JOURNAL OF BACTERIOLOGY, Jan. 1996, p. 266–272
0021-9193/96/$04.00+0

Vol. 178, No. 1

Characterization of PcaQ, a LysR-Type Transcriptional Activator Required for Catabolism of Phenolic Compounds, from *Agrobacterium tumefaciens*

DONNA PARKE*

Department of Biology, Yale University, New Haven, Connecticut 06520-8103

Received 7 August 1995/Accepted 3 November 1995

Previous work demonstrated that catabolism of the phenolic compounds *p*-hydroxybenzoate and protocatechuate via the β-ketoadipate pathway in *Agrobacterium tumefaciens* is mediated by a regulatory gene, *pcaQ*, that acts in *trans* to elicit expression of many of the enzymes encoded by the *pca* genes. There was evidence that five *pca* structural genes are organized in a polycistronic operon transcribed in the order *pcaDCHGB*. The *pcaQ* gene is upstream of this operon. The activator encoded by *pcaQ* was novel in having the metabolite β-carboxy-*cis*,*cis*-muconate as a coinducer. This communication reports the nucleotide sequence of *pcaQ* and identifies its deduced polypeptide product as a member of the LysR family of regulatory molecules. PcaQ has a calculated molecular weight of 33,546, which is consistent with the size of LysR relatives. Like many other LysR members, PcaQ serves as an activator at the level of transcription, it has a conserved amino-terminal domain, and its gene is transcribed divergently from the operon that it regulates and is subject to negative autoregulation. Studies of coinducer specificity identified an unstable pathway metabolite, γ-carboxymuconolactone, as a second coinducer. Analysis of expression from a *pcaD*::*lacZ* promoter probe plasmid revealed that PcaQ and the coinducer exert their effect on a 133-nucleotide region upstream of *pcaD*. The nucleotide sequence of this region in a mutant strain constitutive for enzymes encoded by the *pcaDCHGB* operon identified nucleotides likely to be involved in the *pcaDCHGB* promoter and substantiated the inclusion of five *pca* structural genes in the operon.

The phenolic compounds *p*-hydroxybenzoate and protocatechuate are broken down into tricarboxylic acid cycle intermediates by the protocatechuate branch of the β-ketoadipate pathway in members of the bacterial family *Rhizobiaceae*. Structural *pca* genes for the protocatechuate pathway are clustered in two divergently transcribed regions 4 kb apart in the *Agrobacterium tumefaciens* A348 genome. A previous study presented evidence that five of these genes are organized in an operon transcribed in the order *pcaDCHGB* (15). The divergently transcribed *pcaIJ* genes for the penultimate step in the pathway are regulated independently in response to β-ketoadipate. A region upstream of the *pcaD* gene was identified as a regulatory gene, *pcaQ*, that encodes an activator which acts in *trans* to elicit expression of the enol-lactone hydrolase (PcaD), encoded by *pcaD* (14). In addition, *pcaQ* is required for the expression of a decarboxylase (PcaC), protocatechuate 3,4-dioxygenase (PcaHG), and a lactonizing enzyme (PcaB) (15) (Fig. 1). The initial characterization of *pcaQ* relied on an *Escherichia coli* bioconversion system. *E. coli* does not have enzymes of protocatechuate catabolism, but the introduction of plasmid-borne *pcaHG* genes enables it to convert exogenously supplied protocatechuate into the commercially unavailable metabolite β-carboxy-*cis*,*cis*-muconate intracellularly. By studying the expression of PcaD from an *A. tumefaciens* subclone containing *pcaQ* and *pcaD* in such an *E. coli* background, it was determined that β-carboxy-*cis*,*cis*-muconate induced PcaD but protocatechuate did not (14). Protocatechuate was also found not to induce any *pca* gene products in *A. tumefaciens*.

The expression of enzymes of protocatechuate catabolism in *Rhizobium leguminosarum* biovar trifolii shows a similar pattern of induction, with β-carboxy-*cis*,*cis*-muconate inducing at least two enzymes of the upper part of the pathway and β-ketoadipate inducing PcaIJ (18). In other bacteria characterized with respect to patterns of induction of enzymes of the β-ketoadipate pathway, protocatechuate and/or β-ketoadipate typically serve as inducers (27). A regulator of the protocatechuate branch of the pathway has been characterized at the molecular level. PcaR from *Pseudomonas putida* has β-ketoadipate as the coinducer; it belongs to a small class of regulatory molecules that includes PobR (19). Regulatory molecules that govern the catechol branch of the β-ketoadipate pathway include CatR from *P. putida* and CatM from *Acinetobacter calcoaceticus* (9, 20). The latter polypeptides respond to *cis*,*cis*-muconate and belong to a subgroup within the LysR family of regulatory molecules.

This communication reports the nucleotide sequence of *pcaQ* and the deduced amino acid sequence of its product. Further characterization of PcaQ elucidated the level at which it activates expression of the *pcaDCHGB* operon, how its synthesis is regulated, and its coinducer specificity and provided a preliminary identification of a *pcaDCHGB* promoter region. The characteristics of PcaQ place it in the LysR family of regulators and reveal that it is remote from the regulator controlling the *pca* pathway in *P. putida*.

MATERIALS AND METHODS

Construction of plasmids. Table 1 lists the bacterial strains and plasmids used in this study. Figure 2 shows physical maps of relevant plasmids. Standard techniques of molecular biology were used in plasmid and gene manipulations (21). The source of *lacZ* for operon fusions was pKOK6, which contains sym-

* Mailing address: KBT No. 840, Dept. of Biology, Yale University, P.O. Box 208103, New Haven, CT 06520-8103. Phone: (203) 432-3505. Fax: (203) 432-6161. Electronic mail address: dparke@minerva.cis.yale.edu.

