Catabolism of Naphthalene and Phenanthrene by *Burkholderia* sp. Strain RP007

The University of Waikato
Te Whare Wānanga o Waikato

A thesis

submitted in partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy in Biological Sciences

at the

University of Waikato

by

Andrew David Laurie

University of Waikato

1998

Abstract

Burkholderia sp. strain RP007 was isolated for its ability to utilise phenanthrene as a sole source of carbon and energy. This strain was identified as a *Burkholderia* sp. by biochemical tests, fatty acid analysis, and nucleotide sequencing of the 16S rRNA gene. The ability of *Burkholderia* sp. strain RP007 to mineralise [9-^{14}C]phenanthrene and [1-^{14}C]naphthalene was confirmed.

A gene cluster (*phn*) which represents a new group of upper pathway genes for polycyclic aromatic hydrocarbon (PAH) degradation was cloned from *Burkholderia* sp. strain RP007. Nucleotide sequencing of an 11.5 kb *Hin*dIII fragment cloned from RP007 revealed the presence of nine open reading frames (ORFs). *Escherichia coli* carrying the recombinant *phn* genes was able to transform naphthalene to salicylic acid. The *phn* genes are significantly different in sequence and gene order from the previously described *nah*, *ndo*, *pah* and *dox* genes for PAH degradation. The *phn* locus encodes iron-sulfur protein α and β subunits of a PAH initial dioxygenase, but not the ferredoxin and reductase components. The dihydrodiol dehydrogenase of the RP007 pathway, PhnB, shows greater similarity to analogous dehydrogenases from described biphenyl pathways than to those characterised from naphthalene/ phenanthrene pathways. Furthermore, the RP007 extradiol dioxygenase, PhnC, shows no similarity to other extradiol dioxygenases for naphthalene or biphenyl oxidation, but is a member of the recently proposed Class III extradiol dioxygenases. Upstream of the *phn* catabolic genes are two putative regulatory genes *phnR* and *phnS*. *phnR* is divergently transcribed with respect to the other genes and is similar to the σ^{54}-dependent family of positive transcriptional regulators. *phnS* is a LysR-type transcriptional activator. Expression experiments suggest the *phn* operon is under regulatory control which may involve PhnR and PhnS.

Two loci encoding aromatic *meta* cleavage pathways were also cloned from *Burkholderia* sp. strain RP007. The nucleotide sequence of the catechol 2,3-dioxygenase genes, and adjacent open reading frames, suggest these are *meta* pathways of different classes which specify catabolism of different aromatic substrates. One of these may be the lower pathway for phenanthrene and naphthalene degradation.

The presence of catabolic genes homologous to the *phnAc* gene of *Burkholderia* sp. strain RP007, relative to the archetypal *nahAc* gene of *Pseudomonas putida* G7, was screened for in both culturable and nonculturable microbial communities. Of 77 environmental isolates cultured for their ability to degrade naphthalene or phenanthrene, 28 showed homology to *nahAc*, whilst none of the strains possessed catabolic genes similar to *phnAc*. However, using *phnAc*-specific primers, PCR products homologous to *phnAc* were amplified from DNA extracts of eight hydrocarbon-contaminated soils, demonstrating the widespread distribution of this genotype. Analysis of these *phnAc* PCR products by nucleotide sequencing of selected cloned products revealed very high (98.5–100%) homology to the *phnAc* gene of RP007, suggesting the *phn* genotype is highly conserved. PCR products were also amplified from soil DNA using *nahAc*-specific primers, and sequencing of these products revealed a diversity similar to that amongst previously described *nah*-like genes.

Acknowledgements

Many thanks to Dr Gareth Lloyd-Jones and Dr Jackie Aislabie of Landcare Research, and Prof. Hugh Morgan of the University of Waikato, who were my doctoral supervisors. I am especially grateful to Gareth for his continuous guidance and support both during the experimental stages of this project, and during the write-up period. Thank you also to staff at Landcare Research in Hamilton, in particular to Rhonda Fraser and David Hunter for the isolation and characterisation of bacterial strains described in Chapter 6, and also to Dave for scanning the graphic images for this thesis. I am grateful to Dr Carl E. Cerniglia at the National Centre for Toxicological Research (NCTR), Arkansas, for helping organise my stay at NCTR.

Financial support for this project was provided by the New Zealand Foundation for Research, Science and Technology, and by the New Zealand Lottery Board Commission. Travel to the American Society for Microbiology (ASM) General Meeting (4–8 May 1997) and my stay at NCTR was supported by a Claude McCarthy Fellowship awarded by the New Zealand Vice-Chancellors' Committee, and travel grants from The Royal Society of New Zealand and the New Zealand Society for Biochemistry and Molecular Biology (NZSBMB); their support is gratefully acknowledged.

Thank you to Brigette who often feigned an interest in my work as a cure for insomnia. I also thank my parents who have always been supportive.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Literature Review

1.1 Overview of PAHs

1.1.1 Sources of PAHs

Polycyclic aromatic hydrocarbons (PAHs) consist of two or more fused benzene rings in linear, angular or cluster arrangements (Figure 1.1). PAHs are formed naturally during thermal geologic reactions and volcanic eruptions, during the burning of vegetation in forest and bush fires, and also by some plant and bacterial reactions (Blumer, 1976; Menzie *et al.*, 1992). This gives rise to a general background level of PAHs in soils and sediments (Jones *et al.*, 1989b). Anthropogenic sources have dramatically increased the quantity of PAHs in the environment, with the majority emitted from fossil fuel combustion sources such as automobiles, coking plants, asphalt plants and manufacturing facilities that use fossil fuels (Freeman & Cattell, 1990, Benner *et al.*, 1990, Menzie *et al.*, 1992). PAH contamination is often associated with wood-treatment activities since PAHs are major constituents of creosote (approximately 85% PAHs by weight) and anthracene oil, which are commonly used wood-treatment pesticides (Bos *et al.*, 1984). In grossly contaminated sites the presence of PAHs is usually associated with industrial activities involving the processing, combustion, and disposal of fossil fuels or fossil fuel derived products (eg. coal tar or carbon black) (Nishioka *et al.*, 1986). Fractionation products derived from crude oil such as diesel, petroleum, fuel oil, and lubricating oil contain PAHs. The crude oil source and fractionation process used have an effect on the PAH content of the final fuel products, and higher concentrations of PAHs are associated with the higher boiling-point distillation products (Wilson & Jones, 1993). The source of PAH contamination of soil from industrial sources is often leakage from storage tanks. PAHs are also present in large quantities in municipal and district wastes leading to concern where sewage sludge is disposed to land used for agricultural or other purposes (Jones *et al.*, 1989a) (Table 1.1).

Levels of PAH contamination range from a low of 5 ng g^{-1} of soil in an undeveloped area, to 1.79×10^6 ng $g^{-1}$of soil at an oil refinery. The concentration of PAHs in marine sediments can exceed 10^5 ng g^{-1} in urban estuaries (Cerniglia, 1992).

Natural oil seeps
Refinery and oil storage wastes
Accidental spills from oil tankers and other ships
Municipal and urban wastewater discharge runoff
River-borne pollution
Atmospheric fallout of flyash particles
Petrochemical industrial effluents
Coal tar and other coal processing wastes
Automobile engine exhausts
Combustion of fossil fuels (gasoline, kerosene, coal, diesel fuel)
Smoked, charcoal broiled or pan-fried foods
Tobacco and cigarette smoke
Forest and prairie fires
Rural and urban sewage sludge
Refuse and waste incineration
Coal gasification and liquification processes
Creosote and other wood preservative wastes
Commercial and pleasure boating activities

Table 1.1 Major sources of PAHs in the environment. Adapted from Cerniglia. (1992).

1.1.2 Distribution and fate of PAHs in the environment

PAHs have been detected in soil, air, sediments and water. Due to their hydrophobic properties PAHs tend to adsorb to particles and eventually migrate to sediments in river, lake, estuarine and marine waters (Means *et al.*, 1980). Much of the PAHs in soil accumulates via atmospheric deposition. As a result of increased anthropogenic activity, soil concentrations have tended to increase in the last 100–150 years, particularly in urban areas (Jones *et al.*, 1989a; b). The widespread distribution of PAHs leads to their presence in various consumer products and food and therefore humans are exposed to these chemicals as part of everyday living.

Although many PAHs are included in the 'priority pollutants' listings produced by the US EPA and by the European Commission, there are uncertainties over the significance of PAHs in relation to soil quality and their transfer from soil into the human food chain. For non-smokers dietary intake has been identified as the principal route of human exposure to PAHs, with plant-based foodstuffs constituting roughly 50% of the total PAH intake in a typical UK diet (Jones *et al.*, 1989b).

The possible fates of PAHs in the environment include volatilisation, photo-oxidation, chemical oxidation, bioaccumulation, adsorption to soil particles, leaching, and microbial degradation (Cerniglia, 1992). Biodegradation by natural

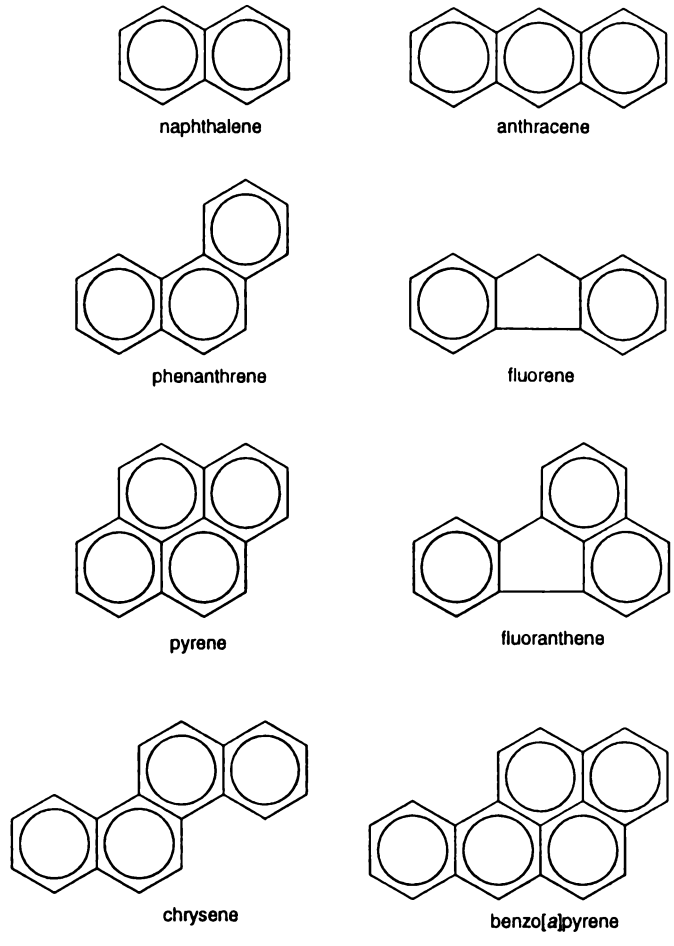


Figure 1.1 Chemical structures of some common PAHs

surface and subsurface soil microorganisms appears to be the process primarily responsible for the removal of PAHs in surface sediments and soil (Sims & Overcash, 1983). These compounds may either be completely degraded (mineralised) or partially transformed by either a microbial community or a single microorganism.

In a study of fourteen PAHs Park *et al.* (1990) found that under unsaturated soil conditions volatilisation was negligible, except for naphthalene and substituted naphthalenes. Mechanisms of abiotic loss appeared to be potentially important for two- and three-ring PAH compounds, but abiotic loss did not play a significant role in determining the fate of PAHs with more than three rings.

