The RcsAB Box

CHARACTERIZATION OF A NEW OPERATOR ESSENTIAL FOR THE REGULATION OF EXOPOLYSACCHARIDE BIOSYNTHESIS IN ENTERIC BACTERIA*  

Markus Wehland and Frank Bernhard‡  
From the Freie Universität Berlin, Institut für Kristallographie, Takustrasse 6, D-14195 Berlin, Germany

The interaction of the two transcriptional regulators RcsA and RcsB with a specific operator is a common mechanism in the activation of capsule biosynthesis in enteric bacteria. We describe RcsAB binding sites in the uexA promoter of the operon for colanic acid biosynthesis in Escherichia coli K-12, in the galf promoter of the operon for K2 antigen biosynthesis in Klebsiella pneumoniae, and in the tviA (vipR) promoter of the operon for Vi antigen biosynthesis in Salmonella typhi. We further show the interaction of RcsAB with the rcsA promoters of various species, indicating that rcsA autoregulation also depends on the presence of both proteins. The compilation of all identified RcsAB binding sites revealed the conserved core sequence TaAGaatatTCctA, which we propose to termed RcsAB box. The RcsAB box is also part of Bordetella pertussis BvgA binding sites and may represent a more distributed recognition motif within the LuxR superfamily of transcriptional regulators. The RcsAB box is essential for the induction of Rcs-regulated promoters. Site-specific mutations of conserved nucleotides in the RcsAB boxes of the E. coli uexA and rcsA promoters resulted in an exopolysaccharide-negative phenotype and in the reduction of reporter gene expression.

Encapsulation by exopolysaccharides (EPS) protects bacteria against a variety of unfavorable environmental conditions. The production of EPS furthermore represents an essential factor in the virulence of bacterial pathogens (1). The biosynthesis of high molecular weight type I EPS in several enteric bacteria like Escherichia coli (2, 3), Salmonella typhi (4), Klebsiella pneumoniae (5, 6), and the plant pathogenic bacteria Erwinia amylovora (7–10) and Pantoea stewartii (11) is controlled by the Rcs (regulation of capsule synthesis) regulatory network.

Two transcriptional regulators, RcsA and RcsB, are supposed to induce EPS biosynthesis cooperatively. The RcsB protein is highly conserved between different species with about 90% identity (10), and it represents the cytoplasmic activator of a classical bacterial two-component system. RcsB might be activated by the membrane-bound receptor RscC (3, 12) via phosphotransfer to highly conserved aspartic acid residues in the N-terminal domain of RscB. RcsA and RcsB are both characterized by a LuxR-type C-terminal DNA binding motif, but RcsA does not contain an N-terminal phosphorylation motif. The RcsA protein is limiting for the induction of EPS biosynthesis and is virtually not detectable in the uninduced cell due to its rapid degradation by the Lon protease (13, 14). The presence of RcsB is absolutely required for capsule biosynthesis, whereas an rcsA minus phenotype can be suppressed by multicopy rcsB (15). RcsA might therefore act as a coinducer of EPS biosynthesis by enhancing the DNA binding activity of RcsB. Recently, genetic evidence for an autoregulation of rcsA expression in E. coli has been reported and a DNA binding activity of RcsA at the rcsA promoter has been discussed (16).

We have previously shown that a heterodimer formed by one copy of RcsA and RcsB binds at corresponding regions approximately 500 bp upstream of the translational start sites of rpsG and cpsA, the first open reading frames (ORF) in the E. amylovora rpsG operon for amylovoran biosynthesis, and in the P. stewartii cps operon for stewartan biosynthesis, respectively (17, 18). The two operators are highly homologous, and the activation of Rcs-dependent promoters by a RcsAB heterodimer as a general mechanism remained unclear. In addition, some evidence for modified Rcs-dependent regulation mechanisms in E. coli and S. typhi has been proposed (4, 16).

In this work we demonstrate that the binding of RcsAB to regulatory DNA regions is a general principle in the Rcs-mediated activation of gene expression. We present RcsAB binding sites identified in the presumed main promoters of EPS biosynthetic operons of E. coli, K. pneumoniae, and S. typhi. RcsAB further modulates the rcsA autoregulation in those species by binding to the rcsA promoters. The compilation of all identified RcsAB binding sites allows us to define the RcsAB box as a new conserved bacterial operator. The RcsAB box was analyzed in vitro and in vivo, and it was found to be essential for full promoter activity in E. coli.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, Oligonucleotides, and DNA Techniques—The E. coli strains X11-Blue (19), BL21, C600, DH5α (Stratagene), and JB3034 (15) and the plasmids pQE30 (Qiagen), pMalC2 (New England Biolabs), and pfdA8 (20) were used for cloning and expression studies. Standard DNA techniques were done as described (21). DNA fragments were amplified from chromosomal DNA of strain X11-Blue with Vent polymerase and suitable primers. The sequences of the oligonucleotides are available upon request.

