The 4-hydroxy-2-oxovalerate aldolase and acetaldehyde dehydrogenase (acylating) encoded by the *nahM* and *nahO* genes of the naphthalene catabolic plasmid pWW60-22 provide further evidence of conservation of *meta*-cleavage pathway gene sequences

Alison Platt,¹ Victoria Shingler,² Stephen C. Taylor³
and Peter A. Williams¹

Author for correspondence: Peter A. Williams. Tel: +44 1248 382363. Fax: +44 1248 370731.
e-mail: P.A.WILLIAMS@BANGOR.AC.UK

¹ Department of Biological Sciences, University of Wales, Bangor, Gwynedd LL57 2UW, UK

² Department of Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden

³ Zeneca BioProducts, PO Box 2, Belasis Avenue, Billingham, Cleveland TS23 1YN, UK

We report the complete nucleotide sequence and over-expression of the *nahOM* genes for the acetaldehyde dehydrogenase (acylating) and the 4-hydroxy-2-oxovalerate aldolase from the *meta* pathway operon of the naphthalene catabolic plasmid pWW60-22 from *Pseudomonas* sp. NCIMB9816. Additional partial sequence analysis of adjacent DNA shows the gene order within the operon to be *nahNLOMK*, identical to the order found for the isofunctional genes in the *meta* pathway operons in the toluene/xylene pathway of TOL plasmid pWW0 and the phenol/methylphenol pathway of pVI150. The deduced amino acid sequences of NahO and NahM were significantly homologous to the equivalent enzymes encoded by other *Pseudomonas meta* pathways, although both were the most divergent in each comparison. The *nahOM* genes were inserted downstream of the T7 promoter in the expression vector pET3a and similar constructs were also made of the isofunctional regions from pVI150 (*dmpFG*) and TOL plasmid pDK1 (*xyIQK*). High expression of all three gene pairs was detected by enzyme assays and by SDS-PAGE.

Keywords: acetaldehyde dehydrogenase (acylating), 4-hydroxy-2-oxovalerate aldolase, *Pseudomonas putida* strain NCIMB9816, *nab* genes, naphthalene catabolism

INTRODUCTION

The *meta*-cleavage pathway from catechol to pyruvate and acetyl-CoA (Dagley & Gibson, 1965; Sala-Trepat & Evans, 1971; Shingler *et al.*, 1992) is the terminal reaction sequence in the catabolism by *Pseudomonas* strains of a wide range of aromatic compounds including toluene and the xylenes (Worsey & Williams, 1975), naphthalene (Davies & Evans, 1964), phenol and the methylphenols

(Sala-Trepat *et al.*, 1972) and biphenyl (Catelani *et al.*, 1971). In all these pathways, the *meta* pathway genes are located at the end of an operon in which the first gene(s) encodes the conversion of a stable compound into catechol or a substituted catechol. Not only is the order of the *meta* pathway genes in these operons conserved (Table 1) but hybridization and nucleotide sequence analysis show that the individual genes share considerable homology (Assinder & Williams, 1988; Harayama *et al.*, 1987; Osborne *et al.*, 1988; Harayama & Rekik, 1990; Shingler *et al.*, 1992; Carrington *et al.*, 1994). The implication of this is that the complete *meta* pathway sequences share a common evolutionary origin and that the operons in which they reside have been formed in a modular manner by recombination between the common sequence and the unrelated upstream gene(s) (Cane & Williams, 1986;

Abbreviations: ADA, acetaldehyde dehydrogenase (acylating); HOA, 4-hydroxy-2-oxovalerate aldolase; LDH, lactate dehydrogenase.

The GenBank accession number for the sequence reported in this paper is U13232.

Table 1. Nomenclature of *meta* pathway genes

The genes in the *meta* pathway operons (columns 1–4) are presented in the order in which they occur. The *bph/tod* genes (columns 5–7) for the hydratase, acylating dehydrogenase and aldolase are as in the *meta* pathway operons but the other genes are in a different order.

Protein	Gene						
	<i>nah</i>	<i>dmp</i>	<i>xyl</i>	<i>bph</i> *	<i>bph</i> †	<i>bph</i> ‡	<i>tod</i>
Ferredoxin-like protein	<i>nahT</i>	<i>dmpQ</i>	<i>xylT</i>	<i>bphT</i>	–	–	
Catechol 2,3-dioxygenase	<i>nahH</i>	<i>dmpB</i>	<i>xylE</i>	<i>bphE</i>	<i>bphC</i> §	<i>bphC</i> §	<i>todE</i>
2-Hydroxymuconic semialdehyde dehydrogenase	<i>nahI</i>	<i>dmpC</i>	<i>xylG</i>	<i>bphG</i>	–	–	–
2-Hydroxymuconic semialdehyde hydrolase	<i>nahN</i>	<i>dmpD</i>	<i>xylF</i>	<i>bphF</i>	<i>bphD</i> §	<i>bphD</i> §	<i>todF</i>
2-Oxopent-4-enoate hydratase	<i>nahL</i>	<i>dmpE</i>	<i>xylJ</i>	<i>bphJ</i>	<i>bphE</i>	<i>bphH</i>	<i>todG</i>
Acetaldehyde dehydrogenase (acylating)	<i>nahO</i>	<i>dmpF</i>	<i>xylQ</i>	<i>bphQ</i>	<i>bphG</i>	<i>bphJ</i>	<i>todI</i>
4-Hydroxy-2-oxovalerate aldolase	<i>nahM</i>	<i>dmpG</i>	<i>xylK</i>	<i>bphK</i>	<i>bphF</i>	<i>bphI</i>	<i>todH</i>
4-Oxalocrotonate decarboxylase	<i>nahK</i>	<i>dmpH</i>	<i>xylI</i>	<i>bphI</i>	–	–	–
4-Oxalocrotonate isomerase	<i>nahJ</i>	<i>dmpI</i>	<i>xylH</i>	<i>bphH</i>	–	–	–

* The biphenyl *meta* pathway operon for catabolism of the benzoate formed from biphenyl by the upper pathway (Carrington *et al.*, 1994).

† The biphenyl 'upper' pathway operon from *Pseudomonas* KKS102 (Kikuchi *et al.*, 1994).

‡ The biphenyl 'upper' pathway operon from *Pseudomonas* LB400 (Hofer *et al.*, 1994).

§ These enzymes are specific for the substrates of the upper pathway but show some homology with the analogous enzymes in the *meta* pathway.

Fig. 1. Reactions catalysed by 4-hydroxy-2-oxovalerate aldolase (NahM) and acetaldehyde dehydrogenase (acylating) (NahO).

(Davies & Evans, 1964). pWW60 underwent two spontaneous structural modifications, the first during conjugational transfer from NCIMB9816 into *P. putida* PaW340, the second being a deletion which resulted in the expression of an otherwise cryptic *meta* pathway operon (Cane & Williams, 1982, 1986). This paper reports the cloning and functional analysis of the *nahOM* genes and their products from plasmid pWW60-22. These genes encode enzymes catalysing the two terminal reactions of the *meta* pathway, namely 4-hydroxy-2-oxovalerate aldolase (HOA) and acetaldehyde dehydrogenase (acylating) (ADA), respectively (Fig. 1).

