Mol Gen Genet (1990) 221:113–120

MGG

The *meta* cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes

Shigeaki Harayama and Monique Rekik
Department of Medical Biochemistry, University of Geneva, 1211 Geneva 4, Switzerland

Summary. The *meta*-cleavage operon of TOL plasmid pWW0 of *Pseudomonas putida* encodes a set of enzymes which transform benzoate/toluates to Krebs cycle intermediates via extradiol (*meta*-) cleavage of (methyl)catechol. The genetic organization of the operon was characterized by cloning of the *meta*-cleavage genes into an expression vector and identification of their products in *Escherichia coli* maxicells. This analysis showed that the *meta*-cleavage operon contains 13 genes whose order and products (in kilodaltons) are *xylX*(57)-*xylY*(20)-*xylZ*(39)-*xylL*(28)-*xylT*(12)-*xylE*(36)-*xylG*(60)-*xylF*(34)-*xylJ*(28)-*xylQ*(42)-*xylK*(39)-*xylI*(29)-*xylH*(4). The *xylXYZ* genes encode three subunits of toluate 1,2-dioxygenase. The *xylL*, *xylE*, *xylG*, *xylF*, *xylJ*, *xylK*, *xylI* and *xylH* genes encode 1,2-dihydroxy-3,5-cyclohexadiene-1-carboxylate dehydrogenase, catechol 2,3-dioxygenase, 2-hydroxymuconic semialdehyde dehydrogenase, 2-hydroxymuconic semialdehyde hydrolase, 2-oxopent-4-enoate hydratase, 4-hydroxy-2-oxovalerate aldolase, 4-oxalocrotonate decarboxylase and 4-oxalocrotonate tautomerase, respectively. The functions of *xylT* and *xylQ* are not known at present. The comparison of the coding capacity and the sizes of the products of the *meta*-cleavage operon genes indicated that most of the DNA between *xylX* and *xylH* consists of coding sequences.

Key words: TOL plasmid – Catabolism – *Pseudomonas putida* – *Meta*-cleavage operon – Maxicells

Introduction

TOL plasmid pWW0 of *Pseudomonas putida* encodes a set of enzymes which oxidize toluene/xylenes to Krebs cycle substrates (Fig. 1). Genes for these enzymes are clustered in two operons (Nakazawa et al. 1980; Inouye et al. 1981; Franklin et al. 1981). One operon encodes a set of enzymes which oxidize toluene/xylenes to benzoate/toluates ("upper" pathway; Harayama et al. 1986b, 1989a; Inouye et al. 1981) whereas the other encodes a series of enzymes which degrade benzoate/toluates to pyruvate, acetaldehyde/propionaldehyde and fumarate/acetate (*meta*-cleavage pathway) via extradiol (*meta*) ring-cleavage of catechols (Harayama et al. 1984, 1986a; Inouye et al. 1981). It has been demonstrated that the structural genes for the *meta*-cleavage pathway enzymes are organized in the following order: promoter – *xylXYZ* (the structural genes for toluate 1,2-dioxygenase) – *xylL* (1,2-dihydroxy-3,5-cyclohexadiene-1-carboxylate dehydrogenase) – *xylE* (catechol 2,3-dioxygenase) – *xylG* (2-hydroxymuconic semialdehyde dehydrogenase) – *xylF* (2-hydroxymuconic semialdehyde hydrolase) – *xylJ* (2-oxopent-4-enoate hydratase) – *xylK* (4-hydroxy-2-oxovalerate aldolase) – *xylI* (4-oxalocrotonate decarboxylase) – *xylH* (4-oxalocrotonate tautomerase; Harayama and Rekik 1989; Harayama et al. 1984, 1986a).

Pathways for aerobic degradation of benzoate are found in many bacterial species. In most species, benzoate is transformed by two enzymatic reactions to catechol, which is subsequently transformed to Krebs cycle substrates via extradiol (*ortho*-) cleavage of the aromatic ring. We have recently found a striking similarity between the DNA sequences of *xylXYZL* responsible for the oxidation of benzoate/toluates to (methyl)catechols and those of the isofunctional genes encoded in the chromosome of *Acinetobacter calcoaceticus* (Neidle et al. 1987, and manuscript in preparation). This observation suggested a mechanism of horizontal spread, among many bacterial species, of a DNA module which encodes the ability to transform benzoate to catechol. Similarly, the reactions which assimilate catechol via *meta*-cleavage were found not only in the TOL-encoded pathway, but also in catabolic pathways for naphthalene and phenol (Harayama and Timmis 1989). Occurrence of common *meta*-cleavage routes in many different catabolic pathways was determined to be a consequence of horizontal spread of an ancestral *meta*-cleavage pathway (Harayama et al. 1987a; Keil et al. 1987; Osborne et al. 1988; Shaw and Williams 1988; Assinder and Williams 1988; Ghosal et al. 1987). It would appear that the TOL *meta*-cleavage pathway genes have been formed from the fusion of two distinct DNA modules, one determining the activities necessary to transform benzoate to catechol and the other encoding enzymes for assimilation of catechol via *meta*-cleavage. To aid understanding of the function and evolution of the *meta*-cleavage pathway, it is important to determine precisely the gene organization of the *meta*-cleavage operon. In this study, we analyzed gene products of the operon in *Escherichia coli* maxicells. This analysis allowed identification of two new genes.

Materials and methods

Strains and plasmids. Strains used were *E. coli* K12 derivatives, K12ΔHIΔtrp [*lacZ*(am) Δ(*bio-uvrB*) Δ*trp*(*E-*