Expression and Purification of Proteins—RcsA proteins were produced with the plasmids pM-RcsAEC, pM-RcsAES, and pM-RcsAES (17, 18) in strain BL21 as C-terminal fusions to the maltose-binding protein. RcsB proteins were produced with the plasmids pQ-RcsBEC and pQ-RcsBES (17) with an N-terminal poly(His)6 tag in the strain JB3034. The proteins were purified as described (17, 18). If appropriate, the purified RcsA and RcsB proteins of E. coli, E. amylovora, and P. stewartii were named according to their origin RcsAEC, RcsAES, RcsAES, RcsBEC, and RcsBES, respectively. A BamHI/HindIII restriction fragment from plasmid pQHB (22) containing the coding region of Bordetella pertussis bvgA was cloned into the expression vector pMalC2, resulting in plasmid pM-bvgA. The BvgA protein was produced from

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed. Tel.: 49-30-838-6702; E-mail: fbern@chemie.fu-berlin.de.

‡ The abbreviations used are: EPS, exopolysaccharide; bp, base pair(s); kb, kilobase pair(s); ORF, open reading frame; PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay; SPR, surface plasmon resonance.
strain DH5α × pM-bvgA and purified by affinity chromatography of the crude extract with a dextrin column.

Electrophoretic Mobility Shift Assay (EMSA)—Radioactive DNA labeling with [α-32P]dATP and the EMSA technique were done as described previously (17, 18). The RcsAB heterodimer was obtained by mixing equimolar concentrations of the two proteins. For the reconstitution of double-stranded DNA fragments, about 1 μg of each of two complementary oligonucleotides were mixed in 200 mM Tris/HCl, pH 7.5, 100 mM NaCl and incubated for 5 min at 95 °C. The mixture was cooled slowly to room temperature and subsequently labeled. Phosphorylation of BvgA was obtained by incubating the protein in 50 mM Tris/HCl, pH 7.0, 20 mM MgCl2, 0.1 mM dithiothreitol, and 20 mM acetylphosphate for 20 min at 28 °C. Bovine serum albumin and dithiothreitol were added to the protein solutions of each of two RcsAB box, was cloned into the vector pBluescript KS II and Pst I sites of the suicide gene into the E. coli C600 by electroporation and selected for kanamycin resistance. The correct insertion of the plasmids by homologous recombination was verified by PCR analysis of the isolated chromosomal DNA. An approximately 1-kb DNA fragment generated by PCR and containing the complete E. coli rcsA gene starting 38 bp upstream of the RcsAB box, was cloned into the vector pBluescript KS I—resulting in plasmid prcsA-WT. Four essential positions of the RcsAB box TAAGC were mutated by PCR upon introduction of an EcoRI restriction site and resulting in the sequence GTCACTCTAAAGAATCTCCTTAAAACCATAATTGAAATGACACTTAAATATATTCTTAAAGAATGCCCCATTGTTATCTTGCCTGCTA

FIG. 1. Analysis of the RcsAB binding site in the E. coli wza promoter. The RcsAB box is underlined twice, and putative promoter consensus sequences are underlined once. Numbers indicate nucleotide positions relative to the transcriptional start site. Fragments analyzed for RcsAB binding in EMSAs are indicated by lines and designated as in the text. Retarded fragments are marked with ‘++’, nonretarded fragments with ‘—’. Bases analyzed by mutagenesis are shown above the sequences, the corresponding positions are given in brackets as mentioned in the text. Up mutations are in bold, and down mutations are italic.

The enzymatic activity of the β-galactosidase was determined with the o-nitrophenyl β-D-galactopyranoside assay after Miller (24). RESULTS

**Location of an RcsAB Binding Site in the E. coli wza Promoter.** The 485-bp PCR fragment Pwza1, containing the putative promoter and the first nine codons of wza, the first ORF in the operon for colanic acid biosynthesis in E. coli, interacted with RcsABEC in EMSAs. Terminal deletions revealed the 55-bp fragment Pwza55, spanning nucleotide positions −119 to −65 relative to the transcriptional start site of wza (Fig. 1) as sufficient for a retardation by the RcsABEC heterodimer (Fig. 2). The extent of retardation was diminished with the 41-bp fragment Pwza41 from positions −119 to −79, and no retardation was observed with the 27-bp fragment Pwza27 spanning...
nucleotide positions 106 to 80. This was also observed with the 38-bp fragment Pwza58 spanning nucleotide positions 119 to 82. This indicated that the 28-bp region spanning nucleotide positions 106 to 79 relative to the transcriptional start site of wza was essential but not sufficient for the binding of RcsABEC (Figs. 1 and 2). An extension of the 3' end of the 55-bp fragment did not further contribute to a better binding of RcsABEC, and the extent of retardation of the 72-bp fragment Pwza72 spanning nucleotide positions 119 to 48 was comparable to that of Pwza55 (data not shown). Incubation at 28 °C compared with 37 °C prior to electrophoresis increased the percent of retardation of the RcsAB EC/DNA complex about 3-fold.

