Five promoters integrate control of the cob/pdu regulon in Salmonella typhimurium.

P Chen, M Ailion, T Bobik, G Stormo and J Roth
J. Bacteriol. **1995, 177(19):5401.**

JOURNAL OF BACTERIOLOGY, Oct. 1995, p. 5401–5410 Vol. 177, No. 19
0021-9193/95/$04.00+0

Five Promoters Integrate Control of the *cob*/*pdu* Regulon in *Salmonella typhimurium*

PING CHEN,[1] MICHAEL AILION,[1] THOMAS BOBIK,[1] GARY STORMO,[2] AND JOHN ROTH[1]*

Biology Department, University of Utah, Salt Lake City, Utah 84112,[1] and Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309[2]

Received 26 April 1995/Accepted 30 July 1995

Propanediol is degraded by a B_{12}-dependent pathway in *Salmonella typhimurium*. The enzymes for this pathway are encoded in a small region (minute 41) that includes the *pdu* operon (controlling B_{12}-dependent degradation of propanediol) and the divergent *cob* operon (controlling synthesis of cobalamin, B_{12}). Expression of both operons is induced by propanediol and globally controlled by the ArcA and Crp systems. The region between the two operons encodes two proteins, PduF, a transporter of propanediol, and PocR, which mediates the induction of the regulon by propanediol. Insertion mutations between the *pdu* and *cob* operons have been characterized, and their exact positions have been correlated with mutant phenotypes. The region includes five promoters, four of which are controlled by the PocR protein and induced by propanediol. The *cob* and *pdu* operons each have one regulated promoter; the *pduF* gene is expressed from two regulated promoters (P1 and P2). The P1 and P2 transcripts extend beyond *pduF* to include the *pocR* gene; thus the PocR protein autoregulates its expression from these promoters. The fifth promoter, P_{Poc}, is adjacent to the *pocR* gene and associated with a Crp binding site. We suggest that all global control of the regulon is exerted by regulating the level of PocR protein at the P1, P2, and P_{Poc} promoters. A putative binding site for the PocR protein has been identified by computer analysis. Eight close matches to this proposed site were found in regions near the four promoters known to be regulated by PocR protein: P_{Pdu}, P1, P2, and P_{Cob}. A three-state model is proposed in which the regulon uses all five of its promoters to control expression.

Salmonella typhimurium can use propanediol as a carbon and energy source via a pathway that requires cobalamin as a cofactor (18, 34). The ability of salmonellae to use propanediol depends on the presence of oxygen and exogenous cobalamin, vitamin B_{12}. This B_{12} requirement reflects the fact that salmonellae synthesize cobalamin only in the absence of oxygen (19). Despite these paradoxical laboratory observations, the genetics and regulatory behavior of the relevant genes suggest that, in nature, salmonellae synthesize cobalamin primarily to support the degradation of propanediol. The genes for propanediol degradation comprise an operon (*pdu*) which is located near a major cluster of B_{12}-synthetic genes, the *cob* operon (20, 28). These operons are transcribed divergently, and both are induced by propanediol and the positive regulatory protein PocR (6, 27). Both operons are subject to global control by the ArcA/ArcB system, which senses anaerobic conditions, and by the Crp/cyclic AMP system which senses a shortage of carbon and energy (2).

The region between the divergent *pdu* and *cob* operons is likely to include features important in the control of this regulon. Genetic studies of this region defined the regulatory gene *pocR* (6, 27). The nucleotide sequence of this region includes two open reading frames (11). One shows homology to the AraC family of regulatory proteins and is likely to encode the PocR regulatory protein; the other gene, *pduF*, appears to encode a diffusion facilitator for propanediol (11, 12).

Here we correlate the phenotypes of insertion mutations lying between the *pdu* and *cob* operons with their genetic map positions and exact sites in the published sequence. We report the mRNA start sites for four promoters: P_{Pdu}, P_{Poc}, and two promoters (P1 and P2) for the *pduF* gene. The mRNA start for the *cob* operon was determined previously (26). Evidence is presented that the two *pduF* promoters (like the *cob* and *pdu* operons) are regulated by the PocR protein and initiate transcripts that include the *pocR* gene. A putative PocR binding site was identified by aligning promoters known to be controlled by the PocR protein. These features are discussed in terms of a three-state model for the control of the regulon.

MATERIALS AND METHODS

Bacterial strains and strain construction. The genotypes of bacterial strains are given in Table 1. All strains were derived from *S. typhimurium* LT2. A transposition-defective derivative of phage Mu (MudJ) was used to create *lacZ* operon fusions (8). The insertion elements Tn*10d*Tc and Tn*10d*Cm are transposition-defective derivatives of transposon Tn*10* (13, 35).

For strain construction, transductional crosses were performed with the high-frequency generalized transducing mutant of phage P22, HT105/1 *int-201* (29). To select drug-resistant transductants (Tc^r, Kn^r, Cm^r), cells and phage were mixed and incubated nonselectively for 30 to 60 min. To remove phage, transductants were streaked on nonselective green indicator plates (9). Purified transductants were tested for phage sensitivity by cross-streaking with the P22 clear-plaque mutant H5.

Media and growth conditions. Rich medium was Difco nutrient broth with 0.5% NaCl. Minimum medium (NCE) was E medium without citrate (5). Carbon sources and electron acceptors were added at the following concentrations: glucose, 0.2% for aerobic culture and 0.4% for anaerobic culture; pyruvate, 0.44%; fumarate, 0.32%; succinate, 1.0%; and propanediol, 0.2%. Antibiotics were added at the concentrations described previously (6). The chromogenic β-galactosidase substrate X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was used at a final concentration of 40 mg/liter.

PCR and cycle sequencing. Oligonucleotides used are listed in Table 2. The primers A-1 and A-2 are directed toward each other and flank a region that includes the *pdu* promoter and its upstream region. The interval between the convergent primers F-1 and F-2 includes the *pduF* gene and its upstream promoter region. The convergent R-1 and R-2 primer pair flanks the *pocR* gene. Since the preceding intervals overlap, insertions at any position in the region can be detected.

The PCR mixtures contained reaction buffer (Promega), deoxynucleoside triphosphates (0.2 mM for each nucleotide), 1 mM $MgCl_2$, two primers (0.5 μM), and *Taq* polymerase (Promega; 0.025 U/μl); the total volume was 50 μl. Templates were either purified chromosomal DNA or cells taken directly from bacterial colonies. The PCR was run for 30 cycles; each cycle was 20 s at 94°C,

* Corresponding author.