The biodegradability of two- and three-ring PAHs has been shown to be extensive, whereas that of four-, five-, and six-ring PAHs is considerably less significant (Herbes & Schwall, 1978).

1.1.3 Physical and chemical properties of PAHs

Each PAH has a unique set of physical and chemical properties, although PAHs in general behave in a similar manner in the environment. The stability of the PAH is indicated by the ring arrangement, linear being the most unstable and angular (rings in step) being the most stable (Blumer, 1976). Both solubility and volatility generally decreases with an increasing number of fused rings (Sims & Overcash, 1983).

The reason why PAHs are so recalcitrant in the environment lies in the inherent stability of the benzene unit. The benzene ring is a highly stable structure and many chemical reactions ultimately lead to its formation. Next to glucosyl residues the benzene ring is the most widely distributed unit of chemical structure in nature (Smith, 1990).

The ring structure of benzene was first proposed by August Kekulé in 1865, although the unusual behaviour and stability of benzene was not fully understood until the development of quantum mechanics in the 1920s. Aromatic compounds such as benzene consist of alternating single and double carbon-carbon bonds. Double bonds are formed from a sigma bond, which is a single bond and is quite strong, and a pi bond which is weak. The electrons in the pi bond are energetically unstable compared with the sigma electrons. Quantum mechanics allows the pi electrons to travel freely between adjacent carbon atoms, therefore benzene can be regarded as having pi bonds of approximately equal strength. The benzene molecule is a hybrid of the two possible configurations of alternating double and single bonds and is referred to as a resonance hybrid (Figure 1.2). Such resonance hybridisation explains the possession of a large (negative) resonance energy by benzene and related compounds which results in a high thermodynamic stability and chemical properties very different from those observed for aliphatic compounds (Aihara, 1992).

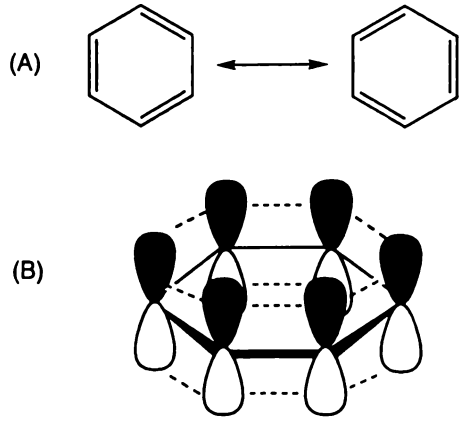


Figure 1.2 (A) The resonance structures of benzene; (B) the overlapping pi orbitals of benzene.

1.1.4 Toxicity and carcinogenicity of PAHs

The widespread presence of PAHs in the environment is of concern because of the acute toxicity of some primarily lower molecular weight PAHs and the carcinogenicity of a number of the larger PAHs. The US EPA has identified 16 unsubstituted PAHs as priority pollutants, some of which are considered to be possible or probable human carcinogens (Menzie *et al.*, 1992).

The link between PAHs and the scientific study of chemical/hydrocarbon carcinogenesis dates back to 1775 when Percival Pott attributed the occurrence of scrotal cancer in chimney sweeps to their occupational exposure to soot (Dipple *et al.*, 1984). Subsequently as new industries became established other industrial carcinogens were recognised (eg. coal tar, paraffin and certain mineral oils).

In the early 1900s, with the first application of rabbit ears, and later mouse skin, as a means of testing for a carcinogenic response, carcinogenic activity could be ascribed to defined chemical compounds. A number of PAHs were identified as carcinogenic agents (Dipple *et al.*, 1984). Lower molecular weight PAHs (up to and including three rings) and a number of larger PAHs (four or more rings) were classed as inactive carcinogens, although for many of these compounds skin papillomas had been reported and in some cases tumours had formed after very high doses. A number of other PAHs of four rings or greater were categorised as having disputed, moderate, or very high carcinogenic activity (Table 1.2). Benzo[*a*]pyrene in particular was attributed very high activity. In many cases the status of low activity

PAH	Toxicity[a]	Genotoxicity[b]	Carcinogenicity[c]
naphthalene	Rat LD_{50}=306–600 ppm	–	–
acenaphthalene	ND	+ Ames	–
anthracene	Mouse LD_{50}=430 ppm	–	–
phenanthrene	Mouse LD_{50}=700 ppm	–	–
benzo[a[pyrene	Mouse LD_{50}=250 ppm	+ Ames	+ Carcinogen, + UDS, + SCE, + CA, + DA
benz[a]anthracene	ND	+ Ames, + SCE	+ CA, + Carcinogen
fluoranthene	Mouse LD_{50}=500 ppm	+ Ames	Weak carcinogen
pyrene	Mouse LD_{50}=514–678 ppm	+/? Ames	+ UDS, + SCE

[a] ND, no data.

[b] Ames, *Salmonella typhimurium* assay; SCE, sister chromatid exchange; +/?, inadequate/inconclusive.

[c] CA, chromosomal aberrations; DA, DNA adducts; UDS, unscheduled chromosomal synthesis; SCE, sister chromatid exchange.

Table 1.2 Toxicological characteristics of selected PAHs. Adapted from Cerniglia & Heitkamp (1989).

carcinogens depends on very limited experimentation as research tends to be directed at those compounds with obviously high activity (Menzie *et al.*, 1992).

It is generally accepted that PAHs *per se* are not carcinogenicly active, but become so after metabolic activation by mammalian microsomal enzymes. PAHs are metabolised to polar oxygenated metabolites to enhance excretion from mammalian cells but which also increases carcinogenicity. The primary reaction PAHs undergo in mammalian cells is epoxidation of the aromatic ring by a cytochrome P-450 monooxygenase to initially produce a dihydrodiol. Other PAH metabolites produced in mammalian cells include epoxides, phenols and their sulfate ester conjugates, quinones, dihydrodiol epoxides, tetraols, and water soluble metabolites such as glucuronide, sulfate, and glutathione conjugates (Pothuluri & Cerniglia, 1994).

Benzo[*a*]pyrene is the most extensively studied, and the most carcinogenic PAH known. In mammalian cells oxidation of the benzo[*a*]pyrene molecule at the 7,8,9,10 postions leads to the formation of biologically active metabolites. Oxidation is highly stereospecific; at the 7,8 position (–)-*trans*-7,8-dihydroxy-7,8-dihydrobenzo[*a*]pyrene is formed which is more potent as a tumour initiator than the (+)-enantiomer. The ultimate carcinogen formed from benzo[*a*]pyrene is benzo[*a*]pyrene 7,8-dihydrodiol-9,10-epoxide-2, which may then interact with DNA, resulting in a highly mutagenic activity toward mammalian cells (Figure 1.3) (Pothuluri & Cerniglia, 1994).

The processes leading to carcinogenesis following metabolic activation of a PAH are

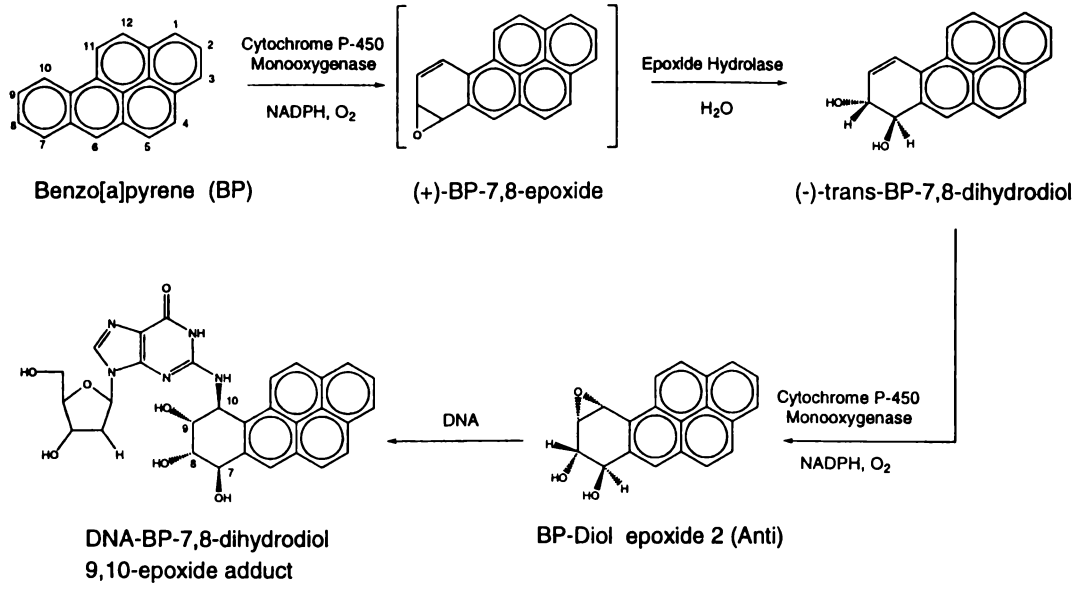


Figure 1.3 The metabolic activation of benzo[*a*]pyrene in mammalian systems. Adapted from Pothuluri & Cerniglia, 1994.

not completely understood. The oxygenated PAH derivative may interact directly with DNA to form a mutagenic adduct as shown in Figure 1.3. Some workers have suggested that the critical molecular event in carcinogenesis might be the formation of a charge-transfer complex between the carcinogen and a cellular receptor.

Other physical and chemical properties of PAHs may explain their carcinogenic activity. Many PAHs exhibit the property of photodynamic action which is the phenomenon whereby a combination of light energy and a chemical sensitiser produces effects which are not induced by either compound alone. (Dipple *et al.*, 1984). Aromatic hydrocarbons are more soluble in purine solutions or in DNA solutions than in water alone, and many researchers have studied the interactions involved in this phenomenon. Most studies support the view that the interaction with DNA involves the insertion of the hydrocarbon molecule between base pairs of DNA. Physical interaction with DNA seems to be related to the size of the hydrocarbons and bears no obvious relationship to their carcinogenic activities (Dipple *et al.*, 1984).

The toxic effect of PAHs results from the interaction of the PAH with cellular constituents other than DNA. The toxicity of the PAH is inversely proportional to the molecular weight of the congener (Table 1.2).

1.2 Bacterial degradation of aromatic compounds

Degradation of aromatic compounds by bacteria has received considerable attention due to the perceived potential for exploiting microorganisms to ameliorate environments contaminated with aromatic compounds. Most of the bacterial strains studied for their ability to degrade aromatic substrates were classified as *Pseudomonas* since this genus has historically been 'a catch-all for many obligately aerobic Gram-negative bacteria which could not be confidently assigned to other groups' (Timmis, 1997). Many *Pseudomonas* strains have been reclassified since the creation of new genera such as *Burkholderia*, *Sphingomonas*, or *Ralstonia*, but are often referred to as 'honorary pseudomonads' to reflect their initial taxonomic designations.

These *Pseudomonas*-like strains were often chosen for study because they were most amenable to culturing in the laboratory, and for studying mechanisms for aromatic degradation. Therefore, much of what is known about microbial degradation of aromatics has come from the study of pseudomonds. The following sections discuss microbial degration of aromatic substrates, but they more accurately refer to aromatic degradation by pseudomonads.

1.2.1 General principles of pathways for aerobic aromatic catabolism

The pathways for the aerobic dissimilation of aromatic substrates may be considered to consist of three parts. In the first stage the substrate undergoes changes to its substituent groups, usually through the introduction of hydroxyl groups by mono- or di-oxygenases, to produce an arene diol intermediate. These dihydroxylated intermediates are substrates for the second stage of the catabolism, the opening of the ring. This occurs by the action of dioxygenases which break one of the carbon-carbon bonds of the ring by the addition of molecular oxygen, producing an unsaturated aliphatic acid. After ring cleavage, the third stage of the catabolism consists of the conversion of the ring cleavage product to small aliphatic compounds which can directly enter central metabolic routes such as the tricarboxylic acid cycle.

