JOURNAL OF BACTERIOLOGY, Mar. 1993, p. 1596–1604
0021-9193/93/061596-09$02.00/0

Vol. 175, No. 6

Cloning and Nucleotide Sequence of the Gene Encoding the Positive Regulator (DmpR) of the Phenol Catabolic Pathway Encoded by pVI150 and Identification of DmpR as a Member of the NtrC Family of Transcriptional Activators

VICTORIA SHINGLER,* MAGDALENA BARTILSON, AND TERRY MOORE

Department of Cell and Molecular Biology, The University of Umeå, S-901 87 Umeå, Sweden

Received 6 November 1992/Accepted 11 January 1993

The catabolic plasmid pVI150 of *Pseudomonas* sp. strain CF600 encodes all the genetic information required for the regulated metabolism of phenol and some of its methyl-substituted derivatives. The structural *dmp* genes of the pathway are clustered in a single operon that lies just downstream of a −24 TGGC, −12 TTGC *nif/ntr*-like promoter sequence. Promoters of this class are recognized by a minor form of RNA polymerase utilizing σ^{54} (NtrA, RpoN). Primer extension analysis demonstrated that the *dmp* operon transcript initiates downstream of the −24, −12 promoter. Transposon insertion mutants, specifically defective in the regulation of the *dmp* operon, were isolated, and complementation of a phenol-utilization regulatory mutant was used to identify the regulatory locus, *dmpR*. The 67-kDa *dmpR* gene product alone was shown to be sufficient for activation of transcription from the *dmp* operon promoter. Nucleotide sequence determination revealed that DmpR belongs to the NtrC family of transcriptional activators that regulate transcription from −24, −12 promoters. The deduced amino acid sequence of DmpR has high homology (40 to 67% identity) with the central and carboxy-terminal regions of these activators, which are believed to be involved in the interaction with the σ^{54} RNA polymerase and in DNA binding, respectively. The amino-terminal region of DmpR was found to share 64% identity with the amino-terminal region of XylR, which is also a member of this family of activators. This region has been implicated in effector recognition of aromatic compounds that is required for the regulatory activity of XylR.

Pseudomonas strain CF600 can efficiently grow on phenol, cresols, and 3,4-dimethylphenol as sole carbon and energy sources (40). All the genetic information for the regulated biodegradation of these aromatic compounds is encoded on a conjugative catabolic plasmid, pVI150. This plasmid, like the CAM and OCT catabolic plasmids, belongs to the incompatibility group P-2 and is very large (40). The biochemical route for the dissimilation of phenolic compounds involves a multicomponent phenol hydroxylase and a subsequent *meta* cleavage pathway (35, 40, 41). These enzymes are encoded by fifteen *dmp* structural genes in a single operon of pVI150 (5, 41). Previous studies have shown that the *dmp* operon is tightly regulated, with very low levels of expression in the absence of pathway substrates (40).

Nucleotide sequence analysis of the entire *dmp* operon identified a single putative promoter sequence 30 bp upstream of the ATG initiation codon of the first gene of the operon (6, 32, 33, 41). The promoter sequence belongs to a distinct class of promoters, σ^{54}-dependent promoters, that have an invariant −24 GG, −12 GC sequence (3, 45). These −24 TGGC, −12 TTGC promoters have little in common with the *Escherichia coli* −35 TTGACA, −10 TATAAT consensus promoter sequence (17) and are recognized by RNA polymerase utilizing the alternative σ^{54} factor encoded by *rpoN* or its homologs (10, 25).

All −24, −12 promoters so far analyzed are regulated by positive transcriptional activators that usually bind to specific DNA sequences located unusually far (100 to 200 bp) upstream of the promoter. The binding sequences are often inverted repeats that can be moved more than 1 kb away without losing their ability to mediate transcription (for reviews, see references 7, 27, and 28). These regions have, for this reason, been called upstream activation sequences (UAS), upstream regulatory sequences, or enhancer sequences.

In the present work, we analyze the regulation of the *dmp* operon. Primer extension mapping was used to demonstrate that in vivo transcription of the *dmp* operon initiates downstream of a −24, −12 promoter. A positive regulator, DmpR, of the *dmp* operon was cloned by complementation of a phenol utilization-deficient regulatory mutant. Nucleotide sequence analysis of the *dmpR* gene revealed that this regulator belongs to the NtrC class of positive transcriptional activators. This family of activators contains many members, including NtrC, NifA, DctD, FhlA, FlbD, HoxA, $HupR_1$, HydG, HrpS, TyrR, and XylR, which regulate expression of genes involved in a variety of physiological processes (see references 27, 28, 37, and 43 and references therein).

MATERIALS AND METHODS

Bacterial strains and culture conditions. *E. coli* DH5 (*endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1* φ80d *lacZ*) (16) was used for construction and maintenance of plasmids, and CSR603 (*phr-1 recA1 uvrA6*) (38) was used for analysis of plasmid-encoded polypeptides. *Pseudomonas* strains used were *Pseudomonas* sp. strain CF600 (phenol and 3,4-dimethylphenol degrader) (40); *Pseudomonas putida* KT2440 (*hsdR*) (13); PB2701, a streptomycin-resistant derivative of KT2440; and PB2701 (pVI150) (40). *E. coli* strains were cultured at 37°C, while *Pseudomonas* strains were grown at 30°C. Luria broth was used as rich medium (30);

* Corresponding author.