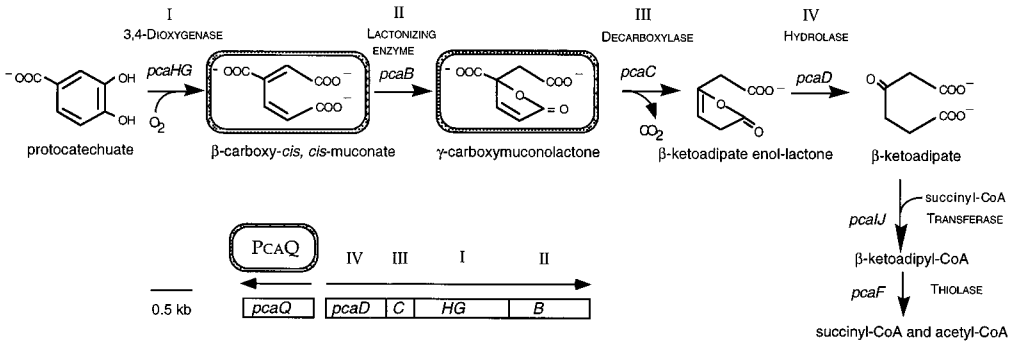

FIG. 1. Protocatechuate branch of the β-ketoadipate pathway in *A. tumefaciens*. The top line of horizontal arrows shows the first four enzymatic steps encoded by the *pca* operon which is under the control of PcaQ. Enzymatic steps include (I) protocatechuate 3,4-dioxygenase (EC 1.13.11.3), (II) β-carboxy-*cis*,*cis*-muconate lactonizing enzyme (EC 5.5.1.2), (III) γ-carboxymuconolactone decarboxylase (EC 4.1.1.44), (IV) β-ketoadipate enol-lactone hydrolase (EC 3.1.1.24), and β-ketoadipate succinyl-coenzyme A transferase (EC 2.8.3.6). Pathway intermediates that serve as coinducers with PcaQ are framed. Shown beneath the pathway is the organization of the *pcaDCHGB* operon, divergently transcribed from the *pcaQ* gene. CoA, coenzyme A.

metrical restriction sites on either side of a sequence that includes a kanamycin resistance marker and a promoterless *lacZ* (7). Two of the *lacZ* fusion plasmids were derivatives of pARO65 and pARO66. Plasmid pARO65 contains a *Hin*cII subclone of *A. tumefaciens* DNA which includes *pcaQ* and *pcaD* inserted into the *Pvu*II sites of pUC18, thereby eliminating the *lac* promoter and the multiple cloning site. A derivative of pARO65, pARO66, contains an Ω element in the unique *Pvu*II site near the middle of *pcaQ*.

The construction of pARO76 entailed removing the Ω element from pARO66 with a *Bam*HI digestion and introducing the *lacZ* cassette from pKOK6 into the *Bam*HI site. This plasmid was used as a vehicle for introducing *lacZ* into *pcaQ* in *A. tumefaciens* cells, described in the next section. The construction of pARO77 involved the insertion of the *lacZ* cassette with *Sal*I ends into the *Xho*I site of pARO65; the *Xho*I site is within *pcaD*, 0.25 kb from the presumed 5′ end of the gene (16).

As a control plasmid and a stepping stone to a broad-host-range *lacZ* fusion plasmid, pARO158 was made with a *Pst*I-*Hin*dIII fragment from pARO148, which included *pcaD* but only the first 33 nucleotides of *pcaQ*. The fragment was inserted into pRK415-1 so that the direction of transcription of *pcaD* was op-

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)[a]	Reference or source
A. tumefaciens strains		
A348	Cm^r Nal^r Rif^r Sm^r; derived from C-58, containing the cryptic plasmid pAtC58 and the octopine Ti plasmid pTiA6	6
ADO2044	Spc^r; *pcaQ1*::Ω mutant of A348	14
ADO2048	Spc^r; *pca-3*; pseudorevertant of ADO2044, which is PcaDHGB[c]	15
ADO2076	Km^r; *pcaQ2*::*lacZ*-Km^r mutant of A348	This study
E. coli strains		
DH5α	F$^-$ ϕ80d*lacZ* ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17* (r_K^- m_K^-) *supE44* λ$^-$ *thi-1 gyrA96 relA1*	Gibco-BRL[b]
S17-1	*recA pro hsdR*; RP4-Tc::Mu-Km::Tn*7* integrated into the chromosome	26
Plasmids[c]		
pARO65	Ap^r; 2.6-kb *Hin*cII fragment (*pcaQ*$^+$ *pcaD*$^+$) in *Pvu*II sites of pUC18	14
pARO66	Ap^r Spc^r; pARO65 with Ω in *Pvu*II site within *pcaQ*	14
pARO76	Ap^r; pARO66 with *lacZ*-Km^r cassette of pKOK6 in *Bam*HI site left by removal of Ω from *pcaQ*	This study
pARO77	Ap^r; pARO65 with *lacZ*-Km^r cassette of pKOK6 in *Xho*I site of *pcaD*	This study
pARO80	Km^r Tc^r; pARO158 with *lacZ*-Km^r cassette of pKOK6 replacing 1.0-kb *Nsi*I fragment	This study
pARO82	Ap^r Spc^r; 3.74-kb *Acc*I-*Nsi*I fragment containing Ω from ADO2048 in pBKS	This study
pARO148	Ap^r; 1.3-kb *Pst*I-*Hin*dIII fragment containing *pcaD* in pUC18	15
pARO158	Tc^r; 1.3-kb *Pst*I-*Hin*dIII fragment of pARO148 in pRK415	This study
pARO528	Ap^r; 2.5-kb *Bgl*II-*Hin*dIII fragment spanning *pcaQ* and *pcaD* in pBKS	This study
pARO531	Tc^r; 1.2-kb *Bgl*II-*Pst*I fragment containing part of *pcaQ* in pRK415	This study
pARO535	Tc^r; 1.45-kb *Bgl*II-*Xho*I fragment containing *pcaQ* in pRK415	14
pARO538	Tc^r; 0.38-kb *Pst*I-*Xho*I fragment containing *pcaQ-pcaD* intergenic region in pRK415	This study
pBKS	Ap^r	Stratagene[d]
pKOK6	Ap^r Cm^r Km^r; 11-kb plasmid containing promoterless *lacZ*-Km^r cassette for constructing operon fusions	7
pRK415	Tc^r; broad host range	5
pUC18 and pUC19	Ap^r	30
pZR9	Ap^r; 5.4-kb *Sau*3AI insertion of *A. calcoaceticus pcaIJFBD* in pUC19	U. Gerischer

[a] Cm, chloramphenicol; Nal, nalidixin; Rif, rifampin; Sm, streptomycin.
[b] Competent *E. coli* DH5α cells were purchased from Gibco-BRL, Gaithersburg, Md.
[c] Excluding the Ω element and *lacZ*-Km^r cassette, insertions in the listed pARO plasmids contain agrobacterial DNA.
[d] Stratagene, La Jolla, Calif.