In summary, aerobic aromatic catabolism consists of a variety of pathways which converge in a limited number of common intermediates, usually dihydroxylated

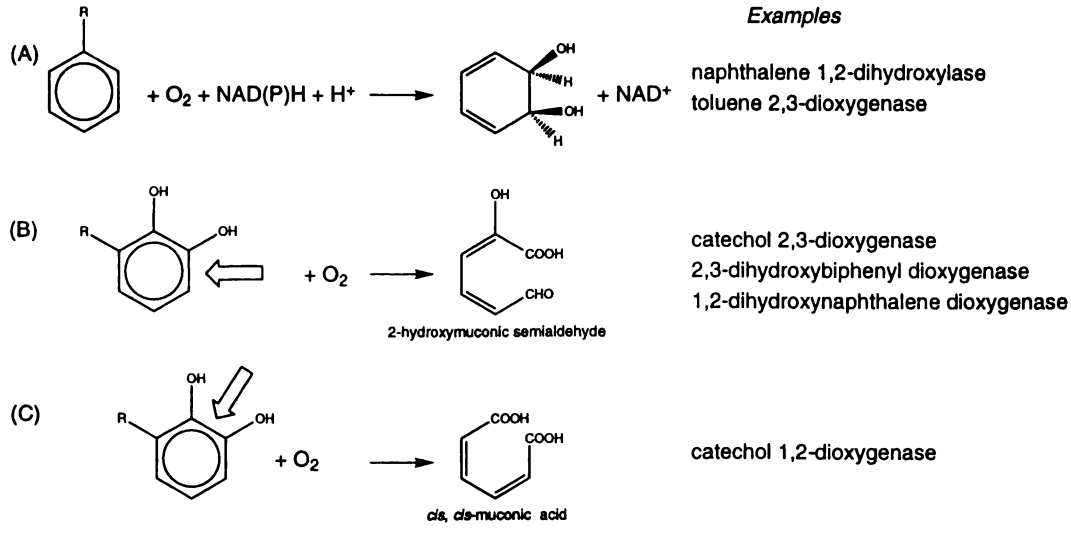


Figure 1.4 Reaction scheme for (A) aromatic ring dioxygenase, (B) extradiol (*meta*) ring cleavage by an extradiol dioxygenase and (C) intradiol (*ortho*) ring cleavage by catechol 1,2-dioxygenase. The block arrows indicate the position of ring cleavage. Adapted from Harayama *et al.*, 1992.

compounds, which are then further assimilated by a small number of common pathways (Williams & Sayers, 1994).

This section describes the biochemistry of oxygenase reactions, and reviews the microbial catabolism of toluene, xylene, and biphenyl, since microbial degradation of PAHs has much in common with these pathways.

1.2.2 Role of oxygenases in aromatic metabolism

Two types of oxygenases, the mono- and dioxygenases, play a role in aromatic catabolism by bacteria. Monooxygenases incorporate one hydroxyl group into substrates by the reduction of two atoms of dioxygen to one hydroxyl group and one H_2O molecule concomitant with oxidation of NAD(P)H. Dioxygenases catalyse the incorporation of both atoms of dioxygen into substrates, and are divided into two classes. Aromatic-ring dioxygenases incorporate two hydroxyl groups into aromatic substrates forming *cis*-diols in a reaction which requires NAD(P)H as an electron donor (Figure 1.4A). Aromatic ring-cleavage dioxygenases incorporate two atoms of dioxygen into a dihydroxylated aromatic substrate and the ring is cleaved (Figure 1.4B,C); this reaction requires no external reductant (Harayama *et al.*, 1992). Ring cleavage dioxygenases are divided into two families dependent on the position the aromatic ring is cleaved. The extradiol (or *meta*) dioxygenases are Fe(II) enzymes

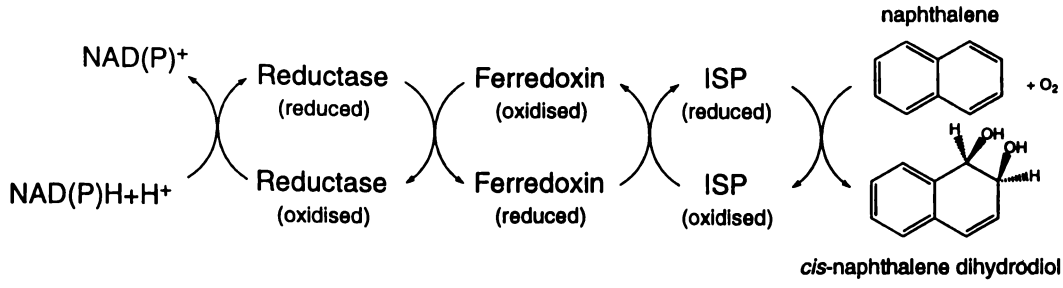


Figure 1.5 Electron transfer scheme for the oxidation of naphthalene to naphthalene *cis*-1,2-dihydrodiol by naphthalene dioxygenase, showing the functions of the reductase, ferredoxin, and ISP (iron-sulfur protein) components of naphthalene dioxygenase. Adapted from Cerniglia, 1992.

and cleave the ring at a bond proximal to one of the hydroxyl groups producing 2-hydroxymuconic semialdehyde (or an analogue) (Figure 1.4B). The intradiol (or *ortho*) dioxygenases cleave between the two hydroxyl groups and produce *cis, cis*-muconic acid (or an analogue) (Figure 1.4C), and contain a nonheme, non-iron-sulfur Fe(III) as a prosthetic group. The catabolic pathway involving *ortho* cleavage is often referred to as the β-ketoadipate pathway.

All oxygenases possess a transition metal, flavin or pteridine cofactor that interacts with dioxygen. This is because the concerted reactions between dioxygen and carbon in organic compounds are spin forbidden and the cofactor is used to overcome this restriction (Harayama *et al.*, 1992).

For the oxygenases that require the NAD(P)H cofactor, the enzyme reaction is separated into two steps: the oxidation of NAD(P)H to generate two reducing equivalents, and the hydroxylation of the substrate. Flavoprotein hydroxylases that catalyse the monohydroxylation of the aromatic ring (eg. salicylate hydroxylase) carry out these two reactions on a single polypeptide chain. In other oxygenases the NAD(P)H oxidation and the hydroxylation reaction are catalysed by two separate polypeptides that are linked by a short electron-transport chain to a cofactor on the hydroxylase component which is reduced. Dioxygen couples with the reduced cofactor and subsequently hydroxylates the substrate.

The electron transport chains associated with oxygenases contain at least two redox centres. The first redox centre is usually a flavin, while the second is an iron-sulfur cluster. The electron transport is initiated by a single two-electron transfer from NAD(P)H to a flavin, followed by two single-electron transfers from the flavin to an iron-sulfur cluster (Harayama *et al.*, 1992). For example, the formation of

naphthalene *cis*-1,2-dihydrodiol from naphthalene is catalysed by a multicomponent enzyme system, designated naphthalene dioxygenase (Ensley *et al.*, 1982). The reductase component, an iron-sulfur flavoprotein, initially accepts electrons from NADH or NAD(P)H and then transfers them to a [2Fe-2S] ferredoxin, designated ferredoxin$_{NAP}$. Ferredoxin$_{NAP}$ in turn reduces the terminal iron-sulfur protein, ISP$_{NAP}$, which catalyses the oxidation of naphthalene to naphthalene *cis*-1,2-dihydrodiol (Figure 1.5) (Cerniglia, 1992).

1.2.3 Catabolic pathways for toluene and xylene

The metabolism of toluene, the xylenes, and other substituted single-ring aromatic compounds by *Pseudomonas* spp. has been widely investigated. Of the four routes for toluene metabolism in pseudomonads that have been described, only the TOL and TOD pathways have been studied in any depth. Both pathways catalyse *meta* (extradiol) cleavage of catechol, but the mechanism by which the initial substrate is transformed to catechol differs. In the TOL pathway, toluene is metabolised to catechol by progressive oxidation of the methyl group to benzoate; this is then converted to catechol in two steps, a dihydroxylation followed by a dehydrogenation/decarboxylation. In contrast, the TOD pathway converts toluene to 3-methyl catechol via successive dihydroxylation and dehydrogenation reactions (Assinder & Williams, 1990; Williams & Sayers, 1994).

Genes of the TOL pathway, designated the *xyl* genes, are located on large plasmids, collectively called the TOL plasmids. The archetypal TOL plasmid is pWW0 (117 kb) from *P. putida* mt-2; numerous other TOL plasmids have been reported (Assinder & Williams, 1990).

The catabolic genes on pWW0 are organised into two operons, an arrangement common in many pathways. The first of the two *xyl* operons is the 'upper pathway' operon, *xylCMABN*, which encodes the enzymes for the conversion of toluene (and its alkyl analogues) to benzoate (and the alkylbenzoates) (Harayama *et al.*, 1989). The 'lower' or '*meta* pathway' operon, *xylXYZLTEGFJQKIH*, codes for the further metabolism of the benzoates to central metabolites through catechol and the *meta* pathway (Harayama & Rekik, 1990). Induction of both operons is by two regulatory proteins which are the products of *xylR* and *xylS* (Burlage *et al.*, 1989).

The *tod* genes for toluene catabolism by the TOD pathway in *P. putida* F1 are

encoded by a single chromosomal operon. The initial enzyme of the pathway, toluene dioxygenase, is encoded by the genes *todC1C2BA* which are located between the two genes for the subsequent transformations in the order *todFC1C2BADEJ* (Zylstra & Gibson, 1989). The locations of the other *meta* pathway genes, or genes for the subsequent catabolism of catechol, have not been reported, nor is the regulation known.

Pseudomonas strains carrying TOL plasmids may also possess chromosomal genes for the β-ketoadipate pathway which catalyses *ortho* cleavage of catechol. *P. putida* mt-2, when grown on benzoate as a sole carbon source, spontaneously mutates to strains which have lost the ability to grow on *m*- or *p*-toluate, but which dissimilate benzoate by the *ortho* pathway. Such mutants grow faster on benzoate than *P. putida* mt-2 because the β-ketoadipate pathway is more efficient than the *meta* pathway conferring a growth advantage on these mutants. The *meta* pathway is able to tolerate alkyl substituents on the catechol ring, unlike the β-ketoadipate pathway; thus it serves as a route for substrates which can be converted to alkylcatechols, in particular the toluates and xylenes (Assinder & Williams, 1990).

The catabolic route used to degrade aromatic substrates in strains of TOL^+ *P. putida* carrying the genetic information for both *ortho* and *meta* pathways is governed by the different regulatory mechanisms of the two pathways (Assinder & Williams, 1990). The first two steps of both pathways involve the oxidation of benzoate to catechol, after which point the two pathways diverge. Expression of the chromosomal β-ketoadipate pathway enzymes is induced by the catechol *ortho* cleavage product *cis*, *cis*-muconate, which is formed by basal levels of catechol 1,2-dioxygenase. In contrast, the plasmid-coded *meta* pathway is induced by benzoate itself. During growth on benzoate, the *meta* pathway is thus fully induced and will rapidly metabolise any catechol formed, preventing formation of *cis*, *cis*-muconate which is necessary for *ortho* pathway induction. Benzoate is therefore degraded preferentially by the *meta* route in strains where the genes for both pathways are present and functional.

