JOURNAL OF BACTERIOLOGY, Feb. 1990, p. 659–669
0021-9193/90/020659-11$02.00/0

Vol. 172, No. 2

RcsB and RcsC: a Two-Component Regulator of Capsule Synthesis in *Escherichia coli*

VALERIE STOUT* AND SUSAN GOTTESMAN

Laboratory of Molecular Biology, National Cancer Institute, Bethesda, Maryland 20892

Received 6 September 1989/Accepted 16 November 1989

Colanic acid capsule synthesis in *Escherichia coli* K-12 is regulated by RcsB and RcsC. The amino acid sequences of these two proteins, deduced from the nucleotide sequence reported here, demonstrate their homology to environmentally responsive two-component regulators that have been reported in both gram-positive and gram-negative bacteria. In our model, RcsC acts as the sensor and RcsB acts as the receiver or effector to stimulate capsule synthesis from *cps* genes. In addition, RcsC shows limited homology to the other effectors in its C terminus. Fusions of *rcsC* to *phoA* that resulted in PhoA$^+$ strains demonstrated that RcsC is a transmembrane protein with a periplasmic N-terminal domain and cytoplasmic C-terminal domain. Additional control of this regulatory network is provided by the dependence on the alternate sigma factor, RpoN, for the synthesis of RcsB. The *rcsB* and *rcsC* genes, which are oriented convergently with their stop codons 196 base pairs apart, are separated by a long direct repeat including two repetitive extragenic palindromic sequences.

Colanic acid capsular polysaccharide is normally synthesized by *Escherichia coli* K-12 in low amounts. Several regulatory mutations that increase the levels of this polysaccharide more than 10-fold have been described (12, 26, 43, 48). These mutations define the negative regulators, *ops*, *lon*, and *rcsC*. Lon is a major ATP-dependent protease in *E. coli* that increases synthesis of the polysaccharide capsule through increased expression of the *cps* genes (43). One of the in vivo substrates of Lon protease is the positive regulator of capsule synthesis, RcsA. Stabilization of RcsA in a *lon* mutant can account for the overproduction of capsular polysaccharide in these cells (41).

The gene for the other negative regulator, *rcsC*, maps adjacent to a second positive regulator, *rcsB*, at 48 min, and they appear to be transcribed toward each other (4). RcsB is a protein with a molecular mass of about 26,000 daltons (Da) and may act as a multimer to allow expression of the capsule genes. RcsC is a 100,000-Da protein. Recessive mutations in *rcsC* (such as *rcsC137*) result in overexpression of capsular polysaccharide but do not act as a bypass for RcsA or RcsB.

There are parallels between the capsule regulatory network and other regulatory pathways in gram-negative bacteria. In the past few years, a new family of regulatory proteins has been identified, members of which control responses to environmental stimuli and are related by protein sequence homology (reviewed in reference 35). This family of two-component regulators is composed of (i) a sensor protein that detects the environmental stimulus and transmits a signal to (ii) an effector or regulator component that elicits a response, usually at the level of transcription. The sensors are often membrane proteins; all share homology in their C-terminal domains over about 250 amino acids. The effectors share homology over about 100 amino acids in the N-terminal domains and are always cytoplasmic. Generally, the genes encoding the sensor and effector pair are located near each other, often in the same operon (2, 5, 7, 11, 13, 29, 33, 35). The cellular functions regulated by members of this family include response to nitrogen and phosphate limitation, osmolarity, anaerobiosis, and chemotaxis.

* Corresponding author.

We have previously suggested that RcsC may act with RcsB in regulating capsule synthesis, in a manner similar to that seen with the two-component regulators (4). Here we show that RcsB and RcsC share sequence homology with other two-component regulator pairs in the appropriate domains. In addition, our results indicate that RcsC is a transmembrane protein with a periplasmic N-terminal domain and a cytoplasmic C-terminal domain.

MATERIALS AND METHODS

Bacterial strains and media. *E. coli* K-12 strains and the plasmids used are listed in Table 1. Generally, cells were grown at 37°C in LB medium with shaking. Strains containing *rcsC* on a multicopy plasmid (e.g., pJB201) could be grown only to an optical density at 600 nm of 0.4 before cells began to lyse. Strains containing the *cpsB10*::*lacZ* fusions were used to monitor capsule expression. This fusion is expressed at high levels in hosts carrying the *rcsC137* mutations and at low levels in *lon* mutant hosts carrying *rcsB* mutations.

DNA manipulation. All DNA manipulations were performed by using techniques described by Maniatis et al. (24). pATC591 was constructed by Angel Torres-Cabassa by inserting a *Hin*dIII linker into the *Sma*I site of pRS591 (38). The 3.6-kilobase-pair (kb) *Hin*dIII-*Bam*HI fragment of pJB100 (4), containing *rcsB*, was transferred into M13mp18 or M13mp19 that was cut with the same enzymes. Similarly, the 4.0-kb *Eco*RI-*Pst*I fragment containing *rcsC* was ligated into M13mp18 and M13mp19. M13mp18 and M13mp19 DNA containing *rcsB* and *rcsC* were sequenced by using deletions that were generated by the cyclone deletion method (6), using the Cyclone kit from International Biotechnologies, Inc. (Fig. 1). To fill gaps in the sequence not covered by the nested set of deletions, oligonucleotides were synthesized by using an Applied Biosystems 380B DNA synthesizer and used as primers. Sequences of the fusion junctions between *rcsB* and *lacZ* and between *rcsC* and *phoA* were determined by using primers within *lacZ* (5′-GGATGTGCTGCAAG GCGA) and Tn*phoA* (5′-GTAATATCGCCCTGAGC), respectively. DNA sequencing was done by the dideoxy-chain termination method (36) and the Sequenase sequencing kit

TABLE 1. Bacterial strains, plasmids, and phage

Strain, plasmid, or phage	Relevant genotype	Reference or source
Strains		
CC117	Δ*phoA20*	25
DO1489	*glnF*::Tn*10* glnG::Tn*5*	Fred Ausubel
EH3742	(F⁺::Tn*5*) *asnA*::Tn*10* *asnB*	Egon Hansen
SG20250	*cps*⁺ Δ*lac*	4; equivalent to MC4100
SG20436	*rcsB18*::ΔTn*10* *lon-100* *cpsB10*::*lac*-Mu *imm*$^{\lambda}$ (F⁺::Tn*5*)	SG20698 × EH3742
SG20698	*lon-100* *rcsB18*::ΔTn*10* *cpsB10*::*lac*-Mu *imm*$^{\lambda}$	4
SG20761	*rcsC137* *cpsB10*::*lac*-Mu *imm*$^{\lambda}$	4
VS20555	(*imm*^{21} *rcsB101*::*lacZ*)	SG20250 + SB72
VS20559	(*imm*^{21} *rcsB101*::*lacZ*) *glnF*::Tn*10*	VS20555 + P1 (DO1489)
Plasmids		
pJB100	*rcsB*⁺ *bla*⁺	4
pJB201	*rcsC*⁺ *bla*⁺	4
pATC591	*bla*⁺	pRS591 + *Hin*dIII linker
pRS591	*bla*⁺	38
Phages		
λRS45	*imm*^{21} Δ*lac*$_{SC}$	38
SB72	*imm*^{21} *rcsB101*::*lacZ*	λRS45 + *rcsB101*::*lacZ*

from U.S. Biochemical Corp. The University of Wisconsin Genetics Computer Group package was used for sequence analysis (9). The accession number for the sequence of RcsB and RcsC in GenBank is M28242.

Isolation of protein fusions. Fusions of *rcsB* to *lacZ* were generated by the cyclone deletion method of Dale et al. (6), using the Cyclone kit. M13mp19 containing *rcsB* on the 3.6-kb *Hin*dIII-*Bam*HI fragment of the *rcsB*⁺ plasmid pJB100 (4) was used as the template in the reaction. Smaller plaques that resulted in a Lac⁺ phenotype were purified, and the approximate size of the bacteriophage DNA was determined. Clones that appeared to have some of the *rcsB* insert deleted were screened for inability to complement the *rcsB* mutation in SG20436 for expression of *cps* genes. To obtain a fusion of *rcsB* with the full *lacZ* gene, instead of the alpha-complementing fragment found on M13mp19, several *rcsB lacZ*⁺ clones were transferred onto a plasmid that contains the full *lacZ* gene (pATC591), using *Hin*dIII and *Eco*RI restriction enzymes. One of the plasmid-borne fusions was crossed onto λRS45 (38), and lysogens of the resulting Lac⁺ phage SB72 were constructed in the appropriate strains. The last nucleotide in the junction of the *rcsB101*::*lacZ* fusion in SB72 is base pair (bp) 426 (Fig. 2). β-Galactosidase assays were performed as previously described (4).