METHODS

Bacterial strains and plasmids. The strains and plasmids used in this work are listed in Table 2 and the extent of the DNA present in key constructs is illustrated in Fig. 2. Strains were maintained on nutrient agar or Sensitivity Test agar (LabM) supplemented with appropriate antibiotics at concentrations

Harayama *et al.*, 1987; Assinder & Williams, 1988; van der Meer *et al.*, 1992; Williams & Sayers, 1994).

The pWW60-22 plasmid was originally derived from pWW60, the plasmid resident in *P. putida* NCIMB9816, in which the naphthalene catabolic pathway was first studied

Table 2. Bacterial strains and plasmids

Strain or plasmid	Relevant properties	Reference/source
<i>Pseudomonas</i>		
<i>Pseudomonas</i> sp. NCIMB9816	Wild-type naphthalene degrader	Assinder & Williams (1988); Cane & Williams (1982, 1986)
<i>P. putida</i> PaW130	Plasmid-free derivative of <i>P. putida</i> mt-2	Keil <i>et al.</i> (1985)
<i>P. putida</i> HS1	Wild-type toluene-xylene degrader	Kunz & Chapman (1981); Shaw & Williams (1988)
<i>E. coli</i>		
ED8654	Host for pKT230 and its derivatives	Murray <i>et al.</i> (1977)
NM522	Host for pUC18 and its derivatives	
DH5aF	Host for pBluescript(SK+) and its derivatives	
BL21(DE3)	Host for pET3a and its derivatives	
Plasmid vectors		
pBluescript(SK+)	3.0 kb Ap ^r cloning and sequencing vector	Stratagene
pBluescript(SK+)- NrNd	Replacement of <i>EcoRV</i> site with <i>NruI</i> - <i>NdeI</i> linker	This study
pET3a	4.4 kb Ap ^r T7 promoter expression vector	Rosenberg <i>et al.</i> (1987)
pKT230	11.9 kb Km ^r Sm ^r broad-host-range cloning vector	Bagdasarian <i>et al.</i> (1981)
pUC18	2.7 kb Ap ^r cloning vector	Vieira & Messing (1982)
Plasmids carrying <i>nah</i> genes		
pWW60-22	85 kb naphthalene catabolic plasmid	Cane & Williams (1986)
pWW60-3026	17.5 kb <i>Bam</i> HI fragment of pWW60-22 in pKT230	Assinder & Williams (1988)
pWW60-3043	7.2 kb <i>Xho</i> I- <i>Hind</i> III fragment of pWW60-3026 in pKT230	This study
pWW60-3046	11.8 kb <i>Xho</i> I- <i>Hind</i> III fragment of pWW60-3026 in pKT230	Assinder & Williams (1988)
pWW60-3055	4.6 kb <i>Sac</i> I- <i>Hind</i> III fragment of pWW60-3026 in pUC18	This study
pWW60-3071	0.45 kb <i>Kpn</i> I- <i>Sma</i> I fragment of pWW60-3055 in pBluescript(SK+)	This study
pWW60-3072.6	0.4 kb <i>Sma</i> I- <i>Pvu</i> II fragment of pWW60-3055 in pBluescript(SK+)	This study
pWW60-3072.9	As above but with the insert in reverse orientation	This study
pWW60-3074	66 bp <i>Nde</i> I- <i>Eco</i> RI PCR-generated fragment of <i>nabO</i> , with the <i>Nde</i> I site at the ATG start, cloned in pBluescript(SK+)-NrNd	This study
pWW60-3076.2	1.5 kb <i>Sma</i> I subclone of pWW60-3055 in <i>EcoRV</i> site of pBluescript(SK+)	This study
pWW60-3076.3	As above but with insert in reverse orientation	This study
pWW60-3078	2.3 kb <i>Kpn</i> I- <i>Pvu</i> II fragment of pWW60-3055 in pBluescript(SK+)	This study
pWW60-3079	150 bp <i>Kpn</i> I fragment of pWW60-3074 carrying 5' end of <i>nabO</i> with the introduced <i>Nde</i> I site, into the <i>Kpn</i> I site of pWW60-3078 reconstituting <i>nabOM</i>	This study
pWW60-3081	2.5 kb <i>Nde</i> I- <i>Bam</i> HI <i>nabOM</i> fragment from pWW60-3079 inserted between these sites downstream of T7 promoter in pET3a	This study
Plasmids carrying <i>xyl</i> genes		
pDK1	125 kb TOL (toluene/xylene) catabolic plasmid	Shaw & Williams (1988)
pDK2-3104	3.4 kb <i>Sst</i> I fragment of pDK1 in pKT230 spanning <i>xylQK</i>	Shaw & Williams (1988)
pDK1-3068	3.4 kb <i>Sst</i> I fragment of pDK2-3104 in pUC18	This study
pDK1-3073	200 bp <i>Nde</i> I- <i>Eco</i> RI PCR-generated fragment of <i>xylQ</i> , with the <i>Nde</i> I site at the ATG start, cloned into pBluescript(SK+)-NrNd	This study
pDK1-3077	2.4 kb <i>Eco</i> RI- <i>Xho</i> I fragment from pDK1-3068 carrying rest of <i>xylQ</i> and <i>xylK</i> inserted between <i>Eco</i> RI and <i>Sma</i> I of pDK1-3073	This study
pDK1-3080	2.6 kb <i>Nde</i> I- <i>Bam</i> HI <i>xylQK</i> fragment of pDK1-3077 ligated between these sites downstream of the T7 promoter of pET3a	This study
Plasmids carrying <i>dmp</i> genes		
pVI316Δ	3.3 kb <i>Nru</i> I- <i>Kpn</i> I fragment spanning <i>dmpFG</i> in pMMB66HEA	Shingler <i>et al.</i> (1992)
pVI150-361	220 bp <i>Nde</i> I- <i>Eco</i> RI PCR-generated fragment of <i>dmpF</i> with the <i>Nde</i> I site at the ATG start, cloned in pBluescript(SK+)-NrNd	This study
pVI150-363	2.5 kb <i>Eco</i> RI- <i>Bam</i> HI fragment of pVI316Δ cloned into pVI150-361 reconstituting <i>dmpFG</i>	This study
pVI150-365	2.7 kb <i>Nde</i> I- <i>Bam</i> HI <i>dmpFG</i> fragment of pVI150-363 cloned between those sites downstream of the T7 promoter of pET3a	This study

($\mu\text{g ml}^{-1}$) of: (for *Escherichia coli*) tetracycline, 15; streptomycin, 15; kanamycin, 15; carbenicillin, 100; (for *Pseudomonas*) streptomycin, 100. Cultures were grown and maintained at 37 °C (for *E. coli*) or 30 °C (for *Pseudomonas* spp.).

DNA manipulations. Recombinant plasmids were constructed by standard procedures. Plasmid DNA was prepared for rapid screening by the method of Holmes & Quigley (1981) and purified by CsCl centrifugation (Guerry *et al.*, 1973) or using the Qiagen method (Qiagen). Restriction digests were performed according to the suppliers instructions and were separated by agarose gel electrophoresis.

Analysis of DNA and protein sequences. Amino acid sequences were aligned by the GCG program PILEUP (Devereux *et al.*, 1984) and the distances calculated from the alignment using the PROTDIST in the PHYLIP package of programs (Felsenstein, 1993).