The two operons on the TOL plasmids are under a complex system of regulatory control which has been the subject of rigorous investigation. The model which has emerged from extensive analysis of pWW0 (see Assinder & Williams, 1990) suggests four promoters are involved in the regulation of the two operons. These are

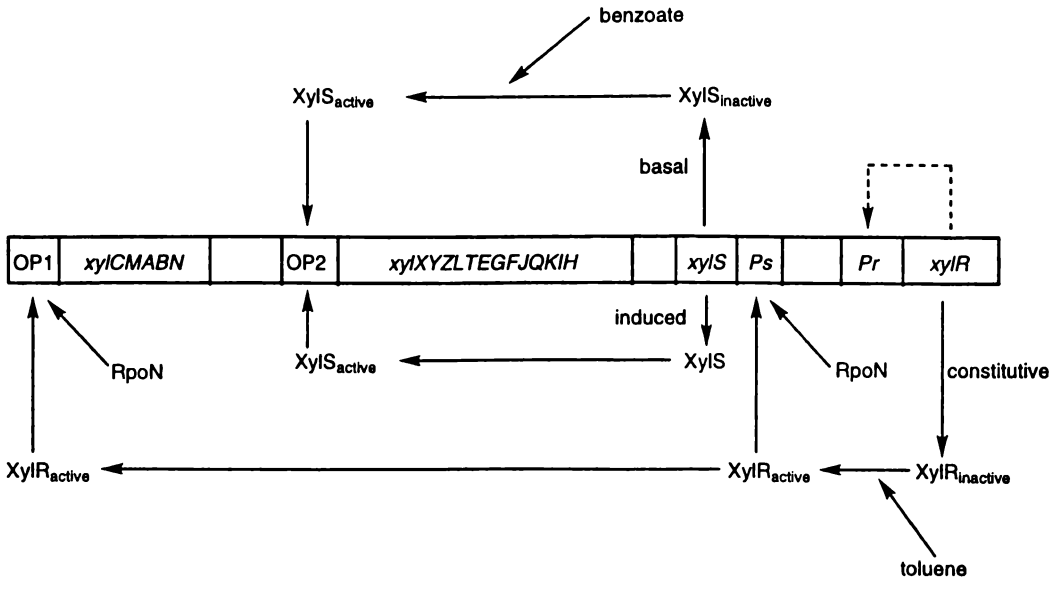


Figure 1.6 Regulatory circuits controlling expression of the pWW0 *xyl* genes in *P. putida*. Solid and broken arrows indicate positive and negative control respectively. Adapted from Assinder & Williams, 1990.

the upper pathway promoter OP1, the *meta* pathway promoter OP2, and the promoters for the two regulatory genes *xylR* and *xylS*, designated *Pr* and *Ps* respectively. *xylS* and *xylR* are located at the downstream end of the *meta* pathway operon, with *xylS* being adjacent to *xylH*.

The key element in the regulatory system is the constitutive expression of *xylR*, and involves a regulatory cascade hinging around the ability of the gene product XylR to interact with both the OP1 and *Ps* promoters. The model suggests that the primary product of *xylR* is an inactive from of the protein, but on interaction with upper pathway substrates it becomes activated (Figure 1.6). It then acts with RpoN, a sigma factor for transcription initiation, to stimulate transcription from the OP1 promoter resulting in expression of the upper pathway genes, and from the *Ps* promoter which expresses *xylS*. The XylS regulatory protein in turn interacts with OP2, thus leading to simultaneous co-induction of both the upper and *meta* pathways (Assinder & Williams, 1990; Nakazawa *et al.*, 1990).

XylS is constitutively expressed at a low basal level and exists in an equilibrium between an inactive and an active form. In the absence of any inducing substrates the equilibrium is shifted toward the inactive form and hence insufficient active XylS is present for activation of the OP2 promoter. In the presence of benzoate, the equilibrium is shifted toward the active form and the OP2 promoter is activated,

switching on the *meta* pathway genes. The same effect is observed in the presence of XylR, but this is due to increased expression of *xylS* which increases the relative levels of active XylS, although the equilibrium position remains unchanged (Assinder & Williams, 1990).

1.2.4 Catabolic pathways for biphenyl

Biphenyl utilising bacteria are ubiquitously distributed in the environment and those isolated and described include a more diverse range of bacteria than for toluene and naphthalene degradation. Biphenyl degraders are of particular interest because the low specificity of many of the early enzymes of the biphenyl pathway enables such strains to carry out partial catabolism of some polychlorinated biphenyl (PCB) congeners, which is of interest for bioremediation purposes (Williams & Sayers, 1994).

The *bph* operon for biphenyl catabolism was initially characterised from chromosomal DNA of *P. pseudoalcaligenes* KF707, and later from *P. putida* KF715 (Furukawa *et al.*, 1990). The conversion of biphenyl to benzoate requires four enzymes which are encoded by a single operon. *bph* genes have been found located on both plasmids and chromosomes, and in most strains the order of genes in this operon is conserved as *bphA1A2A3A4BCD* (Furukawa *et al.*, 1989, Williams & Sayers, 1994). *bphA* encodes a multicomponent dioxygenase which has an analogous mode of action to naphthalene dioxygenase (Figure 1.5). *bphB*, *C* and *D* encode enzymes for the further transformation of the dihydroxylated biphenyl to benzoate (Figure 1.7).

Evidence suggests the *bph* gene cluster is highly conserved amongst biphenyl/PCB degrading bacteria, but that divergent genes not homologous to *bph* probes may specify biphenyl catabolism in some strains. Using the *bph* genes from *P. pseudoalcaligenes* KF707 as probes, significant homology was found with other biphenyl/PCB degrading species of *Pseudomonas*, *Alcaligenes*, and *Achromobacter*. However, a number of strains, including members of the *Moraxella* and *Arthrobacter* genus, showed no homology with the KF707 *bph* probes. From *Rhodococcus* sp. RHA1 was cloned a cluster of biphenyl/PCB degrading genes (Masai *et al.*, 1995). The gene order of this operon was different to that of other biphenyl degraders reported previously, being *bphA1A2A3A4CB*. The RHA1 *bph* genes showed a higher

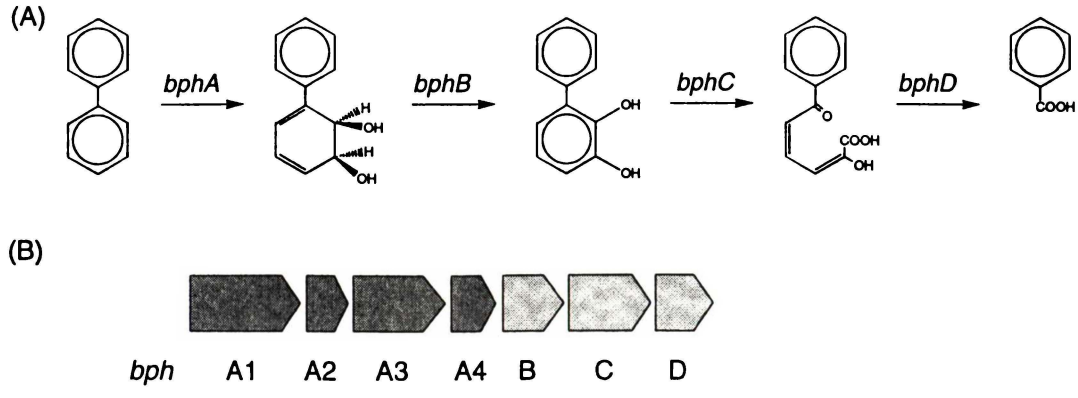


Figure 1.7 (A) Catabolic pathway for the degradation of biphenyl to benzoate. (B) Organisation of the *bph* operon in *P. pseudoalcaligenes* KF707 and *P. putida* KF715 (strain KF707 contains an ORF between *bphC* and *bphD*, the function of which is unknown). Enzymes encoded by the *bph* genes are *bphA*, biphenyl 2,3-dioxygenase; *bphB*, dihydrodiol dehydrogenase; *bphC*, 2,3-dihydroxybiphenyl dioxygenase; *bphD*, hydrolase. Adapted from Furukawa *et al.* (1992).

degree of homology to the *tod* genes of *P. putida* F1 than with the *bph* genes of *Pseudomonas* sp. strains KF707 and KKS102.

The upper pathway and *meta* pathway genes for biphenyl catabolism were located to a plasmid, pWW100, in *Pseudomonas* sp. CB406 (Lloyd-Jones *et al.*, 1994). During growth of this strain on benzoate, spontaneous mutants were observed which had lost the ability to grow on biphenyl. Growth of these mutants on benzoate proceeded via the β-ketoadipate pathway, analogous to strains harbouring the TOL plasmid (Assinder & Williams, 1990). Recombination of the *bph* genes into plasmid RP4 to form a cointegrate plasmid suggests both upper and *meta* pathways for biphenyl metabolism may be present on a transposable element (Lloyd-Jones *et al.*, 1994). The loss of both pathways during growth on benzoate, possibly by a deletion, may also be related to transposition.

1.3 Microbial degradation of PAHs

The initial observation that a PAH could provide the sole source of carbon and energy for bacterial growth came in 1927 when Tausson demonstrated bacterial utilisation of naphthalene (Yen & Serdar, 1988). Since then a large number of bacteria have been isolated which are able to use PAHs as a sole carbon and energy source; some of these are described in Table 1.3.

Microbial degradation of PAHs containing three or fewer rings has been reported for a number of different genera of aerobic bacteria. The pathways for degradation of lower molecular weight PAHs, particularly naphthalene, have been well documented, and in many cases the genes and enzymes involved have been characterised. Larger molecular weight PAHs, however, are more recalcitrant and fewer bacterial strains have been isolated that are able to degrade compounds with four or more rings. Those which have been isolated are mainly Gram-positive bacteria. In some instances the pathways for the degradation of certain four- and five-ring PAHs have been elucidated, but very little is known about the genetics involved.

1.3.1 Naphthalene

Naphthalene was often selected as a model compound for the study of PAH degradation because of its high aqueous solubility, and the ease of isolation of microbes capable of its degradation. The biochemical pathway for naphthalene oxidation by *Pseudomonas* was first reported in 1964 (Davies & Evans, 1964).

Naphthalene catabolism begins with a sequence of reactions analogous to the upper pathway for biphenyl metabolism. The initial step involves the action of naphthalene dioxygenase which incorporates both molecules of molecular oxygen to form (+)-*cis*-1R,2S-dihydroxy-1,2-dihydronaphthalene. This intermediate is dehydrogenated to form 1,2-dihydroxynaphthalene which is enzymatically cleaved by a dioxygenase to yield *cis*-2'-hydroxybenzalpyruvate. An aldolase catalyses the cleavage of *cis*-2'-hydroxybenzalpyruvate to pyruvate and salicylaldehyde; salicylaldehyde is subsequently oxidised to salicylate by a dehydrogenase. Further oxidation of salicylate by salicylate hydroxylase yields catechol, which may undergo either *ortho* or *meta* cleavage (Figure 1.8).