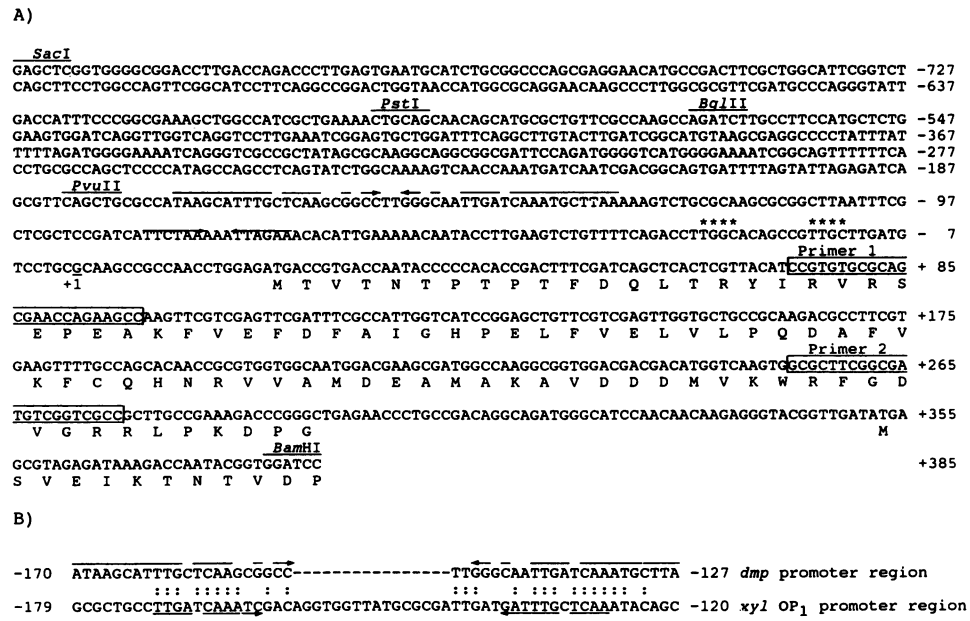

FIG. 1. Regulatory region of the *dmp* operon. (A) A single strand of the nucleotide sequence surrounding the start of the *dmp* operon is shown (32) with translation of *dmpK*, the first gene of the operon, and the first 12 codons of *dmpL*. Asterisks indicate the locations of promoter sequences with homology to −24 TGGC, −12 TTGC promoters (10). Discontinuous arrows indicate the extent of inverted repeat sequences, and boxes indicate the locations of regions complementary to primers used to identify the transcriptional start, which is marked +1. (B) Alignment showing 70% identity of the large inverted repeat shown in panel A with the inverted repeat identified as the binding site for XylR (9, 18, 26). Each of the inverted repeats displays dyad symmetry with the consensus sequence of TTG-TCAA. Dashes indicate the location and size of spaces inserted into the sequence to obtain maximal homology.

minimal medium was M9 salts (30) supplemented with the following carbon sources: phenol, cresols, or 3,4-dimethylphenol (2.5 mM) or benzoate or glucose (5 mM). Methionine assay medium (Difco) was used in maxicell analysis of polypeptides. Selection for resistance markers was at concentrations in micrograms per milliliter for *E. coli* and *Pseudomonas* strains, respectively, as follows: ampicillin (100), carbenicillin (1,000), kanamycin (100 and 100), and streptomycin (100 and 1,000).

Plasmids and strain construction. Plasmids were constructed by using standard recombinant techniques and were introduced into *E. coli* strains by transformation (26) and into *Pseudomonas* strains by electroporation with a Bio-Rad Gene Pulser. Previously published plasmids used in this study are pEJL449 (24), pBluescript SK(+) (Stratagene), pUTmini-Tn*5*Km2 (8), and the broad-host-range vectors pMMB66EHΔ and pMMBHEΔ, which constitutively express DNA cloned into the polycloning cassette located downstream from the *tac* promoter (40). Plasmid pVI299 (41) carries all the structural genes of the *dmp* operon and can express the pathway enzymes from the *lacI*^{q}-regulated *tac* promoter of the vector.

Plasmid pVI349 was constructed by cloning the 1.11-kb *Sac*I-to-*Bam*HI fragment depicted in Fig. 1 into the polycloning site of pBluescript SK(+). Plasmids pVI350Δ to pVI359Δ are based on the *tac* expression vectors pMMB66EHΔ and pMMB66HEΔ, and the extent of CF600-derived DNA cloned into these plasmids is depicted in Fig. 2. Plasmids pVI350Δ and pVI351Δ have the 14.8-kb *Bam*HI fragment from a *Pseudomonas* strain CF600-derived lambda clone, in the two different orientations, in pMMB66EHΔ. Plasmid pVI352Δ was constructed by deleting the internal 6.1-kb *Sac*I fragment of pVI350Δ. Plasmids pVI353Δ and pVI354Δ have the 6.1-kb *Sac*I fragment from pVI350Δ cloned into the polycloning site of pMMB66EHΔ and pMMB66HEΔ, respectively, while pVI355Δ and pVI356Δ carry the 4.1-kb *Sma*I fragment in these vectors. Plasmids pVI357Δ and pVI358Δ were generated by *Bal* 31 deletion of pVI355Δ, and pVI359Δ was constructed by *Bal* 31 deletion of pVI357Δ. The nucleotide sequences of the junction of the *Bal* 31 deletions and of the vector were determined.

Plasmids pVIΩ1 to pVIΩ16 were generated by ligation of total *Sac*I-digested DNA from PB2701 (pVI150)::Ω1 to -16 (see below) and subsequent selection of the kanamycin resistance determinant of the Ω transposon.

The 4.1-kb *Sma*I fragment (coordinates 1.6 to 5.7 [see Fig. 7]), which spans *dmpR*, was cloned into the *Eco*RV site of pBluescript to allow the subsequent isolation of the *dmpR* coding region on a *Not*I fragment (coordinates 3.3 to 5.7, [see Fig. 7]). This *Not*I fragment was subsequently cloned into the unique *Not*I site within the defective transposon mini-Tn*5*Km2, residing on a suicide plasmid (8). KT2440::*dmpR* was constructed by transposition of the *dmpR*-carrying mini-Tn*5*Km2 onto the chromosome by using selection for kanamycin resistance.

Transposon mutagenesis. The Ω transposon located on pEJL449, a pBR322-based plasmid that serves as a *Pseudomonas* suicide donor plasmid, contains a plasmid origin of replication and a kanamycin resistance determinant and is devoid of restriction endonuclease recognition sites for *Sac*I (24). Transposon insertion mutants of PB2701 (pVI150) were generated by introduction of pEJL449 and subsequent selection for the kanamycin resistance determinant of the transposon.

Lambda cloning. A total genomic library of *Pseudomonas* strain CF600 *Bam*HI-digested DNA was obtained with

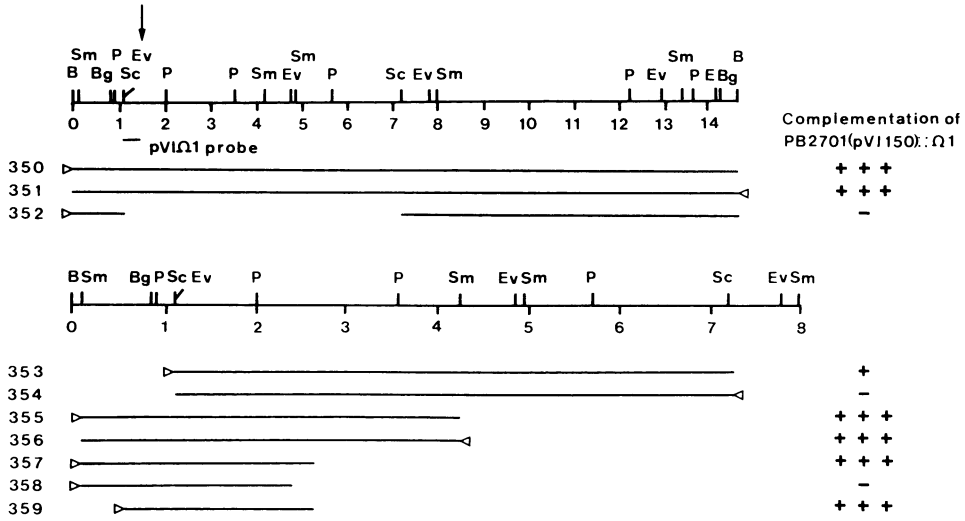

FIG. 2. Deletion and subcloning mapping of the *dmpR* gene. The extent of DNA derived from pVI150 in plasmids pVI350Δ to pVI359Δ is shown. The + and − signs indicate the ability or inability of plasmids to complement PB2701 (pVI150)::Ω1 to restore growth on phenol. Plasmid numbers are given on the left. The arrow indicates the site of insertion of the Ω transposon within PB2701 (pVI150)::Ω1. Restriction endonuclease recognition sites are B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; Ev, *Eco*RV; P, *Pst*I; Sc, *Sac*I; and Sm, *Sma*I.