The alignment of the 55-bp fragment with the RcsAB binding site of the E. amylovora ansA promoter (18) implicated several nucleotides in the region from −115 to −96 relative to the transcriptional start site of wza as putative targets for RcsAB (Fig. 1). Suspected positions were further analyzed by the introduction of putative up and down mutations according to the RcsABansG consensus (Table I). The replacement of the degenerated adenine at position −110 by a conserved guanine in the fragments Pwza72 (G109) and Pwza72 (G107A108) increased the extent of retardation by RcsABEC (Table II). In contrast, the retardation was considerably reduced after replacing the highly conserved thymine at position −112 in fragment Pwza72 (G112) by guanine. The retardation was completely abolished in fragment Pwza72 (C110T109C109) carrying mutations in three conserved positions. A decreased percent of retardation in an EMSA was furthermore observed after the replacement of two less conserved adenines by cytosines in the fragment Pwza72 (C108G106). Interestingly, two mutations adjacent to the putative RcsAB consensus also reduced the extent of retardation of fragment Pwza72 (G106A98) by the two proteins (Table II). This gives evidence for additional DNA/protein interactions neighboring the consensus motif.

To analyze whether the wza promoter exhibits some preference for the recognition by the homologous RcsABEC proteins, we used the 55-bp fragment of the wza promoter as a target in EMSAs with various combinations of the RcsAEC, RcsABEC, RcsAEC, RcsBEA, and RcsBEC proteins (Fig. 2). The wza promoter was recognized by the heterologous Rcs proteins in all combinations tested, and we could not detect significant differences in the extent of retardation of the DNA fragment. We observed no binding of RcsAEC or RcsBEC alone to the 55-bp fragment, or to the 485-bp fragment containing the complete wza promoter in concentrations up to 4.5 μM. In contrast, approximately 0.2 μM of the two proteins together already retarded these DNA fragments in EMSAs. The DNA fragment Pwza55 spanning the nucleotide positions −95 to −48, including an inverted repeat sequence, was retarded neither by RcsAEC nor by RcsABEC (Fig. 2). In addition, the mutation of four bases in the inverted repeat sequence of the fragment Pwza55 (C−68G−69G−66A−65) did not show any effect on the retardation by RcsABEC (Table II).

The binding kinetics of the RcsABEC heterodimer at Pwza72 was analyzed by the surface plasmon resonance technique (Fig. 3). With protein concentrations in a range of 47 nM and 7.5 μM, the $k_d$ was calculated to 5.4 ± 3.3 x 10^{-3} s^{-1} and the $k_d$ to 1.4 ± 0.4 x 10^{-3} s^{-1}, resulting in a $K_D$ of 77 ± 28 nM. The $K_D$ of RcsABEC at the E. coli wza promoter corresponds to the $K_D$ of RcsABEA at the E. amylovora ansG promoter previously determined by the EMSA technique (18). We furthermore analyzed the DNA fragment Pwza72 (G112C109T109C109) containing four point mutations in highly conserved positions (Fig. 1). This fragment was not retarded by RcsAB in an EMSA (Table II). The $k_d$ was calculated to 560 ± 140 M^{-1} s^{-1} and the $k_d$ to 2.4 ± 1.0 x 10^{-2} s^{-1}, resulting in an approximately 10^3-fold increased $K_D$ of 50 ± 30 μM.

**In Vivo Analysis of the E. coli wza RcsA Binding Site by Mutagenesis**—The chromosomal merodiploids MW29 and MW31 of the E. coli K12 strain C600 were constructed after integration of the two plasmids pMW29 and pMW31 by homologous recombination. The wza promoter of strain MW31 was truncated just upstream of the RcsAB box, and strain MW29 contained additionally four point mutations in essential bases within the identified RcsAB binding site. The phenotypes of the mutants were assayed after introduction of plasmid pEA101 containing the E. amylovora rcsA gene, which resulted in the wild type strain C600 in the induction of colanic acid biosynthesis by activation of the wza promoter. The EPS production and the phenotype of the control mutant MW31 × pEA101 was not altered compared with the wild type strain C600 × pEA101 (Table III). Thus, the approximately 450-bp fragment upstream of wza is sufficient for full promoter activity. In contrast, the EPS production of the mutant MW29 × pEA101 was drastically reduced and the mutant showed a butyrous colony type (Table III). These results demonstrate the importance of the identified RcsAB binding site for the activation of colanic acid biosynthesis, and they indicate that RcsAB might also bind in vivo to that region.

**RcsA and RcsB Bind to the rcsA Promoters of E. coli, K. pneumoniae, S. typhi, and E. amylovora**—The autoregulation of E. coli rcsA has been reported previously (16), and we inves-
This work (31) proteins, as neither RcsA nor RcsB alone in concentrations of up to population appears to be dependent on the presence of both pro-
ceeding site was detected at nucleotide positions 2
binding are three conserved purines most likely represented by
rcsA in vitro the 277-bp fragment by RcsABEC (Fig. 4
promoters of other species. We detected putative RcsAB bind-
are located nucleotides if applicable. Nucleotide positions
Consensus in the in vitro selected RcsAB box in the context of the E. amylovora amsg promoter; R = A + G, V = A + C + G, W = A + T, S =
A Correction The RcsAB box is in bold, palindromic nucleotide positions within the RcsAB box are underlined, and the colon represents the center of symmetry.
* Nucleotide positions relative to the translational start site.
* Reference for the corresponding sequences.
* Consensus of the 12 presented sequences showing the degree of conservation: uppercase letter = 70%, lower case letter = 50%, dot < 50%.

