The Structure and Transcriptional Regulation of the Bacterial Degradative Genes for Chlorocatechols and 2,4-Dichlorophenoxyacetate

細菌のクロロカテコール及び2, 4-ジクロロフェノキシ酢酸分解遺伝子群の構造と転写調節

Naoto Ogawa

小川直人

Contents

List of Figures and Tables

Tables

Figures

Abbreviations

Ben: benzoate
3-CB: 3-chlorobenzoate
3-CC: 3-chlorocatechol
CCM: *cis,cis*-muconate
2-CM: 2-chloro-*cis,cis*-muconate
2,4-D: 2,4-dichlorophenoxyacetate
PCBs: polychlorinated biphenyls
4- (or 2-) CB: 4- (or 2-) chlorobenzoate

Introduction

Contamination of environment and its adverse effects are some of the major problems that we are facing today. The environment is the shared gift to all organisms and thus responsibilities of all persons, no matter where they live. Global environmental degradation is so severe, and accelerating so rapidly, that we will soon be forced to accept protection of the environment as our first priority as a matter of sheer survival (Drucker, 1994). Bioremediation is the process of using microorganisms to transform hazardous chemical compounds in an environment to nonhazardous end products. It is an effective method of decontaminating pollutants contaminated environment, while creating less environmental impact than other methods (Bandyopadhyay *et al.*, 1994; Fletcher, 1994). Most of other methods are not cost effective and are inappropriate for the in situ treatment. Some of them are not effective in treating a complex array of different pollutants. Biological treatment appears to offer solutions to these limitations. The basic advantages of bioremediation can be outlined as follows (Bandyopadhyay *et al.*, 1994): 1. Bioremediation is a natural process in which naturally occurring microorganisms are used for treatment of the wastes. 2. The residues or the by-products of biological processes (CO_2, H_2O) are usually geochemically cycled in the environment as harmless products. 3. Microorganisms have a wide range of abilities to metabolize different chemicals. 4. Where possible, the technologies are developed to utilize and improve native microorganisms that have been demonstrated to degrade the pollutants on the site. In some cases, microorganisms known to metabolize the pollutants can be introduced and supplemented to improve biodegradation. 5. For in situ treatment of soils, sludge, and ground water, bioremediation is less expensive and less disruptive than options frequently used for treatment, such as excavation followed by incineration or landfilling.

Since bioremediation harnesses the ability of the degradative enzymes that microorganisms produce, it is critical to study the biochemical and genetic basis of microbial degradation to find ways to improve the degradative ability of microorganisms as well as to understand the evolution of the genes for degradative enzymes. Furthermore, even if the genes are present, functional expression of the genes is essential for the appropriate degradative enzymes to be produced. Various factors present in the environment are known to activate or repress gene expression and thereby modulate microbial activities (Bandyopadhyay *et al.*, 1994; Daubaras and Chakrabarty, 1992). If bioremediation is to succeed, it is imperative that we learn how

the environment controls evolution of new genes in microorganisms and the expression of microbial genes (Daubaras and Chakrabarty, 1992). In view of these interests in future application to bioremediation and in study of evolution, structure of bacterial genes for biodegradation and the expression of the genes are examined and discussed in this study.

Among all types of pollutants, chlorinated aromatic compounds pose one of the most serious contemporary environmental problems worldwide because they have been used in large quantities or released in the environment as solvents, herbicides, pesticides, polychlorinated biphenyls (PCBs) and other chemicals such as dioxins. Although some of these compounds persist in the environment for long periods, others are degraded by microorganisms in soil and water (Alexander, 1981). In the oxidative bacterial degradation of chlorinated aromatics, reactions known collectively as the modified *ortho*-cleavage pathway play a pivotal role by the degradation of chlorocatechols produced through converging pathways from various chlorinated aromatics (Fig. 1) (reviewed in Daubaras and Chakrabarty, 1992; Harwood and Parales, 1996; Reineke, 1998; Schlömann, 1994; van der Meer *et al.*, 1992, van der Meer, 1997). Chlorocatechols, however, are toxic to bacterial cells (Fritz *et al.*, 1992) and cannot be used as substrates for growth of bacteria in most cases. Undisturbed growth of organisms with some chlorinated aromatics requires well balanced production of the enzymes involved in the degradation to avoid accumulation of toxic metabolites such as chlorocatechols (Fritz *et al.*, 1992). Instead, chlorobenzoates which can be growth substrates of bacteria have been used frequently to study the degradation of chlorinated aromatics because they are intermediates of cometabolism of PCBs and are soluble in water (Chatterjee *et al.*, 1981). Among the isomers of chlorobenzoates, 3-chlorobenzoate (3-CB) is degraded through 3- or 4-chlorocatechol via the modified *ortho*-cleavage pathway in the most common cases of aerobic soil bacteria (Fig. 1). The herbicide 2,4-dichlorophenoxyacetate (2,4-D) has been another representative model of biodegradation because it has been a major agrochemical and is soluble in water. 2,4-D is also degraded via the modified *ortho*-cleavage pathway through 3,5-dichlorocatechol (Fig. 1). Thus, bacterial genes for degradation of 3-CB or 2,4-D are studied in this thesis, especially those encoding the enzymes of the modified *ortho*-cleavage pathway because of their central role.

Three evolutionarily related clusters of genes of gram-negative bacteria that encode enzymes in the modified *ortho*-cleavage pathway have been well described. These are *clcABDE*, *tfdCDEF*, and *tcbCDEF*. The *clcABDE* genes are responsible for degradation of 3-chlorocatechol (3-CC) and were cloned from plasmid pAC27 of

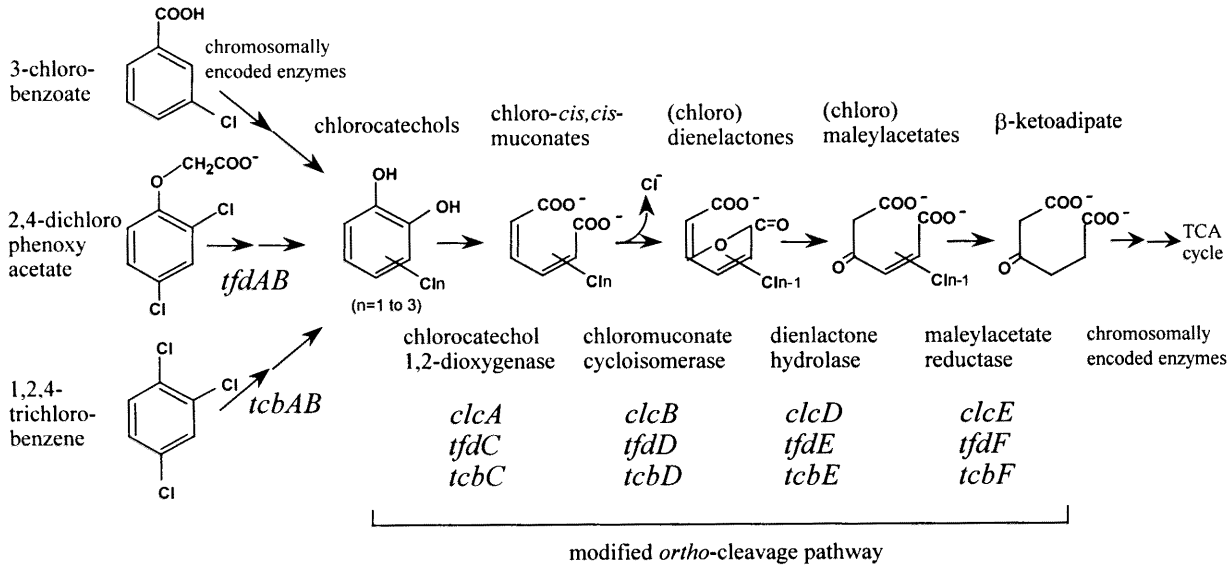


Fig. 1. Pathways for degradation of chlorinated aromatic compounds *via* modified *ortho*-cleavage pathway.

Pseudomonas putida, which is a 3-CB-degrading bacterium (Chatterjee and Chakrabarty, 1982; Chatterjee *et al.*, 1981; Frantz and Chakrabarty, 1987; Kasberg *et al.*, 1997). The *tfdCDEF* genes are present on plasmid pJP4 and are responsible for the degradation of 3,5-dichlorocatechol, which is produced from 2,4-D, by the products of the *tfdAB* genes in *Ralstonia eutropha* (*Alcaligenes eutrophus*) JMP134 (Don and Pemberton, 1981; Don *et al.*, 1985; Perkins *et al.*, 1990). The *tcbCDEF* genes are located on plasmid pP51 and are responsible for the degradation of 3,4,6-trichlorocatechol, generated from 1,2,4-trichlorobenzene by the products of the *tcbAB* genes in *Pseudomonas* sp. P51 (van der Meer *et al.*, 1991a, c). These three gene clusters of gram-negative bacteria have apparently evolved from common ancestral chlorocatechol genes (van der Meer *et al.*, 1991a). Recently, the study of both of the catechol and the chlorocatechol *ortho*-cleavage pathway of *Rhodococcus opacus* 1CP has shown that the chlorocatechol *ortho*-cleavage genes of this strain have evolved independently of those of gram-negative bacteria from common origin of the catechol *ortho*-cleavage genes in all bacteria (Eulberg *et al.*, 1997; Eulberg *et al.*, 1998).

The process of dissemination of the genes for biodegradation is interesting from the viewpoint of study of evolution of bacterial strains capable of biodegradation. The worldwide distribution of genes for chlorocatechol degradation has been speculated by

several isolates from different places (Amy *et al.*, 1985; Bhat *et al.*, 1994; Chatterjee and Chakrabarty, 1983; Chatterjee *et al.*, 1981; Chaudhry and Huang, 1988; Don and Pemberton, 1981; Ka *et al.*, 1994; Mäe *et al.*, 1993; Matheson *et al.*, 1996; Suwa *et al.*, 1996; Top *et al.*, 1995; van der Meer *et al.*, 1991c) and has recently been demonstrated more systematically by the studies of Fulthorpe *et al.* (Fulthorpe *et al.*, 1995; Fulthorpe *et al.*, 1998; Leander *et al.*, 1998). However, as to the means of the dissemination of the gene clusters, there have been only a few examples of identical or highly homologous plasmids which carried the modified *ortho* pathway genes thus indicated the transfer by plasmids (Amy *et al.*, 1985; Chatterjee and Chakrabarty, 1983; Don and Pemberton, 1981). Although the similar operon-like structures of the gene clusters of the modified *ortho*-cleavage pathway suggest they might have spread as units on transposable element, there have been only two documented examples of transposable elements that carry the modified *ortho* pathway genes. One is the ISJP4 composite transposon (Leveau and van der Meer, 1997) and the other is the *clc* element (Ravatn *et al.*, 1998b). The ISJP4 composite transposon carries the genes *tfdS-R-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K* on plasmid pJP4 in *R. eutropha* JMP134. The 105-kb *clc* element has recently been found to carry the *clcR-ABCD* genes in *Pseudomonas* sp. strain B13.

In contrast to the above two transposable elements, this thesis presents the first example of the class I composite transposon containing the modified *ortho* pathway genes which are captured between the two identical insertion sequences (ISs).

A few reports of transposons that carry the genes for degradation of chlorinated aromatics other than 3-CB have appeared in the literature (reviewed in Tsuda, 1996; Tsuda *et al.*, 1999; Wyndham *et al.*, 1994). The *tcbAB* gene cluster of plasmid pP51 encoding enzymes that convert 1,2,4-trichlorobenzene to 3,4,6-trichlorocatechol, is located on a composite transposon (Tn*5280*) with two slightly different insertion sequences (IS*1066* and IS*1067*), in inverted orientation, at its ends (van der Meer *et al.*, 1991d). A 17-kb segment of plasmid pBRC60, specifying enzymes for degradation of 3-CB in *R. eutropha* BR60, was found to be a composite transposon (Tn*5271*) flanked by directly repeated 3.2-kb segments of class II insertion sequences (IS*1071*) (Nakatsu *et al.*, 1991). Strain BR60 of *R. eutropha* is unique in that it degrades 3-CB by way of protocatechuate *meta*-ring-fission, and its catabolic genes are different from the degradative genes of plasmid pAC27 (Nakatsu and Wyndham, 1993; Wyndham *et al.* 1988). The catabolic transposon Tn*4371* (59 kb) found in *R. eutropha* A5 carries the genes to convert biphenyl and 4-chlorobiphenyl to benzoate (Ben) and 4-chlorobenzoate, respectively (Springael *et al.*, 1993). This transposon has an integrase system and conjugative transfer genes, the characteristics of the conjugative

transposons. Although the conjugative transfer of Tn*4371* has not been demonstrated, its transposition between replicons has been observed and the second mobile element, Tn-*bph,* which shares the right half of Tn*4371* was shown to transfer by conjugation (Merlin *et al.*, 1999).

With most of the above examples of the transposons containing the genes for degradation for chlorinated aromatics, identical or highly homologous gene clusters which were not carried by transposons have not been reported. Therefore the process of the dissemination of the degradative genes are not clear in terms of their evolutionary time scale. In the present thesis, the strong homology between the *cbnR-ABCD* genes and the *tcbR-CDEF* genes illustrates the recent horizontal transfer of these genes for the modified *ortho* cleavage pathway.

Genetic rearrangements of the degradative genes by general recombination have been another focus of interest. There are a few reports of genetic recombination related to the degradation of 3-CB. Amplification and deamplification of the 4.3-kb fragment that contains the structural genes for degradation of 3-CC was demonstrated in *Pseudomonas* sp. strain B13 (Rangnekar, 1988). Amplification of cloned fragments containing the *clc* structural genes from plasmid pAC27 and from pJP4 was reported (Ghosal *et al.*, 1985). Inverted duplication of catabolic genes on pJP4 and pre-existing tandem duplication of catabolic genes on a related plasmid, pJP2, were also described (Ghosal and You, 1988; Ghosal *et al.*, 1985). Although each of these events seemed to have been mediated by homologous recombination, the nucleotide sequences responsible for the recombination have not been determined in any of these cases. The amplification and deamplification of the *clc* element in several bacterial strains have been described (Ravatn *et al.*, 1998a). While the deamplification of the element was attributed to homologous recombination, the process of amplification has not been elucidated.

This study underscores the significance of the recombination events in the assembly and reconstruction of the gene clusters for the degradation of chlorinated aromatics with two examples, 3-CC degradative genes on plasmid pENH91 in *R. eutropha* NH9 and 2,4-D degradative genes on plasmid pMAB1 in *Alcaligenes* sp. CSV90.

Diversities are exhibited not only in the overall structures of the gene clusters and the sequences of the respective degradative enzymes but also in the regulation of the expression of the genes. Sequence homology and overall structural similarity among the modified *ortho*-cleavage pathway operons and the catechol *ortho*-cleavage pathway operon *catBCA* (Aldrich *et al.*, 1987; Houghton *et al.*, 1995) indicate an evolutionary relationship (Schlömann, 1994; van der Meer *et al.*, 1992). Each operon has a

lysR-type regulatory gene (Schell, 1993) which is located upstream of and is divergently transcribed from the degradative gene clusters (Coco *et al.*, 1993; Rothmel *et al.*, 1990; van der Meer *et al.*, 1991b). For the *tfdCDEF* operon of plasmid pJP4, however, *tfdT*, the *lysR*-type regulatory gene originally located upstream of the operon, is inactivated by an insertion-sequence element and its function has been taken over by distantly located *tfdR* (Leveau and van der Meer, 1996; You and Ghosal, 1995).

With regard to the regulatory mechanism of the expression of the degradative genes, the operons *catR-BCA* and *clcR-ABD* have been studied both in vivo and in vitro (Chugani *et al.*, 1998, 1997; Coco *et al.*, 1994, 1993; Houghton *et al.*, 1995; McFall *et al.*, 1997a, b, c; Parsek *et al.*, 1995, 1994b,c, 1992; Rothmel *et al.*, 1990; reviewed in McFall *et al.*, 1998). The regulators CatR and ClcR bind specifically to the *catB* and *clcA* promoter regions, respectively, and activate the expression of the degradative genes upon recognition of inducer. The inducers of the *catBCA* and *clcABD* operons have been identified as intermediates of their respective pathways, *cis,cis*,-muconate (CCM) (Parsek *et al.*, 1992) and 2-chloro-*cis,cis*,-muconate (2CM) (McFall *et al.*, 1997c). In the 2,4-D degradation process, 2,4-dichloromuconate, the intermediate produced by TfdC, has been identified as an inducer of the *tfdCDEF* operon in vivo (Filer and Harker, 1997). DNA binding of the regulator TcbR to the *tcbC* promoter has been described (Leveau *et al.*, 1994), and initial in vivo characterization of the role of *tcbR* has been performed (Leveau and van der Meer, 1996; van der Meer *et al.*, 1991b). *P. putida* KT2442 harboring a plasmid containing the *tcbR-CDEF* genes was found to grow on 3-CB while KT2442 containing the *tcbCDEF* genes with inactivated *tcbR* gene grew on 3-CB at a much lower rate (van der Meer *et al.*, 1991b). *P. putida* and *R. eutropha* strains harboring plasmids containing the *tcbR-CDEF* genes showed elevated (chloro)catechol 1,2-dioxygenase activity towards 3-chlorocatechol after cultivation on 3-CB. This activity was not observed with cells of *R. eutropha* containing the *tcbCDEF* genes with *tcbR* inactivated (Leveau and van der Meer, 1996; van der Meer *et al.*, 1991b). These results indicated that *tcbR* was required for the efficient expression of *tcbC* which encodes chlorocatechol dioxygenase. Further analysis of this operon including the identification of inducer is yet to be done.

The *cbnR-ABCD* operon on plasmid pENH91, found in a 3-CB degradative bacterium *R. eutropha* NH9, is highly homologous to the *tcbR-CDEF* operon (95.6%~100% identity at the amino acids level, and identical 150-bp divergent promoter regions) and is responsible for the degradation of 3-CC. This thesis reports the transcriptional regulation of the *cbnR-ABCD* operon, which includes the identification of inducers and a change in the bending angle of the promoter region

upon recognition of inducer. These observations provide a view of a conserved transcriptional mechanism of regulation plus the independent evolution of inducer-recognizing specificity and DNA-binding property among the regulators of the *ortho*-cleavage pathway. Together with the results from in vivo analysis of the transcriptional activation of the *tfdC* and the *tfdA* promoter by *tfdS* of *Alcaligenes* sp. CSV90, the divergence of the LysR-type regulators are discussed in terms of their inducer-recognizing specificity.

Materials and Methods

Bacteria, plasmids, media, chemicals and enzyme assays. The strains, phages, and plasmids used in this study are listed in Table 1. All strains of *Escherichia coli* were grown in Luria-Bertani medium (Sambrook *et al.*, 1989) at 37°C. Strains of *Ralstonia eutropha* and *Pseudomonas putida* were grown at 30°C. Strain NH9 was routinely maintained on basal salts medium [per liter: $(NH_4)_2SO_4$, 1.1 g; K_2HPO_4, 2.29 g; KH_2PO_4, 0.9 g; $MgSO_4 \cdot 7H_2O$, 0.1 g; $MnSO_4 \cdot 4\sim6H_2O$, 0.025 g; $FeSO_4 \cdot 7H_2O$, 0.005 g; L-ascorbic acid, 0.005 g] that contained either 0.1% (ca. 6.39 mM) or 5 mM 3-chlorobenzoate (3-CB) (final pH 7.0). Luria-Bertani medium, nutrient broth medium (per liter: beef extract, 5 g; peptone, 5 g; NaCl, 2.5 g; [pH unadjusted]), glucose-yeast extract (GY) medium [per liter: glucose, 1 g; yeast extract, 1 g; K_2HPO_4, 0.5 g; $MgSO_4 \cdot 7H_2O$, 0.2 g; $Fe_2(SO_4)_3 \cdot 7H_2O$, trace ; (pH 6.8)] were also used for strains of *R. eutropha.*

As a selective medium for the transconjugants of *R. eutropha* NH9D and *P. putida* KT2440, either Davis-minimal medium (K_2HPO_4, 7 g; KH_2PO_4, 2 g; $MgSO_4 \cdot 7H_2O$, 0.1 g; $(NH_4)_2SO_4$, 1 g) or basal salts medium for NH9 (above) was used supplemented with 0.2% sodium citrate. M9 minimal medium (Sambrook *et al.*, 1989) with 0.2% mannitol, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 1 mM thiamine hydrochloride was used as the minimal medium for the selection of recipient strains in conjugations with transfers of plasmids from *R. eutropha* to *E. coli.* To obtain transformants of *P. putida* PRS4020, Pseudomonas Isolation Agar (Difco) was used.

Antibiotics were incorporated into media at the following final concentrations for the strains of *E. coli* and *R. eutropha*: ampicillin (Ap), 100 μg/ml; kanamycin (Km), 50 μg/ml; streptomycin (Sm), 25 μg/ml; tetracycline (Tc), 15 μg/ml. For *P. putida* PRS4020: carbenicillin (Cbpc), 1000 μg/ml; gentamicin (Gm) 10 μg/ml; Tc 200 μg/ml.

The ability of strain NH9 to grow on dichlorobenzene and on 1,2,4-trichlorobenzene was tested basically as described by van der Meer *et al.* (1991c). A preculture of NH9 was inoculated into liquid basal salts medium (above) supplemented either with 3.5 mmol of 1,2- or 1,4-dichlorobenzene/liter or with 3.2 mmol of 1,2,4-trichlorobenzene/liter and was incubated at 30°C.

2-Chloromuconate was produced by the enzymatic conversion of 3-chlorocatechol with chlorocatechol 1,2-dioxygenase (ClcA) by the method of McFall *et al.* (1997b). Quantitative determination of β-galactosidase activity in the reporter assay was performed by the method of Miller (Miller, 1972; McFall *et al.*, 1997c). Each experiment was performed in triplicate.

Table 1. Bacterial strains, phages, and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
E. coli		
B378	DH1(RP4) Ap^r Km^r Tc^r *tra*$^+$	Savard *et al.*, 1986
B387	DH1(pPSA842) Ap^r Sm^r Tc^r *mob*$^+$ *tra* λ *cos*	Savard *et al.*, 1986
JM109	*recA1 supE44 endA1 hsdR17gyrA96 relA1 thi* Δ(*lac-proAB*) F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]	Yanisch-Perron *et al.*, 1985
XLI-Blue	*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac*$^-$ F'[*proAB*$^+$ *lacI*^{q} *lacZ*ΔM15 Tn*10*(Tet^r)]	Bullock *et al.*, 1987
TG1	*supE hsd*Δ*5 thi* Δ(*lac-proAB*) F'[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]	Sambrook *et al.*, 1989
DH5α	*supE44* Δ*lacU169*(ϕ80*lacZ*ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Hanahan, 1985
BL21(DE3)pLysS	*hsdS gal* (λ*cIts*857 *ind1 Sam7 nin5lacUV5*-T7 gene 1) pLysS Cm^r	Novagen
S17-1	C600::RP4 2-(Tc::Mu)(Km::Tn*7*) *thi pro hsdR hsdM*$^+$ *recA*	Simon *et al.*, 1983
P. putida		
KT2440	$r^- m^+$	Bagdasarian, 1981
PRS4020	*catR* ::Gm^r Ben^- derivative of PRS2000	Parales and Harwood, 1993
R. eutropha		
NH9	pENH91, 3-CB^+	This study
NH9D	3-CB^- derivative of NH9	This study
NH9d5	pENH91d5, 3-CB^-	This study
NH9A	pENH91A, 3-CB^+	This study
Phages		
M13mp18,19	*lacZ*α$^+$	Norrander *et al.*, 1983
Plasmids		
pBluescript KS(-), SK(-)	*lacZ*α$^+$ Ap^r	Stratagene
pUC18, 19	*lacZ*α$^+$ Ap^r	Yanisch-Perron *et*

		al., 1985
pPSA842	Ap^r Sm^r Tc^r mob^+ *tra* λ *cos*	Savard *et al.*, 1986
pKT230	Km^r Sm^r mob^+	Bagdasarian, 1981
pDC100	pMMB22 carrying the 4.3-kb *Bgl*II E fragment of pAC27; Ap^r	Frantz *et al.*, 1987
pEKC1	9.2-kb *Sac*I *cbnR-ABCD* insert from pENH91 in pKT230; Km^r	This study
pEUDR2	pUC18 carrying the 3.3-kb *Sal*I-*Sph*I fragment containing IS*1600* (DR2) of pENH91; Ap^r	This study
pELDR1	pBluescript KS(-) carrying the 4.7-kb *Eco*RI fragment containing IS*1600* (DR1) of pENH91; Ap^r	This study
pBLcbn1	2.8-kb *Hin*dIII *cbnR-AB'* insert from pEKC1 in pBluescript KS(-); vector mcs *Eco*RI is located on downstream of *cbnR*; *cbnB* is truncated at aa 270 out of 370; Ap^r	This study
pBLcbn1r	pBLcbn1 insert in reverse orientation; vector mcs *Xho*I and *Sal*I are located on downstream of *cbnR*; Ap^r	This study
pBLcbn2	5.6-kb *Eco*RI-*Xho*I *cbnABCD* insert from pEKC1 in pBluescript KS(-); Ap^r	This study
pBLcbn3	6.7-kb *Eco*RI-*Xho*I *cbnR-ABCD* insert in pBluescript KS(-); Ap^r	This study
pBLcbn3H	6.6-kb *Xba*I-*Xho*I *cbnR-ABCD* insert with 6 His codons on C-terminal of *cbnR* in pBluescript KS(-); Ap^r	This study
pT7-7	T7 promoter expression vector; Ap^r	Tabor and Richardson, 1985
pT7cbnR	956-bp *Nde*I-*Bam*HI PCR-generated fragment containing *cbnR* in pT7-7 under T7 promoter; Ap^r	This study
pT7cbnRHis	911-bp *Nde*I-*Hin*dIII fragment of *cbnR* with 6 His codons on C-terminal generated by PCR in pT7-7 under T7 promoter; Ap^r	This study
pET16b	T7 promoter expression vector for constructing His fusion proteins; Ap^r	Novagen
pEHiscbnR	956-bp *Nde*I-*Bam*HI *cbnR* insert from pT7cbnR in pET16b; Ap^r	This study
pCP13	Broad-host-range cosmid vector; mob^+tra^-; Km^r, Tc^r	Darzins and Chakrabarty, 1984
pCbn13ABCD	5.6-kb *Xba*I-*Xho*I *cbn-ABCD* insert from pBLcbn2 in pCP13; Tc^r	This study
pCbn13RABCD	6.7-kb *Xba*I-*Xho*I *cbnR-ABCD* insert from pBLcbn3 in pCP13; Tc^r	This study
pCbn13RHABCD	6.6-kb *Xba*I-*Xho*I *cbnRHis-ABCD* insert from pBLcbn3H in pCP13; Tc^r	This study
pQF50	Broad-host-range *lacZ* promoter probe vector; Ap^r	Farinha and Kropinski, 1990

pNO50RAB'	2.8-kb *Hin*dIII *cbnR-AB'* insert from pBLcbn1 in pQF50; *cbnB* is truncated at aa 270 out of 370; Ap^r	This study
pNO50AB'	1.7-kb *Nco*I-*Hin*dIII *cbn-AB'* insert in pQF50; Nearly complete *cbnA* promoter region is included; *cbnB* is truncated at aa 270; Ap^r	This study
pNO50RA'	1.4-kb *Sal*I *cbnR-A'* insert from pBLcbn1r in pQF50; *cbnA* is truncated at aa 60 out of 251; Ap^r	This study
pNO50RHAB'	2.6-kb *Hin*dIII *cbnRHis-AB'* insert from pBLcbn3H in pQF50; *cbnB* is truncated at aa 270; Ap^r	This study
pMP7	In vitro transcription vector; Ap^r	Hershberger *et al.*, 1995
pMPcbn1	417-bp *Pst*I-*Bam*HI insert containing the *cbnR-A* divergent promoter region from pBLcbn1 in pMP7; Ap^r	This study
pKS100	4.2-kb *Pst*I-*Bam*HI fragment from pUS1028 in pUC119; Ap^r	Daubaras *et al.*, 1995
pJET41	Supercoiled vector used for estimation of transcript size; Ap^r	Erickson and Gross, 1989

Abbreviations: Ap^r, ampicillin resistant; Gm^r, gentamicin resistant; Km^r, kanamycin resistant; Tc^r, tetracycline resistant; aa, amino acid; mcs, multicloning site; Ben; benzoate; 3-CB, 3-chlorobenzoate; r, host-specific restriction; m, host-specific modification of DNA.

DNA manipulations. Recombinant plasmids from *E.coli* were prepared either by the alkaline lysis method of Kieser (1984) or by the "small-scale preparations" method described by Sambrook *et al.* (1989). Total DNA from *R. eutropha* was prepared as described by Schmidt *et al.* (1986). Other recombinant DNA techniques including digestion with restriction endonucleases, DNA ligations and transformation of *E. coli* with plasmid DNA were performed basically as described by Sambrook *et al.* (1989). Plasmid pUC18, pUC19, pBluescript KS(-) and SK(-) were used routinely for subcloning of DNA fragments. Transformation of *E. coli* JM109, XLI-Blue, or DH5α was carried out either by the method of Hanahan (1983) or by electroporation with a Gene Pulser apparatus (Bio-Rad Laboratories, Richmond, Calif.) (Smith *et al.*, 1990). Transformation of the plasmid pQF50 and its derivatives into *P. putida* PRS4020 were performed as described previously (McFall *et al.*, 1997c). Mobilization of the plasmid pCP13 and its derivatives into *P. putida* PRS4020 and *R. eutropha* NH9D were conducted based on the method of Franklin (1985).

Cloning of genes for the catabolism of 3-CB. Large plasmids from *R. eutropha* were prepared by the method of Casse *et al.* (1979), followed by cesium chloride/ethidium bromide centrifugation. Plasmid pENH91 was digested partially with *Sau*3A to generate fragments of predominantly 30 to 50 kb and subjected to centrifugation in a 10% to 40% sucrose density gradient. Fractions containing DNA fragments of 30 to 50 kb were pooled. This DNA was inserted into *Bam*HI-digested broad-host-range cosmid pPSA842, packaged with a packaging extract (Gigapack Plus; Stratagene) by the procedure described by the manufacturer, and transduced into *E. coli* B378. The individual cosmid clones were mobilized into strain NH9D, a 3-CB$^-$ derivative of NH9 that had been cured of plasmid pENH91, as described by Franklin (1985). Transconjugants were then screened for growth on plates that contained 0.1% 3-CB and streptomycin (25 mg/ml). Plasmid pKT230 was used for subcloning of the genes for catabolic enzymes to test complementation of growth. *P. putida* KT2440 was used as the host strain to test complementation of growth by DNA fragments subcloned in pKT230.

Conjugation and incompatibility test. Conjugation was perfomed as described by Franklin (1985). The conclusion of incompatibility tests was based on the observation of the appearance of the transconjugants of *R. eutropha* that retained the ability to use 3-CB and exhibited the phenotype of antibiotic resistance of introduced plasmid.

Southern hybridization. Southern blotting to nylon membranes (Hybond-N; Amersham International plc, Amersham, UK) was performed by the standard procedure (Sambrook *et al.*, 1989) or with a vacuum blotting apparatus (VacuGene; Pharmacia LKB Biotechnology Inc., Uppsala, Sweden). DNA restriction fragments were labeled by nick-translation or random-priming (Sambrook *et al.*, 1989) with [α-^{32}P]dCTP (Amersham) except for cloning of the 3.7-kb *Sal*I fragment from strain P51. Hybridizations were performed overnight at 42°C as described by Sambrook *et al.* (1989). The hybridization solution consisted of (per 10 ml) 5.08 ml formamide, 0.50 ml Hepes (0.5 M, pH 7.5), 2.25 ml of a solution of nonhomologous DNA in water (1 mg/ml), 0.10 ml 100xDenhardt's solution, 0.56 ml distilled water and 1.50 ml 20 x SSC. Washing was performed under high-stringency conditions: twice for 15 min with 2 x SSC at 65°C, once for 30 min with 2 x SSC, 0.1% SDS at 65°C, and then once for 10 min with 0.1 x SSC at 65°C. Southern hybridization experiments for cloning of the 3.7-kb *Sal*I fragment from strain P51 were performed with a Digoxigenin Labeling Kit (Boehringer Mannheim, Mannheim, Germany) according to the protocol from the manufacturer.

Sequencing. For sequencing of the 3.3-kb *Sal*I-*Sph*I fragment containing DR2 and the relevant fragments containing the degradative genes of pENH91, they were cloned into M13mp18 and M13mp19, and nested deletions were introduced in both directions and clones with inserts of appropriate lengths were selected. Restriction fragments subcloned in pUC vectors and pBluescript vectors were also used for sequencing. Sequences were determined by the dideoxy chain-termination method (Sanger *et al.*, 1977) with automated sequencers (373A [Perkin-Elmer-Applied Biosystems Inc., Foster City, Calif.], ALFred [Pharmacia, Uppsala, Sweden], DSQ-1000L [Shimazu, Kyoto, Japan], and LIC-4200L2 [Li-Cor, USA]), with the dye-primer or dye-terminator kits recommended by the respective manufactures.