Strain	PAHs degraded	Reference
P. putida NCIB 9816	naphthalene, fluorene, phenanthrene	Yang *et al.*, 1994
Sphingomonas sp.	naphthalene	Fredrickson *et al.*, 1995
P. putida G7	naphthalene	Dunn & Gunsalus, 1973
Sphingomonas paucimobilis EPA505	naphthalene, phenanthrene anthracene, fluoranthene, benzo[*b*]fluorene	Mueller *et al.*, 1990
S. paucimobilis	phenanthrene	Weissenfels *et al.*, 1990
P. vesicularis	fluorene	Weissenfels *et al.*, 1990
Alcaligenes denitrificans WW1	naphthalene, phenanthrene, anthracene, fluoranthene	Weissenfels *et al.*, 1990
P. putida OUS82	naphthalene, phenanthrene	Kiyohara *et al.*, 1994
P. fluorescens 5R	naphthalene, phenanthrene, anthracene	Sanseverino *et al.*, 1993a
Mycobacterium sp. PYR-1	naphthalene, phenanthrene, fluoranthene, pyrene, 3-methylcholanthene, 1-nitropyrene, 6-nitrochrysene,	Heitkamp *et al.*, 1988a
P. cepacia F297	naphthalene, phenanthrene, fluorene, anthracene	Grifoll *et al.*, 1995
Mycobacterium sp.	pyrene, benzo[*a*]pyrene	Grosser *et al.*, 1991
Mycobacterium sp. BG1	phenanthrene	Guerin & Jones, 1988
Rhodococcus sp. UW1	phenanthrene, fluoranthene, pyrene,anthracene, chrysene	Walter *et al.*, 1991
Mycobacterium sp.	phenanthrene, fluoranthene, pyrene	Fritzsche, 1994
Mycobacterium sp. BB1	phenanthrene, fluoranthene, pyrene	Boldrin *et al.*, 1993
Arthrobacter sp. F101	fluorene	Grifoll *et al.*, 1992
Staphylococcus auriculans DBF63	fluorene	Monna *et al.*, 1993

Table 1.3 Some of the described bacteria capable of utilising PAHs

Much of the work contributing to the current understanding of the naphthalene catabolic pathway occurred in the laboratory of Gibson. Gibson and co-workers characterised naphthalene dioxygenase, the first enzyme in the naphthalene catabolic pathway, from *P. putida* NCIB 9816 (Ensley *et al.*, 1982). Naphthalene dioxygenase is a three-component oxygenase consisting of a reductase, ferredoxin, and an iron-sulfur protein (ISP). The ISP is a multimer consisting of α and β subunits arranged $\alpha_2\beta_2$ (Mason & Cammack, 1992). The mode of action of naphthalene dioxygenase was described in section 1.2.2 above.

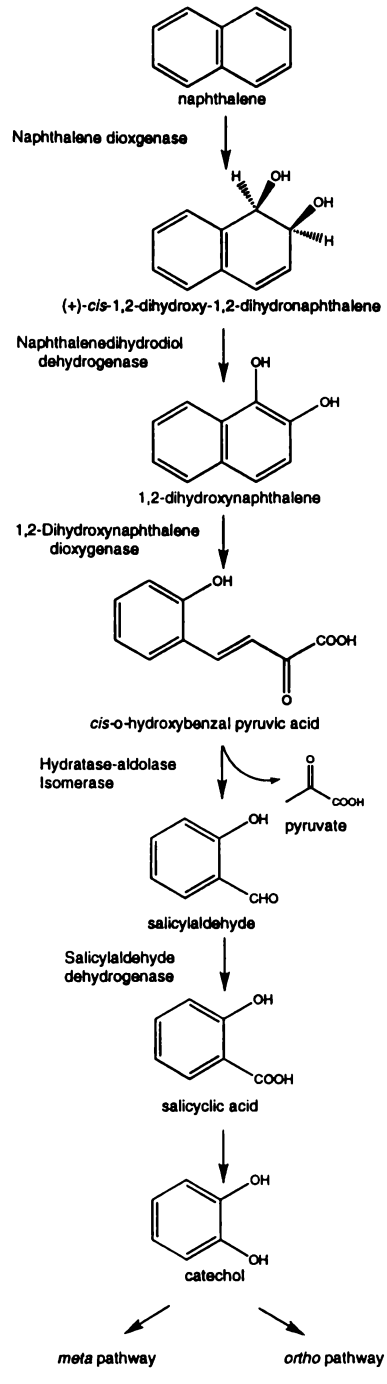


Figure 1.8 Bacterial degradation of naphthalene. Adapted from Pothuluri & Cerniglia (1994).

The second enzyme in the pathway is a *cis*-naphthalene dihydrodiol dehydrogenase which catalyses the formation of 1,2-dihydroxynaphthalene from the *cis*-dihydrodiol. This enzyme was purified from *P. putida* strain NP and requires nicotinamide-adenine dinucleotide (NAD) as an electron acceptor (Patel & Gibson, 1974). The enzyme was shown to oxidise other *cis*-dihydrodiols but no activity was observed

towards *trans*-naphthalene dihydrodiol. This dehydrogenase has a molecular weight of 102,000 and consists of four subunits of 25,000.

The third enzyme in the pathway, 1,2-dihydroxynaphthalene 1,2-dioxygenase, was purified from *P. putida* NCIB 9816 and has a molecular weight of over 275,000. This enzyme is an extradiol dioxygenase which cleaves the aromatic ring. 1,2-dihydroxynaphthalene 1,2-dioxygenase has a subunit molecular weight of 19 kdaltons and requires Fe^{2+} for activity (Patel & Barnsley, 1980).

Some confusion has surrounded the next step of the pathway. Work by Davies & Evans (1964) and Barnsley (1976a) reached different conclusions in their study of the metabolism of 1,2-dihydroxynaphthalene. The Davies-Evans pathway suggests the unstable ring cleavage product, formed from oxidation of 1,2-dihydroxynaphthalene, rearomatises to *cis*-2'-hydroxybenzal pyruvate (cHBPA) which is then metabolised in two sequential steps (hydration and aldol cleavage) to salicylaldehyde and pyruvate. Direct evidence for the hydratase enzyme was not found, nor was the intermediate formed by the hydratase reaction. Barnsley's suggested pathway includes another intermediate: the ring cleavage product spontaneously recyclises to 2-hydroxychromene-2-carboxylate (HCCA) which is converted by an isomerase to cHBPA.

Davies & Evans (1964) and Barnsley (1976a) reached these different conclusions despite taking similar approaches. Both groups were hampered by the instability of the starting substrate, 1,2-dihydroxynaphthalene, which is rapidly and spontaneously oxidised to 1,2-naphthoquinone in water, and by insufficient quantities of the products for rigorous identification.

Eaton & Chapman (1992) used a genetic approach to resolve this confusion. Genes from the naphthalene catabolic plasmid NAH7 of *P. putida* G7 were cloned to form recombinant plasmids which were carried by *P. aeruginosa* PAO1. Clones were obtained which converted the stable substrate, naphthalene, to metabolites which could be accumulated and identified. This approach overcame many of the problems associated with the instability of the intermediate metabolites.

Results obtained by this study indicate that 1,2-dihydroxynaphthalene is cleaved by a dioxygenase which inserts a molecule of oxygen between the angular carbon and carbon 1 to give an unstable ring cleavage product (Figure 1.9). This compound then rearomatises yielding another unstable intermediate, *cis*-2'-hydroxybenzal pyruvate

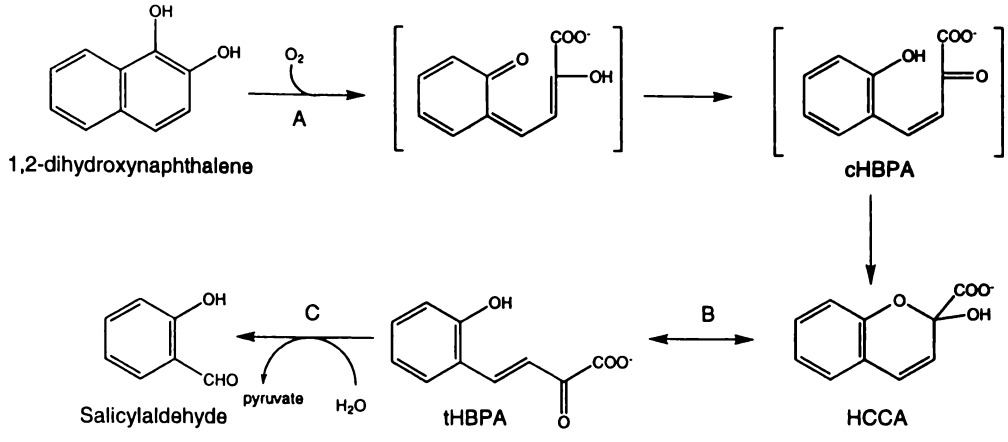


Figure 1.9 Pathway for the plasmid NAH7-encoded metabolism of 1,2-dihydroxynaphthalene to salicylaldehyde as proposed by Eaton & Chapman (1992). Chemical designations: cHBPA, *cis*-2'-hydroxybenzal pyruvate; HCCA, 2-hydroxychromene-2-carboxylate; tHBPA, *trans-o*-hydroxybenzylidenepyruvate. Enzyme designations: A, 1,2-dihydroxynaphthalene dioxygenase; B, 2-hydroxychromene-2-carboxylate isomerase; C, *trans-o*-hydroxybenzylidene-pyruvate hydratase-aldolase. Adapted from Eaton & Chapman (1992).

(cHBPA), which is spontaneously converted to its hemiketal, 2-hydroxychromene-2-carboxylate (HCCA), the first detectable ring cleavage product. HCCA isomerises to an equilibrium mixture containing 45% *trans-o*-hydroxybenzylidene pyruvate in a reaction catalysed by an isomerase. This compound, tHBPA, is metabolised by a hydratase-aldolase enzyme to salicylaldehyde and pyruvate.

Salicylaldehyde was shown to be oxidised to salicylate by an NAD^+-dependant dehydrogenase. In most cases salicylate is oxidised to catechol which is the ring fission substrate for the *meta* or *ortho* pathway. The *meta* cleavage pathway involves extradiol cleavage of catechol with subsequent formation of 2-hydroxymuconic semialdehyde; the *ortho* fission pathway involves intradiol cleavage of the catechol and the product is *cis, cis*-muconic acid. Salicylate may also be metabolised via gentisate; this has been suggested for the metabolism of naphthalene by *Pseudomonas* sp. U2 (Fuenmayor *et al.*, 1998) and *P. fluorescens* and *P. alcaligenes* strains (Yen & Serdar, 1988).

The naphthalene dioxygenase gene is often used as a marker during attempts to clone catabolic operons from wild type strains (Kiyohara *et al.*, 1994; Yang *et al.*, 1994). Naphthalene dioxygenase catalyses the dihydroxylation of indole to indole dihydrodiol, which is spontaneously dehydrated to 3-hydroxyindole, two molecules

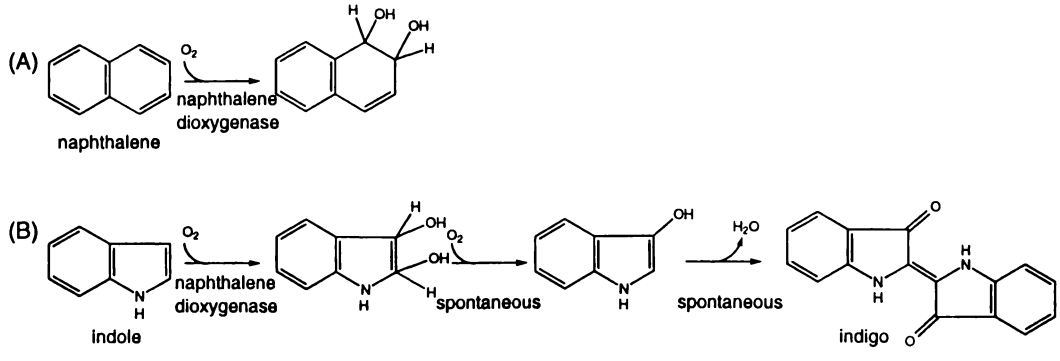


Figure 1.10 Reactions catalysed by naphthalene dioxygenase; (A) oxidation of naphthalene and (B) conversion of indole to indigo. Adapted from Assinder & Williams, 1990.

of which condense to form the blue product indigo (Ensley *et al.*, 1983) (Figure 1.10). *E. coli* clones containing the naphthalene dioxygenase gene will appear as blue colonies when grown on a solid medium supplemented with indole. Blue colonies may also result in the absence of indole as *E. coli* possesses the enzyme tryptophanase which acts on tryptophan to produce indole endogenously. It is interesting to note that xylene oxidase, the product of the TOL *xylMA* gene, also converts indole to indigo (Assinder & Williams, 1990). It is not clear how this occurs since xylene oxygenase is a monooxygenase; it may hydroxylate indole to 3-hydroxyindole which would then dimerise, but this would imply a very broad substrate specificity of the enzyme as indole bears very little in common with toluene.