Offprint requests to: S. Harayama

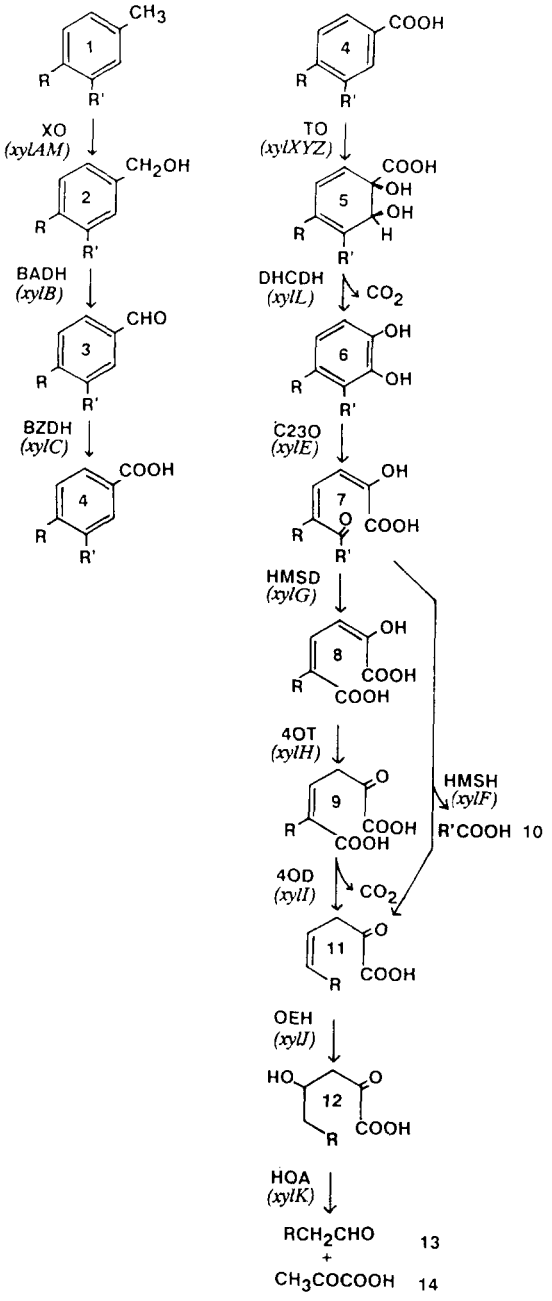

Fig. 1. The TOL plasmid-encoded metabolic pathway for degradation of toluene/xylenes to benzoate/toluates: pathway intermediates, enzymes, and their structural genes. Enzyme abbreviations: XO, xylene monooxygenase (multicomponent); BADH, benzyl alcohol dehydrogenase; BZDH, benzaldehyde dehydrogenase; TO, toluate 1,2-dioxygenase (multicomponent); DHCDH, 1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase; C23O, catechol 2,3-dioxygenase; HMSD, 2-hydroxymuconic semialdehyde dehydrogenase; HMSH, 2-hydroxymuconic semialdehyde hydrolase; 4OT, 4-oxalocrotonate tautomerase; 4OD, 4-oxalocrotonate decarboxylase; OEH, 2-hydroxypent-2,4-dienoate hydratase; HOA, 4-hydroxy-2-oxovalerate aldolase. *xylA* to *xylZ* designate the structural genes for the catabolic enzymes. Compounds: for R=H, R'=H, (1) toluene, (2) benzyl alcohol, (3) benzaldehyde, (4) bentoate, (5) 1,2-dihydroxycyclohexa-2,5-diene-1-carboxylate, (6) catechol, (7) 2-hydroxymuconic semialdehyde, (8) 2-hydroxyhexa-2,4-diene-1,6-dioate (enol form of 4-oxalocrotonate), (9) 2-oxohex-4-ene-1,6-dioate (keto form of 4-oxalocrotonate), (10) formate, (11) 2-oxopent-4-enoate, (12) 4-hydroxy-2-oxovalerate, (13) acetaldehyde, (14) pyruvate; for R=H, R'=CH_3, (1) *m*-xylene, (2) *m*-methylbenzyl alcohol, (3) *m*-methylbenzaldehyde, (4) *m*-toluate, (5) 1,2-dihydroxy-3-methylcyclohexa-3,5-diene-1-carboxylate, (6) 3-methylcatechol, (7) 2-hydroxy-6-oxohepta-2,4-dienoate, (10) acetate, (11) 2-oxopent-4-enoate, (12) 4-hydroxy-2-exovalerate, (13) acetaldehyde, (14) pyruvate; for R=CH_3, R'=H, (1) *p*-xylene, (2) *p*-methylbenzyl alcohol, (3) *p*-methylbenzaldehyde, (4) *p*-toluate, (5) 1,2-dihydroxy-4-methylcyclohexa-3,5-diene-1-carboxylate, (6) 4-methylcatechol, (7) 2-hydroxy-5-methyl-6-oxohexa-2,4-dienoate, (8) 2-oxopent-5-methyl-4-ene-1,6-dioate, (9) 5-methyl-2-oxohex-4-ene-1,6-dioate, (10) formate, (11) 2-hydroxy-*cis*-hex-2,4-dienoate, (12) 4-hydroxy-2-oxohexanoate, (13) propionaldehyde, (14) pyruvate

A)2/λN7(am) *N57*(am) *cI857*(ts) Δ*HI*] and MCL22 [*trkA405 trkD1 thi-1 rha rpsL31* Δ(*kdp-phr*)*214* Δ(*gal-uvrB*) Δ(*srl-recA*)*306*]. Plasmids used were pRME1, pCI857 and pLV85 (Harayama et al. 1986b, 1989a). Derivatives of pLV85 constructed in this study are shown in Fig. 2. Since the present work involves many plasmid constructs, we use their simplified names (A, B, C, etc.) instead of their original names (pGSH2829, 2831, 2830, etc) in the text. Both sets of names are listed in Fig. 2.

DNA and genetic manipulations. Methods for DNA manipulations have been described by Maniatis et al. (1982). The method for Tn*1000* insertion mutagenesis has been described previously (Harayama et al. 1989a).

Identification of gene products using maxicells. In the MCL22 strain containing two compatible plasmids, pCI857 and a pLV85 derivative, genes cloned into pLV85 downstream of the P_2 promoter are inducible at 42° C. Cells of MCL22 containing pCI857 and a pLV85-based plasmid were grown overnight at 30° C in L broth (LB) containing 100 mM KCl (MCL22 is defective in potassium transport), 100 μg/ml ampicillin (selection for a pLV85 derivative) and 25 μg/ml kanamycin (selection for pCI857). A 20 μl aliquot of the overnight culture was transferred into 10 ml of the same medium and the culture was grown to about 3×10^8 cells/ml. Cells were washed and resuspended in 10 ml TRIS-HCl buffer (100 mM, pH 7.3) containing 10 mM $MgSO_4$, and irradiated for 20 s by a germicidal lamp (15 W) at a distance of 30 cm. Irradiated cells were suspended in 10 ml LB containing 100 mM KCl and incubated at 30° C for 1 h, then D-cycloserine was added to 100 μg/ml. The incubation at 30° C was continued overnight and the next morning the same amount of D-cycloserine was added. The incubation at 30° C was continued for 2 h and the cells were pelleted, washed and resuspended to 4×10^9 cells/ml in methionine assay medium (Difco) containing M9 base salt and 0.2% glucose. The cells were incubated at 30° C for 5 min then further incubated for 15 min either at 30° or 42° C. [^{35}S]methionine (1.2 Ci/μmol, Amersham) at a final concentration of 1 μM was added to the bacterial suspension thus treated, and the suspension incubated at the same temperature for 5 min. Unlabeled methionine at a final concentration of 10^{-4} M was then added to the bacterial suspension and after 20 s, trichloroacetic acid to a final concentration of 2% was added to stop the reaction. The cells were then pelleted, washed, and resuspended in 40 μl of sample buffer (8% sodium dodecyl sulfate, 1% β-mercaptoethanol, 5% glycerol, 0.004% bromophenol blue and 0.1 M TRIS-HCl, pH 7.3). Gel electrophoresis and autoradiography were performed as previously described (Harayama et al. 1982).

Enzyme assays. A 1 l sample of K12ΔH1Δtrp cells containing a pLV85 derivative was grown at 30° C in LB containing 100 μg/ml ampicillin. When the cell density had reached approximately 2×10^8 cells/ml, the culture was shifted to 42° C and 250 ml aliquots were withdrawn every hour from 1 to 4 h after the temperature shift. Cells were pelleted, washed and resuspended in 30 ml of 0.1 M potassium phosphate buffer (pH 7.3). Preparation of cell-free extracts and assays for catechol 2,3-dioxygenase, 2-hydroxymuconic semialdehyde dehydrogenase, 2-hydroxymuconic semialdehyde hydratase, 2-oxopent-4-enoate hydratase, 4-oxalocrotonate decarboxylase and 4-oxalocrotonate tautomerase were carried out as previously described (Harayama et al. 1984, 1987b). The assay method for 1,2-dihydroxy-3,5-cyclohexadiene-1-carboxylate dehydrogenase has been described by Reiner (1972).