Nucleotide sequence determinations. Sequences of DNA inserts in either pBluescript(SK+) or pUC18 were determined with a Pharmacia DNA sequencing kit using the appropriate synthetic oligonucleotide primers. Nested deletions were generated using exonuclease III and mung bean nuclease according to the Stratagene sequencing manual. The complete sequence of both strands of the *nabOM* region was determined from the sequences of 45 individual fragments. Custom-synthesized oligonucleotide primers were used to determine the sequences across the junctions of subfragments from this region. Partial sequences outside of *nabOM* were obtained by *Sau3A* digestion of the insert of pWW60-3055 followed by shotgun cloning into the *Bam*HI site of pUC18.

Insertion of genes into high expression vector pET3a. To clone *nabOM* into pET3a, PCR was used to generate a fragment which carried the 5' end of the *nabO* gene but which was modified in the following ways: (a) the bases upstream of the ATG start codon were changed to make a *Nde*I site with a primer 5'-GGCCATATGAGCAAGAACTC, and (b) an *Eco*RI site was incorporated downstream of a *Kpn*I site which is 46 bases into the gene with the reverse primer 5'-GGACGA-ATTCATCACCAGATCG. The PCR fragment was digested with *Nde*I and *Eco*RI to give a 74 bp nucleotide which was inserted into plasmid pBluescript(SK+)-NrnD to form plasmid pWW60-3074. The sequence of the insert was determined to confirm that no mutations had been incorporated. The remainder of the *nabOM* genes 3' from the *Kpn*I site was excised from pWW60-3055 on a 2.3 kbp *Kpn*I-*Pvu*II fragment and ligated into pBluescript(SK+) to give plasmid pWW60-3078. The 5' end of *nabO*, spanning the introduced *Nde*I site, was regenerated by cloning the 150 bp *Kpn*I fragment from pWW60-3074 into the *Kpn*I site in pWW60-3078 in the correct orientation. The resulting plasmid, pWW60-3079, hence encodes *nabOM* with an *Nde*I site overlapping the ATG start codon of *nabO*. Finally, the *nabOM* genes were excised from pWW60-3079 as an *Nde*I-*Bam*HI fragment and inserted downstream of the Φ 10 promoter of pET3a to generate pWW60-3081.

The *xyI*QK and *dmp*FG gene pairs were cloned by a similar strategy using *Eco*RI sites situated 200 and 220 bp, respectively, from the ATG start codons as the distal restriction sites in the first step (see Table 2). The two PCR primers used for *xyI*QK were 5'-GGCCATATGAGCAAGAACTC and 5'-CGATGTCGCGCAATTCGGGC, and for *dmp*FG they were 5'-CCGCATATGAACCAGAACTC and 5'-GATGTCTGGCGAATTCGGGC. The resulting pET3a-based plasmids, pDK1-3080 and pVI150-365, hence have the same regulatory upstream sequences as pWW60-3081.

PCR. This was carried out using *Taq* polymerase from Promega as instructed by the supplier. Reactions were subjected to the

following cycles: 3 min at 94 °C, 5 cycles (30 s) where the temperature was increased from 50–55 °C in 1 °C per cycle, 25 cycles of 94 °C (15 s), 55 °C (30 s) and 72 °C (90 s), one cycle at 72 °C (5 min).

Cultures and extracts for enzyme assays. Cultures of *Pseudomonas* spp. were grown under forced aeration for 24 h in 2 l minimal salts medium containing 10 mM of the carbon source. *E. coli* were grown in 200 ml nutrient broth cultures containing the appropriate antibiotic. When *E. coli* BL21(DE3) harbouring a pET3a recombinant was used, the recombinant-containing strains were freshly regenerated and a lawn of transformants was washed from a plate into 5 ml of warm Luria broth (LB) which was used to inoculate 200 ml LB. Cells were grown to mid-log phase (OD_{600} 1.2; Kontron) when they were induced with 5 mM IPTG for the stated period. Cells were harvested by centrifugation at 1500 g for 30 min, washed, re-centrifuged and the pellet stored in two portions at -20 °C. Frozen cells were resuspended in 5 ml buffer and sonicated for three periods of 30 s with 15 s intervals after which the pellet was removed by centrifugation at 23000 g for 1 h at 4 °C. *E. coli* cells containing pET3a recombinants were harvested by centrifugation at 4500 g for 10 min, washed, re-centrifuged for 1 min in a microfuge and stored in five portions at -20 °C. The frozen pellet was resuspended in 0.5 ml buffer, sonicated for five periods of 8 s and the debris removed by spinning in a microfuge for 30 min. Protein concentration was determined using the standard Biuret method or the bicinchonic acid assay (Pierce) with BSA as standard.

Preparation of substrate for HOA. DL-4-Methyl-2-oxobutylolactone was custom-synthesized by Synlab Organics (London). L-(S)-4-Methyl-2-oxobutylolactone was synthesized biologically as described by Shingler *et al.* (1992). In both cases the lactone was hydrolysed to 4-hydroxy-2-oxovalerate by mild alkaline hydrolysis with stoichiometric quantities of 1 M NaOH for 14 h: only preparations with a final pH $\leq$ 7 were used.

Enzyme assays. HOA was assayed by determining the rate of pyruvate production as measured by A_{340} due to NADH oxidation in the presence of excess lactate dehydrogenase (LDH). A typical assay consisted of 50 mM HEPES buffer, pH 8 (0.94 ml), 500 mM MnCl_2 (2 μl), 10 mg NADH ml^{-1} (21 μl), LDH (25 U) and cell extract (to 1 ml). The reaction was initiated by addition of 30 μl 10 mM substrate. The rate of NADH oxidation due to NADH oxidase activity was also measured in a control reaction lacking only substrate and its value was subtracted from the assay value. ADA was assayed as previously described (Shingler *et al.*, 1992), except that (i) 0.5 M HEPES pH 8 replaced phosphate buffer and (ii) 35 μl 2 M acetaldehyde (in 3 ml) was included to provide saturating substrate under these conditions.

Determination of pyruvate. The concentration of pyruvate was measured after removal of protein by centrifugation in a Microcon 10 tube (Amicon). The filtrate (100 μl) was incubated with 50 mM HEPES buffer, pH 8 (0.875 ml), 10 mg NADH ml^{-1} (21 μl) and LDH (30 U) and A_{340} was measured.

Stereospecificity of HOA reaction. Synthetic DL-substrate (0.5 ml, 10 mM) was incubated with 0.5 ml cell extract, 0.5 mM MnCl_2 and 50 mM HEPES buffer, pH 8 (to 10 ml). Samples were removed and assayed for pyruvate production. When this had ceased, 330 μl 1% H_2O_2 was added to convert unchanged substrate to the corresponding 3-hydroxybutyrate and the mixture was then incubated at 25 °C for 3 min after which 1500 U catalase was added for 1 min. The solution was then incubated at 100 °C for 3 min. Samples (1 ml) were taken and assayed for D-3-hydroxybutyrate using the Boehringer kit according to the supplier's instructions.

Protein electrophoresis. SDS-PAGE and native gel electrophoresis was carried out by standard procedures in a Protean II cell (BioRad). For micro-sequencing of the N-terminal residues, the polyacrylamide gel was electroblotted onto a ProBlott membrane and the proteins were located on the membrane by staining with Coomassie Blue R for 1 min followed by destaining with 50% methanol. Micro-sequencing was carried out Dr M. Wilkinson, Department of Biochemistry, University of Liverpool, UK.