Construction of plasmids for the study of transcriptional activation by *cbnR*. Plasmids to test growth complementation by *cbnR* were constructed as follows: a plasmid, pBLcbn2, containing the *cbnABCD* genes with *cbnA* promoter region in pBluescript KS(-) was constructed using a 5.6-kb insert from pEKC1 which was cut out by a unique *Eco*RI site in the *cbnR-ABCD* operon (This *Eco*RI site which is 4 bp within the N-terminal of *cbnR* will be referred to as *Eco*RI*) and a *Xho*I site located downstream of the *cbnD* gene (Table 1). A 1.1-kb *Eco*RI fragment containing *cbnR* truncated at the *Eco*RI* site was cut out from pBLcbn1 and inserted into the *Eco*RI site of pBLcbn2 to restore the original structure of the *cbnR-ABCD* operon, yielding pBLcbn3. From pBLcbn2 and pBLcbn3, the relevant fragments were excised and inserted into the cloning site of pCP13, yielding pCbn13ABCD and pCbn13RABCD, respectively. The broad-host-range vector pQF50 was used to produce the transcriptional fusion constructs (Table 1, Fig. 12, 24, and 25).

To make constructs for expression of *cbnR* in *E. coli,* DNA fragments containing *cbnR* with *Nde*I and *Bam*HI restriction sites were synthesized by PCR using plasmid pBLcbn1 as a template. A PCR-generated fragment using primers CBNR1 (5'-TTT TCA TAT GGA ATT CCG GCA GCT-3') and CBNR2 (5'-TTT TGG ATC CCT GTC CAG CGT GA-3') was digested and inserted into the cloning sites of pT7-7, yielding plasmid pT7cbnR. To make a construct of *cbnR* with 6 His codons on its carboxyl-terminus (*cbnRHis*), a PCR-generated fragment using primers CBNR1 and CBNRHis1 (5'-TTT TAA GCT TCA ATG ATG ATG ATG ATG ATG GTC CTT CGC GGA TCG CCG CAC GTG TTC CAC GAA CC-3') was digested with *Nde*I and *Hin*dIII and was inserted into pT7-7, yielding pT7cbnRHis. Nucleotide sequences of the PCR-derived inserts of the two plasmids were verified by sequencing analysis. In these two constructs, the initiation codons of *cbnR* or *cbnRHis* were ligated with the *Nde*I site of the vector and thus were located 8-bp downstream of a vector derived

ribosomal binding site, and transcription was initiated from the T7 promoter of the vector. The insert *cbnR* was excised from pT7cbnR by digestion with *Nde*I and *Bam*HI and cloned into pET16b, yielding pEHiscbnR. In this construct, *cbnR* was fused downstream of 10 His codons and a Factor Xa site.

To demonstrate that CbnRHis has the same function as wild type CbnR, plasmids were constructed as follows: A 0.95-kb *Eco*RI*-*Xba*I fragment containing *cbnRHis* truncated near the amino-terminal was excised from a pUC18-based construct which was made using the insert from pT7cbnRHis, and was inserted into *Xba*I-*Eco*RI* site of pBLcbn2. The resulting plasmid, pBLcbn3H, contained a reconstructed *cbnR-ABCD* operon with 6 His codons attached to C-terminal of *cbnR*. Relevant fragments from pBLcbn3H were used (Table 1) to construct pCbn13RHABCD for growth complementation test and pNO50RHAB' for reporter assay in which *cbnB'* gene was connected to a promoterless *lacZ*. A supercoiled template for in vitro transcription assay, designated pMPcbn1, was constructed using the vector pMP7. The inserted fragment spanned positions -163 to +254 with respect to the *cbnA* transcriptional start site.

Purification of protein. Induction and extraction of protein from *E. coli* BL21(DE3)pLysS containing a construct of pT7-7 or pET16b and further purification of the protein with the histidine tag were conducted according to the His•Bind resin manual (Novagen, Madison, Wis.). In order to partially purify CbnR and also to eliminate a contaminating protein of 27 kd from crude lysate of cells with pT7cbnRHis, a heparin-agarose column was used as described previously (Coco *et al.*, 1994). The fractions that showed specific binding to the *cbnA* promoter fragment by gel retardation assay were pooled and concentrated by Centricon 10 (Amicon, Mass.). To prepare the vector (pT7-7) control for pT7cbnR, the fractions corresponding to the ones containing activity with pT7cbnR were used. CbnRHis was purified further by His•Bind resin.

Gel retardation assay. DNA binding reactions were performed in 20 μl volumes consisting of 10 mM HEPES (pH7.9), 10% glycerol, 100 mM KCl, 4 mM spermidine, 0.1 mM EDTA, 0.25 mM dithiothreitol, 1.5 μg of bovine serum albumin, 5 μg of heparin (Sigma, H-3125), approximately 0.1 ng of a DNA fragment and the tested protein. The DNA fragments were synthesized by PCR and labeled internally with [α-^{32}P]dCTP, followed by a purification procedure (Parsek *et al.*, 1994a). The binding reactions were initiated by the addition of CbnR and incubated for 20 min at room temperature. The binding reactions were then electrophoresed through 5% native polyacrylamide gel in 0.5 X TBE buffer for 2 hours at 130 V with circulation of cooling water in Bio-Rad Protean apparatus. The DNA probe containing the 252-bp

cbnR-A promoter region was made by PCR with primers, CBNFT1 (5'-TTG GCT GCT GCA GCC ATG TTC CC-3') and CBNFT2 (5'-AAT GCG GAC GCA ACC TGC TTC ACT CG-3'), and pBLcbn1 as a template. A 336-bp fragment containing a part of hydroxyquinol 1,2-dioxygenase gene from *Burkholderia cepacia* AC1100 was used as a nonspecific DNA binding probe. The fragment was synthesized by PCR using primers, ORF4P1 (5'-GCC TGC AGC GGC CCC TTC CAT GTG-3') and ORF4P2 (5'-GCC TGC AGC CTC AAG CAT TTG ACC-3'), and pKS100 as a template (Daubaras *et al.*, 1995).

S1 nuclease analysis. Bacterial RNA was isolated via the RNeasy total RNA isolation kit (Qiagen) from a 9-ml culture of *P. putida* PRS4020 containing the plasmid pNO50RAB' cultivated in Luria broth supplemented with 5 mM 3-CB. To prepare the DNA probe, primer CBNFT2 was end-labeled with T4 kinase (Gibco BRL, Gaithersburg, Md.) and [γ-^{32}P]ATP followed by PCR synthesis of a 252-bp fragment containing the promoter region with 2nd primer, CBNFT1 and pBLcbn1 as a template. The PCR product was purified by QIA quick PCR purification kit (Qiagen, Chatsworth, Calif.). This double stranded fragment was denatured by the addition of 1/10th volume of 2N NaOH, 2 mM EDTA, pH 8.0 and incubated at room temperature for 5 minutes. The denatured DNA was ethanol precipitated, washed with 70% ethanol and resuspended in water. The labeled DNA fragment was hybridized with the RNA at 45 °C overnight and incubated with S1 nuclease using S1-Assay kit (Ambion, Austin, Tex.) according to the manufacturer's instructions. Sequencing reactions to juxtapose the protected DNA fragment resulting from the S1 reaction were performed using SequiTherm cycle sequencing kit (Epicentre Technologies, Madison, Wis.) with CBNFT2 as the primer.

DNase I protection assay. DNase I footprinting experiments were conducted based on the method described previously (Parsek *et al.*, 1994a), except that the binding reaction was performed in 20 µl solution consisting of 100 mM Tris-HCl (pH7.9), 1 mM EDTA, 4% glycerol, 100 mM KCl, 1 µg of bovine serum albumin, 1 µg of poly(dI-dC), approximately 10 ng of a DNA fragment plus the purified protein indicated. PCR was used to generate a 252-bp fragment spanning the *cbnA* promoter region using pBLcbn1 as a template. In each reaction, one of the primers, CBNFT1 or CBNFT2, was end-labeled with T4 kinase (Gibco BRL) and [γ -^{32}P]ATP as described previously (Parsek *et al.*, 1994a). To locate the footprint, sequencing reactions were performed as described above with either of the end-labeled primers, CBNFT1 or CBNFT2, and pBLcbn1 as the template.

In vitro transcription assays. In vitro transcription assays were performed as described previously (Hershberger *et al.*, 1995; McFall *et al.*, 1997c). The supercoiled template, pMPcbn1 and *E.coli* holo RNA polymerase (Epicentre Technologies) were used. For each reaction, 100 ng of purified CbnRHis and 1 mM each of the chemicals tested as effector molecule were used.

DNA bending by circular permutation gel shift assay. Circular permutation gel shift assays were conducted by the same method as the gel retardation assay with the following modifications: the binding reactions contained 0.2 μg of purified CbnRHis, approximately 0.3 ng labeled DNA fragments and 2 μg heparin. In the inducer containing samples, CCM was added at 1 mM in the binding reactions and at 0.5 mM in the running buffer. Electrophoresis was performed at 130 V for 5 hours. Five DNA fragments of 257 bp containing the CbnR-binding sites at different positions from the ends to the middle were generated and labelled internally by PCR with [α-^{32}P]dCTP using the following primer pairs: MCBD1, 5'-GTT CGA TCC CGC GGT GGC TTC GCT CCA GAA GC-3' and MCBD2, 5'-GCG CCG GCC ATG CCG TCC AAT ACC-3'; MCBD3, 5'-TCC AGG GCT TGC ATC TGC CGC GTG ATGG-3' and MCBD4, 5'-TTG TCG GTT TGC CCG GTCC-3'; MCBD0, 5'-GGC GCT TGG CTG CTG CAG CCA TGT TCCC-3' and CBNFT2 (above); MCBD5, 5'-GAA ATA CTT GAG CTG CCGG-3' and MCBD6, 5'-TTC CGT CAC GCG TTG CTC CGT GAG GG-3'; MCBD7, 5'-GTG CCT ATA TTA CGC AAA CC-3' and MCBD8, 5'-TTG GCC TCG GCC AGT TTC ATC ATG TAG CC-3'. The DNA fragments were purified with QIA quick PCR purification kit (Qiagen). The bending angles were calculated as described previously (McFall *et al.*, 1997b).

Accession number of the nucleotide sequence. The nucleotide sequence of the 6959-bp *Sac*I-*Kpn*I region containing the *cbnR-ABCD* genes and the deduced amino acid sequences are deposited in the DDBJ, GenBank, and EMBL databases under accession number AB019032. The nucleotide sequence of IS*1600* (DR2) and orfL (from the *Sph*I site to outside of the *Kpn*I site of pENH91 in Fig. 4) and the deduced amino acid sequences are deposited under accession number D64144. The nucleotide sequence containing the 1300-bp region from strain P51, homologous to part of IS*1600*, is deposited under accession number AB019033.

Results

1. The structure of the 3-chlorocatechol degradative transposon Tn*5707*

R. eutropha NH9, isolated in Japan, grew on 3-chlorobenzoate (3-CB) abundantly. The genes responsible for 3-CB degradation cloned from a large plasmid of NH9 were suggested to be those for the modified *ortho*-cleavage pathway by a hybridization experiment and thus further characterized. A deletion event of the degradative genes from the plasmid led to the elucidation of the composite transposon structure capturing the degradative genes on the plasmid.

1.1 *R. eutropha* NH9 and its plasmids

R. eutropha NH9 was isolated from a soil sample at our institute under conditions where 3-CB was the sole source of carbon and energy, with subsequent subculturing on 3-CB-containing mineral salts agar plates. Taxonomic identification was carried out by K. Katoh (Personal communication) and was confirmed later by analysis of the sequence of 16s ribosomal DNA.

Mutants of NH9 unable to utilize 3-CB (3-CB$^-$) arose spontaneously and at high frequency (>10%) after repeated subculture (5 times) on glucose-yeast extract (GY) liquid medium. It appeared, therefore, that some of the functions necessary for degradation of 3-CB were plasmid encoded.

With a slight modification to the method of Casse *et al.* (1979), it was shown that NH9 harbored two large plasmids (Fig. 2, Lane 1). The yield of the larger plasmid (designated pENH92) was much lower than that of the smaller one (pENH91). Many of the 3-CB$^-$ mutants obtained as independent colonies after subculture on GY medium had been cured of pENH91 (Fig. 2, Lanes 2-5, 7, 8). Although pENH91 appeared to be retained in one of the 3-CB$^-$ mutants (Lane 6; designated NH9d5), the plasmid seemed to be smaller than the original one. These results suggested that some of the genes essential for the degradation of 3-CB were carried by pENH91.

pENH91 was transferred from wild-type NH9 cells to spontaneous 3-CB-negative Sm^r cells by conjugation. Transfer into the recipient strain occurred at a frequency of 10^{-4} per donor cell.

pENH91 was tested for its incompatibility with pKT230, pPSA842 (IncQ) and RP4 (IncP1). Plasmids pKT230 and pPSA842 were mobilized from *E. coli* S17-1 to NH9, and RP4 was transferred from *E. coli* B378 to NH9. The test was based on the observation of the appearance of transconjugants of *R. eutropha*, which can grow on

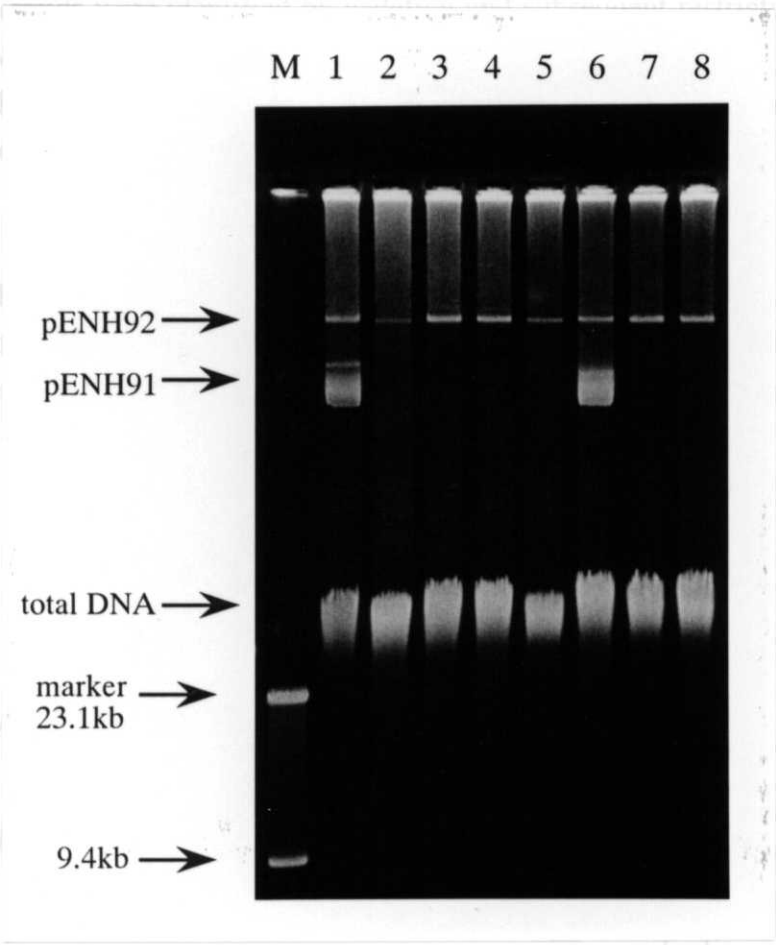


Fig. 2. Agarose gel electrophoresis of the plasmids from strain NH9 and from 3-CB$^-$ segregants (0.7% gel and TBE buffer). Lanes: M, lambda DNA digested with *Hin*dIII; 1, *R. eutropha* NH9; 2 to 8, 3-CB$^-$ segregants of *R. eutropha* NH9.

3-CB agar plates supplemented with either Km (for pKT230 and RP4) or Ap (for pPSA842). Transconjugants resistant to Km and Ap were obtained after introduction of pKT230 and pPSA842, respectively. The presence of the two plasmids (pENH91 plus either pKT230 or pPSA842) was verified after re-introduction of the plasmids into *E. coli*. The plasmids were identified by isolation and subsequent restriction analysis. Thus, pENH91 was found to be compatible with pKT230 and pPSA842. But it was found that pENH91 could not be maintained with RP4 in a stable manner. Thus, pENH91 was considered to be incompatible with RP4 and therfore is likely an IncP group plasmid.

1.2 Cloning of the genes for catabolic enzymes

A library of genes in pENH91 was constructed in *E. coli* B378 by use of the broad-host-range cosmid pPSA842. The individual cosmid clones were mobilized

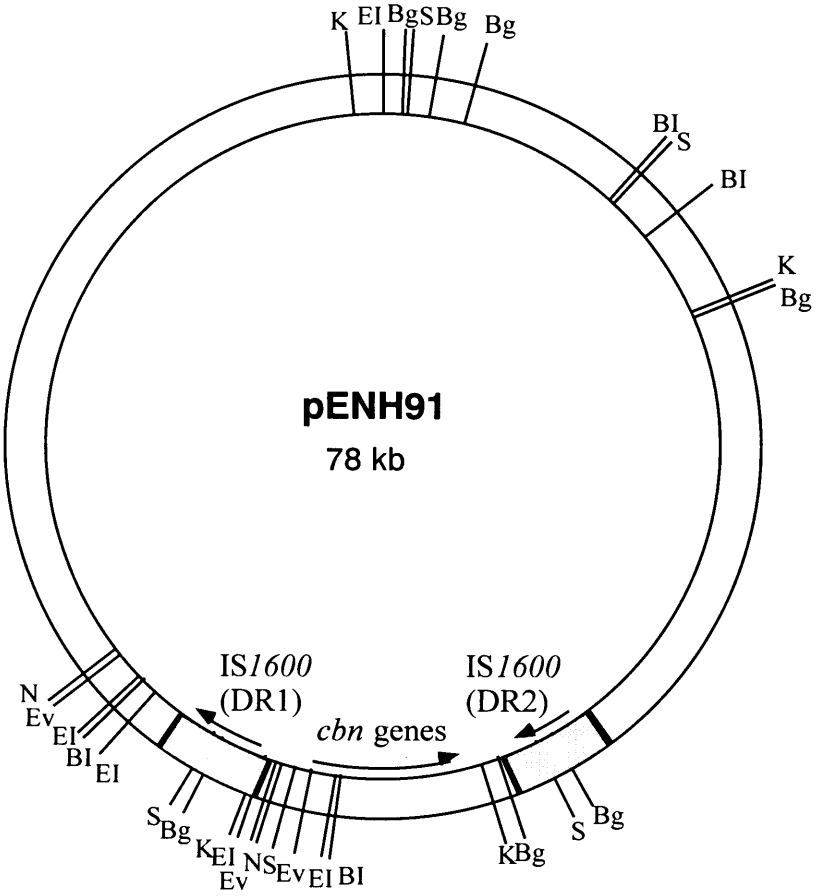


Fig. 3. Restriction map of pENH91. Restriction sites are abbreviated as follows: BI, *Bam*HI; Bg, *Bgl*II; EI, *Eco*RI; Ev, *Eco*RV; K, *Kpn*I; N, *Nhe*I; S, *Sac*I.

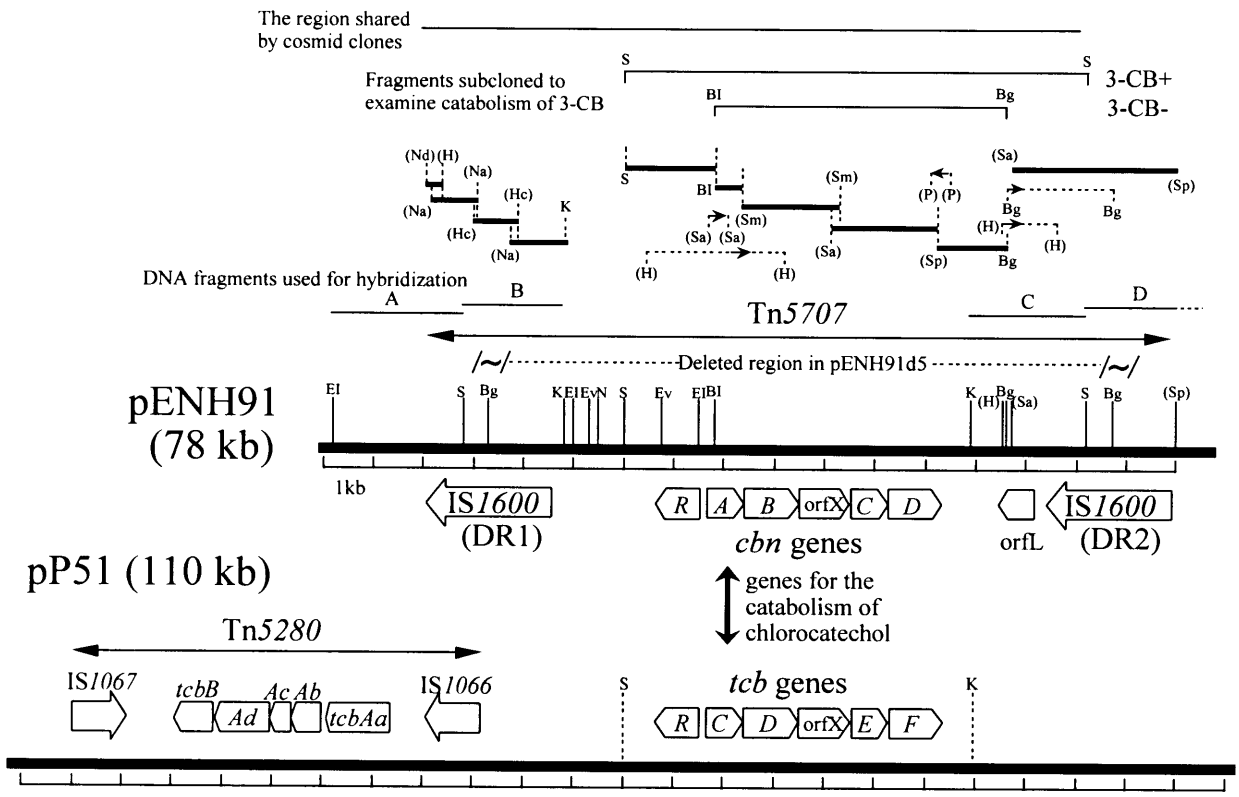


Fig. 4. Schematic representation of regions containing degradative genes and insertion sequences on plasmids pENH91 and pP51. The open arrows for the *cbn* genes, *orfL*, and the *tcb* genes show the locations and the directions of transcription of the open reading frames (orfs). The orientation of open arrows of IS*1600*, IS*1066*, and IS*1067* are in agreement with the direction of transcription of the orfs within the ISs. Horizontal lines above the linear restriction map show the fragments (A and B) used as probes in hybridization experiments and the hybridized fragments (C and D), respectively. The strategies for subcloning and sequencing the catabolic region on plasmid pENH91 are shown above the fragments for hybridization experiments. Fragments shared by cosmid clones or subcloned to examine the 3-CB phenotype are shown by thin solid lines at the top of the figure. The thick solid lines above the map of pENH91 indicate DNA fragments that were sequenced in both directions by using nested sets of deletions or subcloned restriction fragments. The thin dotted lines with small arrows indicate subcloned fragments used to sequence the boundary sites between the sequenced fragments described above. The small arrows indicate the lengths and directions of the sequences determined (5' to 3'). Restriction sites are abbreviated as follows, in addition to those defined in panel a: H, *Hin*dIII; Hc, *Hinc*II; Na, *Nae*I; Nd, *Nde*I; P, *Pst*I; Sa, *Sal*I; Sm, *Sma*I; Sp, *Sph*I. The restriction sites in parentheses are those determined only for subcloning of related fragments; thus, other sites recognized by such enzymes within the linear map were not determined before sequencing. The map of pP51 is based on material in references van der Meer *et al.*, 1991a, b, d, and Werlen *et al.*, 1996. Only the *Sac*I and *Kpn*I sites described in the text are shown for pP51.

from *E. coli* into *R. eutropha* NH9D, a 3-CB$^-$ derivative that had been cured of pENH91, and transconjugants were selected on minimal agar plates that contained 3-CB and streptomycin. Among about 200 clones examined, eight had the 3-CB$^+$ phenotype. A comparison of the restriction maps of the inserts of the positive clones showed that all included a common 13-kb region (Fig. 4). A physical map of pENH91 was constructed by further restriction analysis (Fig. 3). For subcloning, a 9.2-kb *Sac*I fragment (Fig. 4) from this 13-kb region was inserted into the broad-host-range vector pKT230 to yield pEKC1, which was mobilized into *R. eutropha* NH9D (3-CB$^-$) and *P. putida* KT2440. Cells of both strains harboring pEKC1 grew on 3-CB-supplemented mineral salts agar plates. A 5.8-kb *Bam*HI-*Bgl*II fragment from within the 9.2-kb *Sac*I fragment did not confer the 3-CB$^+$ phenotype on either strain. These results showed that the genes for catabolism of 3-CB were located within the 9.2-kb *Sac*I fragment.

1.3 Hybridization with chlorocatechol degradative genes from *P. putida*

To identify the determinants responsible for catabolism of 3-CB on pENH91, hybridization experiments were conducted with the chlorocatechol catabolic genes of pAC27 (Chatterjee and Chakrabarty, 1982), which had been originally isolated from *P. putida* AC866, as a probe. The 4.3-kb fragment containing the *clcABD* genes of pAC27, isolated from an *Eco*RI digest of pDC100, hybridized to the 5.8-kb *Bam*HI-*Bgl*II fragment from within the 9.2-kb *Sac*I fragment (data not shown). Therefore, the catabolic genes on pENH91 appeared to encode enzymes of the modified *ortho*-pathway.

1.4 Deletion of the catabolic genes from pENH91 by homologous recombination

A 12.5-kb region containing the 9.2-kb *Sac*I fragment (Fig. 4) was found to have been deleted from the plasmid harbored by one of the 3-CB-negative segregants (Fig. 2, Lane 6 [designated pENH91d5].) (A result of hybridization experiments is shown in lanes 3 of the Fig. 8a and b). Two fragments of pENH91, which encompassed either the left or right boundary region of the deleted segment, showed strong homology to each other in a hybridization experiment (Fig. 5; fragments B and D in Fig. 4). A 2.5-kb *Eco*RI-*Sac*I fragment (fragment A in Fig. 4) hybridized with a 2.3-kb *Kpn*I-*Sac*I fragment (fragment C) (data not shown). These results indicated that the two homologous fragments were in the same orientation and suggested that the deletion event had been due to homologous recombination between the two elements. On the restriction map (Fig. 4), the left homologous sequence was designated as DR1, and the right homologous sequence was designated as DR2. When the 2.0-kb *Sac*I-*Kpn*I

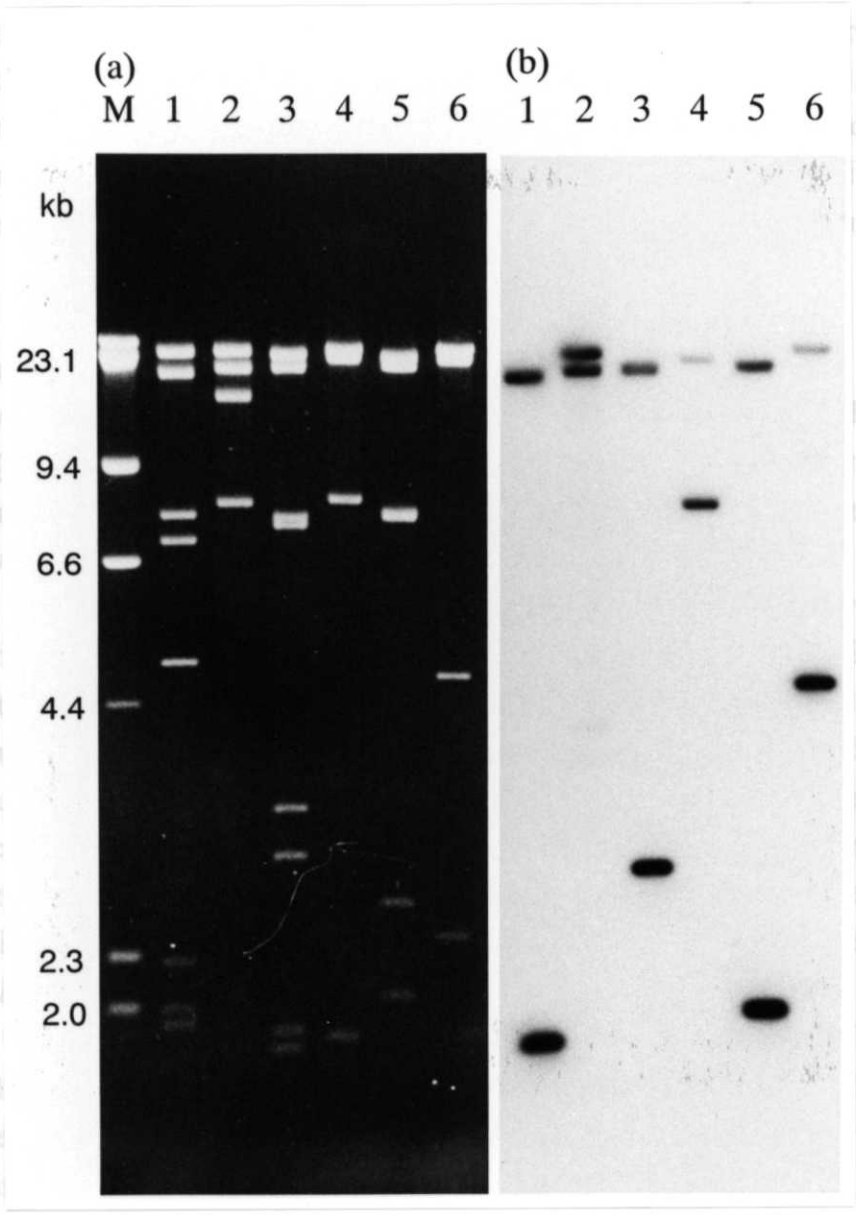


Fig. 5. Southern hybridization of fragments of pENH91 with the 2-kb *Sac*I-*Kpn*I fragment, designated B, as the labeled probe. (a) Agarose gel electrophoresis of fragments of pENH91 (0.7% gel and TBE buffer). Lanes: M, lambda DNA digested with *Hin*dIII as size standards; 1 to 6, pENH91 DNA digested with restriction enzymes (1, *Sac*I and *Kpn*I; 2, *Kpn*I; 3, *Sac*I and *Bam*HI; 4,*Bam*HI; 5, *Sac*I and *Eco*RI; 6, *Eco*RI.). (b) Southern hybridization of a filter prepared from the gel in panel a and probed with the 2-kb *Sac*I-*Kpn*I fragment B of pENH91 located on the left part of the linear map shown in Fig. 4. (the lanes correspond to those in panel a) In each lane, there are two hybridized bands. One contains fragment B used as the probe, and the other contains fragment D located on the far right of the linear map in Fig. 4.

fragment (fragment B in Fig. 4) isolated from pENH91 was used as a probe against pENH91d5, only one hybridized band appeared for pENH91d5 instead of two bands for pENH91 (data not shown). Further restriction analysis with pENH91d5 showed that this plasmid retained one sequence of the direct repeat elements and thus confirmed the deletion of catabolic region was due to homologous recombination between the two elements.

1.5 Identification of the direct repeats, DR1 and DR2, as identical ISs

After the boundary regions of DR1 and DR2 were determined by hybridization, a 3.3-kb *Sal*I-*Sph*I fragment containing DR2 was subcloned into M13 vectors and pUC18 (pEUDR2) and was sequenced. DR2 was 2,520 bp long including the inverted repeats (Fig. 6). Two long open reading frames (ORFs) were found on one DNA strand of DR2 (Fig. 6 and Fig. 7a). The amino acid sequences deduced from the two ORFs showed the highest homology to IstA and IstB of IS*1326* (Brown *et al.*, 1996) of the IS*21* family (Fig. 7a). Since DR2 was found to have the perfect structure of an IS of the IS*21* group, it was designated IS*1600* and the two genes were designated *istA* and *istB* (Fig. 6 and 7a). IS*1600* contained a set of 16-bp inverted repeats at its ends, with two mismatches (Fig. 7b). Interestingly, although the terminal inverted repeats of IS*1326* (26 bp) were considerably longer than those of IS*1600*, some nucleotides of IS*1600* proximal to the inverted repeats matched the nucleotides within the inverted repeats of IS*1326* (shown by asterisks in Fig. 7b).

A 4.7-kb *Eco*RI fragment containing DR1 was cloned from plasmid pENH91 (Fig. 4) to yield plasmid pELDR1. Sequencing analysis revealed that DR1 was identical to DR2. There were no target-site duplications flanking the extreme ends of the 15-kb region containing the two elements or the termini of either DR1 or DR2.

There were two possible initiation codons for *istA* of IS*1600*. This gene had the capacity to encode a protein of 518 (or, alternatively, 515) amino acids, if translation was initiated at the GTG codon at position106 (or position 115; Fig. 6). Each of these codons was preceded by a potential ribosome-binding site (SD1, SD2). *istB* encoded a putative protein of 264 amino acids, if the ATG codon of position 1,652 (Fig. 6) was used as an initiation codon. This codon was preceded by a possible ribosome-binding site (SD3).