TABLE 1. Strains used in this study

Strain	Genotype
TT10852	*metE205 ara-9 cob-24*::Mu*dJ*
TT17074	*Del1077* (*metE*) *ara-9 pdu-352*::Tn*10d*Tc (*zeb-3718*::Tn*10d*Tc)[a]
TT17075	*Del1077* (*metE*) *ara-9 pdu-353*::Tn*10d*Tc (*zeb-3719*::Tn*10d*Tc)[a]
TT17076	*Del1077* (*metE*) *ara-9 pdu-354*::Tn*10d*Tc (*zeb-3720*::Tn*10d*Tc)[a]
TT17078	*Del1077* (*metE*) *ara-9 pocR1*::Tn*10d*Tc
TT17080	*Del1077* (*metE*) *ara-9 pocR3*::Tn*10d*Tc
TT17089	*Del1077* (*metE*) *ara-9 pdu-201*::Tn*10d*Tc
TT17090	*Del1077* (*metE*) *ara-9 pdu-202*::Tn*10d*Tc
TT17091	*Del1077* (*metE*) *ara-9 pdu-203*::Tn*10d*Tc
TT17092	*Del1077* (*metE*) *ara-9 pdu-204*::Tn*10d*Tc
TT17107	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ* (*zeb-3722*::Mu*dJ*)[a]
TT17167	*Dup1736* [(*pocR12*)∗*MudA*∗(*hisD9952*)] *metE205 ara-9*
TT17173	*Dup1737* [(*cob-24*)∗*MudA*∗(*hisD9952*)] *metE205 ara-9*
TT17176	*Dup1737* [(*cob-24*)∗*MudA*∗(*hisD9952 pocR8*::Tn*10d*Tc)] *cob-364*::Tn*10d*Cm *metE205 ara-9*
TT17859	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ pocR1*::Tn*10d*Tc
TT17861	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ pdu-353*::Tn*10d*Tc
TT17862	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ pdu-354*::Tn*10d*Tc
TT17863	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ pdu-353*::Tn*10d*Tc *pocR15*::Tn*10d*Cm
TT17864	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ arcA*201::Tn*10d*Tc
TT17865	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ crp*-773::Tn*10*
TT17866	*Del1077* (*metE*) *ara-9 pdu-351*::Mu*dJ crp*-773::Tn*10 arcA*201::Tn*10d*Tc
TT18121	*Dup1737* [(*pduF354*::Tn*10d*Tc *cob-24*)∗*MudA*∗(*hisD9952*)] *metE205 ara-9*
TT18122	*Dup1737* [(*pduF354*::Tn*10d*Tc *cob-24*)∗*MudA*∗(*hisD9952 pocR15*::Tn*10d*Cm)] *metE205 ara-9*
TT18123	*Dup1736* [(*pocR12*)∗*MudA*∗(*hisD9952 pduF353*::Tn*10d*Tc)] *metE205 ara-9*
TT18124	*Dup1736* [(*pocR12*)∗*MudA*∗(*hisD9952 pduF354*::Tn*10d*Tc)] *metE205 ara-9*
TT18125	*Dup1736* [(*pduF353*::Tn*10d*Tc *pocR12*)∗*MudA*∗(*hisD9952*)] *metE205 ara-9*
TT18126	*Dup1736* [(*pduF354*::Tn*10d*Tc *pocR12*)∗*MudA*∗(*hisD9952*)] *metE205 ara-9*

[a] These *zeb* numbers are original allele designations assigned to these *pdu* insertions (6); these mutations were renamed when they proved to affect the newly discovered *pduF* gene.

20 s at 55°C, and 2 min at 72°C. The products of PCR were visualized on agarose gels (1%) by ethidium bromide staining.

Cycle sequencing was done with the *fmol* DNA sequencing system from Promega. Templates were PCR-amplified DNA purified with the GIBCO BRL glass MAX spin cartridge system. Primers were end labeled with [γ-^{32}P]ATP with T4 polynucleotide kinase. The sequencing reaction mixtures were analyzed on 6% acrylamide sequencing gels within 24 h.

For each Tn*10* insertion in a given region, four reactions were set up, each with a genomic primer at one side of the region and a Tn*10* primer homologous to one end of the element. Depending on the orientation of the Tn*10* element, only two reaction mixtures were expected to yield a product. As an example, mutant *pduF354* gave PCR products in two sets of reactions (with primer pair T-R and F-1 and with primer pair T-L and F-2). The three other Tn*10* insertions in this region are all in an orientation opposite to that of mutant *pduF354* since only reactions with oligonucleotide T-R and primer F-2 and oligonucleotide T-L and primer F-1 gave PCR products. The approximate location of the insertions was deduced by the length of the PCR products. For a Tn*10* insertion between primers F-1 and F-2 (separated by 1.5 kb), each amplified fragment should be 99 bases longer than the distance between primer F-1 (or F-2) and the insertion site, since the 5′ end of each Tn*10* primer is located 99 bases within the element.

TABLE 2. Oligonucleotides used in PCRs

Oligonucleotide	Sequence	Position of 5′ base[a] in *pdu-cob* sequence
A-1	5′-GAGCTCGAATTCGCTAACCGG-3′	1
A-2	5′-CGCCTTTAGAGAATCATTCATGTG-3′	1800
F-1	5′-GCTGACTTGCTGGTTGAGCAT-3′	1320
F-2	5′-ATGATAAAACCCCTCAGTTAA	2800
P-1	5′-GGCGCGTGATTTCGGTCCCAA-3′	2400
P-2	5′-AACTGGTGCAGTTTATGCCCT-3′	3850
T-I	5′-GACAAGATGTGTATCCACCTTAAC-3′	NR
T-L	5′-ACCTTTGGTCACCAACGCTTTTCC-3′	NR
T-R	5′-TCCATTGCTGTTGACAAAGGGAAT-3′	NR
MuL	5′-ATCCCGAATAATCCAATGTCC-3′	NR
MuR	5′-GAAACGCTTTCGCGTTTTTCGTGC-3′	NR

[a] Nucleotide positions are as described previously (11). NR, not relevant.

The insertion *pduF351*::Mu*dJ* was previously isolated and mapped between the *pdu* operon and the *pocR* gene, and its orientation was determined (6). The position of this insertion was identified by PCR with oligonucleotide primers MuL and MuR (which correspond to the left and right ends of the Mu transposon) in combination with the primers F-1 and F-2 (Table 2). The precise position of the insertion was determined by sequencing, performed as described above.

Preparation of RNA for primer extension. Total RNA was isolated by the hot-phenol-extraction method (1). Cells were grown on minimal medium with 1% succinate and 0.2% propanediol, sedimented by centrifugation, and resuspended in 2 ml of buffer (0.02 M sodium acetate [pH 5.5], 0.5% sodium dodecyl sulfate, and 1 mM EDTA) before phenol extraction at 60°C. Primer extension reactions with reverse transcriptase were performed as described previously (31). The oligonucleotide primers used to identify transcriptional starts were 5′-GTT GCATAAAGACCTCGCATG-3′ for the *pdu* operon; 5′-CACTATAAAAAGT GCCGCCTG-3′ for the *pduF* gene; and 5′-CATGATAAAACCCCTCAGTTA ATT-3′ for the *pocR* gene.

Total RNA was isolated from *S. typhimurium* LT2 and from the *his-cob* deletion mutant CRR299 (TT11855), which showed no runoff product. For strain LT2, the intensity of the reverse transcript signal seen for RNA from a propanediol-induced culture was 8- to 15-fold higher than that for RNA from an uninduced culture. The abundance of the *pocR* mRNA from cells grown aerobically in glucose was 6- to 10-fold higher than that of the other mRNAs of the *pdu*/*cob* regulon. This is consistent with the finding that basal expression of a *pocR-lac* fusion is higher than that of fusions to the other transcripts studied here.

Assay of β-galactosidase activity. The growth conditions (both aerobic and anaerobic) for cell cultures have been described previously (2, 6). The assay of β-galactosidase was by the method of Miller (23), and activities are reported in Miller units.

Binding site alignments. To identify common sequence patterns that might represent binding sites for the PocR protein, computerized searches were performed with the Consensus program (16, 17). Gaps were not permitted, but a variety of pattern widths were tested, as well as whether the binding sites were symmetric. Multiple sites were allowed to occur within each region (see Results).