1.3.2 Phenanthrene and anthracene

Bacterial oxidation of anthracene and phenanthrene, and the subsequent degradative pathway, follow the same principles as that described for naphthalene. Many naphthalene degrading strains are also able to degrade phenanthrene and anthracene due to the relaxed substrate specificity of the naphthalene-catabolic enzymes which allows attack of these similar structured substrates.

Phenanthrene is degraded by bacteria through one of two routes via a common intermediate, 1-hydroxy-2-naphthoic acid. In one route 1-hydroxy-2-naphthoic acid is decarboxylated oxidatively to 1,2-dihydroxynaphthalene, which is further degraded via salicylate as for naphthalene degradation. The enzymatic steps for the

decomposition of the first and second rings of phenanthrene via the salicylate route resemble those for the first ring of naphthalene. It has been demonstrated that a common set of enzymes was responsible for the decomposition of the first and second rings of phenanthrene, as well as for the first ring of naphthalene (Kiyohara *et al.*, 1994; Takizawa *et al.*, 1994; Yang *et al.*, 1994).

In the other route, the ring of 1-hydroxy-2-naphthoic acid is cleaved and further metabolised via *o*-phthalate. *Aeromonas*, *Alcaligenes*, *Micrococcus* and *Vibrio* strains have been described which degrade phenanthrene via the *o*-phthalate route (Pothuluri & Cerniglia, 1994), and the enzymes which transform 1-hydroxy-2-naphthoic acid to *o*-phthalate have been characterised from *Nocardioides* sp. KP7 (Iwabuchi & Harayama, 1997).

The stereochemistry of the initial dihydroxylation of phenanthrene has been elucidated. Dihydroxylation occurs at the 1,2- and 3,4- positions to form (+)-*cis*-1R,2S-dihydroxy-1,2-dihydrophenanthrene and (+)-*cis*-3R,4S-dihydroxy-3,4-dihydrophenanthrene (Jerina *et al.*, 1976; Koreeda *et al.*, 1978). The major isomer formed is (+)-*cis*-3R,4S-dihydroxy-3,4-dihydrophenanthrene.

Several species of *Pseudomonas* and *Sphingomonas yanoikuyae* B1 (formerly *Beijerinckia* sp.) are able to oxidise anthracene in the 1,2-position to form (+)-*cis*-1R,2S-dihydroxy-1,2-dihydroanthracene (Jerina *et al.*, 1976). This is followed by a dehydrogenation step to form 1,2-dihydroxyanthracene. A *Pseudomonas* sp. was found to cleave 1,2-dihydroxyanthracene by a dioxygenase to yield *cis*-4-(2-hydroxynaph-3-yl)-2-oxo-but-3-enoic acid (Evans *et al.*, 1965). Further metabolism of this ring fission product leads to the formation of 2-hydroxy-3-naphthoic acid. This compound is further metabolised through salicylate and catechol by the enzyme of the naphthalene pathway. *S. yanoikuyae* B1 uses the same enzyme system for naphthalene and anthracene oxidation (Kim *et al.*, 1997b). This strain grows on naphthalene, and although it is unable to grow on anthracene it is able to transform it to some extent. An insertional mutant of *S. yanoikuyae* B1 which did not express 2-hydroxychromene-2-carboxylate isomerase (the isomerase of the naphthalene pathway) accumulated 6,7-benzocounarin, implicating the naphthalene pathway enzymes in the metabolism of anthracene.

1.3.3 Pyrene

Much of the work pertaining to microbial degradation of high molecular weight PAHs has focused on a few selected strains which have a very versatile catabolic ability.

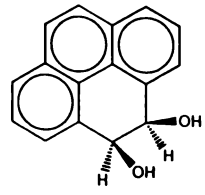
Mycobacterium sp. PYR-1 was isolated from marine sediments collected near a point source for petrogenic chemicals in southern Texas (Heitkamp & Cerniglia, 1988). This strain mineralised four-ring PAHs faster than two- or three-ringed PAHs, and PAH degrading enzymes were shown to be inducible. *Mycobacterium* sp. PYR-1 utilises a wide range of PAHs for growth (Table 1.3), which may reflect a relaxed substrate specificity of the initial dioxygenase enzymes.

Pyrene-induced *Mycobacterium* sp. PYR-1 cultures mineralised 5% of pyrene after 6 hours and reached a maximum of 48% mineralisation within 72 hours when grown in a media supplemented with organic nutrients (Heitkamp *et al.*, 1988a; b). The initial oxidative attack on pyrene by *Mycobacterium* sp. PYR-1 resulted in the formation of dihydrodiols at the 4,5 position (Figure 1.11). Both *cis*- and *trans*-4,5-dihydrodiols were formed suggesting multiple pathways for the initial oxidative attack on pyrene, as dioxygenases catalyse the formation of *cis*-dihydrodiols only. Experiments with $^{18}O_2$ confirmed this, showing that pyrene *cis*-dihydrodiol formation was a dioxygenase catalysed reaction and that pyrene *trans*-dihydrodiol formation was catalysed by monooxygenase enzymes. It is believed *cis*-dihydrodiols undergo rearomatisation to hydroquinine derivatives, which are precursors for ring cleavage and further metabolism (Heitkamp *et al.*, 1988b).

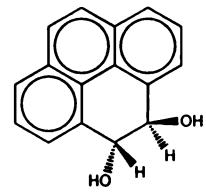
The inducible pyrene degrading enzymes of this *Mycobacterium* may be either plasmid or chromosome mediated. Initial efforts by Heitkamp *et al.* (1988b) to isolate and identify plasmids in *Mycobacterium* sp. PYR-1 were not successful because of very high activities of nuclease enzymes, which degraded the bacterial DNA during processing.

Walter *et al.* (1991) reported the degradation of pyrene by *Rhodococcus* sp. UW1 which was isolated from contaminated soil using conventional enrichment techniques. This isolate was able to utilise pyrene as sole carbon and energy source; it mineralised 72% of the pyrene within 2 weeks, and showed a maximum degradation rate of 0.08 mg pyrene ml^{-1} day^{-1}, while growing with a doubling time of 30 hours.

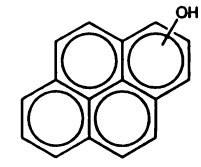
Ring-oxidation products:



cis-4,5-pyrenedifydrodiol

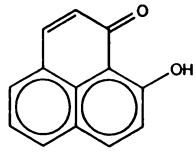


trans-4,5-pyrenedihydrodiol

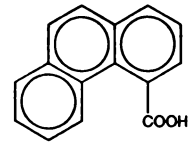


pyrenol

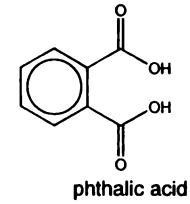
Ring-fission products:



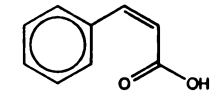
4-hydroxyperinaphthenone



4-phenanthroic acid



phthalic acid



cinnamic acid

Figure 1.11 Structures of identified pyrene metabolites produced by *Mycobacterium* sp. PYR-1. Adapted from Heitkamp *et al.*, 1988b.

During growth a metabolite was detected in the culture fluid and characterisation by UV and mass spectrometry indicated this metabolite resulted from a recyclisation of the direct *meta* ring-fission product of pyrene after dihydroxylation in either the 1,2- or the 4,5- position.

The proposed initial attack on pyrene (Figure 1.12) proceeds via dihydroxylation followed by ring cleavage. It was not clear whether ring cleavage occurs in the 1,2- or the 4,5 position, so two possible pathways for the initial oxidation and ring fission of pyrene by *Rhodococcus* sp. UW1 are suggested. The proposed attack is analogous to the microbial degradation of naphthalene and phenanthrene. The uncertainty as to whether the oxidation of pyrene proceeds via attack at the 1,2- or 4,5 position could be explained by a relaxed substrate specificity for the initial dioxygenase system responsible, such that pyrene, a highly symmetrical and compact PAH, is hydroxylated either in the 1,2- or 4,5 positions. This hypothesis is supported by the wide PAH-utilisation spectrum of *Rhodococcus* sp. UW1. Besides pyrene, phenanthrene, anthracene, fluoranthene and chrysene were dihydroxylated by this strain, and these compounds were utilised as sole source of carbon and energy (Table 1.3).

For cultures of *Rhodococcus* sp. UW1, pH 7.0 and 30°C were optimum for cell growth and pyrene degradation, which agrees with previous studies of the influence

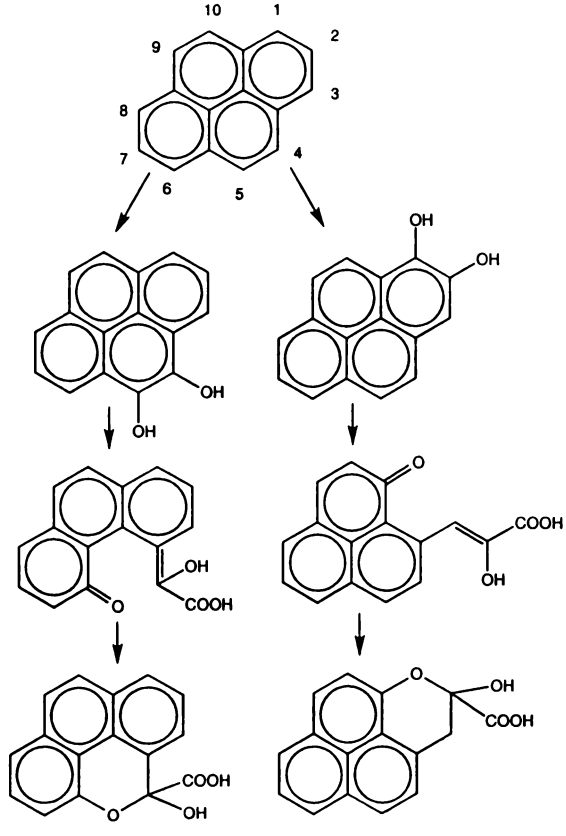


Figure 1.12 Initial reactions in the degradation of pyrene by *Rhodococcus* sp. UW1. Adapted from Walter *et al.*, 1991.

of pH and temperature on PAH degradation by pure bacterial cultures (Weissenfels, 1990). In contrast to the bacterial culture, the initial dioxygenase showed a high tolerance against extreme pH values and temperatures. Measurements of the oxygen-consumption rates of pyrene-induced resting cells showed maximal activity occurred at pH 7.2 and 45°C. This is not unusual as a number of other authors have described ring-fission dioxygenases with high acid and heat tolerance.