EMBL3 (15). Generation, screening, isolation, and purification of recombinant phage identified with a 370-bp probe derived from pVIΩ1 was as previously described (5).

Primer extension. The 5′-mRNA start of the *dmp* operon transcript was determined by primer extension analysis essentially as previously described (4). The oligonucleotides (primer 1, 5′-GGCTTCTGGTTCGCTGCGCACACGG-3′, and primer 2, 5′-GGCGACCGACATCGCCGAAGCGC-3′) were 5′ end labelled with [γ-^{32}P]ATP (Amersham) and annealed to 50 μg of total RNA prepared from *Pseudomonas* sp. strain CF600 grown on minimal medium containing phenol as the sole carbon source. The annealing buffer conditions used were 50 mM Tris-HCl (pH 8.3)–60 mM NaCl–10 mM dithiothreitol–1 mM EDTA.

Nucleotide sequence determinations. Nucleotide sequences were determined directly from plasmids with Pharmacia's DNA sequencing kit (no. 27-168201). To determine the sequence of the 2.1-kb DmpR-encoding region, the DNA was subcloned into the polycloning site of the pBluescript SK(+) sequencing vector and ordered deletion libraries were generated with exonuclease III and mung bean nuclease essentially as described in the Stratagene Exo/Mung DNA sequencing manual. Each part of both strands was sequenced. The nucleotide sequence of one strand from the equivalent DNA region derived from PB2701 (pVI150) was determined with custom-designed oligonucleotides.

Enzymes. DNA restriction and modifying enzymes were purchased from Boehringer Mannheim or New England Biolabs and used as recommended by the manufacturer.

Analysis of plasmid-encoded polypeptides. *E. coli* CSR603, harboring the different plasmids, was prepared, labelled with L-[^{35}S]methionine (Amersham), and analyzed essentially as described previously (38) by 10-to-20%-gradient sodium dodecyl sulfate-polyacrylamide gel electrophoresis (29). Size estimations of polypeptides were performed by using Pharmacia's LMW calibration kit (no. 17-044601).

Nucleotide sequence accession number. The sequence data in this paper have been submitted to the EMBL data library under accession no. X68033.

RESULTS

Determination of the 5′-mRNA start of the *dmp* operon transcript. The 15 genes of the phenol catabolic pathway encoded by pVI150 lie in a single, tightly regulated operon (40, 41). Nucleotide sequence analysis of the entire *dmp* operon identified a single putative promoter sequence 30 bp upstream of the ATG initiation codon of the first gene of the operon (Fig. 1). The promoter sequence belongs to the class of *nif/ntr*-like σ^{54}-dependent −24 TGGC, −12 TTGC promoters. This group of promoters also includes the XylR-regulated *xylCAB* operon and *xylS* promoters of the toluene-xylene catabolic plasmid pWW0 (21, 22, 36).

To identify the in vivo transcriptional start of the *dmp* operon transcript, primer extension analysis was performed with RNA isolated from phenol-induced *Pseudomonas* strain CF600. Two oligonucleotides, designated primer 1 and primer 2, located 54 and 374 bp, respectively, downstream of the ATG codon of the first gene of the pathway (Fig. 1A), were used. A single band, corresponding to a G residue 19 bp upstream of the ATG codon of the first gene of the operon, was detected with both primers (Fig. 3 and data not shown). These results place the start of the operon transcript at a position consistent with initiation of transcription from the −24, −12 promoter (Fig. 1A).

Isolation and identification of regulatory mutants. Transposon mutagenesis was used to investigate the regulatory circuit that controls the expression of the *dmp* operon-encoded enzymes. Transposon insertion mutants of PB2701 (pVI150) were generated with the synthetic Ω transposon from pEJL449 as described in Materials and Methods. This synthetic Ω transposon has a plasmid origin of replication and a kanamycin resistance marker and is devoid of recognition sequences for a number of restriction enzymes. These properties facilitated the subsequent cloning of the DNA surrounding the point of insertion. Transposon insertion derivatives specifically deficient in phenol utilization were identified by their inability to grow on minimal medium containing phenol as the sole carbon source while retaining

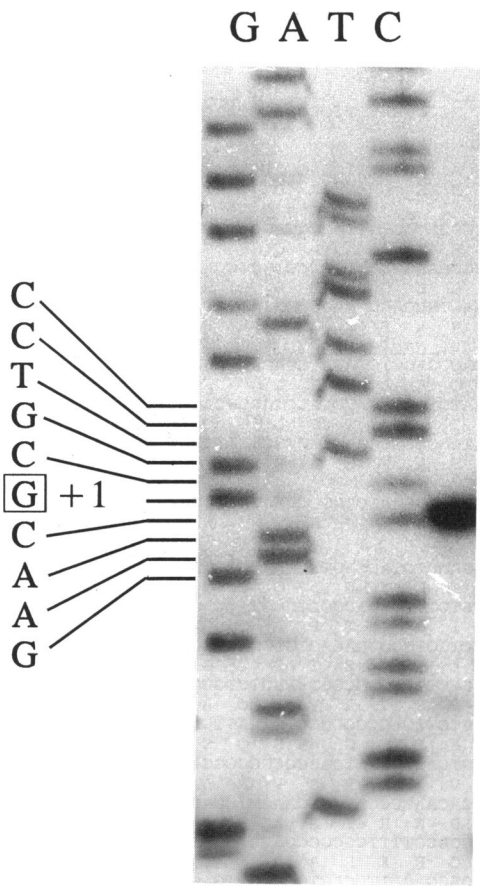

FIG. 3. Mapping of the 5'-mRNA start of the *dmp* operon transcript. Primer extension analysis of RNA isolated from phenol-grown *Pseudomonas* strain CF600 was performed with [γ-^{32}P]ATP-end-labelled primer 1 as described in Materials and Methods. The reaction was run next to sequence reactions performed on pVI349 with primer 1 and α-^{35}S-dCTP. The shift between the band of the primer extension reaction (shown on the right) and the corresponding band of the sequence reaction is due to the presence of an additional phosphate group on the sequencing primer.

the ability to grow on minimal medium containing benzoate or glucose.