Once putative sites were identified, they were converted to a specificity matrix (32) that was used to search the entire region for additional sites with the program PatSer. The specificity matrix is determined by the following method. The aligned sites give the frequency at which each base occurs at each position, f_{b,i}. The genome has an occurrence of each base of p_{b} (for *S. typhimurium* this is approximately 0.25 for each base). The elements of the specificity matrix are log (f_{b,i}/p_{b}). These values are related to the log-likelihood of observ-

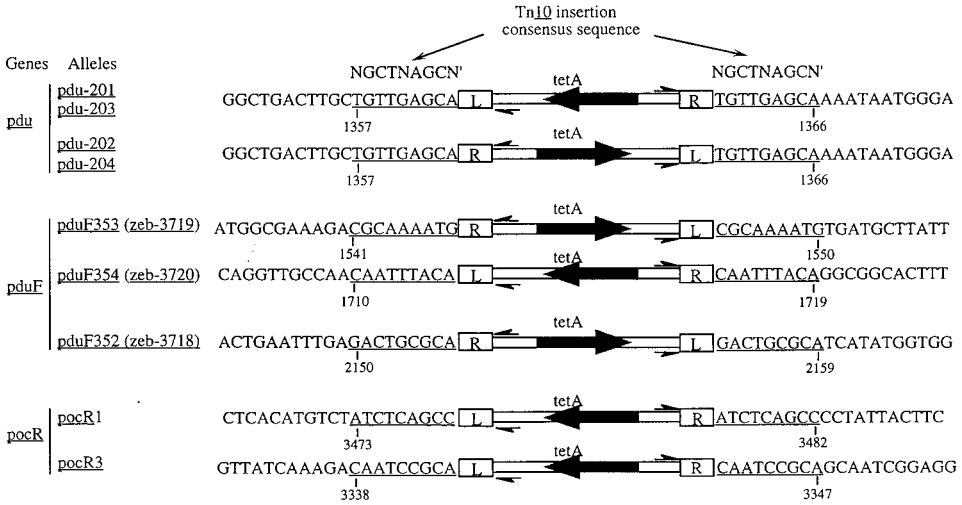

FIG. 1. Insertion sites of Tn*10d*Tc. The underlined sequences indicate the duplicated 9-bp target sequences flanking each Tn*10d*Tc insertion. Numbers indicate the positions of insertions in the sequence. The letters L and R indicate the 66-bp inverse repeats at the left and right ends of the Tn*10d*Tc element. Half arrows above and below the Tn*10d*Tc element represent oligonucleotides used for the PCR. The consensus target sequence for Tn*10* insertions is shown above the target site repeats. The thick shaded arrow within the Tn*10d*Tc element represents the *tetA* gene and its direction of transcription.

ing the particular frequencies of bases by chance, given the genomic base probabilities. The score for any potential binding site is calculated as the sum of the specificity matrix values corresponding to that sequence (32). These programs are available by anonymous ftp at the electronic mail address beagle.colorado.edu in pub/Consensus.

RESULTS

Localization of insertion mutations between the *pdu* and *cob* operons (the control region). Four classes of insertion mutations have been found to affect the *pdu*/*cob* regulon (6, 11). These mutations and their phenotypes are listed below in map order (counterclockwise in the chromosome).

(i) *pdu*. Mutations within the *pdu* operon cause a simple inability to degrade propanediol, which was scored as a failure to produce acid on MacConkey agar containing propanediol and B_{12} (6, 18).

(ii) *pduF*. This gene was identified in the sequence of the control region as an open reading frame whose encoded amino acid sequence is 65% identical to the sequence of the glycerol diffusion facilitator (*glpF*) (11). Mutations shown here to affect this gene have no growth phenotype but were located genetically within the region between the *pdu* operon and the *pocR* gene (6). Mutations mapping in this same general region are slightly impaired for the induction of the *cob* operon by propanediol (27). We designated this region of the sequence *pduF* in the belief that it encodes a diffusion facilitator for propanediol. This assignment has been confirmed by direct assay (12).

(iii) *pocR*. These mutations cause a defect in both cobalamin synthesis and propanediol degradation. Mutants with this designation show a recessive defect in transcription of both the *cob* and *pdu* operons (6, 27). This region of the sequence encodes a protein homologous to regulatory proteins of the AraC family (11, 28).

(iv) *cob*. Mutations in the *cob* operon eliminate only the ability to synthesize cobalamin (19, 20). This sequence of this operon has been reported previously (10, 28). Some mutations in this region have been sequenced previously (3, 24).

While the phenotypes, general map positions, and DNA sequences have suggested the arrangement listed above, there has been no direct correlation of *pdu*, *pduF*, or *pocR* mutations with exact positions in the sequence. We have identified such insertion sites.

Each insertion mutation affects the sequence at a position expected from the mutant phenotype. The exact position of each insertion was determined by sequencing the ends of the PCR fragments as described in Materials and Methods. Figure 1 shows the exact base-pair positions of Tn*10* insertions. For all Tn*10* insertions, there is a 9-bp direct repeat of the target sequence, with one copy at each end of the inserted material; these show agreement with the Tn*10* consensus target sequence (21). The positions of these insertions vis-à-vis other features of the region are summarized in Fig. 2A.

The four promoter-proximal *pdu* insertions are all located at the same site in the mRNA leader region 70 bp upstream of the *pduA* coding sequence and downstream of the mRNA start point (see below). Both orientations are represented. This appears to be a hot spot for Tn*10* insertion. The target sequence at this site differs by only one base pair from the Tn*10* target consensus (21, 22).

Four Tn*10* insertions were localized in the *pduF* region. One insertion is between promoters P1 and P2 (*pduF353*::Tn*10d*Tc); one insertion is between promoter P2 and the *pduF* coding sequence (*pduF354*::Tn*10d*Tc). The remaining two insertions are in the *pduF* gene (*pduF352*::Tn*10d*Tc and *pduF355*::Tn*10d*Tc). The *pduF355* insertion was shown by PCR to affect the *pduF* coding sequence, but its exact nucleotide position was not determined.

The *pocR* insertion mutations, which cause a defect in the expression of both the *cob* and *pdu* operons, all affect the open reading frame whose inferred product is similar to that of the AraC regulatory protein, supporting the idea that these mutations eliminate the production of the positive activator of the *cob*/*pdu* regulon.

Determination of transcription start sites for the *pdu* operon, the *pduF* gene, and the *pocR* gene. The 5′ ends of transcripts originating in the control region were determined by primer extension as described in Materials and Methods. Results are depicted in Fig. 3, and start sites are diagrammed in Fig. 2A.

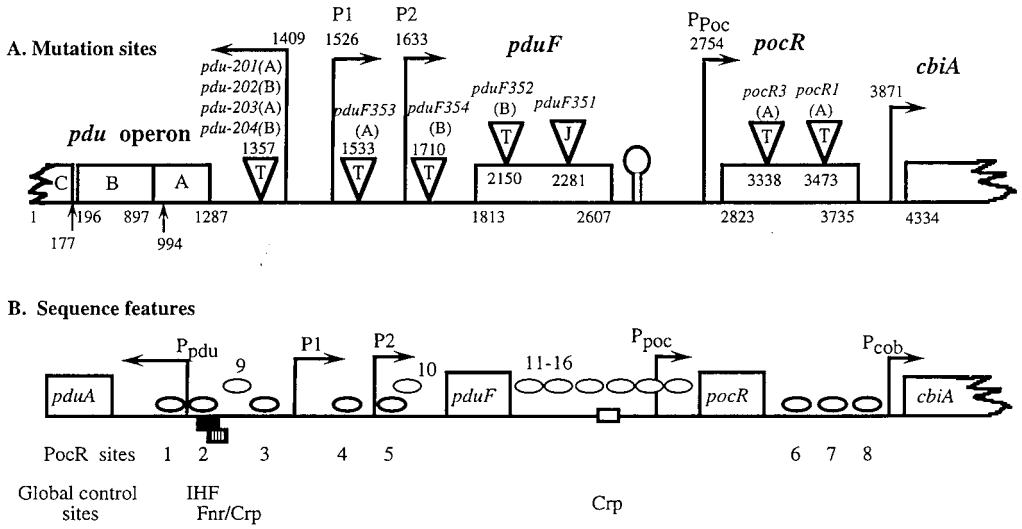

FIG. 2. (A) Exact positions of promoters, insertions, and a putative terminator in the control region. The letters C, B, and A in the *pdu* operon designate the first three genes. Arrows represent transcript start sites. Triangles marked with T or J are Tn*10d*Tc or MudJ insertions, respectively. The symbols (A) and (B) near Tn*10* insertions indicate the orientation of the *tetA* gene in each Tn*10* element (clockwise and counterclockwise, respectively). (B) Putative binding sites for PocR and global regulatory proteins. The banded rectangle indicates a putative hybrid Fnr-Crp binding site; the black rectangle is an IHF binding site; the white rectangle is the Crp binding site; and ovals are putative *pocR* binding sites. Darkly drawn ovals near the horizontal line are sites with a close match to the consensus; lighter ovals farther from the line are sites with a weaker match to the consensus.

