Journal of Bacteriology, Feb. 1990, p. 922–931
0021-9193/90/020922-10$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 2

Nucleotide Sequencing and Characterization of *Pseudomonas putida* *catR*: a Positive Regulator of the *catBC* Operon Is a Member of the LysR Family

RANDI KUBRICK ROTHMEL,[1] TERI L. ALDRICH,[1]† JOHN E. HOUGHTON,[2] WAYNE M. COCO,[1] L. NICHOLAS ORNSTON,[2] AND A. M. CHAKRABARTY[1]*

Department of Microbiology and Immunology, University of Illinois College of Medicine, Chicago, Illinois 60612,[1] and Department of Biology, Yale University, New Haven, Connecticut 06511[2]

Received 9 August 1989/Accepted 17 November 1989

Pseudomonas putida utilizes the _catBC_ operon for growth on benzoate as a sole carbon source. This operon is positively regulated by the CatR protein, which is encoded from a gene divergently oriented from the _catBC_ operon. The _catR_ gene encodes a 32.2-kilodalton polypeptide that binds to the _catBC_ promoter region in the presence or absence of the inducer _cis-cis_-muconate, as shown by gel retardation studies. However, the inducer is required for transcriptional activation of the _catBC_ operon. The _catR_ promoter has been localized to a 385-base-pair fragment by using the broad-host-range promoter-probe vector pKT240. This fragment also contains the _catBC_ promoter whose −35 site is separated by only 36 nucleotides from the predicted CatR translational start. Dot blot analysis suggests that CatR binding to this dual promoter-control region, in addition to inducing the _catBC_ operon, may also regulate its own expression. Data from a computer homology search using the predicted amino acid sequence of CatR, deduced from the DNA sequence, showed CatR to be a member of a large class of procaryotic regulatory proteins designated the LysR family. Striking homology was seen between CatR and a putative regulatory protein, TfdS.

Pseudomonads can detoxify chemical wastes because of their ability to utilize many natural and synthetic compounds. Pure cultures capable of dissimilating simple nonchlorinated and chlorinated compounds have been isolated (9, 12, 16, 20, 34). In order to expand the substrate ranges of these organisms so that they can degrade more complex and toxic compounds, it is necessary to understand both the structural and regulatory features of catabolic genes as well as their regulation. While some *Pseudomonas putida* genes that encode enzymes for the dissimilation of benzoate and 3-chlorobenzoate have been fairly well characterized (1, 2, 17, 19, 49), the structures and functions of the regulatory genes involved are not well understood.

The *P. putida catBC* operon provides a good model for examining the regulation of catabolic genes. The *catB* and *catC* genes encode *cis,cis*-muconate-lactonizing enzyme I (EC 5.5.1.1.) and muconolactone isomerase (EC 5.3.3.4.), respectively. Both of these genes are required for the dissimilation of benzoate (34) (Fig. 1). The *catBC* operon is coordinately regulated and requires the product of its regulatory gene for induction (48). This regulatory gene, *catR*, also exerts positive control over the expression of the *catA*-encoded pyrocatechase I. The *catR* gene maps upstream of the *catBC* operon in *P. putida* PRS2000 (48, 50; J. E. Houghton, E. J. Hughes, and L. N. Ornston, Abstr. Annu. Meet. Am. Soc. Microbiol. 1988, K4, p. 207). Insertion of Tn*5* in this region, creating *P. putida* PRS3026, results in the inability of the bacteria to grow on benzoate as a sole carbon source (Houghton et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1988).

Previously, we reported on the cloning of the *catBC* operon from *P. putida* RB1 (1, 2). This clone, pTA3, contains 257 nucleotides upstream of the structural gene for *catB*. Sequencing of a subclone containing *catB* and upstream DNA revealed a truncated protein of 40 amino acids encoded divergently from *catB*. Two independent computer searches (S. Henikoff, personal communication; I. P. Crawford, personal communication) predicted that this region encodes a potential regulatory gene having considerable homology to the N terminus of a number of regulatory proteins from various gram-negative bacteria, including *Enterobacter cloacea* AmpR (24), *Escherichia coli* LysR (43), *Rhizobium* sp. NodD (14), *E. coli* and *Salmonella typhimurium* CysB (35), *S. typhimurium* MetR (38), *Pseudomonas aeruginosa* TrpI (10), and *P. putida* NahR (42). All of these regulatory proteins have been grouped into a large family of bacterial regulatory proteins, the LysR family (21).

A number of similarities exist between these proteins, including their size of approximately 300 amino acids. Comparisons of protein sequences show strong homologies within the N terminus, which includes a helix-turn-helix motif implicated in DNA binding. The homologies drop off at the C-terminal region and the mid-region, which may be the region involved in interaction with the inducer molecule. To determine whether *catR* is a member of the LysR family, we have cloned the gene from *P. putida* RB1 and compared its DNA and amino acid sequence with those of other family members. This comparison indicated that CatR is a LysR family member that is closely related to a second predicted regulatory protein, TfdO or TfdS (E. J. Perkins, personal communication; B. Kaphammer and R. H. Olsen, Abstr. Annu. Meet. Am. Soc. Microbiol. 1989, K-59, p. 254). Our evidence presented in this work indicates that like other divergently translated LysR proteins, CatR binds in *trans* to the *catR-catBC* promoter-control region in the presence or absence of inducer but only activates the *catBC* operon in the presence of the inducer, *cis-cis*-muconate. Localization of the *catR* promoter and the proximity between the CatR

* Corresponding author.
† Present address: Department of Genetics, University of Washington, Seattle, WA 98195.