1.3.4 Fluoranthene

Sphingomonas paucimobilis EPA505 was isolated from a bacterial consortium enriched from creosote contaminated soil (Mueller *et al.*, 1990). The enzymic activities in extracts of fluoranthene-induced cells indicate a *meta* ring-fission involved in the degradation process. This led to the prediction of a biodegradative pathway for fluoranthene which involves dihydroxylation of fluoranthene at the 9,10-position to yield 9,10-dihydroxyfluoranthene. The pathway may then follow the principles of PAH degradation as previously discussed with respect to naphthalene and phenanthrene, which would involve a second dihydroxylation of the 9,10-fluoranthenedihydrodiol, resulting in *meta* cleavage of the aromatic ring. An aldolase reaction on the ring fission product would give pyruvate and 7-hydroxy-8-acenaphthylenealdehyde. The latter compound was not detected but instead 7-hydroxyacenaphthylene occurred. The formation of this compound requires a 1-carbon excision from the aromatic aldehyde but the mechanism of this reaction remains unclear. A similar 1-carbon excision has been reported by Heitkamp *et al.* (1988b) who found 4-phenanthroic acid as a major metabolite of pyrene degradation by *Mycobacterium* sp. PYR-1 (Figure 1.11). 7-hydroxyacenaphthylene is transformed to 4,5-benzocoumarine according to the suggested pathway for fluoranthene catabolism. This requires the insertion of one oxygen atom into the carbon ring system; this type of reaction has not been previously reported for the degradation of a PAH (Weissenfels *et al.*, 1991).

Mycobacterium sp. PYR-1 has been found to mineralise fluoranthene significantly, following enzymatic induction by pyrene (Heitkamp & Cerniglia, 1988; Heitkamp *et al.*, 1988b). Kelley & Cerniglia (1991) reported *Mycobacterium* sp. PYR-1 degrades 95% of fluoranthene within a 24 hour period after an initial 6 to 12 hour lag phase when supplemented with organic nutrients. The major fluoranthene metabolite formed was identified as 9-fluorenone-1-carboxylic acid. Kelley *et al.* (1993) studied fluoranthene metabolism by *Mycobacterium* sp. PYR-1 further, and identified ten fluoranthene metabolites isolated from culture extracts. On this basis a pathway for fluoranthene degradation by *Mycobacterium* sp. PYR-1 was proposed (Figure 1.13). The initial attack of fluoranthene appears to occur by a dioxygenase at the 1,2-position followed by extradiol ring cleavage. However, the identification of

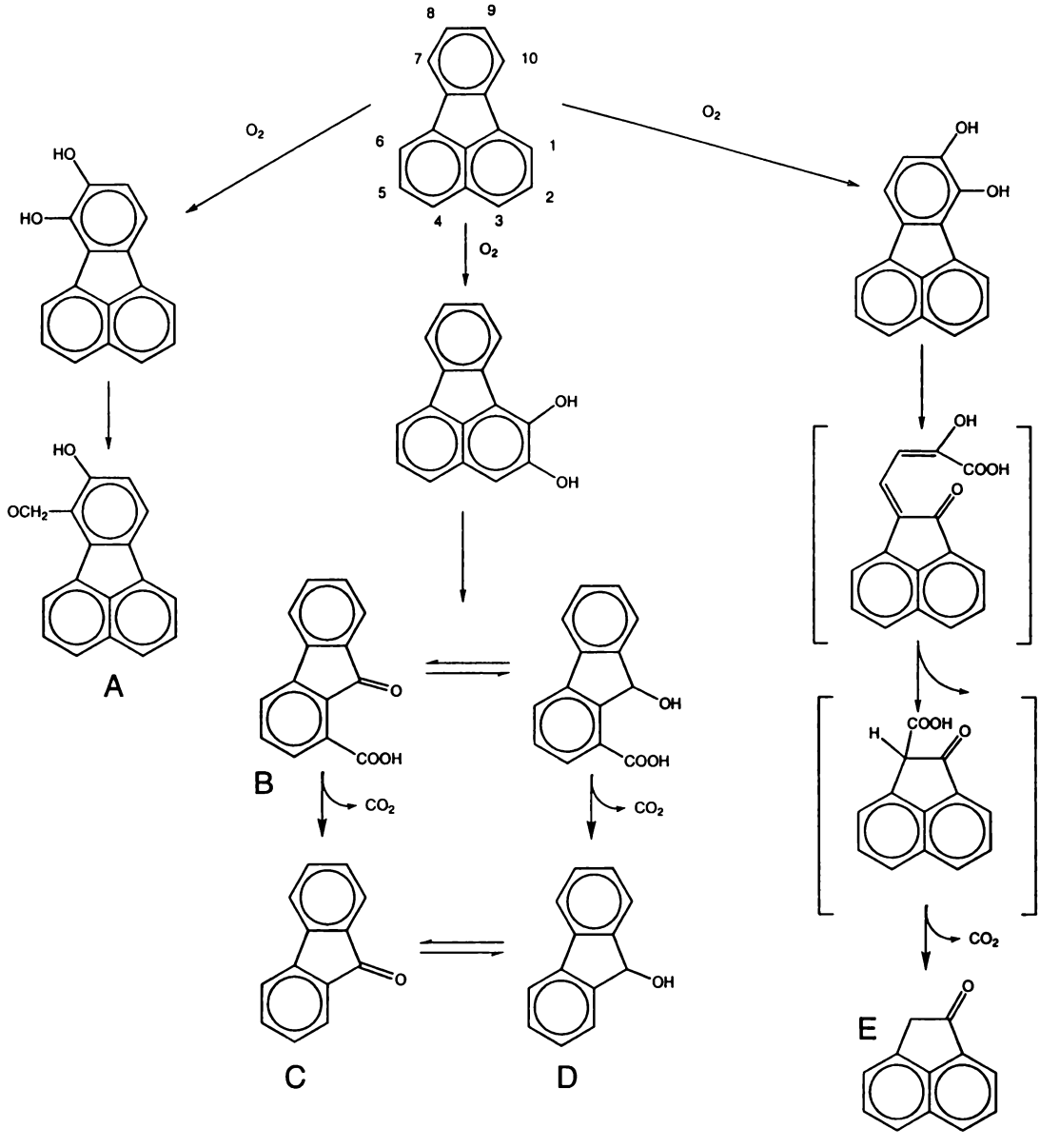


Figure 1.13 Pathways proposed for the metabolism of fluoranthene by *Mycobacterium* sp. PYR-1. Metabolites are: A, 7-methoxy-8-hydroxy-fluoranthene; B, 9-fluorenone-1-carboxylic acid; C, 9-fluorenone; D, 9-hydroxyfluorene; E, 1-acenaphthenone; (compounds in brackets were not identified). Adapted from Kelley *et al.* (1993).

metabolites that retain the intact fluorene configuration (ie. dihydroxylation at the 1,2 position), as well as the identification of an initial oxygenated metabolite in which the single benzene ring has been attacked, indicate that several pathways operate simultaneously on the degradation of fluoranthene by *Mycobacterium* sp. PYR-1 (Kelley *et al.*, 1993). This observation may be related to evidence that pyrene degradation by this strain involves the action of both mono- and dioxygenases (Heitkamp *et al.*, 1988b).

1.3.5 Benzo[*a*]pyrene

Very little is known about the microbial degradation of the high molecular weight PAHs such as benzo[*a*]pyrene. Although bacteria capable of utilising five-ring PAHs as the sole carbon and energy source are limited, some microorganisms can oxidise these very insoluble PAHs when they are grown on an alternative carbon source. Gibson *et al.* (1975) described a mutant strain of *S. yanoikuyae* B1 which, when grown on succinate in the presence of biphenyl, oxidised benzo[*a*]pyrene to *cis*-9,10-dihydroxy-9,10-dihydrobenzo[*a*]pyrene and *cis*-7,8-dihydroxy-7,8-dihydrobenzo-[*a*]pyrene. The ring cleavage products and the reaction sequence for bacterial mineralisation have not yet been identified.

1.3.6 Benz[*a*]anthracene

The bay-region 1,2-position is the preferred site of attack for the bacterial oxidation of benzo[*a*]anthracene (Gibson & Subramanian, 1984) although oxidation at the 8,9- and 10,11- positions has been reported (Gibson *et al.*, 1975). The absolute stereochemistry of the *cis*-1,2-, *cis*-8,9-, and *cis*-10,11-dihydrodiols formed by oxidation of benzo[*a*]anthracene by *S. yanoikuyae* B1 has been reported by Jerina *et al.* (1984).

Mahaffey *et al.* (1988) studied the bacterial oxidation of benz[*a*]anthracene after induction with biphenyl, *m*-xylene, and salicylate. Biotransformation studies with benz[*a*]anthracene showed that after 14 hours 56% was converted to an isomeric mixture of three *o*-hydroxypolyaromatic acids by *S. yanoikuyae* B1. The major metabolite was identified as 1-hydroxy-2-anthranoic acid; the two minor metabolites were identified as 2-hydroxy-3-phenanthroic acid and 3-hydroxy-2-phenanthroic acid. The formation of the two minor metabolites supposedly occurred by oxidative cleavage of the arene diol formed at the 10,11- and 8,9-positions of benz[*a*]anthracene. Both of these acids were apparently formed by a series of reactions analogous to those proposed for the oxidation of anthracene to 2-hydroxy-3-naphthoic acid (Evans *et al.*, 1965). Mineralisation experiments with [12-^{14}C]benzo[*a*]anthracene showed formation of $^{14}CO_2$ leading Mahaffey *et al.* (1988) to suggest the hydroxy acids can be oxidised whereby at least two rings of the molecule can be further degraded.

P. putida NCIB 9816 transforms benz[*a*]anthracene to some extent due to the relaxed

substrate specificity of the naphthalene catabolic enzymes. Salicylate enhances the degradation of benz[*a*]anthracene by this strain (Barnsley, 1975a), probably because of the increased expression of naphthalene degrading enzymes induced by salicylate (Barnsley, 1976b).

1.4 Genetics of PAH degradation

Considerable research has focused on the genetic basis of naphthalene catabolism by bacteria. This information has contributed to a growing understanding of the metabolism of aromatic compounds by bacteria, and concurs with findings for toluene and biphenyl. An understanding of the role of plasmids in carrying catabolic genes has also emerged, and possible mechanisms for regulation of catabolic pathways have been suggested.

1.4.1 Naphthalene catabolic plasmids

Plasmids carrying naphthalene catabolic genes represent a class of well-documented catabolic plasmids. A number of different plasmids encoding naphthalene degradation have been isolated from strains of *P. putida* and *P. fluorescens* (Yen & Serdar, 1988). Naphthalene catabolic plasmids belong to either compatibility group P7 or P9 and are all self-transmissible. Available data suggest that all of the naphthalene catabolic plasmids encode a single upper pathway and most of them specify the same lower pathway for salicylate degradation. Naphthalene catabolic plasmids are all quite large (81–173 kb).

The first and best studied naphthalene catabolic plasmid is NAH7 which is an 83 kb plasmid isolated from *P. putida* G7. Dunn & Gunsalus (1973) attributed growth on naphthalene to the presence of a plasmid. *P. putida* G7 was shown to grow on naphthalene but this phenotype was lost spontaneously and the frequency of this loss was increased by treatment with mitomycin C. The Nah^+ (naphthalene degradation) phenotype could be transferred to other Nah^- heterologous fluorescent pseudomonads by conjugation or transduction (Dunn & Gunsalus, 1973).