Transposon insertion into genes involved in a positive regulatory circuit or into genes encoding enzymes of the catabolic pathway would result in the inability of the bacteria to grow on phenol. To distinguish between the two different types of insertion mutants, DNA surrounding the site of insertion was cloned from 16 phenol utilization-deficient mutants to generate pVIΩ1 to pVIΩ16. These plasmids were subsequently analyzed for their ability to hybridize to pathway-specific DNA probes derived from structural *dmp* genes (data not shown). Four phenol utilization-deficient mutants, designated PB2701 (pVI150)::Ω1 to -4, which harbored the catabolic plasmid pVI150 but did not have the transposon inserted into the structural genes, were identified.

Identification of the *dmpR* coding region. In addition to a pathway-specific plasmid-encoded regulator(s), host-encoded proteins such as alternative sigma factors and histone-like proteins may interact with the regulatory system. To determine whether the Ω transposon had inserted into the chromosome or the catabolic plasmid pVI150 in the phenol utilization-deficient mutants, DNA from pVIΩ1 to pVIΩ4 was hybridized with total DNA from PB2701 and PB2701 (pVI150). The plasmid pVIΩ1 was found to hybridize only to the DNA of the strain harboring pVI150, while pVIΩ2 to pVIΩ4 hybridized to DNA from both strains (data not shown). Hence, the transposon had inserted into the catabolic plasmid pVI150 in the PB2701 (pVI150)::Ω1 strain, while it had inserted into the chromosomal DNA in PB2701 (pVI150)::Ω2 to -4. The plasmid-located regulatory gene inactivated in PB2701 (pVI150)::Ω1 was designated *dmpR*.

To isolate the region of DNA present in pVIΩ1 in its wild-type configuration, a lambda library of total *Pseudomonas* sp. strain CF600 DNA was constructed. Approximately 2×10^4 plaques were screened by hybridization to a 370-bp fragment derived from pVIΩ1. The restriction map of the 14.8-kb *Bam*HI fragment from two identical hybridizing phages is shown in Fig. 2 along with the location of the pVIΩ1-specific probe.

The 14.8-kb *Bam*HI fragment, from one of the hybridizing phages identified above, was subcloned in both orientations downstream of the *tac* promoter of the broad-host-range expression vector pMMB66EHΔ. The resulting plasmids, pVI350Δ and pVI351Δ, were analyzed for their ability to complement the defect in the mutant strain PB2701 (pVI150)::Ω1. Complementation was scored by testing for restoration of growth on phenol. Both plasmids were found to complement the defect in PB2701 (pVI150)::Ω1 (Fig. 2). To further define the complementing region, deletion and subcloning analysis were used to generate pVI352Δ to pVI359Δ. The extent of the DNA inserted in these plasmids and their ability to complement PB2701 (pVI150)::Ω1 are summarized in Fig. 2. The results localized the complementing region to a 2.1-kb DNA region encoded by pVI359Δ.

Nucleotide sequence of the *dmpR* region. The nucleotide sequence of the 2.1-kb insert of pVI359Δ was determined as described in Materials and Methods. The sense strand of this 2,153-bp *dmpR*-encoding region is shown in Fig. 4. Translation of a single long open reading frame, from nucleotides 245 to 1,933, which would encode a 563-amino-acid-long polypeptide of 63.2 kDa, is also shown. An alternative translational start, located at nucleotide 320, would encode a 538-amino-acid-long polypeptide of 60.2 kDa. It is not clear at present whether both or only one of the putative translational start points is used (see below). The nucleotide sequence of the coding region is G+C rich (61.3%), as has been found for other *dmp* genes and *Pseudomonas* genes in general (see reference 41 and references therein). In contrast, the upstream region has a lower G+C content (51.3%). The nucleotide sequence immediately downstream of the open reading frame does not contain any obvious terminator-like sequences (34) but does contain a long T track.

The deduced amino acid sequence of DmpR has 67% identity with that of the XylR regulator of the TOL catabolic plasmid (23; see below; Fig. 5). In view of this homology, the upstream regulatory regions of these two genes were compared, but little overall homology was detected. The nucleotide sequence upstream of the coding region contains a number of sequences with homology to *E. coli* and *P. putida* promoters (Fig. 4). However, determination of which of these sequences are functional awaits analysis of *dmpR* transcription.

Comparison of the deduced amino acid sequence of DmpR with the GenBank and EMBL data bases (release 31) demonstrated extensive homology between DmpR and the NtrC family of transcriptional activators. XylR, NtrC, and NtrA members of this family have previously been aligned (23). DmpR was found to be most homologous to XylR, and the