Results

Our previous studies showed (Harayama et al. 1984, 1986a) that the *Eco*RI-E (fifth largest *Eco*RI) segment of pWW0 contains the catabolic genes, *xylYZLEGFJ*. This segment was cloned into the expression vector, pLV85, in an orientation which allowed expression of *xylYZLEGFJ* from the lambda P_L promoter. This construct (plasmid A; Fig. 2) was introduced into the *E. coli* host, K12ΔH1Δtrp, which carries the temperature-sensitive lambda repressor gene, *cI857*, on its chromosome. At 42° C, the following enzyme activities were induced: 1,2-dihydroxy-3,5-cyclohexadiene-1-carboxylate dehydrogenase, catechol 2,3-dioxygenase, 2-hydroxymuconic semialdehyde dehydrogenase, 2-hydroxymuconic semialdehyde hydrolase and 2-oxopent-4-enoate hydratase (data not shown).

Tn*1000* insertion mutants of plasmid A were isolated and the sites of these insertions were located on the restriction map of the TOL *Eco*RI-E segment. From the genetic map established previously (Harayama et al. 1984, 1986a, 1987b), the genes inactivated by these insertions could be deduced: Tn*1000* insertions 1 and 2 were in *xylZ*, 3 and 4 in *xylL*, 5 in *xylE*, 6 in *xylG* and 7 in *xylJ* (Fig. 2). By measuring enzyme activities expressed at 42° C in K12ΔH1Δtrp cells containing Tn*1000* insertion derivatives of plasmid A (A-1, A-2, etc.), it was demonstrated that these

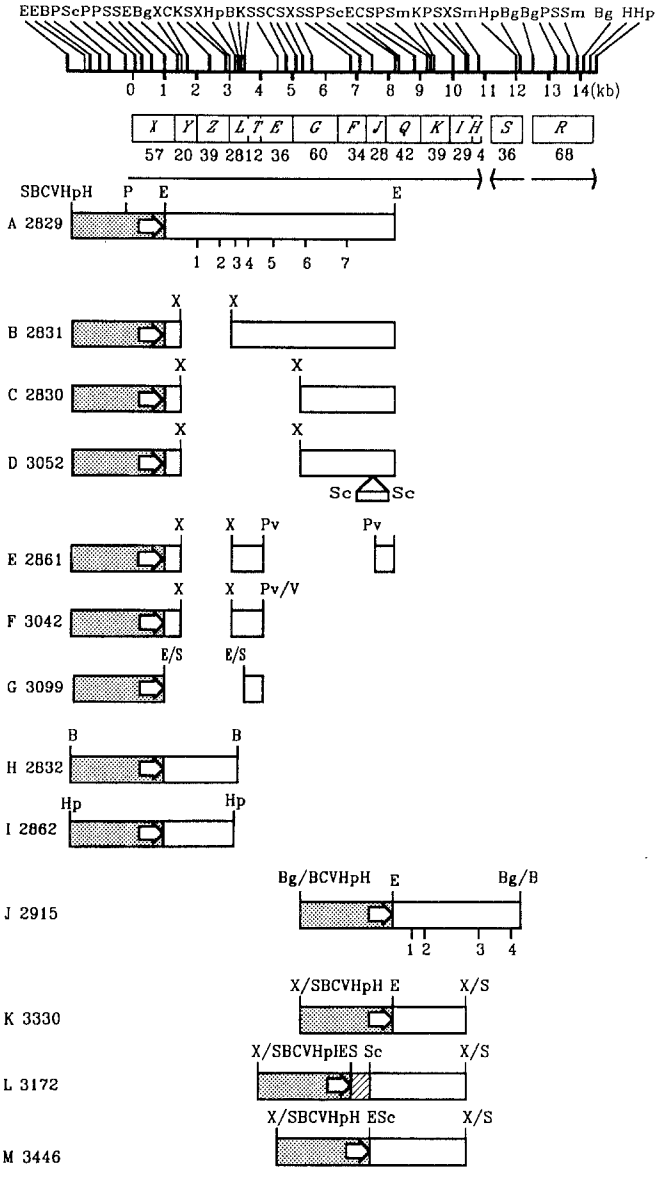

Fig. 2. pLV85-based hybrid plasmids containing the TOL *meta*-cleavage genes. The uppermost part shows the restriction sites of TOL plasmid pWW0 and location of the *meta*-cleavage pathway genes. The transcriptional start site of the operon (Inouye et al. 1984; Mermod et al. 1984) is given as the zero coordinate of the restriction map. B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; K, *Kpn*I; P, *Pst*I; S, *Sal*I; Sc, *Sac*I; Sm, *Sma*I; X, *Xho*I. The functions of *xylX*, *Y*, *Z*, *L*, *E*, *G*, *F*, *J*, *K*, *I* and *H* are indicated in Fig. 1 whereas those of *xylT* and *xylQ* are not known. The *xylS* and *xylR* genes are regulatory genes (Inouye et al. 1986, 1988). The sizes of the products in kDa are indicated under the gene names. The structures of the hybrid plasmids used in this study are indicated under the map. *Open* and *shaded boxes* show TOL plasmid pWW0 DNA and the pLV85 expression vector DNA, respectively. *Arrows* in the *shaded boxes* indicate the direction of transcription from the P_L promoter. Plasmid A was constructed by cloning the TOL *Eco*RI-E segment into pLV85. Tn*1000* insertion sites in plasmids A1 to A7 (derived from plasmid A) are indicated. Plasmid B was constructed from plasmid A by deletion of its 1.5 kb *Xho*I segment, whereas plasmid C was constructed by deletion of both the 1.5 and 2.3 kb *Xho*I segments of plasmid A. Plasmid D is a derivative of plasmid C carrying a kanamycin resistance cassette inserted into its unique *Sac*I site. Plasmid E was isolated from plasmid B by deleting three *Pvu*II (Pv) segments of the latter plasmid (*Pvu*II sites on the TOL DNA are not shown on the restriction map). Plasmid F was constructed from plasmid E by digestion with *Pvu*II and *Eco*RV (V) followed by ligation (*Eco*RV sites on the TOL DNA are not shown on the restriction map). Plasmid G was constructed by digesting plasmid F with *Eco*RI and *Sal*I followed by Klenow-filling and ligation. Plasmids H and I were constructed from plasmid A by deletion of its 5 kb *Bam*HI segment and its 5.3 kb *Hpa*I segment, respectively. Plasmid J was constructed by cloning a 4 kb *Eco*RI-*Bgl*II segment of TOL pWW0 into pLV85. From plasmid J, Tn*1000* insertion derivatives J1, J2, J3 and J4 were isolated. Their Tn*1000* insertion sites are indicated. Plasmid K was constructed by inserting a 2.3 kb *Eco*RI-*Xho*I segment of TOL pWW0 between the *Eco*RI and *Sal*I cloning sites of pLV85. Plasmid L was constructed by inserting the 3 kb *Sac*I-*Xho*I segment of TOL pWW0 into pGSH2833, a pLV85 derivative containing TOL upper operon DNA (Harayama et al. 1989a). In plasmid L, a 0.6 kb DNA segment derived from the upper operon of TOL pWW0 (*striped box*) has been inserted between the vector DNA and the *meta*-cleavage operon DNA. In plasmid M, the same 3 kb *Sac*I-*Xho*I segment was cloned directly into pLV85 between the *Eco*RI and *Sal*I sites using an *Eco*RI-*Sac*I adaptor