RESULTS

Cloning and sequencing strategy

The starting point for cloning and sequencing was plasmid pWW60-3026 (Fig. 2), the 17.5 kb insert of which carries the *meta* pathway genes including the *nabM* gene for HOA approximately located to a region of about 2–3 kb (Assinder & Williams, 1988). Recombinant plasmids carrying various small inserts from this region were constructed but aldolase assays of extracts of *Pseudomonas* and *E. coli* hosts failed to show consistent activity. However, it was shown that in the *meta* pathway

of the phenol-catabolic plasmid pVI150 from *Pseudomonas* sp. strain CF600, activity of HOA was found only when it was co-expressed with the product of the adjacent gene, ADA (Shingler *et al.*, 1992). Thus, the entire region was sequenced to locate the *nabM* gene and to identify whether a putative *nabO* gene for ADA was present in an analogous position.

The 4.6 kb *SacI*–*HindIII* region of pWW60-3043 was cloned into pUC18 to form pWW60-3055 and, after digestion with *Sau3A*, fragments were shotgun-cloned into pUC18 and sequenced on one strand (Fig. 2). In addition, the termini of subclones pWW60-3071, -3072 and -3076 were sequenced. Comparison of the sequences obtained were made with that of the complete *dmp* gene sequence from plasmid pVI150 (Shingler *et al.*, 1992). Homologies varied from 63 to 88% (Fig. 2) and showed the presence on pWW60-3055 of genes for HOA, ADA as well as for 2-hydroxymuconic semialdehyde hydrolase (*nabN*), 2-oxopent-4-enoate hydratase (*nabL*) and 4-oxalocrotonate decarboxylase (*nabK*). The gene order was deduced to be *nabNLOMK*, identical to that found for the

Fig. 2. Restriction map of the insert of pWW60-22 DNA carrying *nah meta* pathway operon in recombinant plasmid pWW60-3026 and of important subclones. Transcription of the operon is from left to right (Assinder & Williams, 1988). The orientation of the vector promoter in each construct is shown by the full arrow. The shaded boxes represent the partial single-stranded nucleotide sequences determined from shotgun subclones of plasmid pWW60-3055 digested with *Sau3A*. The direction of the arrows beside each box shows the direction of the sequence determination and the associated figure is the percentage homology with the equivalent *dmp* gene. The double-headed arrow shows the region sequenced on both strands and presented in Fig. 3. The deduced positions of the corresponding *nah* genes is shown below. Restriction sites: B, BamHI; H, HindIII; K, KpnI; P, PvuII; S, SacI; Sm, SmaI; X, XhoI.

Xpn I 100
GGTACCCCTGAGCCGGTCAAGGCCGCGACTTCATCGCGCTCGAGATCGCGGCATCGGCAGCGCCTCGGTGCGCTTCACCTGATCGAACAGAGGACAA
 NahL: V P L E P V K A G D F M R V E I G G I G S A S V R F T *
 Xpn I 200
 ACCCATGAGCAAGAACTCAAGCGGCCATAGGCCCGGCAATATCGGTACCGATCTGGTGATGAAGATGCTCGGTTCGGAGTGGATTGAGCCCGTG
 NahO: M S K K L K A A I I G P G N I G T D L V M K M L R S E W I E P V
 300
 TGGATGGTGGCATGAGCCCACTCGAGCGGCTCAAACGCGCCGCGAATTTCGGCATGAAGACACAGCCGAAGGCGTGCAGCGGCTGCTCCCGCAG
 W M V G I D P N S D G L K R A R D F G M K T T A E G V D G L L P H
 400
 TGCTGGACGACGACATCGGCTTCGAGCGCCACTCGGCTATGTGCATGCGGAGAATAGCCGCAAGCTCAACGCGCTTGGCGTGTGATGGTTCGA
 V L D D D I R I A F D A T S A Y V H A E N S R K L N A L G V L M V D
 500
 CCTGACCCCGCGCCATCGGCCCCCTACTGCGTGGCGCGGTCAACCTCAAGCAGCATGTGCGCGCCCTGGAATGAACOTCAACATGGTTCACCTGCGGC
 L T P A A I G P Y C V P P V N L K Q H V G R L E M N V N M V T C G
 Sma I
 GGCCAGGCCACCATCCCATGGTGGCGCGGTGTCCCGCGTGCAGCGGCTGGCTACCGCGAGATCGTGCCACCGCTCTCCGCGCTCGGTCGGCGCGG
 G Q A T I P M V A A V S R V Q P V A Y A E I V A T V S S R S V G P
 700
 GCAGCGCAAGAACATCGAGCGATTCAACCGCACCACCGCGCGCCATCGAGCAGGTGCGCGGCCAGGGAAGGCAAGCGATCATCGTCATCAACCC
 G T R K N I D E F T R T T A G A I E Q V G G A R E G K A I I V I N P
 800
 GGCCGAGCCCGCTGATGATGCGGACACCATCCACTGCTGACCGAGCGGACCGGCTGCGATCACCGCTTCGGTTCACCGGATGATCGCC
 A E P P L M M R D T I H C L T D S E P D Q A A I T A S V H A M I A
 900
 GAGGTGCAAGAAATACGTGCCCGCTACCGCTGAAGAACGCCCGGTGTTCGAGCGCAACCGCGTGTGATCTTCATGGAAGTCGAAGGCGCTGGCGGACT
 E V Q K Y V P G Y R L K N G P V F D G N R V S I F M E V E G L G D
 1000
 ACCTGCCCAAGTACGCGCAACCTCGACATCATGACCGCGCGCGCTGCGTACCGCGGAGATGTTCGCGGAGAAATGCGCGCGCCACCATTCAACT
 Y L P K Y A G N L D I M T A A A L R T G E M P A E E I A A G T I Q L
 1100
 GCGCGTTCGCGACATCGCGCTGGCTTGGAGTAGCACCATGAATTTGACCGCAAGAGCGTCACTGCGACGATGAGCTGCGCGAGCGGATGCGCAG
 P R R D I A L A * NahM: M N L H G K S V I L H D M S L R D G M H
 1200
 CCAAGCGCCACGATCAGCTGGAGCAGATGGTGGCGTGCACCGCGCTCGATCAAGCGGTATGCGCGTGTGATGAGATCACCCACGCGGAGCGCT
 A K R H Q I S L E Q M V A V A T G L D Q A G M P L I E I T H G D G L
 1300
 CGCGCGTGTGATCAACTACGGCTTCGCGGCCACAGTGAAGAGTACCTGCGCGCGTGTATCCGCGAGTCAAGCAGGCCAAAGTTCGCGCGCTG
 G G R S I N Y G P P A H S D E E Y L R A V I P Q L K Q A K V S A L
 1400
 CTGCTGCCCGCATCGGCACCGTCGACCACTGAAGATGGCCCTGGACTGCGCGCTCTGACTATTGCGGTGGCCACCCACTGTACCGAGGCGGATGTCT
 L L P G I G T V D H L K M A L D C G V S T I R V A T H C T E A D V
 1500
 CCGAGCAGCATCGGATGCGCGCAAGCTGGGGTGCACACCGTGGCTTCCTGATGATGGCGCATGATGACGCGCGAGAAAGTCTCGGAGCAGGC
 S E Q H I G M A R K L G V D T V G F L M M A H M I S A E K V L E Q A
 1600
 CAAGCTGATGGAAGCTATGGTGCCTGCTACTGCAACGACTCGCGCGGCTACATGCTGCGCTGATGAAGTCAAGCAGAGAAATCGCGCTTCCTGCGC
 K L M E S Y G A N C I Y C T D S A G Y M L P D E V S E K I G L R
 1700
 GCGAGCTGAACCCGCGCACCGAAGTGGCTTCCACGCGCACCAACATGGGCATGGCTATCGCCAACTGCTGGCGCATCGAAGCGGCTGCGCGC
 A E L N P A T E V G F H G H H N M G M A I A N S L A A I E A G A A
 1800
 GCATCGAGCGCTGGTGGCGCGCTCGCGCGCGTGGCGCAACCCCGTGAAGTGTTCGTGCGAGTGTGCAACGATGGCGGTGAGACCGCAT
 R I D G S V A G L G A G A G N T P L E V F V A V C K R M G V E T G I
 1900
 CGACCTGTACAAGATCATGGAGCTGGCGAGGACCTGGTGGTGGCGATGATGATCAGCGGATCCGCGTGCAGCGGAGCGCCCTGACCTGGGCTACGCC
 D L Y K I M D V A E D L V V P M M D Q P I R V D R D A L T L G Y A
 2000
 GGGGTGTACAGCTCGTTCCTGCTGTTCGCCAGCGCGCGAGAGAAATATGGCGTGTGCGCGCGGATCTGCTGGAAGTGGGCGCGCGCGCACCG
 G V Y S S F L L F A Q R A E K Y G V S A R D I L V E L G R R G T
 Sma I 2100
 TCGGTGGCCAGGAGATGATCGAAGACCTCGCCCTGGACATGGCGCGGCGCGTCAAGCAGAGGTTGAGCGCATGAACCGTACCCCTGACCCGCGAA
 V G G Q E D M I E D L A L D M A R A R Q Q Q K V S A *
 NahK: M N R T L T R E
 2180
 CAGGTGCTGGCCCTGGCGGAGCAGATCGAAAACGCGAGCTGAATGTCACGACATCGGCAAGGTGACCAACGATTTTCC
 Q V L A L A E H I E N A E L D V H D I P K V T N D Y