The G+C content was 63.6% for IS*1600*, 52.1% for IS*1326*, and 52.3% for IS*21*. The G+C content of IS*1600* was similar to but slightly lower than the G+C content of the total genomic DNA of *R. eutropha* (66.3 to 66.8% [Davis *et al.*, 1969]). The G+C contents of the third positions in the codons used for IstA and IstB of IS*1600* were

83.4% and 80.4%, respectively. Those for IstA and IstB of IS*21* were 62.6 and 60.9%, respectively. By contrast to the genes of IS*21*, the codon usage in IstA and IstB of IS*1600* was highly biased in favor of G or C in the third position. This phenomenon has also been observed in other genes from *R. eutropha* (Andersen and Caton, 1987; Hein and Steinbüchel, 1994; Priefert *et al.*, 1991).

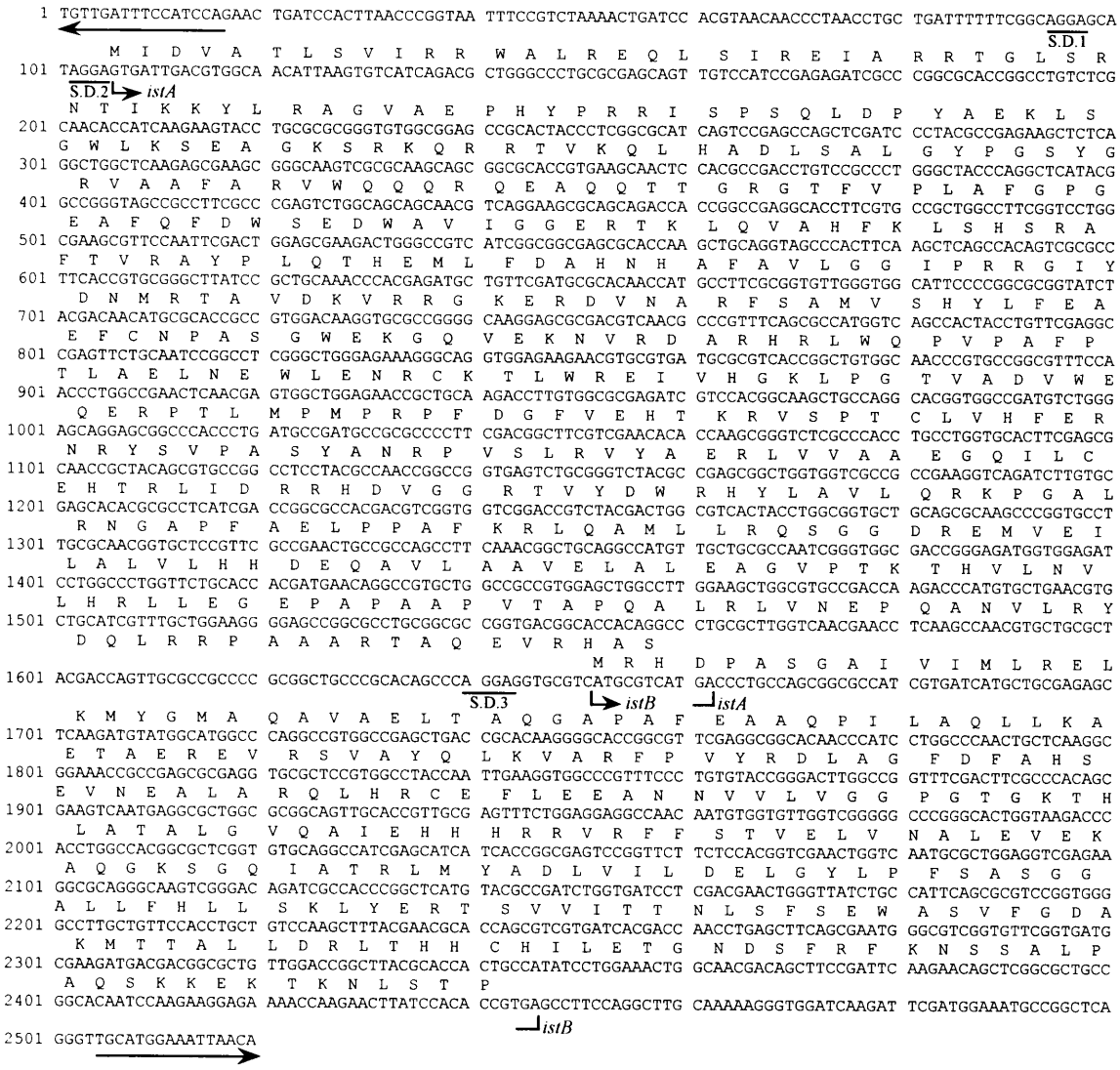
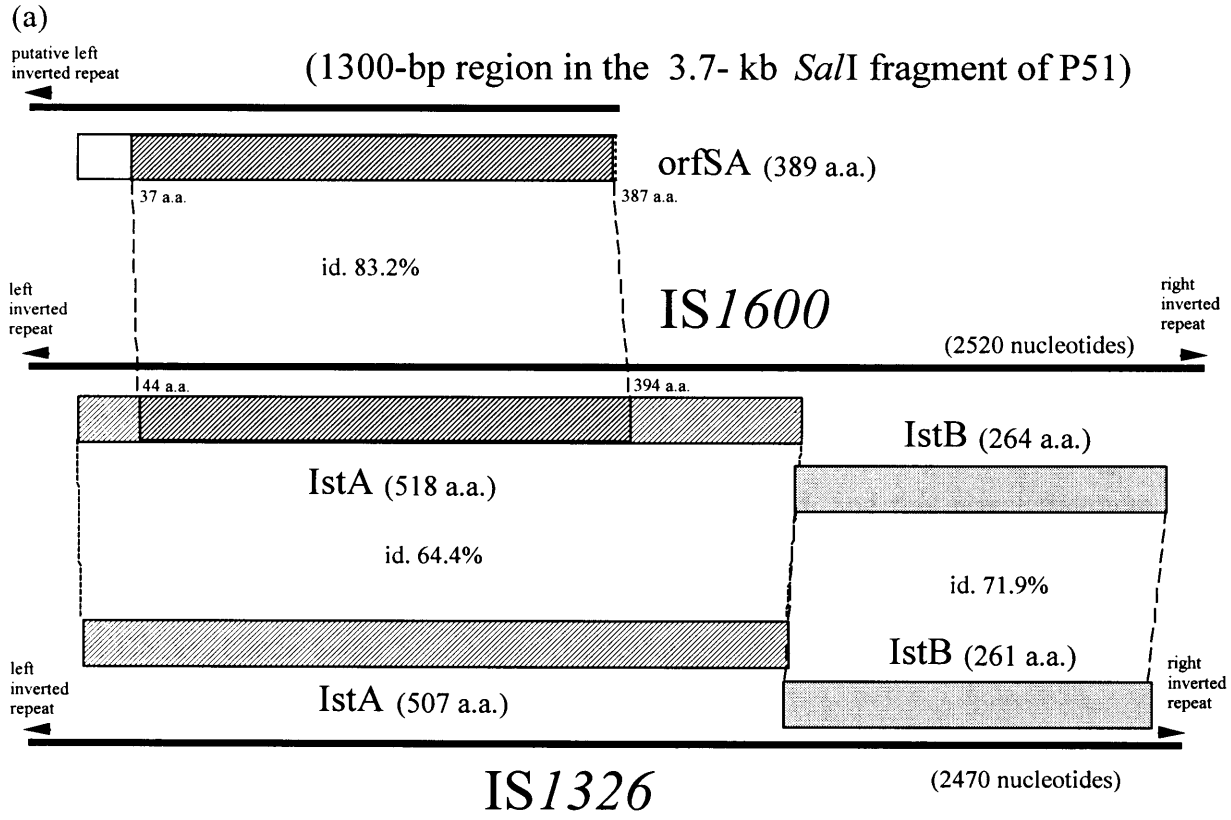


Fig. 6. Nucleotide sequence of IS*1600*. Long arrows show inverted repeats. The predicted amino acid sequences of *istA* and *istB* are shown above the first nucleotide of each codon. Shine-Dalgarno sequences are indicated as S.D.1 to -3.



(b)

```
                    left terminal                                  right terminal

P51SalI fragment    TGTTGATTTCCACCCggaagtgaccc~
                                  *** *** **
IS1600              TGTTGATTTCCATCCAgaactgatcc~   ~ggctcagggtTGCATGGAAATTAACA
                                  ** *** **      ** ***    *
IS1326              TGTTGAGTTGCATCTAAAATTGACCC~   ~GGGTCAAATTTGGGTGCAAATCAACA

IS21                TGTCAGCGCCA~                                ~TGGCGTTGACA
```

Fig. 7. (a) Schematic representation of the orfs in related IS or IS-like elements. The hatched areas indicate the regions with homology at the amino acid level. Percentages of identical amino acids (id.) are shown. (b) Terminal regions of the elements shown in panel a and IS*21*. The solid arrows indicate the inverted repeats and their orientations (Brown *et al.*, 1996; Reimmann *et al.*, 1989). The dotted arrow below the nucleotide sequence of the *Sal*I fragment from P51 shows a putative left inverted repeat that was delineated on the basis of similarity to the sequence of IS*1600*. Note that here the left and right designations refer to the directions of the orfs found in IS*1600*, as discussed in the text, and thus, they are the reverse of the directions shown in Fig. 4. The asterisks indicate identical nucleotides in the juxtaposed sequences that are not included in the inverted repeats of IS*1600* and the *Sal*I fragment from P51.

1.6 Duplication of the catabolic genes

After successive transfers of NH9 to fresh 3-CB-containing liquid medium at approximately two-week intervals for a year, the cells in the culture were compared with NH9 by hybridization experiments with both plasmid DNAs and total DNAs to examine whether any recombination events had occured. The cells and the catabolic plasmid from the one-year-old subculture were designated strain NH9A and pENH91A, respectively.

The 5.8-kb *Bam*HI-*Bgl*II fragment of the cloned 9.2-kb *Sac*I fragment was labeled and used as a probe to examine *Eco*RI digests of plasmid DNAs and total DNAs from the two strains (Fig. 8). In the *Eco*RI digest of pENH91, a 40-kb *Eco*RI fragment containing the probe region was found to have hybridized (Lane 1). In the *Eco*RI digest of pENH91A, an additional 10-kb fragment appeared, and this novel fragment also hybridized strongly with the probe (Lane 2). Hybridization with *Eco*RI digests of total cellular DNA from the two strains gave the same patterns as those of the respective plasmid DNAs (Lanes 4 and 5). These results suggested that the novel 10-kb *Eco*RI fragment on pENH91A also contained the catabolic genes.

Several restriction profiles of pENH91 and pENH91A were compared. One *Nhe*I site, one *Bam*HI site, two *Eco*RV sites, and two *Eco*RI sites were present within the intervening region between DR1 and DR2 on pENH91, but there were no sites for any of these four restriction endonucleases in the direct repeat sequences (Fig. 4). On the contrary, *Bgl*II sites were found in the direct repeats of pENH91 as well as within the intervening region. All the fragments that were present in the digests of pENH91 were observed in the corresponding restriction profile of pENH91A. In addition, restriction profiles of pENH91A with *Nhe*I, *Bam*HI, *Eco*RV, or *Eco*RI gave an additional fragment (12.5 kb, 12.5 kb, 11 kb, and 10 kb, respectively). The profiles of *Bgl*II digests of the two plasmids were identical (data not shown).

A 1.0-kb *Hin*dIII-*Bgl*II fragment of IS*1600* (DR2), when used as a probe, hybridized with two fragments that contained either DR2 or DR1 in each digest and with the additional fragments of pENH91A listed above (data not shown). The intensities of the three hybridization bands in each of digests of pENH91A with *Nhe*I, *Bam*HI, *Eco*RV, or *Eco*RI were nearly equal. This result suggested that the molar amount of each of the three fragments in pENH91A was equal. Conversely, the hybridization patterns were the same for the two plasmids in the *Bgl*II digests.

When the 5.8-kb *Bam*HI-*Bgl*II fragment of the cloned 9.2-kb *Sac*I fragment was used as a probe, it hybridized with the novel fragments that appeared in the digests of pENH91A as well as with the fragments that initially contained the 5.8-kb region.

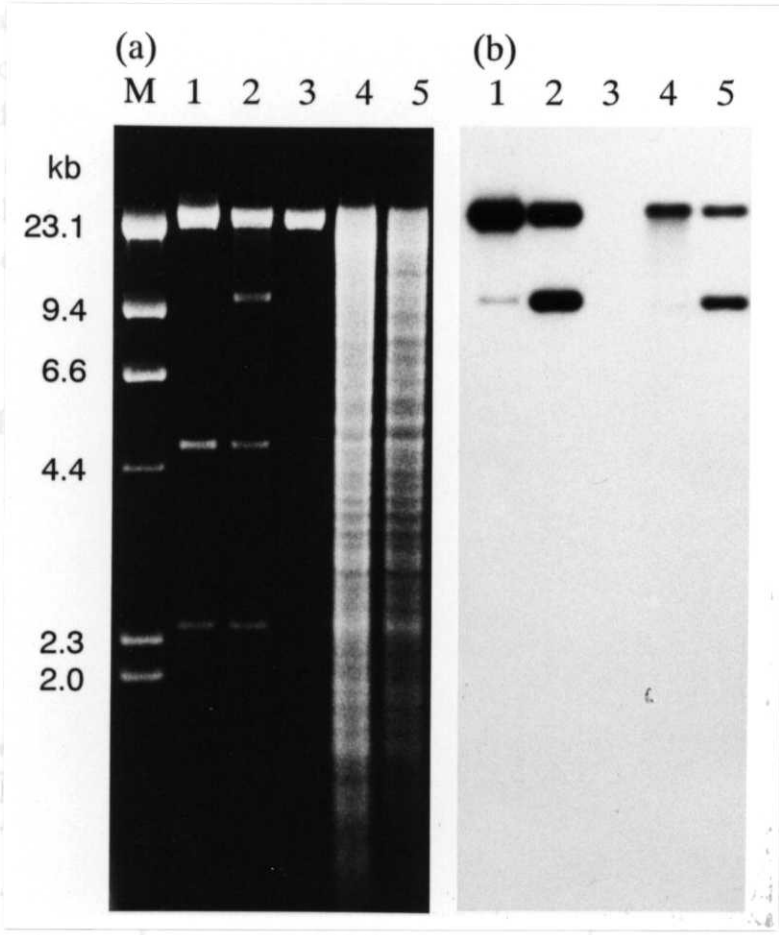


Fig. 8. Southern hybridization of restriction digests of plasmid DNA and total DNA from NH9 and derivative strains with the 5.8-kb *Bam*HI-*Bgl*II fragment being used as the labeled probe. (a) Agarose gel electrophoresis (0.7% gel and TBE buffer) of *Eco*RI digests of the samples of DNA. Lanes; 1, pENH91; 2, pENH91A; 3, pENH91d5; 4, total DNA from NH9; 5, total DNA from NH9A; M, fragments of lambda DNA digested with *Hin*dIII as size standards. (b) Autoradiogram of the corresponding Southern blot after hybridization with the 5.8-kb *Bam*HI-*Bgl*II fragment as the labeled probe. The lanes correspond to those in panel a.

Hybridization patterns with the *Bgl*II digests of the two plasmids were the same; this probe hybridized only with the 10.4-kb fragment that included the 5.8-kb region.

These results indicated that the novel fragment contained both the region that encoded the catabolic enzymes and the direct repeat and that there were doublet band(s) in the electrophoretic profile of the *Bgl*II digest of pENH91A. If the novel fragment had been located at a distant site on the plasmid, for example, by transposition, it would have altered the sizes of the original fragments of pENH91. Because all the restriction fragments of pENH91 were retained in pENH91A, it was clear that the novel fragment was located close to the original region.

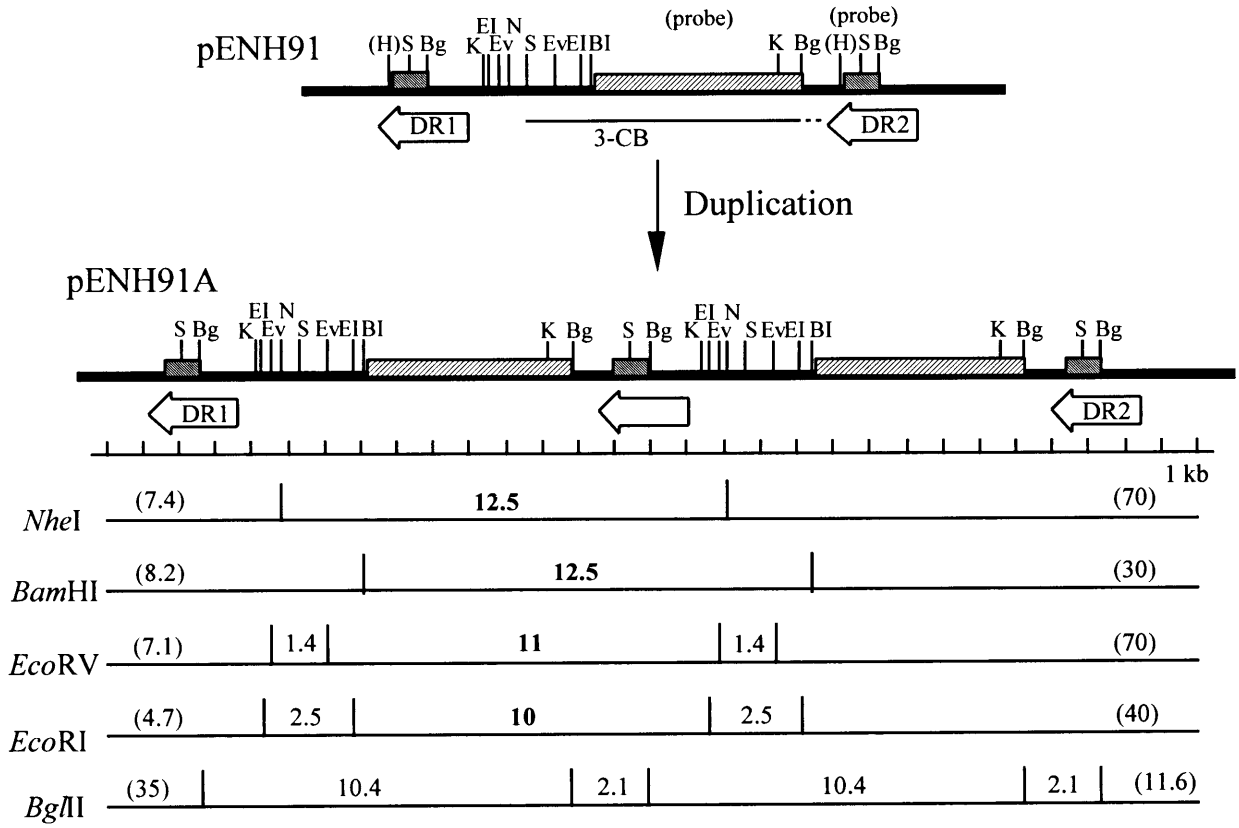


Fig. 9. Restriction map of the duplicated region of pENH91A. Hatched areas indicate the fragments that were used as probes in hybridization experiments and the corresponding hybridized regions. The numbers on the map for each restriction enzyme are the lengths of fragments in kilobases. The fragments with numbers in bold type are the fragments that were not seen in the restriction profiles of pENH91. The fragments with numbers in parentheses are those whose entire lengths are not included in this figure. The abbreviations for the restriction sites are the same as those in the legend to Fig. 3 and 4.

The recombination event must have occurred with one of the following two patterns: i) -DR-CB-DR-CB-DR-, by homologous recombination between the direct repeats; or ii) -DR-CB-DR-DR-CB-DR-, by replicative transposition (DR and CB indicate IS*1600* and the intervening region of 3-CB-catabolic genes, respectively).

The length (12.5 kb) of the novel fragments generated by digestion with either *Nhe*I or *Bam*HI was equal to the sum of the lengths of the intervening region (10 kb) and a direct repeat (2.5 kb). This result is in accord with pattern (i). In addition, the identical profiles obtained with the *Bgl*II digests of the two plasmids allow us to exclude pattern (ii). If pattern (ii) were to describe what had occurred, the *Bgl*II digest of pENH91A should have given an additional fragment of (at least) 2.5 kb, derived from the junctional region (-DR-DR-).

The lengths of the novel fragments generated by the four enzymes other than *Bgl*II corresponded to the lengths of fragments that would be produced by tandem duplication and not by inverted duplication.

These considerations confirmed that tandem duplication of the 12.5-kb region that contained the catabolic genes took place via reciprocal recombination between the direct repeats, as illustrated in Fig. 9.

1.7 Determination of sequences of genes for degradative enzymes

Sequencing analysis of the 9.2-kb *Sac*I fragment revealed seven long open reading frames (orfs; Fig. 4). Six of the orfs formed a cluster and exhibited strong homology to orfs in the following clusters of chlorocatechol-degradative genes (in the order of relatedness): i) the *tcbR-CDXEF* genes on plasmid pP51 in *Pseudomonas* sp. P51 (van der Meer *et al.*, 1991a,b); ii) the *clcR-ABXDE* genes on plasmids pAC27 in *P. putida* AC866 and pWR1 in *Pseudomonas* sp. B13 (Coco *et al.*, 1993; Frantz and Chakrabarty, *et al.*, 1987; Kasberg *et al.*, 1997); and iii) the *tfdR* and *tfdCDEF* genes on plasmid pJP4 in *R. eutropha* JMP134 (Matrubutham and Harker, 1994; Perkins *et al.*, 1990) (X denotes the third orf in the clusters of *tcbCDEF* and *clcABDE* degradative genes; the functions of these genes are unknown). In particular, the extent of homology between six orfs of NH9 and orfs of the *tcbR-CDXEF* genes of *Pseudomonas* sp. strain P51 was very great (Table 2).

From the high homology to known chlorocatechol-degradative genes and by analogy to the pathways formed by the products of these gene clusters (van der Meer *et al.*, 1991a), it was apparent that the sequenced degradative genes of strain NH9 encoded enzymes of the modified *ortho*-cleavage pathway (Fig. 10). Since strain NH9D (a derivative of strain NH9 that had been cured spontaneously of plasmid pENH91) grew

Table 2. Homology between *cbn* genes and *tcb* genes

	cbnR	*cbnA*	*cbnB*	orf3*cbn*	*cbnC*	*cbnD*
Corresponding genes	(66.1%)	(65.0%)	(63.3%)	(64.3%)	(61.0%)	(62.9%)
and G+C content	*tcbR*	*tcbC*	*tcbD*	orf3*tcb*	*tcbE*	*tcbF*
	(66.2%)	(65.2%)	(63.7%)	(63.3%)	(61.0%)	(62.9%)
Length Nucleotides	885 bp	756 bp	1113 bp	1011 bp	717 bp	1059 bp
Amino acids	294 aa	251 aa	370 aa	336 aa	238 aa	352 aa
Nucleotide level (%)	99.9	98.0	88.9	97.9	99.9	100
Amino acid level (%)	99.7	95.6	97.0	97.0	99.6	100

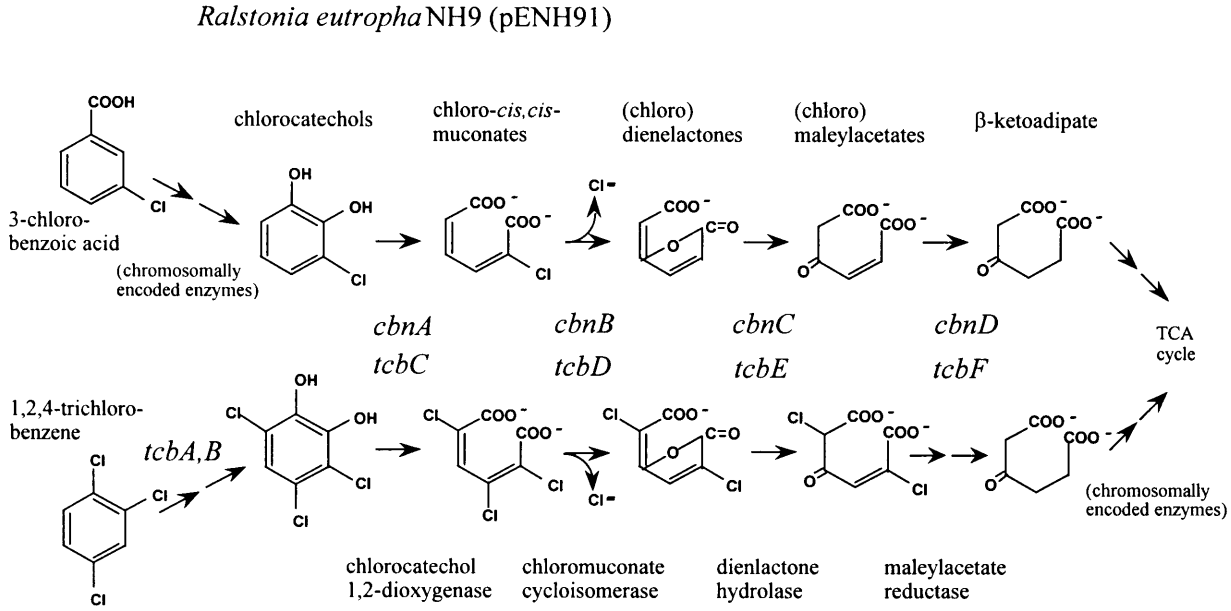


Fig. 10. Pathways for degradation of 3-chlorobenzoic acid by *Ralstonia eutropha* NH9 and for degradation of 1,2,4-trichlorobenzene by *Pseudomonas* sp. P51. The pathway for degradation of 1,2,4-trichlorobenzene by *Pseudomonas* sp. P51 was first described by van der Meer *et al.* (1991c).

on benzoate, strain NH9 was assumed to harbor genes for benzoate 1,2-dioxygenase and 1,2-dihydro-1,2-dihydroxybenzoate dehydrogenase, which might convert (chloro)benzoate into (chloro)catechol, either on its chromosome (Johnson and Stanier, 1971) or on the additional plasmid pENH92 (Fig. 2).

Since the enzymes encoded by the chlorocatechol-degradative genes of NH9 were responsible for degradation of 3-CB, the genes were designated *cbnR-ABCD*, with *cbnA*, *B* and *C* encoding chlorocatechol 1,2-dioxygenase, chloromuconate cycloisomerase, and dienelactone hydrolase, respectively (van der Meer *et al.*, 1991a, c). Recent studies suggested that *cbnD*, corresponding to *tcbF*, might encode maleylacetate reductase (Kasberg *et al.*, 1995; Kasberg *et al.*, 1997; Seibert *et al.*, 1993). *cbnR* was presumed to be a regulatory gene that belonged to the *lysR* family (Schell, 1993; van der Meer *et al.*, 1991b). orfX, between *cbnB* and *cbnC*, corresponded to the third orf in the *tcbCDEF* and *clcABDE* gene clusters, whose products have unknown functions. The strongly conserved amino acid sequences encoded by orfX in *cbnABXCD* and in *tcbCDXEF* suggest that the products of these orfs might play a role that is indispensable for the function of the gene cluster in some as yet unknown fashion.

The extent of the homology of each gene in the *cbnR-ABXCD* cluster to the corresponding gene in the *clcR-ABXDE* cluster (Coco *et al.*, 1993; Frantz and Chakrabarty, *et al.*, 1987; Kasberg *et al.*, 1997) ranged from 59% to 72% at the nucleotide level and from 51% to 76% at the amino acid level. Homology to the *tfdR* and *tfdCDEF* genes of strain JMP134 (Matrubutham and Harker, 1994; Perkins *et al.*, 1990) was 58% to 66% at the nucleotide level and 52% to 67% at the amino acid level. The degradative genes *cbnA*, *cbnB* and *cbnR* are considered to be evolutionarily related to the functionally similar genes in the catechol *ortho*-cleavage pathway, namely, *catA*, *catB*, and *catR* (Schlömann, 1994). The homology between the *cbnR-AB* genes and the corresponding *cat* genes of *P. putida* PRS2000 (Houghton, 1995) and *Pseudomonas* strain RB1 (Aldrich *et al.*, 1987; Rothmel *et al.*, 1990) was 51% to 57% at the nucleotide level and 31% to 45% at the amino acid level.

1.8 Comparison between the *cbn* and *tcb* gene clusters

The regions containing the chlorocatechol-degradative genes of the two plasmids, namely, the 6959-bp *Sac*I-*Kpn*I regions of pENH91 and pP51, were compared with one another (Fig. 4, 11 and Table 2). All of the corresponding orfs were of the same respective lengths (Table 2). Hence, all the differences between the coding regions of the *cbnR-ABCD* and *tcbR-CDEF* genes were substitutions. The *cbnA* and *cbnB* genes overlapped by 4 bp. One nucleotide was present between *cbnB* and the third orf in the *cbn* degradative operon. The intergenic region between the third orf of *cbn* and *cbnC* consisted of 21 bp. *cbnC* overlapped with *cbnD* by 4 bp. All of these structural features of the *cbn* genes were the same as those of the corresponding regions of the *tcb* genes. The nucleotide sequence of the promoter region between *cbnR* and *cbnA* (150

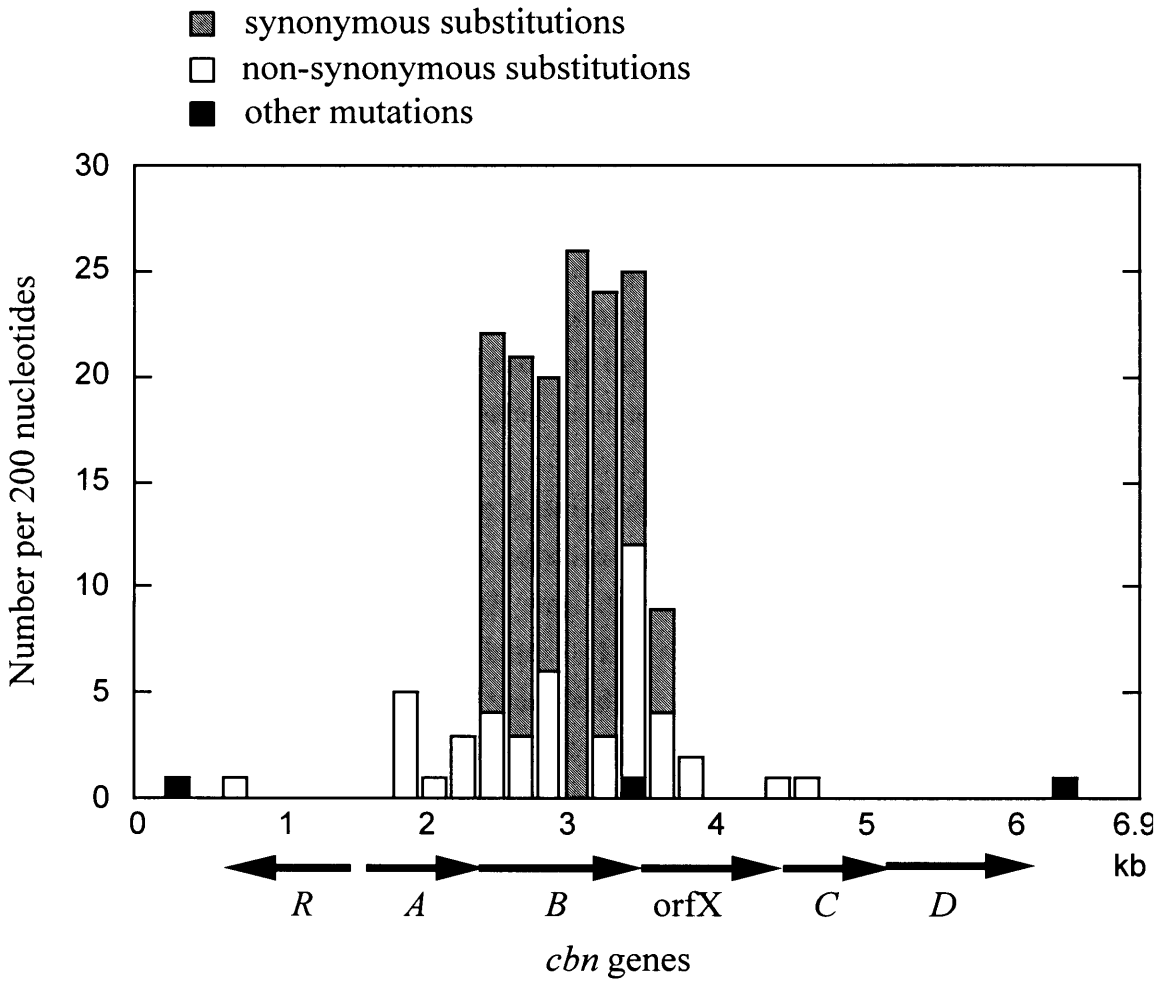


Fig. 11. Numbers of nucleotide substitutions and other mutations in the region of the *cbn* gene cluster in comparison to the region of the *tcb* gene cluster. The numbers per 200 nucleotides are shown, counted from the *Sac*I site in the 6959-bp region that contains the *cbn* gene cluster. The corresponding regions of nucleotides in the two clusters are as follows: *cbn* 1 to 287 and *tcb* 1 to 287, *cbn* 289 to 6578 and *tcb* 288 to 6577, and *cbn* 6579 to 6959 and *tcb* 6579 to 6959.

bp) was identical to that between *tcbR* and *tcbC*.