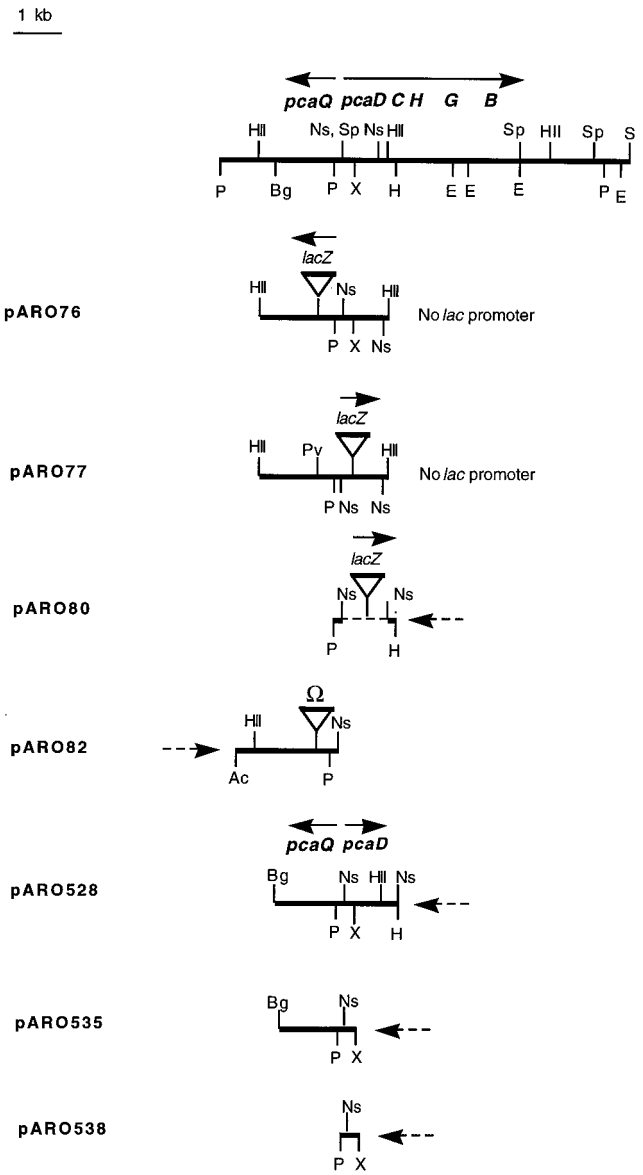

FIG. 2. Subclones of the *pca* genetic region of *A. tumefaciens* used to characterize *pcaQ* and its product, aligned below a restriction map of the region. The direction of transcription is shown with an arrow above the genes. The direction of transcription from the *lac* promoter is denoted by a dashed arrow. Restriction enzyme sites are abbreviated as follows: Ac, *Acc*I; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; HII, *Hin*cII; Ns, *Nsi*I; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Sp, *Sph*I; X, *Xho*I.

posite to the direction of transcription from the *lac* promoter of the vector. A 1.0-kb *Nsi*I fragment internal to *pcaDC* (16) was removed from pARO158, and the *lacZ* cassette, with *Pst*I ends, was inserted into the *Nsi*I site, forming pARO80. Between the downstream end of the *lacZ*-Km^r cassette and vector sequences lie 117 bp of a truncated *pcaC* (16).

Strain ADO2048 is a pseudorevertant of the *pcaQ*::Ω strain ADO2044, which retains the omega element in *pcaQ* but expresses PcaD, PcaHG, and PcaB constitutively (15). The presumptive mutation unique to strain ADO2048 was cloned from the strain by isolating a 3.74-kb *Acc*I-*Nsi*I fragment, which was inserted into pBKS, creating pARO82. The fragment selected bore the omega spectinomycin-streptomycin (Spc-Str) resistance marker, and it was presumed to include the intergenic region of *pcaQ* and *pcaD*.

Table 1 describes the other plasmids constructed for this study. Plasmids were maintained in *E. coli* cells with 12.5 μg of tetracycline (Tc) ml^{-1}, 30 μg of kanamycin (Km) ml^{-1}, and 100 μg of ampicillin (Ap) ml^{-1}.

Generation and characterization of a *pcaQ*::*lacZ* mutant strain of *A. tumefaciens*. A strain of *A. tumefaciens* with a reporter gene in *pcaQ* was desired in order to measure chromosomal *pcaQ* promoter activity in the presence or absence of PcaQ. This strain, ADO2076, was constructed by the electroporation of *A. tumefaciens* A348 with pARO76. Colonies which were resistant to 100 μg of kanamycin ml^{-1} but sensitive to 100 μg of ampicillin ml^{-1} were presumed to contain the *lacZ* cassette with the Km^r marker gene in *pcaQ* but not vector DNA; they were purified and further characterized. Strain ADO2076 had the properties of a previously characterized *pcaQ*::Ω strain, ADO2044 (14). ADO2076 failed to grow on minimal medium (MM) containing quinate, shikimate, *p*-hydroxybenzoate, or protocatechuate as carbon sources, and it accumulated protocatechuate on MM plates containing *p*-hydroxybenzoate and the protocatechuate-complexing chromophore *p*-toluidine (13). The introduction of pARO535, which contains the entire *pcaQ* gene, into ADO2076 enabled the strain to grow at the expense of shikimate in the presence of tetracycline at 1 μg ml^{-1}. pARO531, which contains a truncated *pcaQ* gene, did not complement the mutation in ADO2076 under the same growth conditions.

DNA sequencing. Plasmid DNA used in sequencing was derived from pARO528, a pBKS plasmid with a 2.55-kb *Bgl*II-*Hin*dIII insertion of *A. tumefaciens* DNA. Nested deletions of this plasmid were created from both ends of the insertion by exonuclease III digestion (21). The nucleotide sequence was determined in both directions with pARO528 and its deletion derivatives. Dideoxy chain termination sequencing reactions (22) with ^{35}S-dATP (Amersham, Arlington Heights, Ill.) and Sequenase 2.0 (United States Biochemical Corp., Cleveland, Ohio) were carried out on double-stranded plasmid DNA.

Transfer of plasmids into *A. tumefaciens* by conjugation. Plasmids pARO158, pARO80, pARO535, and pARO531 were transferred from *E. coli* S17-1 cells to *A. tumefaciens* by conjugation. The S17-1 transformants were grown in liquid Luria-Bertani medium (LB) (25) overnight and rinsed with LB, and a 50-μl aliquot of S17-1 donor cells was mixed on agar-solidified medium with 100 μl of a recipient cell culture. Conjugations were carried out overnight on LB at 30°C. Transconjugants were selected and purified on MM (17) with 10 mM succinate. The level of tetracycline used to select for the pRK415-based plasmids in *A. tumefaciens* was 1 μg ml^{-1}.

Growth conditions and β-galactosidase assays. For autoregulation studies, *A. tumefaciens* ADO2076 cultures containing either pARO531 or pARO535 were grown at 30°C in 5 ml of LB with tetracycline and kanamycin to an optical density of 0.6 at 620 nm. Cells were rinsed and resuspended in equal volumes of MM containing 10 mM arabinose or 10 mM arabinose plus 5 mM quinate. Following incubation at 30°C for different lengths of time, cells were rinsed and resuspended in equal volumes of MM containing 20 mM glucose at 4°C.