Fig. 3. Complete nucleotide sequence of the 2180 bp from the 3' end of *nahL* to the 5' end of *nahK* spanning *nahOM* and the deduced amino acid sequences of the gene products. The amino acid sequences are shown in one letter codes beneath the corresponding codons and the stop codons are marked with an asterisk. The position of the important restriction sites are underlined and denoted above the corresponding nucleotides.

isofunctional *meta* pathway genes (Table 1) on the phenol catabolic plasmid pV1150 (Shingler *et al.*, 1992) and on the TOL plasmid pWW0 (Harayama & Rekik, 1990, 1993).

Sequence analysis

A segment spanning 2180 bp of the *nahOM* region (Fig. 3) was completely sequenced on both strands using a combination of nested deletions and synthetic oligo-nucleotide primers. The ORF of NahO (921 bp) has 307 residues, giving a theoretical molecular mass of 32.9 kDa.

The molecular mass of the derived NahM sequence (1038 bp, 346 residues) is 37.1 kDa. The G + C content of both genes is high (65 mol %) and typical of *Pseudomonas* and the main contributory factor is the preferential use of codons with C or G in the third position (West & Igleski 1988) which, in the case of *nahO*, amounted to 88.4 mol % and, for *nahM*, 88.9 mol %. Upstream of *nahO* is the 3' end of an ORF of 27 codons, giving 25 amino acid residues which were identical to the C-terminus of DmpE and which we suggest is part of *nahL*. Similarly, downstream of *nahM* is the 5' end of an overlapping ORF of 34

Fig. 4. SDS-PAGE of extracts of *E. coli* BL21(DE3) containing recombinant plasmids carrying adjacent genes for ADA and HOA proteins coupled to T7 promoter in vector pET3a. Gels were stained with Coomassie Blue. Lanes: 1, pET3a, 6 µg protein; 2, pWW60-3081, 3 µg protein; 3, pVI150-365, 6 µg protein; 4, purified DmpFG, 2.5 µg; 5, pVI150-365, 10 µg protein; 6, pDK1-3080, 10 µg protein; 7, pWW60-3081, 10 µg protein; 8, pVI150-365, set as standard activity of HOA; 9, pDK1-3080, diluted to give same HOA activity as lane 8; 10, pWW60-3081, diluted to give same ADA activity as lane 8; 11, pWW60-3081, diluted to give same HOA activity as lane 8. Molecular mass markers are in lanes labelled M. The molecular masses of the subunits (determined from the gel/predicted from their sequences) are: NahO, 35.6/32.9; DmpF, 34.2/32.7; XylQ, 36.5/32.9; NahM, 36.5/37.1; DmpG, 38.5/37.5; XylK, 38.2/38.0.

codons, giving a protein sequence 30 residues of which are identical to the N-terminus of DmpH and which we suggest is *nabK*.

Construction of plasmids for over-expressing HOA and ADA proteins

Because of difficulties of expressing the aldolase activity from small subclones of pWW60-3026 (see above), we used the deduced sequence to construct strains which over-expressed this enzyme. For comparison, similar constructs were made using the isofunctional genes from the TOL plasmid pDK1 (Kunz & Chapman, 1981; Shaw & Williams, 1988) and the phenol/methylphenol catabolic plasmid pVI150 (Shingler *et al.*, 1992). The nucleotide sequence of the relevant part of pDK1 was provided by R. C. Benjamin (personal communication). Because of the demonstrated physical association of the HOA and ADA proteins (DmpG and DmpF) from pVI150 (Powlowski *et al.*, 1993) and the need to co-express the two genes to get activity (Shingler *et al.*, 1992), we chose to construct plasmids containing the adjacent gene pairs, *nabOM* from pWW60-22, *xylQK* from pDK1 and *dmpFG* from pVI150. The expression vector chosen was pET3a (Rosenberg *et al.*, 1987) which contains the $\Phi 10$ promoter from phage T7 upstream of a unique *NdeI* site. When present in host strain *E. coli* BL21(DE3), with a chromosomal copy of the T7 polymerase gene under control of the *lacUV5* promoter, genes cloned with their ATG start codon as part of the *NdeI* site are strongly induced by IPTG. This was possible for all three gene pairs, *nabOM*, *dmpFG* and *xylQK*, because there is no other *NdeI* site within these genes.

Electrophoretic analysis of expressed HOA and ADA proteins

Extracts of *E. coli* BL21(DE3) carrying plasmids pWW60-3081, pDK1-3080 or pVI150-365 were run on SDS-polyacrylamide gels. In each case, the extracts contained

two novel proteins of approximately the same molecular mass as predicted from the deduced amino acid sequences of the corresponding HOA and ADA proteins (Fig. 4). The bands for NahO and NahM were consistently more intense than those for XylQ/XylK or DmpF/DmpG, suggesting that there was more soluble protein in the extracts, correlating with the higher specific activities measured (see below).