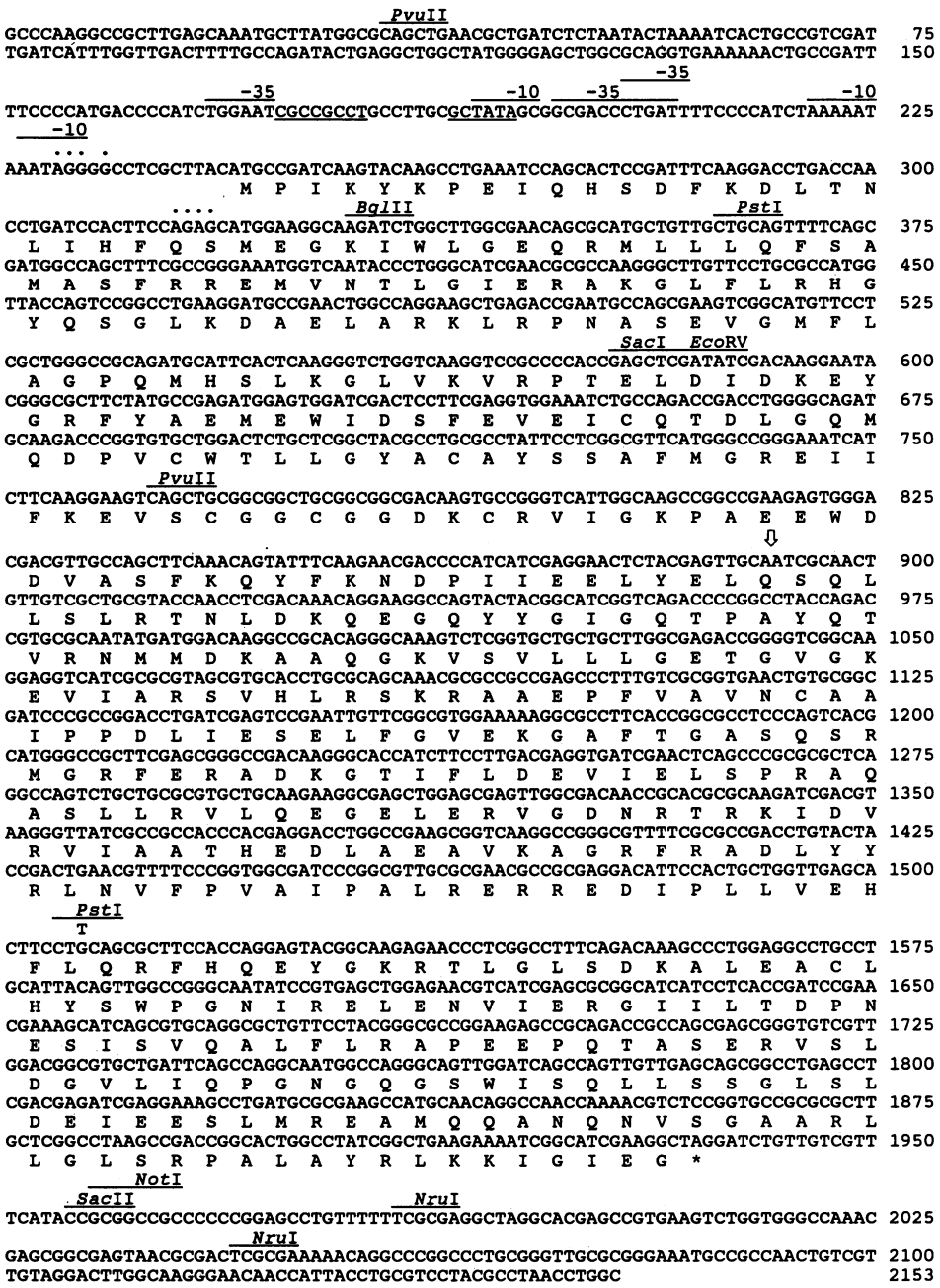

FIG. 4. Nucleotide sequence of the *dmpR* region. The sense strand of the *Pseudomonas* strain CF600-derived 2,153-bp region of pVI359Δ is shown with translation of a large open reading frame. The amino acid sequence is shown in the one-letter code, and an asterisk indicates the stop codon. Dots indicate regions upstream of the putative translational start site that are complementary to the 3′ end of the 16S rRNA of *P. aeruginosa* (39). Over- and underlined bases indicate regions with homology to the *E. coli* −35 TTGACA <17 +/−1> −10 TATAAT (17) and *P. putida* A-AGGC-T <7–12> GC(T/A)ATA (31) consensus promoter sequences. The arrow indicates the location of the insertion site of Ω1. The nucleotide above the sequence at position 1507 indicates the G-to-T point mutation found in the equivalent region from PB2701 (pVI150), which results in disruption of the wild-type *PstI* site. After the last base of the termination codon, the sequence derived from pVI150 of *Pseudomonas* sp. strain CF600 (shown) and that derived from PB2701 (pVI150) completely diverge.

alignment of DmpR with XylR is shown in Fig. 5. The sequence is divided into four domains on the basis of analysis of the most extensively studied members of this group of proteins, namely, NtrC and NifA (11).

Plasmids that express the entire *dmpR* gene are capable of efficiently complementing the defect in PB2701 (pVI150):: Ω1. Plasmid pVI353Δ, however, which has a part of the 5′ coding region deleted, is still able, albeit inefficiently, to complement. The DNA present in this plasmid was cloned on a *Sac*I fragment starting at nucleotide 578 in Fig. 4. The sequence of the DNA downstream of this *Sac*I site has an ATG initiation codon (nucleotide 620) preceded by a possible *Pseudomonas aeruginosa* ribosome binding site (AAGGA, nucleotides 593 to 597). A protein initiating at this

FIG. 5. Homology of DmpR and XylR (top and bottom sequences, respectively) (23). Identical residues are indicated by asterisks, and dashes indicate the location and size of gaps inserted to achieve maximal alignment. The two proteins share 67% overall identity. Domains are described in the text and share the following identities: A, 64%; B, 52%; C, 79%; and D, 65%.

ATG would not contain the first 125 amino-terminal residues. Partial complementation by pVI353Δ might be interpreted in either of two ways: (i) the amino-terminal residues are not critical for the protein function, or (ii) deletion of this region produces a protein that mimics inefficiently the activated form of the regulator.

Estimation of the size of the DmpR protein. To investigate the size of the *dmpR* gene product, plasmids based on the *tac* expression vector pMMB66EHΔ were used. Plasmids pVI355Δ to pVI359Δ carry different portions of the *dmpR* coding region downstream of the *tac* promoter of the vector (Fig. 2). These plasmids were introduced into the *E. coli* maxicell strain CSR603, and plasmid-encoded polypeptides were analyzed. The results of these experiments are shown in Fig. 6.

Plasmids pVI355Δ, pVI357Δ, and pVI358Δ carry the entire coding region of the *dmpR* gene in the correct orientation for transcription from the *tac* promoter. These plasmids mediate production of two novel polypeptides of 67 and 65 kDa (lanes 2, 4, and 6), while a plasmid carrying DNA deleted in the 3′ end of the *dmpR* coding region mediated production of a smaller truncated polypeptide (lane 5). Plasmid pVI355Δ and pVI356Δ carry the same *dmpR*-encoding DNA fragment, and both complement PB2701 (pVI150):: Ω1 efficiently. However, pVI356Δ has the DNA oriented

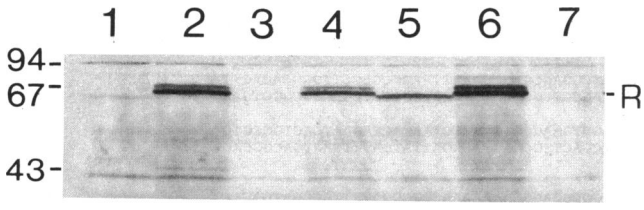

FIG. 6. Expression of *dmpR*. Maxicell analysis of plasmid-encoded polypeptides separated by 10-to-20%-gradient sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Lanes: 1 and 7, the control vector pMMB66EHΔ; 2, pVI355Δ; 3, pVI356Δ; 4, pVI357Δ; 5, pVI358Δ; 6, pVI359Δ. Molecular mass standards are given in kilodaltons. R indicates the location of two polypeptides of 67 and 65 kDa. Plasmid pVI356Δ, lane 3, expresses very low levels of a 67-kDa polypeptide that was visible on the original autoradiogram.

such that the *dmpR* gene could not be transcribed from the *tac* promoter, suggesting that in this plasmid *dmpR* is expressed solely from its own promoter. This plasmid expresses very low levels of a 67-kDa polypeptide that was visible on the original autoradiogram. The polypeptide produced by pVI356Δ is indistinguishable, on the basis of size, from the larger of the two polypeptides produced by pVI355Δ. These data suggest that the larger, 67-kDa polypeptide is the normal product of the *dmpR* gene. The presence of the smaller, 65-kDa polypeptide might reflect processing of DmpR when it is overexpressed or translation from an alternative start point from transcripts initiating from the *tac* promoter of the expression vector. Similarly, overexpression of the *xylR* gene has been reported to give rise to a 64-kDa polypeptide in addition to the normal 67-kDa product (23). The observed size of 67 kDa is in good agreement with the 63.2 kDa predicted from the deduced amino acid sequence of DmpR.