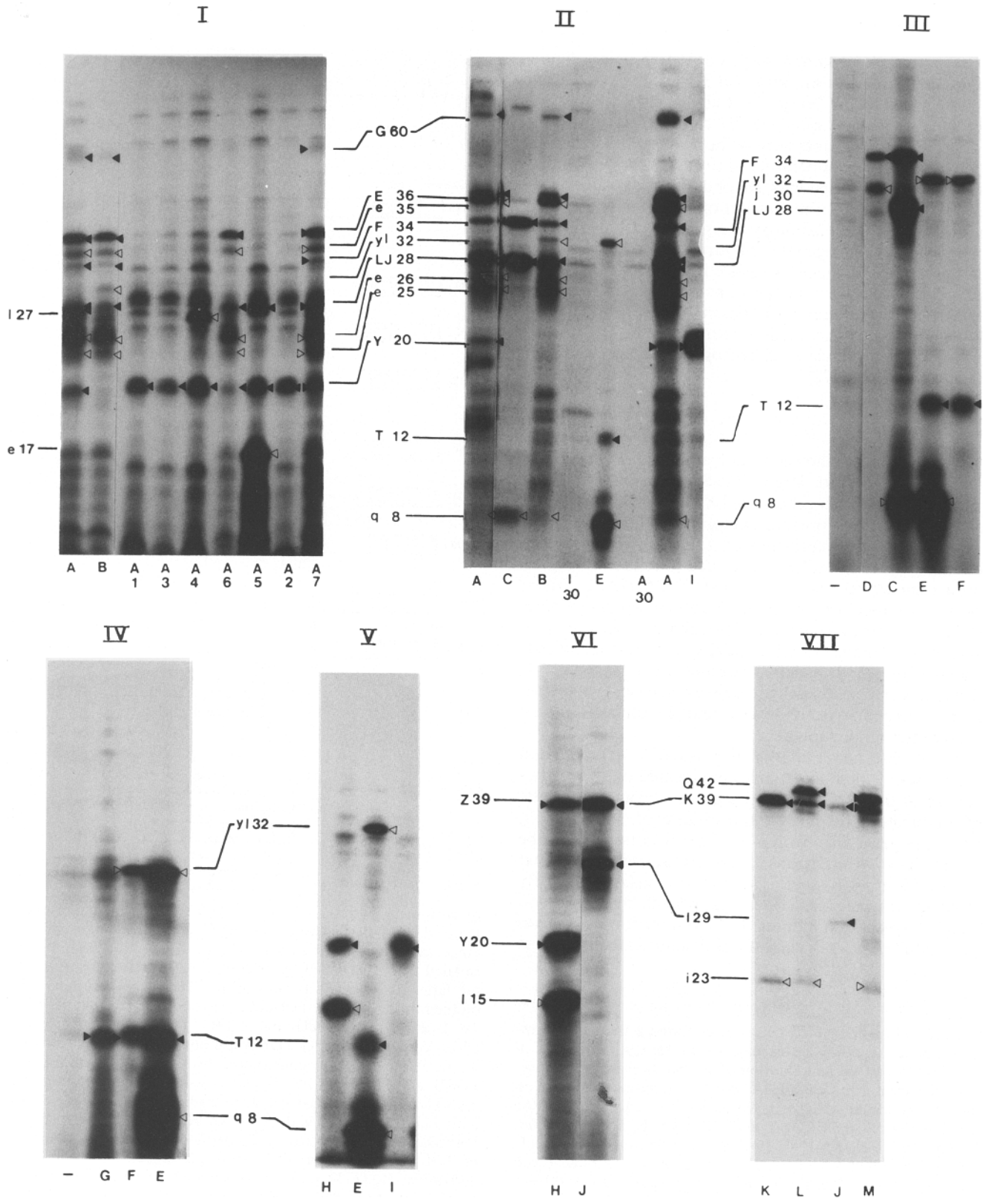

Fig. 3. Autoradiograms of sodium dodecyl sulfate polyacrylamide gels containing [^{35}S]-methionine-labeled proteins synthesized in maxicells. The *xyl* genes cloned into pLV85 were, if not stated otherwise, induced at 42° C and their products were examined in maxicells. The symbol (◀) indicates the *xyl* gene products whereas the symbol (◁) indicates truncated or degraded products of the *xyl* genes. On the left side of the gel, molecular weights of gene products and gene names are indicated. *Upper case letters* indicate that these are true products whereas *lower case letters* mean that they are either truncated or degraded products. (I) Lane 1, plasmid A; lane 2, plasmid B; lane 3, plasmid A-1; lane 4, plasmid A-3; lane 5, plasmid A-4; lane 5, plasmid A-6; lane 6, plasmid A-5; lane 7, plasmid A-2; lane 8, plasmid A-7. (II) Lane 1, plasmid A; lane 2, plasmid C; lane 3, plasmid B; lane 4, plasmid H; lane 5, plasmid F; lane 6, plasmid A (30° C); lane 7, plasmid A; lane 8, plasmid I. (III) Lane 1, pLV85; lane 2, plasmid D; lane 3, plasmid C; lane 4, plasmid E; lane 5, plasmid F. (IV) Lane 1, pLV85; lane 2, plasmid G; lane 3, plasmid F; lane 4, plasmid E. (V) Lane 1, plasmid H; lane 2, plasmid F; lane 3, plasmid I. (VI) Lane 1, plasmid H; lane 2, plasmid J. (VII) Lane 1, plasmid K; lane 2, plasmid L; lane 3, plasmid J; lane 4, plasmid M

Tn*1000* insertion mutations exerted polar effects on downstream genes (data not shown). Plasmid A and its Tn*1000* insertion derivatives were used to transform the maxicell strain MCL22 containing pCI857. MCL22 containing pCI857 and plasmid A produced at least eight different proteins at 42° C (Fig. 3-I, lane 1). The $XylY^+Z^-L^-E^-G^-F^-J^-$ (A-1 and A-2) and $XylY^+Z^+L^-E^-G^-F^-J^-$ (A-3 and A-4) mutants of plasmid A synthesized a 20 kDa polypeptide (Fig. 3-I). Therefore, the 20 kDa polypeptide is the *xylY* product, as was previously determined (Harayama et al. 1986a). One polypeptide of 27 kDa in size was synthesized from the $XylY^+Z^+L^-E^-G^-F^-J^-$ plasmid A-4 but not from the parental plasmid A (Fig. 3-I), and was presumably a truncated product from the *xylL* sequence. The *xylZ* product which could be synthesized from the $XylZ^+$ plasmids (A, A-3 to A-7) but not from the $XylZ^-$ plasmids (A-1 and A-2) was not identified from this autoradiogram. A 28 kDa polypeptide was produced from $XylY^+Z^+L^+E^-G^-F^-J^-$ (A-5) but not from $XylY^+Z^+L^-E^-G^-F^-J^-$ (A-3 and A-4) plasmids (Fig. 3-I). Therefore, the *xylL* product is most likely this 28 kDa protein. Plasmid A-5 synthesized, in addition to the *xylL* product, a protein of 17 kDa which may be a truncated product from the *xylE* sequence. This band was smeared, suggesting that this truncated protein is rapidly degraded in vivo.