The effectiveness of γ-carboxymuconolactone as a coinducer was determined with *E. coli* DH5α cells harboring pARO77. Cultures were grown overnight in LB with kanamycin, rinsed, and resuspended in equal volumes of MM. A single culture was used for one set of experiments by diluting the cells 1:5 with MM containing supplements and by distributing 1-ml aliquots to growth tubes. The MM supplements were 20 mM glucose, thiamine at 1 μg ml^{-1}, arginine at 0.1 mg ml^{-1}, and kanamycin at 20 μg ml^{-1}. Diluted cells were incubated at 37°C in a rotary shaker for 1 h. At this time, a 50-μl dose of inducer or control solution, described below, was added to the cells; additions were repeated every 30 min for 4 h. Cells were harvested as described above for the autoregulation experiments.

A 10 mM solution of γ-carboxymuconolactone was prepared by adding 5 μmol of β-carboxy-*cis*,*cis*-muconate, 50 μmol of Tris-HCl at pH 7.5, and 12 U of β-carboxy-*cis*,*cis*-muconate lactonizing enzyme (PcaB) from an extract of DH5α (pZR9) in a total volume of 500 μl. The solution was left at room temperature for 5 min, filter sterilized, and kept on ice for immediate use. Control mixtures were prepared analogously, and equal volumes of solution were aliquoted to parallel cultures in a given experiment. Control solutions had separate components of the γ-carboxymuconolactone mix in the same concentrations and at the same final volume: a solution of Tris-HCl buffer, a solution of β-carboxy-*cis*,*cis*-muconate in the buffer, and a solution of PcaB in the buffer.

To investigate *trans* activation of the *pcaD* promoter, the pARO80 *pcaD*::*lacZ* construction was used. Turbid cultures were diluted 1:50 in fresh MM and grown in the presence of inducing or noninducing carbon sources. Tetracycline was used to maintain pARO80. *A. tumefaciens* cells were harvested at an optical density at 620 nm of 0.5 to 0.7. The harvesting procedure was similar to that used for *E. coli* cells.

For β-galactosidase (LacZ) assays, cells were permeabilized by adding 2 drops of chloroform and 1 drop of 0.1% sodium dodecyl sulfate to 1 ml of cells. The method of Miller (8) was followed to assay for LacZ activity.

Nucleotide sequence accession number. The DNA sequence of the *pcaQ* sequence reported here has been submitted to GenBank (accession number U32867).

RESULTS

DNA sequence of *pcaQ* and the deduced amino acid sequence of PcaQ. The DNA sequence assembled from nested deletions of pARO528 was analyzed in all six reading frames. An open reading frame (ORF) which has homology to *pcaD* and *catD* of *A. calcoaceticus* was identified (16). A 1,098-bp segment of DNA extending from an *Nsi*I site presumed to be