For the *pdu* operon, a single transcript start site is initiated with a G residue located 122 bp upstream of the *pduA* coding sequence. There are two mRNA start sites for the *pduF* gene; the P1 and P2 start sites are located 287 and 180 bp before the start site of the *pduF* coding sequence. Only one transcription start site is found near the *pocR* gene; this start site is located 50 bp before the start site of the *pocR* coding sequence.

Induction of the *pduF* gene by propanediol. Expression of the *pduF* gene is induced by propanediol under both aerobic and anaerobic conditions (Table 3, line 1). Data on the behavior of a *cob* fusion are included for purposes of comparison (Table 3, line 2). The basal level of *pduF*::*lacZ* transcription is higher during aerobic growth on succinate or anaerobic growth on pyruvate-fumarate than it is during aerobic growth on glucose. Maximal induction by propanediol was obtained under conditions of anaerobic respiration on pyruvate-fumarate. This pattern of expression is very similar to that found previously for the *cob* and *pdu* operons and the *pocR* gene (2, 6).

Role of the PocR protein in *pduF* gene control. The induc-

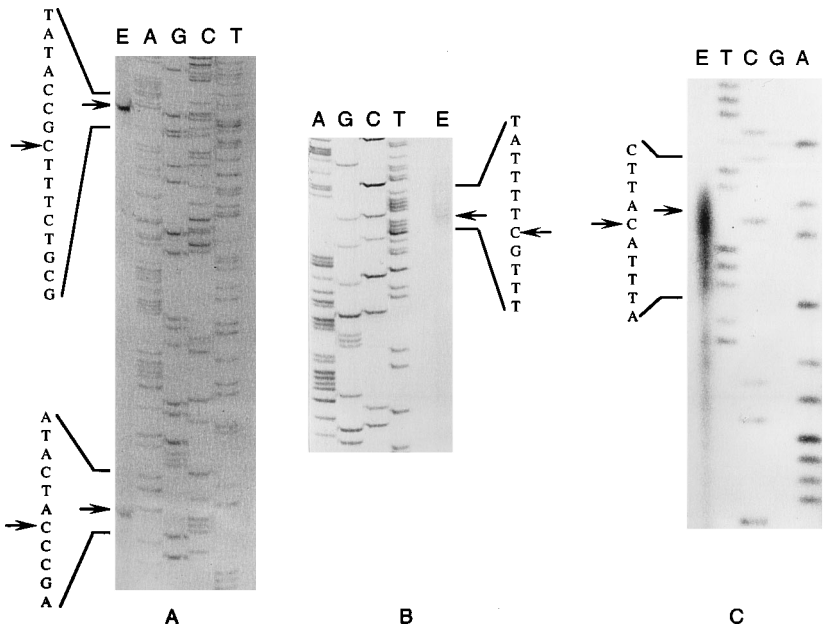

FIG. 3. Primer extension analysis for the *pduF* gene (A), the *pdu* operon (B), and the *pocR* gene (C). Lane E contains cDNA from reverse transcription. Lanes A, G, C, and T are termination positions of dideoxy sequencing ladders primed by the same primer used for reverse transcription. Arrows indicate the position of the message start.

TABLE 3. Regulation of the *pduF* gene

Line	Strain	Relevant genotype	β-Galactosidase levels (Miller units) in cells grown									
			Aerobically on:						Anaerobically on:			
			Glucose		Succinate		Pyruvate		Glucose		Pyruvate-fumarate	
			Without PD[a]	With PD	Without PD	With PD	Without PD	With PD	Without PD	With PD	Without PD	With PD
1	TT17107	*pduF*::*lac*	2	8	25	320	ND[b]	ND	5	25	9	540
2	TT10852	*cob*::*lac*	3	15	30	510	ND	ND	1	24	67	550
3	TT17107	*pduF*::*lac*	2	18	12	450	ND	ND	4	42	45	630
4	TT17859	*pduF*::*lac pocR*	3	2	4	6	ND	ND	3	5	5	4
5	TT17862	*pduF*::*lac pdu-354*::Tn*10*	3	2	2	3	ND	ND	4	3	4	5
6	TT17861	*pduF*::*lac pdu-353*::Tn*10*	4	24	2	120	ND	ND	2	38	4	250
7	TT17863	*pduF*::*lac pdu-353*::Tn*10 pocR*	3	2	2	1	ND	ND	2	4	5	4
8	TT17107	*pduF*::*lac*	3	14	18	420	8	36	5	26	8	846
9	TT17864	*pduF*::*lac arcA*	3	6	8	127	1	13	2	5	5	168
10	TT17865	*pduF*::*lac crp*	2	2	NG[c]	NG	3	1	1	2	4	18
11	TT17866	*pduF*::*lac arcA crp*	2	2	NG	NG	1	2	2	1	1	1

[a] PD, propanediol.
[b] ND, not done.
[c] NG, no growth; *crp* mutants cannot use succinate as a sole carbon source.

tion of the *pduF* gene by propanediol requires the PocR protein (Table 3, compare lines 3 and 4). The *pduF354*::Tn*10d*Tc mutation (between the P2 *pduF* promoter and the *pduF* coding sequence) eliminates *pduF* expression (Table 3, line 5). This suggests that *pduF* transcription starts to the left of this point and is consistent with the identified P1 and P2 promoters being the biologically relevant promoters for the *pduF* gene. The *pduF353*::Tn*10d*Tc mutation (between the P1 and the P2 *pduF* promoters) should block the P1 transcript but allow transcription of the *pduF* gene from the P2 promoter. This insertion leaves the *pduF* gene inducible but reduces the maximal expression level; the residual induction, presumably from the P2 promoter, requires PocR protein (Table 3, lines 6 and 7). Thus, both the P1 and P2 promoters appear to contribute to *pduF* expression, and at least P2 is regulated by PocR protein. Barring other effects of these Tn*10d*Tc insertions, the control of P1 would be seen as the difference between values in Table 3, line 1 (with both P1 and P2) and those in line 4 (with P2 alone). These differences are small but suggest that promoter P1 (like P2) may be controlled by the PocR protein.