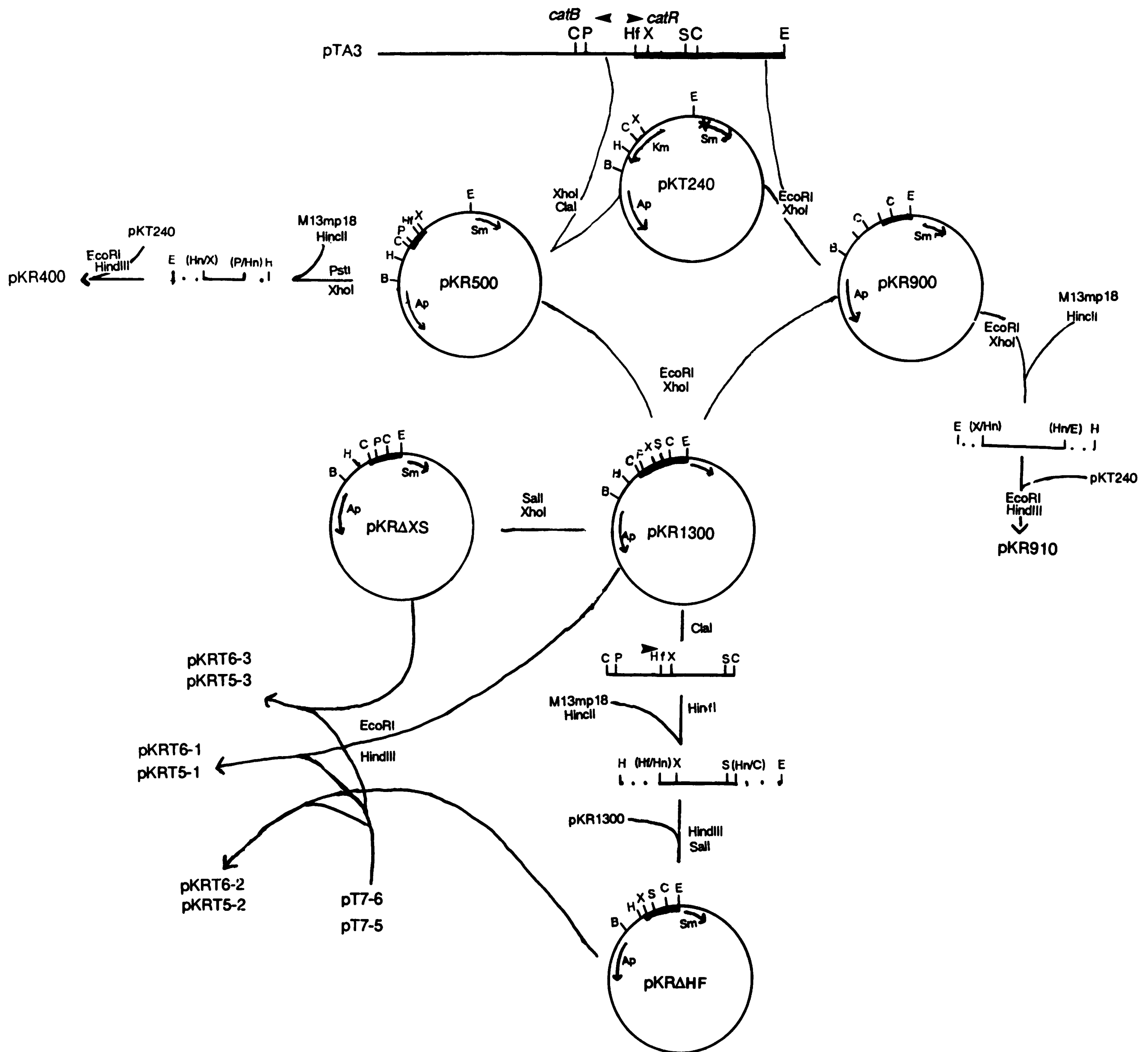

FIG. 1. Cloning of the *catR* constructs used in this study. Restriction sites are abbreviated as follows: B, *Bam*HI; C, *Cla*I; E, *Eco*RI; Hf, *Hin*fI; H, *Hin*dIII; Hn, *Hin*cII; P, *Pst*I; S, *Sal*I; X, *Xho*I. Only the restriction sites used in cloning are shown. There are multiple sites for *Hin*fI, *Pst*I, and *Sal*I in the recombinant plasmid pTA3. The directions of the *catR* and *catB* genes are indicated by the arrowheads. The thick line indicates the *catR* gene and the inserted DNA in each of the clones. The thin lines in the plasmid constructs are vector DNA. ....., Multiple cloning site of M13mp18. Asterisks indicate that the aminoglycoside phosphotransferase gene lacks its own promoter and that streptomycin resistance is seen only when a promoter is cloned upstream. Restriction sites that were lost during cloning are in parentheses.

translational start and the *catBC* promoter suggest that the two promoters overlap one another.

MATERIALS AND METHODS

Bacteria, plasmids, bacteriophage, and media. The strains used in this study are listed in Table 1. Cultures were grown on either Luria broth or basal salts medium (BSM) supplemented with glucose and/or benzoate as described previously (2). Antibiotics were added as required for selection as follows: ampicillin, 75 mg/liter; kanamycin, 50 mg/liter. Streptomycin was added at varying concentrations ranging between 100 and 8,000 mg/liter.

Nucleic acid preparations and DNA sequencing. Plasmid DNA was isolated from 40-ml cultures by alkaline lysis as described in Maniatis et al. (29) for a large-scale preparation, except that after isopropanol precipitation the DNA was phenol-chloroform extracted twice, precipitated with ethanol, and treated with RNase. RNA was isolated from heat-induced *E. coli* K38 (pGP1-2) by the guanidinium-isothiocyanate, hot-phenol extraction method as described previously (3). DNA sequencing was carried out by using the previously described dideoxy analysis (3, 31).

Molecular cloning. Restriction enzymes were purchased from Bethesda Research Laboratories, Gaithersburg, Md.,

TABLE 1. Bacterial strains, plasmids, and bacteriophages

Strain, plasmid, or phage	Relevant characteristics	Source or reference
E. coli		
C600	*F⁻ thi-1 thr-l leuB6 lacY1 tonA21 supE44* λ⁻	4
HB101	*F⁻ hsdS20* (r⁻ m⁻) *recA13 ara-14 proA2 lacY1 galK2 poL20 xyl-5 mtl-1 supE44* λ⁻	29
JM109	Δ(*pro-lac*) *recA1 thi-1 supE endA gyrA96 hsdR relA1* (F' *traD36 proAB lacI*ᑫ *lacZ*Δ*M15*)	51
K38		Stan Tabor
P. putida		
PRS2000	*per-1103*	48
PRS3026	*per-1103 catR*::Tn*5* Km^r	E. J. Hughes
Plasmids and phages		
pKT240	Km^r Ap^r	5
pRK2013	Km^r *mob*⁺	15
pTA3	Ap^r *catBC*⁺ ben⁺	2
pTAΔXH	Ap^r p_{catBC}	3
M13mp18	p_{lac} *lacZ*α⁺	33
M13mp19	p_{lac} *lacZ*α⁺	33
pT7-5	Ap^r	Stan Tabor
pT7-6	Ap^r	Stan Tabor
pKR400	Ap^r Km^r p_{catR}	This study
pKR500	Ap^r Km^r p_{catR}	This study
pKR900 and pKR910	Ap^r Km^r	This study
pKR1300	Ap^r Km^r CatR⁺	This study
pKRΔHf	Ap^r Km^r	This study
pKRT5-1 and pKRT6-1	Ap^r	This study
pKRT5-2 and pKRT6-2	Ap^r	This study
pKRT5-3 and pKRT6-3	Ap^r	This study

and were used in accordance with the instructions of the manufacturer. Vector DNA was treated with calf intestinal alkaline phosphatase purchased from Boehringer Mannheim Biochemicals, Indianapolis, Ind. Ligations using T4 DNA ligase (New England BioLabs, Beverly, Mass.) were done at 15°C overnight. DNA was transformed into *E. coli* C600 or JM109 cells and screened by mini-prep analysis (22). *P. putida* PRS3026, a *catR* mutant that is unable to grow on benzoate as a sole carbon source (Houghton et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1988) was used for complementation tests as described previously (2). Broad-host-range plasmids were mated into *P. putida* strains, and exconjugants were selected by antibiotic resistance as described previously (1, 2).

The *catR* gene was subcloned from the plasmid pTA3 (2) into the broad-host range vector pKT240 (5) in two subsequent ligations (Fig. 1). First, a 480-base-pair (bp) *Xho*I-*Cla*I fragment from pTA3 that overlapped the *catB* gene was ligated into pKT240, creating pKR500 (Fig. 1). Second, the upstream 816-bp *Eco*RI-*Xho*I fragment used to construct pKR900 and pKR910 was ligated into pKR500 to give pKR1300. Construction of the pT7-5 and pT7-6 clones used for protein analysis is shown in Fig. 1. The plasmids pKRT5-1 and pKRT6-1 were generated by cloning the 1,700-bp *Eco*RI-*Hin*dIII fragment from pKR1300 containing the entire *catR* gene, including the upstream *catR* promoter and control region (Fig. 1). pKRT5-2 and pKRT6-2 carry a 1,000-bp insert in which the *catR* promoter and control region have been deleted. The deletion was constructed (Fig. 1) as follows. An 830-bp *Cla*I fragment was isolated from pKR1300 and recut with *Hin*fI (located 15 bases upstream of the *catR* Shine-Dalgarno site). The ends were then filled in and blunt end ligated into M13mp18 at the *Hin*cII site. DNA was isolated from the mixed phage clones, and a 300-bp *Hin*dIII-*Sal*I fragment which contained the N-terminal region of the structural *catR* gene without any promoter region was isolated. The C terminus was regenerated by cloning this fragment into pKR1300 at the *Hin*dIII-*Sal*I site to give pKRΔHf. The 1,000-bp *Eco*RI-*Hin*dIII fragment from pKRΔHf was cloned into pT7-5 and pT7-6, creating pKRT5-2 and pKRT6-2. The final *catR* pT7 constructs, pKRT5-3 and pKRT6-3 (Fig. 1), simply contained a 150-bp *Xho*I-*Sal*I deletion internal to the predicted protein sequence.