To obtain alkaline phosphatase fusions to *rcsC*, the procedure of Manoil and Beckwith (25) was modified to screen for *rcsC* mutant isolates. An early-stationary-phase culture of CC117 containing the *rcsC*⁺ plasmid pJB201 (4) was infected with λTn*phoA* at a multiplicity of infection of 1 and divided among 10 tubes to ensure isolation of independent transpositions. The samples were grown and plated on LB plates containing 300 μg of kanamycin per ml, 125 μg of ampicillin per ml, and 40 μg of 5-bromo-4-chloro-3-indolyl phosphate (XP; Sigma Chemical Co.) per ml. Colonies that grew on these plates (of which about 1% were blue colonies) were scraped and pooled. Plasmid DNA was isolated and transformed into CC117, selecting on LB containing 30 μg of kanamycin per ml and XP (and screening for ampicillin resistance). Plasmid DNA was isolated from about 30 Kan^r Amp^r blue colonies and about 20 Kan^r Amp^r white and very light blue colonies. To screen these plasmids for loss of *rcsC*, SG20761 was used as a transformation recipient to check for complementation of the *rcsC137* mutation by the incoming plasmid. The plasmids that did not complement were subjected to restriction enzyme analysis to determine the location and orientation of the insertion. Kan^r Amp^r *rcsC* mutant

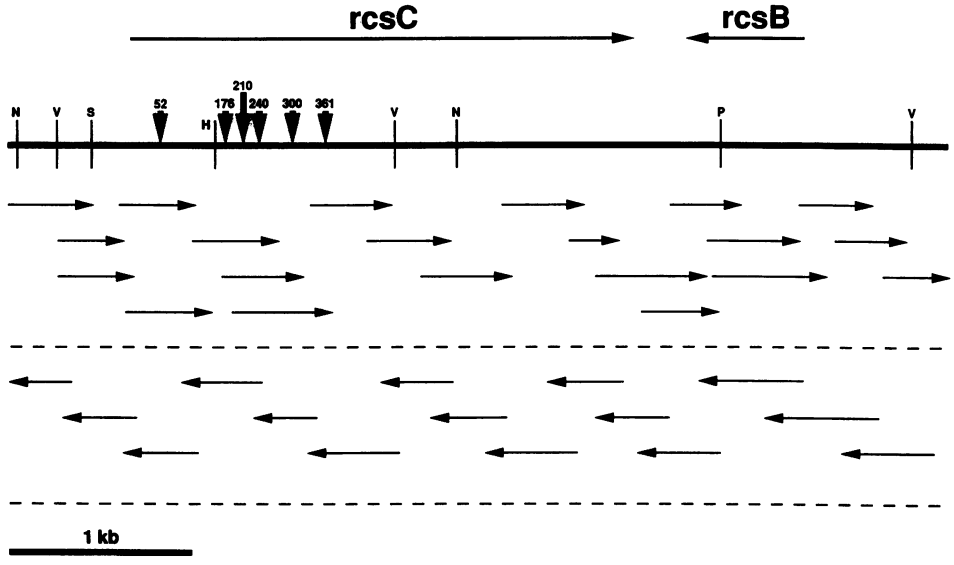

FIG. 1. Abbreviated restriction map and sequencing strategy of the *rcsB* and *rcsC* region. The letters on the bold line indicate restriction sites: N, *Nru*I; V, *Pvu*I; S, *Sph*I; H, *Hin*dIII; P, *Pst*I. The arrows on the bold line indicate alkaline phosphatase fusion junctions to RcsC. The numbers above the fusion junctions indicate the number of the last RcsC amino acid in the fusion. The sequencing strategy shown in the lower portion indicates the direction and length of each fragment that was sequenced.

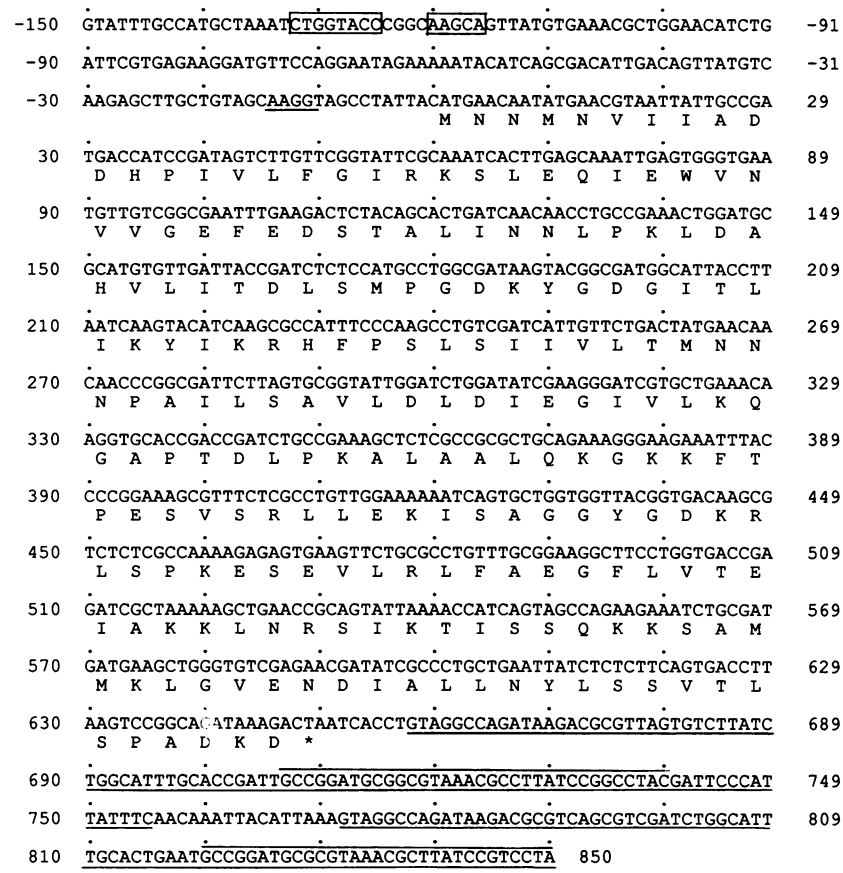

FIG. 2. Nucleotide sequence of *rcsB*. The predicted amino acid sequence of RcsB is indicated in one-letter code below the appropriate nucleotide bases. The putative ribosome-binding site is underlined at bp −14. The boxed areas are the RpoN-like promoter at bp −132. The underlined regions at the end of the ORF are the direct repeats, and the overlined nucleotides are the REP sequences.

plasmids were tested for the production and size of fusion proteins by Western blot (immunoblot) analysis (16), using anti-alkaline phosphatase antisera (from Phillip Bassford). Alkaline phosphatase assays were performed as described by Manoil and Beckwith (25).

RESULTS

Nucleotide sequences of *rcsB* and *rcsC*. Previously, we had defined the DNA region containing *rcsB* and *rcsC* by using transposable elements and complementation analysis (4). *rcsB* and *rcsC* lie adjacent to each other and are transcribed toward one another (Fig. 1). The sequence of this region is shown in Fig. 2 and 3. An open reading frame (ORF) of 651 bases, corresponding to the expected location and orientation of *rcsB*, is shown in Fig. 2. This ORF defines a protein with a predicted molecular mass of 23,656 Da, which is close to the observed 26,000-Da mass (4). This reading frame was confirmed by sequencing across the junctions of several *rcsB*::*lacZ* constructs that resulted in $RcsB^-$ $LacZ^+$ fusion proteins. The Lac^+ phenotype indicates that the reading frames of *rcsB* and *lacZ* must be identical. Two possible ATG translation start sites were identified from the sequence in the ORF corresponding to *rcsB*. The translation initiation codon was assigned to the ATG codon at bp 1 because its upstream region (AAGG) at −14 to −11 (underlined in Fig. 2) had a better match to the consensus ribosome-binding site (37) than did that for the ATG at bp 10.