When the products from the $XylE^+$ plasmid A-6 and the $XylE^-$ plasmid A-5 were compared (Fig. 3-I), the $XylE^+$ plasmid A-6 synthesized 36, 35, 26 and 25 kDa polypeptides in addition to the polypeptides synthesized from the $XylE^-$ plasmid A-5. We interpreted this result to mean that the *xylE* product is a 36 kDa polypeptide as has been demonstrated in previous studies (Nakai et al. 1983a, b) and that the 35, 26 and 25 kDa polypeptides were degraded products from this 36 kDa polypeptide.

Two polypeptides of 60 and 34 kDa in size were synthesized from plasmids A and A-7 (Fig. 3-I, lanes 1 and 9) but not from plasmid A-6 (Fig. 3-I, lane 6). Therefore, these polypeptides are encoded in the *xylGF* region. To pinpoint the genes for these polypeptides, deletion derivatives were isolated from plasmid A (Fig. 2). Plasmid B was constructed by deletion of a 1.5 kb *Xho*I segment of plasmid A. As described below, the *Xho*I site at the 3 kb coordinate is inside the *xylL* gene. Therefore, plasmid B carries intact *xylEGFJ*. Similarly, plasmid C was constructed by deletion of both the 1.5 and the 2.3 kb *Xho*I segments of plasmid A, hence carries only the *xylFJ* genes. Plasmid C synthesized three polypeptides whose sizes were 34, 28 and 8 kDa (Fig. 3-II, lane 2). This plasmid neither synthesized the 20 kDa *xylY* product nor the 36, 35, 26 or 25 kDa *xylE* products as expected, but synthesized a 28 kDa protein in spite of the fact that the *xylL* gene was deleted from this plasmid. We concluded from this observation that the *Eco*RI-E segment encodes a second 28 kDa protein in addition to the *xylL* product. Plasmid B produced polypeptides of 60, 36, 35, 32, 26 and 25 kDa in size (Fig. 3-I, lane 2; Fig. 3-II, lane 3) in addition to the gene products synthesized from plasmid C. The 32 kDa polypeptide synthesized from plasmid B was not produced from plasmid A (compare Fig. 3-II, lane 3 and lanes 1, 7), therefore this band may correspond to a polypeptide synthesized from a fusion of *xylY* and *xylL*. Polypeptides of 36, 35, 26 and 25 kDa and those smaller than 18 kDa synthesized from plasmid B would be the *xylE* products as described above. A 60 kDa polypeptide synthesized from $XylE^+G^+F^+J^+$ plasmid B (Fig. 3-II, lane 3) could not be seen in the products from $XylF^+J^+$ plasmid C (Fig. 3-II, lane 2). Therefore, it was concluded that the *xylG* product is the 60 kDa protein.

To identify structural genes for the 34, 28 and 8 kDa polypeptides synthesized from plasmid C (Fig. 3-II, lane 2; Fig. 3-III, lane 3), the *neo* gene cassette of pRME1 was inserted into the unique *Sac*I site of plasmid C, and this construct (called plasmid D) was introduced into the maxicell strain MCL22 containing pCI857. Plasmid D synthesized the 34 kDa polypeptide but not the 28 and 8 kDa polypeptides (Fig. 3-III, lane 2) suggesting that the structural gene for the 34 kDa polypeptide is located upstream of the *Sac*I site at the 7.5 kb coordinate whereas the structural genes for the 28 and 8 kDa polypeptides are located at or downstream from the *Sac*I site. In our previous study, the *xylF* gene was mapped upstream of the *Sac*I site (Harayama et al. 1984), therefore, the 34 kDa polypeptide is the *xylF* product. Plasmid D synthesized a 30 kDa protein in addition to the 34 kDa *xylF* protein. This polypeptide was not synthesized from plasmid C. The 30 kDa protein therefore may be a truncated product of the *xylJ* gene fused with a sequence on the *neo* gene cassette.

Plasmid C has four *Pvu*II sites. Three *Pvu*II segments internal to the cloned DNA of plasmid C were deleted to construct plasmid E. The latter plasmid synthesized 32, 12 and 8 kDa polypeptides (Fig. 3-III, lane 4). The 28 kDa *xylL* product was not produced from this plasmid. This observation indicated that *xylL* is defective in plasmid D and therefore that the *Xho*I site at the 3.0 kb coordinate is inside the *xylL* gene. From this conclusion, we determined that the 32 kDa polypeptide synthesized from plasmid E was the product of the *xylY-L* fusion. When the 0.5 kb of DNA between the *Pvu*II site at the 7.7 kb coordinate and the *Eco*RI site at the 8.2 kb coordinate was deleted from plasmid E to construct plasmid F (Fig. 2), the latter plasmid did not synthesize an 8 kDa polypeptide (Fig. 3-III, lane 5). Therefore, the DNA sequence which encodes this 8 kDa polypeptide was located at the right end of the *Eco*RI-E segment. As described above, plasmid D synthesized the 34, 28 and 8 kDa polypeptides. The 34 kDa polypeptide was the *xylF* product and the structural gene for the 8 kDa polypeptide was located at the very right end of the *Eco*RI-E segment. Therefore, the genes in this region are organized in the order: 34 kDa gene (*xylF*) – 28 kDa gene – 8 kDa gene. In our previous study (Harayama et al. 1984), the *xylJ* gene was mapped to a region which included the structural gene for the 28 kDa polypeptide. Therefore, this 28 kDa polypeptide is the *xylJ* product. Plasmids E and F both synthesized a 12 kDa polypeptide in addition to the 32 kDa protein synthesized from the *xylY-L* fusion (Fig. 3-IV, lanes 4 and 3). The 12 kDa polypeptide was also synthesized from plasmid G (Fig. 3-IV, lane 2). Therefore, this polypeptide is coded for by a gene located between *xylL* and *xylE*. The existence of a gene which is located between *xylL* and *xylE* and which encodes a polypeptide of 12 kDa was confirmed by DNA sequencing of this region (unpublished). This gene has not been identified in previous studies and was named *xylT*.