Protein expression. A T7 expression system developed by Stan Tabor (personal communication) was used to express CatR. The three *catR* constructs cloned into the vectors pT7-5 and pT7-6 (see above) were transformed into *E. coli* K38 (pGP1-2) cells. Protein was expressed from these cells by using a heat-inducible T7 RNA polymerase gene located on the plasmid pGP1-2 (45). The [^{35}S]methionine-labeled polypeptides were run on a 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel by the method of Laemmli (26).

Binding studies. Cell protein extracts from *E. coli* K38 (pGP1-2) harboring the pT7-5 or pT7-6 clones were made from 100-ml cultures grown in Luria broth for 3 h at 30°C, heat-induced (42°C) for 45 min, and grown at 30°C for an additional hour. Cells were harvested at 8,000 × *g* for 10 min, suspended to 5 ml in a buffer consisting of 50 mM Tris hydrochloride (pH 8.0)–5 mM dithiothreitol–1 mM phenylmethylsulfonyl fluoride, and sonicated for 2 min with 30-s bursts. The lysate was centrifuged at 40,000 × *g* for 30 min. The supernatant was centrifuged and spun at 105,000 × *g* for 1 h, and the resulting supernatant was stored at −70°C in 50% glycerol. Protein content was determined by the method of Bradford (7). Gel retardation studies were performed by the method of Fried and Crothers (18). A 485-bp DNA fragment containing the *catRBC* promoter control region was used for the gel retardation studies. This fragment was labeled by primer extension by using single-stranded DNA isolated from the M13 clone TAΔXH:18 (3) and [α-^{32}P]dCTP. The double-stranded DNA fragment was isolated as a *Hin*dIII-*Eco*RI fragment. *cis,cis*-muconate (a generous gift from Celgene Corp.) was added to binding reactions at a final concentration of 2 mM.

Dot blot analysis. Amersham nylon membrane (Hybond-M) was used as directed for analysis of mRNA isolated from *E. coli* K38 (pGP1-2) harboring the pT7-6 clones. The *catR* probe was labeled by primer extension of an M13mp19 clone with [α-^{32}P]dCTP. The *catR* insert was cut out as a *Hin*dIII-*Eco*RI fragment and purified on a 1% agarose gel.

RESULTS

Cloning and sequencing *catR*. Wheelis and Ornston (48) showed that the *catBC* operon required the product of a regulatory gene for induction. Recently, this regulatory gene, *catR*, was cloned from upstream of the *catBC* operon in *P. putida* PRS2000 (Houghton et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1988). To determine the location of *catR* in *P. putida* RB1, a previously cloned 6.8-kilobase *Eco*RI-*Hin*dIII chromosomal segment from *P. putida* RB1, designated pTA3 (2), which was capable of complementing both *catB* and *catC* mutations, was tested for its ability to

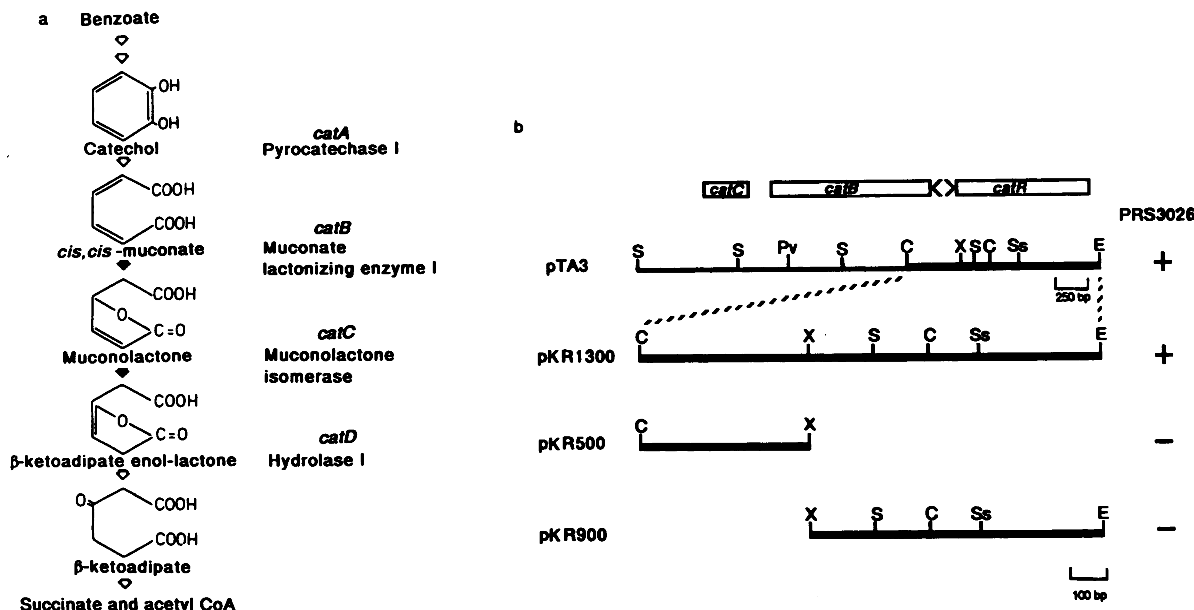

FIG. 2. (a) Catabolic pathway of benzoate degradation. The genes and corresponding enzymes for utilization of catechol are indicated. Filled arrows indicate the two enzymatic steps encoded by the *catBC* operon. (b) The organization of the *catR* and *catBC* operons and subcloning of *catR*. Coding regions are indicated (▭), as is the direction of transcription (< and >). Positive (+) and negative (−) complementation of the *catR* mutation in PRS3026 by the recombinant plasmids are shown. Restriction sites are abbreviated as given in the legend to Fig. 1; others are Pv, *Pvu*II; Ss, *Sst*II.

complement the *catR* mutant PRS3026. Growth of exconjugants on benzoate as a sole source of carbon confirmed that *catR* was encoded on pTA3 (Fig. 2). Subcloning this region resulted in locating *catR* on a 1,300-bp *Eco*RI-*Cla*I fragment upstream of *catB*. This 1,300-bp fragment, when cloned into pKT240, fully complemented the *catR* mutation (Fig. 2). pKT240 vectors containing the 816-bp *Eco*RI-*Xho*I (pKR900) or the 480-bp *Xho*I-*Cla*I (pKR500) fragment were unable to complement PRS3026, confirming that *catR* extends into both of these fragments.