| Gene | Species | RcsAB box<sup>a</sup> | Location<sup>b</sup> | Ref<sup>c</sup> |
|------|---------|-----------------|-----------------|-------|
| wza | E. coli K12 | aacc taagaa:actcctaa aaaa | -452/-439 (25) |
| galQ (orfL) | K. pneumoniae K2 | aaaa taagaa:atcctaa cttc | -181/-168 (26) |
| tsvA (svaR) | S. typhi | cgaat taagaa:atcctaa tttt | -322/-309 (27) |
| amsg | E. amylovora | atat tgaagatt:aatcctaa tttt | -550/-537 (28) |
| cpsA | P. stewartii | aaca tggataa:atcctga tttt | -537/-524 (18) |
| rcsA | E. coli K12 | atcc taagaa:atcctga aaaa | -264/-251 (29) |
| rcsA | S. typhi | ttac tgaagatt:aatcctaa aaaa | -265/-252 (4) |
| rcsA | K. pneumoniae | gaag tgaagaa:atcctaa aagt | -257/-244 (5) |
| rcsA | E. amylovora | aatt taagaa:atcctaa tctt | -318/-305 (7–9) |
| bvgA | B. parapertussis | gat taagaa:tttcctaa tttt | -175/-162 (30) |
| bvgA | B. pertussis | gaat tcaagaa:tttcctaa tttt | -176/-163 (30) |
| fha | B. pertussis | tgac tgaagaa:tttcctaa caag | -165/-152 (30) |
| Consensus<sup>d</sup> | | aaa Taagaa:atcctaa .ttt | This work |
| RcsAB<sub>amsG</sub><sup>e</sup> | | TRVGAAM:AWSTYSGR | (18) |
| luxI box | | acc tgaagaa:tttcctaa ggt | (91) |

<sup>a</sup> The RcsAB box is in bold, palindromic nucleotide positions within the RcsAB box are underlined, and the colon represents the center of symmetry.
<sup>b</sup> Nucleotide positions relative to the translational start site.
<sup>c</sup> Reference for the corresponding sequences.
<sup>d</sup> Consensus of the 12 presented sequences showing the degree of conservation: uppercase letter = 70%, lower case letter = 50%, dot < 50%.
<sup>e</sup> Consensus of the in vitro selected RcsAB box in the context of the E. amylovora amsg promoter; R = A + G, V = A + C + G, W = A + T, S = G + C, Y = C + T.

| Fragment<sup>a</sup> | wza promoter sequence<sup>b</sup> | Retardation<sup>c</sup> | % |
|-----------------|-----------------|-----------------|---|
| Pu224<sub>277</sub> | TAAAGAAACTCCTAAA | 25.2 ± 3.4 |
| Pu224<sub>277</sub> (G<sup>c</sup>118) | GAAAGAAACTCCTAAA | 3.6 ± 0.1 |
| Pu224<sub>277</sub> (C<sup>c</sup>110 $\cdot$ 108C<sup>c</sup>106) | TACATAACACTCCTAAA | 0 |
| Pu224<sub>277</sub> (G<sup>c</sup>109) | TAAAGAAACTCCTAAA | 33.6 ± 1.2 |
| Pu224<sub>277</sub> (G<sup>c</sup>109A<sup>c</sup>108) | TAAAGAAACTCCTAAA | 30.3 ± 2.2 |
| Pu224<sub>277</sub> (G<sup>c</sup>92C<sup>c</sup>96) | TAAAGAAACTCCTAGCG | 8.6 ± 1.4 |
| Pu224<sub>277</sub> (G<sup>c</sup>92C<sup>c</sup>96T<sup>c</sup>108C<sup>c</sup>107) | TAAAGAAACTCCTAAAG | 0 |
| Pu224<sub>277</sub> (G<sup>c</sup>91C<sup>c</sup>90) | TAAAGAAACTCCTAAA-N<sup>c</sup>1-GC | 14.3 ± 2.1 |
| Pu224<sub>277</sub> (G<sup>c</sup>79G<sup>c</sup>68G<sup>c</sup>68) | TAAAGAAACTCCTAAA-N<sup>c</sup>15-GCGA | 22.2 ± 1.8 |

<sup>a</sup> DNA fragments of 72 bp were used for the EMSA.
<sup>b</sup> Relative to the wza transcriptional start site from position -112 to -96 and further 3’ located nucleotides if applicable. Nucleotide positions analysed by mutagenesis are underlined.
<sup>c</sup> Protein concentrations used for EMSA analysis were 2 μM RcsAB<sub>EC</sub>.

tigated whether the activation of rcsA promoters is also di-
rected via DNA binding of RcsAB. The 277-bp PCR fragment
Pr<sub>rcsA<sub>EC</sub>277</sub> containing the intergenic region between the E. coli
filR and rcsA genes including the start codon of rcsA<sub>EC</sub> was
clearly retarded by RcsAB<sub>EC</sub> (Fig. 4A). A putative RcsAB bind-
ing site was detected at nucleotide positions -264 to -251
relative to the translational start site of rcsA<sub>EC</sub> (Table I). Ac-

The compilation of the RcsAB binding sites of the four rcsA
promoters and of the promoters of wza, amsg, and cpsA (Table I)
revealed several highly conserved positions within a core
sequence of 14 bp, which we will further term RcsAB box.

