to determine if higher concentrations of MBP-Rns facilitated the formation of homodimers (Fig. 2B). However, we found no evidence of MBP-Rns dimers, even when the MBP-Rns concentration was raised to 12 μM. In contrast, SpaT dimers were detected with 6 μM (Fig. 2B). These results indicate that full-length Rns protein is a monomer under these conditions and are consistent with the results for cIDBD heterologous fusions, which demonstrated that the NTD of Rns does not dimerize in vivo.

Amino terminus of Rns is required for DNA binding. Since the NTD of Rns does not contain a homodimerization interface, we designed a series of plasmids with progressive deletions of the Rns NTD to determine if the NTD is necessary for the regulator’s activity. However, only three plasmids expressed stable proteins, as determined by Western blotting using a polyclonal antibody to a His6 epitope tag (Fig. 3A). These plasmids, along with a plasmid expressing Rns or a plasmid carrying rns::kan, were then transformed into lysogens harboring msp-lacZ and cfaAp-lacZ reporter prophage to determine their ability to activate Rns-dependent promoters. In contrast to Rns, none of the deletion constructs were able to activate either promoter because the β-galactosidase expression was essentially the same as that of the rns::kan negative controls (Fig. 3B). We also found that none of the truncated proteins were able to efficiently repress nlpAp. With His6-Rns(100-265) we did observe that the expression of β-galactosidase from nlpAp was slightly less than that of the negative control strain carrying rns::kan; however, the difference between the two strains is within the uncertainty of the measurements and is therefore probably not significant (Fig. 3B). These results suggest that the truncated proteins cannot activate msp and cfaAp because they cannot bind DNA, and they eliminate the possibility that the truncated proteins bind DNA but fail to activate msp and cfaAp because they lack an activation domain.

We also determined that the amino terminus of Rns is required for DNA binding in vitro by DNase I footprinting. At the CFA/I pilus promoter 250 nM MBP-Rns is sufficient to saturate both of the previously reported (30) Rns binding sites, cfaAo1, and cfaAo2, as well as an additional binding site, cfaAo3,
further upstream (Fig. 4). In contrast, none of the binding sites were occupied by MBP-Rns(80-265) or MBP-Rns(128-265) even at concentrations as high as 2.5 M (Fig. 4). As a result of different cloning strategies, the two latter MBP fusions have truncations that differ from the His6 epitope-tagged truncations reported above. Nevertheless, our conclusions from both sets of experiments are the same. The loss of its amino terminus substantially reduces the DNA binding affinity of Rns. Based upon DNase I footprinting, we estimated that the reduction is greater than 1 order of magnitude.

Isolation of mutations that abolish Rns positive autoregulation. Since deletion analysis of Rns revealed that the first 60 amino acids of the regulator are essential for DNA binding, we next sought to identify specific residues required for Rns activity by selecting for randomly generated mutations that abolish Rns positive autoregulation. To avoid the isolation of null mutations, our selection strategy also required that these mutations did not abolish Rns-dependent expression from the CFA/I pilus promoter. We reasoned that these types of mutations could be isolated because the DNA sequences of Rns binding sites at cfaAp are not identical to those at rmsp (30). In addition, Rns binds to a site adjacent to the −35 hexamer at the pilus promoter but does not bind near the −35 hexamer of rmsp (28, 30). This suggested that Rns autoregulation may involve at least some unique residues that are not required for activation of the pilus promoter.

The Rns expression plasmid pEU2030 was randomly mutagenized by propagation in the mutator strain XL1-Red (Stratagene) and then transformed into a reporter strain that harbored an rmsp-lacZ reporter prophage and a plasmid-borne cfaAp-tetA reporter. The phenotype of this strain is Lac in the absence of Rns but LacTet when it is transformed with pEU2030. Transformants were plated on Lac indicator media containing tetracycline, and several LacTet colonies were isolated for further analysis. False positives were eliminated by transforming naive rmsp-lacZ and cfaAp-lacZ reporter strains with plasmids recovered from LacTet isolates. Plasmids that resulted in Rns-dependent expression of β-galactosidase from the pilin promoter but not from rmsp were sequenced to determine the mutation(s) that each carried.

Our analysis revealed two individual mutations, I14T and N16D, near the amino terminus of Rns that abolished positive autoregulation (Fig. 5). This was expected since the mutations were selected because of their inability to activate rmsp. However, the effects of I14T and N16D at rmsp could not have been the result of poor protein expression or protein instability given that each mutation increased Rns-dependent expression from the CS2 pilus promoter, cotBp, compared to Rns. In the case of N16D, the increase was more than 150%. Rns(N16D)
also repressed nlpAp as efficiently as Rns, although the repression by Rns(I14T) was less than that by the wild-type regulator (Fig. 5). Our initial selection also required that the mutations did not abolish Rns-dependent expression from the CFA/I pilus promoter, and as expected, Rns(I14T) activated cfaAp as well as Rns. However, the N16D mutation reduced expression from the CFA/I pilus promoter by as much as 70% compared to Rns (Fig. 5). Evidently, the reduced level of expression from cfaAp was sufficient for growth of Rns-dependent tetracycline-resistant colonies in our initial selection experiment.