Nucleotide substitutions between *cbnA* and *tcbC* resulted in a slight decrease in the percentage homology between the deduced amino acid sequences (Table 2). Eleven of 13 nucleotide substitutions caused non-synonymous substitutions at the amino acid level. However, the four amino acids that are supposed to coordinate the ferric ion at the active site were conserved in CbnA (Tyr-130, Tyr-164, His-188, and His-198), as were another 28 amino acids that are conserved among catechol 1,2-dioxygenases (Nakai *et al.*, 1995). Although homology at the nucleotide level was the lowest among the cycloisomerase genes *cbnB* and *tcbD* (Table 2), the majority of nucleotide

substitutions between *cbnB* and *tcbD* were synonymous substitutions (in 102 out of 113 codons) and reflected the high frequency of nucleotide substitutions at the third base in the codon (107 out of 124 nucleotides; Fig. 11). Consequently, the homology at the amino acid level between CbnB and TcbD remained high, in contrast to the case of CbnA and TcbC (Table 2). There was one nucleotide substitution both between *cbnR* and *tcbR* and between *cbnC* and *tcbE* , resulting in one amino acid substitution in each pair. The nucleotide sequence of *cbnD* was identical to that of *tcbF*. All of the putative ribosome-binding sites for each of the *cbn* genes were located at the same respective positions and all had the same sequences as those for the *tcb* genes (van der Meer *et al.*, 1991a, b).

The nucleotide sequences of the flanking regions of the two gene clusters within the *Sac*I-*Kpn*I fragments were nearly identical: the 637-bp nucleotide sequence of the downstream flanking region of *cbnR* was identical to that of *tcbR* except that one nucleotide was missing from the latter. The 618-bp nucleotide sequence of the downstream flanking region of *cbnD* was identical to that of *tcbF* except for insertion of one nucleotide in the latter case. These results suggested that the highly homologous regions of the two plasmids extended beyond both the *Sac*I site and the *Kpn*I site (Fig. 4).

1.9 An additional orf in the region between *cbnD* and DR2

In a 2-kb region between *cbnD* and DR2, we found an orf of considerable length that exhibited some similarity to known proteins at the amino acid level (Fig. 4; designated orfL). The deduced amino acid sequence of orfL (240 a.a.) was homologous to many bacterial polypeptides that are known to be components of membrane-bound transport systems for amino acids. In particular, LivF of *E. coli* (Adams *et al.*, 1990) and BraG of *Pseudomonas aeruginosa* PAO (Hoshino and Kose, 1990) were 42% and 39% homologous to orfL at the amino acid level, respectively. orfL seemed to have complete length in comparison with LivF (237 a.a.) and BraG (233 a.a.). Both LivF and BraG are located at the downstream end (in the direction of transcription) in their respective gene clusters. Because of the location and apparent direction of transcription of orfL in the region bracketed by the two ISs, it appeared as if orfL might have been separated from the rest of the genes in the cluster by an excision event involving IS*1600* (DR2).

1.10 Cloning and sequencing of a DNA fragment from strain P51 with homology to IS*1600*

It was examined whether the *tcbR-CDEF* gene cluster was also associated with an IS*1600* (or IS*1600*-like) sequence on plasmid pP51. Total DNA was extracted from cells of strain P51 that had been grown on Luria-Bertani (LB) medium (Sambrook *et al.*, 1989) and subjected to Southern hybridization experiments. The presence of plasmid pP51 DNA in the total DNA was confirmed with a 3.5-kb *Bam*HI-*Pst*I fragment containing a part of the *cbn* gene cluster as the probe. Then the same membrane was used in a hybridization experiment with a 2.3-kb *Hin*dIII-*Sph*I fragment containing a part of IS*1600* (DR2) as the probe. A hybridizing band was observed and this fragment was cloned from the total DNA of strain P51 into pUC19 as a 3.7-kb *Sal*I fragment. In a subsequent hybridization experiment, using the cloned 3.7-kb *Sal*I fragment as the probe, the *Sal*I-digested total DNA from strain P51 (which retained the *tcb* genes) and from derivative strains of P51 that had been cured spontaneously of the plasmid P51 after successive cultures on LB liquid medium gave the same pattern of signals. Therefore, it appeared that the cloned 3.7-kb *Sal*I fragment resided on the chromosome of strain P51.

Sequence of a ca. 2-kb region of the cloned 3.7-kb *Sal*I fragment revealed that the region homologous to IS*1600* extended for 1,300 bp, with nucleotide homology of 81 %. The 1,300-bp region was flanked by nonhomologous marginal regions of ca. 0.2 kb and 0.5 kb on each side. Thus, the homologous region had not been truncated by the cloning procedure using the *Sal*I sites. The homologous 1,300-bp region started with a 15-bp sequence that resembled the left inverted repeat of IS*1600* (Fig. 7b). It was followed by a 115-bp intervening region and then by an open reading frame (designated orfSA; 1170 bp) which showed the highest homology with a part of IstA of IS*1600* (83% at amino acid level; Fig. 7a and 26) in the database. In addition to the fact that the 1,300-bp region seemed to lack the 3' portion of *istA*, it was obviously not followed by *istB* and an inverted repeat at the other end. Thus, this fragment seemed to be a remnant of an IS.

Summary of results section 1: A ca. 5.7-kb six gene cluster responsible for chlorocatechol degradation was cloned from the transmissible plasmid pENH91 of 3-CB degradative bacterium *R. eutropha* NH9: the *cbnABCD* operon encoding the degradative enzymes (including *orfX* of unknown function) and the divergently transcribed *cbnR* encoding the LysR-type transcriptional regulator of the *cbn* operon. The *cbnR-ABCD* genes showed the highest homology to the *tcbR-CDEF* genes on

plasmid pP51 of the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. P51, which was isolated in the Netherlands (89% to 100% identity at the nucleotide level). The *cbnR-ABCD* genes were found to be located between two directly-oriented identical insertion sequences (ISs) of 2,520 bp, designated IS*1600*, thus forming a composite transposon designated Tn*5707* (ca. 15 kb). Both deletion and duplication events of the *cbnR-ABCD* gene cluster on the plasmid pENH91 were mediated by homologous recombination between the two elements of IS*1600*.

2. Transcriptional activation of the 3-chlorocatechol degradative genes *cbnR-ABCD*

The structure of the *cbnR-ABCD* genes suggested that CbnR was a LysR-type transcriptional activator for the expression of the *cbnA* promoter by analogy with the *tcbR-CDEF* genes. The function of CbnR with relation to the expression of the *cbnA* promoter was examined in vivo and in vitro.

2.1 Complementation by CbnR of growth on 3-CB of bacteria harboring *cbnABCD* genes

The *cbnABCD* genes encode enzymes of the modified *ortho*-cleavage pathway which degrade 3-chlorocatechol converted from 3-CB (Fig. 10). To examine if *cbnR* was required for growth of NH9 on 3-CB, a complementation test was conducted. Either pCbn13RABCD or pCbn13ABCD which retained *cbnA* promoter region was mobilized into *R. eutropha* NH9D, a derivative of NH9 which is cured of the *cbnR-ABCD* genes, or *P. putida* PRS4020, a *catR* knockout strain. The growth of the resultant four strains on 3-CB were compared together with strains containing the vector control, pCP13. Strains NH9D and PRS4020 with pCbn13RABCD showed growth on the 3-CB plate, while the strains with pCbn13ABCD or pCP13 did not. This result indicated that *cbnR* was necessary for the growth of these bacteria on 3-CB.

2.2 Activation of the *cbnA* promoter by CbnR during growth in the presence of benzoate or 3-chlorobenzoate

It was presumed that CbnR may activate the *cbnA* promoter based upon the regulatory systems of other *ortho* cleavage operons (Leveau *et al.*, 1994; McFall *et al.*, 1998; van der Meer *et al.*, 1991b). To analyze the function of *cbnR* for the expression of the degradative genes in vivo, *P. putida* PRS4020 (*catR* knockout strain) was used as a host for reporter plasmids (Fig. 12). The cells were grown in basal synthetic medium (BSM) (Aldrich *et al.*, 1987) supplemented with 10 mM glucose, 10 mM glucose and 5 mM Ben, or 10 mM glucose and 5 mM 3-CB for 18 hours at 30°C. When the cells containing pNO50RAB' were grown on glucose with either Ben or 3-CB, the transcription from *cbnA* promoter was activated 17 fold and 6.8 fold, respectively, compared to the cells grown on glucose alone (Table 3). The activation in the presence of Ben or 3-CB was not seen for the cells with pNO50AB' which did not contain *cbnR*. These results indicated that *cbnR* was a positive regulator of the transcription from *cbnA* promoter.

Cells containing pNO50RA' did not induce substantially in the presence of 3-CB. When the same cells were grown in the presence of Ben, the transcription was activated 12-fold. The intact *cbnA* gene was necessary for the activation upon addition of 3-CB and also for the higher activation upon addition of Ben observed with pNO50RAB'. 3-CB and Ben are converted to 3-chlorocatechol and catechol, respectively, by enzymes encoded on the host chromosome, and CbnA (chlorocatechol dioxygenase) can further convert these (chloro)catechols to 2-chloro-*cis,cis*-muconate (2-CM) and *cis,cis*-muconate (CCM), respectively. The above results suggested that 2-CM and CCM were the inducers of the *cbnA* promoter.

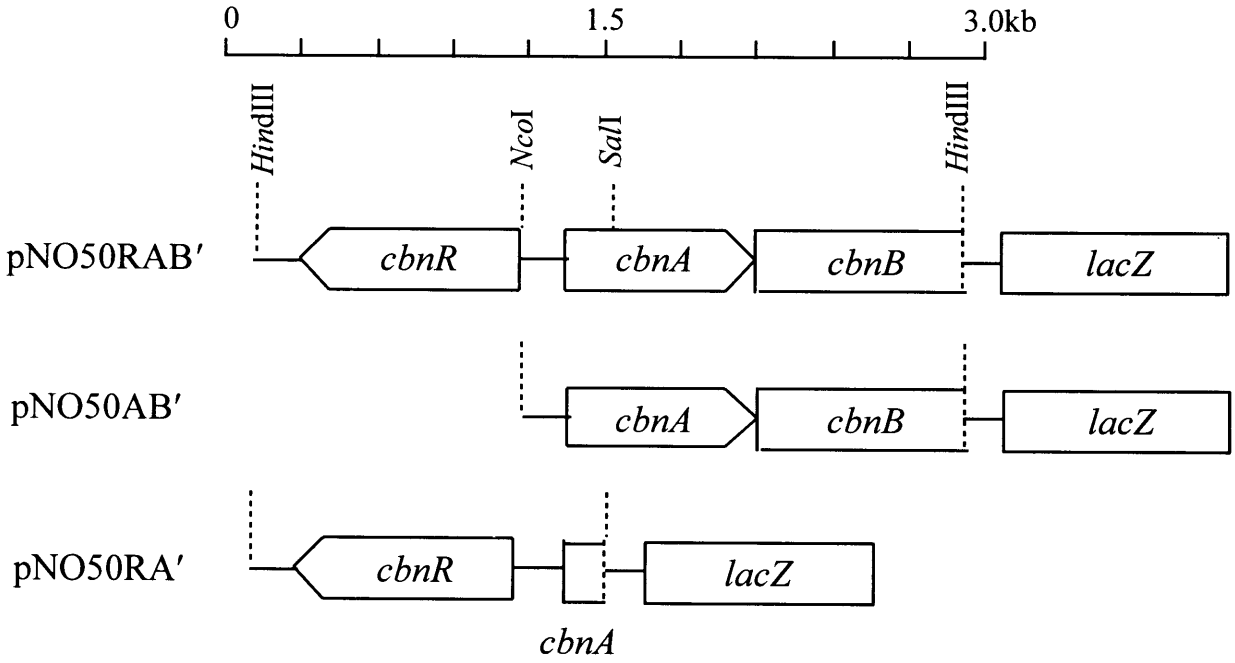


Fig. 12. Diagram of the inserts of the constructs used for the *cbnA* promoter activity assay.

Table 3. The effect of Ben and 3-CB on transcriptional activation at the *cbnA* promoter

P.putida PRS4020 /plasmid construct	Glucose 10mM	Glucose 10mM + Benzoate 5mM	Glucose 10mM + 3-CB 5mM
vector control (pQF50)	3.5 (±0.2)	1.8 (±0.2)	2.0 (±0.2)
pNO50RAB'	61.8 (±2.3)	1038.1 (±46.8)	424.9 (±22.4)
pNO50AB'	116.1 (±3.2)	56.5 (±3.9)	71.1 (±5.5)
pNO50RA'	21.6 (±0.7)	261.9 (±27.0)	56.4 (±8.7)

β-galactosidase units: nmol/min/mg extract; (±): standard deviation

2.3 Purification of CbnRHis protein

Specific binding of CbnR to the *cbnA* promoter was demonstrated by gel retardation assay using the crude protein from *E. coli* BL21(DE3)pLysS containing pT7cbnR. In order to simplify protein purification, a plasmid expressing *cbnR* with 6 His codons on its C-terminal, pT7cbnRHis, was constructed. The crude soluble protein from *E. coli* BL21(DE3)pLysS containing pT7cbnRHis showed binding activity to *cbnA* promoter, although a considerable portion of the protein was secluded in inclusion bodies. When the crude protein was directly purified with the nickel affinity column, two major bands appeared in the SDS-PAGE profile. One of the bands apparently corresponded to the calculated molecular weight of CbnRHis (32.9 kD), and the other was a polypeptide of about 27 kD. This 27-kD protein was also produced from the cells of the vector control. To remove this contaminating protein, a heparin-agarose column was employed followed by the nickel affinity column, and the 32.9-kD protein was recovered at more than 90% purity (Fig. 13, lane 3). The amino acid sequence of the N-terminal of this protein was found to be identical to that of the deduced sequence of CbnRHis. The function of CbnRHis was evaluated with respect to partially-purified wild-type CbnR in vivo and in vitro in the following sections.

2.4 Confirmation of the function of CbnRHis

Partially purified CbnR and purified CbnRHis were subjected to gel retardation assays with a 252-bp fragment containing the *cbnA* promoter region. The two proteins showed the same retardation mobility of the *cbnA* promoter fragment (Fig. 14, lanes 3-5 and lanes 6-8). To test if CbnRHis had the same function as CbnR in vivo, growth complementation tests and reporter analysis were conducted. In growth complementation, strains NH9D and PRS4020 harboring pCbn13RHABCD grew on 3-CB as did strains with pCbn13RABCD. In reporter analysis, PRS4020 containing pNO50RHAB' exhibited activation of transcription in the presence of either Ben or 3-CB to the same levels as PRS4020 containing pNO50RAB' did (data not shown). These results indicated that CbnRHis had the same function as CbnR in these in vitro and in vivo experiments.

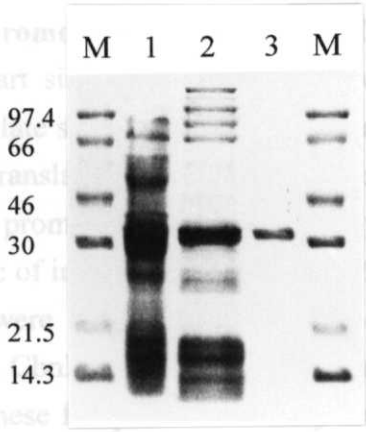


Fig. 13. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis profile of purification of CbnRHis. Lanes: 1, total protein from induced cells of BL21(DE3)pLysS containing pT7cbnRHis (100 μg); 2, protein purified by heparin-agarose column (50 μg); 3, CbnRHis purified by heparin-agarose column and nickel column (5 μg); M, protein molecular mass markers. Sizes are shown in kilodaltons.

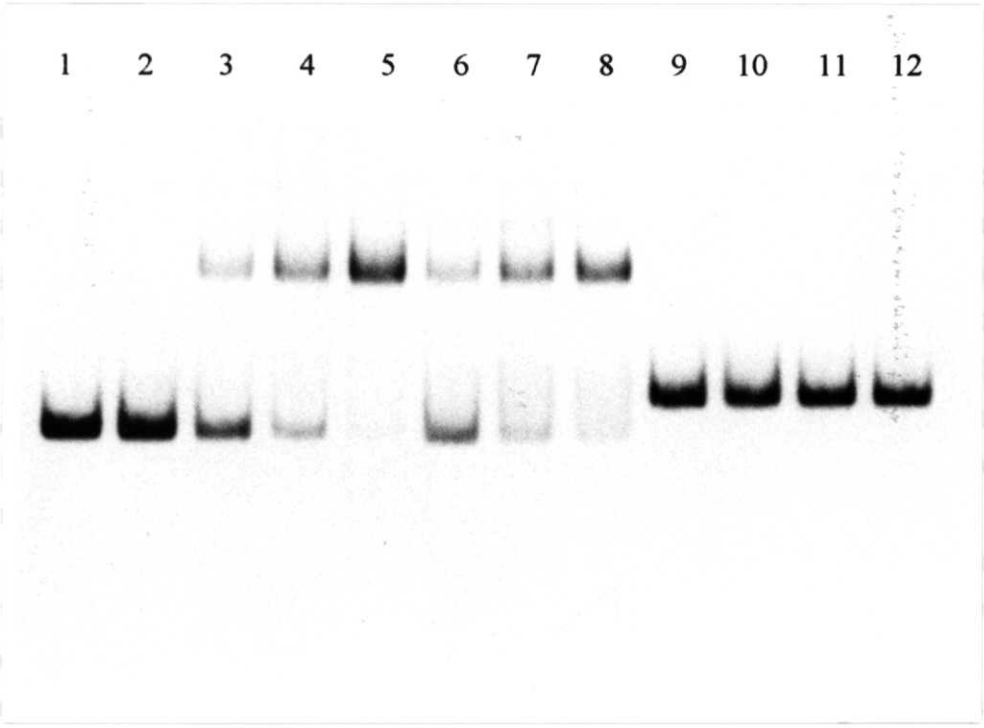


Fig. 14. Gel retardation assay demonstrating specific binding of CbnR and CbnRHis to the *cbnA* promoter region. A 252-bp fragment containing the *cbnA* promoter region was used as probe in lanes 1 to 8. A 336-bp fragment containing part of the hydroxyquinol 1,2-dioxygenase gene of *Burkholderia cepacia* AC1100 was used in lanes 9 to 12. Lanes: 1 and 9, no protein; 2 and 10, partially purified protein from BL21(DE3)pLysS/pT7-7 at 0.5 μg; 3, 4, 5, and 11, partially purified CbnR at 0.05, 0.1, 0.5, and 0.5 μg, respectively; 6, 7, 8, and 12, CbnRHis at 0.05, 0.1, 0.5, and 0.5 μg, respectively.

2.5 Features of the *cbnA* promoter region bound by CbnR

The transcriptional start site of *cbnA* promoter was determined by S1 nuclease analysis; the A on the template strand is marked by an arrow in Fig. 15 and was located 47-bp upstream from the translational start codon of *cbnA* gene. The precise site of CbnR binding on the *cbnA* promoter was determined by DNase I protection assay (Fig. 16 and 17). In the absence of inducer, approximately 60-bp from -76 to -19 (relative to the +1 of transcription) were protected from DNase I cleavage by CbnR. Both partially purified CbnR and CbnRHis exhibited the same protected patterns on the sense and anti-sense strands. These footprints were very similar to those of TcbR bound to the *tcbC* promoter (Leveau *et al.*, 1994). The addition of CCM or 2CM up to 1 mM in the binding reaction did not change the footprinting pattern of CbnR to *cbnA* promoter. Besides the inverted repeat containing T-N_{11}-A located in the RBS (Schell, 1993), there was a second inverted repeat (Fig. 17) whose location is similar to those described recently by the study of BlaA, a LysR type regulator for β-lactamase genes of *Streptomyces cacaoi* (Magdalena *et al.*, 1997).

2.6 Identification of the inducers by in vitro transcription assay

The results of the LacZ assays suggested that CCM and 2-CM act as the inducers of the *cbnA* promoter. Considerable activity, however, was also shown when the cells containing a truncated CbnA gene (pNO50RA') were grown in glucose with Ben (Table 3). The possibility that either Ben or catechol may also serve as inducers could not be excluded. To examine the effect of these compounds on the transcriptional activation by CbnR, in vitro transcription assays were performed with purified CbnRHis and potential inducer compounds. Figure 18 shows that *cbnA* transcripts were produced only when either CCM or 2-CM was added to the reaction (lanes 12 and 14). No transcript was produced with other compounds including Ben (lane 4) or catechol (lane 8). Therefore Ben or catechol does not serve as an inducer for the *cbnA* promoter. The transcriptional start sites of the products with CCM or 2-CM were determined by S1 nuclease assay and were confirmed to be the same as that derived by in vivo assay (data not shown).

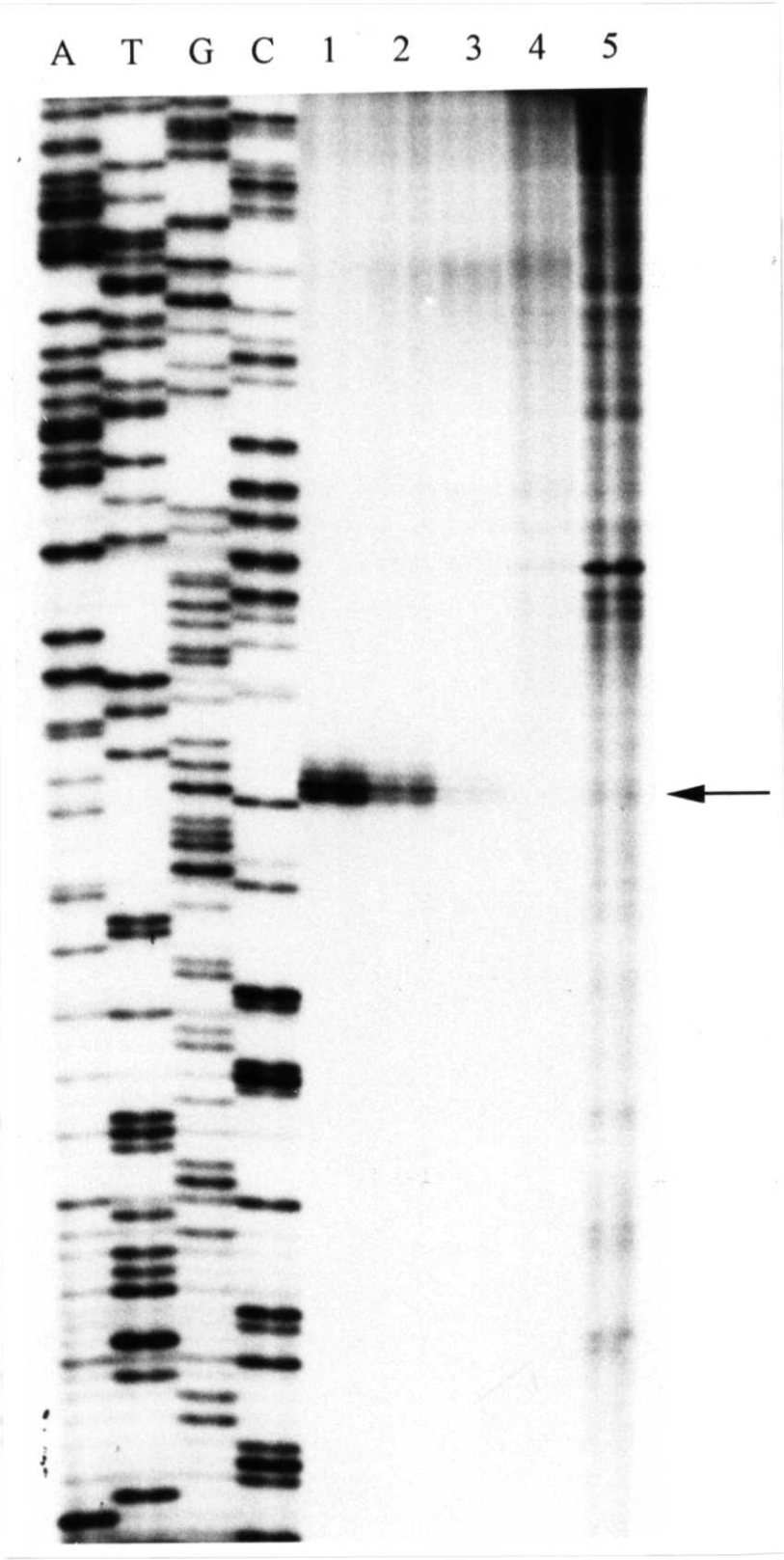


Fig. 15. S1 nuclease assay to determine the transcriptional start site of *cbnA*. Sequencing reactions of the bottom strand of the *cbnA* promoter region are shown as A, T, G, and C. The products of S1 nuclease assay are shown in lanes 1 and 2 (marked by arrow). In lanes 1, 2, and 3, 20, 10, 3.3 μg of in vivo-derived RNA were used, respectively. In lanes 4 and 5, 10 μg of transfer RNA were used for each lane. In lane 5, S1 nuclease was not added.

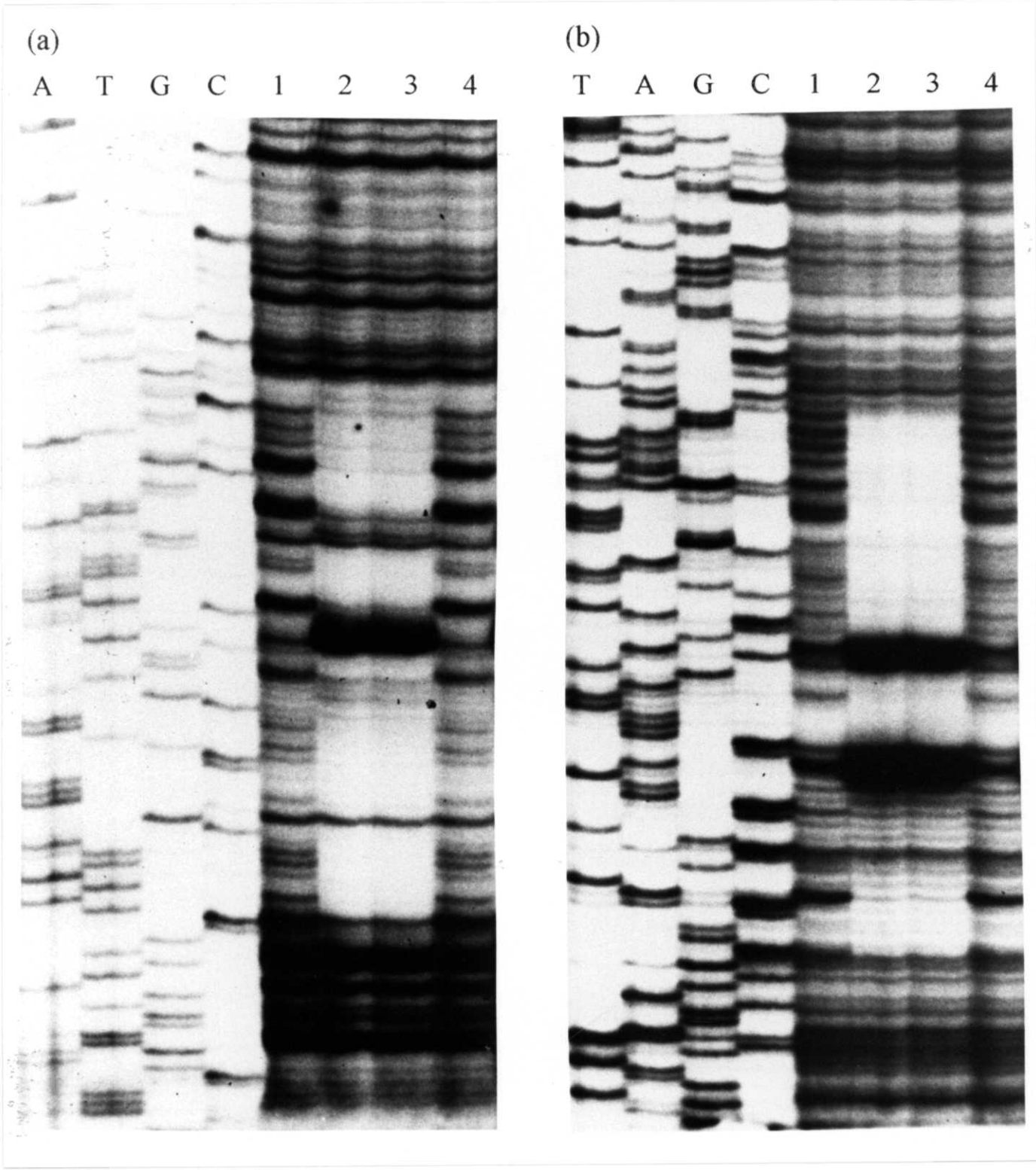


Fig. 16. DNase I footprint of CbnRHis and CbnR on the *cbnA* promoter region. (a and b) Top strand (a) and bottom strand (b) of the *cbnA* promoter region. Lanes: 1, no protein; 2, CbnRHis (0.5 μg); 3, partially purified CbnR (1 μg); 4, partially purified protein from BL21(DE3)pLysS/pT7-7 (1 μg).

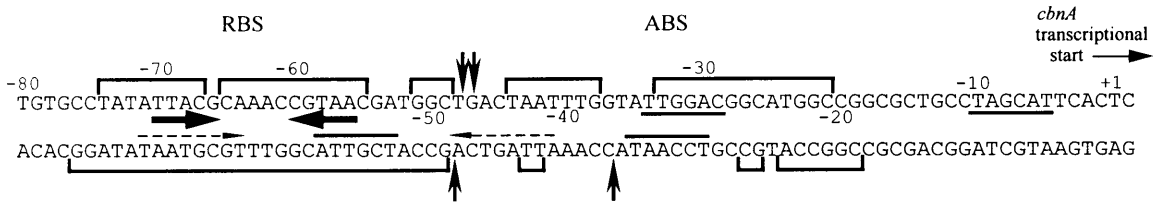


Fig. 17. Schematic diagram of the *cbnA* promoter region protected from DNase I digestion by CbnR(His). The protected nucleotides are shown by the brackets. The vertical arrows indicate sites of hypersensitivity to DNase I digestion. The thick horizontal arrows show the inverted repeats containing T-N_{11}-A, regarded as a motif of LysR regulatory systems (Schell, 1993). The horizontal dotted arrows above the bottom strand indicate the second imperfect inverted repeat similar to those described recently (Magdalena *et al.*, 1997). The horizontal solid lines under the top strand and above the bottom strand indicate the -35 and -10 regions of the *cbnA* promoter and the divergently transcribed *cbnR* promoter. The regions from nucleotides -76 to -49 and from -44 to -19 are suggested to be the recognition binding site (RBS) and the activation binding site (ABS), respectively (McFall *et al.*, 1998). The numbering is relative to the transcriptional start site of *cbnA*.

2.7 Bending of the *cbnA* promoter region

The presence of hypersensitive sites in the center of the footprint region of the *cbnA* promoter suggested that changes in DNA conformation might occur upon binding of CbnRHis. To examine this potential change with or without inducer, circular permutation gel shift assays were performed. Lanes 1 to 5 of Fig. 19 indicate that binding of CbnRHis caused bending of the promoter fragment; the bending angle was estimated to be 78°. When CbnRHis was bound to the promoter in the presence of CCM, relaxation of the bending angle to 54° was observed (lanes 6 to 10). The bending angles of three independent experiments varied by no more than 7.7%. Partially purified CbnR gave the same bending angles as CbnRHis (data not shown.)

Summary of results section 2: CbnR was demonstrated to regulate the expression of the *cbnA* promoter positively by in vivo and in vitro experiments. The inducers for the expression of the *cbnA* promoter by CbnR were found to be CCM and 2-CM. Specific binding of CbnR protein to the *cbnA* promoter region was demonstrated by gel shift and DNaseI footprinting analysis. In the absence of inducer, a region of ca. 60 bp from -20 to -80 upstream of the *cbnA* transcriptional start point was protected from DNaseI cleavage by CbnR, with a region of hypersensitivity to DNase I cleavage clustered at -50. Circular permutation gel shift assays demonstrated that CbnR bent the *cbnA* promoter region to an angle of 78° and that this angle was relaxed to 54° upon the addition of inducer, CCM.