N-terminal sequences of NahM and NahO proteins

To confirm the amino acid sequences deduced from the nucleotide sequences, blots of SDS-PAGE gels were made and the N-terminal sequences determined. The sequence of NahO (ADA) was found to be SKKLKAAII suggesting that the methionine has been removed, whereas the sequence of NahM (HOA) was MNLHGK?V. Residue 7 could not be clearly identified in the latter case because of the small amounts of protein on the membrane.

Enzyme assay of HOA and ADA activities

Two substrates were available for direct enzyme assay of aldolase activity. Both were prepared by mild alkaline hydrolysis of 4-methyl-2-oxobutylolactone which in one case was chemically synthesized and in the other was prepared biologically from catechol using treated extracts of phenol-grown strain CF600. Even under optimal conditions, the yield of pyruvate obtained by extensive hydrolysis of the biologically produced substrate by active extracts containing HOA was always $\leq 75\%$ of the expected theoretical yield. This was interpreted as indicating that either the hydrolysis was incomplete or that the substrate was unstable once it had been formed in the hydrolysis.

Using extracts of cells carrying a range of recombinant plasmids expressing HOA, the rate of conversion of the synthetic substrate was consistently less (10–30%) than that found with the biological substrate (data not shown).

Table 3. Specific activities of HOA and ADA in extracts of cells harbouring recombinant plasmids

Plasmid	Host*	Inducer	Specific activity (mU mg ⁻¹)†		Ratio (ADA:HOA)
			ADA	HOA‡	
pWW60-3026	<i>E. coli</i>	None	20	20	1.0
	<i>P. putida</i>	None	ND	21	
		Salicylate	ND	260	
pWW60-3046	<i>E. coli</i>	None	25	20	1.3
	<i>P. putida</i>	None	ND	25	
		Salicylate	ND	250	
pWW60-3043	<i>E. coli</i>	None	15	15	1.0
pWW60-3055	<i>E. coli</i>	None	4	21	0.2
pWW60-3081	<i>E. coli</i>	None	20 000	2300	8.9
		IPTG (4 h)	9600	460	2.1
		IPTG (6 h)	13 000	1700	8.1
pVI150-365	<i>E. coli</i>	None	2500	470	5.2
		IPTG (4 h)	510	120	4.3
		IPTG (6 h)	1200	280	4.4
pDK1-3068	<i>E. coli</i>	None	23	46	0.5
pDK1-3080	<i>E. coli</i>	None	1600	340	4.6
		IPTG (4 h)	1300	230	5.4
		IPTG (6 h)	1300	310	4.2

ND, Not determined.

* *E. coli* strains ED8654, NM522 or BL21(DE3) as appropriate (see Methods); *P. putida* PaW130.

† Mean of three assays.

‡ With biological substrate.

The specific activities of both HOA and ADA expressed from the *nabOM* genes of pWW60-3081 were considerably higher than those of the other recombinants carrying the same region of pWW60-22 (Table 4). However, the HOA activity was only 10-fold greater than the activity induced by salicylate in *P. putida* carrying pWW60-3026, the insert of which has the entire *meta* pathway operon of pWW60-22 in addition to the *nabR* gene, the product of which recognizes salicylate as co-inducer (Assinder & Williams, 1988).

The specific activities of HOA or ADA in extracts of *E. coli* BL21(DE3) harbouring any of the pET3a-based recombinant plasmids were depressed rather than stimulated by growth in the presence of IPTG (Table 4). This result was unexpected since the host strain encodes T7 polymerase under the control of the IPTG-inducible *lacUV5* promoter. In this system, low-level transcription from this promoter does occur even in the absence of IPTG resulting in some transcription from the promoter on pET3a. We assume that, in cells grown in the presence of IPTG, expression from the T7 promoter of the pET3a-based constructs was so high that the proteins formed insoluble inclusion bodies; these were indeed observed in intact cells under the light microscope.

The specific activities of both HOA and ADA in extracts of strain BL21(DE3) containing plasmid pWW60-3081 were consistently higher than that of the two equivalent

pVI150-365 or pDK1-3080 constructs (Table 4). Because the regulatory signals and vectors were identical in the three constructs it is likely that their rates of transcription were comparable, if not identical. The explanation for the higher activity of the pWW60-3081 construct must be either that the two *nab* enzymes have intrinsically greater turnover numbers or that they form inclusion bodies to a lesser extent under the conditions of high expression.

The ratio of ADA:HOA activity from all the pET3a recombinants was 5–10-fold higher than those from *dmpFG* and *xylQK* (Table 4). This would not be expected if the proteins formed a stoichiometric complex as has been shown for DmpF and DmpG (Powlowski *et al.*, 1993). This ratio of activities was also found to vary with the length of time between induction by IPTG and harvesting the cells (Table 3). In cultures carrying pWW60-3081, the activity ratio of ADA:HOA was up to 21 in cells harvested 4 h after induction, but 6 h after induction the ratio was only 8. A possible explanation is that the HOA activity is dependent upon the formation of a complex with the ADA, but not *vice versa*, and as a consequence ADA can be active either outside of a complex or in a complex with inactive HOA molecules. The higher relative intensity of the ADA protein band over that of HOA in extracts (Fig. 4) suggested that excess ADA could account for this variability in the pET3a expression system.

Table 4. Distances between derived amino acid sequences from *Pseudomonas meta* pathways

ADA protein sequences	BphG*	TodI†	XylQ‡	XylQ§	DmpF	NahO¶
BphG*	0.000	0.142	0.232	0.242	0.235	0.597
TodI†		0.000	0.262	0.271	0.269	0.620
XylQ‡			0.000	0.037	0.097	0.539
XylQ§				0.000	0.094	0.539
DmpF					0.000	0.541
NahO¶						0.000
HOA protein sequences	BphF*	TodH†	XylK‡	XylK§	DmpG	NahM¶
BphF*	0.000	0.140	0.295	0.249	0.226	0.636
TodH†		0.000	0.310	0.256	0.248	0.692
XylK‡			0.000	0.084	0.185	0.610
XylK§				0.000	0.127	0.610
DmpG					0.000	0.625
NahM¶						0.000

* From *Pseudomonas* strain KKS102 (accession no. D16407).

† From *P. putida* F1 (accession no. U09250).

‡ From TOL plasmid pDK1 (R. Benjamin, personal communication).

§ From TOL plasmid pWW0 (accession no. M94186).

|| From plasmid pVI150 (accession no. X60835).

¶ This study.

The HOA expressed by pWW60-3081 showed a sharp pH optimum of 8.5 and was stimulated about 1.6-fold at pH 8 by the addition of 1 mM Mn²⁺ and to a lesser extent (1.3-fold) by 1 mM Mg²⁺. Overnight dialysis against 5 mM EDTA reduced the activity by about twofold.

Stereospecificity of the HOA reaction

The absolute configuration of the C-4 of the 4-hydroxy-2-oxovalerate substrate was determined through the complete conversion of the racemic synthetic substrate with a cell extract. The amount of pyruvate formed was equivalent to 38% conversion (or 76% of one enantiomer). The reaction mixture was then incubated with H₂O₂ to decarboxylate the remaining substrate to 3-hydroxybutyrate without loss of the configuration at C-4. Excess H₂O₂ was removed with catalase and the reaction was assayed with D-3-hydroxybutyrate dehydrogenase. The amount of D-3-hydroxybutyrate corresponded to between 48 and 49.4% of the original racemic substrate present in the reaction. In a control experiment in which the cell extract was excluded, no pyruvate was produced but the quantity of D-3-hydroxybutyrate produced was the same. The enzyme therefore appears to attack only the L-enantiomer.