The DNA sequence of the 1,300-bp region was determined, and the results are shown in Fig. 3. The nucleotide composition of this fragment is 66% G+C, which is very typical for *P. putida* (32). Translation of the DNA sequence in all six reading frames indicated that there is a probable open reading frame of 867 nucleotides in a divergent orientation from the *catB* gene (Fig. 2). The predicted *catR* Shine-Dalgarno sequence (GGAGG) is located 114 bp upstream of the Shine-Dalgarno sequence (GGA) of the *catBC* regulon. The total distance between these two translational start sites is thus 140 nucleotides. This predicted open reading frame potentially encodes a polypeptide of 32,200 daltons (Da) and corresponds to the protein molecular mass observed on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel from minicell analysis (see below). It is interesting that another large open reading frame (1,377 bp) was predicted in the same orientation and in the same reading frame as *catB*. This open reading frame would start 676 nucleotides (bp 741 in the *catR* sequence [Fig. 3]) upstream from the *catBC* transcriptional start. However, there is no obvious Shine-Dalgarno sequence for this open reading frame, and no protein was observed in this orientation in *E. coli* minicells (below). It is therefore likely that this 1,300-bp region codes for only one polypeptide divergently translated from the *catBC* operon.

Localization and transcriptional activation of the *catR* promoter. The promoter of *catR* was localized to the 480-bp *Xho*I-*Cla*I fragment by determining the level of streptomycin resistance of *P. putida* PRS2000 harboring the pKT240 promoter probe vectors pKR500, pKR1300, pKR900, and pKR910. It has been previously shown (1, 3) that cloning a fragment containing promoter activity upstream of the promoterless aminoglycoside phosphotransferase gene present in pKT240 will allow the host cells to grow with high concentrations of streptomycin. Only pKR500 and pKR1300 conferred resistance to 2,000 μg of streptomycin to PRS2000 on BSM-benzoate plates. Both of these vectors have the 480-bp *Xho*I-*Cla*I fragment (Fig. 1), which contains the *catR*/*catBC* promoter region, cloned upstream of the promoterless aminoglycoside phosphotransferase gene in the orientation allowing transcription from the *catR* promoter. Previous results (1, 3) showed that cloning this fragment in the opposite orientation (pTAΔXH) directing the *catBC* promoter 5′ to the aminoglycoside phosphotransferase gene resulted in PRS2000 cells conferring streptomycin resistance up to 8,000 μg/ml on BSM-benzoate plates. pKR500 was further subcloned as a 385-bp *Xho*I-*Pst*I fragment, designated pKR400, which also allowed the PRS2000 host to grow in the presence of 2,000 μg of streptomycin per ml on BSM-benzoate plates.

On BSM-glucose plates, PRS2000 cells harboring pKR500 or pKR400 were unable to grow with streptomycin concentrations greater than 250 μg/ml. In addition, pKR500 and pKR400, like pTAΔXH, were unable to confer streptomycin resistance to *E. coli* C600 cells, which normally lack *catR*. Neither pKR900 nor pKR910, with the 816-bp *Eco*RI-*Xho*I fragment cloned in either orientation, was able to confer resistance to more than 250 μg of streptomycin per ml to PRS2000 on BSM-benzoate or BSM-glucose plates. None of the constructs enabled PRS3026 to grow on BSM-glucose or BSM-benzoate plates containing more than 500 μg of streptomycin per ml when mobilized into the *catR* mutant PRS3026.

Expression of CatR. Several *E. coli* T7 expression vectors

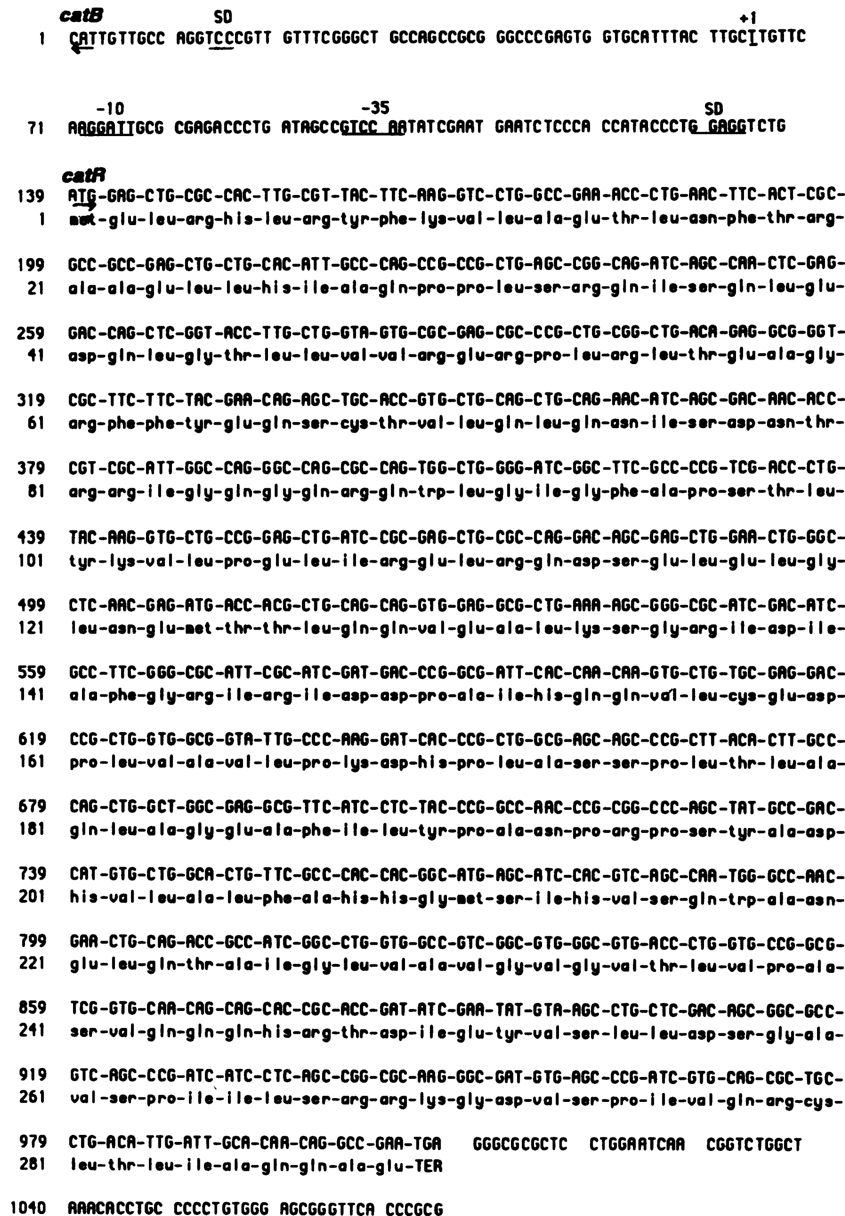

FIG. 3. Nucleotide sequence and translation of *catR*. The top number in each line indicates the nucleotide number starting with the AUG of *catB*; the bottom number indicates the amino acid number of the CatR protein. The promoter region of *catBC* operon is indicated by the underlined −35 and −10 region. The transcriptional start site of *catBC* at +1 is also underlined. The Shine-Dalgarno sequences for both the *catB* and *catR* genes are indicated, and the initiation codons are shown (→ and ←).

were constructed to analyze both CatR production and to determine the role of the control region in *catR* expression (Fig. 4). The results of the minicell analysis (Fig. 4) indicated that *catR* does encode for a polypeptide of approximately 30 kDa, which is in agreement with the predicted size by DNA sequence analysis (32.2 kDa). The *catR* gene is expressed in the opposite orientation from the *catBC* operon, as CatR-specific protein is detected only from the pT7-6 constructs and not from pT7-5 constructs. In addition, when the *catRBC* insert from pTA3 was used, maxicell analysis showed only the expected production of CatB and CatC (unpublished results). This indicated that the predicted open reading frame starting from within the *catR* gene and reading through *catB* does not code for a functional protein. The two other protein bands seen on the gel, in the pKRT6-2 lane (Fig. 4), likely correspond to a readthrough product and a truncated protein, both artifacts of expressing a *P. putida* gene in an *E. coli* system. It is possible that a hairpin structure forms in the RNA with the sequence CGGCGTCG GTGCAAC AGCAGCUCCGCUCCG (the underlined ribonucleotides indicate potential secondary structure) located in the C-terminal region 5′ to Pro-239. This high-GC-content hairpin region may be difficult for the *E. coli* transcriptional machinery to read, resulting in a truncated protein. Very little *catR* gene product can be detected from the pKRT6-1 clone, which contains the structural gene of *catR* as well as the *catR*/*catBC* promoter control region. This interference in protein production suggests that the control region plays a role in autoregulation of the *catR* gene.