While no σ^{70} consensus promoter sequence was found upstream of the *rcsB* ORF, a sequence with homology to the consensus RpoN-activated promoter was detected at bp −132 to −116 (boxed in Fig. 2 and compared with the consensus sequence in Fig. 4). RpoN is an alternative sigma factor found in gram-negative organisms that acts at the promoters of genes that regulate various metabolic functions and flagellum synthesis (17, 27, 34). The homology in the promoter of *rcsB* suggested that its expression might be dependent on the presence of RpoN. In fact, the expression of *rcsB* (measured by using *rcsB*::*lacZ* lysogens) was decreased more than 300-fold in mutants defective in RpoN (0.6 U of β-galactosidase specific activity versus 195 U for the $rpoN^+$ strain). Therefore, the promoter of *rcsB* is dependent on the alternate sigma factor, RpoN.

The distance between the translational stop codons of *rcsB* and *rcsC* is 196 bases (shown at the end of the sequences in Fig. 2, 3, and 5). The CTA at 848 to 850 in Fig. 2 corresponds to the inverse complement of the stop codon of *rcsC*, and the last three bases (TTA) of Fig. 3 correspond to the stop codon of *rcsB* in the complementary strand. Within this 196-bp region is a 98-bp direct repeat (underlined in Fig. 2; 658 to 755 and 772 to 866) separated by a 16-bp A+T-rich sequence and perfect 12-bp inverted repeats, GCCAGATAAGAC (Fig. 2; 662 to 673, 693 to 682, and 776 to 787). The direct repeat contains just 10 mismatches and three single-base-pair gaps. Within each 98-bp direct repeat is a 34-bp sequence (overlined in Fig. 2 [bp 707 to 740 and 820 to 851] and shown in Fig. 5) that is homologous to the

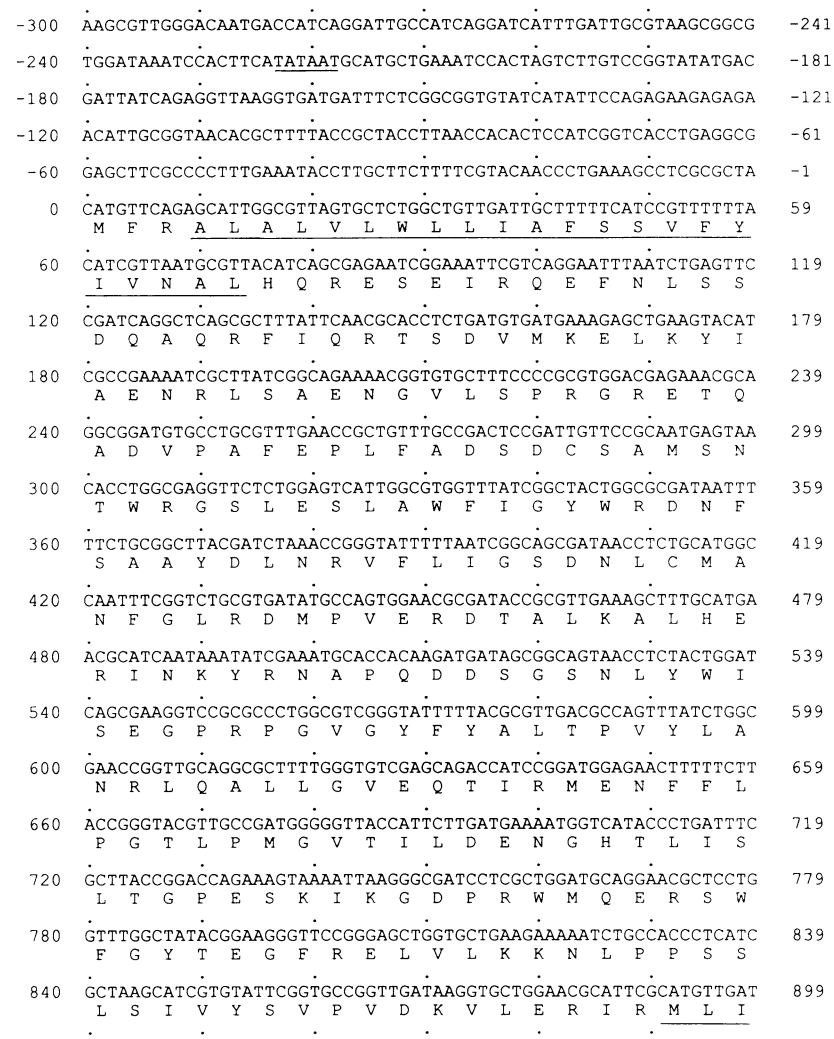

FIG. 3. Nucleotide sequence of *rcsC*. The predicted amino acid sequence of RcsC is indicated in one-letter code below the appropriate nucleotide bases. The underlined nucleotides are part of a possible promoter. The underlined amino acids are the hydrophobic stretches that could be transmembrane domains.

consensus repetitive extragenic palindromic (REP) sequence (Fig. 5) (39).

There is a 2,802-bp ORF on the complementary strand to *rcsB*, in the position expected for *rcsC* (Fig. 3). The predicted molecular mass of 104,456 Da agrees with the previously estimated size of 100,000 Da (4). As with *rcsB*, the reading frame of *rcsC* was confirmed by sequence analysis of fusion junctions between *rcsC* and Tn*phoA* that had resulted in RcsC$^-$ PhoA$^+$ fusion proteins (see below).

There is no sequence upstream of the *rcsC* ORF that resembles the RpoN-like promoter of *rcsB*. In addition, mutations in *rpoN* have no effect on *rcsC*::*phoA* expression (data not shown). A possible promoter, containing a −10 region that matches perfectly with the consensus (TATA AT), is located at −217 to −223 (underlined in Fig. 3), although no corresponding −35 region was detected. The promoter for *rcsC* is clearly regulated differently than the promoter for *rcsB*.

The deduced amino acid sequence of RcsC contains two hydrophobic regions that stretch from the Ala at position 4 to Leu-25 and from Met-298 to Ala-319 (underlined in Fig. 3 and horizontally striped in Fig. 5). Computer analysis (using PC Gene) showed that these hydrophobic regions can function as transmembrane domains, indicating that RcsC may act as an inner membrane protein.

Tn*phoA* fusion analysis of *rcsC*. To test whether RcsC functions in the membrane, protein fusions of RcsC to alkaline phosphatase were generated as described in Materials and Methods. Since alkaline phosphatase functions only when localized to the periplasm and Tn*phoA* is missing the signal sequence necessary for transport to the periplasm, the only way protein fusions can result in active alkaline phosphatase is if RcsC can provide a signal to transport the alkaline phosphatase across the inner membrane (25). Five different stable RcsC$^-$ PhoA$^+$ fusion proteins were mapped to the N terminus of RcsC, and the one stable RcsC$^-$ PhoA$^-$ fusion protein was mapped downstream from the active fusions. The exact junctions of the *rcsC*::*phoA* fusions were determined by sequencing through the junctions. The sizes of the fusion proteins (determined by Western blot analysis) agreed with the restriction mapping and sequencing data.