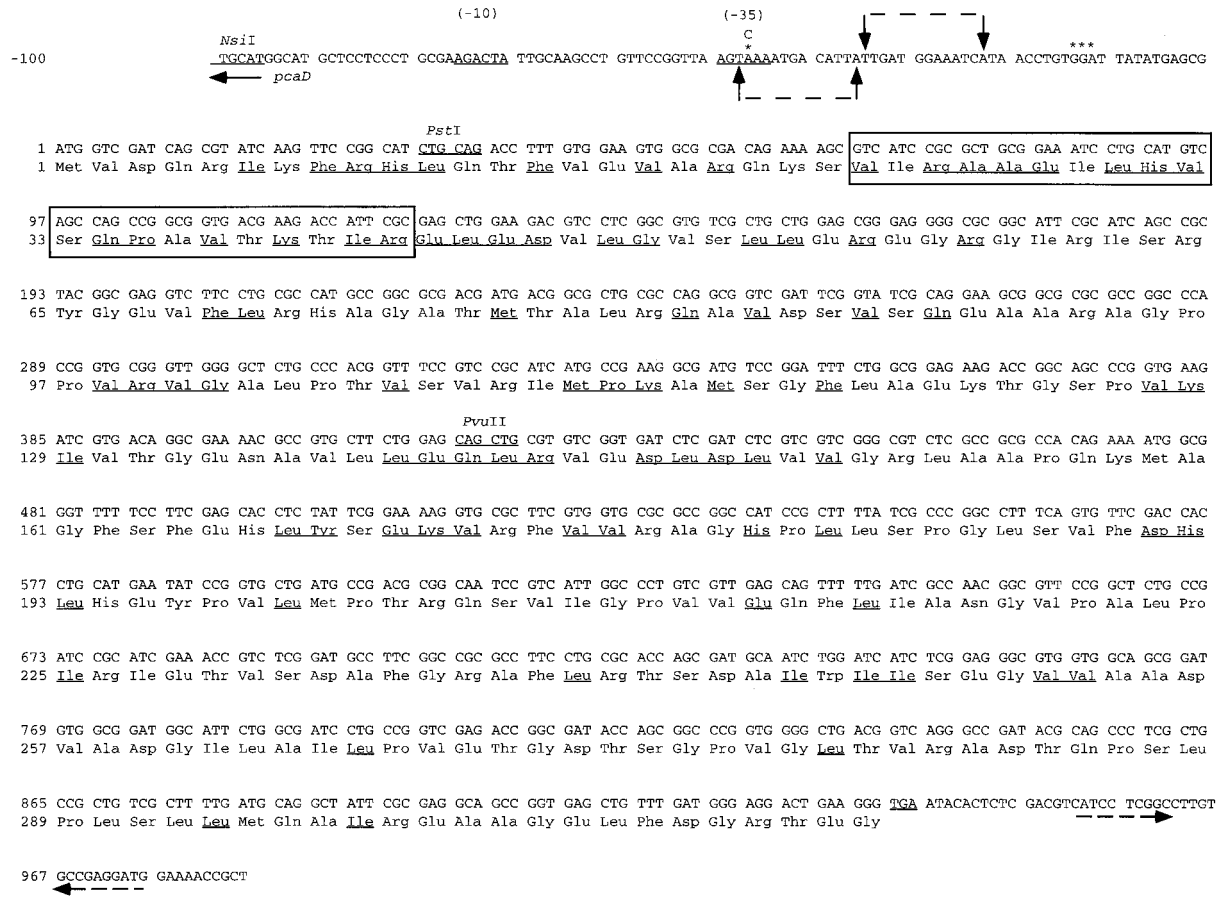

FIG. 3. Nucleotide and deduced amino acid sequences of the *pcaQ* ORF from *A. tumefaciens* A348. A putative ribosomal binding site is marked with three asterisks. A possible ATG translational start site for the divergently transcribed *pcaD* gene is located at nucleotide −96 on the complementary strand. (Note that 21 nucleotides upstream of that codon lies another potential initiation codon, GTG.) A probable ρ-dependent termination structure (10-nucleotide stem–5-nucleotide loop–10-nucleotide stem; estimated $\Delta G = -22$ kcal [−92.048 kJ] at 25°C), which is 15 nucleotides beyond the underlined TGA stop codon, is marked with divergent dashed arrows. This palindrome is 22 nucleotides from the second stop codon of the *pobA* gene (16). Restriction sites for *Nsi*I (partial site shown), *Pst*I, and *Pvu*II are underlined. Amino acids 23 to 42 that make up a conserved LysR-type helix-turn-helix motif (4) are boxed. Underlined amino acid residues are similar in >50% of 22 LysR family members, as analyzed by Viale et al. (28). An up promoter mutation from adenine to cytosine is noted with an asterisk at position −47. The locations of likely −10 and −35 promoter sequences for *pcaD* on the opposite strand are underlined. Two pairs of inverted repetitions that possess the LysR-class consensus motif upstream of *pcaD* and *pcaQ* are indicated by arrows linked by dashes. The pairs of short inverted repetitions occur within the lower set of arrows and outside the upper pair.

at the initiation codon of the *pcaD* gene to the upstream region previously identified as containing *pcaQ* (14) is shown in Fig. 3. The sequence was verified on both strands, and its analysis led to the assignment of an ORF to *pcaQ* (Fig. 3). Previous work showed that a functional PcaQ was produced from a 1.53-kb *Xho*I-*Bgl*II fragment (pARO535) but not from a 1.2-kb *Pst*I-*Bgl*II fragment (pARO531), in which the *Xho*I and *Pst*I sites are on the *pcaD* side of the fragment (14). In addition, the insertion of an omega element at a unique *Pvu*II site inactivated PcaQ. These results indicated that the PcaQ translational start site is upstream of the *Pst*I site and that the ORF extends beyond the *Pvu*II site (Fig. 3). When the nucleotide sequence divergently transcribed from *pcaD* was translated in all three reading frames, only one gave an extended ORF that fulfilled the requirements defined by the *Pst*I and *Pvu*II sites. The only possible start codon in that reading frame upstream of the *Pst*I site is 11 nucleotides downstream of a plausible prokaryotic ribosome-binding sequence, GGA (Fig. 3). The translational stop codon for the *pcaQ* ORF is 71 nucleotides away from the stop codon for another ORF coding for *pobA* (16). As shown in Fig. 3, a possible ρ-dependent transcription termination structure lies between the two genes.

The *A. tumefaciens pcaQ* gene encodes a polypeptide of 311 amino acids with a sequence-calculated molecular mass of 33.5 kDa. The G+C content of *pcaQ* is 64%, higher than that of the overall *A. tumefaciens* genome (60 to 62%) (11) and higher than that of the adjacent genes *pobA* and *pcaD*, which have G+C contents of 60% and 61%, respectively (16). Anomalously low Lys-to-Arg ratios were observed in many LysR family members (28). A low Lys to Lys-plus-Arg ratio (0.25%) is present in PcaQ also. However, similarly low ratios in the products of the genes adjacent to *pcaQ*, 0.27% and 0.29% in PobA and PcaD, respectively (16), supports the suggestion that the low ratios are related to high G+C content (23).

Sequence homology of PcaQ to members of the LysR family of regulatory proteins. The GenBank sequence library was scanned for sequences homologous to PcaQ. Polypeptides identified as homologs were members of the LysR family of bacterial regulators (4, 23). Underlined amino acid residues in Fig. 3 share a similarity coefficient above a log-odds threshold

TABLE 2. *trans* autorepression of the *pcaQ* promoter

Time after resuspension in minimal medium (h)	Carbon source	β-Galactosidase activity in ADO2076 (*pcaQ*::*lacZ*)[a] with:		% Activity[b]
		pARO538 (PcaQ$^-$)	pARO535 (PcaQ$^+$)	
0	Arabinose	153 (1.7)	4 (0.04)	2.6
25	Arabinose	487 (2.6)	7 (0.24)	1.4
25	Arabinose + quinate	779 (4.0)	13 (0.13)	1.7

[a] Average of values in Miller units from three independent cultures with standard deviations in parentheses.
[b] Level of repressed activity divided by the amount of unrepressed activity.

TABLE 3. Effects of different exogenous coinducers on induction of β-galactosidase synthesis (*pcaD*::*lacZ*) by PcaQ in *E. coli* DH5α (pARO77) cells

Addition(s) to medium	β-Galactosidase activity (Miller units ± SD)[a]
Tris-HCl	180 ± 4.0
PcaB enzyme extract in Tris-HCl	156 ± 0.3
β-Carboxy-*cis*,*cis*-muconate in Tris-HCl	235 ± 6.0
γ-Carboxymuconolactone[b]	21,159 ± 383

[a] Averages of two independent experiments.
[b] With added PcaB enzyme extract and β-carboxy-*cis*,*cis*-muconate in Tris-HCl.

of 10 with aligned residues of more than half of 22 other LysR family members (28).

Transcriptional regulation of *pcaQ*. To study the expression of *pcaQ*, *A. tumefaciens* ADO2076 was constructed. This strain carries a *pcaQ*::*lacZ* chromosomal fusion. Two plasmid constructions, based on the broad-host-range plasmid pRK415-1, were introduced into this strain to study the effect of PcaQ on its own synthesis. Plasmid pARO535 contains the entire *pcaQ* gene; pARO538 contains only 33 bp of *pcaQ* at the 5′ end (Fig. 2 and 3). The latter plasmid was a suitable control because it provided the *pcaQ-pcaD* intergenic region of pARO535 in equal copy numbers. Table 2 shows that 25 h after the resuspension of *A. tumefaciens* cells in MM containing arabinose, activity from the *pcaQ* promoter in the presence of a complete PcaQ polypeptide was autorepressed to less than 2% of the constitutive level found in the pARO538 control lacking PcaQ. A similar relative level of autorepression was observed under conditions for the induction of *pcaD*. However, the absolute value of β-galactosidase transcription from the *pcaQ* promoter in ADO2076 (pARO535 or pARO538) under inducing conditions was 160 to 186% the level of noninducing conditions. It is not clear what causes the stimulation of transcription under inducing conditions, but it may reflect a secondary effect of aromatics on the cells' physiology. Given the low level of stimulation, it appears that *pcaQ* is similar to other LysR family members in being expressed at a low constitutive level.