Binding studies and autoregulation. In order to determine

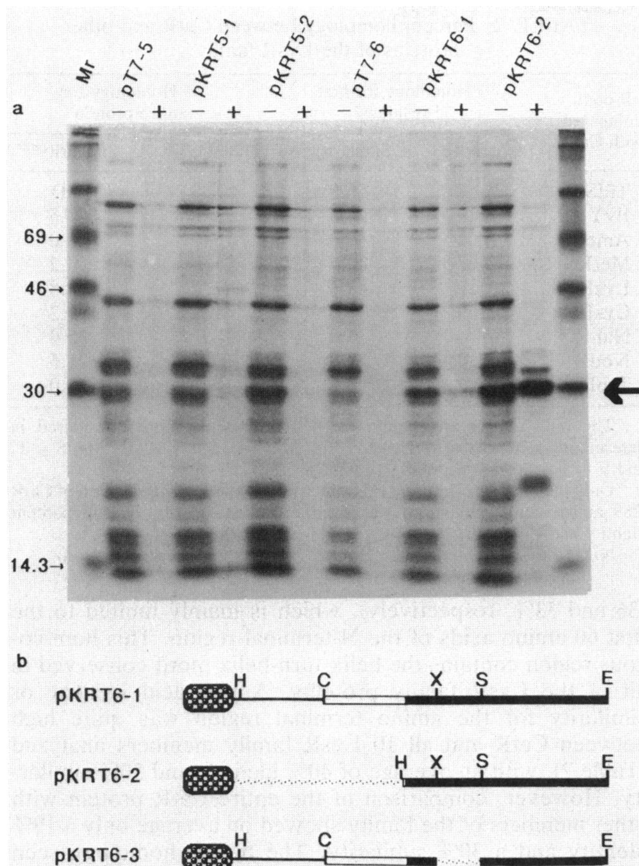

FIG. 4. (a) Minicell analysis using pT7-5 and pT7-6 vectors and *catR* clones without (−) and with (+) heat induction as described in the text. Molecular mass markers are shown on the left (in kilodaltons). Bold arrow on the right shows the approximately 30-kDa major protein product produced from the *catR* gene from induced cells. (b) Restriction map of *catR* clones used in minicell analysis as well as binding studies and dot blot analysis. Symbols: ▩, T7 Φ-10 promoter; ——, pKT240 vector DNA from cloning; ☐, *catRBC* promoter control region; ■, structural gene for *catR*; ⋯⋯, deletions. Abbreviations for restriction sites are the same as in Fig. 1.

whether the 30-kDa polypeptide expressed from pKRT6-2 binds with the *catR*/*catBC* control region, gel retardation experiments were conducted. The results strongly suggest that the CatR protein binds the promoter control region, as indicated by the retarded band seen only in the binding reactions containing the CatR protein (Fig. 5a and b). The amount of binding was greatly reduced when extracts from cells harboring pKRT6-1 were used, even though the protein concentration was equivalent in each sample. This supports the notion that *catR* gene expression is autoregulated by the promoter control region. As expected, no binding was seen with extracts from cells harboring pKRT6-2, a deletion mutant of *catR* (Fig. 3). CatR specifically binds the 485-bp control region, as shown by competition binding reactions (Fig. 5b). When nonradioactive supercoiled plasmid pKR500, which contains the control region, was added to the binding reaction in excess of the ^{32}P-labeled DNA, the amount of protein bound to the radioactive probe was greatly reduced. This effect was not seen when cold vector DNA was added to the reaction (results not shown). In addition, CatR was able to bind the control region with or without the inducer *cis*,*cis*-muconate (results not shown).

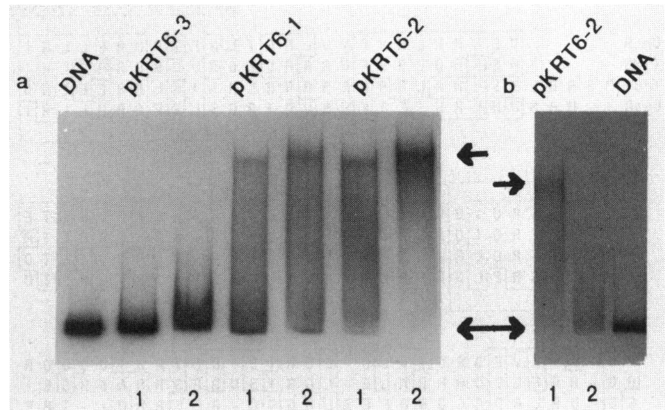

FIG. 5. (a) Binding analysis using *E. coli* extracts from cells harboring the *catR* clones depicted in Fig. 3. The lane labeled DNA is the control with no crude extract added. Lane 1 in each set shows the result when the probe was incubated with 5 μg of a crude extract, as indicated at the top, and lane 2 shows the result obtained with 10 μg of the same crude cell extract. (b) Lanes 1, Result obtained with 10 μg of crude cell extract; lanes 2, result obtained by incubating the probe with both 10 μg of crude cell extract and a 15-fold excess of cold supercoiled DNA containing the *catRBC* promoter-control region. In both panels the lower arrow indicates the unretarded band and the upper arrows indicate the gel-retarded band.

The presence of *cis*,*cis*-muconate apparently does not affect the level or amount of CatR binding.

As mentioned above, *catR* expression may be autoregulated by the interaction of CatR at the promoter control region. Binding CatR to this region affects the level of mRNA synthesis, as seen by the dot blot analysis (Fig. 6). RNA isolated from cells harboring pT7-6, pKRT6-1, pKRT6-2, or pKRT6-3 was probed with ^{32}P-labeled *catR*-specific DNA. The amount of *catR* mRNA transcribed from pKRT6-2 appeared to be much higher than that transcribed from pKRT6-1. Analysis of transcription in the absence of a

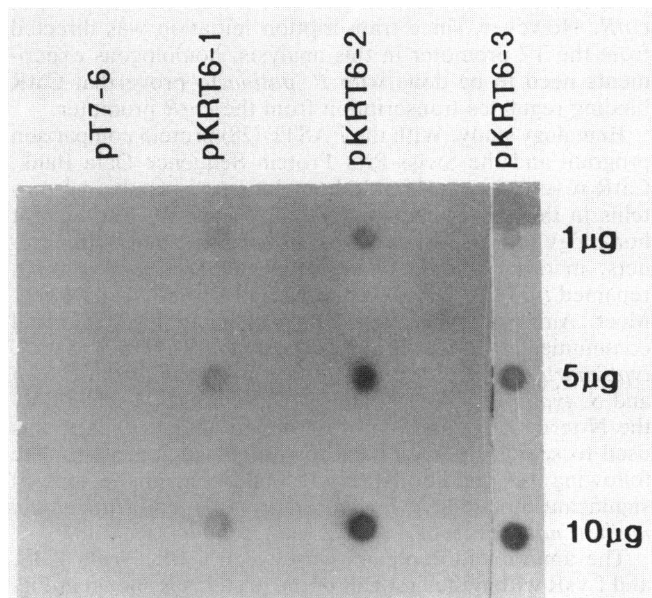

FIG. 6. Dot blot analysis. The amount of RNA probed for each clone is indicated. The 816-bp *Eco*RI-*Xho*I fragment from pKR900 was labeled with ^{32}P and used as a probe.