The numbers defining the six fusions shown in Fig. 1 and 5 indicate the number of the last amino acid of RcsC present in each fusion. The alkaline phosphatase production from

```
900  CCTTAACGCAATTTTGCTGAATGTGCTTGCCGGAGCTGCATTGTTTACTCTCGCACGGAT  959
       L  N  A  I  L  L  N  V  L  A  G  A  A  L  F  T  L  A  R  M
960  GTACGAGCGACGTATTTTCATTCCGGCGGAAAGCGACGCCCTGCGACTGGAAGAACATGA  1019
       Y  E  R  R  I  F  I  P  A  E  S  D  A  L  R  L  E  E  H  E
1020 GCAGTTCAATCGCAAGATTGTCGCCTCCGCGCCAGTGGGTATCTGCATTTTGCGTACCGC  1079
       Q  F  N  R  K  I  V  A  S  A  P  V  G  I  C  I  L  R  T  A
1080 TGATGGCGTCAATATTTTAAGTAACGAACTGGCGCATACCTATCTCAATATGCTTACGCA  1139
       D  G  V  N  I  L  S  N  E  L  A  H  T  Y  L  N  M  L  T  H
1140 TGAGGACCGCCAACGACTGACGCAAATTATCTGTGGGCAGCAGGTCAATTTTGTTGATGT  1199
       E  D  R  Q  R  L  T  Q  I  I  C  G  Q  Q  V  N  F  V  D  V
1200 CCTGACCAGCAACAATACCAATCTGCAAATCAGCTTCGTCCATTCGCGCTATCGTAATGA  1259
       L  T  S  N  N  T  N  L  Q  I  S  F  V  H  S  R  Y  R  N  E
1260 AAACGTGGCCATTTGTGTGCTGGTGGATGTTTCTTCGCGCGTGAAGATGGAAGAGTCGTT  1319
       N  V  A  I  C  V  L  V  D  V  S  S  R  V  K  M  E  E  S  L
1320 GCAGGAGATGGCACAAGCAGCGGAACAGGCGAGCCAGTCAAAATCGATGTTCCTTGCCAC  1379
       Q  E  M  A  Q  A  A  E  Q  A  S  Q  S  K  S  M  F  L  A  T
1380 CGTCAGTCATGAGCTGCGAACGCCGCTGTATGGCATTATCGGTAACCTGGATCTGTTGCA  1439
       V  S  H  E  L  R  T  P  L  Y  G  I  I  G  N  L  D  L  L  Q
1440 AACCAAAGAGTTACCGAAAGGCGTCGATCGGCTGGTGACGGCAATGAACAACTCTTCCAG  1499
       T  K  E  L  P  K  G  V  D  R  L  V  T  A  M  N  N  S  S  S
1500 CCTGTTGTTGAAAATTATCAGCGATATTCTCGATTTCTCGAAGATTGAATCGGAACAGTT  1559
       L  L  L  K  I  I  S  D  I  L  D  F  S  K  I  E  S  E  Q  L
1560 GAAGATCGAACCGCGTGAGTTTTCACCGCGTGAAGTGATGAACCACATCACCGCCAACTA  1619
       K  I  E  P  R  E  F  S  P  R  E  V  M  N  H  I  T  A  N  Y
1620 TTTACCGCTGGTGGTACGCAAGCAGTTAGGCTTGTACTGCTTTATTGAACCGGATGTGCC  1679
       L  P  L  V  V  R  K  Q  L  G  L  Y  C  F  I  E  P  D  V  P
1680 AGTGGCCTTAAATGGCGACCCGATGCGTTTACAGCAGGTCATCTCCAACCTGTTGAGTAA  1739
       V  A  L  N  G  D  P  M  R  L  Q  Q  V  I  S  N  L  L  S  N
1740 CGCCATAAAATTCACCGATACCGGCTGTATAGTTTTGCATGTTCGCGCGGATGGCGATTA  1799
       A  I  K  F  T  D  T  G  C  I  V  L  H  V  R  A  D  G  D  Y
1800 TCTCTCTATCCGTGTTCGCGATACCGGCGTGGGGATACCGGCGAAAGAAGTGGTGCGCTT  1859
       L  S  I  R  V  R  D  T  G  V  G  I  P  A  K  E  V  V  R  L
1860 GTTTGATCCCTTCTTCCAGGTCGGAACGGGCGTACAGCGTAATTTCCAGGGGACCGGTCT  1919
       F  D  P  F  F  Q  V  G  T  G  V  Q  R  N  F  Q  G  T  G  L
1920 GGGTCTGGCGATTTGTGAAAAACTGATCAGCATGATGGACGGCGATATCTCGGTAGATTC  1979
       G  L  A  I  C  E  K  L  I  S  M  M  D  G  D  I  S  V  D  S
1980 AGAACCGGGAATGGGCAGCCAGTTTACCGTGCGTATTCCGTTGTACGGCGCTCAGTACCC  2039
       E  P  G  M  G  S  Q  F  T  V  R  I  P  L  Y  G  A  Q  Y  P
2040 GCAGAAAAAAGGCGTGGAAGGGTTGAGTGGTAAACGCTGCTGGCTGGCGGTCCGCAATGC  2099
```

FIG. 3—*Continued*

the fusion proteins and their predicted locations in the protein are shown in Fig. 6. The levels of alkaline phosphatase production from the five PhoA$^+$ fusions varied from 27 to 46 U, with the higher levels found in the fusions closer to the N terminus (Fig. 6). The four PhoA$^+$ fusions closest to the N terminus of RcsC (at amino acids 52, 176, 210, and 240) mapped between the two hydrophobic, membrane-spanning regions; the PhoA$^+$ fusion that produced the lowest amount of alkaline phosphatase (amino acid 300) mapped just two amino acids beyond the start of the second hydrophobic region; and the only PhoA$^-$ fusion (amino acid 361) mapped between the second hydrophobic region and the C terminus of RcsC (Fig. 5 and 6). The PhoA$^-$ fusion produced just 2 U of alkaline phosphatase (Fig. 6). These results suggest that the N-terminal region of RcsC is located in the periplasm and that the C-terminal region is located in the cytoplasm and are consistent with the model of membrane topology for RcsC shown in Fig. 6.

Homology of RcsB and RcsC to two-component regulators. By analogy with the nitrogen regulatory circuit, we have suggested that RcsB and RcsC may act as a two-component regulatory circuit with homology to other two-component regulators (4). These circuits have been defined as (i) a class of environmental sensors that share homology and act on (ii) a class of effectors (or regulators) that share homology (reviewed in reference 35). The predicted amino acid sequences for RcsB and RcsC were used to search the National Biomedical Research Foundation data base with the FASTP program (9).

RcsC was found to share homology in the C-terminal region with the conserved regions of the sensors CpxA (1), PhoR (23), EnvZ (35), and PhoM (2). Other members of this sensor group include NtrB (35), VirA (35), DctB (33), FixL (7), UhpB (11), CheA (29), NarX (40), and DegS (13). All of these sensors share homology in their C-terminal regions, including a conserved histidine (amino acid 139 of RcsC) and conserved GXGLGLAI (amino acids 637 to 644 of RcsC). The region of homology is shown schematically in Fig. 5, and the amino acid sequence of RcsC is compared with sequences of four other sensors in Fig. 7. RcsC, however, is unusual in that the final 125 amino acids share homology with the N termini of the effectors as shown in Fig. 7 (and the vertically striped region of RcsC in Fig. 5).