A conclusion from the results is that effector-bound PcaQ is active in the repression of the *pcaQ* gene. The high degree of autorepression observed for *pcaQ*, compared, for example, with that of other LysR-type systems (1), may be accounted for by the fact that, although multicopy plasmids provided PcaQ$^+$ or PcaQ$^-$ in *trans*, the *pcaQ*-reporter gene fusion was inserted into the chromosome and was not plasmid borne, as has been the case with the results of some analogous studies. Support for this conclusion came from studies of *pcaQ*::*lacZ* expression in *E. coli* cells. *E. coli* DH5α carrying the high-copy-number plasmid pARO76 plus the low-copy-number plasmid pARO535 (PcaQ$^+$) had 32% the level of β-galactosidase expression observed in *E. coli* DH5α (pARO76 and pARO538; PcaQ$^-$) (data not shown). Thus, the degree of autorepression was lower in a situation in which *pcaQ*::*lacZ* was present in high copy number and PcaQ synthesis was proportionately low.

PcaQ activates at the level of transcription with two coinducers, β-carboxy-*cis*,*cis*-muconate and γ-carboxymuconolactone. Catabolism of β-carboxy-*cis*,*cis*-muconate gives rise to an unstable intermediate, γ-carboxymuconolactone (Fig. 1). Evidence from *A. tumefaciens pcaC* mutant strains suggested that not only β-carboxy-*cis*,*cis*-muconate but also γ-carboxymuconolactone could act as coinducers with PcaQ. To test this directly, a transcriptional fusion between *pcaD* and *lacZ* was constructed in a pUC plasmid with the *lac* promoter deleted. The plasmid, designated pARO77, contains a complete *pcaQ* gene as well as the *pcaQ-pcaD* intergenic region (Fig. 2). pARO77 was introduced into *E. coli* DH5α cells, and the cells were exposed to aliquots of γ-carboxymuconolactone. The lactone was generated enzymatically from β-carboxy-*cis*,*cis*-muconate, and controls included additions of the components that made up the γ-carboxymuconolactone mix, including one with β-carboxy-*cis*,*cis*-muconate. The data in Table 3 show that the lactone was clearly a coinducer of *pcaD*::*lacZ*, while, as expected, carboxymuconate apparently failed to permeate the *E. coli* cells effectively. When cells exposed to the γ-carboxymuconolactone mix were resuspended in LB rather than MM, the level of induction by the lactone fell from greater than 100-fold to only 5-fold (data not shown). Since *lacZ* is joined to *pcaD* in a transcriptional fusion, the results of Table 3 also reveal that PcaQ activates at the level of transcription. This conclusion was supported by the results with another *lacZ* fusion plasmid, pARO80, in *A. tumefaciens*, described below.

Analysis of a pseudorevertant of the *pcaQ*::Ω strain ADO2044. A phenotypic revertant of ADO2044, strain ADO2048, which had low-level constitutive expression of PcaD, PcaHG, and PcaB (PcaC was not assayed), was described previously (15). Under noninducing conditions, the levels of PcaD, PcaHG, and PcaB were 32 to 39% of the specific activities found in fully induced wild-type cells. Strain ADO2048 was presumed to be a pseudorevertant because it still had the Spc^r of the omega element. The ADO2048 mutant was analyzed by sequencing and functional assay. For the latter, a broad-host-range plasmid, pARO158, was constructed such that the direction of transcription of *A. tumefaciens pcaD* from its own promoter would be opposite to the direction of transcription from the *lac* promoter of the plasmid. To monitor the activation of *pcaD* transcription, the promoterless *lacZ* cassette was fused to the beginning of the *pcaD* gene in pARO158, creating pARO80 (Fig. 2), and this plasmid was introduced into different *A. tumefaciens* strains. The expression of *pcaD*::*lacZ* in these strains relied on the ability of PcaQ transcribed from the chromosome to act in *trans* on the *pcaD* promoter.

Table 4 shows that in the parental strain A348 (pARO80), PcaQ activated in *trans* only in the presence of the coinducer; there was negligible LacZ activity with the A348 (pARO158) control. In ADO2044 (pARO80), PcaQ is truncated and no activation occurred. The results with ADO2048 (pARO80) were similar to those of ADO2044, indicating that a functional PcaQ is not present in the strain. In addition, the results indicate that the secondary mutation which causes constitutive synthesis of some enzymes of protocatechuate catabolism does not act in *trans*. This finding lends support to the hypothesis that the secondary mutation lies in a promoter that is common to the *pcaDCHGB* genes.

TABLE 4. Abilities of different *A. tumefaciens* strains to activate transcription of *pcaD*::*lacZ* in *trans*

Strain	Plasmid carried by strain	Carbon source	β-Galactosidase activity[a]
A348	pARO158	Arabinose + quinate	<1
A348	pARO80	Arabinose	92 (2.0)
A348	pARO80	Arabinose + quinate	4,794 (4.5)
ADO2044	pARO80	Arabinose	97 (1.0)
ADO2044	pARO80	Arabinose + quinate	149 (0.2)
ADO2048	pARO80	Arabinose + quinate	192 (3.0)

[a] Averages of values in Miller units from three independent cultures with standard deviations in parentheses, except for A348 (pARO158), which was done in duplicate.

Restriction analysis and sequencing of the pARO82 clone isolated from ADO2048 verified that it indeed contained *pcaQ*::Ω and the intergenic *pcaQ-pcaD* region. A transversion mutation from adenine to cytosine was identified at nucleotide −47 in the intergenic region (Fig. 3).

DISCUSSION

Verification of the *pcaQ* ORF. The insertion in plasmid pARO531 differs from that in pARO535 in lacking a *Pst*I-*Xho*I DNA segment extending from within the predicted 5′ end of *pcaQ* into *pcaD*; both insertions share a *Bgl*II end located 1.2 kb from the *Pst*I site. Only the larger pARO535 complemented a *pcaQ*::Ω mutant (14). Sequencing the conserved *pobA* gene revealed that the *Bgl*II site was located within the gene, 0.2 kb from its 3′ end (16). Thus, the results of sequence analysis showing *pcaQ* as a 0.93-kb gene are consistent with the results of the complementation studies (14).

Two metabolites serve as coinducers with PcaQ. A reversible biochemical reaction (PcaB), commercial unavailability, and instability of metabolites all complicated the study of PcaQ coinducer identity. Experiments characterizing this aspect of PcaQ were carried out with *E. coli* cells in order to avoid biochemical conversion in *A. tumefaciens*. The results from these studies were consistent with the phenotypes of *A. tumefaciens pcaHG* and *pcaC* mutants (15). The high level of PcaQ activity in *E. coli* indicates that its synthesis and function does not require a chaperone or RNA polymerase not present in *E. coli* DH5α cells under standard conditions of growth.