DISCUSSION

Detailed genetic analysis of *Pseudomonas meta* pathway operons has shown that the gene order is conserved downstream from the gene for a ferredoxin-like protein

(*xylT*, *dmpO*, *nabT*) and the equivalent genes show high levels of homology (Ghosal *et al.*, 1987; Harayama *et al.*, 1987; Horn *et al.*, 1991; Shingler *et al.*, 1992; Harayama & Rekik, 1993; Carrington *et al.*, 1994; Lau *et al.*, 1994). The most detailed such comparison has been with the *xyl* genes from TOL plasmid pWW0 (Nakai *et al.*, 1983; Horn *et al.*, 1991; Harayama & Rekik, 1993) and the *dmp* genes from phenol plasmid pVI150 (Bartilson & Shingler, 1989; Nordlund & Shingler, 1990; Shingler *et al.*, 1992), the close identity of which has been demonstrated by their complete nucleotide sequence. Co-linearity of DNA hybridization has also been used for the *meta* pathway operons of naphthalene plasmid pWW60-22 (Assinder & Williams, 1988) and biphenyl plasmid pWW110 (Carrington *et al.*, 1994). From sequence analysis, the present study has added additional evidence for the conserved order and homology of the *meta* pathway genes on pWW60-22 by demonstrating that the *nabNLOMK* genes are both homologous to and in the same order as the equivalent *dmpDEFGH* and *xylFJQKI*. This finding is in contrast to the reported order for *nab* genes on plasmid NAH7, as determined by transposon mutagenesis, where *nabJ* appeared to be upstream of *nabK* (Yen & Gunsalus, 1982). Our results confirm the presence in the *nab* pathway in strain NCIMB9816 of genes for (i) 2-hydroxymuconic semialdehyde hydrolase (*nabN*), the presence of which within the same regulon had been in some doubt from earlier biochemical studies (Catterall *et al.*, 1971), and (ii) ADA (*nabO*) which has only recently been shown to be part of the *meta* pathway (Shingler *et al.*, 1992). In these three pathways (*nab*, *dmp*, *xyl*) there is now

incontrovertible evidence that the *meta* pathway genes share a common ancestry and this adds weight to the idea that the lower pathway operons have evolved by recombination between this common sequence and the gene(s) which is upstream of them.

Comparison of the deduced amino acid sequences of NahO and NahM with those for other ADA and HOA proteins in the databases shows them to be closely related (Table 4). In the comparisons (data not shown), 150 identical residues are shared between the HOA proteins and between the ADA proteins, and in both sets of comparisons the identity is fairly evenly spread over the entire length of the sequences. The exception is NahO which appears to have suffered a deletion of about 18 amino acids close to its C terminus compared to the other homologous proteins and also has an extended C terminus of 9–11 residues.

The high degree of identity of the *meta* operon enzymes (NahOM, XylQK, DmpFG) with enzymes from the *bph* upper pathway operon and the *tod* operon suggests that other recombination events involving the co-acquisition of genes as a 'module' may have taken place in the evolution of aromatic pathways. In biphenyl catabolism, hydrolysis of the ring cleavage product of the first ring releases 2-oxopent-4-enoate, which is also a metabolite of the *meta* pathway being the substrate for the hydratase (XylJ, DmpE, NahL). Further catabolism of this compound requires the last three enzymes (hydratase, aldolase, acylating dehydrogenase; see Table 1) of the *meta* pathway. Recent evidence has shown that in *Pseudomonas* strains KKS102 and LB400, the operon encoding the early steps of biphenyl catabolism has also acquired homologues of these three genes (Hofer *et al.*, 1994; Kikuchi *et al.*, 1994) which are in the same order as in the *meta* pathway. However, the block of three genes are located at different positions within the upper pathway operons of the two strains. In the case of the *tod* pathway, which encodes an alternative route for toluene catabolism, the genes encoding enzymes for the primary attack on toluene appear to be very closely related to the operon for primary attack on the biphenyl molecule. In addition, this operon also possesses these three *meta* pathway genes (*todGIH*) (Table 1; Lau *et al.*, 1994) in the same order as they are found in the *meta* pathway operons but once again with different neighbours to those of the biphenyl upper pathway operons. Numerical assessment of the homologies within the family (Table 4) indicates that the NahOM proteins are the most distantly related of the pairs of isofunctional enzymes within these pathways. This suggests that the recombination event in which the *bph* and the related *tod* operons acquired the three *meta* pathway genes occurred subsequent to the assembly of the ancestral *meta* pathway operons.

In view of the homologies between the NahO and NahM proteins and the other isofunctional *meta* pathway enzymes, it is not surprising that the enzymic properties are similar. The pH optimum of around pH 8.5, the stimulation of the HOA by Mn^{2+} and Mg^{2+} , and the stereospecificity of the reaction appear to be similar to those of the previously studied enzymes (Collinsworth *et*

al., 1973; Burlingame & Chapman, 1983; Powlowski *et al.*, 1993). The *nahM*-encoded HOA also shares the same property of being able to cleave the racemic synthetic substrate at a significantly slower rate than the biological (L) enantiomer (Powlowski *et al.*, 1993), possibly because of inhibition by the D-enantiomer. A possible additional similarity is implied by our inability to co-express HOA in the absence of the ADA gene as found by Shingler *et al.* (1992), which they have shown to be due to the association of the two proteins in a complex (Powlowski *et al.*, 1993).

ACKNOWLEDGEMENTS

We would like to thank Drs Linda Shaw, Justin Powlowski and Bob Holt for their scientific and technical help, Richard Ogden for computing advice, Dr Mark Wilkinson for the N-terminal amino acid sequences and Dr R. Benjamin for releasing the sequence of the pDK1 *meta* pathway operon. This work was funded with a CASE studentship between the Science and Engineering Research Council and Zeneca Bioproducts and by the Swedish Research Councils for Natural and Engineering Sciences.