From the searches using RcsB, it was found that the N-terminal half of RcsB shared homology with the effectors CheY (29), Spo0A (10), OmpR (35), and NtrC (35). Other

```
        Q  K  K  G  V  E  G  L  S  G  K  R  C  W  L  A  V  R  N  A
2100 GTCGCTCTGTCAGTTCCTGGAAACCAGTTTGCAGCGCAGCGGCATCGTCGTTACAACATA 2159
      S  L  C  Q  F  L  E  T  S  L  Q  R  S  G  I  V  V  T  T  Y
2160 CGAAGGGCAGGAACCGACTCCCGAAGATGTGTTGATCACTGACGAGGTAGTGAGTAAAAA 2219
      E  G  Q  E  P  T  P  E  D  V  L  I  T  D  E  V  V  S  K  K
2220 ATGGCAGGGCAGAGCGGTAGTGACCTTCTGTCGTCGCCATATTGGTATTCCGCTGGAGAA 2279
      W  Q  G  R  A  V  V  T  F  C  R  R  H  I  G  I  P  L  E  K
2280 AGCGCCAGGGGAGTGGGTACACAGTGTGGCTGCTCCGCATGAGCTACCGGCATTGTTGGC 2339
      A  P  G  E  W  V  H  S  V  A  A  P  H  E  L  P  A  L  L  A
2340 GCGTATTTATTTGATCGAGATGGAGAGCGACGATCCTGCTAACGCTCTGCCGTCAACGGA 2399
      R  I  Y  L  I  E  M  E  S  D  D  P  A  N  A  L  P  S  T  D
2400 CAAAGCGGTCAGCGATAATGACGATATGATGATTCTGGTCGTGGATGATCATCCGATTAA 2459
      K  A  V  S  D  N  D  D  M  M  I  L  V  V  D  D  H  P  I  N
2460 CCGGCGTTTGCTGGCAGATCAGTTGGGATCGTTGGGCTATCAATGTAAAACCGCGAATGA 2519
      R  R  L  L  A  D  Q  L  G  S  L  G  Y  Q  C  K  T  A  N  D
2520 TGGCGTCGATGCGCTTAATGTACTTAGCAAGAATCATATTGATATCGTGCTTAGCGACGT 2579
      G  V  D  A  L  N  V  L  S  K  N  H  I  D  I  V  L  S  D  V
2580 CAACATGCCAAATATGGATGGTTACCGCTTGACGCAACGCATTCGTCAGTTGGGACTGAC 2639
      N  M  P  N  M  D  G  Y  R  L  T  Q  R  I  R  Q  L  G  L  T
2640 GTTGCCGGTAATCGGAGTAACTGCTAATGCGTTGGCTGAAGAGAAGCAGCGGTGTCTGGA 2699
      L  P  V  I  G  V  T  A  N  A  L  A  E  E  K  Q  R  C  L  E
2700 GTCCGGTATGGACAGCTGCCTGTCGAAGCCGGTAACGCTGGATGTGATAAAACAGAGCCT 2759
      S  G  M  D  S  C  L  S  K  P  V  T  L  D  V  I  K  Q  S  L
2760 GACGTTATATGCCGAGAGGGTCAGGAAATCGCGGGATTCGTAGGACGGATAAGCGTTTAC 2819
      T  L  Y  A  E  R  V  R  K  S  R  D  S  *
2820 GCGCATCCGGCATTCAGTGCAAATGCCAGATCGACGCTGACGCGTCTTATCTGGCCTACT 2879
2880 TTAATGTAATTTGTTGAAATAATGGGAATCGTAGGCCGGATAAGGCGTTTACGCCGCATC 2939
2940 CGGCAATCGGTGCAAATGCCAGATAAGACACTAACGCGTCTTATCTGGCCTACAGGTGAT 2999
3000 TA 3001
```

FIG. 3—*Continued*

effectors include PhoB (35), VirG (35), SfrA (35), DctD (33), FixJ (7), UhpA (11), CheB (29), Spo0F (42), AlgR (8), DegU (13), and NarL (40). All of the effectors share homology in their first 110 amino acids (shaded region of RcsB in Fig. 5; Fig. 7), including two conserved aspartic acid residues (amino acids 11 and 56 of RcsB). By using the first ATG codon in the ORF of *rcsB* (at bp 1 in Fig. 2) as discussed above, RcsB can be better aligned with the other regulators at the N terminus. The effectors can be subdivided into several groups on the basis of their C-terminal domain homology (8a). OmpR, SfrA, PhoB, and VirG are one such group. NtrC and DctD are another group. Spo0A and AlgR are a third group. The fourth group comprises FixJ, NarL, DegU, and UhpA. RcsB shares homology in its C terminus with this fourth group (Fig. 7).

DISCUSSION

Transcriptional regulation of the *cps* genes required for capsule synthesis in *E. coli* is dependent on at least three regulators, RcsA, RcsB, and RcsC. RcsA is an unstable positive regulator that is degraded by Lon protease (41). In a previous paper (4), we suggested that RcsC may function as a sensor because mutations in *rcsC* had effects on capsule synthesis reminiscent of the effects on nitrogen regulation of

```
RpoN Promoter Consensus   C T G G C A C N N N N C T T G C A
                          : : : :  : :      :    : : :
Putative RcsB Promoter    C T G G T A C C C G G C A A G C A
```

FIG. 4. Sequence comparison of RpoN promoter consensus with *rcsB* promoter. N indicates any nucleotide.

mutations in *glnL*, the sensor protein in the nitrogen regulation circuit in *E. coli* (22). For example, while mutations can be isolated in either sensor that result in constitutive expression of the regulated genes, null mutations in either gene have relatively little effect on expression of the regulated genes. In both cases, an associated positive regulatory gene maps nearby: *glnG* for the nitrogen assimilatory system and *rcsB* for capsule synthesis.

We show here that RcsC and RcsB belong to a family of two-component regulatory pairs that respond to environmental stimuli in a variety of bacterial organisms (reviewed in reference 35). These pairs are composed of a sensor protein and a regulator (or effector) protein. A general model has been proposed, based on the biochemical results in the nitrogen and chemotaxis regulation systems (14, 30, 46). In this model, the N terminus of the sensor receives a signal from the environment and transmits this signal to the conserved C terminus of the protein. The C terminus first undergoes autophosphorylation at a conserved histidine residue (45) and subsequently acts as a kinase to phosphorylate the conserved N terminus of the effector protein. The phosphorylated effector protein then effects a response, either by binding DNA to mediate transcription of one or more genes or by turning on the flagellum motor.

All of the sensor proteins in these regulatory pairs except for NtrB have two hydrophobic transmembrane domains and therefore have been predicted to be membrane proteins, although only the VirA, EnvZ, and CpxA sensor proteins have been physically localized to the inner membrane (19, 28, 44). From protein sequence analysis, it has been suggested that the N-terminal domain of the sensor is located in

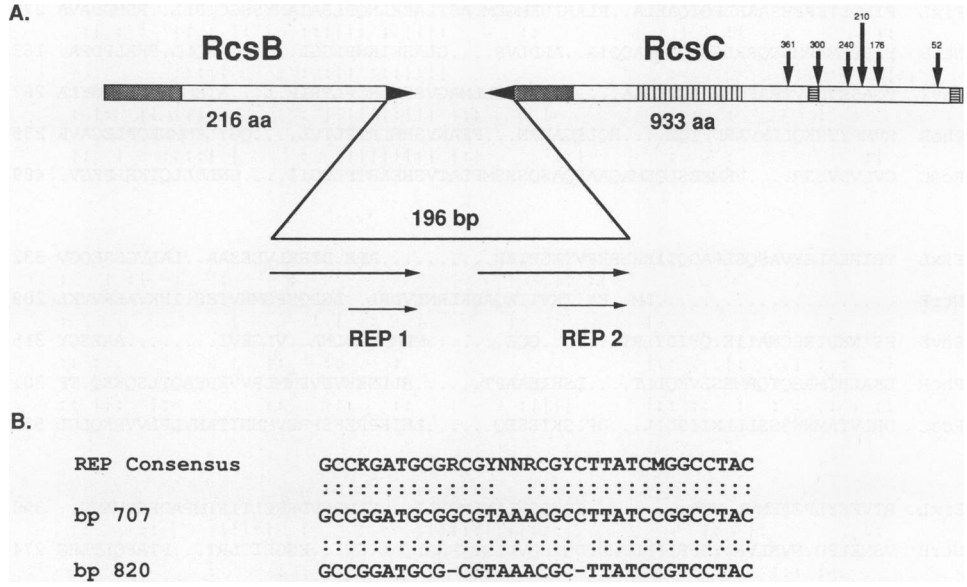

FIG. 5. (A) Orientation of the RcsB and RcsC region. The shaded regions of RcsB and RcsC indicate the domains that share homology with other effector proteins. Symbols: ▥, domain of RcsC that shares homology with the sensor proteins; ▤, predicted transmembrane domains of RcsC. Alkaline phosphatase fusion junctions are indicated above the RcsC protein, with their numbers as in Fig. 1. The longer arrows below the proteins indicate the 98-bp direct repeats (underlined in Fig. 2). (B) REP1 and REP2, compared with the REP consensus. Ambiguous bases in the consensus: K, G or T; R, G or A; Y, C or T; M, C or A; N, any base.

the periplasm and the C-terminal domain is located in the cytoplasm (35). This predicted topology has been confirmed for CpxA by using alkaline phosphatase fusions (44).

The sequence prediction of two transmembrane domains in RcsC is consistent with the membrane topology shown in Fig. 6 and is similar to that of the other sensors (35). The $PhoA^+$ and $PhoA^-$ fusions of RcsC to alkaline phosphatase strongly suggest this proposed topology, in which the N-terminal 298-amino-acid domain of RcsC is periplasmic and a C-terminal 614-amino-acid domain is cytoplasmic.