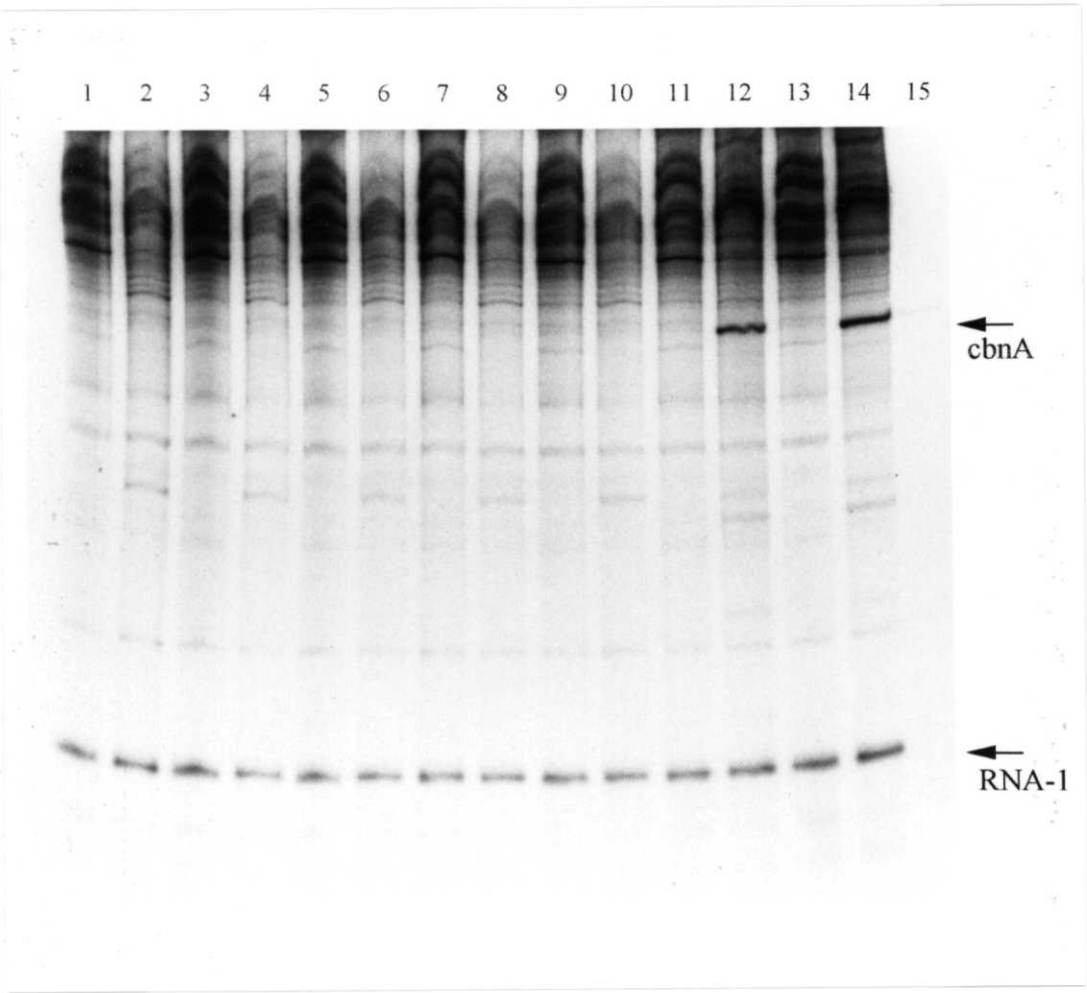


Fig. 18. In vitro transcription assay demonstrating the requirement of both CbnR and inducer. In lanes 1 to 14, 0.1 μg CbnRHis was used in the lanes with even numbers while it was not added to the lanes with odd numbers. Lanes: 1 and 2, no chemicals; 3 and 4, Ben; 5 and 6, 3-CB; 7 and 8, catechol; 9 and 10, 3-chlorocatechol; 11 and 12, CCM; 13 and 14, 2-CM; 15, marker (a transcript of 428 bases derived from plasmid template pJET41) (Erickson and Gross, 1989; McFall *et al.*, 1997c) is shown as a faint band near the arrow marked cbnA. All of the chemicals were used at 1 mM concentrations. RNA-1 is the transcript from the ColE1 ori of the supercoiled plasmid pMP7.

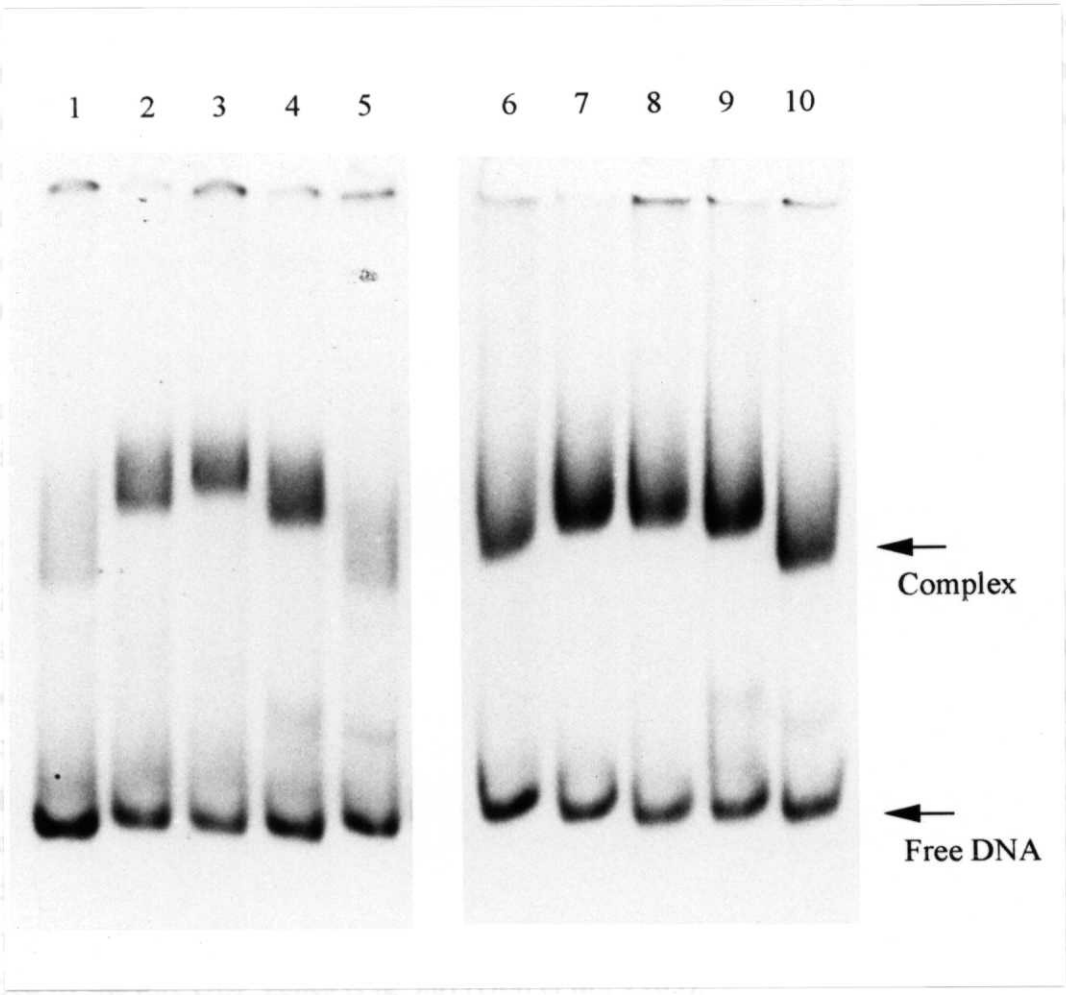


Fig. 19. Circular permutation gel shift assay demonstrating CbnRHis bending of the *cbnA* promoter region in the absence (lanes 1 to 5) and presence (lanes 6 to 10) of 1 mM CCM. The ca. 60-bp region that showed a footprint by CbnR was distributed in different positions from the ends (lanes 1 and 6, left ends; lanes 5 and 10, right ends) to the center (lanes 3 and 8) of 257-bp fragments. The calculated bending angles were 78° and 54° in the absence and presence of CCM, respectively.

3. The structure and transcriptional acivation of the 2,4-dichlorophenoxyacetate degradative genes of *Alcaligenes* sp. CSV90

The *tfdC* gene encoding the dichlorocatechol dioxygenase cloned from the large plasmid pMAB1 of the 2,4-dichlorophenoxyacetate (2,4-D) degradative bacterium *Alcaligenes* sp. CSV90 has been found to be identical with that of the well-characterized plasmid pJP4 of the 2,4-D degradative bacterium *R. eutropha* JMP134. The sequence of the upstream region of the *tfdC* gene on pMAB1, however, has turned out to be different from that on pJP4 (Bhat, 1994; Leveau, 1996). Therefore, the structure of the 2,4-D degradative genes on pMAB1 was examined in comparison with that of pJP4. Also, the expression of the degradative genes was analyzed in vivo using the same system as that used to test the expression of the *cbnA* promoter to compare the function of the regulators.

3.1 The structure of the 2,4-D degradative genes from *Alcaligenes* sp. CSV90

2,4-D is one of the major herbicides and has been used as a model compound for the study of the evolution of the degradative genes for chlorinated aromatic compounds. In the well studied catabolic pathway of 2,4-D by aerobic bacteria, the side chain is firstly removed by the 2,4-D/α-ketoglutarate dioxygenase encoded by the *tfdA* gene (Fukumori and Hausinger, 1993a, b) and the resulting 2,4-dichlorophenol is converted to 3,5-dichlorocatechol (3,5-DC) by the 2,4-dichlorophenol hydroxylase encoded by the *tfdB* gene (Perkins *et al.*, 1990). 3,5-DC is then degraded to a common intermediate by the enzymes of the chlorocatechol (modified) *ortho*-cleavage pathway encoded by the *tfdCDEF* genes (Fig. 20) (Don *et al.*, 1985).

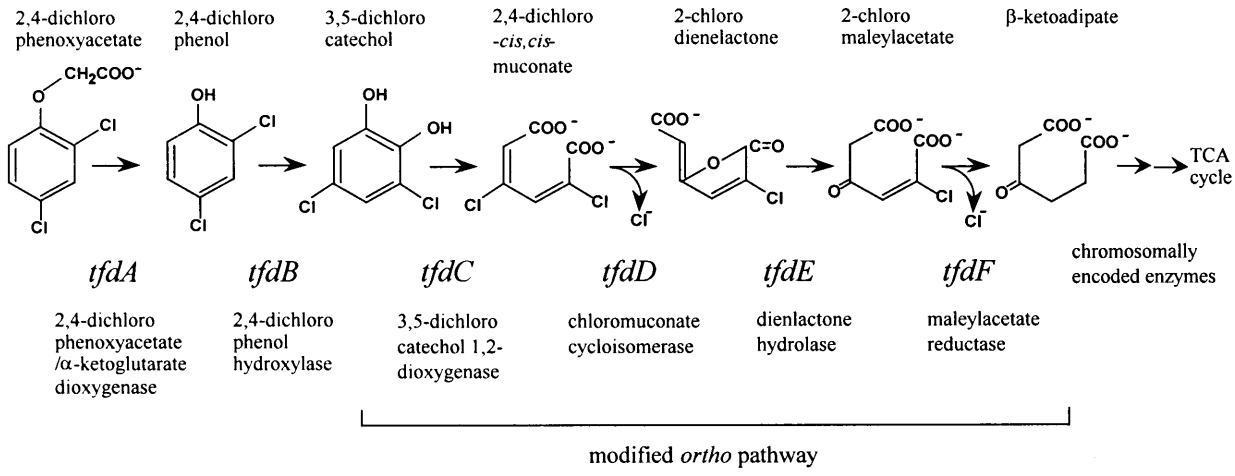


Fig. 20. Pathway for degradation of 2,4-dichlorophenoxyacetate.

2,4-D degradative bacterium *Alcaligenes* sp. CSV90 was found to harbor a 90-kb plasmid pMAB1 and the *tfdC* gene encoding 3,5-dichlorocatechol dioxygenase was cloned from the plasmid (Bhat *et al.*, 1993, 1994). The nucleotide sequence of the *tfdC* gene and a part of the *tfdD* gene contained in the cloned 1.6-kb *Hin*dIII fragment from pMAB1 was identical to the corresponding region of the *tfdCD* genes of pJP4 (Bhat *et al.*, 1994; this study). Sequencing analysis of the downward flanking region of the *tfdC* gene of pMAB1 revealed the presence of the complete *tfdD* gene and the *tfdEF-B* genes which were nearly identical to those of pJP4. At the upstream region of the *tfdC* gene of pMAB1, a gene apparently encoding a LysR-type regulator was found in the opposite orientation, which was highly homologous to the *tfdT* gene

```
1st Nucleotide Sequence: tfdT(pMAB1), 888 bp
2nd Nucleotide Sequence: tfdT(pJP4),  687 bp

  Unit Size to Compare = 2
  Pick up Location     = 1

[100.0% / 633 bp]     INT/OPT.Score : <   2532/  2532 >

     1' ATGGAAATAA GACAGTTGAA ATACTTCGTC GCGGTCGCGG AGGCGGGAGG TTTCGGAACA GCGGCACAGA GGATGCACAT
        ********** ********** ********** ********** ********** ********** ********** **********
     1" ATGGAAATAA GACAGTTGAA ATACTTCGTC GCGGTCGCGG AGGCGGGAGG TTTCGGAACA GCGGCACAGA GGATGCACAT

    81' ATCGCAGCCA CCGCTTACTC GTCAGATCCA GGCCTTGGAA CGCGATATTG GGGCAAAGCT TTTCGAGCGG ACGGCACGGG
        ********** ********** ********** ********** ********** ********** ********** **********
    81" ATCGCAGCCA CCGCTTACTC GTCAGATCCA GGCCTTGGAA CGCGATATTG GGGCAAAGCT TTTCGAGCGG ACGGCACGGG

   161' GCGTCGAACT CACCGCCGCC GGAAAGGTGT TTCTTGATGA CGCGCGTCAG TTGCTCGCGC TTGTGCAGCG GTCTTCGCGG
        ********** ********** ********** ********** ********** ********** ********** **********
   161" GCGTCGAACT CACCGCCGCC GGAAAGGTGT TTCTTGATGA CGCGCGTCAG TTGCTCGCGC TTGTGCAGCG GTCTTCGCGG

   241' CGATCGCAAG CTGCCGCTAG AGGCGAATCG GGTGAGCTGA AGCTCGTCTA CTTTGGGACT CCTGTCTTTG AGACGGTTCC
        ********** ********** ********** ********** ********** ********** ********** **********
   241" CGATCGCAAG CTGCCGCTAG AGGCGAATCG GGTGAGCTGA AGCTCGTCTA CTTTGGGACT CCTGTCTTTG AGACGGTTCC

   321' GGCGTTCGTC AGAACGTTCC TCGCAACGTA TCCCGACGCC ACCGTCGCCG TGTCTCACAT GACCAAGGAA GCTCAGCTCG
        ********** ********** ********** ********** ********** ********** ********** **********
   321" GGCGTTCGTC AGAACGTTCC TCGCAACGTA TCCCGACGCC ACCGTCGCCG TGTCTCACAT GACCAAGGAA GCTCAGCTCG

   401' AGTCTCTTCT TTCAGGGGTG GTGGACATCG GATTCGGGCG CTTCTATCCG GTGACCGAAG GCGTGTCGAG TTGGAACATT
        ********** ********** ********** ********** ********** ********** ********** **********
   401" AGTCTCTTCT TTCAGGGGTG GTGGACATCG GATTCGGGCG CTTCTATCCG GTGACCGAAG GCGTGTCGAG TTGGAACATT

   481' GGGACGGAGA CGCTTCACGT CGCCGCTGCC GATCCGTGGG ACACACGCGT TTCCCGTGCC CGTGCAGTTG TAGACCTACT
        ********** ********** ********** ********** ********** ********** ********** **********
   481" GGGACGGAGA CGCTTCACGT CGCCGCTGCC GATCCGTGGG ACACACGCGT TTCCCGTGCC CGTGCAGTTG TAGACCTACT

   561' CGATGTGCCC CTCATTCTTT ATCCACGCGG CGACCGACCA AGTTTCGCCG ACAAGGTAGT TTCCATATTT AGAACGGAAG
        ********** ********** ********** ********** ********** ********** ********** ***
   561" CGATGTGCCC CTCATTCTTT ATCCACGCGG CGACCGACCA AGTTTCGCCG ACAAGGTAGT TTCCATATTT AGAGACCGTT

   641' GAGTGGAGCC AAAAATCGCA GCCGAGGTCG AAGATGTCAA CGCGGCCCTC GGGCTCGTCG CGGCCGGCGC CGGCGTCACG

   641" TCAAAAAGAG CCCCGGGGGG CTGTTAACCT TTTCGGCAAG CTGTTAA
```

Fig. 21. Alignment of the nucleotide sequences of the *tfdT* genes from pMAB1 and pJP4. The underline indicates 33 nucleotides of the *tfdT* gene from pMAB1 that are not highly similar to either corresponding portions of *tfdT* (pJP4) or *tfdR* (pJP4).

located at the corresponding part of pJP4. The 633-bp nucleotide sequence of the 5'-end of the gene on pMAB1, designated *tfdT*, was identical to the corresponding part of the *tfdT* gene of pJP4 (Fig. 21). However, in contrast to the incomplete length of the *tfdT* gene (687 bp) of pJP4 which was disrupted by the ISJP4 composite transposon (Leveau and van der Meer, 1996), the *tfdT* gene on pMAB1 had a length of 888 bp thus

```
1st Nucleotide Sequence: tfdT(pMAB1), 888 bp
2nd Nucleotide Sequence: tfdR(pJP4),  888 bp

  Unit Size to Compare = 2
  Pick up Location     = 1

[66.9% / 893 bp]     INT/OPT.Score : <   1692/  1728 >

     1' ATGGAAATAA GACAGTTGAA ATACTTCGTC GCGGTCGCGG AGGCGGGAGG TTTCGGAACA GCGGCACAGA GGATGCACAT
        *****  *   ***** *     ** *****  ** * ***** *** ***      ****  *  ** ** * *  ** **** **
     1" ATGGAGTTTC GACAGCTTCG CTATTTCGTT GCTGCCGCGG AGGAGGGCAA CGTCGGTGCC GCCGCGCGGC GGCTGCATAT

    81' ATCGCAGCCA CCGCTTACTC GTCAGATCCA GGCCTTGGAA CGCGATATTG GGGCAAAGCT TTTCGAGCGG ACGGCACGGG
         ** *****  *** * ** * * ***** **  **  * *** *   ** * * * *    * *  ********  *  ** ** *
     81" TTCCCAGCCC CCGGTCACGC GACAGATTCA CGCGCTCGAA CAGCATCTGG GCGTGTTGTT GTTCGAGCGC AGCGCGCGCG

    161' GCGTCGAACT CACCGCCGCC GGAAAGGTGT TTCTTGATGA CGCGCGTCAG TTGCTCGCGC TTGTGCAGCG GTCTTCGCGG
         ****  * ** ***  ***** **    * ** **** ** **  ** ** *    ******  * * *  * *    **   *
    161" GCGTGCAGCT CACGCCCGCC GGGGCCGCGT TTCTCGAAGA TGCACGGCGC ATGCTCGAAC TGGGTCGGAC TTCCGTGGAC

    241' CGATCGCAAG CTGCCGCTAG AGGCGAATCG GGTGAGCTGA AGCTCGTCTA CTTTGGGACT CCTGTCTTTG AGACGGTTCC
         ** ** *  * * ***    *  *****     **  * **   *  *** *** * * ** **   *  ***    **** ** **
    241" CGGTCCCGCG CCGCCAGCCG GGGCGAGATC GGCCAACTCG ATATCGGCTA CCTCGGCACG GCGATCTACC AGACCGTCCC

    321' GGCGTTCGTC AGAACGTTCC TCGCA-ACGT ATCCCGACGC CACCGTCGCC GTGTCTCACA TGACCAAGGA AGCTCAGCTC
         *** **  **      **** *  ****  **    ** * **  **  *  *  ** * *  * ** ******      *** **
    321" GGCATTGCTC CATGCGTT-C ACGCAGGCGG TCCCGGGGGC GACGCTGTCT CTGGCCCTGA TGCCCAAGGT GCGGCAGATC

    400' GAGTCTCTTC TTTCAGGGGT GGTGGACATC GGATTCGGGC GCTTCTATCC GGTGACCGAA GGCGTGTCGA GTTGGAACAT
         *** * ** *  * * **      *  *  ** **  **** * ******* **     *       **  *  *  * **** **
    400" GAGGCCCTGC GTGCCGGCAC CATCCATCTC GGTGTCGGCC GCTTCTACCC CCAGGAGCCT GGAAT--CAC GGTGGAGCAC

    480' TGGGAC--GG AGACGCTTCA CGTCGC--CG CTGCCGATCC GTGGGACACA CGCGTTTCCC GTGCCCGTGC AGTTGTAGAC
           * **   * *   ***  *   ****  **     *** *   ** *   **    *   *   *  **   **  * **  *
    478" CTGCACTACG AACGGCTGTA TATCGCAGCG GGTTCGAGCA TTGCGCGCCA GCTGAGACAG GATCCGACGC TGCTGCGGCT

    556' CTACTCGATG TGCCCCTCAT TCTTTATCCA CGCGGCGACC GACCAAGTTT CGCCGACAAG GTAGTTTCCA TATTTAGAAC
         * *      * *  ***** * *****  **     *  *   * ** ***** *** *** *  **  *  **  *  *  *
    558" CAAGAGCGAG T--CCCTCGT TCTTTTCCCC AAGGAGGGGA GGCCGAGTTT CGCTGACGAA GTGATCGCCT TGATGCGCCG

    636' GGAAGGAGTG GAGCCAAAAA TCGCAGCCGA GGTCGAAGAT GTCAACGCGG CCCTCGGGCT CGTCGCGGCC GGCGCCGGCG
         **  ** **  *****      *  * **  ********* ********** ********** ********** **********
    636" GGCCGGGGTC GAGCCGCGCG TGACGGCGAT TGTCGAAGAT GTCAACGCGG CCCTCGGGCT CGTCGCGGCC GGCGCCGGCG

    716' TCACGCTGGT CCCGGCCTCG GTGGCCGCGA TTCGGCGGCC CTTCGTCCGG ACGATGGAGA TGGCCGATGC GAGCGCCAAG
         ********** ********** ********** ********** ********** ********** ********** ***** ****
    716" TCACGCTGGT CCCGGCCTCG GTGGCCGCGA TTCGGCGGCC CTTCGTCCGG ACGATGGAGA TGGCCGATGC GAGCGACAAG

    796' GTGCCGGTCA GCCTGACTTA CCTGACCGAC TCTCGCGTAC CCGTGCTTCG CGCATTTCTC GATGTCGCAA GACGCGGGAA
         ********** ********** ********** ********** ********** ********** ********** **********
    796" GTGCCGGTCA GCCTGACTTA CCTGACCGAC TCTCGCGTAC CCGTGCTTCG CGCATTTCTC GATGTCGCAA GACGCGGGAA

    876' AGGACAGAAA TAG
         ********** ***
    876" AGGACAGAAA TAG
```

Fig. 22. Alignment of the nucleotide sequences of the *tfdT* gene from pMAB1 and the *tfdR* gene from pJP4. The underline indicates 33 nucleotides of the *tfdT* gene from pMAB1 that are not highly similar to either corresponding portions of *tfdT* (pJP4) or *tfdR* (pJP4).

seemed to encode a full-length LysR-type regulator. Curiously, the 222-bp nucleotide sequence of the 3'-end of the *tfdT* gene of pMAB1 was nearly identical to the corresponding part of the *tfdR* (or *tfdS*) gene of pJP4 (Fig. 22). Nucleotide sequence of 33 bp between the two parts in the *tfdT* gene of pMAB1 (underlined in both Fig. 21 and 22) did not exhibit high similarity to known sequences of pJP4 or examined sequence of pMAB1 shown in Fig. 23. Further sequencing analysis of the downward flanking region of the *tfdT* gene on pMAB1 revealed another *lysR*-type regulatory gene *tfdS* and the *tfdA* gene, both of which were identical to those of pJP4 (Fig. 23).

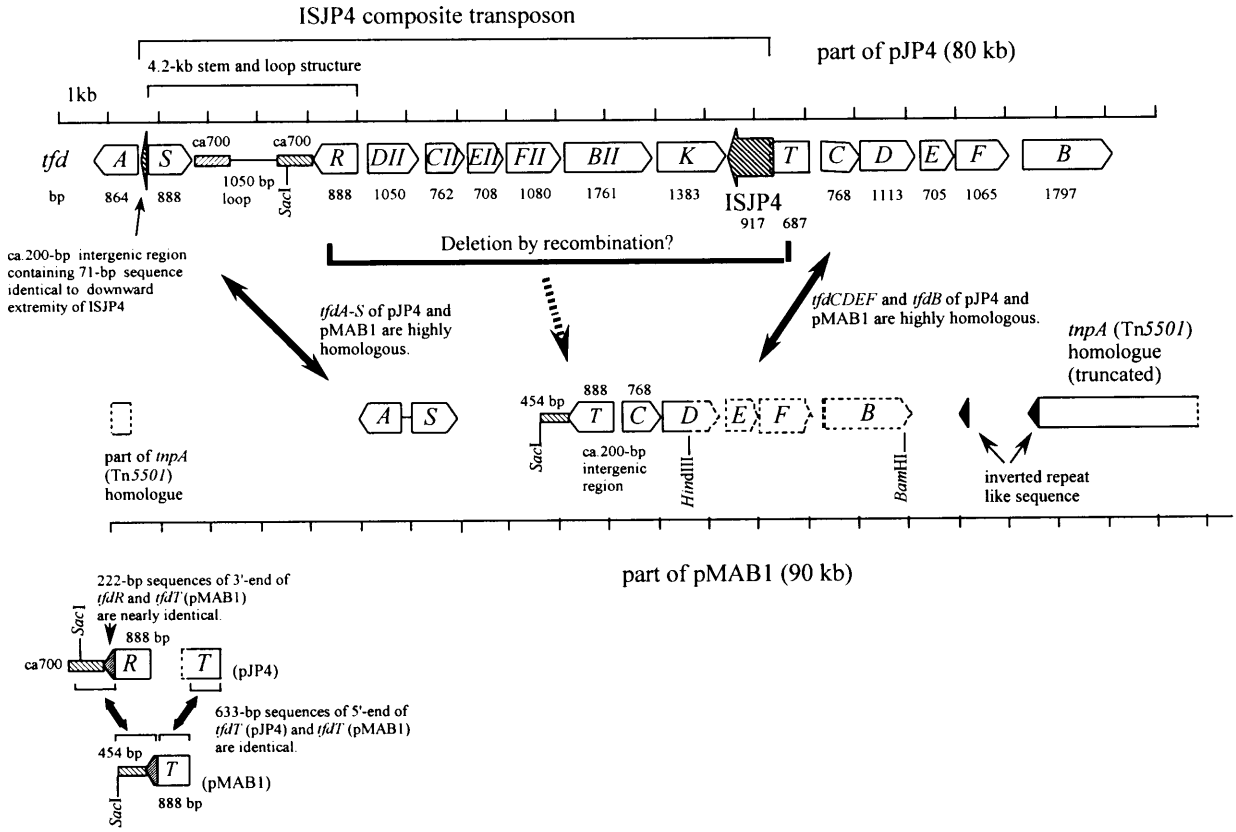


Fig. 23. Schematic representation of regions containing 2,4-D degradative genes on pJP4 and pMAB1.

3.2 Activation of the *tfdC* promoter by TfdS

There was possibility that the two LysR-type regulators may act on either of the *tfdC* promoter and the *tfdA* promoter. To differentiate the possible role of the *tfdT* and *tfdS* genes in transcriptional regulation of the promoters of the degradative genes, several constructs for each promoter were made including those in which either of the two *lysR*-type genes was disrupted and used for reporter analysis (Fig. 24 and 25). The same method was used as that used for the analysis of activation of the *cbnA* promoter. It was presumed 2-CM could be an inducer for the *tfdC* promoter because this strain can degrade 3-CB by using the enzymes encoded by the *tfdCDEF* genes which are homologous to the *clcABDE* genes and the *cbnABCD* genes both regulated by related LysR-type activators recognizing 2-CM as inducer (McFall *et al.*, 1997c; this study).

When the cells containing pCS50STCD' were grown on glucose in the presence of 3-CB, the transcription from the *tfdC* promoter was increased 2-fold compared to that in the cells grown on glucose alone (Table 4). This elevation in the presence of 3-CB was also seen for the cells with pCS50ST'CD' (*tfdT* disrupted) but was not seen for the cells with pCS50S'TCD' (*tfdS* truncated) or pCS50TCD' containing only *tfdT*. These results indicated that *tfdS* activated the *tfdC* promoter while *tfdT* did not.

Cells containing pCS50TC' or pCS50STC' exhibited very low basal level activity when grown on glucose alone, compared to those containing the constructs with the full length *tfdC* gene, and did not show any increased transcriptional activity in the presence of either 3-CB or Ben. These results indicated neither of Ben, catechol, 3-CB, or 3-CC could induce the transcriptional activation of the *tfdC* promoter by *tfdS* or *tfdT*. The results endorsed that 2-CM was the inducer for the activation of the *tfdC* promoter shown with pCS50STCD' and pCS50ST'CD'. The high basal level observed with the cells containing the constructs with truncated *tfdD* gene could be ascribed to the absence of a cognate regulator of the *tfdC* promoter and to some *cis* effect of the region of the *tfdC* gene (see Discussion).

In order to examine the effect of the distance between the *tfdS* gene and the *tfdC* promoter on the transcriptional activity, pCS50SCD'1 and pCS50SrCD'1 were constructed and used for reporter analysis. While the cells with pCS50SCD'1 did not show substantial difference in the activation by 3-CB from those with pCS50STCD' and pCS50ST'CD', the cells with pCS50SrCD'1 induced the highest activity among all the experimental sections with various constructs in the presence of 3-CB (5-fold activation compared to the cells grown with glucose alone) (Table 4). These results suggested the possibility of the effect of local DNA structure (see Discussion).

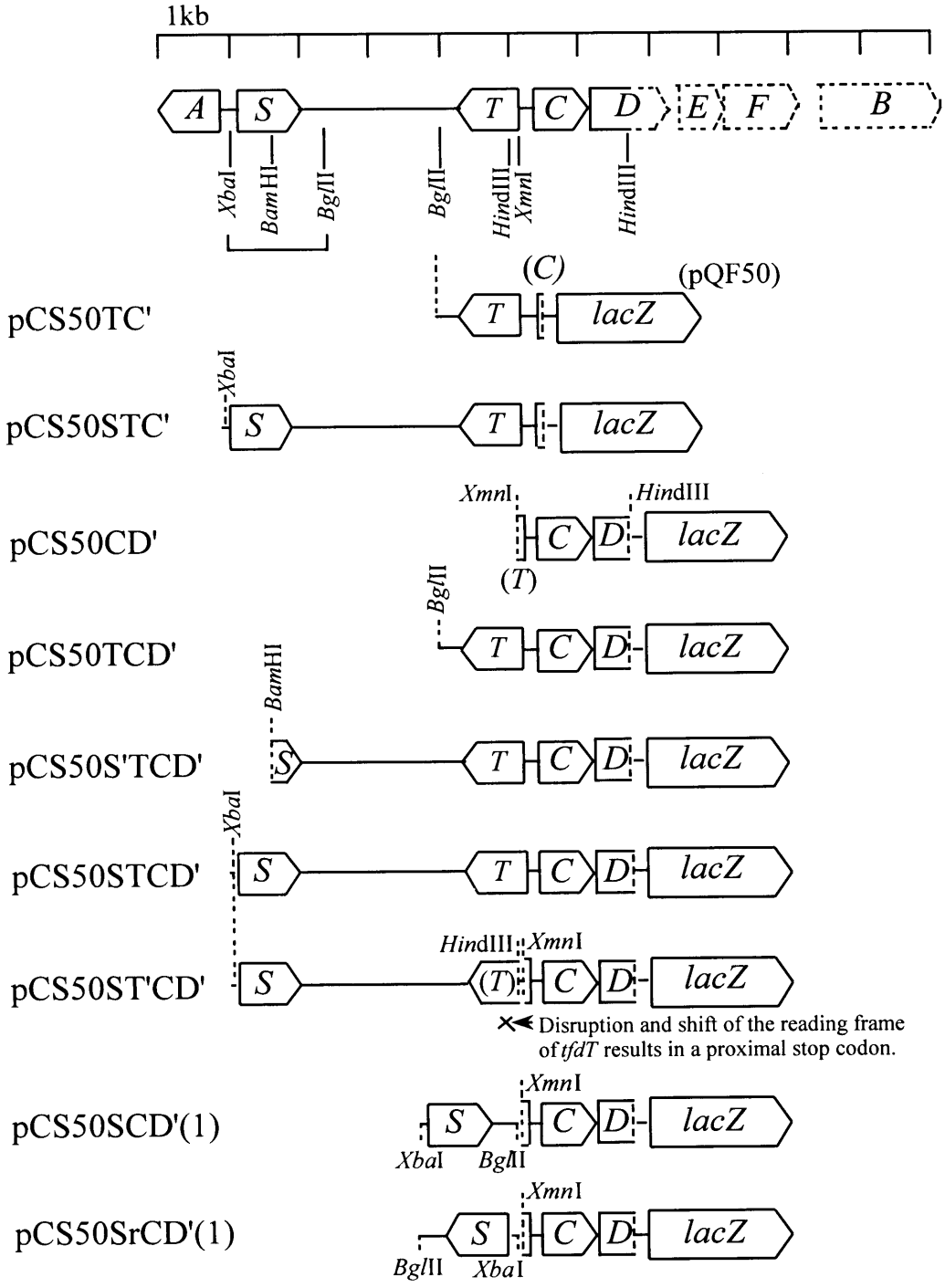


Fig. 24. Diagram of the inserts of the constructs used for activity assay of the *tfdC* promoter of pMAB1.