Effect of the Arc and Crp global regulatory systems on expression of the *pduF* gene. Previous studies have shown that two global regulatory systems (Crp and Arc) control the *cob*, *pdu*, and *pocR* transcripts (2, 4). These two global control systems also affect expression of the *pduF* gene (Table 3, lines 8 to 11). When cells were grown aerobically on pyruvate, the *arcA* mutation caused a small reduction in the induced level of the *pduF* transcription, while a *crp* mutation caused a 15-fold decrease; this is expected since Crp is the major activator under aerobic conditions. During anaerobic growth on glucose, the *arcA* mutation strongly reduced the maximal induction level and a *crp* mutation had a smaller effect (Table 3, lines 9 and 10); this is consistent with ArcA acting without Crp under anaerobic conditions on a good carbon source and with both Crp and ArcA acting together during anaerobic growth on poor carbon sources. The *arcA* and *crp* mutations together essentially eliminate expression under all growth conditions tested (Table 3, line 11). Thus, the maximal induction of the *pduF* gene (like those of the *pocR*, *cob*, and *pdu* transcripts) requires both the Arc and the Crp global systems.

We found previously that *arc* and *crp* mutations had additive effects on the induced expression level of the *cob*, *pdu*, and *pocR* genes during anaerobic respiration (2). This additive quality was not seen for the *pduF* gene (described above); the *arcA* mutation reduced expression fivefold, and the *crp* mutation decreased expression 45-fold. We believe that the behavior of *pduF* fusions is complex, because the P1 and P2 promoters normally contribute to the expression of PocR, the positive regulator. This contribution is prevented by the polar effects of the *pduF*::MudJ(*lac*) insertion used to monitor *pduF* transcription. Thus, in the *pduF* fusion strains, the *pocR* gene is expressed only by its own promoter, P_{poc}. Evidence to support this is described below. The large effect of the *crp* mutation seen above may reflect in part the effect of the Crp protein on the P_{poc} promoter.

Evidence for transcription of *pocR* from *pduF* promoters (P1 and P2). Insertion mutations in the *pduF* gene have previously been seen to reduce maximal induction of the *cob* operon (27). This effect could be due to the restricted transport of propanediol (the lack of the PduF transporter). Alternatively, the reduction could reflect a deficit in PocR protein caused by polar effects of the *pduF* insertion on the *pocR* gene. The latter effect would be expected if some expression of *pocR* depended on transcripts that originated at the P1 and P2 promoters of the *pduF* gene. A complementation test should distinguish between these two possibilities. A simple transport defect should be corrected by supplying a wild-type copy of the *pduF* gene in *trans*. However, the polar effect of a *pduF* insertion on the adjacent *pocR* gene should not be *trans*-correctable by a $pduF^+$ allele.

This complementation test was made with strains containing a duplication of the region that includes the *pduF*, *pocR*, and *cob* promoters. The genotype of the basic strain is diagrammed in Fig. 4. This strain carried a *cob*::*lac* fusion at the duplication join point; this fusion served as a reporter of operon expression; each copy of the duplicated region included both the *pduF* gene and the *pocR* gene. Starting with this strain, we could place a *pduF* insertion *cis* to either a $pocR^+$ or a *pocR* mutant allele and we could provide a $pduF^+$ allele in *trans* for the complementation test. The *cob* operon expression in these strains is shown in Table 4.

In Table 4, it can be seen that *cob* operon induction was

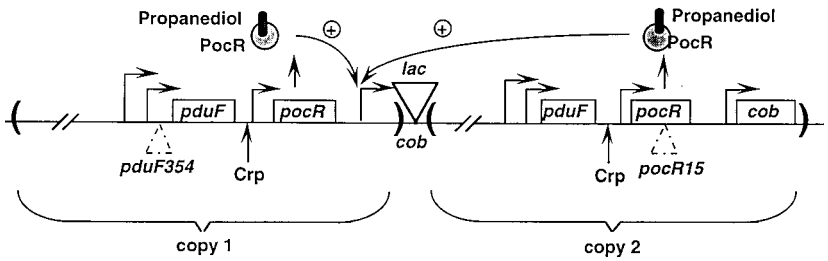

FIG. 4. The structure of duplication strains used in Table 4. The inverted triangle indicates a Mud::*lac* insertion at the duplication join point; this insertion creates a *cob-lac* operon fusion. Dotted and dashed triangles indicate the position at which Tn*10d*Tc insertions were placed in some strains. Straight arrows indicate the direction and start of transcription. Curved arrows indicate the position at which PocR acts to stimulate the *cob* operon. Open rectangles delineate genes. The PocR protein and propanediol are represented by shaded circles and smaller shaded cylinders, respectively. Plus signs in circles indicate that PocR protein acts as a positive regulatory protein.

reduced by a *pduF*::Tn*10* insertion only if that insertion was placed *cis* to the only functional *pocR* gene (Table 4, last line). Both a *pocR* insertion (line 2) and a *pduF* insertion (line 3) are fully recessive when tested individually, but they do not complement each other in the diploid; this is expected if the *pduF* insertion caused a defect in the expression of the adjacent $pocR^+$ allele. The results suggest that a significant amount of *pocR* transcription is initiated at the P1 and/or P2 promoters. Since these promoters are regulated by PocR (see above), it seems likely that they are responsible for the autoinduction of *pocR*, which has been reported previously (6).

Insertions within the P2 transcript eliminate autoinduction of PocR. The previous results suggest that some transcripts of the *pocR* gene might originate at the P1 and P2 promoters of the *pduF* gene. We tested this possibility in duplication merodiploids with a *pocR*::*lac* fusion at the join point and with a second, wild-type $pocR^+$ allele. Insertions of Tn*10* were placed downstream of the P1 or P2 promoters, either *cis* or *trans* to the *pocR*::*lac* fusion. The structure of these strains is shown in Fig. 5, and the ability of the strains to control *pocR* transcription is shown in Table 5.

In Table 5, it can be seen that blocking both P1 and P2 transcripts with a Tn*10* insertion, placed *cis* to the *pocR* reporter fusion, eliminated autoinduction but left considerable *pocR* gene expression, presumably from the P_{Poc} promoter. A smaller effect was seen for the *cis* Tn*10* insertion that blocked only P1. The effect of Tn*10* insertions placed in *trans* is probably due to a reduction in total PocR production. These effects of *pduF* insertions on *pocR* expression cannot be due to the impairment of propanediol transport, since all of these strains carry a functional *pduF* gene.

In these assays it should be noted that the uninduced levels of *pocR* gene expression are quite high, even in strains with a *pduF* insertion. This suggests that the P_{Poc} promoter is active but is not autoregulated. Other results suggest that the P_{Poc} promoter is stimulated by the Crp protein, which might explain the high basal levels of transcription under these growth conditions. In any case, the *pocR* gene is transcribed from three promoters, two of which (P1 and P2) are subject to control by PocR protein.

Efforts to identify possible binding sites for PocR. The data presented, with previously published data, suggest that four promoters are controlled by the PocR protein, the *cob* operon promoter, the *pdu* operon promoter, and the P1 and P2 promoters of the *pduF* gene (2, 6, 26, 27). To identify potential binding sites for the PocR protein within the regulatory regions of the *pdu* and *cob* operons, we used the program Consensus (16, 17). This program finds sites that, when aligned, have high information content values. The search procedure is described in Materials and Methods.

Initially, we considered only two fragments of DNA, the intergenic region between the *pduA* and *pduF* genes, which includes the P_{pdu}, P1, and P2 promoters, and the region between the *pocR* and *cbiA* genes, which includes the P_{cob} promoter. We first tested sites in the size range of 15 to 25 bases, which did not require symmetry. The most significant patterns were, in fact, approximately symmetric. We then required symmetry and tested a wider range of site widths of up to 32 bases long. It is common for prokaryotic regulatory proteins, including AraC, a homolog of PocR (11), to be dimeric and to bind at long symmetrical sites. In some analyses, we included the region immediately upstream of the *pocR* gene, even though there is no evidence that the PocR protein regulates its own synthesis from this closest promoter, P_{Poc}. In these latter analyses, we did not require that sites from the *pocR* region contribute to the aligned set of binding sites, but we allowed that to occur if it would increase the information content of the set.