Previous studies with *E. coli* identified β-carboxy-*cis*,*cis*-muconate as a coinducer of PcaQ (14). Further characterization has revealed that the removal of a carboxyl group from the compound by PcaB gives rise to a second coinducer, γ-carboxymuconolactone. The carboxyl group of γ-carboxymuconolactone is highly labile. At 30°C and pH 6.0, the half-life of the lactone is approximately 30 min (12). This instability was circumvented by providing doses of the lactone at 30-min intervals, and under these conditions, the expression of β-galactosidase from the *pcaD* promoter increased over 100-fold. The exposure of *E. coli* cells carrying *pcaQ* and an agrobacterial *pcaD* promoter probe (pARO77) to the two compounds supplied exogenously demonstrated the impermeability of the cells to the tricarboxylic acid and the effectiveness of the lactone as coinducer.

The observation of dual coinducers governing enzymes of the β-ketoadipate pathway is not unprecedented. However, the work with PcaQ has demonstrated unequivocably that two catabolite coinducers mediate transcriptional activation of the *pca* pathway in association with the same regulatory protein. Two structurally dissimilar compounds, benzoate and *cis*,*cis*-muconate, were found to induce the *catA* gene from *A. calcoaceticus* (10). Both *p*-hydroxybenzoate and β-ketoadipate were observed to induce PcaB in *Pseudomonas cepacia* (31). The latter result is intriguing in light of the characterization of *pcaR* from *P. putida* (19). PcaR, which has β-ketoadipate as the sole coinducer, is a homolog of PobR from *A. calcoaceticus*, the regulator of *p*-hydroxybenzoate hydroxylase, for which *p*-hydroxybenzoate is the coinducer (2, 19).

Relationship of PcaQ to other members of the LysR family of regulatory proteins. PcaQ shares a number of characteristics with many members of the ever-enlarging LysR family of transcriptional regulator proteins (23). Its size of 311 residues falls in the range observed for other LysR-type regulators (276 to 324 residues). PcaQ is responsive to coinducers, activating transcription from the *pcaD* promoter in their presence by about 100-fold over that of the uninduced level. Transcription of *pcaQ* is divergent from the promoter activated by the PcaQ-coinducer complex, and it is negatively regulated by its gene product.

LysR-type regulators show a high degree of homology in their amino-terminal domains where the helix-turn-helix DNA-binding region is located. In a proposed consensus sequence, the helix-turn-helix motif is located 23 residues from the amino-terminal end of the polypeptide (4). PcaQ possesses a sequence homologous to this motif 23 amino acids from the initial methionine residue (Fig. 3), and it shows a high degree of conservation of the domain organization found in many LysR members (23). An area of divergence, however, lies in the carboxy-terminal domain (residues 236 to 246), where Gly-239, Gly-241, and Pro-246 are substituted with Ser-242, Ala-244, and Ser-249 in PcaQ.

Analysis of the relationship of PcaQ to other LysR-type proteins was based on a composite alignment of CatR subgroup members (1) with 21 other members of the LysR family (28). It was difficult to place PcaQ precisely in a phylogenetic tree of representative LysR-type proteins (24), except to conclude that it lies outside the NodD, SyrM, and NahR groups. Its overall amino acid sequence identity with each of these three proteins is under 16%, and the amino-terminal portions of the aligned polypeptides (residues 6 to 70 of PcaQ) show levels of identity of 25% or less. By contrast, an alignment of these 64 residues of PcaQ with some other LysR family members (1, 28) revealed closer relationships: RbcR and TcbR, 45.3%; LysR and GltC, 40.6%; and CatR, ClcR, and OxyR, 39%. Of these, only two showed an overall sequence identity of greater than 20% when aligned with PcaQ: OxyR at 22.8% and RbcR at 20.7%. The others ranged between 15.5% (CatR) and 19.3% (LysR). It was possible to align homologs with residues 6 to 59 of PcaQ without introducing any gaps. Nucleotide sequences underlying this alignment were analyzed for *catR*, *tcbR*, *rbcR*, *oxyR*, and *gltC*. The greatest nucleotide identity was found with *rbcR* (63.5%) and then with *oxyR* (57.4%), *catR* (54.3%), *tcbR* (53.7%), and *gltC* (50%).

Evidence that the *pcaDCHGB* genes share a common promoter in the *pcaQ-pcaD* intergenic region. Within the postulated *pcaDCHGB* operon controlled by *pcaQ*, only a few nucleotides separate *pcaD* and *pcaC* and there is a 1-nucleotide overlap between *pcaC* and *pcaH* (15, 16). In keeping with the tight clustering of *pcaDCHGB*, *pcaIJ*, and *pobA* genes (15, 16) and economy of intergenic nucleotides, there are 95 (or possibly 120) nucleotides in the intergenic region of *pcaQ* and *pcaD*. Strain ADO2048, which is *pcaQ*::Ω, has a secondary suppressor mutation that acts in *cis*, resulting in partially constitutive synthesis of the enzymes encoded by the *pcaDCHGB* operon. The DNA sequence of the *pcaQ-pcaD* intergenic region of this strain is consistent with the strain containing a

promoter up mutation. A point mutation in ADO2048 alters the putative −35 sequence (5′-TTTACT-3′) to 5′-TTGACT-3′, which more closely resembles the canonical −35 sequence (5′-TTGACA-3′) of σ^{70} promoters (Fig. 3). Pinpointing the mutation in ADO2048 within the intergenic region, together with other experimental results (15), has defined conclusively that the *pcaDCHGB* genes are within a common operon which is controlled by *pcaQ*.

In the *pcaD* promoter-probe plasmid pARO80, only 33 bp of *pcaQ*, 95 bp of intergenic DNA, and 5 bp presumed to be the 5′ end of *pcaD* make up the upstream agrobacterial sequence fused to the promoterless *lacZ*. Given the results with wild-type and *pcaQ* mutant *A. tumefaciens* strains carrying pARO80 under inducing and noninducing conditions, at least one binding site specific for PcaQ can be inferred to lie in the 133-bp region upstream of *lacZ*. A LysR promoter motif has been proposed (3) and supported by numerous studies (29), namely, T-N_{11}-A, with the T and A being part of a short inverted repeat. As has been noted with other LysR-type regulatory targets, such as that of RcbR (28), the intergenic region of *pcaQ* and the *pcaDCHGB* operon that it regulates is A+T-rich, having a G+C content of only 43%. In the region between −20 and −53 bp upstream of *pcaQ* (Fig. 3), the G+C content is only 21%. Within this section of DNA two sequences having characteristics of a LysR-type target consensus motif are located: TAA-7 bp-TTA, 48 nucleotides upstream of the putative *pcaD* translation start site, and TTAT-11 bp-ATAA, 61 nucleotides upstream of the site (Fig. 3). No other such motifs exist in the intergenic region, although sites are found after the *pcaD* initiation codon and one occurs 123 bp from the translational start site of *pcaQ*.