Table 4. The effect of Ben and 3-CB on transcriptional activation at the *tfdC* promoter of *Alcaligenes* sp. CSV90 (pMAB1)

P.putida PRS4020 /plasmid construct	Glucose 10mM	Glucose 10mM + Benzoate 5mM	Glucose 10mM + 3-CB 5mM
pCS50TC'	7.9 (±2.0)	5.7 (±1.8)	5.1 (±2.2)
pCS50STC'	8.6 (±1.7)	7.9 (±1.4)	7.7 (±2.1)
pCS50CD'	136.3 (±10.3)	93.0 (±3.1)	72.2 (±5.4)
pCS50TCD'	109.0 (±4.2)	84.2 (±8.6)	91.4 (±16.6)
pCS50S'TCD'	191.3 (±20.6)	141.7 (±5.0)	109.0 (±11.9)
pCS50STCD'	115.8 (±10.9)	139.8 (±10.9)	234.7 (±12.7)
pCS50ST'CD'	104.4 (±8.5)	136.7 (±12.3)	225.6 (±21.7)
pCS50SCD'1	160.7 (±3.8)	152.5 (±7.6)	284.2 (±15.5)
pCS50SrCD'1	126.5 (±4.1)	118.7 (±11.0)	630.5 (±36.6)

β-galactosidase units: nmol/min/mg extract; (±): standard deviation

3.3 Activation of the *tfdA* promoter by TfdS

Transcriptional activation of the *tfdA* promoter was tested using PRS4020 cells with each of the four constructs depicted in Fig. 25. When the cells containing pCS50CTSA' or pCS50CT'SA' (*tfdT* disrupted) were grown on glucose with 3-CB, the transcription from the *tfdA* promoter was increased 37- and 38-fold, respectively, compared to those in the respective cells grown on glucose alone (Table 5). In contrast, the cells with pCS50CTS'A' (*tfdS* disrupted) did not exhibit any activation upon the addition of 3-CB. These results indicated that *tfdS* was the positive regulator for the transcription at the *tfdA* promoter. No activation observed for the cells with pCS50C'TSA' (*tfdC* truncated) in the presence of 3-CB further indicated that 2-CM was the inducer for the activation of the *tfdA* promoter by *tfdS*.

The basal level obtained with the cells grown on glucose alone was quite low throughout the four experimental sections. Three sections containing the constructs with the intact *tfdS* gene showed a constant level of activation in the presence of Ben (11- to 15- fold), whether or not the construct contained the full length of the *tfdC* gene. This activation in the presence of Ben was not induced for the cells with pCS50CTS'A' (*tfdS* disrupted). These results could suggest that either Ben, catechol or CCM might serve as inducer for the transcriptional activation at the *tfdA* promoter by *tfdS*.

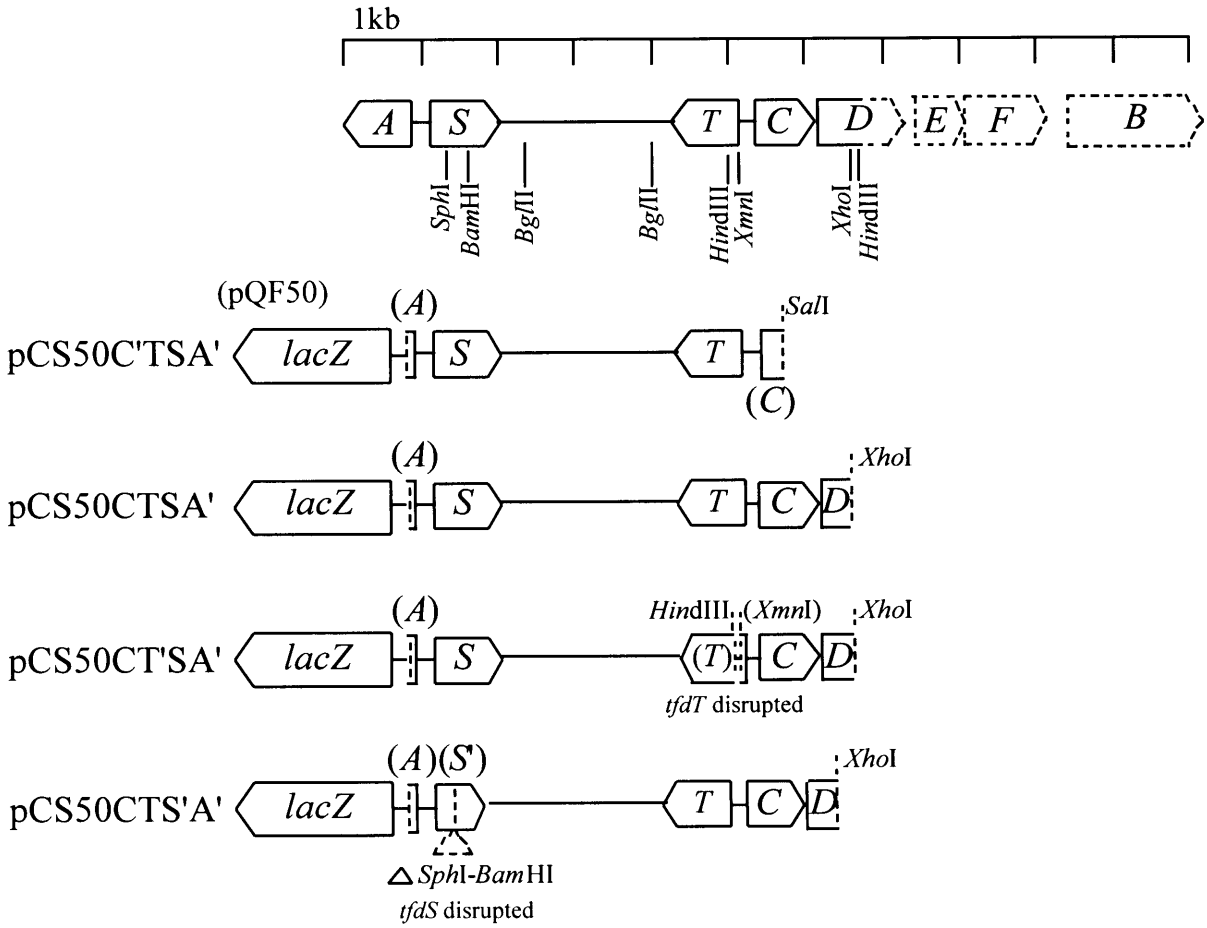


Fig. 25. Diagram of the inserts of the constructs used for activity assay of the *tfdA* promoter of pMAB1.

Table 5. The effect of Ben and 3-CB on transcriptional activation at the *tfdA* promoter of *Alcaligenes* sp. CSV90(pMAB1)

P.putida PRS4020 /plasmid construct	Glucose 10mM	Glucose 10mM + Benzoate 5mM	Glucose 10mM + 3-CB 5mM
pCS50C'TSA'	2.0 (±0.3)	29.2 (±2.3)	4.9 (±1.1)
pCS50CTSA'	2.6 (±0.3)	29.3 (±0.5)	97.5 (±7.7)
pCS50CT'SA'	2.1 (±0.2)	32.7 (±0.7)	80.2 (±10.5)
pCS50CTS'A'	2.3 (±0.1)	2.2 (±0.1)	1.8 (±0.2)

β-galactosidase units: nmol/min/mg extract; (±): standard deviation

Summary of results section 3: The *tfdA-S-T-CDEFB* genes on plasmid pMAB1 (90kb) from *Alcaligenes* sp. CSV90 have turned out to be highly homologous (>98% at the nucleotide level) with the corresponding genes of the 2,4-D degradative gene cluster (*tfdA-S-R-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K-T'-CDEFB*) on plasmid pJP4 (80kb) of *R. eutropha* JMP134 which contained the additional genes (*tfdR-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K*). The genes *tfdS* and *tfdT* apparently encoded LysR-type transcriptional regulators. These results indicated evolutionarily close relationship between the two degradative gene regions, suggesting recent genetic recombination event(s). Transcriptional fusion assays showed that the *tfdC* promoter was not regulated by *tfdT* but that it was activated by the distantly-located *tfdS* in the presence of 2-chloromuconate which is an analogous compound of 2,4-dichloromuconate, the metabolite from 3,5-dichlorocatechol by TfdC. TfdS also activated the *tfdA* promoter in the presence of (chloro)muconate.

Discussion

1. The structure of the degradative genes for 3-chlorocatechol and 2,4-dichlorophenoxyacetate

1.1 A novel 3-chlorocatechol catabolic plasmid pENH91

Plasmid pENH91 from *R. eutropha* strain NH9 that carries genes for catabolism of chlorocatechols was characterized in this study. The restriction pattern of pENH91 (Fig. 3) was different from those of other plasmids with chlorocatechol degradative genes (Chatterjee and Chakrabarty, 1984; Chaudhry and Huang, 1988; van der Meer *et al.*, 1991c). NH9 could not grow on 2,4-D, and the restriction profile of pENH91 was different from those of 2,4-D degrading plasmids (Bhat *et al.*, 1994; Don and Pemberton, 1985; Mäe *et al.*, 1993). Therefore, it was apparent that pENH91 was a novel chlorocatechol degradative plasmid.

1.2 The structure of IS*1600* and the chlorocatechol catabolic transposon Tn*5707*

The chlorocatechol degradative genes on plasmid pENH91 were found to be located between two copies of IS*1600* thus the ca. 15-kb region formed the composite class I-type transposon Tn*5707* (Fig. 3 and 4).

The homologies and the overall structural similarity among IS*1600*, IS*1326*, and IS*21* indicated an evolutionary relationship (Fig. 7 and 26).

The genes *istA* and *istB* of IS*1600* could encode transposase/cointegrase and a helper protein, respectively, by analogy with those of IS*21* (Schmid *et al.*, 1998, 1999). The *istA* gene of IS*21* encodes both transposase and cointegrase. While the transposase of IS*21* is the full-length product (46 kDa) of the *istA* gene, its cointegrase (45 kDa) is translated from an internal translational start of *istA* and devoids of eight amino acids of the amino (N-) terminus of the transposase. The functional specialization of the transposase was attributed to the eight amino acid residues (MLSREDYF) and it was suggested that the highly polar terminus might help the protein to bring the ends of IS*21* together in a putative mechanism for simple insertion (transposition of one copy of IS*21* element)(Schmid *et al.*, 1998). The N-terminus of IstA of IS*21* contains six amino acids (M, S, R, E, D, and probably Y) that are polar or charged. The N-terminus of the deduced amino acid sequence of IstA of IS*1600* is MIDVATLS (Fig. 6). Four amino acids (M, D, T, and S) out of the eight are polar or charged. Alternatively, if the translation starts at the second GTG codon, the N-terminus is MATLSVIR and four amino acids (M, T, S, and R) are polar or charged.

If highly polar N-terminus is important for the function as transposase, this might suggest either that the transposase encoded by the *istA* gene of IS*1600* might not be highly effective or that catalytic activity of cointegrase could be dominant at its full-length product of *istA* of IS*1600*.

Although the terminal inverted repeats of IS*1600*, IS*1326*, and IS*21* differed in length (Fig. 7b), there was some similarity between the nucleotide sequences. Besides the high similarity between the inverted repeats of IS*1600* and IS*1326*, the inverted repeats of the three ISs conserved a dinucleotide CA-3' at their extreme ends. This dinucletide is known to be significantly preferred for cleavage and joining by retroviral integrases containing the catalytic domain of DDE amino acid residues during integration process of retroviruses to host DNA (Goff, 1992; Katz and Skalka, 1994; Hindmarsh and Leis, 1999). The DDE triad, which is proposed to coordinate a divalent metal ion essential for catalytic activity, are conserved among the IstAs of the ISs of IS*21* family as well as among transposases of many bacterial ISs and retroviral integrases (Haren *et al.*, 1999). The dinucleotide has also been suggested to be critical for the cointegrate formation of plasmids containing two copies of IS*21* (Haas *et al.*, 1996) and thus could be important for the function of IstA of IS*1600*.

Because pENH91 was found not to coexist stably with RP4, RP4 was introduced into NH9 by conjugation, expecting to obtain isolates in which the catabolic genes were transposed onto either RP4 or the chromosome of NH9. Although a few colonies appeared on plates prepared with 3-CB and Km, none of these isolates was stable and no evidence of transposition has yet been obtained. Thus, the transposability of the 15-kb region remains to be demonstrated.

The structure of the 15-kb region might indicate that this region had been transposed into a region of a plasmid that was an antecedent of pENH91. But this region was not flanked by duplicated sequences of the target site, which was supposed to be generated by transposition. Some explanations are possible for this structure. (i) Integration of an antecedent of pENH91 with one copy of IS*1600* element into a target site flanking chromosomal chlorocatechol genes by replicative transposition, followed by intramolecular transposition of the element into a site beyond the chlorocatechol genes and deletion. This mechanism has been proposed for mobilization of catabolic genes into plasmids (Wyndham *et al.*, 1994; Di Gioia *et al.*, 1998) and accounts for the lack of target-site duplications. (ii) Insertion of an IS*1600* element proximal to chlorocatechol genes of an antecedent of pENH91 that had originally contained the chlorocatechol genes, followed by duplication of the element. This case has been observed for the 2,4,5-trichlorophenoxyacetate genes and associated insertion sequences

(Haugland *et al.*, 1990). This mechanism can account for the observed 15-kb structure of pENH91 if the transposition of the element did not generate target-site duplication, as was found for a certain kind of IS elements (Diaz-Aroca *et al.*, 1987; Joset and Guespin-Michel, 1993).

1.3 Duplication and deletion of the catabolic genes on the plasmid pENH91

Duplication and deletion of the degradative genes were demonstrated to be consequences of recombination between the two IS*1600* elements.

There have been a few reports of the amplification of genes for 3-CB (or 3-CC) degrading enzymes. The nucleotide sequences involved in these recombination events have not been reported. Amplification of the 4.3-kb region in *Pseudomonas* sp. strain B13 seemed to be necessary to avoid the accumulation of intermediate products of the degradation of 3-CB (Rangnekar, 1988). Similar amplification of a cloned 4.2-kb fragment from pAC27 and a 15-kb fragment from pJP4 were also ascribed to the necessity for high gene-dosage, for example, in order to compensate for the absence of a positive regulatory element (Ghosal *et al.*, 1985). In the case of the inverted duplication on pJP4 and the pre-existing tandem duplication on pJP2, the advantage was attributed to mutation of the duplicated copy with resultant acquisition of new catabolic phenotypes, rather than to a gene-dosage effect (Ghosal and You, 1988). Recently, tandem amplification of the 105-kb *clc* element containing the *clcR-ABCD* genes on the chromosome of *P. putida* F1 has been described (Ravatn *et al.*, 1998a). The amplification was prerequisite for the growth of the strain on chlorobenzene. It seemed that the strain could obtain enough enzymatic activity for the efficient conversion of chlorocatechol, which is a critical step in the chlorobenzene degradation pathway, only by expressing the *clc* genes from multiple copies (Ravatn *et al.*, 1998a). While the genetics and function of the recombinase of the *clc* element as the bacteriophage P4 type integrase have been described (Ravatn *et al.*, 1998b), the process of the amplification to more than two copies at one locus has not been fully characterized as to whether it is due to the replicative transposition of the *clc* element or due to homologous recombination between directly oriented *clc* elements.

Amplification of chromosomal genes due to general recombination mediated by homologous IS elements have been reported for *Salmonella typhimurium* under laboratory selection condition (Haack and Roth, 1995) and for *E. coli* strain W3110 (Hesslinger and Sawers, 1998). The finding of duplication of the genes in NH9(A) is rather unique in that it was found in *Ralstonia* and on a plasmid. Interestingly, the transcription of the IS*200*-encoded orf increased the frequency of homologous

recombination between directly repeated chromosomal IS*200* elements (Haack and Roth, 1995). In the case of amplification of the *cbn* genes on pENH91, IstA of IS*1600* might play a role, as well as the recombinase of the *clc* element for its tandem amplification.

No major differences between NH9 and NH9A have yet been observed when features of their growth in 3-CB-containing liquid medium or their tolerance to 3-CB are compared. In Fig. 8b (lanes 1 and 4) (plasmid DNA and total DNA of NH9, respectively), very weakly hybridized bands can be seen at the same size of the extra fragment of the duplicated plasmid of NH9A. This suggested that culture of NH9 would contain NH9A at a low ratio. However, the fact that NH9A became dominant after 1 year of successive subculturing indicates that duplicated catabolic genes provide cells with an advantage under our laboratory conditions. The functional significance of the gene duplication on pENH91 could be slightly higher growth rate by virtue of efficient conversion of chlorocatechol resulting from increased production of the enzymes which could be possible only by duplication of the genes.

1.4 The similarity between the *cbnR-ABXCD* genes and the *tcbR-CDXEF* genes and their different roles in their respective genetic backgrounds

The chlorocatechol-degradative genes of NH9 (*cbnR-ABCD*) were found to be highly homologous to those of *Pseudomonas* sp. P51 (*tcbR-CDEF*) (van der Meer *et al.*, 1991a, b). The lengths of the corresponding orfs, the overlaps of orfs, and the intervening sequences between orfs were the same in the two gene clusters. Highly conserved nucleotide sequences and the identical overall structures of the two gene clusters indicated that the horizontal transfer and the divergence at the nucleotide level of the two clusters had occurred relatively recently in the evolutional history of the clusters of genes in the modified *ortho* pathway (Schlömann, 1994).

The frequency of nucleotide divergence between the regions that contained *cbnB* and *tcbD* was significantly higher than that between the other corresponding regions of the two clusters (Table 2 and Fig. 11). This point is of particular interest since the genes for chloromuconate cycloisomerase are generally more conserved than the other three genes in the cluster of modified *ortho* pathway genes (van der Meer *et al.*, 1991a). If selective pressure on the genes caused such divergence, several features should be noted here. Even though the two clusters were located on plasmids, the G+C content of each gene in the two clusters, which ranged from 61% to 66% in either cluster (Table 2), cannot be regarded as being significantly different from the chromosomal G+C content of *R. eutropha* (66.3% to 66.8 %) (Davis *et al.*, 1969) or of species of

Pseudomonas [e.g., *P. putida*, 59.6% to 63.4 % (Mandel, 1966)]. Furthermore, the G+C contents of *cbnB* (63.3%) and *tcbD* (63.7%) are similar and in the middle range of those of all the genes in the respective clusters. Therefore, the nucleotide divergence between *cbnB* and *tcbD* cannot be attributed to "GC pressure" from the host. Alternatively, one possible explanation for the rapid divergence of *cbnB* and *tcbD* is that some pattern of biased codon usage forced the bacteria in which these genes were located to replace nucleotides in an effort to adapt to their genetic background, for example, the pool of transfer RNAs. However, no significant differences in codon-usage patterns between *cbnB* and *tcbD* or between *cbnB* (*tcbD*) and the other genes in the clusters have been found. At present, it is unknown whether any selective pressure forced the rapid nucleotide divergence of *cbnB* and *tcbD*.

The fact that numerous nucleotide substitutions were not limited to the *cbnB* (*tcbD*) gene but spanned a region of ca. 1.3 kb that contained *cbnB* (*tcbD*) (Fig. 11) suggests that this divergence might have been caused by physical conditions in this region of DNA and not by the genetic nature of *cbnB* (*tcbD*). The rate of substitutions is high throughout *cbnB* (*tcbD*). Substitutions in the flanking genes tended to be localized in regions proximal to *cbnB* (*tcbD*): most of the nucleotide substitutions (18 out of 21) between orfXs (next to *cbnB* or *tcbD*) were located within 138 bp of the 5' portion of orfXs (1011 bp), and 5 out of 14 nucleotide substitutions between *cbnA* and *tcbC* were located within 24 bp of the 3' region of the genes (756 bp). In the other corresponding parts of the two 6959-bp *Sac*I-*Kpn*I fragments, the rate of mutations, including substitutions, insertions and deletions, was one in several hundred base pairs, which might reflect the basal rate of spontaneous mutation. Some unidentified local structure of the DNA in the region containing *cbnB* (*tcbD*) might have made this region more vulnerable to substitutions during replication. The high homology kept at amino acid level might reflect some constraint for the function of chloromuconate cycloisomerases.

There was a difference between the chloroaromatic compounds that two strains NH9 and P51 to utilize as substrates for growth. Strain P51 grew either on 1,2- or on 1,4-dichlorobenzene and it also grew on 1,2,4-trichlorobenzene (van der Meer *et al.*, 1991c). Strain NH9 was tested for growth on these compounds, but it failed to grow in liquid medium in the presence of these chlorobenzenes. On the other hand, strain P51 did not grow on 3-CB (van der Meer *et al.*, 1991c). This discrepancy in growth substrates might be explained by the difference in available "upper-pathway" enzymes between the two strains. Strain NH9 probably synthesizes enzymes that convert 3-CB to chlorocatechol but not the enzymes that convert chlorobenzenes to any

further-metabolizable compounds. Strain P51 has been reported to synthesize the enzymes that convert chlorobenzenes to the corresponding chlorocatechols (van der Meer *et al.*, 1991c; Werlen *et al.*, 1996), but it does not seem to synthesize enzymes that convert chlorobenzoates to chlorocatechols. The recruitment of the homologous chlorocatechol-degradative gene clusters by the two strains, which resulted in the difference in utilizable chloroaromatics as growth substrates, illustrates the economy whereby bacteria adapt to xenobiotic compounds.

1.5 The diverse origins of the catabolic transposons carrying chlorocatechol-degradative genes

In contrast to reports on well-described catabolic transposons, such as the toluene transposons on TOL plasmids and the catabolic genes mobilized by IS*1071* (Nakatsu *et al.*, 1991) (for reviews, see Di Gioia *et al.*, 1998; Tsuda, 1996; Tsuda *et al.*, 1999; Wyndham *et al.*, 1994), there have been only two documented examples of transposable elements that carries genes for the modified *ortho* pathway; one is the ISJP4 composite transposon (Leveau and van der Meer, 1997) and the other is the 105-kb *clc* element (Ravatn *et al.*, 1998b). These two and Tn*5707* are diverse in terms of the origin of the transposable elements.

The insertion sequence ISJP4 copy A and its incomplete copy C captured the genes *tfdS-R-*$D_{II}C_{II}E_{II}F_{II}B_{II}K$ to form a composite transposon on plasmid pJP4 in *R. eutropha* JMP134 (Leveau and van der Meer, 1997; Leveau *et al.*, 1998). Although the transposition of this whole composite transposon has not been observed, transposition of an artificial composite transposon comprising two copies of ISJP4 has been detected (Leveau and van der Meer, 1997). With regard to their characteristics as transposable elements, there are some differences between Tn*5707* and the composite transposon formed by ISJP4. IS*1600* of Tn*5707* belongs to the IS*21* family while ISJP4 belongs to the IS*5* group of the IS*4* family. The ISJP4 transposon was apparently transposed to pJP4 as a composite transposon, as indicated by the presence of target-site duplication at both ends (Leveau and van der Meer, 1997). By contrast, the absence of such duplication at both ends of Tn*5707* suggests that Tn*5707* may not transpose as this composite unit. Instead this feature suggests mobilization of the chromosome that resulted from plasmid integration and subsequent excision mediated by IS*1600* (Wyndham *et al.*, 1994). The existence of orfL in Tn*5707* suggests that the origin of the region carried by Tn*5707* might have been a bacterial chromosome since genes for amino acid-transport systems have been found on them (Adams *et al.*, 1990; Hoshino and Kose, 1990). The catabolic genes carried by the ISJP4 transposon are different

from those carried by Tn*5707* as follows. Although the *tfdD*$_{II}$*C*$_{II}$*E*$_{II}$*F*$_{II}$ genes are the most homologous to the *tcbCDEF* genes, the homology between the corresponding genes ranged from 58% to 70% at the nucleotide level and from 27% to 65% at the amino acid level. The ISJP4 transposon contains duplicated regulatory genes, *tfdR* and *tfdS* (Matrubutham and Harker, 1994; You and Ghosal, 1995), as well as additional genes, namely, *tfdB*$_{II}$ that might encode chlorophenol monooxygenase and *tfdK* that encodes an active transporter of 2,4-D (Leveau *et al.*, 1998; Perkins *et al.*, 1990).

The 105-kb *clc* element has recently been found to carry the *clcR-ABCD* genes in *Pseudomonas* sp. strain B13 (Ravatn *et al.*, 1998a). This element also takes the form of the plasmid pB13 especially at stationary phase of the growth of the cell and integrates into chromosomes of various bacterial strains at glycine tRNA structural genes using a site-specific recombinase, Int-B13, which belongs to the bacteriophage P4 integrase family (Ravatn *et al.*, 1998b). The *clc* element has been proposed to be considered as a "degradation island" (Ravatn *et al.*, 1998b).

In addition to the differences in inherent characteristics between Tn*5707* and the other two chlorocatechol transposons, the strong homology between the two clusters, *cbnR-ABCD* and *tcbR-CDEF*, illustrates the role of the IS elements in the recent dissemination of genes in the modified *ortho* pathway.

On the phylogenetic tree of IstAs of the IS*21* family, IstA of IS*1600* formed a distinct cluster together with orfSA from strain P51, IstA of IS*1326*, and NmoT (Fig. 26). The branching point of the IstA of IS*1326* and the other three elements suggests that these four elements diverged relatively recently in the evolution of the members of IS*21* family. IS*1326* was found in integrons in antibiotic-resistant clinical isolates (Brown *et al.*, 1996). NmoT is a putative transposase that corresponds to IstAs and was found proximal to nitrilotriacetate-degradative genes in *Chelatobacter heintzii* (Xu *et al.*, 1997). The various origins of the four elements indicate the recent wide distribution of the related IS (-like) elements among bacteria, which in turn raises the possibility that these IS (-like) elements might have been involved in recent genetic rearrangements of various kinds. Recent findings of the ISs of IS*21* family including the highly reiterated IS*1631* on the chromosome of *Bradyrhizobium japonicum* (Isawa *et al.*, 1999) and several members distributed on the chromosome of *Mycobacterium tuberculosis* H37Rv (Gordon *et al.*, 1999) further illustrate the extensive spread of the ISs of the family.

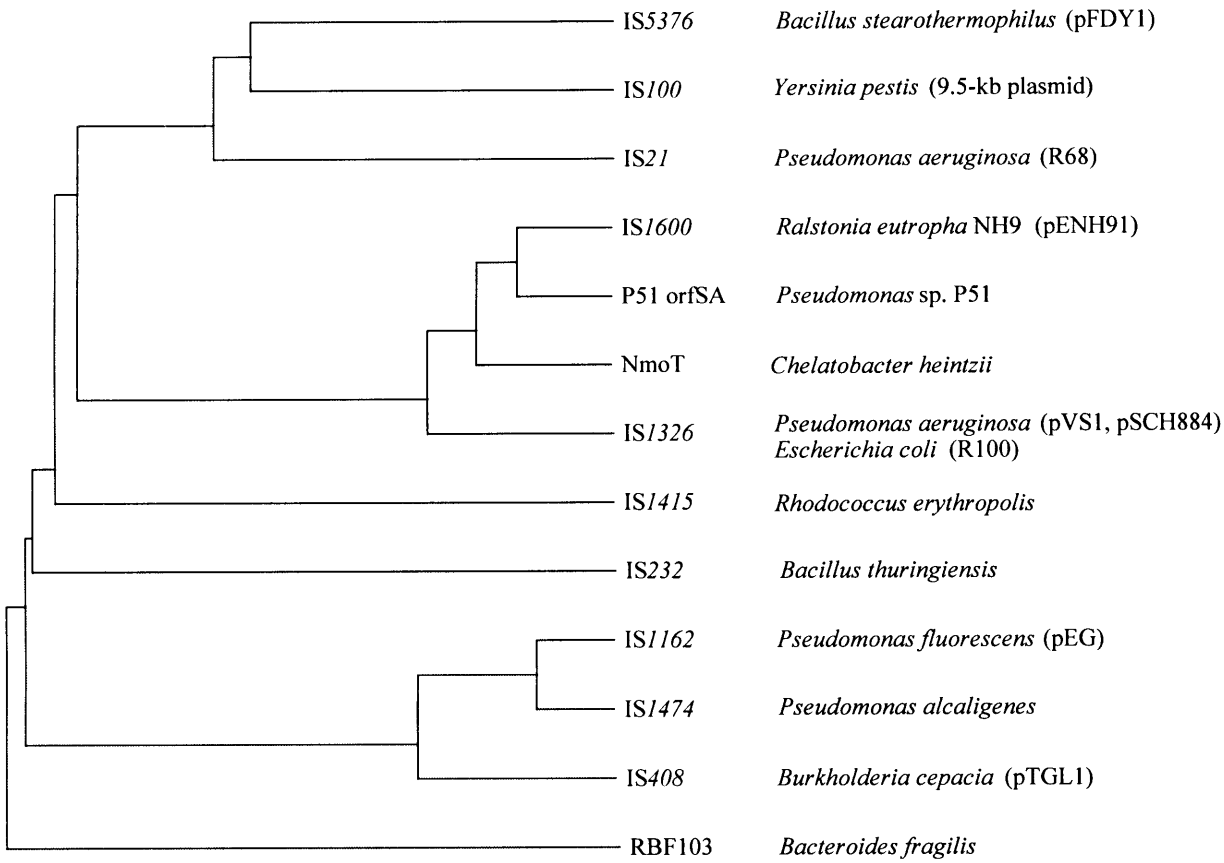


Fig. 26. Phylogenetic tree of members of the IS*21* family based on the alignment of IstA or IstA-like proteins. The alignment was performed with ClustalW software and adjusted manually to incorporate the results reported by Haas *et al.* (1996). The tree was constructed with the Genetyx software program (Software Development Co., Tokyo, Japan) by the unweighted-pair group method with mathematical averages. Accession numbers (and the references) except those for IS*1600* and P51 orfSA are as follows (from top to bottom): X67861 (Xu *et al.*, 1993), Z32853 (Podladchikova *et al.*, 1994), X14793 (Reimmann *et al.*, 1989), L49438 (Xu *et al.*, 1997), U38187 (Brown *et al.*, 1996), AF002247 (Nagy *et al.*, 1997), M38370 (Menou *et al.*, 1990), X79443 (Solinas, *et al.*, 1995), U67315 (Yeo and Poh, 1997), L09108 (Byrne and Lessie, 1994), and U05888 (Rogers *et al.*, 1994).

1.6 The structure of the 2,4-D degradative gene cluster on plasmid pMAB1

In the 2,4-D degradative bacterium *Alcaligenes* sp. CSV90, the degradative genes are located within a ca. 12-kb region on its plasmid pMAB1 (90kb) and are organized as two clusters, *tfdA-S* genes and *tfdT-CDEFB* genes. The genes *tfdS* and *tfdT* apparently encode LysR-type transcriptional regulators and are located in the opposite orientation from the degradative gene(s) in their respective units. The *tfdA-S-T-CDEFB* genes from CSV90 have turned out to be highly homologous (>98% at the nucleotide level) with the corresponding genes of the 2,4-D degradative gene cluster (*tfdA-S-R-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K-T'-CDEFB*) from *R. eutropha* JMP134 (pJP4 [80kb])

which contained the additional genes (*tfdR-$D_{II}C_{II}E_{II}F_{II}B_{II}$K*).

In contrast to the apparent difference between the backbones of the plasmids pMAB1 and pJP4, the above result indicates evolutionarily close relationship between the two degradative gene regions, suggesting recent genetic recombination event(s) (Fig. 23). The structure of the gene clusters on pMAB1 and the high similarities of 5' portion and 3' end of the *tfdT* gene of pMAB1 to 5' portion of the *tfdT* gene of pJP4 and 3' end of the *tfdR* gene of pJP4, respectively, could suggest that reciprocal recombination between the *tfdR* and *tfdT* genes on pJP4 resulted in the apparent structure on pMAB1. However, the presence of the 33-bp nucleotide sequence of unknown origin between the two parts in the *tfdT* gene of pMAB1 excludes the possibility of this simple scheme. The process of formation of the gene clusters on pMAB1 remains to be elucidated and so is the question whether *tfdT* of pMAB1 is a pseudogene or not.

One of the possible advantages of the regulons with LysR-type regulators which share divergent promoter regions with the regulated promoters could be their self-containment (Beck and Warren, 1988). The function of the regulons with LysR-type regulators could be less vulnerable to change of their location during genetic recombination events because of this autonomy; for example, they could be resistant to read-through effect by transcription from neighboring genes. The structures of gene modules for 2,4-D degradation on pMAB1 and pJP4 illustrate this advantage; each of the modules *tfdA-S* and *tfdR-$D_{II}C_{II}E_{II}F_{II}B_{II}$K* could be independently functional.

As to the *tfd(T)-CDEF-B* module either on pMAB1 and pJP4, although the cognate regulatory gene *tfdT* seems to be non-functional for both, the activation of the *tfdC* promoter has been functionally fulfilled by the TfdR(S) on pJP4 (Leveau *et al.*, 1996) and by TfdS on pMAB1 (this study). This was possible because of the recruitment of closely related regulators, which allow cross-activation upon recognition of the pathway intermediate (Leveau *et al.*, 1996; Filer and Harker, 1997; this study).