The predicted sites are listed in Fig. 6 together with a consensus sequence defined with the eight strongest sites (16 half-sites). In total, 15 significant matches were found between the

TABLE 4. Effects of a *pduF* insertion on the inducibility of a *cob-24*::*lac* fusion

Strain	Genotype and positions of mutations relative to *cob-24*::*lac* fusion		β-Galactosidase activity (Miller units) in cells grown			
			Aerobically on succinate		Anaerobically on pyruvate-fumarate	
	cis	*trans*	Without PD[a]	With PD	Without PD	With PD
TT17173	$pduF^+$ $pocR^+$	$pduF^+$ $pocR^+$	24	424	64	454
TT17176	$pduF^+$ $pocR^+$	$pduF^+$ *pocR*	26	305	111	538
TT18121	*pduF* $pocR^+$	$pduF^+$ $pocR^+$	20	268	106	483
TT18122	*pduF* $pocR^+$	$pduF^+$ *pocR*	16	44	42	214

[a] PD, propanediol.

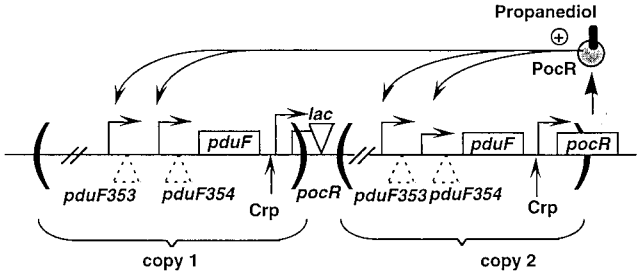

FIG. 5. The structure of duplication strains used in Table 5. See the legend to Fig. 4 for an explanation of the symbols.

pdu and *cob* operons. Figure 2B presents the position of these sites vis-à-vis other features of the region. Even when the P_{Poc} region (between the *pduF* and *pocR* genes) was included in the analysis, the program found eight strong sites that occur only in the *pduA-pduF* and *pocR-cbiA* regulatory regions. The former region contains five predicted sites, and the latter region contains three sites. The most significant site width is 30 bases, with an information content of about 18 bits (when corrected for the sample size). This amount of information is fairly typical of sites for prokaryotic-positive regulatory proteins (30). The graph of information content versus base position within the site (Fig. 6) indicates that, if these sites are correct, the PocR protein probably binds to three consecutive major grooves of binding site DNA, as was suggested for the homologous AraC protein (15).

The scores for the eight strongest sites range from 24.6 to 29.8. In the entire 22 kb of known sequence of the *cob*/*pdu* regulon (11, 28), only two other sites have a score higher than 18; these are both within the *pduA-pduF* regulatory region and may be significant but weaker binding sites for the protein. Average scores of random sequences are slightly less than 0. It is interesting that the region around the *pocR* promoter (not known to be controlled by PocR protein) has six fairly regularly spaced weak sites, with scores in the range of 13.0 to 15.6. The possible significance of these sites is discussed below.

DISCUSSION

We have determined the positions of insertion mutations and transcript start points in the region between the *pdu* and *cob* operons. The phenotype of each sequenced insertion mutation allows the assignment of a function to each open reading frame in this region. The *pdu* operon, the *cob* operon and the *pocR* gene each have one transcriptional start site. The *pduF* gene has two promoters. We cannot exclude the possibility that the runoff reverse transcripts seen are due to the processing of a longer mRNA or to secondary structures in mRNA that block reverse transcriptase. However, we think it likely that the identified sites represent message start sites, since each site is found near sequences with strong similarities to sigma 70 promoters and because genetic data indicate the existence of a promoter in the general region of each identified site.

Data presented here suggest that significant transcription of *pocR* is initiated at the sites of the P1 and P2 promoters and extends across the *pduF* and *pocR* genes. This was surprising, since the region immediately downstream of the *pduF* gene includes a stem-loop structure resembling a Rho-independent message termination site (11). The results presented here suggest that this stem-loop is not a transcription terminator or that it causes only incomplete termination under the growth conditions tested.

The existence of the *pduF* gene was first inferred from sequence data (11), and its function (propanediol transport) was initially assigned on the basis of its close similarity to the GlpF protein, which facilitates the diffusion of glycerol (14, 33). We have recently shown directly that a *pduF* mutation reduces but does not eliminate propanediol entry (12). Apparently, these alternative means of uptake are strong enough that *pduF* mutants show no growth phenotype attributable to the transport deficit. Some insertions in the general region between the *pdu* and *cob* operons were found to reduce the maximum induction of *cob-lac* or *pdu-lac* fusions by about twofold (27). Evidence presented here suggests that this defect is due to the impaired expression of the positive regulatory protein PocR. We conclude that the P1 and P2 transcripts contribute to *pocR* expression and that *pduF* insertions reduce the level of PocR protein.

A predicted binding site for PocR protein was identified by the alignment of the regions near each of the promoters known to be controlled by PocR protein. Eight strong matches to this sequence were found, five between the *pduA* and *pduF* genes and three immediately upstream of the *cob* promoter. Weaker matches to this sequence were also found, including six closely spaced copies immediately upstream of the *pocR* gene (Fig. 2B). While we cannot be certain that these sites are all significant, many are in positions appropriate for the activation or repression of transcription. Computerized searches such as this cannot replace direct binding studies, but it should be remembered that direct binding tests are informative only if they give positive results. Many binding sites can be missed by direct assays if proper test conditions are not known; binding sites are extremely difficult to identify by genetic analysis. The computer searches suggest models that can be tested more directly and point out sites that should be scrutinized most carefully.

The site suggested for PocR binding has several features that are consistent with a role in regulation. The site resembles that recognized by the AraC protein, a homolog of PocR. Similarities include large size (30 bp), imperfect matches to the consensus, and three blocks of eight highly informative bases separated by three much less informative bases. In the case of AraC, these eight-base blocks were suggested to represent regions of the major groove of DNA that are recognized by the protein. It should be noted that later analyses of the AraC binding site suggest important asymmetrical aspects of this site (7, 25).

The positions of strong binding sites upstream of the *cob* promoter and near the P1 and P2 promoters of the *pduF* gene all seem reasonable for use by a transcription activator. Bind-

TABLE 5. Effects of a *pduF* insertion on autoinduction of the *pocR* gene

Strain[a]	Genotype and position of mutation relative to *pocR*::*lac* fusion		β-Galactosidase activity (Miller units) in cells grown			
			Aerobically on succinate		Anaerobically on pyruvate-fumarate	
	cis	*trans*	Without PD[b]	With PD	Without PD	With PD
TT17167	*pduF*$^+$	*pduF*$^+$	192	468	384	515
TT18123	*pduF*$^+$	*pduF353*	212	413	427	657
TT18124	*pduF*$^+$	*pduF354*	180	315	449	641
TT18125	*pduF353*	*pduF*$^+$	223	329	397	563
TT18126	*pduF354*	*pduF*$^+$	214	192	347	253

[a] The genotype of each strain was *pocR*::*lac*/*pocR*$^+$.
[b] PD, propanediol.