**DISCUSSION**

All members of the AraC/XylS family contain dual HTH motifs that provide binding site specificity by placing at least one recognition helix in the major groove of DNA (1, 15, 20, 34). For the majority of family members that consist of two domains, the HTH motifs are usually in the CTD of the protein. In some cases this domain also contains all of the residues necessary for DNA binding and transcriptional activation. For example, MarA and SoxS are roughly one-half the size of a typical family member and are equivalent to the CTD of Rns. Nevertheless, they are able to bind DNA and activate transcription. It has also been shown that the CTDs of XylS and RhaS are sufficient to activate transcription (19, 41). Although the CTD of MelR has also been shown to bind DNA, it is unable to activate transcription (16). In contrast, we have shown that Rns does not bind DNA after removal of residues from its amino terminus. The inability of the Rns CTD to bind DNA cannot be explained by the absence of a dimerization interface because we have also shown that the NTD of Rns does not homodimerize and that full-length Rns is not a multimer in vitro. This differs from the findings for several other family members for which it has been demonstrated that the NTD is sufficient for dimerization, such as AraC, UreR, ToxT, and XylS (4, 31, 33, 35). Because its NTD is unable to homodimerize, Rns may be analogous to PerA, another two-domain virulence regulator belonging to the AraC/XylS family that has also been shown to be a monomer (17).

The NTD of Rns may also lack the other function usually associated with the NTDs of AraC/XylS family members because there is no evidence that it contains a binding site for an exogenous ligand. In general, family members that respond to effector ligands activate the expression of genes that encode proteins that transport, catabolize, or otherwise modify the effector ligand. For example, MarA and SoxS are roughly one-half the size of a typical family member and are equivalent to the CTD of Rns. However, the NTD of Rns lacks the two functions usually associated with the NTDs of AraC/XylS family members, we have shown that it is nevertheless essential for the protein’s activity and have identified two residues, I14 and N16, that are essential for positive autoregulation. These residues are conserved in nearly all of the regulators with which Rns is functionally interchangeable, including the ETEC virulence regulators CfaD and CsvR and AggR from EAEC. They are also conserved in the uncharacterized regulator HdaR (accession no. BAFl33878) from an unusual strain of enterohemorrhagic *E. coli* that also exhibits some of the traits of EAEC. Surprisingly, these residues are not conserved in VirF from *S. flexneri*. Unlike Rns, VirF is not autoregulatory (9); nevertheless, VirF can activate rnsAp in a heterologous system (26). It is not yet known whether VirF activates rnsAp via a similar mechanism, albeit with alternative residues, or through a mechanism that is different than that of Rns autoregulation. However, in the case of Rns autoregulation, I14T and N16D are probably not positive control mutations because they are not transdominant over wild-type Rns (Bodero and Munson, unpublished data).

Curiously, we have also found that both I14T and N16D increase the activity of Rns at the CS2 pilus promoter but not at the CFA/I pilus promoter, even though both promoters contain Rns binding site adjacent to the −35 hexamer. The N16D mutation also had no effect on the ability of Rns to repress nlpAp, while I14T partially relieved repression. It seems unlikely that I14T and N16D disrupt the overall organization of Rns; otherwise, their effects would be similar at each promoter or at least at promoters with similar arrangements of binding sites, such as cfaAp and cotBp. A more plausible explanation for our results is that I14 and N16 are involved in DNA binding either directly or, more likely, through an interaction between the amino terminus of Rns and its CTD. This interaction may have some similarities to the interaction between the amino-terminal arm of AraC and its CTD, which holds the protein in a conformation that binds distally spaced sites (37). In our system, we propose that an interaction between the NTD and CTD of Rns is necessary for the proper spatial orientation of the HTH motifs. This would explain the inability of the Rns CTD to bind DNA in the absence of its NTD even though Rns is not a dimer. Mutations I14T and N16D may alter the domain-domain interaction and subtly change the orientation of CTD residues that make base-specific contact. Since the sequences of the various Rns binding sites are not identical (30), these changes may decrease the occupancy of some binding sites while having minimal impact on the occupancy of other binding sites. Measuring the binding
affinities of Rns, Rns(I14T), and Rns(N16D) to specific sites should allow this proposed mechanism to be evaluated in future studies. Nevertheless, the current results demonstrate that the amino terminus of Rns is essential for its activity even though the NTD of Rns lacks the ability to homodimerize and probably does not bind an effector ligand.

ACKNOWLEDGMENTS

We thank James C. Hu for providing bacterial strains, λ phage, and cI expression plasmids. This research was supported by NIH NIAID Public Health Service award AI057648.

REFERENCES