Recent study with the roles of the *tfd* modules, I and II, in 3-CB degradation process showed TfdF of module I is not sufficient for the efficient degradation of 3-CB by strain JMP134. Instead $TfdF_{II}$ or host-encoded isoenzyme could play the role and this scheme was suggested to be applicable to 2,4-D degradation process as well (Pérez-Pantoja *et al.*, 2000). The *tfdF* gene of pMAB1 is most probably identical to that of pJP4 module I. The ca. 22-kb region of pMAB1 containing the 2,4-D degradative genes was delimited by fragments homologous to parts of Tn*5501* (Lauf *et al.*, 1998) (Fig. 23). It remains to be elucidated if strain CSV90 produces another maleylacetate reductase encoded elsewhere on pMAB1 or on its chromosome.

2. Transcriptional activation of the degradative genes for 3-chlorocatechol and 2,4-dichlorophenoxyacetate

2.1 Transcriptional activation of the *cbnA* promoter by CbnR

Transcriptional activation of the *cbnA* promoter by *cbnR* has been characterized both in vivo and in vitro. In growth complementation studies, *cbnR* was demonstrated to be necessary for growth on 3-CB for strains *R. eutropha* NH9D (3-CB$^-$ derivative of NH9) and *P. putida* PRS4020 (*catR* knockout strain), when they harbor the *cbnABCD* degradative genes. CbnR has also been demonstrated to be a positive regulator by transcriptional reporter and in vitro transcription assays.

Both CCM and 2-CM have been identified as inducer molecules. In reporter analysis, the PRS4020 cells containing a plasmid with truncated *cbnA* gene, pNO50RA', exhibited some elevated transcription in the presence of Ben. Therefore, it remained possible that the *cbnA* promoter could be induced by Ben or catechol. However, neither Ben nor catechol induce transcriptional activation by purified CbnRHis in vitro. The reason for the in vivo activation of the pNO50RA' construct is unclear. It is possible that the activity observed may have been induced by CCM supplied by the host-encoded catechol dioxygenase (CatA). Although *catR*, the activator of the host *catA* gene (Houghton *et al.*, 1995), is disrupted in PRS4020, there may be some low level constitutive expression of *catA* which results in the conversion of Ben to CCM allowing the activation of the *cbnA* promoter. It is also possible that there are additional CbnR binding sites in the *cbnA* gene which add to the complexity of the *cbnA* promoter regulation. Both repressor and activator internal binding sites have been observed with other LysR family members including CatR (Chugani *et al.*, 1998; Cowan *et al.*, 1993). Although the activity observed with construct pNO50RA' in the presence of Ben cannot be explained, the in vitro transcription assays clearly demonstrated that CCM and 2-CM function as inducers for *cbnA* promoter expression.

It is puzzling that CbnR recognizes CCM as well as 2-CM as inducing molecules. The enzymes of the *cbn* operon can catabolize catechol to a certain extent. Chlorocatechol 1,2-dioxygenase (CbnA) can utilize catechol as a substrate (Dorn and Knackmuss, 1978; Reineke, 1998), and chloromuconate cycloisomerase (CbnB) has maintained some activity against nonchlorinated CCM (Reineke, 1998; Schmidt and Knackmuss, 1980). However, 3-oxoadipate enol-lactone, the product from CCM by CbnB, is presumed not to be catabolized further by the enzymes encoded by *cbnCD* (Schlömann, 1994). CatR and ClcR activate their regulated promoters significantly only in the presence of the intermediates of their respective pathway, CCM for CatR

(Parsek *et al.*, 1992) and 2-CM for ClcR (McFall *et al.*, 1997c). The *clcABDE* operon is expressed when cells are grown in benzoate because CatR binds to and activates the *clcA* promoter (McFall *et al.*, 1997b, c; Parsek *et al.*, 1994b). However, CatR cannot bind to or activate the *cbnABCD* promoter (unpublished observations). It is possible that in addition to utilizing chlorocatechols as carbon substrates, the chlorocatechol dioxygenase enzymes act as important detoxifying agents. Catechol could be toxic to pseudomonads as chlorocatechols are (Fritz *et al.*, 1992). Because CatR cannot activate the expression of CbnA to detoxify catechol as it would ClcA, CbnA expression by CbnR upon recognition of CCM could be beneficial for the cell to decompose catechol rapidly.

The *clcABDE* operon was isolated from 3-chlorobenzoate degrading bacteria (Chatterjee *et al.*, 1981) and thus functions to degrade 3-chlorocatechol in the original isolate. The high homology between the *cbnABCD* and the *tcbCDEF* genes suggests that the *cbnABCD* genes may have evolved to degrade more highly chlorinated catechols compared to the *clcABDE* genes. The recognition of more highly substituted catechols or corresponding muconates by the regulators of the *ortho*-cleavage pathway is mostly unknown. It has been demonstrated that 2,4-dichoromuconate acts as an inducer of *tfdCDEF* expression (Filer and Harker, 1997). However, the range of inducing molecules for these operons has not been explored. Since the products from the *tcbCDEF* operon degrade dichlorocatechols and 3,4,6-trichlorocatechol, TcbR (CbnR) probably recognizes corresponding chloromuconates as inducer. Nevertheless, CbnR also has the ability to recognize CCM as an inducer, which is characteristic of CatR.

The specific binding of CbnRHis and CbnR to the *cbnA* promoter has been shown by gel retardation and DNase I protection assays. CbnRHis had the same function as CbnR in growth complementation, reporter assay, and bending assay of *cbnA* promoter with and without inducer. Therefore, functional equality of CbnRHis with CbnR has been verified in vivo and in vitro. Similar observation about a His tag on C-terminal has been reported for NhaR, a LysR type regulater of Na^+/H^+ antiporter gene of *Escherichia coli* (Carmel *et al.*, 1997), suggesting the unaltered function of a C-terminally His-tagged LysR-type protein. On the other hand, the crude soluble protein from *E. coli* BL21(DE3)pLysS containing pEHiscbnR did not show binding activity to *cbnA* promoter (data not shown). Although the N-terminal part of LysR family proteins functions as the DNA binding domain, adding additional amino acids on the N-terminal does not necessarily abolish binding activity. The addition of a polypeptide to the N-terminal of SpvR, a LysR-type regulator of virulence genes of

Salmonella dublin, showed little effect on the specific binding to the regulated promoters and activation of the promoters (Grob and Guiney, 1996). It is not known whether the protein produced from the construct pEHiscbnR does not have the binding activity to *cbnA* promoter or is exclusively retained in the inclusion bodies.

Using established conditions for gel retardation assays (Parsek *et al.*, 1994a; van der Meer *et al.*, 1991b), crude soluble protein from BL21(DE3)pLysS containing either pT7cbnR or pT7cbnRHis always formed aggregates with the *cbnA* promoter fragment in the well of the electrophoresis gel. The addition of heparin in the binding reaction dissolved the aggregate and the protein/DNA complex was able to migrate in the gel. It was presumed that CbnR formed these aggregates because of its highly basic nature [calculated pI=10.3 by Genetyx Software (Software Development Co., Tokyo, Japan)]. Heparin has been used successfully in the study of the formation of splicing complex on mRNAs where it was presumed to quench the nonspecific binding of components in the nuclear extract with the highly negatively charged RNA (Konarska and Sharp, 1986). In our study, it is possible that the negatively charged heparin was able to cancel the electrostatic force among the complexes of CbnR(His)/DNA but not the specific binding between CbnR(His) and DNA fragment.

The length and the location of the footprinting pattern of CbnR to *cbnA* promoter, with hypersensitive sites localized in the center, resemble those of CatR to the *catB* promoter (in the presence of the inducer CCM) (Parsek *et al.*, 1992) and of ClcR to the *clcA* promoter (Coco *et al.*, 1994; McFall *et al.*, 1997c). Thus, it is suggested that the region from nucleotides -76 to -49 is the recognition binding site (RBS) and the region from -44 to -19 is the activation binding site (ABS) (McFall *et al.*, 1998) as shown in Fig. 6c. RBS contains the conserved T-N_{11}-A motif critical for binding by the LysR-type regulators (McFall *et al.*, 1998; Schell, 1993). Further binding of the regulator to ABS helps RNA polymerase bind to the promoter region, thus to activate transcription (McFall *et al.*, 1998). Although both CbnR and CatR recognize CCM as an inducer, the extended footprinting pattern of CbnR in the absence of the inducer, encompassing both RBS and ABS of the *cbnA* promoter, resembles that of ClcR to the *clcA* promoter (Coco *et al.*, 1994; McFall *et al.*, 1997c). This is in contrast to the restricted footprinting pattern of CatR to RBS of the *catB* promoter in the absence of CCM (Parsek *et al.*, 1992). The DNA-binding domain of the LysR-type regulators is believed to be located in the N-terminus and the regions responsible for inducer recognition are believed to be located in the middle of the protein (Schell, 1993). The difference in extent of the footprinting pattern without CCM between *cbn* and *cat*

suggests that the evolution of the DNA-binding domain and the inducer recognition region(s) could be discrete.

The ABS-binding pattern of CbnR without inducer seemed similar to that of ClcR. However, the addition of inducer, CCM or 2CM, to the binding reaction did not cause a change in the footprinting pattern of CbnR to the *cbnA* promoter while addition of 2CM caused a shift in the footprinting pattern of ClcR to the ABS of the *clcA* promoter region (McFall *et al.*, 1997c). It remains to be elucidated if the apparent lack of change in the footprinting pattern of CbnR is the intrinsic characteristic of CbnR-*cbnA* promoter system or due to experimental conditions.

The relaxation of the bending angles of *cbnA* promoter region bound by CbnR(His) upon the addition of inducer was similar to those shown by CatR (Parsek *et al.*, 1995) and ClcR (McFall *et al.*, 1997b). The footprinting patterns and changes in DNA bending upon addition of inducers of the *cat, clc* and *cbn* promoters suggest that the outline of the transcriptional mechanism of these regulatory systems is conserved. This study further indicates independent evolution of the regulatory genes in terms of their DNA binding and recognition of inducer molecules.

2.2 Transcriptional activation of the *tfdC* promoter and the *tfdA* promoter by TfdS in vivo

Transcriptional fusion assays showed that the *tfdC* promoter was not regulated by *tfdT* but that it was activated by the distantly-located *tfdS* in the presence of 2-CM (Fig. 24 and Table 4) which is an analogous compound of 2,4-dichloromuconate, the metabolite from 3,5-dichlorocatechol by TfdC (Fig. 20). TfdS also activated the *tfdA* promoter in the presence of 2-CM (Fig. 25 and Table 5). These results indicated the degradative abilities of the two units, *tfdA* and *tfdCDEFB*, though recruited separately, are interconnected functionally by one regulator, TfdS, to constitute the 2,4-D degradative pathway in CSV90.

In the assay of the activity of the *tfdC* promoter, high basal level was observed for the cells containing the constructs with the full length *tfdC* gene in contrast to the very low basal level exhibited by the cell containing the constructs with only 5'-terminus of the *tfdC* gene. High basal level might be the result of the following two factors. (i) The absence of intact cognate regulatory gene (putative original *tfdT* gene) which should regulate the promoter strictly thus repress the transcription when substrate is not available might allow leaky transcription. (ii) Some structure resulting from the sequence of the *tfdC* (or truncated *tfdT*) gene might make the promoter region more vulnerable to contact by RNA polymerase. This tendency was parallel with that of the

activation of the *cbnA* promoter by CbnR (See the result for pNO50AB' [Table 3]). In any case, the alternative regulator, *tfdS*, could regulate the *tfdC* promoter only in a relaxed manner.

In the presence of 3-CB, the PRS4020 cells containing pCS50SrCD'1 exhibited much higher activity than those containing pCS50SCD'1. In fact, it was the highest activity observed among the constructs for the assay of the *tfdC* promoter activity. In the construct pCS50SrCD'1, the promoter regions of the two genes, *tfdS* and *tfdC*, are located closely but not overlapped. The primary reason for this high activity could be the contribution of the read-through effect of the transcription from the *tfd(S-)A* promoter, as well as the transcription from the *tfdC* promoter, for the expression of the β-galactosidase gene in the construct pCS50SrCD'1. In this respect, it is of interest that the activity exhibited by the cells with either of the constructs, pCS50CTSA' or pCS50CT'SA', in the presence of 3-CB was much lower (Table 5) compared to that by the cells with pCS50SrCD'1. The expression of the *tfdA* promoter was regulated by TfdS strictly (see below). In the constructs, pCS50SCD'1 and pCS50SrCD'1, the –10 portion of the putative promoter sequence for *tfdA* (thus, for *tfdC* in pCS50SrCD'1) was altered from TAGACT to TAGAGC by the cloning procedure using the overlapping *Xba*I site (TCTAGA), while –35 portion of the *tfdA* promoter and the whole putative promoter for *tfdS* was retained. If the short region originally located between the *tfdA* gene and the *Xba*I site (Fig. 24), which was absent from pCS50SrCD'1, was not critical for the strict regulation by TfdS, the *tfdS-A* promoter in the construct pCS50SrCD'1 should be regulated strictly, too, and the activity with this construct in the presence of 3-CB could be sum of those from the constructs pCS50SCD'1 and pCS50CTSA' (or pCS50CT'SA'). The high activity by pCS50SrCD'1 with 3-CB seems to suggest some unknown multiple effect resulting from the activation of both *tfdA* promoter and *tfdC* promoter. Some of the other reasons why divergent promoter region(s) could contribute to the transcriptional activation, even though the promoters are not shared, are as follows (Beck and Warren, 1988); (a) Regulatory proteins can act more efficiently in *cis*, which allows effective control at low concentrations of the regulatory molecules. (b) Multiple promoters in divergent promoter regions might attract RNA polymerase independently, resulting in a higher local concentration of the protein. (c) Binding of RNA polymerase to a promoter deforms DNA thus affect DNA topology. The different activities exerted by 3-CB from pCS50SCD'1 and pCS50SrCD'1 could not likely to be attributed to the point (a) but possibly to (b) and (c). The fact that the activity exhibited by pCS50SCD'1 was closer to those by pCS50STCD' or pCS50ST'CD'could also argue the inappropriateness of applying (a) to explain the

results in this case. If there is *cis* effect of the regulatory protein in the sense that it can work at lower concentration, it does not seem relevant at least within this distance of ca. 3 kb in such a case like *tfdS-tfdC* of pMAB1. Instead, the possible reason could be that the proximal location of the two promoter regions could attract TfdS and RNA polymerase more efficiently and also might have changed DNA topology around this region and made the *tfdC* promoter region more amenable to contacts by TfdS and RNA polymerase.

In contrast to the leaky control of the *tfdC* promoter by *tfdS*, the regulation of the *tfdA* promoter by *tfdS* is rather strict. This is of particular interest considering the possibility that *tfdS* might not be the original cognate regulator for the *tfdA* gene. The structure of the gene clusters on pJP4 seems to indicate *tfdR* belongs to *tfd* chlorocatechol gene module II. It is possible that the duplication event, which gave rise to the *tfdS* gene from the *tfdR* gene, resulted in simultaneous duplication of the (*tfdR*-)$tfdD_{II}$ promoter region for the *tfdA* gene on pJP4 and the subsequent structure of the *tfdS-A* region of pMAB1, provided that the structure of the gene cluster on pMAB1 is the descendant resulted from recombination event(s) happened to the gene clusters of pJP4.

The apparent discrepancy between the substrate of TfdA (i.e. 2,4-D) and the inducers recognized by TfdS(R) (2-CM [this study] and, for the transcriptional activation of the *tfdC* promoter in JMP134, 2,4-dichloromuconate [Filer and Harker, 1997]) seems to endorse the above possibility that *tfdS* might not be the original cognate regulator for the *tfdA* gene. 2,4-Dichloromuconate could also be the inducer for the transcriptional activation of the *tfdA* gene in 2,4-D pathway in strain JMP134, while 2,4-D itself does not induce the expression of the *tfdA* gene (Filer and Harker, 1997). Recent study with TfdR of 2,4-D pathway encoded by plasmid pEST4011 showed that this regulator which shared 90% amino acid identity with TfdR of pJP4 activated the cognate *tfdC* promoter with 2,4-dichloromuconate as an effector molecule (Vedler *et al.*, 2000).

The transcription at the *tfdA* promoter of pMAB1 was weakly activated by TfdS in the presence of Ben (Table 5). Since the cells with pCS50C'TSA' exhibited comparable activity with those by pCS50CTSA' and pCS50CT'SA', the true effector molecule could be either Ben, catechol, or CCM. However, a recent study of the strain JMP134 in vivo strongly suggested CCM could induce the expression of the *tfdA* gene (Radnoti de Lipthay *et al.*, 1999). Thus, it is highly likely that CCM is the inducer also for the *tfdA* gene of pMAB1. The activity shown with pCS50C'TSA' might have been exerted by CCM produced by the constitutive avtivity of host encoded CatA.

Substrate(s) of TfdA in nature except 2,4-D and its analogues like phenoxyacetate and other chlorophenoxyacetates (Fukumori and Hausinger, 1993b) is not known, nor is the origin of the gene. It would be interesting if gene(s) homologous to *tfdA* but with different substrates could be found or if cognate regulator(s) could be found for the *tfdA* gene since these could give a clue to the origin of *tfdA*. In the case of *tfdA* and *tfdS* on pMAB1, this fortuitous combination contributed to the assembly of the efficient 2,4-D degradative pathway.

Although the expression of the *tfd* module II in strain JMP134 has been detected at mRNA level (Leveau *et al.*, 1999), the detail of its regulation is not known. Since both of the *tfdA-S* genes and the tfd(*T*)-*CDEF-B* (module I) genes are nearly identical between pMAB1 and pJP4, further study with the genes of pMAB1 and comparative study with strains CSV90 and JMP134 could cast light upon the overall regulation of the expression of the 2,4-D degradative genes.

2.3 Divergence of the LysR-type transcriptional regulators of the (chloro)benzoate /(chloro)catechol degradative pathway

A phylogenetic tree was constructed with the amino acid sequences of LysR-type transcriptional regulators which activate the promoters of the degradative genes for (chloro)catechols or (chloro)benzoates of Gram negative bacteria and for which information of the inducers is available (Fig. 27). NahR was used as the out-group.

The regulators are apparently divided into two large groups. One group (group I; upper portion of the tree) comprises CbeR(NK8), CatR(RB1), BenM(ADP1) and CatM(ADP1). These regulators activate the expression of catechol *ortho*-clevage genes or benzoate dioxygenase genes. The other group (group II; lower portion except NahR) consisted of the regulators for the expression of the chlorocatechol *ortho*-clevage genes. The chlorocatechol *ortho*-clevage genes are considered to have diverged from catechol *ortho*-clevage genes after the common ancestor had recruited a LysR-type regulator. The amino acid sequences of the regulators have accordingly diverged, which is reflected as the two large clusters as a whole. However, there seem to be distinctive characteristics for each group which may not be explicable only by co-evolution with the cognate degradative genes.

While the regulators, ClcR, TfdS(R), and CbnR, of group II seem to have specialized to recognize chloromuconates or CCM, the other regulators seem to acquire the ability to recognize Ben or 3-CB fortuitously. According to the results of the reporter assay for the *tfdA* promoter, TfdS of pMAB1 might recognize either Ben, catechol, or CCM as well as 2-CM, as inducer. However, since the activation in the

presence of Ben is significantly lower than that exerted with 2-CM and transcription at the *tfdC* promoter by TfdS was activated only with the full length *tfdC* gene and in the presence of 3-CB, it was apparent that 2-CM (or 2,4-dichloromuconate) was the favored inducer for TfdS. Moreover, in a reporter analysis using plasmid constructs with *tfdS-A* region of pMAB1 and *E. coli* DH5α as a host, the activity of the *tfdA* promoter was not exerted either by Ben or catechol (data not shown). ClcR, TfdS and CbnR form a sub-cluster among the members of group II. Considering that CbnR recognize CCM and 2-CM equally well as inducer, ClcR and TfdS might have been more specialized for 2-CM. These three regulators of group II seem to have highly specialized for (chloro)muconates or seem to have lost the flexibility of the structure to resume the recognition of (chloro)benzoate. Since the inducer-recognizing specificity could be altered by subtle change in the amino acid sequence of the regulator (Cebolla *et al.*, 1997), the specialization of these three regulators is rather conspicuous. The three regulators might have suffered some drastic change in their primary structure or in higher-order structure during the evolution.

CbeR of *Burkholderia* sp. NK8 recognize Ben and 3-CB to induce the expression of the cognate catechol genes and the chlorobenzoate dioxygenase genes (Francisco, Ogawa, Suzuki, and Miyashtia, submitted). The recognition of 3-CB by CbeR was necessary for the cell to degrade this compound because of the structure of the gene cluster. It should be noted that the regulators, ClcR, TfdS and CbnR, may not have had a chance to be in such a situation that necessitate them to alter the amino acid seuence to recognize Ben or 3-CB. TfcR of strain NK8 is a regulator for chlorocatechol degradative genes which are similar to *tfd* genes (Liu, Ogawa, and Miyashita, in preparation). Transcriptional fusion study suggested TfcR probably recognize 3-CB and 3-chlorocatechol to activate the cognate promoter thus implying the diversity of the inducer recognition among the group II. It remains to be elucidated if the apparent specialization of CbnR, ClcR and probably TfdS(R) is the result of some irreversible change in their higher level structure or the transient characteristic in the evolution that could be altered by some change in their primary structure. The present study demonstrated the diversity of inducer recognition by the LysR-type regulators involved in the degradation of (chloro)benzoates/(chloro)catechols and the specialization of ClcR, TfdS, and CbnR to recognize (chloro)muconates.

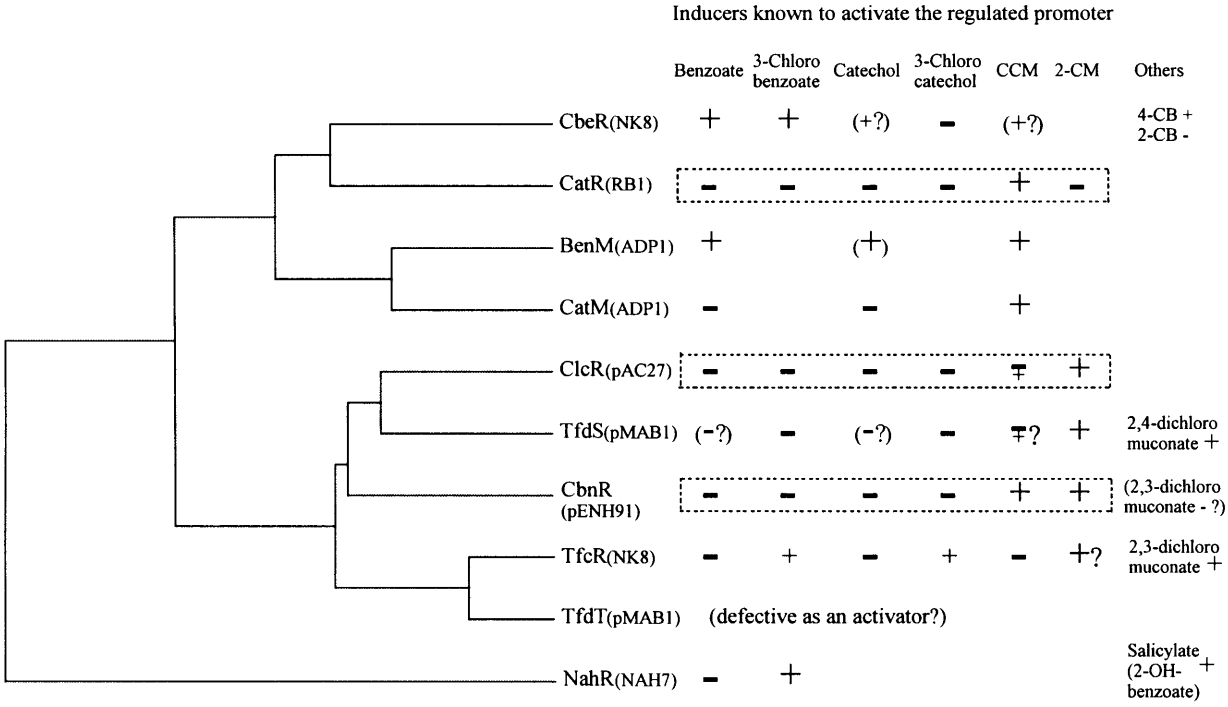


Fig. 27. Phylogenetic tree of LysR-type regulators of (chloro)benzoate/ (chloro) catechol degradative pathway based on the deduced amino acid sequences.

CbeR (NK8); catechol/(chloro)benzoate operon of *Burkholderia* sp. NK8 (Francisco, Ogawa, Suzuki, and Miyashtia, submitted)

CatR (RB1); catechol operon of *Pseudomonas putida* RB1 (Rothmel *et al.*, 1990; M33817)

BenM (ADP1); benzoate operon of *Acinetobacter* sp. ADP1 (Collier *et al.*, 1998; AF009224)

CatM (ADP1); catechol operon of *Acinetobacter* sp. ADP1 (Neidle *et al.*, 1989; AF009224)

ClcR (pAC27); chlorocatechol operon from pAC27 of *Pseudomonas putida* AC866 (Coco *et al.*, 1993; L06464)

TfdS (pMAB1); 2,4-D/chlorocatechol operon from pMAB1 of *Alcaligenes* sp. CSV90 (this study). Identical with TfdS and TfdR of 2,4-D/chlorocatechol operon from pJP4 of *Ralstonia eutropha* JMP134. (Matrubutham and Harker, 1994; M98445, You and Ghosal, 1995; S80112)

CbnR (NH9); chlorocatechol operon from pENH91of *Ralstonia eutropha* NH9 (this study; AB019032). Nearly identical with TcbR of chlorocatechol operon from pP51 of *Pseudomonas* sp. P51 (van der Meer *et al.*, 1991b)

TfcR (NK8); chlorocatechol operon from a large plasmid of *Burkholderia* sp. NK8 (Liu, Ogawa, and Miyashita, in preparation)

TfdT (pMAB1); 2,4-D/chlorocatechol operon from pMAB1 of *Alcaligenes* sp. CSV90 (Bhat *et al.*, 1994; D16356, this study)

NahR (NAH7); naphthalene operon from plasmid NAH7 of *Pseudomonas putida* (Schell and Sukordhaman, 1989; J04233)

Acknowledgements

I would like to thank, firstly, Dr. Kiyotaka Miyashita of the Laboratory of Soil General Microbiology, the National Institute of Agro-Environmental Sciences for guiding and supporting the research from the initial stage and for helpful discussions all through the research years.

My greatest appreciation goes to Prof. Ananda M. Chakrabarty of the University of Illinois at Chicago for introducing me to the study of regulation of the *ortho*-cleavage pathway and molecular biology of Pseudomonads as well as for valuable discussions and suggestions. I am also grateful to the members of his laboratory for helpful discussions and assistance and especially thank Dr. Sally M. McFall (Current affiliation: North Western University) for guiding the research and everything with research and life in U.S. My memory of research life in Chicago will never fade away.

I am deeply indebted to Prof. Teruko Nakazawa and Prof. Masataka Tsuda * of the Department of Microbiology, Yamaguchi University School of Medicine (* Current affiliation: Institute of Genetic Ecology, Tohoku University) for teaching biology and genetics of Psudomonads as well as for providing the strain CSV90 and the plasmids.

Thanks are due to the colleagues of our laboratory of Soil General Microbiology, especially to Dr. Perigio B. Francisco, Jr., Dr. Shenghao Liu, and Dr. Katsuhisa Suzuki for helpful discussions and for sharing their research life and to Toshiko Kajiwara and B.S. Naho Hara (the Department of Bioengineering, Nagaoka University of Technology) for assistance with the experiments.

I also thank Dr. Jan Roelof van der Meer for providing the strain P51 and for the kind and quick responses to my inquiries.

Finally, I would like to thank Prof. Satoshi Matsumoto of the University of Tokyo for encouragement and guiding with this thesis.

References

1. **Adams, M. D., L. M. Wagner, T. J. Graddis, R. Landick, T. K. Antonucci, A. L. Gibson, and D. L. Oxender.** 1990. Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli*. J Biol Chem. **265**(20):11436-43.
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**(2-3):185-95.
3. **Alexander, M.** 1981. Biodegradation of chemicals of environmental concern. Science. **211**(4478):132-8.
4. **Amy, P. S., J. W. Schulke, L. M. Frazier, and R. J. Seidler.** 1985. Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl Environ Microbiol. **49**(5):1237-45.
5. **Andersen, K., and J. Caton.** 1987. Sequence analysis of the *Alcaligenes eutrophus* chromosomally encoded ribulose bisphosphate carboxylase large and small subunit genes and their gene products. J Bacteriol. **169**(10):4547-58.
6. **Bagdasarian, M., R. Lurz, B. Rückert, F. C. H. Franklin, M. M. Bagdasarian, J. Frey, and K. N. Timmis.** 1981. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene. **16**(1-3):237-47.
7. **Bandyopadhyay, S., S. K. Bhattacharya, and P. Majumdar.** 1994. Engineering aspects of bioremediation. p. 55-75. *In* D. L. Wise and D. J. Trantolo (ed.), Remediation of hazardous waste contaminated soils. Mercel Dekker, Inc, New York.
8. **Beck, C. F., and R. A. J. Warren.** 1988. Divergent promoters, a common form of gene organization. Microbiol Rev. **52**(3):318-26.
9. **Bhat, M. A., T. Ishida, K. Horiike, C. S. Vaidyanathan, and M. Nozaki.** 1993. Purification of 3,5-dichlorocatechol 1,2-dioxygenase, a nonheme iron dioxygenase and a key enzyme in the biodegradation of a herbicide, 2,4-dichlorophenoxyacetic acid (2,4-D), from *Pseudomonas cepacia* CSV90. Arch Biochem Biophys. **300**(2):738-46.
10. **Bhat, M. A., M. Tsuda, K. Horiike, M. Nozaki, C. S. Vaidyanathan, and T. Nakazawa.** 1994. Identification and characterization of a new plasmid carrying genes for degradation of 2,4-dichlorophenoxyacetate from *Pseudomonas cepacia* CSV90. Appl Environ Microbiol. **60**(1):307-12.
11. **Brown, H. J., H. W. Stokes, and R. M. Hall.** 1996. The integrons In0, In2, and In5 are defective transposon derivatives. J Bacteriol. **178**(15):4429-37.

12. **Bullock, W. O., J. M. Fernandez, and J. M. Short.** 1987. XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. BioTechniques. **5**:376-9.

13. **Byrne, A. M., and T. G. Lessie.** 1994. Characteristics of IS*401*, a new member of the IS*3* family implicated in plasmid rearrangements in *Pseudomonas cepacia*. Plasmid. **31**(2):138-47.

14. **Carmel, O., O. Rahav-Manor, N. Dover, B. Shaanan, and E. Padan.** 1997. The Na^+-specific interaction between the LysR-type regulator, NhaR, and the *nhaA* gene encoding the Na^+/H^+ antiporter of *Escherichia coli*. Embo J. **16**(19):5922-9.

15. **Casse, F., C. Boucher, J. S. Julliot, M. Michel, and J. Denarie.** 1979. Identification and characterization of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis. J Gen Microbiol. **113**:229-42.

16. **Cebolla, A., C. Sousa, and V. de Lorenzo.** 1997. Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading *Pseudomonas* define protein sites involved in binding of aromatic inducers. J Biol Chem. **272**(7):3986-92.

17. **Chatterjee, D. K., and A. M. Chakrabarty.** 1983. Genetic homology between independently isolated chlorobenzoate- degradative plasmids. J Bacteriol. **153**(1):532-4.

18. **Chatterjee, D. K., and A. M. Chakrabarty.** 1982. Genetic rearrangements in plasmids specifying total degradation of chlorinated benzoic acids. Mol Gen Genet. **188**(2):279-85.

19. **Chatterjee, D. K., and A. M. Chakrabarty.** 1984. Restriction mapping of a chlorobenzoate degradative plasmid and molecular cloning of the degradative genes. Gene. **27**(2):173-81.

20. **Chatterjee, D. K., S. T. Kellogg, S. Hamada, and A. M. Chakrabarty.** 1981. Plasmid specifying total degradation of 3-chlorobenzoate by a modified *ortho* pathway. J Bacteriol. **146**(2):639-46.

21. **Chaudhry, G. R., and G. H. Huang.** 1988. Isolation and characterization of a new plasmid from a *Flavobacterium* sp. which carries the genes for degradation of 2,4-dichlorophenoxyacetate. J Bacteriol. **170**(9):3897-902.

22. **Chugani, S. A., M. R. Parsek, and A. M. Chakrabarty.** 1998. Transcriptional repression mediated by LysR-type regulator CatR bound at multiple binding sites. J Bacteriol. **180**(9):2367-72.

23. **Chugani, S. A., M. R. Parsek, C. D. Hershberger, K. Murakami, A. Ishihama, and A. M. Chakrabarty.** 1997. Activation of the *catBCA* promoter: probing the interaction of CatR and RNA polymerase through in vitro transcription. J Bacteriol. **179**(7):2221-7.