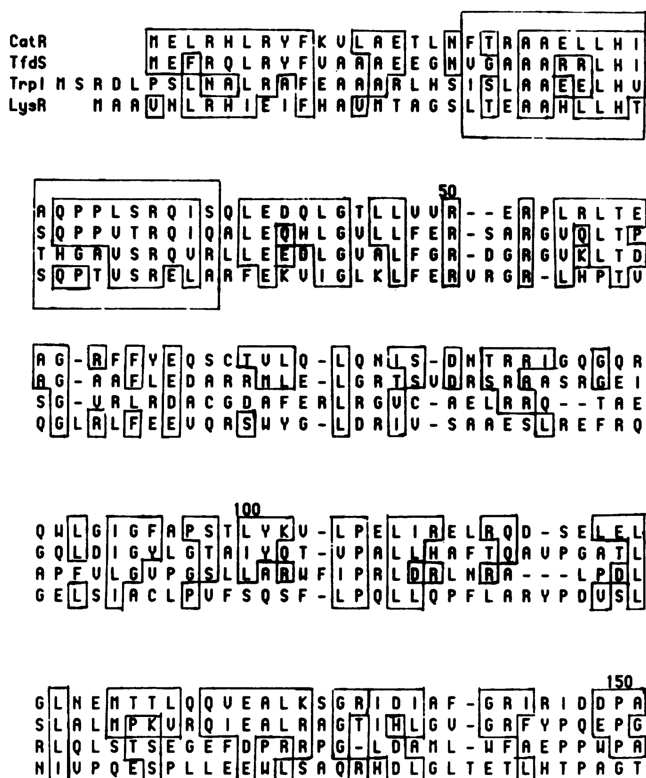

FIG. 7. Amino acid homology between CatR, TfdS, LysR, and TrpI for the first 180 amino acid residues. Boxed areas indicate amino acid similarities. The large boxed area in the amino-terminal region (top) indicates the helix-turn-helix motif. The only conservative amino acid changes that were allowed were L = I = V = M, D = E, R = K, H = Q, S = T, and F = Y.

TABLE 2. Percent homology between CatR and other proteins of the LysR family

Protein compared with CatR	% Homology for first 75 residues		% Homology for entire protein	
	Identity	Similarity[a]	Identity[b]	Similarity[a,b]
TfdS	49.3	56.0	ND[c]	ND
IlvY	48.0	57.3	23.9	31.5
AmpR	31.0	40.0	16.6	28.0
MetR	34.7	42.7	23.5	33.2
LysR	34.7	44.0	23.9	30.8
CysB	36.0	48.0	21.0	33.2
NahR	24.0	36.0	15.0	25.0
NodD	21.0	30.0	13.8	21.5
TrpI	30.6	45.3	15.0	25.0

[a] The only conservative amino acid changes that were considered in determining similarity were L = I = V = M, D = E, R = K, H = Q, S = T, and F = Y.

[b] Only the number of amino acids which overlapped with the length of CatR (289 amino acids maximum) was used in calculating the overall percent identity and similarity.

[c] ND, Not done (C-terminal region of TfdS has not been completed).

functional CatR protein, using pKRT6-3, indicated that mRNA production is nearly equal to that from pKRT6-2. These results confirm that CatR binding to the *catR/catBC* promoter control region interferes with the transcription of *catR*. However, since transcription initiation was directed from the T7 promoter in this analysis, homologous experiments need to be done with *P. putida* to prove that CatR binding regulates transcription from the *catR* promoter.

Homology study. With the FASTP (28) protein comparison program and the Swiss-Prot Protein Sequence Data Bank, CatR revealed considerable homology to a number of proteins in the LysR regulatory family (Table 2). The highest homology scores included the following seven protein products, in order of homology: *Alcaligenes eutrophus tfdO*, renamed *tfdS* (37, 44; Kaphammer and Olsen, Abstr. Annu. Meet. Am. Soc. Microbiol. 1989); E. J. Perkins, personal communication); *E. coli ilvY* (46); *E. cloacea ampR* (24); *S. typhimurium metR* (38); *E. coli lysR* (43); *cysB* from *E. coli* and *S. typhimurium* (35); and *E. coli leuO* (21). When only the N-terminal region (50 or 75 amino acids) of CatR was used to search the data bank for homologous proteins, the following two additional LysR family members showed significant homology: *E. coli antO* (21) and *Rhizobium meliloti nodD* (14).

The amino acid comparison between CatR, TrpI, TfdS, and LysR within the first half of the proteins is shown in Fig. 7. There is considerable homology, about 52%, between the first 150 amino acids of the CatR protein and that of TfdS. This contrasts with the homology seen with LysR and TrpI (36 and 33%, respectively), which is mainly limited to the first 60 amino acids of the N-terminal region. This homologous region contains the helix-turn-helix motif conserved in all of the LysR family proteins. Amino acid identity or similarity for the amino terminal region was quite high between CatR and all 10 LysR family members analyzed (Table 2), with an average of 40% identity and 50% similarity. However, comparison of the entire CatR protein with other members of the family showed on average only a 19% identity and a 30% similarity. The higher homology seen between CatR and TfdS within the first 150 amino acids compared with other members of the family suggests that these proteins evolved from a common ancestor.

DISCUSSION

A regulatory gene product, designated CatR as proposed by Wu et al. (50), is required for transcriptional activation of the *catBC* operon in the presence of the inducer *cis,cis*-muconate. In this paper, we describe the cloning of *catR*, which was localized to a 1,300-bp region upstream of the *catBC* operon by its ability to complement the *catR* mutant PRS3026. From DNA sequence analysis we predicted that the *catR* gene encoded a protein of 32.2 kDa translated in a divergent orientation from the *catBC* operon, and this was confirmed by minicell analysis. The amino acid sequence of the CatR protein showed significant homology to a number of other regulatory proteins of the LysR regulatory family (21).

Members of the LysR regulatory family have several common features. All are approximately 30 kDa in size and have a helix-turn-helix motif in the N-terminal region implicated in binding DNA sequences (13, 36). These regulatory proteins are often translated in a divergent orientation from the operons they regulate, two exceptions being CysB from *E. coli* and *S. typhimurium* (35) and the recently reported OxyR from *E. coli* and *S. typhimurium* (11). Most appear to be positive regulators, and many have been shown to regulate their own expression (11, 27, 30, 40, 41, 43, 47). This mode of gene regulation does not appear to be limited to gram-negative bacteria, as a LysR family member (MleR) required for malolactic fermentation has been isolated from the gram-positive bacterium *Lactococcus lactis* (39). This implies that this family of regulatory proteins represents a very general mechanism for gene regulation.