Plasmid H is a derivative of plasmid A from which a 5 kb *Bam*HI segment was deleted. From the known physical and functional maps (Harayama et al. 1984) this plasmid was expected to carry *xylY* and *xylZ*. Plasmid I is also a derivative of plasmid A from which a 5.3 kb *Hpa*I seg-

ment was deleted. In maxicells, plasmid H synthesized 20 and 15 kDa polypeptides (Fig. 3-V, lane 1) whereas plasmid I only produced a 20 kDa polypeptide (Fig. 3-V, lane 3). The 20 kDa polypeptide was the *xylY* product as described above. Since the 15 kDa polypeptide was not synthesized from plasmid A, it may be a truncated product of *xylL*.

The *xylZ* gene product was not detected in these autoradiograms. In some experiments, however, a 39 kDa protein was detected as a product from plasmid H (Fig. 3-VI, lane 1). The difficulty in detecting this product may be due to problems in migration of this protein in sodium dodecyl sulfate polyacrylamide gels. For instance, separation of the same radioactive sample on two different gels sometimes resulted in detection of the 39 kDa protein on one but not on the other gel. We could not find gel conditions which allowed reproducible detection of the *xylZ* product. Nevertheless, we have recently determined the complete nucleotide sequence of the *xylZ* gene and its product was deduced to be a 36 kDa protein (unpublished).

To examine genes and their products downstream of *xylJ*, a 4 kb *Eco*RI-*Bgl*II segment in the vicinity of the *Eco*RI-E segment was cloned into pLV85 to construct plasmid J. From our previous studies (Harayama et al. 1984; Harayama and Rekik 1989), three genes, *xylK*, *xylI* and *xylH* (in that order) have been mapped on the 4 kb *Eco*RI-*Bgl*II segment. In maxicells, plasmid J produced two polypeptides of 29 and 39 kDa in size (Fig. 3-VI, lane 2; Fig. 3-VII, lane 3). To locate further the structural genes for these two polypeptides, we isolated plasmids J-1 to J-4 which are Tn*1000* insertion derivatives of plasmid J. The insertion sites of Tn*1000* in these mutant derivatives are presented in Fig. 2. From the functional maps of this region that have been constructed in our previous studies (Harayama et al. 1984; Harayama and Rekik 1989), insertions 1 and 2 were mapped upstream of, or within the *xylK* gene. Insertion 3 was located within the *xylH* gene and insertion 4 mapped downstream of this gene. Since Tn*1000* exerts a polar effect, the phenotypes of plasmids J-1, J-2, J-3 and J-4 were expected to be $XylK^-I^-H^-$, $XylK^-I^-H^-$, $XylK^+I^+H^-$ and $XylK^+I^+H^+$, respectively. Assays for the XylI and XylH enzymes in cell-free extracts prepared from plasmids J-1 to J-4 confirmed this notion (data not shown). As expected, plasmids J-1 and J-2 did not synthesize any *xyl* gene products in maxicells, however both plasmids J-3 and J-4 synthesized 39 and 29 kDa polypeptides (data not shown). Since *xylH* is defective in J-3, none of these two polypeptides could have been the *xylH* product. Attempts to identify in maxicells the *xylH* product using different constructs carrying *xylH* also failed (unpublished) but the *xylI* and *xylK* gene products were confirmed using another construct, plasmid K. Plasmid K is a derivative of pLV85 containing a 4 kb *Eco*RI-*Xho*I segment covering the *xylK* region (Fig. 2). The *xylI* activity was not expressed from plasmid K. This $XylK^+XylI^-$ plasmid synthesized the 39 kDa polypeptides but not the 29 kDa polypeptide (Fig. 3-VII, lane 1). Therefore, the *xylI* and *xylK* products were identified as the 29 and 39 kDa polypeptides, respectively.

The analysis of gene products from the *Eco*RI-E segment described above detected an 8 kDa polypeptide gene product encoded at the right end of this segment. This protein was, most likely, the truncated product of a gene encompassing the *Eco*RI site at the 8.2 kb coordinate. To examine this possibility, a 3 kb *Sac*I-*Xho*I segment spanning the 7.5 kb and 10.5 kb coordinates was cloned to construct plasmids L and M. Plasmid L was constructed by cloning of this segment into pGSH2833, a pLV85-based hybrid plasmid containing a sequence of the upper operon of the TOL pWW0 plasmid (Harayama et al. 1986b). Consequently, plasmid L contains a 0.6 kb segment derived from the upper operon DNA from which no gene product is synthesized (Harayama et al. 1989a). Plasmid M was constructed by cloning the 3 kb *Sac*I-*Xho*I segment into the multiple cloning sites of pLV85 between the *Eco*RI and *Sal*I sites. In this construction, an *Eco*RI-*Sac*I adaptor (GAATTCGAGCTC) was inserted between the *Eco*RI cloning site of pLV85 and the *Sac*I site of the TOL DNA. From both plasmids L and M, two polypeptides of 42 and 39 kDa in size were synthesized (Fig. 3-VII, lanes 2 and 4). Above, we identified the 39 kDa polypeptide as the *xylK* product. The 42 kDa polypeptide should not be encoded at the right end of the TOL DNA in plasmids L and M because plasmid K did not synthesize this product. Therefore, the gene for the 42 kDa polypeptide must be located upstream of *xylK*. This protein is not a product synthesized from the left end of the *Sac*I-*Xho*I segment fused with an upstream sequence because it was synthesized from both plasmids L and M, in which the upstream sequences are different. The existence of the 42 kDa gene was not expected from the known enzyme activities of the *meta*-cleavage pathway. The new gene was named *xylQ*. The protein of 23 kDa in size synthesized from plasmids K, L and M (Fig. 3-VII, lanes 1, 2, and 4) may be a truncated product of *xylI*. Therefore, the arrangement of genes in this region is: 42 kDa gene (*xylQ*) – 39 kDa gene (*xylK*) – 29 kDa gene (*xylI*).

Discussion

In our previous study, we demonstrated that the first three genes of the *meta*-cleavage operon, *xylX*, *xylY* and *xylZ*, encode three different subunits of toluate 1,2-dioxygenase and that the products of *xylX* and *xylY* are 57 and 20 kDa polypeptides, respectively (Harayama et al. 1986a). In this study, the product of *xylZ* was determined to be a 39 kDa polypeptide. Therefore, toluate 1,2-dioxygenase is composed of subunits of 57, 20 and 39 kDa. In *Pseudomonas* and several other soil bacteria, chromosomally encoded benzoate 1,2-dioxygenase is ubiquitous. This enzyme catalyzes a similar reaction to that catalyzed by TOL-encoded toluate 1,2-dioxygenase but its substrate specificity is narrower than that of the TOL-encoded enzyme (Reineke and Knackmuss 1978). Yamaguchi and Fujisawa (1978, 1980, 1982) have purified benzoate 1,2-dioxygenase from *P. putida*, and the enzyme was found to be composed of three different subunit proteins of about 50, 20 and 38 kDa in size. The 50 and 20 kDa components constitute an oxygenase, whereas the 39 kDa component has NADH-cytochrome *c* reductase activity. The similarity in the molecular weights of the three subunit proteins of toluate 1,2-dioxygenase with those of benzoate 1,2-dioxygenase suggests the evolutionary relatedness of these two enzymes. Our unpublished data, in fact, demonstrate an evolutionary relationship between TOL-encoded toluate 1,2-dioxygenase and benzoate 1,2-dioxygenase. Thus, we can infer the biochemical functions of the *xylXYZ* products: the *xylX* and *xylY*

products are two components of an oxygenase and the *xylZ* product is an NADH-cytochrome *c* reductase.