The predicted primary sequence of RcsC has homology to the other sensors within the 250-amino-acid stretch in the C-terminal cytoplasmic domain (2, 7, 11, 13, 29, 33, 35, 40). This domain, which is homologous among all of the sensors, is predicted by the model to transmit information via phosphorylation to the homologous N terminus of the regulators. RcsC differs from most of the other known sensors in having

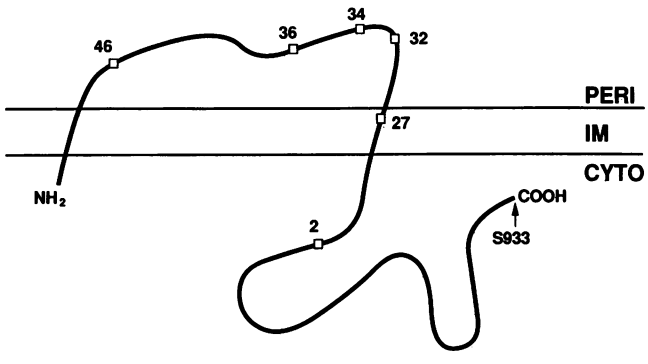

FIG. 6. Predicted membrane topology of RcsC. Small boxes within protein indicate points of alkaline phosphatase fusions to RcsC and their corresponding levels of alkaline phosphatase (in units of activity). PERI, IM, and CYTO indicate periplasm, inner membrane, and cytoplasm, respectively. S933 refers to the carboxy-terminal amino acid, serine.

an extra 260 amino acids. The 125 amino acids at the C terminus of RcsC share limited homology to the conserved N-terminal domain of the effectors. VirG is the only other sensor that contains a similar C-terminal domain with limited homology to the conserved domain of the effectors (8a). This C-terminal domain of RcsC could be acting as a titrating acceptor, thereby modulating the phosphorylation action of RcsC on RcsB.

The primary sequence of RcsB demonstrated homology to all of the other effectors in the conserved N terminus. RcsB also has more extended homology in its C-terminal domain to the subfamily of regulators containing FixJ, DegU, NarL, and UhpA. These subfamilies may relate to both the function of these proteins and their interactions with sigma factors. For example, the regulators in the subfamily consisting of NtrC and DctD are involved in cellular response to low external levels of nitrogen and act with RpoN (σ^{54}) to activate their target genes (34). No common factor has been defined among the members of the subfamily that includes RcsB, although RcsB and UhpA are both involved with sugar transport in an indirect or direct manner. UhpA is a positive regulator for a gene (*uhpT*) required for uptake of hexose phosphates in *E. coli* (11), and RcsB is a positive regulator of the capsule genes that are necessary for synthesis of a surface polysaccharide of *E. coli*. It is interesting to note that the two effectors involved in activating capsule synthesis (RcsB in *E. coli* and AlgR in *Pseudomonas aeruginosa* [8]) are members of different subfamilies of effectors.

Sequence analysis of the *rcsB-rcsC* region provided evidence for additional control of this regulatory circuit in two ways. Synthesis of *rcsB* is dependent on the alternate sigma factor, RpoN (σ^{54}). RpoN is the sigma factor that is required for transcription of physiologically diverse genes (reviewed in reference 18). Transcription inititation of σ^{54}-dependent genes appears to require an activator protein in addition to the sigma factor (18). Transcription of the positive regulator, *rcsB*, is reduced more than 300-fold in an *rpoN* mutant,

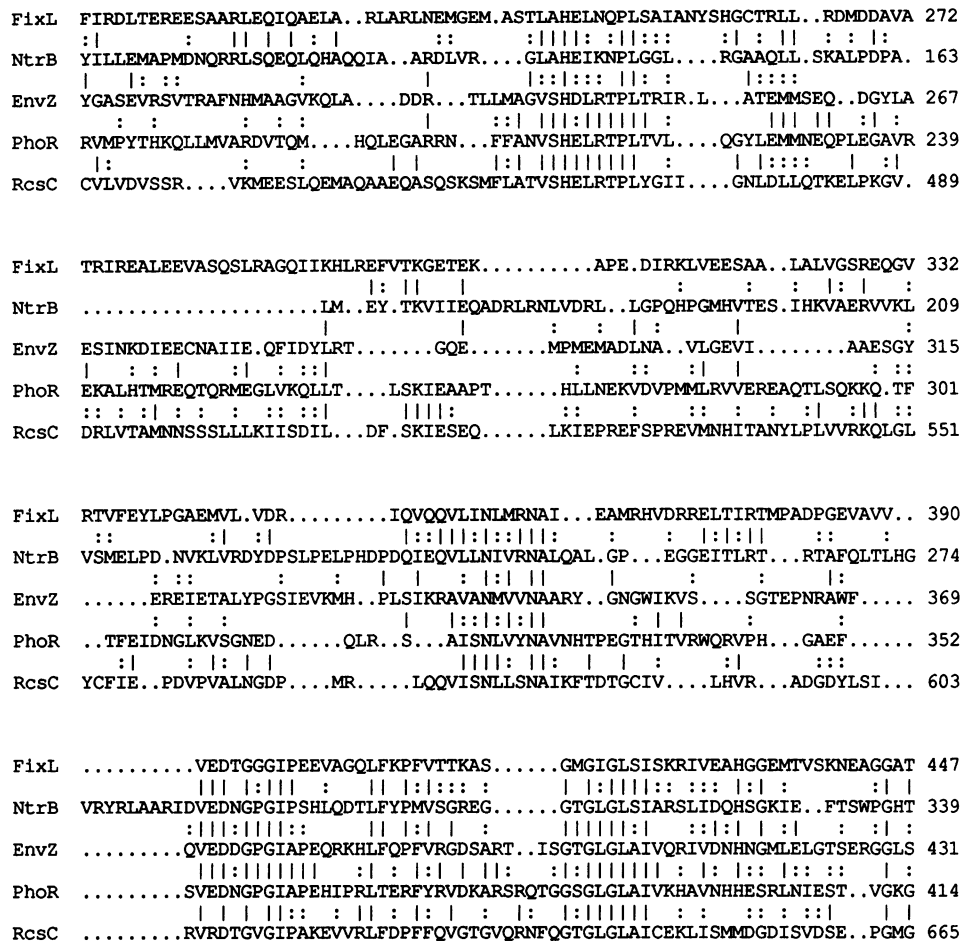

FIG. 7. Sequence alignments of RcsC with other sensors (*Klebsiella pneumoniae* NtrB, *Rhizobium meliloti* FixL, and *E. coli* PhoR and EnvZ) and alignments of RcsB and RcsC with other effectors (*K. pneumoniae* NtrC, *R. meliloti* FixJ, and *E. coli* UhpA). Residues that are identical (|) and residues that belong to the same family (:) are indicated. The families of amino acids are ILMVFYW, AGPST, QNED, and HKR (32).

although limiting nitrogen conditions do not alter synthesis (V. Stout and S. Gottesman, manuscript in preparation). If activation of *rcsB* occurs in a manner similar to activation of the other RpoN-utilizing genes, then another factor may be required for activation. This factor cannot be either RcsA or RcsC, since they do not affect the synthesis of RcsB as tested with *rcsB*::*lacZ* fusions (Stout and Gottesman, in preparation).

The promoter for *rcsC* has not yet been defined, but it is apparently not RpoN dependent. The relatively low levels of alkaline phosphatase produced by the *rcsC*::*phoA* fusions support the sequence analysis of *rcsC*, in which only poor promoters were found.

The direction of transcription of *rcsB* and *rcsC* relative to each other differs from that of other sensor-regulator pairs and may provide additional control of these two regulators. All of the other pairs that map close together are transcribed in the same direction, often as members of the same operon (for example, DctB-DctD and OmpR-EnvZ) (28, 33), whereas RcsB and RcsC are made in a converging fashion, with a 196-bp spacer between the two stop codons. Within this spacer region are 98-bp direct repeats, each of which contain a REP sequence.