Analysis of the sequence of PcaQ has demonstrated that it is unrelated to PobR and PcaR, regulators of *p*-hydroxybenzoate and protocatechuate catabolism in bacteria of the γ subgroup of proteobacteria. The finding that another member of the α subgroup, *R. leguminosarum* biovar trifolii, has a distinctive pattern of induction by β-carboxy-*cis*,*cis*-muconate and β-ketoadipate (18) similar to that found in *A. tumefaciens*, raises the possibility that the regulators which respond to the two coeffectors have been conserved in the two species. The construction of the PcaQ target-promoter probe vector pARO80 used in the analysis of ADO2048 will facilitate a study of whether PcaQ is conserved in *R. leguminosarum* and related rhizobia, in which genes for the catabolism of protocatechuate appear to be universally present.

ACKNOWLEDGMENTS

This research was supported by Department of Energy grant DOE88ER13947.

I thank L. N. Ornston for sharing the facilities of his laboratory and β-carboxy-*cis*,*cis*-muconate and U. Gerischer and L. N. Ornston for the subclone pZR9.

REFERENCES

1. **Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty.** 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. J. Bacteriol. **175:**417–427.
2. **DiMarco, A. A., B. Averhoff, and L. N. Ornston.** 1993. Identification of the transcriptional activator *pobR* and characterization of its role in the expression of *pobA*, the structural gene for *p*-hydroxybenzoate hydroxylase in *Acinetobacter calcoaceticus*. J. Bacteriol. **175:**4499–4506.
3. **Goethals, K., M. Van Montagu, and M. Holsters.** 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. Proc. Natl. Acad. Sci. USA **89:**1646–1650.
4. **Henikoff, S., G. W. Haughn, J. M. Calva, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85:**6602–6606.
5. **Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger.** 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene **70:**191–197.
6. **Knauf, V. C., and E. W. Nester.** 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid **8:**45–54.
7. **Kokotek, W., and W. Lotz.** 1989. Construction of a *lacZ*-kanamycin-resistance cassette, useful for site-directed mutagenesis and as a promoter probe. Gene **84:**467–471.
8. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Neidle, E. L., C. Hartnett, and L. N. Ornston.** 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. **170:**4874–4880.
10. **Neidle, E. L., and L. N. Ornston.** 1987. Benzoate and muconate, structurally dissimilar metabolites, induce expression of *catA* in *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**414–415.
11. **Normone, W. M., and J. R. Brown.** 1970. Guanine plus cytosine (G+C) composition of bacteria, p. H24–H74. *In* H. A. Sober (ed.), CRC handbook of biochemistry: selected data for molecular biology, 2nd ed. The Chemical Rubber Co., Cleveland.
12. **Ornston, L. N., and R. Y. Stanier.** 1966. The conversion of catechol and protocatechuate to β-ketoadipate by *Pseudomonas putida*. I. Biochemistry. J. Biol. Chem. **241:**3776–3786.
13. **Parke, D.** 1992. Application of *p*-toluidine in chromogenic detection of catechol and protocatechuate, diphenolic intermediates in catabolism of aromatic compounds. Appl. Environ. Microbiol. **58:**2694–2697.
14. **Parke, D.** 1993. Positive regulation of phenolic catabolism in *Agrobacterium tumefaciens* by the *pcaQ* gene in response to β-carboxy-*cis*,*cis*-muconate. J. Bacteriol. **175:**3529–3535.
15. **Parke, D.** 1995. Supraoperonic clustering of *pca* genes for catabolism of the phenolic compound protocatechuate, in *Agrobacterium tumefaciens*. J. Bacteriol. **177:**3808–3817.
16. **Parke, D.** Unpublished data.
17. **Parke, D., and L. N. Ornston.** 1984. Nutritional diversity of *Rhizobiaceae* revealed by auxanography. J. Gen. Microbiol. **130:**1743–1750.
18. **Parke, D., F. Rynne, and A. Glenn.** 1991. Regulation of phenolic catabolism in *Rhizobium leguminosarum* biovar trifolii. J. Bacteriol. **173:**5546–5550.
19. **Romero-Steiner, S., R. E. Parales, C. S. Harwood, and J. E. Houghton.** 1994. Characterization of the *pcaR* regulatory gene from *Pseudomonas putida*, which is required for the complete degradation of *p*-hydroxybenzoate. J. Bacteriol. **176:**5771–5779.
20. **Rothmel, R. K., T. L. Aldrich, J. E. Houghton, W. M. Coco, L. N. Ornston, and A. M. Chakrabarty.** 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the LysR family. J. Bacteriol. **172:**922–931.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
23. **Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47:**597–626.
24. **Schlaman, H. R. M., R. J. H. Okker, and B. J. J. Lugtenberg.** 1992. Regulation of nodulation gene expression by NodD in rhizobia. J. Bacteriol. **174:**5177–5182.
25. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Bio/Technology **1:**784–791.
27. **Stanier, R. Y., and L. N. Ornston.** 1973. The β-ketoadipate pathway, p. 89–151. *In* A. H. Rose and D. W. Tempest (ed.), Advances in microbial physiology, vol. 9. Academic Press, London.
28. **Viale, A. M., H. Kobayashi, T. Akazawa, and S. Henikoff.** 1991. *rcbR*, a gene coding for a member of the LysR family of transcriptional regulators, is located upstream of the expressed set of ribulose 1,5-bisphosphate carboxylase/oxygenase genes in the photosynthetic bacterium *Chromatium vinosum*. J. Bacteriol. **173:**5224–5229.
29. **Von Lintig, J., D. Kreusch, and J. Schroder.** 1994. Opine-regulated promoters and LysR-type regulators in the nopaline (*noc*) and octopine (*occ*) catabolic regions of Ti plasmids of *Agrobacterium tumefaciens*. J. Bacteriol. **176:**495–503.
30. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
31. **Zylstra, G. J., R. H. Olsen, and D. P. Ballou.** 1989. Cloning, expression, and regulation of the *Pseudomonas cepacia* protocatechuate 3,4-dioxygenase genes. J. Bacteriol. **171:**5907–5914.