The RcsAB Box Is Essential for rcsA Autoregulation in E. coli.—The plasmids pacsA-WT, containing the cloned E. coli
rcsA gene with an approximately 300-bp upstream region in-
cluding the RcsAB box and pacsA-M4, containing mutations in four
elemental positions of the RcsAB box, were transformed in the E. coli strain DH5α, and the bacteria were analyzed for
their phenotype. The plasmid pacsA-WT with the wild type
rcsA promoter increased the EPS production and resulted in a
fluid colony type. This was obviously due to an increased
rcsA copy number (Table III). In contrast, the colonies re-
mained butyrous with plasmid pacsA-M4 containing the four
point mutations in the RcsAB box, and the EPS production was
dramatically decreased compared with strain DH5α × prcsA-WT. These results show that the RcsAB box is also essential for the rcsA autoregulation in vivo in E. coli.

In contrast, the introduction of plasmid prcsA-M4 in the lon minus strain SG1087 resulted in a fluidal phenotype and in an increased EPS production (Table III). Accordingly, the expression of a cpsB::lacZ fusion in the lon minus strain JB3034 clearly increased upon introduction of plasmid prcsA-M4 (Table III), but was still lower compared with that of strain JB3034 × prcsA-WT. The absence of the Lon protease increases the half-life of RcsA. The background expression of rcsA with plasmid prcsA-M4 is then obviously sufficient for the activation of the colanic acid biosynthetic operon.

**Identification of an RcsAB Box in the Gene Cluster for K2 Antigen Expression in K. pneumoniae, and for Vi Antigen Expression in S. typhi**—The rcsA and rcsB genes have also been described from the enteric bacteria *S. typhi* and *K. pneumoniae*. If the corresponding capsular polysaccharide biosynthetic gene clusters were regulated by RcsAB, a RcsAB box should be present in the main promoters. The first ORF of the *S. typhi* Vi antigen cluster encodes for the putative regulator TviA (VipR) and a putative RcsAB box was found at position −322 to −309 relative to the translational start site of *tviA* (Table I). The 60-bp fragment *PviA* spanning nucleotides −347 to −288 was retarded by RcsABEC in EMSAs (Fig. 6). This indicates that the *tviA* RcsAB box is recognized in vitro by the two proteins.

16 ORFs are described for the *K. pneumoniae* K2 antigen biosynthesis gene cluster. An approximately 0.9-kb intergenic region containing a putative s54-dependent promoter precedes orf3, but the DNA fragment does not contain a RcsAB box. However, a RcsAB box is located at positions −218 to −168 relative to the translational start site of *orf3*, but the DNA fragment does not contain a RcsAB box. The Heterodimer RcsAB and the Transcriptional Regulator BvgA of Bordetella pertussis Recognize Similar DNA Sequences—Similarity searches with the programs MEME and MAST (32) revealed potential RcsAB boxes within the regulatory regions *bugA* and *fhaB* of *B. pertussis* and *Bordetella parapertussis* (Table I). The two promoters are reported to be activated by BvgA, a transcriptional regulator of the LuxR superfamily, whose DNA binding domain is homologous to that of RcsA and

**Fig. 3. SPR analysis of the RcsAB binding properties.** SPR measurements were performed with the immobilized DNA fragments P\textsubscript{wza\textsubscript{72}} (solid line) and P\textsubscript{wza\textsubscript{72}(G\textsubscript{2}112C\textsubscript{2}109T\textsubscript{2}108C\textsubscript{2}107)} (dotted line). Resonance units with the two fragments were determined with a random 72-bp DNA fragment as a control; the beginning of the injection was defined as 0 s. In the presented diagrams, RcsB was used at 750 nM and RcsA at 3.75 μM.

**TABLE III**

Phenotype of mutated RcsAB boxes in E. coli

| Strain<sup>a</sup> (genotype) | Plasmid | Colony type<sup>b</sup> | EPS production<sup>c</sup> | cpsB expression<sup>d</sup> |
|-------------------------------|---------|------------------------|----------------------------|---------------------------|
| C600 (wt)                     |         | B                      | 6.0 ± 0.4                  | —                         |
| MW31                          | pEA101  | B                      | 15.2 ± 0.3                 | —                         |
| MW29                          | pEA101  | B                      | 0.3 ± 0.1                  | —                         |
| DH5α (wt)                     | prcsA-WT| B                      | 5.8 ± 0.6                  | —                         |
| SG1087 (rcsA, lon)            | prcsA-WT| F                      | 6.0 ± 0.4                  | 2 ± 0.1                   |
| JB3034 (rcsA, lon, cpsB::lacZ)| prcsA-WT| —                      | 382 ± 84                   | 67 ± 38                   |
|                              | prcsA-M4| —                      |                            |                          |

<sup>a</sup> Relevant genotype.

<sup>b</sup> Colony type determined after 24 h of growth on LB-agar at 37 °C.