REP sequences are palindromes that are reiterated up to 1,000 times in *E. coli* chromosomal DNA (reviewed in reference 15). It has been proposed that these sequences may be involved in stabilizing mRNA that ends with a REP sequence. Since the messages for both *rcsB* and *rcsC* end in the region where the REP sequences lie, one or both may be affected by this sequence. It has also been suggested that DNA gyrase binds to REP sequences and makes cuts near them (47). It is possible that gyrase is required to relieve the strain that would result from converging unwinding DNA templates. It has been shown that positive supercoils are generated in front of advancing DNA polymerase (20) and suggested that this effect can be more dramatic when the gene product is a membrane protein (21).

The unique orientation of *rcsB* and *rcsC* may be a reflection of a requirement for independent control of the levels of synthesis of the two genes. The latter suggestion is supported by the different promoters of the two genes.

The control of capsule synthesis appears to involve an intricate and complex network that includes RcsC and RcsB. If RcsC is acting as a sensor, we do not know what RcsC is sensing in the environment. However, we have suggested that temperature may be important, since more capsule is made at lower temperatures than at 37°C and *rcsC137* mutants show temperature-independent capsule synthesis (T. Klopotowski and S. Gottesman, unpublished results). It is clear that syntheses of *rcsC* and *rcsB* are not regulated by

```
FixL  FRFTLPAYLDERIVAND*                                                                  464
NtrB  EFSVYLPIRK*                                                                         349
EnvZ  IRAWLPVPVTRGTTKEG*                                                                  450
PhoR  TRFSFVIPERIAKNSD*                                                                   431
RcsC  SQFTVRIPLYGAQYPQKKGVEGLSGKRC // RIYLIEMESDDPANALPSTDKAVSDNDDMMILVVDDHPINRR         822
NtrC                                                    MQRGIAWIVDDDSSIRW                 17
FixJ                                                    MTDYTVHIVDDEEPVRK                 17
UhpA                                                      MITVALIDDHLIVRS                 15
RcsB                                                    MNNMNVIIADDHPIVLF                 17

RcsC  LLADQLGSLGYQCKTA..NDGVDALNVLSKNHIDIVLSDVNMPN...MDGYRLTQRIRQLGLTLPVIGVTANAL 891
NtrC  VLERALTGAGL..SCTTFESGNEVLDALTTKTPDVLLSDIRMPG...MDGLALLKQIKQRHPMLPVIIMTAHSD  86
FixJ  SLAFMLTMNGF..AVKMHQSAEAFLAFAPDVRNGVLVTDLRMPDMSGVELLRNLGDLKINIPSI.........  79
UhpA  GFAQLLGLEPDLQVVAEFGSGREALAGLPGRGVQVCICDISMP......DISGLELLSQLPKGMATIMLSVHDS  83
RcsB  GIRKSLEQIEWVNVVGEFEDSTALINNLPKLDAHVLITDLSMPGDKYGDGITLIKYIKRHFPSLSIIVLTMNNN  91

RcsC  AEEKQRCLESGMDSCLSKPVTLDVIKQSLTLYAERVRKSRDS*                                   933
NtrC  ...LDAAVSAYQQGAFDYLPKPFDIDEAVALVDRAISHYQEQQQPRNAPINSPTAD.....IIGERPAMQDV   150
FixJ  ..VITGHGDVPMAVEAMKAGAVDFIEKPFEDTVIIEAIERASEHLVVAEADVDDANDIRARLQTLSERERQV   149
UhpA  PALVEQALNAGARGFLSKRCSPDELIAAVHTVATGGCYLTPDIAIKLASGRQDP.....LTKRERQV   145
RcsB  PAILSAVLDLDIEGIVLKQGAPTDLPKALAALQKGKKFTPESVSRLLEKISAGGYGDKR.....LSPKESEV   158

NtrC  FRIIGRLSRSSISVLINGESGTGKELVAHALHRHSPRAKAPFIALNMAAIPKDLIESELFGHEKGA...      216
FixJ  LSAVV.....AGLPNKSIAYDLDISPRTVEVHRANVMAKMKAKSLPHLVRMALAGGFGPS*              204
UhpA  AEKLA.....QGMAVKEIAAELGLSPKTVHVHRANLMEKLGVSNDVELARRMFDGW*                  196
RcsB  LRLFA.....EGFLVTEIAKKLNRSIKTISSQKKSAMMKLGVENDIALLNYLSSVTLSPADKD*          217
```

FIG. 7—*Continued*

temperature when tested with *rcsC*::*phoA* and *rcsB*::*lacZ* fusions (data not shown).

Because insertion mutants of *rcsC* still show a dependence on RcsA and RcsB for capsule synthesis and there is still a temperature effect on capsule synthesis (4), it is possible that RcsC is a redundant component of this regulatory circuit. This is very similar to the nitrogen-regulatory circuit in which null mutations of the sensor, *ntrB*, result in a phenotype very similar to wild type (3). An explanation for these results has been suggested by the work of Ninfa et al. (31), in which they were able to demonstrate cross-talk between the nitrogen assimilation and chemotaxis systems in vitro. The sensor of one system was shown to phosphorylate the effector of another system.

The homologies of RcsC and RcsB with members of the sensor-effector family of regulators and the genetics of the capsule regulatory network suggest the model shown in Fig. 8. In this model, RcsC receives a signal that could be a shift to a lower temperature and transmits that signal to RcsB, perhaps through phosphorylation (although no direct interaction between RcsC and RcsB has been demonstrated yet). Activated RcsB could then either directly or indirectly stimulate expression of the *cps* genes, to make colanic acid polysaccharide. Mutations like *rcsC137*, which increase capsule synthesis, originally were used to define RcsC as a negative regulator. The results presented here and summarized in Fig. 8 suggest instead that RcsC regulates RcsB

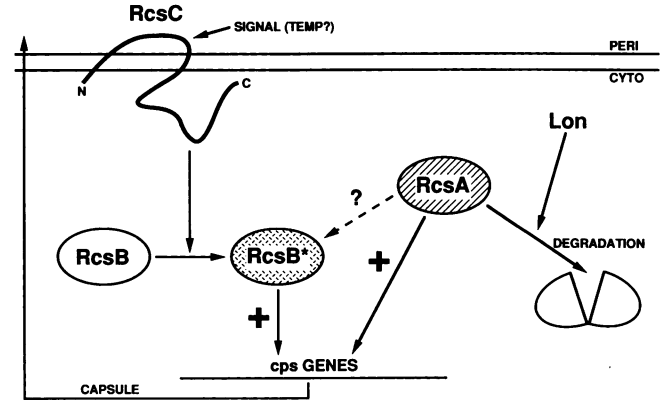

FIG. 8. Model of capsule regulation in *E. coli*. Periplasm and cytoplasm are indicated by PERI and CYTO, respectively.

activity both positively and negatively. The *rcsC137* mutation may be locked in a state that leads to constitutive activation of RcsB. RcsA is rapidly degraded by the Lon protease and is normally limiting in cells (41). When Lon is not present, enough RcsA accumulates to stimulate *cps* gene expression either directly or indirectly by functioning in the RcsB-RcsC interaction to allow RcsB to stimulate capsule expression. None of the three regulators appear to affect the synthesis of themselves or the other regulators (Stout and Gottesman, in preparation), indicating that any interactions among the regulators probably occur at the protein level.

ACKNOWLEDGMENTS

We thank Nancy Trun, Janine Trempy, David Gutnick, and Michael Maurizi for helpful discussions and Phillip Bassford for the alkaline phosphatase antisera. We thank Vojo Deretic for providing unpublished information and the suggestion that RcsC has homology with the effectors as well as the sensors. Tn*phoA* primer sequence was provided by Dana Boyd.