24. **Coco, W. M., M. R. Parsek, and A. M. Chakrabarty.** 1994. Purification of the LysR family regulator, ClcR, and its interaction with the *Pseudomonas putida clcABD* chlorocatechol operon promoter. J Bacteriol. **176**(17):5530-3.

25. **Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty.** 1993. Nucleotide

sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. J Bacteriol. **175**(2):417-27.

26. **Collier, L. S., G. L. Gaines, III, and E. L. Neidle.** 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J Bacteriol. **180**(9):2493-501.

27. **Cowan, J. M., M. L. Urbanowski, M. Talmi, and G. V. Stauffer.** 1993. Regulation of the *Salmonella typhimurium metF* gene by the MetR protein. J Bacteriol. **175**(18):5862-6.

28. **Darzins, A., and A. M. Chakrabarty.** 1984. Cloning of genes controlling alginate biosynthesis from a mucoid cystic fibrosis isolate of *Pseudomonas aeruginosa*. J Bacteriol. **159**(1):9-18.

29. **Daubaras, D., and A. M. Chakrabarty.** 1992. The environment, micorobes and bioremediation: microbial activities modulated by the environment. Biodegradation. **3**:125-135.

30. **Daubaras, D. L., C. D. Hershberger, K. Kitano, and A. M. Chakrabarty.** 1995. Sequence analysis of a gene cluster involved in metabolism of 2,4,5-trichlorophenoxyacetic acid by *Burkholderia cepacia* AC1100. Appl Environ Microbiol. **61**(4):1279-89.

31. **Davis, D. H., M. Doudoroff, and R. Y. Stanier.** 1969. Proposal to reject the genus *Hydrogenomonas*: taxonomic implications. Int J Syst Bacteriol. **19**:375-90.

32. **Di Gioia, D., M. Peel, F. Fava, and R. C. Wyndham.** 1998. Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America. Appl Environ Microbiol. **64**(5):1940-6.

33. **Diaz-Aroca, E., M. V. Mendiola, J. C. Zabala, and F. de la Cruz.** 1987. Transposition of IS*91* does not generate a target duplication. J Bacteriol. **169**(1):442-3.

34. **Don, R. H., and J. M. Pemberton.** 1985. Genetic and physical map of the 2,4-dichlorophenoxyacetic acid- degradative plasmid pJP4. J Bacteriol. **161**(1):466-8.

35. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J Bacteriol. **145**(2):681-6.

36. **Don, R. H., A. J. Weightman, H. -J. Knackmuss, and K. N. Timmis.** 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J Bacteriol. **161**(1):85-90.

37. **Dorn, E., and H. -J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Substituent effects on 1,2-dioxygenation of catechol. Biochem J. **174**(1):85-94.

38. **Drucker, M. P.** 1994. Ethical responsibilities of environmental engineers. p. 1-7. *In* D. L. Wise

and D. J. Trantolo (ed.), Remediation of hazardous waste contaminated soils. Marcel Dekker, Inc., New York.

39. **Erickson, J. W., and C. A. Gross.** 1989. Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternate σ factor involved in high-temperature gene expression. Genes Dev. **3**(9):1462-71.

40. **Eulberg, D., L. A. Golovleva, and M. Schlömann.** 1997. Characterization of catechol catabolic genes from *Rhodococcus erythropolis* 1CP. J Bacteriol. **179**(2):370-81.

41. **Eulberg, D., E. M. Kourbatova, L. A. Golovleva, and M. Schlömann.** 1998. Evolutionary relationship between chlorocatechol catabolic enzymes from *Rhodococcus opacus* 1CP and their counterparts in proteobacteria: sequence divergence and functional convergence. J Bacteriol. **180**(5):1082-94.

42. **Farinha, M. A., and A. M. Kropinski.** 1990. Construction of broad-host-range plasmid vectors for easy visible selection and analysis of promoters. J Bacteriol. **172**(6):3496-9.

43. **Filer, K., and A. R. Harker.** 1997. Identification of the inducing agent of the 2,4-dichlorophenoxyacetic acid pathway encoded by plasmid pJP4. Appl Environ Microbiol. **63**:317-20.

44. **Fletcher, R. D.** 1994. Practical considerations during bioremediation. p. 39-53. *In* D. L. Wise and D. J. Trantolo (ed.), Remediation of hazardous waste contaminated soils. Mercel Dekker, Inc., New York.

45. **Franklin, F. C. H.** 1985. Broad host range cloning vectors for gram negative bacteria, p. 165-84. *In* D. M. Glover (ed.), DNA cloning vol. I. IRL Press, Oxford.

46. **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 U S A. **84**(13):4460-4.

47. **Frantz, B., K. -L. Ngai, D. K. Chatterjee, L. N. Ornston, and A. M. Chakrabarty.** 1987. Nucleotide sequence and expression of *clcD*, a plasmid-borne dienelactone hydrolase gene from *Pseudomonas* sp. strain B13. J Bacteriol. **169**(2):704-9.

48. **Fritz, H., W. Reineke, and E. Schmidt.** 1992. Toxicity of chlorobenzene on *Pseudomonas* sp. strain RHO1, a chlorobenzene-degrading strain. Biodegradation. **2**:165-70.

49. **Fukumori, F., and R. P. Hausinger.** 1993. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J Bacteriol. **175**(7):2083-6.

50. **Fukumori, F., and R. P. Hausinger.** 1993. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J Biol Chem. **268**(32):24311-7.

51. **Fulthorpe, R. R., C. McGowan, O. V. Maltseva, W. E. Holben, and J. M. Tiedje.** 1995. 2,4-Dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl

Environ Microbiol. **61**(9):3274-81.

52. **Fulthorpe, R. R., A. N. Rhodes, and J. M. Tiedje.** 1998. High levels of endemicity of 3-chlorobenzoate-degrading soil bacteria. Appl Environ Microbiol. **64**(5):1620-7.

53. **Ghosal, D., and I.-S. You.** 1988. Gene duplication in haloaromatic degradative plasmids pJP4 and pJP2. Can J Microbiol. **34**(6):709-15.

54. **Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty.** 1985. Genes specifying degradation of 3-chlorobenzoic acid in plasmids pAC27 and pJP4. Proc Natl Acad Sci U S A. **82**(6):1638-42.

55. **Goff, S. P.** 1992. Genetics of retroviral integration. Annu Rev Genet. **26**:527-44.

56. **Gordon, S. V., B. Heym, J. Parkhill, B. Barrell, and S. T. Cole.** 1999. New insertion sequences and a novel repeated sequence in the genome of *Mycobacterium tuberculosis* H37Rv. Microbiology. **145**(Pt 4):881-92.

57. **Grob, P., and D. G. Guiney.** 1996. In vitro binding of the *Salmonella dublin* virulence plasmid regulatory protein SpvR to the promoter regions of *spvA* and *spvR*. J Bacteriol. **178**(7):1813-20.

58. **Haack, K. R., and J. R. Roth.** 1995. Recombination between chromosomal IS*200* elements supports frequent duplication formation in *Salmonella typhimurium*. Genetics. **141**(4):1245-52.

59. **Haas, D., B. Berger, S. Schmid, T. Seitz, and C. Reimmann.** 1996. Insertion sequence IS*21*: related insertion sequence elements, transpositional mechanisms, and application to linker insertion mutagenesis, p. 238-49. *In* T. Nakazawa, K. Furukawa, D. Haas, and S. Silver (ed.), Molecular biology of pseudomonads. American Society for Microbiology, Washington, D.C.

60. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109-35. *In* D. M. Glover (ed.), DNA cloning vol. I. IRL Press, Oxford.

61. **Haren, L., B. Ton-Hoang, and M. Chandler.** 1999. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol. **53**:245-81.

62. **Harwood, C. S., and R. E. Parales.** 1996. The β-ketoadipate pathway and the biology of self-identity. Annu Rev Microbiol. **50**:553-90.

63. **Haugland, R. A., U. M. X. Sangodkar, and A. M. Chakrabarty.** 1990. Repeated sequences including RS1100 from *Pseudomonas cepacia* AC1100 function as IS elements. Mol Gen Genet. **220**(2):222-8.

64. **Hein, S., and A. Steinbüchel.** 1994. Biochemical and molecular characterization of the *Alcaligenes eutrophus* pyruvate dehydrogenase complex and identification of a new type of dihydrolipoamide dehydrogenase. J Bacteriol. **176**(14):4394-408.

65. **Hershberger, C. D., R. W. Ye, M. R. Parsek, Z. -D. Xie, and A. M. Chakrabarty.** 1995. The *algT*(*algU*) gene of *Pseudomonas aeruginosa*, a key regulator involved in alginate biosynthesis, encodes an alternative σ factor (σ^E). Proc Natl Acad Sci U S A. **92**(17):7941-5.

66. **Hesslinger, C., and G. Sawers.** 1998. The tdcE gene in Escherichia coli strain W3110 is separated from the rest of the tdc operon by insertion of IS5 elements. DNA Seq. **9**(3):183-8.

67. **Hindmarsh, P., and J. Leis.** 1999. Retroviral DNA integration. Microbiol Mol Biol Rev. **63**(4):836-43.

68. **Hoshino, T., and K. Kose.** 1990. Cloning, nucleotide sequences, and identification of products of the *Pseudomonas aeruginosa* PAO *bra* genes, which encode the high-affinity branched-chain amino acid transport system. J Bacteriol. **172**(10):5531-9.

69. **Houghton, J. E., T. M. Brown, A. J. Appel, E. J. Hughes, and L. N. Ornston.** 1995. Discontinuities in the evolution of *Pseudomonas putida cat* genes. J Bacteriol. **177**(2):401-12.

70. **Isawa, T., R. Sameshima, H. Mitsui, and K. Minamisawa.** 1999. IS*1631* occurrence in *Bradyrhizobium japonicum* highly reiterated sequence-possessing strains with high copy numbers of repeated sequences RSα and RSβ. Appl Environ Microbiol. **65**(8):3493-501.

71. **Johnson, B. F., and R. Y. Stanier.** 1971. Dissimilation of aromatic compounds by *Alcaligenes eutrophus*. J Bacteriol. **107**(2):468-75.

72. **Joset, F., and J. Guespin-Michel.** 1993. Transposable elements, p. 132-64, Prokaryotic genetics. Blackwell Scientific Publications, Oxford.

73. **Ka, J. O., W. E. Holben, and J. M. Tiedje.** 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. Appl Environ Microbiol. **60**(4):1106-15.

74. **Kasberg, T., D. L. Daubaras, A. M. Chakrabarty, D. Kinzelt, and W. Reineke.** 1995. Evidence that operons *tcb*, *tfd*, and *clc* encode maleylacetate reductase, the fourth enzyme of the modified *ortho* pathway. J Bacteriol. **177**(13):3885-9.

75. **Kasberg, T., V. Seibert, M. Schlömann, and W. Reineke.** 1997. Cloning, characterization, and sequence analysis of the *clcE* gene encoding the maleylacetate reductase of *Pseudomonas* sp. strain B13. J Bacteriol. **179**(11):3801-3.

76. **Katz, R. A., and A. M. Skalka.** 1994. The retroviral enzymes. Annu Rev Biochem. **63**:133-73.

77. **Kieser, T.** 1984. Factors affecting the isolation of CCC DNA from *Streptomyces lividans* and *Escherichia coli*. Plasmid. **12**(1):19-36.

78. **Konarska, M. M., and P. A. Sharp.** 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell. **46**(6):845-55.

79. **Lauf, U., C. Müller, and H. Herrmann.** 1998. The transposable elements resident on the plasmids of *Pseudomonas putida* strain H, Tn*5501* and Tn*5502*, are cryptic transposons of the Tn3 family. Mol Gen Genet. **259**(6):674-8.

80. **Leander, M., T. Vallaeys, and R. Fulthorpe.** 1998. Amplification of putative chlorocatechol dioxygenase gene fragments from α- and β-*Proteobacteria*. Can J Microbiol. **44**(5):482-6.

81. **Leveau, J. H. J., W. M. de Vos, and J. R. van der Meer.** 1994. Analysis of the binding site of the LysR-type transcriptional activator TcbR on the *tcbR* and *tcbC* divergent promoter sequences. J Bacteriol. **176**(7):1850-6.

82. **Leveau, J. H. J., F. Konig, H. Fuchslin, C. Werlen, and J. R. van der Meer.** 1999. Dynamics of multigene expression during catabolic adaptation of *Ralstonia eutropha* JMP134 (pJP4) to the herbicide 2, 4- dichlorophenoxyacetate. Mol Microbiol. **33**(2):396-406.

83. **Leveau, J. H. J., and J. R. van der Meer.** 1997. Genetic characterization of insertion sequence ISJP4 on plasmid pJP4 from *Ralstonia eutropha* JMP134. Gene. **202**(1-2):103-14.

84. **Leveau, J. H. J., and J. R. van der Meer.** 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4) [published erratum appears in J Bacteriol 1997 Mar;179(6):2096]. J Bacteriol. **178**(23):6824-32.

85. **Leveau, J. H. J., A. J. B. Zehnder, and J. R. van der Meer.** 1998. The *tfdK* gene product facilitates uptake of 2,4-dichlorophenoxyacetate by *Ralstonia eutropha* JMP134(pJP4). J Bacteriol. **180**(8):2237-43.

86. **Mäe, A. A., R. O. Marits, N. R. Ausmees, V. M. Kôiv, and A. L. Heinaru.** 1993. Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J Gen Microbiol. **139**:3165-3170.

87. **Magdalena, J., C. Gérard, B. Joris, M. Forsman, and J. Dusart.** 1997. The two β-lactamase genes of *Streptomyces cacaoi*, *blaL* and *blaU*, are under the control of the same regulatory system. Mol Gen Genet. **255**(2):187-93.

88. **Mandel, M.** 1966. Deoxyribonucleic acid base composition in the genus *Pseudomonas*. J Gen Microbiol. **43**:273-92.

89. **Matheson, V. G., L. J. Forney, Y. Suwa, C. H. Nakatsu, A. J. Sexstone, and W. E. Holben.** 1996. Evidence for acquisition in nature of a chromosomal 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase gene by different *Burkholderia* spp. Appl Environ Microbiol. **62**:2457-63.

90. **Matrubutham, U., and A. R. Harker.** 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J Bacteriol. **176**(8):2348-53.

91. **McFall, S. M., B. Abraham, C. G. Narsolis, and A. M. Chakrabarty.** 1997a. A tricarboxylic acid cycle intermediate regulating transcription of a chloroaromatic biodegradative pathway: fumarate-mediated repression of the *clcABD* operon. J Bacteriol. **179**(21):6729-35.

92. **McFall, S. M., S. A. Chugani, and A. M. Chakrabarty.** 1998. Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme. Gene. **223**(1-2):257-67.

93. **McFall, S. M., T. J. Klem, N. Fujita, A. Ishihama, and A. M. Chakrabarty.** 1997b. DNase I

footprinting, DNA bending and *in vitro* transcription analyses of ClcR and CatR interactions with the *clcABD* promoter: evidence of a conserved transcriptional activation mechanism. Mol Microbiol. **24**(5):965-76.

94. **McFall, S. M., M. R. Parsek, and A. M. Chakrabarty.** 1997c. 2-Chloromuconate and ClcR-mediated activation of the *clcABD* operon: in vitro transcriptional and DNase I footprint analyses. J Bacteriol. **179**(11):3655-63.

95. **Menou, G., J. Mahillon, M. M. Lecadet, and D. Lereclus.** 1990. Structural and genetic organization of IS*232*, a new insertion sequence of *Bacillus thuringiensis*. J Bacteriol. **172**(12):6689-96.

96. **Merlin, C., D. Springael, and A. Toussaint.** 1999. Tn*4371*: A modular structure encoding a phage-like integrase, a *Pseudomonas*-like catabolic pathway, and RP4/Ti-like transfer functions. Plasmid. **41**(1):40-54.

97. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, New York.

98. **Nagy, I., G. Schoofs, J. Vanderleyden, and R. De Mot.** 1997. Transposition of the IS*21*-related element IS*1415* in *Rhodococcus erythropolis*. J Bacteriol. **179**(14):4635-8.

99. **Nakai, C., H. Uyeyama, H. Kagamiyama, T. Nakazawa, S. Inouye, F. Kishi, A. Nakazawa, and M. Nozaki.** 1995. Cloning, DNA sequencing, and amino acid sequencing of catechol 1,2-dioxygenases (pyrocatechase) from *Pseudomonas putida* mt-2 and *Pseudomonas arvilla* C-1. Arch Biochem Biophys. **321**(2):353-62.

100. **Nakatsu, C., J. Ng, R. Singh, N. Straus, and C. Wyndham.** 1991. Chlorobenzoate catabolic transposon Tn*5271* is a composite class I element with flanking class II insertion sequences. Proc Natl Acad Sci U S A. **88**(19):8312-6.

101. **Nakatsu, C. H., and R. C. Wyndham.** 1993. Cloning and expression of the transposable chlorobenzoate-3,4-dioxygenase genes of *Alcaligenes* sp. strain BR60. Appl Environ Microbiol. **59**(11):3625-33.

102. **Neidle, E. L., C. Hartnett, and L. N. Ornston.** 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J Bacteriol. **171**(10):5410-21.

103. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene. **26**(1):101-6.

104. **Parales, R. E., and C. S. Harwood.** 1993. Regulation of the *pcaIJ* genes for aromatic acid degradation in *Pseudomonas putida*. J Bacteriol. **175**(18):5829-38.

105. **Parsek, M. R., W. M. Coco, and A. M. Chakrabarty.** 1994a. Gel-shift assay and DNase I footprinting in analysis of transcriptional regulation of biodegradative genes. Methods Mol Genet. **3**:273-90.

106. **Parsek, M. R., M. Kivisaar, and A. M. Chakrabarty.** 1995. Differential DNA bending introduced by the *Pseudomonas putida* LysR-type regulator, CatR, at the plasmid-borne *pheBA* and chromosomal *catBC* promoters. Mol Microbiol. **15**(5):819-28.

107. **Parsek, M. R., S. M. McFall, D. L. Shinabarger, and A. M. Chakrabarty.** 1994b. Interaction of two LysR-type regulatory proteins CatR and ClcR with heterologous promoters: functional and evolutionary implications. Proc Natl Acad Sci U S A. **91**(26):12393-7.

108. **Parsek, M. R., D. L. Shinabarger, R. K. Rothmel, and A. M. Chakrabarty.** 1992. Roles of CatR and *cis,cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. J Bacteriol. **174**(23):7798-806.

109. **Parsek, M. R., R. W. Ye, P. Pun, and A. M. Chakrabarty.** 1994c. Critical nucleotides in the interaction of a LysR-type regulator with its target promoter region. *catBC* promoter activation by CatR. J Biol Chem. **269**(15):11279-84.

110. **Pérez-Pantoja, D., L. Guzmán, M. Manzano, D. H. Pieper, and B. González.** 2000. Role of $tfdC_ID_IE_IF_I$ and $tfdD_{II}C_{II}E_{II}F_{II}$ gene modules in catabolism of 3-chlorobenzoate by *Ralstonia eutropha* JMP134(pJP4). Appl Environ Microbiol. **66**(4):1602-8.

111. **Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin.** 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J Bacteriol. **172**(5):2351-9.

112. **Podladchikova, O. N., G. G. Dikhanov, A. V. Rakin, and J. Heesemann.** 1994. Nucleotide sequence and structural organization of *Yersinia pestis* insertion sequence IS*100*. FEMS Microbiol Lett. **121**(3):269-74.

113. **Priefert, H., S. Hein, N. Krüger, K. Zeh, B. Schmidt, and A. Steinbüchel.** 1991. Identification and molecular characterization of the *Alcaligenes eutrophus* H16 *aco* operon genes involved in acetoin catabolism. J Bacteriol. **173**(13):4056-71.

114. **Radnoti de Lipthay, J., T. Barkay, J. Vekova, and S. J. Sørensen.** 1999. Utilization of phenoxyacetic acid, by strains using either the *ortho* or *meta* cleavage of catechol during phenol degradation, after conjugal transfer of *tfdA*, the gene encoding a 2,4-dichlorophenoxyacetic acid / 2-oxoglutarate dioxygenase. Appl Microbiol Biotechnol. **51**(2):207-14.

115. **Rangnekar, V. M.** 1988. Variation in the ability of *Pseudomonas* sp. strain B13 cultures to utilize *meta*-chlorobenzoate is associated with tandem amplification and deamplification of DNA. J Bacteriol. **170**(4):1907-12.

116. **Ravatn, R., S. Studer, D. Springael, A. J. B. Zehnder, and J. R. van der Meer.** 1998a. Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. strain B13. J Bacteriol. **180**(17):4360-9.

117. **Ravatn, R., S. Studer, A. J. B. Zehnder, and J. R. van der Meer.** 1998b. Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase *clc* element of *Pseudomonas* sp. strain B13. J Bacteriol. **180**(21):5505-14.

118. **Reimmann, C., R. Moore, S. Little, A. Savioz, N. S. Willetts, and D. Haas.** 1989. Genetic structure, function and regulation of the transposable element IS*21*. Mol Gen Genet. **215**(3):416-24.

119. **Reineke, W.** 1998. Development of hybrid strains for the mineralization of chloroaromatics by patchwork assembly. Annu Rev Microbiol. **52**:287-331.

120. **Rogers, M. B., T. K. Bennett, C. M. Payne, and C. J. Smith.** 1994. Insertional activation of *cepA* leads to high-level β-lactamase expression in *Bacteroides fragilis* clinical isolates. J Bacteriol. **176**(14):4376-84.

121. **Rothmel, R. K., T. L. Aldrich, J. E. Houghton, W. M. Coco, L. N. Ornston, and A. M. Chakrabarty.** 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the LysR family. J Bacteriol. **172**(2):922-31.

122. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

123. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. **74**(12):5463-7.

124. **Savard, P., L. Peloquin, and M. Sylvestre.** 1986. Cloning of *Pseudomonas* sp. strain CBS3 genes specifying dehalogenation of 4-chlorobenzoate. J Bacteriol. **168**(1):81-5.

125. **Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47**:597-626.

126. **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**(4):1952-9.

127. **Schlömann, M.** 1994. Evolution of chlorocatechol catabolic pathways. Conclusions to be drawn from comparisons of lactone hydrolases. Biodegradation. **5**(3-4):301-20.

128. **Schmid, S., B. Berger, and D. Haas.** 1999. Target joining of duplicated insertion sequence IS*21* is assisted by IstB protein in vitro. J Bacteriol. **181**(7):2286-9.

129. **Schmid, S., T. Seitz, and D. Haas.** 1998. Cointegrase, a naturally occurring, truncated form of IS*21* transposase, catalyzes replicon fusion rather than simple insertion of IS*21*. J Mol Biol. **282**(3):571-83.

130. **Schmidt, E., and H. -J. Knackmuss.** 1980. Chemical structure and biodegradability of halogenated aromatic compounds. Conversion of chlorinated muconic acids into maleylacetic

acid. Biochem J. **192**(1):339-47.

131. **Schmidt, E. L., M. J. Zidwick, and H. M. Abebe.** 1986. *Bradyrhizobium japonicum* serocluster 123 and diversity among member isolates. Appl Environ Microbiol. **51:**1212-5.

132. **Seibert, V., K. Stadler-Fritzsche, and M. Schlömann.** 1993. Purification and characterization of maleylacetate reductase from *Alcaligenes eutrophus* JMP134(pJP4). J Bacteriol. **175**(21):6745-54.

133. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. Bio/Technology. **1:**784-91.

134. **Smith, M., J. Jessee, T. Landers, and J. Jordan.** 1990. High efficiency bacterial electroporation: $1x10^{10}$ *E. coli* transformants/μg. Focus. **12:**38-40.

135. **Solinas, F., A. M. Marconi, M. Ruzzi, and E. Zennaro.** 1995. Characterization and sequence of a novel insertion sequence, IS*1162*, from *Pseudomonas fluorescens*. Gene. **155**(1):77-82.

136. **Springael, D., S. Kreps, and M. Mergeay.** 1993. Identification of a catabolic transposon, Tn*4371*, carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. J Bacteriol. **175**(6):1674-81.

137. **Suwa, Y., A. D. Wright, F. Fukimori, K. A. Nummy, R. P. Hausinger, W. E. Holben, and L. J. Forney.** 1996. Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase from *Burkholderia* sp. strain RASC. Appl Environ Microbiol. **62**(7):2464-9.

138. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci U S A. **82**(4):1074-8.

139. **Top, E. M., W. E. Holben, and L. J. Forney.** 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl Environ Microbiol. **61**(5):1691-8.

140. **Tsuda, M.** 1996. Catabolic transposons in pseudomonads, p. 219-28. *In* T. Nakazawa, K. Furukawa, D. Haas, and S. Silver (ed.), Molecular biology of pseudomonads. American Society for Microbiology, Washington, D. C.

141. **Tsuda, M., H. M. Tan, A. Nishi, and K. Furukawa.** 1999. Mobile catabolic genes in bacteria. J Biosci Bioeng. **87**(4):401-10.

142. **van der Meer, J. R.** 1997. Evolution of novel metabolic pathway for the degradation of chloroaromatic compounds. Antonie Leeuwenhoek. **71:**159-78.

143. **van der Meer, J. R., W. M. de Vos, S. Harayama, and A. J. B. Zehnder.** 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol Rev. **56**(4):677-94.

144. **van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991a. Sequence

analysis of the Pseudomonas sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. J Bacteriol. **173**(8):2425-34.

145. **van der Meer, J. R., A. C. J. Frijters, J. H. J. Leveau, R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991b. Characterization of the Pseudomonas sp. strain P51 gene *tcbR*, a LysR-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon, and analysis of the regulatory region. J Bacteriol. **173**(12):3700-8.

146. **van der Meer, J. R., A. R. W. van Neerven, E. J. de Vries, W. M. de Vos, and A. J. B. Zehnder.** 1991c. Cloning and characterization of plasmid-encoded genes for the degradation of 1,2-dichloro-, 1,4-dichloro-, and 1,2,4-trichlorobenzene of *Pseudomonas* sp. strain P51. J Bacteriol. **173**(1):6-15.

147. **van der Meer, J. R., A. J. B. Zehnder, and W. M. de Vos.** 1991d. Identification of a novel composite transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. J Bacteriol. **173**(22):7077-83.

148. **Vedler, E., V. Kõiv, and A. Heinaru.** 2000. TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011. Gene. **245**(1):161-8.

149. **Werlen, C., H. -P. E. Kohler, and J. R. van der Meer.** 1996. The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. J Biol Chem. **271**(8):4009-16.

150. **Wyndham, R. C., A. E. Cashore, C. H. Nakatsu, and M. C. Peel.** 1994. Catabolic transposons. Biodegradation. **5**(3-4):323-42.

151. **Wyndham, R. C., R. K. Singh, and N. A. Straus.** 1988. Catabolic instability, plasmid gene deletion and recombination in *Alcaligenes* sp. BR60. Arch Microbiol. **150**(3):237-43.

152. **Xu, K., Z. Q. He, Y. M. Mao, R. Q. Sheng, and Z. J. Sheng.** 1993. On two transposable elements from *Bacillus stearothermophilus*. Plasmid. **29**(1):1-9.

153. **Xu, Y., M. W. Mortimer, T. S. Fisher, M. L. Kahn, F. J. Brockman, and L. Xun.** 1997. Cloning, sequencing, and analysis of a gene cluster from *Chelatobacter heintzii* ATCC 29600 encoding nitrilotriacetate monooxygenase and NADH:flavin mononucleotide oxidoreductase. J Bacteriol. **179**(4):1112-6.

154. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene. **33**(1):103-19.

155. **Yeo, C. C., and C. L. Poh.** 1997. Characterization of IS*1474*, an insertion sequence of the IS*21* family isolated from *Pseudomonas alcaligenes* NCIB 9867. FEMS Microbiol Lett.

149(2):257-63.

156. **You, I.-S., and D. Ghosal.** 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetate catabolic plasmid pJP4. Mol Microbiol. **16**(2):321-31.

Summary

Ralstonia eutropha NH9, isolated in Japan, degrades 3-chlorobenzoate (3-CB) via the modified *ortho*-cleavage pathway. A ca. 5.7-kb six gene cluster responsible for chlorocatechol degradation was cloned from the transmissible plasmid pENH91 of NH9: the *cbnABCD* operon encoding the degradative enzymes (including *orfX* of unknown function) and the divergently transcribed *cbnR* encoding the LysR-type transcriptional regulator of the *cbn* operon.

The *cbnR-ABCD* genes were found to be located between two directly-oriented identical insertion sequences (ISs) of 2,520 bp, designated IS*1600*, thus forming a composite transposon designated Tn*5707* (ca. 15 kb).

The *cbnR-ABCD* genes showed the highest homology to the *tcbR-CDEF* genes on plasmid pP51 of the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. P51, which was isolated in the Netherlands (89% to 100% identity at the nucleotide level). The structure of the operon, including the lengths of open reading frames and intervening sequences, was completely conserved between the *cbn* and *tcb* genes. Most nucleotide substitutions were localized within and proximal to the *cbnB* (*tcbD*) gene. The difference in utilizable chloroaromatics as growth substrates between the two strains seemed to be due to differences in enzymes that convert substrates to chlorocatechols.

Although the *tcbR-CDEF* genes were not associated with similar ISs, a DNA fragment homologous to IS*1600* was cloned from the chromosome of strain P51. The sequence of the fragment suggested that it might be a remnant of an IS. The two sequences, together with IS*1326* and *nmoT*, formed a distinct cluster on a phylogenetic tree of the IS*21* family. The diversity of the sources of these IS or IS-like elements suggests the prevalence of ISs of this type among related bacteria.

In one of the 3-CB$^-$ segregants, the plasmid had undergone deletion of a segment of about 12.5 kb that covered the catabolic genes. The deletion event seemed to be the result of reciprocal recombination between two elements of IS*1600*. During repeated subculturing of NH9 on liquid media with 3-CB, the culture was taken over by a derivative strain (designated NH9A), in which the degradative plasmid carried a duplicate copy of the 12.5-kb region that contained the *cbnR-ABCD* genes. The duplication of these genes seemed again to have been mediated by recombination between the direct repeat sequences.

Transcriptional fusion studies demonstrated that *cbnR* regulates the expression of *cbnABCD* positively in the presence of either 3-CB or benzoate which are catabolized via 3-chlorocatechol and catechol, respectively. In vitro transcription assays

confirmed that 2-chloro-*cis,cis*-muconate (2-CM) and *cis,cis*-muconate (CCM), intermediate products from 3-chlorocatechol and catechol, were inducers of this operon. This inducer-recognizing specificity is different from those of the homologous catechol (*catBCA*) and chlorocatechol (*clcABD*) operons of *Pseudomonas putida*, in which only the intermediates of the regulated pathway, CCM for *catBCA* and 2-CM for *clcABD*, act as significant inducers.

Specific binding of CbnR protein to the *cbnA* promoter region was demonstrated by gel shift and DNaseI footprinting analysis. In the absence of inducer, a region of ca. 60 bp from -20 to -80 upstream of the *cbnA* transcriptional start point was protected from DNaseI cleavage by CbnR, with a region of hypersensitivity to DNase I cleavage clustered at -50. Circular permutation gel shift assays demonstrated that CbnR bent the *cbnA* promoter region to an angle of 78° and that this angle was relaxed to 54° upon the addition of inducer. While a similar relaxation of bending angles upon the addition of inducer molecules observed with the *catBCA* and the *clcABD* promoters may indicate a conserved transcriptional activation mechanism of *ortho*-cleavage pathway genes, CbnR is unique in having a different specificity of inducer recognition and the extended footprint as opposed to the restricted footprint of CatR without CCM.

In the 2,4-dichlorophenoxyacetate (2,4-D) degradative bacterium *Alcaligenes* sp. CSV90, the degradative genes were located within a ca. 12-kb region on its plasmid pMAB1 (90kb) and were organized as two gene clusters, *tfdA-S* and *tfdT-CDEFB*. The genes *tfdS* and *tfdT* apparently encoded LysR-type transcriptional regulators and were located in the opposite orientation from the degradative gene(s) in their respective units. The *tfdA-S-T-CDEFB* genes from CSV90 have turned out to be highly homologous (>98% at the nucleotide level) with the corresponding genes of the 2,4-D degradative gene cluster (*tfdA-S-R-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K-T'-CDEFB*) from *R. eutropha* JMP134 (pJP4[80kb]) which contained the additional genes (*tfdR-*$D_{II}C_{II}E_{II}F_{II}B_{II}$*K*). These results indicated evolutionarily close relationship between the two degradative gene regions, suggesting recent genetic recombination event(s).

Transcriptional fusion assays showed that the *tfdC* promoter was not regulated by *tfdT* but that it was activated by the distantly-located *tfdS* in the presence of 2-CM which is an analogous compound of 2,4-dichloromuconate, the metabolite from 3,5-dichlorocatechol by TfdC. TfdS also activated the *tfdA* promoter in the presence of (chloro)muconate. It was suggested that the three regulators, CbnR, TfdS, and ClcR, have specialized to recognize (chloro)muconates as inducer molecules while some of other LysR-type regulators of (chloro)benzoate/(chloro)catechol pathway seemed to recognize (chloro)benzoates or chlorocatechols as inducer molecule(s).