So far, two enzymes isofunctional to the *xylL* product, 1,2-dihydroxy-3,5-cyclohexadiene-1-carboxylate dehydrogenase, have been purified from *A. calcoaceticus* and *Alcaligenes eutrophus* (Neidle et al. 1987; Reiner 1972). The molecular weights of the purified enzymes were similar to that of the TOL enzyme. However, their molecular weights do not correspond to those of naphthalane dihydrodiol dehydrogenase from a *Pseudomonas* strain (Patel and Gibson 1974) and *cis*-benzene dihydrodiol dehydrogenase from *Pseudomonas* strains (Irie et al. 1987; Axcell and Geary 1973) which catalyze similar but different reactions. We do not know whether the diversity in molecular weight of dihydrodiol dehydrogenases in different organisms suggests evolutionarily independent origins of these functionally similar enzymes.

Between *xylL* and *xylE*, we found a gene, *xylT*, which encodes a polypeptide of 12 kDa in size. Our DNA sequence data in this region also confirmed the existence of an open reading frame which can synthesize a polypeptide of this size (unpublished). At present, we are not aware of the function of this gene.

We identified at least four polypeptides of 36, 35, 26 and 25 kDa in size as the products of *xylE*. Since catechol 2,3-dioxygenase in its active form contains ferrous ion as a prosthetic group, the 36 and 35 kDa polypeptides may represent the same protein, each containing a different state of iron, whereas the 26 and 25 kDa polypeptides may be degradative products of the higher molecular weight polypeptides. We do not know whether the apparent instability of the catechol 2,3-dioxygenase polypeptide was the result of artificial expression of this enzyme in *E. coli* or of the native characteristics of the enzyme.

The *xylF* product, 4-hydroxymuconic semialdehyde hydrolase, has been purified by Duggleby and Williams (1986) from the same TOL-bearing organism. They reported that the molecular weight of the native enzyme, composed of two identical subunits, was 65000. This result is in agreement with our present results in which the monomer of this protein has a molecular weight of 34000. Isofunctional enzymes purified from other sources were smaller (Bayly and Di Berardino 1978).

We have recently purified 2-oxopent-4-enoate hydratase (the *xylJ* product) and 4-oxalocrotonate decarboxylase (the *xylI* product). These two enzymes are apparently associated with each other in vivo and are co-purified by several different chromatographic procedures. The purest samples contained two polypeptides of 28 and 29 kDa in sizes (Harayama et al. 1989b). From the present study, we could identify the 28 kDa protein as 2-oxopent-4-enoate hydratase and the 29 kDa protein as 4-oxalocrotonate decarboxylase. 2-oxopent-4-enoate hydratase has also been purified from *P. putida* NCIB 10015 by Collinsworth et al. (1973) and its molecular weight has been determined to be 28 kDa.

Our genetic analysis (Harayama and Rekik 1989) identified the region encoding a 39 kDa polypeptide as the structural gene for 4-hydroxy-2-oxovalerate aldolase (*xylK*). Between *xylJ* and *xylK*, we found one gene encoding a 42 kDa polypeptide. The function of the 42 kDa gene, named *xylQ*, is not known at present.

By transposon insertion and deletion analysis, we have located *xylH*, the structural gene for 4-oxalocrotonate tautomerase, at the end of the *meta*-cleavage operon (Harayama et al. 1984). In the present study, we could not identify the *xylH* product. Nevertheless, our recent studies have demonstrated that 4-oxalocrotonate tautomerase is a small polypeptide of 4 kDa in size. The failure to detect the *xylH* product in maxicells may be due to either the low translational efficiency of the *xylH* gene in *E. coli* or, since 4-oxalocronate tautomerase is a relatively small protein, low methionine content of the *xylH*-encoded protein. The positive regulatory gene of the *meta*-cleavage operon, *xylS*, has been mapped downstream of *xylH* (Inouye et al. 1986). Between *xylH* and *xylS* there is only a 0.35 kb interval. Therefore, the *xylH* gene may be the last gene of the *meta*-cleavage operon.

When coding capacities and gene product sizes were compared in the *meta*-cleavage operon, most of the DNA in this region was estimated to be composed of coding sequences. This compact structure of the *meta*-cleavage operon contrasts with the structure of the upper pathway operon, where a 1.5 kb non-coding sequence was found at the beginning of the operon (Harayama et al. 1989a).

Overall, the *meta*-cleavage operon is composed of 13 genes and extends over 10 kb of DNA, making it one of the biggest operons found in bacteria. This fact poses several questions as to the transcription and stability of the *meta*-cleavage operon mRNA, such as whether any specific structure at the 3′ end of the mRNA exists which might increase its stability (Belasco and Higgins 1988) or whether there is a secondary promoter inside the operon. The second question raises the possibility that the *meta*-cleavage operon was formed by fusion of two DNA molecules, comprising *xylXYZL(T)* and *xyl(T)EGFJQKIH*, which each possessed its own promoter.

Two new genes, *xylT* and *xylQ*, were found in this study but their functions were not identified. Site-directed mutagenesis of these genes is currently under way in our laboratory to clarify their physiological roles.

Acknowledgements. We wish to thank Dr. B. Angus for careful reading of the manuscript and Miss F. Rey for secretarial work. We wish to thank Professor L.N. Ornston and Dr. E. Neidle for communicating the *benABC* sequence prior to publication. This work was supported by Grant No. 3.546-0.86 from Swiss National Science Foundation.

References

Assinder SJ, Williams PA (1988) Comparison of the *meta* pathway operons on NAH plasmid pWW60-22 and TOL plasmid pWW53-4 and its evolutionary significance. J Gen Microbiol 134:2769–2778

Axcell BC, Geary PJ (1973) The purification and some properties of the enzyme *cis*-1,2-dihydroxycyclohexa-3,5-diene (nicotinamide adenine dinucleotide) oxidoreductase (*cis*-benzene glycol dehydrogenase). Biochem J 136:927–934

Bayly RC, Di Berardino D (1978) Purification and properties of 2-hydroxy-6-oxo-2,4-heptadienoate hydrolase from two strains of *Pseudomonas putida*. J Bacteriol 134:30–37

Belasco JG, Higgins CF (1988) Mechanisms of mRNA decay in bacteria: a prospective. Gene 72:15–23

Collinsworth WL, Chapman PJ, Dagley S (1973) Stereospecific enzymes in the degradation of aromatic compounds by *Pseudomomans putida*. J Bacteriol 113:922–931

Duggleby CJ, Williams PA (1986) Purification and some properties of the 2-hydroxy-6-oxohepta-2,4-dienoate hydrolase (2-hydroxymuconic semialdehyde hydrolase) encoded by the TOL plasmid pWW0 from *Pseudomonas putida* mt-2. J Gen Microbiol 132:712–726

Franklin FCH, Bagdasarian M, Bagdasarian MM, Timmis KN (1981) Molecular and functional analysis of the TOL plasmid pWW0 from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring *meta* cleavage pathway. Proc Natl Acad Sci USA 78:7458–7462