Location of *dmpR*. Computer-assisted searches of nucleotide sequence data bases revealed that the *dmpR* upstream sequence overlaps the previously reported *dmp* operon upstream region and that these two genetic units are linked and divergently transcribed. Figure 7 shows the contiguous restriction map and the relative locations of the *dmp* genes and illustrates the divergent transcriptional directions of *dmpR* and the *dmp* operon. The nucleotide sequence of the regulatory region common to the two transcriptional units is shown in Fig. 8, along with the relative locations of putative regulatory signals.

The restriction map of the *Bam*HI-to-*Pvu*II fragment (coordinates 5.9 to 18.1, Fig. 7) had previously been determined from DNA cloned from pVI150 isolated from *Pseudomonas* sp. strain CF600. The restriction map of an overlapping 13.5-kb *Eco*RI fragment cloned from pVI150 isolated from the strain PB2701 (pVI150) had been aligned with the *Bam*HI-to-*Pvu*II fragment to gain a contiguous map (5) (Fig. 7). Since the strain PB2701 (pVI150) was generated by conjugation of pVI150 from *Pseudomonas* strain CF600 to PB2701 and has the same growth substrate range and induction profile as the wild-type strain (40), it was initially surprising that the restriction maps of the DNA from the two strains were not identical. Although the maps are the same

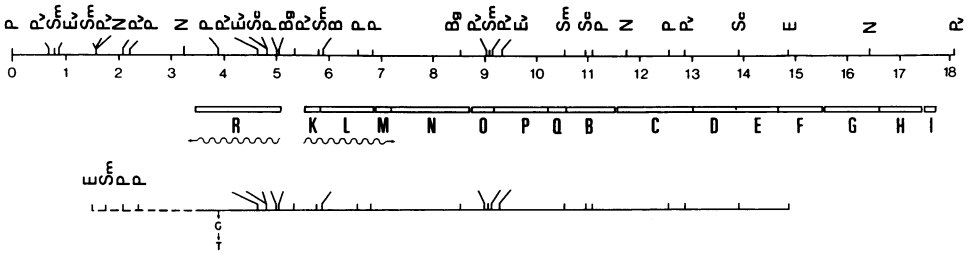

FIG. 7. Location of the *dmpR* gene. The upper part of the figure illustrates the restriction map of pVI150 DNA derived from *Pseudomonas* strain CF600. Only the region that overlaps the *dmpR* and *dmp* operon genes (KLMNOPQBCDEFGHI) is shown. Restriction site coordinates are given in kilobases and abbreviated as in the legend to Fig. 3, except as follows: N, *Not*I, and Pv, *Pvu*II. Wavy lines indicate the direction of transcription. The lower part of the figure represents the restriction map of an overlapping *Eco*RI fragment cloned from PB2701 (pVI150) (5). The dashed line indicates where the restriction maps diverge, and the arrow indicates the location of the point mutation within the *Pst*I site (Fig. 4).

over the majority of their lengths, they start to differ from within the *dmpR* region (Fig. 7). The 4.5-kb *Eco*RI-to-*Bam*HI fragment from PB2701 (pVI150) DNA spans the region that differs; nevertheless, this region was found to be able to complement the *dmpR* insertion mutant PB2701 (pVI150)::Ω1, demonstrating that a functional DmpR was produced despite the difference in restriction maps (data not shown). Determination of the nucleotide sequence of this region showed that the integrity of DmpR was intact (Fig. 4). A silent point mutation (G to T at position 1507, Fig. 4) was found to remove a *Pst*I site within the coding region, while the nucleotide sequence from the first base after the termination codon was completely different. Southern blotting analysis of the DNA from *Pseudomonas* strain CF600 and PB2701 (pVI150), using the 0.2-kb *Sac*I-to-*Pst*I fragment (coordinates 4.8 to 5.0, Fig. 7), demonstrated that only one configuration of the DNA was present in each strain (data not shown). Therefore, we conclude that a genetic rearrangement has taken place, either during or after construction of PB2701 (pVI150), which leaves the integrity of the *dmp* regulatory locus and structural operon intact but accounts for the difference in the restriction maps.

***trans* activation of the *dmp* operon promoter.** Plasmid pVI299 carries the *dmp* operon and its promoter as a 13-kb *Bgl*II-to-*Pvu*II fragment (coordinates 5.1 to 18.1, Fig. 7), orientated so that the pathway genes can be expressed from the IPTG (isopropyl-β-D-thiogalactopyranoside)-inducible *tac* promoter of the expression vector. Expression of the pathway enzymes on this plasmid has been shown to be dependent on the vector-located *tac* promoter (41). Hence, pseudomonads that harbor this plasmid are only able to grow on phenol and its methyl-substituted derivatives in the presence of IPTG. To investigate whether *dmpR* is sufficient to allow expression from the *dmp* operon promoter, KT2440::*dmpR*, which expresses DmpR from the chromosome, was constructed as described in Materials and Methods. KT2440::*dmpR* harboring pVI299 was found to be able to grow on phenol and its methyl-substituted derivatives in the absence of IPTG. These data suggest that DmpR alone is sufficient to allow expression of the pathway enzymes from the *dmp* operon promoter in a *Pseudomonas* background. Some of the NtrC family of transcriptional activators have been shown to be part of two-component regulatory systems composed of a surface-located sensor that transfers a signal to a cytosolic transcriptional activator (reviewed in reference 42). Since DmpR is sufficient to activate the *dmp* operon promoter, it seems likely that DmpR does not require a specific sensor.

DISCUSSION

In this study, we investigated the regulation of growth on phenol mediated by the *dmp* operon of pVI150. The *dmp* operon encodes all the enzymes required for phenol catabolism and is divergently transcribed from a closely linked regulatory locus, *dmpR*. The *dmpR* locus was identified by generation of transposon insertion regulatory mutants. Nucleotide sequence determination of the *dmpR* coding region showed that the *dmpR* gene product belongs to a class of transcriptional activators that, in conjunction with σ^{54}, regulate transcription of −24, −12 promoters. This is consistent with our finding that the start of the *dmp* operon transcript lies downstream of a −24, −12 promoter sequence.