Although the precise molecular mechanism for gene regulation by members of the LysR family is unknown, activation is mediated by protein binding to the promoter-control region of the genes they regulate (8, 23, 27, 42, 47, 52). A number of proteins, including TrpI, NahR, NodD, AmpR, and IlvY, are known to bind a specific DNA sequence within the intercistronic region; however, there appears to be no consensus DNA-binding site among the target DNAs. As with other LysR members, gel retardation studies using CatR showed that the target for DNA binding is within the promoter-control region. Although the exact DNA sequence for CatR binding has not been determined, it has been localized to a 385-bp region which includes the 140-bp span between the translational start sites of the *catR* and *catBC* regulons. Since CatR also regulates *catA* expression, it would be predicted that a second binding site exists upstream of the *catA* coding region. To determine whether CatR binding overlaps the *catR* or the *catBC* promoter, additional studies need to be conducted, including mRNA mapping to determine the transcriptional start site as well as footprinting analysis. These types of studies are under way. In addition, mutational analysis will help determine the exact contact sites that exist between CatR and its target DNA sequence.

Like many of the LysR regulatory proteins, CatR binds to its promoter-control region in the presence or absence of inducer. In the presence of an inducer molecule, the binding conformation presumably alters, priming the DNA for transcription initiation (21, 47). Binding studies using *Trp*I and its target DNA support the idea of a conformational change in binding when an inducer is present (M. Chang and I. P. Crawford, Abstr. Annu. Meet. Am. Soc. Microbiol. 1988, H-98, p. 186). In the absence of its inducer (indoleglycerol phosphate), *Trp*I binds upstream of the *trpBA* operon, overlapping its own promoter region. The binding pattern alters in the presence of inducer so that the protected region extends downstream towards the *trpBA* promoter region. Conformational changes in binding are also indicated by an altered footprint pattern of IlvY and its target DNA in the presence of inducer (46, 47). However, alteration in binding patterns is not always indicated by gel retardation or footprinting analysis, as it is with NahR and AmpR (27, 41). Further investigation of protein-DNA interaction using CatR and other LysR family members is required to clarify this discrepancy.

The proposed conformational change in binding of the regulatory protein to the target DNA is mediated by inducer interacting with the regulatory protein. This interaction likely occurs within the C-terminal region of the regulatory protein, as indicated by studies on several NodD regulatory proteins (25). Only circumstantial evidence exists because of the divergence in amino acid homology that is found at the C-terminal region among members of the LysR family. Because this region contains a unique primary structure, it is a likely candidate for interaction with an individual inducer molecule. It is interesting that there are large regions of amino acid similarities in the C-terminal region between CatR and ClcR, the regulatory protein needed for induction of the genes required for 3-chlorocatechol degradation (unpublished results). This is not surprising, as the likely inducer for the *clcABD* operon is 2-chloro-*cis*,*cis*-muconate, which structurally resembles *cis*,*cis*-muconate. Because TfdS regulates a similar degradative pathway, a comparison of CatR with TfdS may also show a high degree of homology within the C terminus.

In addition to regulating structural genes required for both biosynthesis and degradation, LysR proteins have been shown to regulate their own expression. As mentioned above, in many cases the target DNA-binding site overlaps with the promoter of the regulatory gene. Therefore, binding the regulatory protein to this site represses its own gene transcription and leads to the activation of the structural gene in the presence of inducer (21). The advantages of gene regulation using divergent promoters have been recently reviewed (6). The transcriptional studies done with *E. coli* suggest that CatR regulates its own expression. Although the promoter of the *catR* gene has not been determined, the proximity of the transcriptional start site of *catB* with the translational starting codon for CatR strongly suggests that the two divergent promoters overlap one another. The reduced protein level seen in minicell analysis and the reduced RNA level seen in dot blot analysis when both CatR and its binding site are intact suggests that CatR is autoregulated. When the binding site is deleted or when CatR is nonfunctional, both protein and RNA levels increase. These experiments, however, need to be repeated in *P. putida* with a reporter gene to measure *catR* transcription.

There may be additional regulation of *catR* transcription, as indicated by the plating experiments using pKT240 as a promoter probe. Preliminary results indicate that transcription from *catR* can be activated. In the absence of benzoate, PRS2000 cells harboring pKR400 were unable to grow on streptomycin levels above the background level, whereas in the presence of benzoate, cells could grow on up to 2,000 μg of streptomycin per ml. This indicates that *catR* is not constitutively expressed, at least not at a level that can be detected by the promoter probe pKT240. Experiments are under way to repeat this analysis with a more sensitive promoter probe containing a β-galactosidase gene so that the product can be assayed. This apparent activation may also depend on the presence of the CatR protein itself, as high levels of streptomycin resistance were not seen in PRS3026. Additional factors or protein products may be required for the regulation of the *catR* gene itself. Transcription of other LysR proteins has been shown to be regulated. For example, *metR* not only regulates its own expression but is also repressed by the product of the *metJ* gene (8, 30). It is possible that the level of CatR within the cell is precisely controlled by dual regulation. Further studies need to be conducted to determine how *catR* itself is regulated.

There are now over 20 proteins that have been grouped into the LysR family on the basis of their amino acid homology to one another. It is becoming increasingly apparent that there are subgroups within the family which show more homology to one another than to other members. For instance, NahR and NodD are very homologous to one another and bind to DNA in a region of sequence homology (41). CatR appears to be part of a subgroup consisting of CatR, TfdS, and ClcR. The amino acid sequence homology throughout the length of these proteins is striking, and they therefore likely evolved from a common ancestor. Future experiments will be geared towards determining how much functional homology has been preserved between these proteins in hopes of understanding how these proteins have evolved to perform their particular role in regulating catabolic genes. It will also be of interest to determine what changes were made in the protein sequence to enable these regulatory genes to be activated by similar but different inducers. Understanding how proteins have evolved to accommodate new substrates will be invaluable for developing organisms that have expanded substrate ranges for degrading more complex and toxic compounds.

ACKNOWLEDGMENTS

This investigation was supported by a grant from the National Science Foundation (DMB-87 21743) and in part by Public Health Service grant ES04050 from the National Institute of Environmental Health Sciences to A.M.C., by the Celgene Corporation, by Public Health Service grant GM33377 from National Institutes of Health, and by an Army Research Office grant to L.N.O.

We thank T. May for critical reading of the manuscript.