Ghosal D, You I-S, Gunsalus IC (1987) Nucleotide sequence and expression of gene *nahH* of plasmid NAH7 and homology with gene *xylE* of TOL plasmid pWW0. Gene 55:19–28

Harayama S, Rekik M (1989) A simple procedure for transferring genes cloned in *Escherichia coli* vectors into other Gram-negative bacteria: phenotypic analysis and mapping of TOL plasmid gene *xylK*. Gene 78:19–27

Harayama S, Timmis KN (1989) Catabolism of aromatic hydrocarbons by *Pseudomonas*. In: Hopwood DA, Chater KF (ed) Genetics of bacterial diversity. Academic Press, New York, pp 151–174

Harayama S, Engström P, Wolf-Watz H, Iino T, Hazelbauer GL (1982) Cloning of *trg*, a gene for a sensory transducer in *Escherichia coli*. J Bacteriol 152:372–383

Harayama S, Lehrbach PR, Timmis KN (1984) Transposon mutagenesis analysis of *meta*-cleavage pathway operon genes of the TOL plasmid of *Pseudomonas putida* mt-2. J Bacteriol 160:251–255

Harayama S, Rekik M, Timmis KN (1986a) Genetic analysis of a relaxed substrate specificity aromatic ring dioxygenase, toluate 1,2-dioxygenase, encoded by TOL plasmid pWW0 of *Pseudomonas putida*. Mol Gen Genet 202:226–234

Harayama S, Leppik RA, Rekik M, Mermod N, Lehrbach PR, Reineke W, Timmis KN (1986b) Gene order of the TOL catabolic plasmid upper pathway operon and oxidation of both toluene and benzyl alcohol by the *xylA* product. J Bacteriol 167:455–461

Harayama S, Rekik M, Wasserfallen A, Bairoch A (1987a) Evolutionary relationships between catabolic pathways for aromatics: conservation of gene order and nucleotide sequences of catechol oxidation genes of pWW0 and NAH7 plasmids. Mol Gen Genet 210:241–247

Harayama S, Mermod N, Rekik M, Lehrbach PR, Timmis KN (1987b) Roles of the divergent branches of the *meta*-cleavage pathway in the degradation of benzoate and substituted benzoates. J Bacteriol 169:558–564

Harayama S, Rekik M, Wubbolts M, Rose K, Leppik RA, Timmis KN (1989a) Identification of five genes and their products in the upper pathway operon of TOL plasmid pWW0 from *Pseudomonas putida*. J Bacteriol 171:5048–5055

Harayama S, Rekik M, Ngai K-L, Ornston LN (1989b) Physically associated enzymes produce and metabolize 2-hydroxy-2,4-dienoate, a chemically unstable intermediate formed in catechol metabolism via *meta*-cleavage in *Pseudomonas putida*. J Bacteriol 171:6251–6258

Inouye S, Nakazawa A, Nakazawa T (1981) Molecular cloning of TOL genes *xylB* and *xylE* in *Escherichia coli*. J Bacteriol 145:1137–1143

Inouye S, Nakazawa A, Nakazawa T (1984) Nucleotide sequence of the promoter region of the *xylDEGF* operon on TOL plasmid of *Pseudomonas putida*. Gene 29:323–330

Inouye S, Nakazawa A, Nakazawa T (1986) Nucleotide sequence of the regulatory gene *xylS* on the *Pseudomonas putida* TOL plasmid and identification of the protein product. Gene 44:235–242

Inouye S, Nakazawa A, Nakazawa T (1988) Nucleotide sequence of the regulatory gene *xylR* of the TOL plasmid from *Pseudomonas putida*. Gene 66:301–306

Irie S, Doi S, Yorifuji T, Takagi M, Yano K (1987) Nucleotide sequencing and characterization of the genes encoding benzene oxidation enzymes of *Pseudomonas putida*. J Bacteriol 169:5174–5179

Keil H, Keil S, Williams PA (1987) Molecular analysis of regulatory and structural *xyl* genes of the TOL plasmid pWW53-4. J Gen Micobiol 133:1149–1158

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Mermod N, Lehrbach PR, Reineke W, Timmis KN (1984) Transcription of the TOL plasmid toluate catabolic pathway operon of *Pseudomonas putida* is determined by a pair of coordinately and positively regulated overlapping promoters. EMBO J 3:2461–2466

Nakai C, Hori K, Kagamiyama H, Nakazawa T, Nozaki M (1983a) Purification, subunit structure, and partial amino acid sequence of metapyrocatechase. J Biol Chem 258:2916–2922

Nakai C, Kagamiyama H, Nozaki M, Nakazawa T, Inouye S, Ebina Y, Nakazawa A (1983b) Complete nucleotide sequence of the metapyrocatechase gene on the TOL plasmid of *Pseudomonas putida* mt-2. J Biol Chem 258:2923–2928

Nakazawa T, Inouye S, Nakazawa A (1980) Physical and functional mapping of RP4-TOL plasmid recombinants: analysis of insertion and deletion mutants. J Bacteriol 144:222–231

Neidle EL, Shapiro MK, Ornston LN (1987) Cloning and expression in *Escherichia coli* of *Acinetobacter calcoaceticus* genes for benzoate degradation. J Bacteriol 169:5496–5503

Osborne DJ, Pickup RW, Williams PA (1988) The presence of two complete homologous *meta* pathway operons on TOL plasmid pWW53. J Gen Microbiol 134:2965–2975

Patel TR, Gibson DT (1974) Purification and properties of (+)-*cis*-naphthalene dihydrodiol dehydrogenase of *Pseudomonas putida*. J Bacteriol 119:879–888

Reineke W, Knackmuss H-J (1978) Chemical structure and biodegradativity of halogenated aromatic compounds. Substituent effects on 1,2-dioxygenation of benzoic acid. Biochim Biophys Acta 542:412–423

Reiner AM (1972) Metabolism of aromatic compounds in bacteria. Purification and properties of the catechol-forming enzyme, 3,5-cyclohexadiene-1,2-diol-1-carboxylic acid (NAD^+) oxidoreductase (decarboxylating). J Biol Chem 247:4960–4965

Shaw LE, Williams PA (1988) Physical and functional mapping of two cointegrate plasmids derived from RP4 and TOL plasmid pDK1. J Gen Microbiol 134:2463–2474

Yamaguchi M, Fujisawa H (1978) Characterization of NADH-cytochrome *c* reductase, a component of benzoate 1,2-dioxygenase system from *Pseudomonas arvilla* C-1. J Biol Chem 253:8848-8853

Yamaguchi M, Fujisawa H (1980) Purification and characterization of an oxygenase component in benzoate 1,2-dioxygenase system from *Pseudomonas arvilla* C-1. J Biol Chem 255:5058–5063

Yamaguchi M, Fujisawa H (1982) Subunit structure of oxygenase component in benzoate-1,2-dioxygenase system from *Pseudomonas arvilla* C-1 . J Biol Chem 257:12497–12502

Communicated by K. Isono

Received June 30, 1989 / November 24, 1989