Site	Sequence of half-sites	Position of 5'base	Score
#1	TTGAG CATCA GAAAA TAATG GGAAA GCAAT	1363	26.7
	ATTGC TTTCC CATTA TTTTC TGATG CTCAA		
#2	TTTGC TTTTA TAACT TATTG ATAAA TTACA	1408	29.2
	TGTAA TTTAT CAATA AGTTA TAAAA GCAAA		
#3	ATTTC ATTTT TACAT TTTTG TTAGT GACGA	1478	24.6
	TCGTC ACTAA CAAAA ATGTA AAAAT GAAAT		
#4	TTGGC TTGTG CAATA TTTTA TTATT TTCAC	1569	27.1
	GTGAA AATAA TAAAA TATTG CACAA GCCAA		
#5	ATGGG CTGCT AAAAA TATCG GAAAA GCAAA	1631	24.8
	TTTGC TTTTC CGATA TTTTT AGCAG CCCAT		
#6	TTATG AATAG CAATA TTTTG CTATT GCCGA	3730	29.8
	TCGGC AATAG CAAAA TATTG CTATT CATAA		
#7	TTTTG TATCT TAATC ATTTC AGAAA GAAAA	3773	29.4
	TTTTC TTTCT GAAAT GATTA AGATA CAAAA		
#8	TCTGG TGTAA CAATA AATTG TCATA GCGCA	3806	27.2
	TGCGC TATGA CAATT TATTG TTACA CCAGA		
Weaker sites			
#9	TTGTC TATAT GATTA TTTTG TTTTT ATTAT	1450	19.6
#10	ATCGG AAAAG CAAAA TTTTG CTATC GCAAC	1647	21.3
Near P_{Poc}			
#11	TCTGC TTTCA CGAAC TGATC AAAAC GTCTC	2670	13.0
#12	CCAGC AATGG GAGAC TTTTG ATATT GTCAT	2700	15.0
#13	AGGAT AATGG TAATA AAAAG CGAAT GTAAA	2749	14.9
#14	TTATC AGGGC CAGGA TAATG GTAAT AAAAA	2788	15.6
#15	TCTGA ACTCA GAACT CATTA ATAAA ATCGC	2841	13.7
#16	TCATG GTGAT GAAAT ATCTG AGCTA TTCAA	2919	13.4

HALF-SITE CONSENSUS SEQUENCE (BOTH HALVES OF 8 BEST SITES)

A	3	0	1	3	2	5	6	0	6	6	1	15	14	7	11
C	0	3	1	0	9	2	1	0	5	2	9	0	1	1	1
G	1	2	6	8	5	0	1	2	1	3	2	1	0	0	0
T	12	11	8	5	0	9	8	14	4	5	4	0	1	8	4

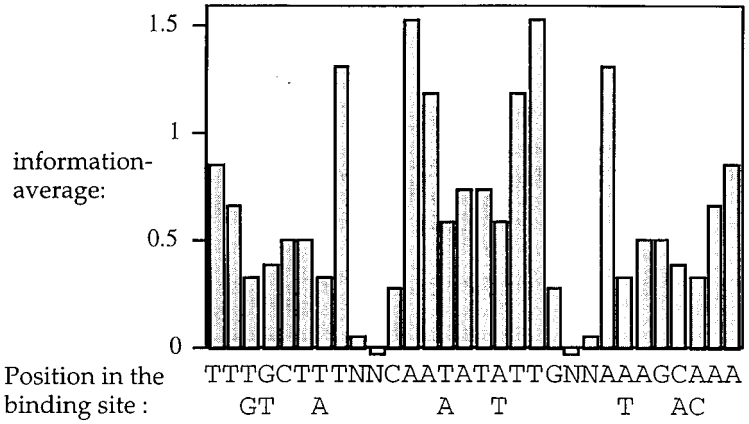

FIG. 6. Predicted binding sites for the PocR protein. The identified strong (no. 1 through 8) and weak (no. 9 and 10) sites are listed at the top with the positions of their 5′ bases and their specificity matrix scores. The binding sites near the P_{Poc} promoter (no. 11 through 16) are also listed. Underlining of bases indicates matches to the consensus sequence. In the middle of the figure, alignment of the two halves of the eight strongest sites is used to identify the consensus sequence. Below is a graph of the information content of sites versus base positions in the proposed PocR binding site. The height of a bar indicates how heavily this position contributed to an identification of the site.

ing of PocR protein to the two strong sites that flank the *pdu* message start point seems more likely to prevent transcription. These sites may serve to allow maximal shutoff. The IHF and Crp-Fnr sites near this promoter may also hinder the activity of the *pdu* promoter. It seems likely that the activation of the *pdu* promoter is achieved with the weak site (no. 9) and the strong site (no. 3) slightly farther from the *pdu* promoter (Fig. 2B).

Six weak matches to the predicted binding site for PocR protein were found near the P_{Poc} promoter, which is not known to be controlled by PocR. These weak sites might be populated, perhaps cooperatively, at a high concentration of PocR protein to limit its synthesis. Alternatively, the PocR protein may have a slightly different binding site specificity in the absence of propanediol, and these sites could be the preferred binding sites in the absence of an inducer and serve to repress the *pocR* gene when its product is not needed.

We suggest a three-state model for control of the *pdu*/*cob* regulon. This model provides a framework for thinking about the five promoters, the proposed PocR binding sites, and the available genetic data. The model (Fig. 7) accounts for most of the available data and makes many testable predictions.

We propose an off condition in which all the promoters are

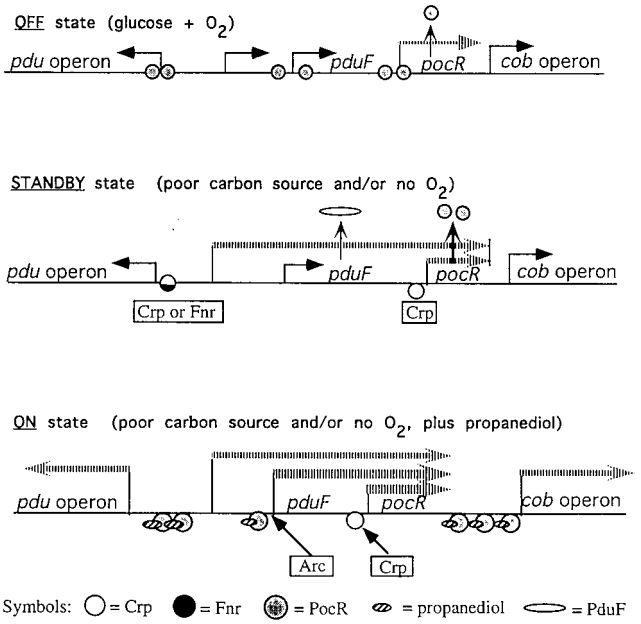

FIG. 7. Model for regulation of the *pdu*/*cob* regulon. Dark arrows indicate the direction of transcription. Arrows constructed with vertical lines indicate transcripts actively formed under the indicated condition.

inactive, except for a basal level of P_{Poc} promoter activity. The off condition might exist during aerobic growth in glucose without inducer. Under these conditions, the low basal level of the PocR protein may repress the expression of some promoters (P_{pdu}, P2, and P_{Poc}).

The second state (standby) would exist when cells grow without propanediol under global conditions that are appropriate for regulon expression (e.g., aerobic growth on a poor carbon source or anaerobic respiration). Under standby conditions, the P1 promoter is activated by IHF with Fnr or Crp proteins acting alternatively at their nearby binding site; this binding occludes the *pdu* promoter. The Crp protein also stimulates the *pocR* promoter, P_{Poc}. In the standby state, the active promoters (P1, P_{Poc}) express genes for the transporter (PduF) and positive regulator (PocR), preparing cells to take up and respond to the inducer (propanediol) if it should appear. (The propanediol "antenna" is up.)