<sup>c</sup> After 24 h of growth on LB-agar at 37 °C and estimated with the anthron assay. Means of at least three determinations. —, not applicable.

<sup>d</sup> β-Galactosidase units estimated after Miller (24). Means of at least three determinations.
RcsB. The two 50-bp DNA fragments \( P_{bvgA\_BP} \) and \( P_{bvgA\_BA} \), containing the putative RcsAB boxes of the \( B. \) pertussis and \( B. \) parapertussis \( bvg \) promoters, respectively, and the 50-bp fragment \( P_{fha} \) containing the putative RcsAB box of the \( B. \) pertussis \( fha \) promoter, were retarded by RcsAB in EMSAs (Fig. 7). The fragment \( P_{fha} \) was also retarded in vitro by the purified BvgA protein (Fig. 7), indicating that the homology of the DNA binding domains of RcsAB and BvgA is sufficient to result in the recognition of similar DNA sequences. The \( luxI \) box, a potential binding site for the LuxR protein, is not related to the RcsAB box (Table 1), and it was not retarded by RcsAB in an EMSA (data not shown).

**DISCUSSION**

The RcsAB box has now been identified in promoters of EPS biosynthetic operons of five different species. It is always present in the promoter region preceding the first ORF, while the organization of the operons is quite variable. The \( E. \) amylovora \( amsG \) and \( P. \) stewartii \( cpaS \) genes are homologous and encode for a putative UDP-galactose lipid-carrier transferase (28). The first ORF of the \( E. \) coli colanic acid biosynthetic operon, \( uza \), encodes for a putative outer membrane lipoprotein (33) and shows homologies to \( amsH \), the second ORF in the \( ams \) operon. The \( tviA \) (\( vrp \)) gene is so far unique to \( S. \) typhi and encodes for a putative regulator protein (4, 34). The \( orfI \) of \( K. pneumoniae \) encodes for a GalF homologue, while \( galF \) represents the last ORF of the \( E. \) coli \( uza \) operon (33). The location of the RcsAB box is therefore correlated to the putative main promoters of EPS biosynthetic operons and not to a specific gene.

The regulation of colanic acid biosynthesis in \( E. \) coli K12 by Rcs proteins has been shown (35). A fragment of approximately 470 bp of the \( uza \) promoter was found to be sufficient for an RcsAB dependent activation (25). This is in agreement with the phenotype of our mutant MW31. The characterized RcsAB box is located just at the 5′-end of this essential promoter region. As observed with the \( amsG \) promoter (18), RcsA alone was not able to bind to the \( uza \) promoter. In addition, an inverted repeat sequence, previously suspected to be an RcsA binding site (16), does not seem to be essential for the in vitro binding of RcsAB. However, some beneficial effects for the RcsAB binding at the \( uza \) promoter could not be ruled out as the deletion of sequences including the repeat resulted in a reduced extent of retardation by RcsAB. A guanine, which has been shown to be important for RcsAB binding at the \( amsG \) promoter (18), is replaced by an adenine in the RcsAB box of the \( uza \) promoter. In addition, the minimal size of the RcsAB binding site at the \( uza \) promoter is obviously larger compared with the previously identified 23-bp sites in the \( E. \) amylovora \( amsG \) and \( P. \) stewartii \( cpaS \) promoters (18). The reduced binding of RcsAB to the degenerated box might therefore be stabilized by additional DNA/protein contacts. The phenotype of the mutant MW29 demonstrates an essential role of the RcsAB box for EPS production in vivo, and it is in full accordance with the results obtained by the in vitro binding studies.

The \( rcs \) regulation of EPS biosynthesis was so far demonstrated in about 10 serotypes of \( K. pneumoniae \) (5, 6) and in about 20 serotypes of \( E. \) coli (3), all producing structurally different capsular polysaccharides. While the complete sequence is so far only available for the serotype \( K. pneumoniae \) K2 biosynthetic operon, the presence of a RcsAB box can be expected also in the promoters of the other operons. Elevated copies of RcsB increased EPS biosynthesis in \( E. \) coli K30, but the phenotype of \( rcsA \) and \( rcsB \) mutants indicated that both genes are not essential for low level EPS biosynthesis (3). The large noncoding upstream regions of \( orf3 \) of the K2 biosynthetic operon and of its homologue \( orfX \) of the K30 biosynthetic operon contain a putative \( a^{+} \)-dependent promoter, while the only detectable RcsAB box in the K2 operon is present upstream of \( galF \) (\( orf1 \)). RcsAB might therefore activate the expression of \( galF \) (\( orf1 \)) and possibly \( orf2 \), whereas the further downstream located genes might depend upon the regulation of additional mechanisms. The product of \( galF \) (\( orf1 \)) might increase the biosynthesis of activated sugar precursors, and the homology of \( orf2 \) to the EPS related gene \( orf1 \) of \( Aeromonas \) \( hydrophila \) indicates some involvement of its gene product in K2 polysaccharide biosynthesis.