LITERATURE CITED

1. **Aldrich, T. L., and A. M. Chakrabarty.** 1988. Transcriptional regulation, nucleotide sequence, and localization of the promoter of the *catBC* operon in *Pseudomonas putida*. J. Bacteriol. **170**:1297–1304.
2. **Aldrich, T. L., B. Frantz, J. F. Gill, J. J. Kilbane, and A. M. Chakrabarty.** 1987. Cloning and complete nucleotide sequence determination of the *catB* gene encoding *cis,cis*-muconate lactonizing enzyme. Gene **52**:185–195.
3. **Aldrich, T. L., R. K. Rothmel, and A. M. Chakrabarty.** 1989. Identification of nucleotides critical for activity of the *Pseudomonas putida catBC* promoter. Mol. Gen. Genet. **218**:266–271.
4. **Appleyard, R. K.** 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. Genetics **39**:440–450.
5. **Bagdasarian, M. M., E. Amann, R. Lurz, B. Ruckert, and M. Bagdasarian.** 1983. Activity of the hybrid *trp-lac* (*tac*) promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host-range, controlled-expression vectors. Gene **26**:273–282.
6. **Beck, C. F., and R. A. J. Warren.** 1988. Divergent promoters, a common form of gene organization. Microbiol. Rev. **52**:318–326.
7. **Bradford, M. M.** 1976. Rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
8. **Cai, X.-Y., M. E. Maxon, B. Redfield, R. Glass, N. Brot, and H. Weissbach.** 1989. Methionine synthesis in *Escherichia coli*: effect of the MetR protein on *metE* and *metH* expression. Proc. Natl. Acad. Sci. USA **86**:4407–4411.
9. **Chakrabarty, A. M., J. S. Karns, J. J. Kilbane, and D. K. Chatterjee.** 1984. Selective evolution of genes for enhanced degradation of persistent, toxic chemicals, p. 43–53. *In* W. Arber, K. Illmensee, W. J. Peacock, and P. Starlinger (ed.), Genetic manipulation: impact on man and society. Cambridge University Press, Cambridge.
10. **Chang, M., A. Hadero, and I. P. Crawford.** 1989. Sequence of the *Pseudomonas aeruginosa trpI* activator gene and relatedness of *trpI* to other procaryotic regulatory genes. J. Bacteriol. **171**:172–183.
11. **Christman, M. F., G. Storz, and B. N. Ames.** 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. USA **86**:3483–3488.
12. **Cook, A. M., H. Grossenbacher, and R. Hutter.** 1985. Isolation and cultivation of microbes with biodegradative potential. Experientia **39**:1191–1198.
13. **Dodd, I. B., and J. B. Egan.** 1987. Systematic method for the detection of potential λ Cro-like DNA-binding regions in protein. J. Mol. Biol. **198**:557–564.
14. **Egelhoff, T. T., R. F. Fisher, T. W. Jacobs, J. T. Mulligan, and S. R. Long.** 1985. Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA **4**:241–248.
15. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76**:1648–1652.
16. **Frantz, B., and A. M. Chakrabarty.** 1986. Degradative plasmids in *Pseudomonas*, p. 295–323. *In* L. N. Ornston and J. R. Sokatch (ed.), The bacteria, vol. 10. The biology of *Pseudomonas*. Academic Press, Inc., New York.
17. **Frantz, B., and A. M. Chakrabarty.** 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA **84**: 4460–4464.
18. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
19. **Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty.** 1985. Microbial degradation of halogenated compounds. Science **228**:135–142.
20. **Harder, W.** 1981. Enrichment and characterization of degrading organisms, p. 77–96. *In* T. Lesinger, A. M. Cook, R. Hutter, and J. Nuesch (ed.), Microbial degradation of xenobiotics and recalcitrant compounds. Academic Press, Inc. (London), Ltd., London.
21. **Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85**:6602–6606.
22. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
23. **Hong, G. F., J. E. Burn, and A. W. Johnston.** 1987. Evidence that DNA involved in the expression of nodulation (*nodD*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. Nucleic Acids Res. **15**:9677–9690.
24. **Honoré, N., M. H. Nicolas, and S. T. Cole.** 1986. Inducible cephaloporinase production in clinical isolates of *Enterobacter cloacae* is controlled by a regulatory gene that has been deleted from *Escherichia coli*. EMBO J. **5**:3709–3714.
25. **Horvath, B., C. W. B. Bachem, J. Schell, and A. Kondorosi.** 1987. Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. EMBO J. **6**:841–848.
26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
27. **Linduist, S., F. Lindberg, and S. Normark.** 1989. Binding of the *Citrobacter freundii* AmpR regulator to a single DNA site provides both autoregulation and activation of the inducible *ampC* β-lactamase gene. J. Bacteriol. **171**:3746–3753.
28. **Lippman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Maxon, M. E., B. Redfield, X.-Y. Cai, R. Shoeman, K. Fujita, W. Fisher, G. Stauffer, H. Weissbach, and N. Brot.** 1989. Regulation of methionine synthesis in *Escherichia coli*: effect of the MetR protein on the expression of the *metE* and *metR* genes. Proc. Natl. Acad. Sci. USA **86**:85–89.
31. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
32. **Normore, W. M.** 1976. Guanine-plus-cytosine (GC) composition of the DNA of bacteria, fungi, algae and protozoa, p. 65–235. *In* G. D. Fasman (ed.), Handbook of biochemistry and molecular biology. Nucleic acids, vol. 2. The Chemical Rubber Company, Cleveland.
33. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
34. **Ornston, L. N., and R. Y. Stanier.** 1964. Mechanism of β-ketoadipate formation by bacteria. Nature (London) **204**:1279–1283.
35. **Ostrowski, J., G. Jagura-Burdzy, and N. M. Kredich.** 1987. The sequences of the *cysB* regions of *Salmonella typhimurium* and *Escherichia coli*. J. Biol. Chem. **262**:5999–6005.
36. **Pabo, C. O., and R. J. Sauer.** 1984. Protein DNA recognition. Annu. Rev. Biochem. **53**:293–321.
37. **Perkins, E. J., G. W. Bolton, M. P. Gordon, and P. F. Lurquin.** 1988. Partial nucleotide sequence of the chlorocatechol degradative operon tfdCDEF of pJP4 and similarity to promoters of the chlorinated aromatic degradative operons *tfdA* and *clcABD*. Nucleic Acids Res. **16**:7200.
38. **Plaman, L. S., and G. V. Stauffer.** 1987. Nucleotide sequence of the *Salmonella typhimurium metR* gene and the *metR-metE*

control region. J. Bacteriol. **169:**3932–3937.

39. **Renault, P., C. Gaillardin, and H. Heslot.** 1989. Product of the *Lactococcus lactis* gene required for malolactic fermentation is homologous to a family of positive regulators. J. Bacteriol. **171:**3108–3114.
40. **Schell, M. A., and E. F. Poser.** 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. J. Bacteriol. **171:**837–846.
41. **Schell, M. A., and M. Sukordhaman.** 1989. Evidence that the transcription activator encoded by the *Pseudomonas putida nahR* gene is evolutionarily related to the transcription activators encoded by the *Rhizobium nodD* genes. J. Bacteriol. **171:**1952–1959.
42. **Schell, M. A., and P. Wender.** 1986. Identification of the *nahR* gene product and nucleotide sequences required for its activation of the *sal* operon. J. Bacteriol. **166:**9–14.
43. **Stragier, P., and J. C. Patte.** 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. J. Mol. Biol. **168:** 333–350.
44. **Streber, W. R., K. N. Timmis, and M. H. Zenk.** 1987. Analysis, cloning and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169:**2950–2955.
45. **Tabor, S., and C. A. Richardson.** 1985. T7RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
46. **Wek, R. C., and G. W. Hatfield.** 1986. Nucleotide sequence and *in vivo* expression of *ilvY* and *ilvC* genes in *Escherichia coli* K12. J. Biol. Chem. **261:**2441–2450.
47. **Wek, R. C., and G. W. Hatfield.** 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. J. Mol. Biol. **203:**643–663.
48. **Wheelis, M. L., and L. N. Ornston.** 1972. Genetic control of enzyme induction in the β-ketoadipate pathway of *Pseudomonas putida*: deletion mapping of *cat* mutations. J. Bacteriol. **109:**790–795.
49. **Wheelis, M. L., and R. Y. Stanier.** 1970. The genetic control of dissimilatory pathways in *Pseudomonas putida*. Genetics **66:** 245–266.
50. **Wu, C. H., M. K. Ornston, and L. N. Ornston.** 1972. Genetic control of enzyme induction in the beta-ketoadipate pathway of *Pseudomonas putida*: Two-point crosses with a regulatory mutant strain. J. Bacteriol. **109:**796–802.
51. **Yanisch-Perron, C., J. Viera, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene **33:**103–119.
52. **You, I. S., D. Ghosal, and I. C. Gunsalus.** 1988. Nucleotide sequence of plasmid NAH7 gene *nahR* and DNA binding of the *nahR* product. J. Bacteriol. **170:**5409–5415.