When propanediol appears, the cell shifts to the on state. The inducer enters and interacts with PocR protein to stimulate the P1 and P2 promoters, increasing inducer entry and induction capacity. We propose that in the on state, Arc protein stimulates the P2 promoter and Crp stimulates the P_{Poc} promoter; thus these global controls would independently stimulate PocR production. As the level of the PocR-propanediol complex increases, the divergent *pdu* and *cob* operons are induced. Having separate *pocR* promoters activated by Arc or Crp would account for the additive effects of these global regulatory proteins on induced expression levels of the *pdu*, *cob*, and *pocR* transcripts (2). In the model, global regulators act directly only to increase the PocR level. We propose that PocR protein (with bound propanediol) acts alone to activate the *pdu* and *cob* promoters.

ACKNOWLEDGMENTS

We thank Elliot Altman and Jeff Lawrence for help with PCR amplification of sequences.

The work was supported by NIH grant GM34804.

REFERENCES

1. **Aiba, H., S. Adhya, and B. de Crombrugghe.** 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:**11905–11910.
2. **Ailion, M., T. A. Bobik, and J. R. Roth.** 1993. Two global regulatory systems (Crp and Arc) control the cobalamin/propanediol regulon of *Salmonella typhimurium*. J. Bacteriol. **175:**7200–7208.
3. **Ailion, M., and J. R. Roth.** Unpublished results.
4. **Andersson, D. I.** 1992. Involvement of the Arc system in redox regulation of the Cob operon in *Salmonella typhimurium*. Mol. Microbiol. **6:**1491–1494.
5. **Berkowitz, D., J. M. Hushon, J. H. J. Whitfield, J. R. Roth, and B. N. Ames.** 1968. Procedure for identifying nonsense mutations. J. Bacteriol. **96:**215–220.
6. **Bobik, T. A., M. Ailion, and J. R. Roth.** 1992. A single regulatory gene integrates control of vitamin B_{12} synthesis and propanediol degradation. J. Bacteriol. **174:**2253–2266.
7. **Carra, J. H., and J. H. Schleif.** 1993. Variation of half-site organization and DNA looping by AraC protein. EMBO J. **12:**35–44.
8. **Castilho, B. A., P. Olfson, and M. J. Casadaban.** 1984. Plasmid insertion mutagenesis and *lac* gene fusion with mini-Mu bacteriophage transposons. J. Bacteriol. **158:**488–495.
9. **Chan, R. K., D. Botstein, T. Watanabe, and Y. Ogata.** 1972. Specialized transduction of tetracycline resistance by phage P22 in *Salmonella typhimurium*. II. Properties of a high-frequency-transducing lysate. Virology **50:**883–898.
10. **Chen, P., M. Ailion, N. Weyand, and J. R. Roth.** 1995. The end of the *cob* operon: evidence that the last gene (*cobT*) catalyzes synthesis of the lower ligand of vitamin B_{12}, dimethylbenzimidazole. J. Bacteriol. **177:**1461–1469.
11. **Chen, P., D. I. Andersson, and J. R. Roth.** 1994. The control region of the *cob*/*pdu* regulon. J. Bacteriol. **176:**5474–5482.
12. **Chen, P., and J. R. Roth.** Unpublished data.
13. **Elliott, T., and J. R. Roth.** 1988. Characterization of Tn*10d*-Cam: a transposition-defective Tn*10* specifying chloramphenicol resistance. J. Bacteriology **213:**332–338.
14. **Heller, K. B., E. C. C. Lin, and T. H. Wilson.** 1980. Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. J. Bacteriol. **144:**274–278.
15. **Hendrickson, W., and R. Schleif.** 1985. A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site. Proc. Natl. Acad. Sci. USA **82:**3129–3133.
16. **Hertz, G. Z., G. W. Hartzell, and G. D. Stormo.** 1990. Identification of consensus patterns in unaligned DNA sequences known to be functionally related. Comput. Appl. Biosci. **6:**81–92.
17. **Hertz, G. Z., and G. D. Stormo.** 1995. Identification of consensus patterns in unaligned DNA and protein sequences known to be functionally related, p. 199–214. *In* Bioinformatics and genome research: Proceedings of the Third International Conference on Bioinformatics and Genome Research. World Scientific Publishing, Singapore.
18. **Jeter, R. M.** 1990. Cobalamin-dependent 1,2-propanediol utilization by *Salmonella typhimurium*. J. Gen. Microbiol. **136:**887–896.
19. **Jeter, R. M., B. M. Olivera, and J. R. Roth.** 1984. *Salmonella typhimurium* synthesizes cobalamin (vitamin B_{12}) de novo under anaerobic growth conditions. J. Bacteriol. **159:**206–213.
20. **Jeter, R. M., and J. R. Roth.** 1987. Cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium*. J. Bacteriol. **169:**3189–3198.
21. **Kleckner, N.** 1979. DNA sequence analysis of Tn10 insertions: origin and role of 9bp flanking repetitions during Tn10 translocation. Cell **16:**711–720.
22. **Kleckner, N., D. A. Steele, K. Reichardt, and D. Botstein.** 1979. Specificity of insertion by the translocatable tetracycline-resistance element Tn*10*. Genetics **92:**1023–1040.
23. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Raux, E., A. Lanois, F. Levillayer, M. Warren, E. Brody, A. Rambach, and C. Thermes.** Vitamin B_{12} biosynthesis in *Escherichia coli*, retrograde evolution by genetic engineering. Submitted for publication.
25. **Reeder, T., and R. Schleif.** 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. J. Mol. Biol. **231:**205–218.
26. **Richter-Dahlfors, A. A., and D. I. Andersson.** 1991. Analysis of an anaerobically induced promoter for the cobalamin biosynthetic genes in *Salmonella typhimurium*. Mol. Microbiol. **5:**1337–1345.
27. **Rondon, M. R., and J. C. Escalante-Semerena.** 1992. The *poc* locus is required for 1,2-propanediol-dependent transcription of the cobalamin biosynthetic (*cob*) and propanediol utilization (*pdu*) genes of *Salmonella typhimurium*. J. Bacteriol. **174:**2267–2272.
28. **Roth, J. R., J. G. Lawrence, M. Rubenfield, S. Kieffer-Higgins, and G. M. Church.** 1993. Characterization of the cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium*. J. Bacteriol. **175:**3303–3316.
29. **Schmieger, H.** 1971. A method for detection of phage mutants with altered transducing ability. Mol. Gen. Genet. **110:**378–381.
30. **Schneider, T. D., G. D. Stormo, L. Gold, and A. Ehrenfeucht.** 1986. Infor-

mation content of binding sites on nucleotide sequences. J. Mol. Biol. **188:** 415–431.
31. **Stirling, D. A., C. S. J. Hulton, L. Waddell, S. F. Park, G. S. A. B. Stewart, I. R. Booth, and C. F. Higgins.** 1989. Molecular characterization of the *proU* loci of *Salmonella typhimurium* and *Escherichia coli* encoding osmoregulated glycine betaine transport systems. Mol. Microbiol. **3:**1025–1038.
32. **Stormo, G. D.** 1990. Consensus patterns in DNA. Methods Enzymol. **183:** 211–221.
33. **Sweet, G., C. Gandor, R. Voegele, N. Wittekindt, J. Beuerle, V. Truniger, E. C. C. Lin, and W. Boos.** 1990. Glycerol facilitator of *Escherichia coli*: cloning of *glpF* and identification of the *glpF* product. J. Bacteriol. **172:**424–430.
34. **Toraya, T., S. Honda, and S. Fukui.** 1979. Fermentation of 1,2-propanediol and 1,2-ethanediol by some genera of *Enterobacteriaceae*, involving coenzyme B_{12}-dependent diol dehydratase. J. Bacteriol. **139:**39–47.
35. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32:**369–379.