The Vi antigen, a linear homopolymer of \( \alpha-1,4 \) 2-deoxy-2-N-acetylgalactosamine uronic acid, is produced by all strains of \( S. \) typhi, and \( S. \) paratyphi as well as from some strains of \( S. \) dublin and \( Cibrobacter freundii \) (36, 37). The RcsB-dependent biosynthesis of the Vi polysaccharide has been reported (4, 38), and the \( rcsA \) and \( rcsB \) genes have been isolated from \( S. \) typhi (4). However, an involvement of RcsA in the...
regulation of Vi antigen biosynthesis could not be shown (4). Our findings indicate a potential molecular target for the interaction of RcsA together with RcsB upstream of tviR (vipR) at the presumed main promoter for Vi antigen biosynthesis. The interaction of the putative regulator TviA (VipR) with its own promoter, containing the RcsAB box, and with RcsB (4) has been proposed (34). It will be interesting to elucidate the in vivo role of the RcsAB box in this complex regulation mechanism.

The detection of an RcsAB box in corresponding regions in the rcsA promoters of E. coli, K. pneumoniae, S. typhi, and E. amylovora agrees with a previously proposed autoregulation mechanism of rcsA expression in E. coli (16). The self-activation of the rcsA promoter might counteract a silencing mechanism on rcsA expression by the histone-like protein H-NS (39), and it could be important for the rapid increase in EPS biosynthesis as a fast response on environmental stimuli. We first present evidence that the rcsA self-activation is dependent on both RcsA and RcsB. The rcsA self-activation is not confined to E. coli but might be a general mechanism in enteric bacteria. We also demonstrate that the RcsAB box, previously identified by in vitro DNA/protein interaction studies, is essential for full activity of the rcsA promoter in vivo. An inverted repeat sequence, present upstream of the RcsAB box, is dispensable at least for the in vitro binding of RcsAB. In addition, the rcsA promoter remained highly active in plasmid prcsA-WT despite the truncation of this sequence. The chromosomal rcsA expression in an E. coli lon strain was rcsB-dependent, while the RcsA protein could be detected in the same strain containing multicopy rcsA (40). The mechanism of autoregulation requires some leakiness of the rcsA promoter. Its activity might therefore be enhanced but not strictly depend on the binding of RcsAB. This is in accordance with the described cpsB induction by the high copy plasmid prcsA-M4, containing a mutated RcsAB box, in an E. coli lon strain. The leaky expression from multicopy rcsA is obviously sufficient to activate some EPS biosynthesis, when the half-life of RcsA is drastically increased by the lon mutation (14).

The consensus RcsAB box with the sequence TaAGaatatTC-ta was determined out of 12 identified wild type RcsAB boxes from the promoters of K. pneumoniae, Pseudomonas aeruginosa (P. aeruginosa), and E. coli. The overall identity of the three promoters is 42.7%. Identical positions are marked by asterisks; the rcsA start codons, putative ribosomal binding sites (RBS), and putative promoter consensus sequences are underlined. The RcsAB boxes are shaded; nucleotides in italic above the aligned sequences indicate mutated positions in plasmid prcsA-M4.
attracting the RNA polymerase or in stabilizing the transcriptional complex.

A consensus motif present in the C-terminal DNA binding domains of transcriptional regulators of the LuxR superfamily indicates some common mechanisms in protein/DNA recognition (42). While the lux box, a 20-bp repeat centered at approximately −40 bp from the lux transcriptional start proposed to be involved in the DNA binding of LuxR from *Vibrio fischeri* (31, 43), is not related to the RcsAB box, similar DNA sequences are recognized by RcsAB and by BvgA, a response regulator of *B. pertussis* virulence factor genes. The DNA binding of BvgA at the RcsAB boxes in the *pha* and the *bvg* promoters has previously been shown (44–46). The BvgA protein is also known to bind to the *ptx* promoter controlling the expression of pertussis toxin, which does not contain a RcsAB box. However, BvgA binds about 10 times weaker to that promoter and evidence for a different mechanism of DNA recognition has been proposed (45–47). The BvgA binding sites at the *pha* and *bvg* promoters are located from approximately −100 to −70 upstream of the transcriptional start sites, which exactly corresponds to the location of the RcsAB box in the *E. coli uwa* and *rcsA* promoters. The interaction of the α-subunit of the RNA polymerase with BvgA has been demonstrated (43–45), and a similar function for RcsAB might be proposed. The homology of the DNA binding domains of BvgA and RcsAB in addition to the observed recognition of similar DNA sequences could point to a common phylogenetic origin. A RcsAB boxlike sequence might therefore be recognized also by some other members of the LuxR superfamily.

Acknowledgments—We thank C. Langner and R. Diehl for technical assistance. We are grateful to W. Saenger for helpful discussions and critical review of the manuscript, and to J. Schneider-Mergener and C. Landgraf for providing the BIACore technique and for helpful assistance. We further thank W. Schröder for providing us with oligonucleotides and C. Kiecker for help in the initial part of this project.