The deduced amino acid sequence of DmpR has strong

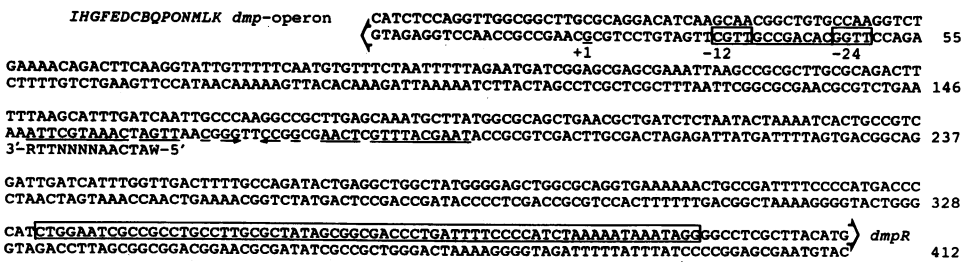

FIG. 8. Divergent regulatory region of *dmpR* and the *dmp* operon. The 412-bp region between the *dmpR* and *dmp* operon coding regions is shown. +1 indicates the location of the transcriptional start of the *dmp* operon. The boxes indicate the locations of regions that have homology to known promoter sequences identified in Fig. 1 and 4. Long discontinuous arrows indicate the extent of an inverted repeat that has homology to the XylR binding sequence. The consensus IHF recognition sequence 5′-WATCAANNNNTTR-3′ (W, A or T; R, A or G) (14) is shown aligned with a region overlapping the long inverted repeat.

homology to the NtrC class of transcriptional activators. The proteins of this family have been divided into four domains on the basis of the extensive studies of NtrC and NtrA (11). Alignment of DmpR with the XylR transcriptional activator of this family revealed 67% identity over the entire length of the proteins (Fig. 5). The central domain C (residues 234 to 472) is the most conserved region, exhibiting between 41 and 79% identity between DmpR and other members of the family. A potential nucleotide binding fold (46) is located within domain C of these proteins. The probable role of this domain in the interaction with σ^{54} RNA polymerase, and binding and hydrolysis of ATP to allow formation of open transcriptional complexes, has recently been reviewed (27, 28).

The carboxy-terminal domain D (residues 519 to 558) is the next most highly conserved region. Identity of 23 to 65% between DmpR and other members of the family was found. This region contains a helix-turn-helix DNA binding motif found in a number of transcriptional activators and repressors (11). The number of residues separating domains C and D is highly variable within this family, and these regions bear little sequence identity.

The B domain (Q linker) is a short hydrophilic region which is particularly rich in Gln (Q) residues and probably serves as an interdomain linker between the C and A domains (11). With the exception of XylR, the A domain of DmpR (residues 1 to 210) does not share homology with other members of the NtrC family or with other proteins in the data bases. In other members of the family, the amino-terminal domains are the site of transfer of signals that are received via sensory proteins (11, 42).

The amino-terminal domain common to XylR and DmpR exhibits 64% identity, suggesting a common role for this domain. Neither XylR nor DmpR appears to require specific sensory proteins for their activity since they can function in the absence of any other plasmid-encoded regulatory protein. The XylR protein is constitutively expressed and can bind to its DNA recognition sequence in the absence of aromatic coinducer molecules but does not function in transcriptional activation unless effector molecules are present (1, 9). Aromatic effector molecules most probably interact through the amino-terminal region to activate the regulatory activity of these proteins. XylR has broad effector specificity that includes some aromatic compounds that are not substrates or intermediates of the toluene-xylene catabolic pathway that it regulates (2). Despite the broad effector specificity of XylR, phenol does not function as an effector molecule for this protein. It will be of interest to study the effector specificity of DmpR in the light of the high homology between the amino-terminal sequences of these two proteins.

The NtrC family of activators generally bind further than 100 bp from the σ^{54}-dependent promoters which they regulate (9, 27, and 28 and references therein). Two inverted repeats upstream of the *dmp* operon promoter are shown in Fig. 1A. The repeat centered at −76 bp is similar to the symmetrical recognition sequence involved in binding of *E. coli* repressors and activators (12). The longer inverted repeat, centered at −150 bp, has 70% identity and positional similarity to the inverted repeat identified as the binding site of the XylR positive activator of the *xylCAB* operon (OP_1) and the *xylS* gene of the TOL catabolic plasmid pWW0 (1, 9, 18, 20) (Fig. 1B). Taken together, these observations suggest that the long inverted repeat upstream of the *dmp* operon is probably the binding site for DmpR.

The location of the UAS binding sites physically separates activator proteins and the closed transcriptional complex formed between the RNA polymerase and the promoter. Models involving loop formation of the intervening DNA, to allow interaction of the upstream-bound activator and the promoter-bound polymerase, have been proposed (44). A role for integration host factor (IHF) in transcriptional regulation of *nif* and *xyl* genes has recently been demonstrated (see references 9 and 19). This histone-like, DNA-binding protein binds DNA with sequence specificity (14). IHF, bound to sites between the UAS and the promoter, is thought to stabilize loop formation and thus aid the interaction between the UAS-bound activator protein and the σ^{54} RNA polymerase. A sequence with high identity to the IHF binding consensus sequence is present within the *dmp* operon regulatory region (Fig. 8), although it is located further upstream from the promoter than has been found in other systems.

The nucleotide sequence upstream of the *dmp* operon contains putative regulatory signals similar to those identified as involved in the regulation of expression from promoters recognized by the minor form of RNA polymerase containing σ^{54} (Fig. 8): namely, a −24, −12 promoter sequence, a putative UAS within a long inverted repeat, and a consensus IHF binding sequence. In addition to the *dmpR* mutant strain, three other independent phenol-regulatory mutants that were located within chromosomal genes of the host bacterium were identified. Although it is not known what gene(s) is inactivated in these strains, in view of the above observations, likely candidates include genes encoding a σ^{54} factor and an IHF homolog. The nature of the defects in these chromosomal mutants is currently being investigated.

ACKNOWLEDGMENTS

We thank Justin Powlowski and Martin Gullberg for critical reading of the manuscript.

This work was supported by the Swedish Research Council for Engineering Science (TRF 91-419) and the Swedish Natural Science Research Council (NFR 9840-301).

REFERENCES

1. **Abril, M.-A., M. Buck, and J. L. Ramos.** 1991. Activation of the *Pseudomonas* TOL plasmid upper pathway operon. Identification of binding sites for the positive regulator XylR and for the integration host factor protein. J. Biol. Chem. **266:**15832–15838.
2. **Abril, M.-A., C. Michan, K. N. Timmis, and J. L. Ramos.** 1989. Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for the degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. J. Bacteriol. **171:**6782–6790.
3. **Ausubel, F. M.** 1984. Regulation of nitrogen fixation. Cell **37:**5–6.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology, vol. 1. John Wiley and Sons, Inc., New York.
5. **Bartilson, M., I. Nordlund, and V. Shingler.** 1990. Location and gene organization of the dimethylphenol catabolic genes of *Pseudomonas* CF600. Mol. Gen. Genet. **220:**294–300.
6. **Bartilson, M., and V. Shingler.** 1989. Nucleotide sequence and expression of the catechol 2,3-dioxygenase-encoding gene of phenol catabolizing *Pseudomonas* CF600. Gene **85:**233–238.
7. **Collado-Vides, J., B. Magasanik, and J. D. Gralla.** 1991. Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. **55:**371–394.
8. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn*5* transposon derivatives for insertional mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J. Bacteriol. **172:**6568–6572.