REFERENCES

- Assinder, S. J. & Williams, P. A. (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.
- Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J. & Timmis, K. N. (1981). Specific-purpose plasmid cloning vectors. 2. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**, 237–247.
- Bartilson, M. & Shingler, V. (1989). Nucleotide sequence and expression of the catechol 2,3-dioxygenase encoding gene of phenol-catabolizing *Pseudomonas* CF600. *Gene* **85**, 233–238.
- Burlingame, R. & Chapman, P. J. (1983). Stereospecificity in *meta*-fission catabolic pathways. *J Bacteriol* **155**, 424–426.
- Cane, P. A. & Williams, P. A. (1982). The plasmid-coded metabolism of naphthalene and 2-methylnaphthalene in *Pseudomonas* strains: phenotypic changes correlated with structural modification of the plasmid pWW60-1. *J Gen Microbiol* **128**, 2281–2290.
- Cane, P. A. & Williams, P. A. (1986). A restriction map of naphthalene catabolic plasmid pWW60-1 and the location of some of its catabolic genes. *J Gen Microbiol* **132**, 2919–2929.
- Carrington, B., Lowe, A., Shaw, L. E. & Williams, P. A. (1994). The lower pathway operon for benzoate catabolism in biphenyl-utilizing *Pseudomonas* sp. strain IC and the nucleotide sequence of the *bphE* gene for catechol 2,3-dioxygenase. *Microbiology* **140**, 499–508.
- Catelani, D., Sorlini, C. & Treccani, V. (1971). The metabolism of biphenyl by *Pseudomonas putida*. *Experientia* **27**, 1173–1174.
- Catterall, F. A., Sala-Trepat, J. M. & Williams, P. A. (1971). The coexistence of two pathways for the metabolism of 2-hydroxymuconic semialdehyde in a naphthalene-grown pseudomonad. *Biochem Biophys Res Commun* **43**, 463–469.
- Collinsworth, W. L., Chapman, P. J. & Dagley, S. (1973). Stereospecific enzymes in the degradation of aromatic compounds by *Pseudomonas putida*. *J Bacteriol* **113**, 933–931.
- Dagley, S. & Gibson, D. T. (1965). The bacterial degradation of catechol. *Biochem J* **95**, 466–474.
- Davies, J. I. & Evans, W. C. (1964). Oxidative metabolism of naphthalene by soil *Pseudomonas*. *Biochem J* **91**, 251–261.

- Felsenstein, J. (1993). *PHYLIP (Phylogeny Inference Package)*, version 3.5. Seattle, WA: Department of Genetics, University of Washington.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**, 387–395.
- Ghosal, D., You, I. S. & Gunsalus, I. C. (1987). Nucleotide sequence and expression of gene *nabH* of plasmid NAH7 and homology with gene *xyIE* of TOL pWW0. *Gene* **55**, 19–28.
- Guerry, P., Leblanc, D. J. & Falkow, S. (1973). General method for the isolation of plasmid deoxyribonucleic acid. *J Bacteriol* **116**, 1064–1066.
- Harayama, S. & Rekik, M. (1990). The *meta*-cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. *Mol & Gen Genet* **221**, 113–120.
- Harayama, S. & Rekik, M. (1993). Comparison of the nucleotide sequences of the *meta*-cleavage pathway genes of TOL plasmid pWW0 from *Pseudomonas putida* with other *meta*-cleavage genes suggests that both single and multiple nucleotide substitutions contribute to enzyme evolution. *Mol & Gen Genet* **239**, 81–89.
- Harayama, S., Rekik, M., Wasserfallen, A. & Bairoch, A. (1987). 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.
- Hofer, B., Backhaus, S. & Timmis, K. N. (1994). The biphenyl polychlorinated biphenyl-degradation locus (*bph*) of *Pseudomonas* sp. LB400 encodes 4 additional metabolic enzymes. *Gene* **144**, 9–16.
- Holmes, D. S. & Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* **114**, 193–197.
- Horn, J. M., Harayama, S. & Timmis, K. N. (1991). DNA-sequence determination of the TOL plasmid (pWW0) *xyl*/GFJ genes of *Pseudomonas putida* – implications for the evolution of aromatic catabolism. *Mol Microbiol* **5**, 2459–2474.
- Keil, H., Keil, S., Pickup, R. W. & Williams, P. A. (1985). Evolutionary conservation of genes coding for *meta* pathway enzymes within TOL plasmids pWW0 and pWW53. *J Bacteriol* **164**, 887–895.
- Kikuchi, Y., Yasukochi, Y., Nagata, Y., Fukuda, M. & Takagi, M. (1994). Nucleotide sequence and functional analysis of the *meta*-cleavage pathway involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. strain KKS102. *J Bacteriol* **176**, 4269–4276.
- Kunz, D. A. & Chapman, P. J. (1981). Isolation and characterization of spontaneously occurring TOL plasmid mutants of *Pseudomonas putida* HS1. *J Bacteriol* **146**, 952–964.
- Lau, P. C. K., Bergeron, H., Labbe, D., Wang, Y., Brousseau, R. & Gibson, D. T. (1994). Sequence and expression of the *todGIH* genes involved in the last 3 steps of toluene degradation by *Pseudomonas putida* F1. *Gene* **146**, 7–13.
- van der Meer, J. R., Devos, W. M., Harayama, S. & Zehnder, A. J. B. (1992). Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev* **56**, 677–694.
- Murray, N. E., Brammar, W. J. & Murray, K. (1977). Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Mol & Gen Genet* **150**, 53–61.
- Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, T., Inouye, S., Ebina, Y. & Nakazawa, A. (1983). Complete nucleotide sequence of the metapyrocatechase gene on the TOL plasmid of *Pseudomonas putida* mt-2. *J Biol Chem* **258**, 2923–2928.
- Nordlund, I. & Shingler, V. (1990). Nucleotide-sequences of the *meta*-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxymuconic semialdehyde hydrolase from *Pseudomonas* CF600. *Biochim Biophys Acta* **1049**, 227–230.
- Osborne, D. J., Pickup, R. W. & Williams, P. A. (1988). The presence of two complete homologous *meta* pathway operons on TOL plasmid pWW53. *J Gen Microbiol* **134**, 2965–2975.
- Powlowski, J., Sahlman, L. & Shingler, V. (1993). Purification and properties of the physically associated *meta*-cleavage pathway enzymes 4-hydroxy-2-ketovalerate aldolase and aldehyde dehydrogenase (acylating) from *Pseudomonas* sp strain CF600. *J Bacteriol* **175**, 377–385.
- Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J. & Studier, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA-polymerase. *Gene* **56**, 125–135.
- Sala-Trepat, J. M. & Evans, W. C. (1971). The *meta* cleavage of catechol by *Azotobacter* species. *Eur J Biochem* **20**, 400–413.
- Sala-Trepat, J. M., Murray, K. & Williams, P. A. (1972). The metabolic divergence in the *meta* cleavage of catechols by *Pseudomonas putida* NCIB10015. *Eur J Biochem* **28**, 347–356.
- Shaw, L. E. & Williams, P. A. (1988). Physical and functional mapping of two cointegrate plasmids derived from RP4 and TOL plasmid pDK1. *J Gen Microbiol* **134**, 2463–2474.
- Shingler, V., Powlowski, J. & Marklund, U. (1992). Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp strain CF600. *J Bacteriol* **174**, 711–724.
- Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113–130.
- Vieira, J. & Messing, J. (1982). The pUC plasmids, an m13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.
- West, S. E. H. & Iglewski, B. H. (1988). Codon usage in *Pseudomonas aeruginosa*. *Nucleic Acids Res* **16**, 9323–9335.
- Williams, P. A. & Sayers, J. R. (1994). The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. *Biodegradation* **5**, 195–217.
- Worsey, M. J. & Williams, P. A. (1975). Metabolism of toluene and xylenes by *Pseudomonas putida* (arvilla) mt-2: evidence for a new function of the TOL plasmid. *J Bacteriol* **124**, 7–13.
- Yen, K. M. & Gunsalus, I. C. (1982). Plasmid gene organization – naphthalene salicylate oxidation. *Proc Natl Acad Sci USA* **79**, 874–878.

Received 23 January 1995; revised 16 May 1995; accepted 22 May 1995.