REFERENCES

1. Leigh, J. A., and Coplin, D. L. (1992) *Annu. Rev. Microbiol.* 46, 307–346
2. Gottesman, S., Trisler, P., and Torres-Cabassa, A. (1985) *J. Bacteriol.* 162, 1111–1119
3. Jayaratne, P., Keenleyside, W. J., MacLachlan, P. R., Dodgson, C., and Whitfield, C. (1993) *J. Bacteriol.* 175, 5384–5389
4. Virlogeux, I., Waxin, H., Keshibich, C., Lee, J. O., and Popoff, M. Y. (1996) *J. Bacteriol.* 178, 1691–1698
5. Allen, P., Hart, C. A., and Saunders, J. R. (1987) *J. Gen. Microbiol.* 133, 331–340
6. Mc Callum, K. L., and Whitfield, C. (1991) *Infect. Immun.* 59, 494–502
7. Bernhard, F., Poetter, K., Geider, K., and Coplin, D. L. (1990) *Mol. Plant Microbe Interact.* 3, 429–437
8. Chatterjee, A., Chua, W., and Chatterjee, A. K. (1990) *Mol. Plant Microbe Interact.* 3, 144–148
9. Coleman, M., Pearce, R., Hitchin, E., Busfield, F., Mansfield, J. W., and Roberts, J. S. (1990) *J. Gen. Microbiol.* 136, 1799–1806
10. Bereswill, S., and Geider, K. (1997) *J. Bacteriol.* 179, 1554–1561
11. Poetter, K., and Coplin, D. L. (1991) *Mol. Gen. Genet.* 229, 155–160
12. Stout, V., and Gottesman, S. (1990) *J. Bacteriol.* 172, 659–669
13. Torres-Cabassa, A. S., and Gottesman, S. (1987) *J. Bacteriol.* 169, 981–989
14. Stout, V., Torres-Cabassa, A., Maurizi, M. R., Gutnick, D., and Gottesman, S. (1991) *J. Bacteriol.* 173, 1738–1747
15. Brill, A. J., Quiman-Walshe, C., and Gottesman, S. (1988) *J. Bacteriol.* 170, 2599–2611
16. Ebel, W., and Trempe, J. E. (1999) *J. Bacteriol.* 181, 577–584
17. Kelm, O., Kiecker, C., Geider, K., and Bernhard, F. (1997) *Mol. Gen. Genet.* 256, 72–83
18. Wehland, M., Kiecker, C., Coplin, D. L., Kelm, O., Saenger, W., and Bernhard, F. (1999) *J. Biol. Chem.* 274, 3300–3307
19. Bullock, W. O., Fernandez, J. M., and Stuart, J. M. (1987) *BioTechniques* 5, 376–379
20. Geider, K., Hohmeyer, C., Haas, R., and Meyer, T. F. (1985) *Gene (Amst.*) 33, 341–349
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
22. Boucher, P. E., Menozzi, F. D., and Locht, C. (1994) *J. Mol. Biol.* 241, 363–377
23. Dubois, M., Giles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, P. (1956) *Anal. Chem.* 28, 350–356
24. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Stout, V. (1996) *J. Bacteriol.* 178, 4273–4280
26. Arakawa, Y., Wacharotayankun, R., Natsuka, T., Ito, H., Kato, N., and Ohta, M. (1995) *J. Bacteriol.* 177, 1788–1796
27. Hashimoto, Y., Li, N., Yokoyama, H., and Ezaki, T. (1993) *J. Bacteriol.* 175, 4456–4465
28. Bugert, P., and Geider, K. (1995) *Mol. Microbiol.* 15, 917–935
29. Blatntier, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Bode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* 277, 1453–1474
30. Scharf, L., Prugnola, A., Arici, B., and Rappuoli, R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 6753–6757
31. Devine, J. H., Shadel, G. S., and Baldwin, T. O. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 5688–5692
32. Bailey, T. L., and Gribskov, M. (1998) *Bioinformatics* 14, 48–54
33. Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996) *J. Bacteriol.* 178, 4885–4893
34. Hashimoto, Y., Khan, A. Q., and Ezaki, T. (1996) *J. Bacteriol.* 178, 1430–1436
35. Gottesman, S., and Stout, V. (1991) *Mol. Microbiol.* 5, 1599–1606
36. Baker, E. E., Whiteside, R. E., Basch, R., and Derow, M. A. (1959) *Biochimica et Biophysica Acta* 5, 577–565
37. Stout, V. (1996) *J. Bacteriol.* 178, 5688–5692
38. Balley, T. L., and Gribskov, M. (1998) *Bioinformatics* 14, 48–54
39. Dierksen, K. P., and Trempy, J. E. (1996) *J. Bacteriol.* 178, 2483–2489
40. Henikoff, S., Wallace, J. C., and Brown, J. P. (1990) *Methods Enzymol.* 183, 111–132
41. Stevens, A. M., Dolan, K. M., and Greenberg, E. P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 12619–12623
42. Boucher, P. E., Morakami, K., Ishihama, A., and Stibitz, S. (1997) *J. Bacteriol.* 179, 1755–1763
43. Zu, T., Manetti, R., Rappuoli, R., and Scarlato, V. (1996) *Mol. Microbiol.* 21, 557–565
44. Roy, C. R., and Falkow, S. (1991) *J. Bacteriol.* 173, 2385–2392
45. Stibitz, S. (1994) *J. Bacteriol.* 176, 5615–5621