9. **de Lorenzo, V., M. Herrero, M. Metzke, and K. N. Timmis.** 1991. An upstream XylR- and IHF-induced nucleoprotein complex regulates the σ^{54}-dependent Pu promoter of TOL plasmid. EMBO J. **10:**1159–1167.
10. **Dixon, R.** 1986. The *xylABC* promoter from *Pseudomonas putida* TOL plasmid is activated by nitrogen fixation regulatory genes in *Escherichia coli*. Mol. Gen. Genet. **203:**129–136.
11. **Drummond, M., P. Whitty, and J. Wooton.** 1986. Sequence domain relationships of *ntrC* and *nifA* from *Klebsiella pneumoniae*: homologies to other regulatory proteins. EMBO J. **5:**441–447.
12. **Ebright, R. H.** 1986. Proposed amino acid-base pair contact for 13 sequence DNA binding proteins, p. 207–229. *In* D. Oxender (ed.), Protein structure, folding and design. Alan R. Liss, New York.
13. **Franklin, F. C. H., M. Bagdasarian, M. M. Bagdasarian, and K. N. Timmis.** 1981. Molecular and functional analysis of the TOL plasmid pWW0 from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring *meta*-cleavage pathway. Proc. Natl. Acad. Sci. USA **78:**7458–7462.
14. **Friedman, D. I.** 1988. Integration host factor: a protein for all reasons. Cell **55:**545–554.
15. **Frischauf, A.-M., H. Lehrach, A. Poustaka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:**827–842.
16. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109–136. *In* D. M. Glover (ed.), DNA cloning, vol. 1. A practical approach. IRL Press Ltd., Oxford.
17. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
18. **Holtel, A., M.-A. Abril, S. Marques, K. N. Timmis, and J. L. Ramos.** 1990. Promoter-upstream activator sequences are required for expression of the *xylS* gene and upper-pathway operon of the *Pseudomonas* TOL plasmid. Mol. Microbiol. **4:**1551–1556.
19. **Hoover, T. R., E. Santero, S. Porter, and S. Kustu.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. Cell **63:**11–22.
20. **Inouye, S., M. Gomada, U. M. X. Sangodkar, A. Nakazawa, and T. Nakazawa.** 1990. Upstream regulatory sequences for transcriptional activator XylR in the first operon of xylene metabolism of the TOL plasmid. J. Mol. Biol. **216:**251–260.
21. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1984. Nucleotide sequence surrounding the transcription initiation site of *xylABC* operon of TOL plasmid. Proc. Natl. Acad. Sci. USA **81:**1688–1691.
22. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1987. Overproduction of the *xylS* gene product and activation of the *xylDLEGF* operon on the TOL plasmid. J. Bacteriol. **169:**3587–3592.
23. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1988. Nucleotide sequence of the regulatory gene *xylR* of the TOL plasmid from *Pseudomonas putida*. Gene **66:**301–306.
24. **Joseph-Liauzun, E., R. Fellay, and M. Chandler.** 1989. Transposable elements for efficient manipulation of a wide range of Gram-negative bacteria: promoter probes and vectors for foreign genes. Gene **85:**83–89.
25. **Köhler, T., S. Harayama, J.-L. Ramos, and K. T. Timmis.** 1989. Involvement of *Pseudomonas putida* RpoN sigma factor in regulation of various metabolic functions. J. Bacteriol. **171:**4326–4333.
26. **Kushner, S. R.** 1978. An improved method for transformation of *Escherichia coli* with ColE1 derived plasmids, p. 17–23. *In* H. W. Boyer and S. Nicosia (ed.), Genetic engineering. Elsevier/North-Holland Publishing Co., Amsterdam.
27. **Kustu, S., A. K. North, and D. S. Weiss.** 1991. Prokaryotic transcriptional enhancers and enhancer-binding proteins. Trends Biochem. Sci. **16:**397–402.
28. **Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53:**367–376.
29. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
31. **Mermod, N., P. R. Lehrbach, W. Reineke, and K. N. Timmis.** 1984. Transcription of the TOL plasmid toluate catabolic pathway operon of *Pseudomonas putida* is determined by a pair of co-ordinately and positively regulated overlapping promoters. EMBO J. **3:**2461–2466.
32. **Nordlund, I., J. Powlowski, and V. Shingler.** 1990. Complete nucleotide and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J. Bacteriol. **172:**6826–6833.
33. **Nordlund, I., and V. Shingler.** 1990. Nucleotide sequences of the *meta*-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxymuconic semialdehyde hydrolase from *Pseudomonas* CF600. Biochim. Biophys. Acta. **1049:**227–230.
34. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55:**339–372.
35. **Powlowski, J., and V. Shingler.** 1990. In vitro analysis of polypeptide requirements of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J. Bacteriol. **172:**6834–6840.
36. **Ramos, J. L., N. Mermod, and K. N. Timmis.** 1987. Regulatory circuits controlling transcription of TOL plasmid operon encoding *meta*-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. Mol. Microbiol. **1:**293–300.
37. **Richaud, P., A. Colbeau, B. Toussaint, and P. M. Vignais.** 1991. Identification and sequence analysis of the $hupR_1$ gene, which encodes a response regulator of the NtrC family required for hydrogenase expression in *Rhodobacter capsulatus*. J. Bacteriol. **173:**5928–5932.
38. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1978. Simple method for the identification of plasmid-coded proteins. J. Bacteriol. **137:**692–693.
39. **Shine, J., and L. Dalgarno.** 1975. Determination of cistron specificity in bacterial ribosomes. Nature (London) **254:**34–38.
40. **Shingler, V., F. C. H. Franklin, M. Tsuda, D. Holroyd, and M. Bagdasarian.** 1989. Molecular analysis of a plasmid encoded phenol hydroxylase of *Pseudomonas* CF600. J. Gen. Microbiol. **135:**1083–1092.
41. **Shingler, V., J. Powlowski, and U. Marklund.** 1992. Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. J. Bacteriol. **174:**711–724.
42. **Stock, J. B., A. M. Stock, and J. M. Mottonen.** 1990. Signal transduction in bacteria. Nature (London) **344:**395–400.
43. **Stoker, K., W. N. M. Reijnders, L. F. Oltmann, and A. H. Stouthamer.** 1989. Initial cloning and sequence of *hydHG*, an operon homologous to *ntrBC* and regulating the labile hydrogenase activity in *Escherichia coli* K-12. J. Bacteriol. **171:**4448–4456.
44. **Su, W., S. Porter, S. Kustu, and H. Echols.** 1990. DNA-looping and enhancer activity: association between DNA bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. Proc. Natl. Acad. Sci. USA **87:**5504–5508.
45. **Thöny, B., and H. Hennecke.** 1989. The $-24/-12$ promoter comes of age. FEMS Microbiol. Rev. **63:**341–358.
46. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.