V.S. was supported by Public Health Service grant GM12497 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. **Albin, R., R. Weber, and P. M. Silverman.** 1986. The Cpx proteins of *Escherichia coli* K12. J. Biol. Chem. **261**:4698–4705.
2. **Amemura, M., K. Makino, H. Shinagawa, and A. Nakata.** 1986. Nucleotide sequence of the *phoM* region of *Escherchia coli*: four open reading frames may constitute an operon. J. Bacteriol. **168**:294–302.
3. **Backman, K. C., Y.-M. Chen, S. Ueno-Nishio, and B. Magasanik.** 1983. The product of *glnL* is not essential for regulation of bacterial nitrogen assimilation. J. Bacteriol. **154**:516–519.
4. **Brill, J. A., C. Quinlan-Walshe, and S. Gottesman.** 1988. Fine-structure mapping and identification of two regulators of capsule synthesis in *Escherichia coli* K-12. J. Bacteriol. **170**:2599–2611.
5. **Comeau, D. E., K. Ikenaka, K. Tsung, and M. Inouye.** 1985. Primary characterization of the protein products of the *Escherichia coli ompB* locus: structure and regulation of synthesis of the OmpR and EnvZ proteins. J. Bacteriol. **164**:578–584.
6. **Dale, R. M. K., B. A. McClure, and J. P. Houchins.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing mitochondrial 18s rDNA. Plasmid **13**: 31–40.
7. **David, M., M.-L. Daveran, J. Batut, A. Dedieu, O. Domergue, J. Ghai, C. Hertig, P. Boistard, and D. Kahn.** 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell **54**:671–683.
8. **Deretic, V., R. Dikshit, W. M. Konyecsni, and A. M. Chakrabarty.** 1989. The *algR* gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. J. Bacteriol. **171**:1278–1283.

8a. **Deretic, V., W. M. Konyecsni, C. D. Mohr, D. W. Martin, and N. S. Hibler.** 1989. Common denoninators of promoter control in *Pseudomonas* and other bacteria. Biotechnology **7**:1249–1254.

9. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
10. **Ferrari, F. A., K. Trach, D. LeCoq, J. Spence, E. Ferrari, and J. Hoch.** 1985. Characterization of the *spo0A* locus and its deduced product. Proc. Natl. Acad. Sci. USA **82**:2647–2651.
11. **Friedrich, M. J., and R. J. Kadner.** 1987. Nucleotide sequence of the *uhp* region of *Escherichia coli*. J. Bacteriol. **169**:3556–3563.
12. **Gottesman, S., P. Trisler, and A. S. Torres-Cabassa.** 1985. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12: characterization of three regulatory genes. J. Bacteriol. **162**:1111–1119.
13. **Henner, D. J., M. Yang, and E. Ferrari.** 1988. Localization of *Bacillus subtilus sacU* (*hy*) mutations to two linked genes with similarities to the conserved procaryotic family of two-component signaling systems. J. Bacteriol. **170**:5102–5109.
14. **Hess, T. F., N. Oosawa, N. Kaplan, and M. I. Simon.** 1988. Phosphorylation of three proteins in signalling pathway of bacterial chemotaxis. Cell **53**:79–87.
15. **Higgins, C. F., R. S. McClaren, and S. F. Newbury.** 1988. Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion?—a review. Gene **72**:3–14.
16. **Katayama, Y., S. Gottesman, J. Pumphrey, S. Rudikoff, W. P. Clark, and M. Maurizi.** 1988. The two-component, ATP-dependent *clp* protease of *Escherichia coli*. J. Biol. Chem. **263**: 15226–15236.
17. **Kohler, M., S. Harayama, J. L. Ramos, and K. N. Timmis.** 1989. Involvement of *Pseudomonas putida* RpoN sigma factor in regulation of various metabolic functions. J. Bacteriol. **171**: 4326–4333.
18. **Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of sigma-54 (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53**: 367–376.
19. **Leroux, B., M. F. Yanofsky, S. C. Winans, J. E. Ward, S. F. Ziegler, and E. W. Nester.** 1987. Characterization of the *virA* locus of *Agrobacterium tumefaciens*: a transcriptional regulator and host range determinant. EMBO J. **6**:849–856.
20. **Liu, L. F., and J. C. Wang.** 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA **84**:7024–7027.
21. **Lodge, J. K., T. Kazic, and D. E. Berg.** 1989. Formation of supercoiling domains in plasmid pBR322. J. Bacteriol. **171**: 2181–2187.
22. **Magasanik, B., and F. C. Neidhardt.** 1987. Regulation of carbon and nitrogen utilization, p. 1318–1325. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
23. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *phoR* gene, a regulatory gene for the phosphate regulon of *Escherichia coli*. J. Mol. Biol. **192**:549–556.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82**: 8129–8133.
26. **Markovitz, A.** 1977. Genetics and regulation of bacterial capsular polysaccharide synthesis and radiation sensitivity, p. 415–462. *In* I. Sutherland (ed.), Surface carbohydrates of the prokaryotic cell. Academic Press, Inc. (London), Ltd., London.
27. **Minnich, S. A., and A. Newton.** 1987. Promoter mapping and cell cycle regulation of flagellin gene transcription in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **84**:1142–1146.
28. **Mizuno, T., E. T. Wurtzel, and M. Inouye.** 1982. Osmoregulation of gene expression. II. DNA sequence of the *envZ* gene of the *ompB* operon of *Escherichia coli* and characterization of its product. J. Biol. Chem. **257**:13692–13698.
29. **Mutoh, N., and M. I. Simon.** 1986. Nucleotide sequence corresponding to five chemotaxis genes in *Escherichia coli*. J. Bacteriol. **165**:161–166.
30. **Ninfa, A. J., and B. Magasanik.** 1986. Covalent modification of the *glnG* product, NRI, by the *glnL* product, NRII, regulates the transcription of the *glnALG* operon in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:5909–5913.
31. **Ninfa, A. J., E. G. Ninfa, A. N. Lupas, A. Stock, B. Magasanik, and J. Stock.** 1988. Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proc. Natl. Acad. Sci. USA **85**:5492–5496.
32. **Nixon, B. T., C. W. Ronson, and F. M. Ausubel.** 1986. Two-component regulatory systems responsive to environmental

stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. Proc. Natl. Acad. Sci. USA **83**:7850–7854.

33. **Ronson, C. W., P. M. Astwood, B. T. Nixon, and F. M. Ausubel.** 1987. Deduced products of C4 dicarboxylic acid transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products. Nucleic Acids Res. **15**: 7921–7934.
34. **Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel.** 1987. *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J. Bacteriol. **169**:2424–2431.
35. **Ronson, C. W., B. T. Nixon, and F. M. Ausubel.** 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell **49**:579–581.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
37. **Shine, J., and L. Dalgarno.** 1974. The 3'-terminal sequence of *Escherichia coli* 16s ribosome RNA; complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
38. **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.
39. **Stern, M. J., G. F.-L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37**:1015–1026.
40. **Stewart, V., J. Parales, and S. M. Merkel.** 1989. Structure of genes *narL* and *narX* of the *nar* (nitrate reductase) locus in *Escherichia coli* K-12. J. Bacteriol. **171**:2229–2234.
41. **Torres-Cabassa, A. S., and S. Gottesman.** 1987. Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. J. Bacteriol. **169**:981–989.
42. **Trach, K., J. W. Chapman, P. J. Piggot, and J. A. Hoch.** 1985. Deduced product of the stage 0 sporulation gene *spo0F* shares homology with the Spo0A, OmpR, and SfrA proteins. Proc. Natl. Acad. Sci. USA **82**:7260–7264.
43. **Trisler, P., and S. Gottesman.** 1984. *lon* transcriptional regulation of genes necessary for capsular polysaccharide synthesis in *Escherichia coli* K-12. J. Bacteriol. **160**:184–191.
44. **Weber, R., and P. M. Silverman.** 1988. The Cpx proteins of *Escherichia coli* K12. Structure of the CpxA polypeptide as an inner membrane component. J. Mol. Biol. **203**:467–478.
45. **Weiss, V., and B. Magasanik.** 1988. Phosphorylation of nitrogen regulator I (NRI) of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85**:8919–8923.
46. **Wylie, D., A. Stock, C.-Y. Wong, and J. Stock.** 1988. Sensory transduction in bacterial chemotaxis involves phosphotransfer between Che proteins. Biochem. Biophys. Res. Commun. **151**: 891–896.
47. **Yang, Y., and G. F.-L. Ames.** 1988. DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. Proc. Natl. Acad. Sci. USA **85**:8850–8854.
48. **Zinkewich-Peotti, K., and J. M. Fraser.** 1988. New locus for exopolysaccharide overproduction in *Escherichia coli* K-12. J. Bacteriol. **170**:1405–1407.