Transconjugants which could use naphthalene grew on salicylate and were shown to carry catechol 2,3-dioxygenase, the initial enzyme of the *meta* cleavage pathway, whereas cured strains grew on neither salicylate nor naphthalene, and lacked catechol 2,3-dioxygenase, but retained catechol 1,2-dioxygenase and the aromatic β-ketoadipate pathway enzymes (Dunn & Gunsalus, 1973). Hence it was assumed that the aromatic *meta* pathway genes were present on a plasmid.

NAH7 contains two *nah* operons (Figure 1.14); the first operon includes genes *nahAaAbAcAdBCDEF*, coding for the conversion of naphthalene to salicylate (upper

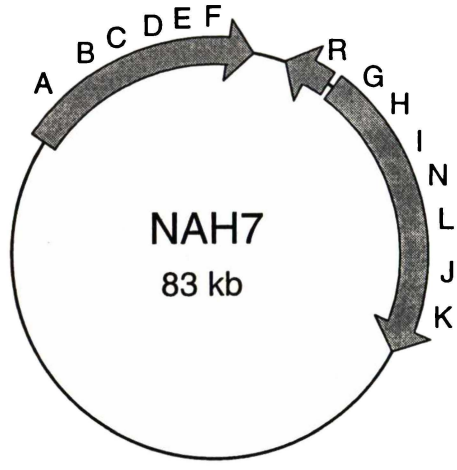


Figure 1.14 The gene organisation of NAH7. The naphthalene catabolic genes are marked by capital letters; transcription of all genes is in a clockwise direction except for *nahR*. *EcoR1* restriction sites are shown; numbers denote kilobases from the *EcoR1* site adjacent *nahA*. Adapted from Yen & Serdar (1988).

pathway), and the second operon includes genes *nahGTHIJK*, coding for the oxidation of salicylate via the catechol *meta* cleavage pathway to acetaldehyde and pyruvate (lower pathway) (Yen & Gunsalus, 1982). The gene order of the first operon has been reported as *nahABCDEF* (Yen & Gunsalus, 1982), *nahABCDF* (Schell, 1983) and *nahABFCED* (Eaton & Chapman, 1992), but recent evidence suggests it is *nahABFCQED* (Eaton, 1994) where *nahQ* is a gene of unknown function associated with naphthalene catabolism.

The organisation of these genes into two distinct groups indicates that each of the two *nah* operons controls the generation of pyruvate from the cleavage of one of the two naphthalene ring structures. Inactivation of the first operon confers a $Nah^{-}Sal^{+}$ phenotype, and inactivation of the second operon confers a $Nah^{+}Sal^{-}$ phenotype. Transcription of the first operon is from the *nahA* to the *nahF* gene, while the second operon is transcribed from *nahG* to *nahK*. Although the NAH7 plasmid is 83 kb in size, restriction mapping of NAH7::Tn5 mutants placed all of the characterised *nah* genes within a 30 kb region of the NAH7 plasmid (Yen & Gunsalus, 1982).

The naphthalene catabolic determinant of *P. putida* NCIB 9816 is also plasmid-borne. The original NCIB 9816 strain was isolated from soil (Evans *et al.*, 1965) and since then four variants of *P. putida* NCIB 9816 are in circulation among different

laboratories, Yen & Serdar (1988) differentiated the four variants by designating them NCIB 9816-1 to NCIB 9816-4 based on differences in the regulation of the naphthalene catabolic pathway and the number of plasmids in each of these strains.

P. putida NCIB 9816-4 has been the subject of most study and contains a single 81 kb plasmid, pDTG1, to which the upper pathway naphthalene catabolic genes have been localised. NCIB 9816-3 possesses two plasmids; the large plasmid was designated NAH2 (Conners & Barnsley, 1982) or pWW60 (Cane & Williams, 1982). A smaller cryptic plasmid was reported by both groups. Although there is a discrepancy in the reported sizes of NAH2 (117 kb) and pWW60 (87 kb), both groups have shown that this plasmid is involved in the metabolism of naphthalene and salicylate. pWW60 appears to carry the genes encoding catechol *meta* cleavage enzymes, but their expression is largely blocked by a small insertion between the *nahG* and *nahH* genes. The presence of this insertion apparently channels the catechol formed from a plasmid mediated naphthalene degradation pathway through the chromosomally-encoded catechol *ortho* cleavage pathway.

A number of other naphthalene catabolic plasmids have been isolated. These include plasmid pBS4 which was originally identified in *P. fluorescens* BS29, which utilises naphthalene as sole carbon and energy source. After growth in the presence of naphthalene, whole cells and crude cell extracts of this strain were found to oxidise naphthalene and salicylate, but not catechol, indicating salicylate is degraded via the gentisic acid pathway (Yen & Serdar, 1988). Plasmid NPL-1 was identified in *P. putida* 12A which degrades naphthalene. NPL-1 carries a naphthalene upper pathway, but not lower pathway genes, for which expression depends on a host-supplied regulatory element (Yen & Serdar, 1988).

Some plasmids have been described which specify only decomposition of salicylate through the *meta* cleavage pathway and have no naphthalene upper pathway functionality; these are termed the SAL plasmids. *P. putida* R1 utilises salicylate as a sole carbon and energy source, which was attributable to a 85 kb plasmid designated SAL1 (Chakrabarty, 1972). SAL1 is a mutant of NAH7 and, although it encodes both upper and lower pathways for naphthalene and salicylate catabolism, the upper pathway is blocked by a 4.6 kb insertion

1.4.2 *nah*-like catabolic genes

The first report of the nucleotide sequence of naphthalene catabolic genes was from *P. putida* NCIB 9816 (Kurkela *et al.*, 1988). Based on the ability to oxidise indole the genes encoding naphthalene dioxygenase were cloned and the nucleotide sequence of a 2.5 kb region was determined. Three open reading frames (ORFs) were identified and designated *ndoA*, *ndoB*, and *ndoC*, although the functions of these genes were not determined. Some homology was observed to the components of benzene dioxygenase from *P. putida*, the only other initial dioxygenase for aromatic degradation to have been sequenced at that time. The complete nucleotide sequence of the *nahAaAbAcAd* genes of the NAH7 plasmid of *P. putida* G7 was reported in 1993 (Simon *et al.*, 1993). The N-terminal amino acid sequence was determined from the purified proteins for each of these genes which was used to locate the ORFs for each gene and to assign each a function. The *nahAa* gene encodes a reductase, *nahAb* encodes a ferredoxin, and the *nahAc* and *nahAd* genes encode the large (α) and small (β) subunits of the iron-sulfur protein (ISP). The nucleotide sequence of the *nahAa* and *nahAb* genes of the pDTG1 naphthalene-catabolic plasmid of *P. putida* NCIB 9816-4 was also determined and found to be 94% identical at the amino acid level to the analogous genes from NAH7. Based on this new information the *ndo* genes of *P. putida* 9816 were assigned functions: *ndoA* is the ferredoxin gene, and *ndoB* and *ndoC* are the α and β subunits of ISP. The *ndo* genes were 96% homologous to the *nahA* genes of *P. putida* G7. The reductase gene was not present on the 2.5 kb fragment from *P. putida* 9816 which expressed naphthalene dioxygenase activity in *E. coli*, which suggests *E. coli* contains a nonspecific reductase which can transfer electrons to the ferredoxin of naphthalene dioxygenase. The nucleotide sequence of other genes of the NAH7 plasmid of *P. putida* G7 has been reported: that of the naphthalene 1,2-dioxygenase gene, *nahC* (Harayama & Rekik, 1989), and the 2-hydroxychromene-2-carboxylate isomerase (*nahD*), hydratase-aldolase (*nahE*), and salicylaldehyde dehydrogenase (*nahF*) genes (Eaton, 1994).

Since the characterisation of the *nah* genes from *P. putida* G7 and the NCIB 9816 strains, several other catabolic pathways have been described which show high similarity to the *nah* genes. The genes of these pathways were given designations congruent with the substrate on which the strain was isolated, which is misleading

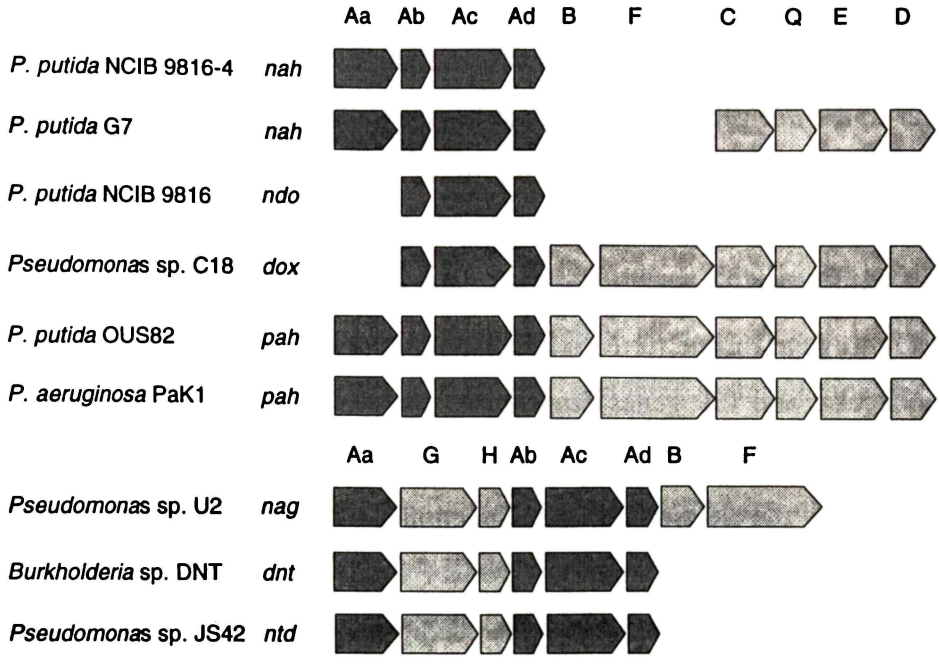


Figure 1.15 Summary of known nucleotide sequences for naphthalene/ phenanthrene and (di)nitrotoluene catabolic loci showing the conservation of gene order. Dark grey shading indicates initial dioxygenase genes, other catabolic genes have light grey shading.

considering the predicted amino acid sequences of the individual genes show >90% similarity to the *nah* genes, and the gene order is highly conserved between these loci (Figure 1.15). Thus, the *dox* genes were isolated from *Pseudomonas* sp. C18 that is able to degrade dibenzothiophene (Denome *et al.*, 1993), but since this strain will also utilise naphthalene the *dox* operon is in fact a naphthalene catabolic pathway. The *pah* genes of *P. putida* OUS82 (Takizawa *et al.*, 1994) and *P. aeruginosa* PaK1 (Takizawa *et al.*, 1996) specify the degradation of phenanthrene and are also >90% similar to the *nah* genes of *P. putida* G7. Since the *ndo*, *nah*, *dox*, and *pah* genes are so homologous they are often referred to as '*nah*-like' genes.

Three other loci for aromatic degradation have been reported in which the characterised initial dioxygenase genes show high homology to those of *P. putida* G7, but are less similar to the *nah* genes than *dox* and *pah*. The plasmid-encoded *dnt* locus from *Burkholderia* sp. DNT specifies the oxidation of dinitrotoluene (Suen *et al.*, 1996). The *dnt* dioxygenase components are 59–78% similar to *nah*-like sequences, and when expressed in *E. coli* were able to oxidise naphthalene. Naphthalene dioxygenase does not oxidise dinitrotoluene which suggests the *dnt*

system has diverged from a *nah*-like ancestor and acquired a greater substrate range. The *ntd* genes for 2-nitrotoluene degradation by *Pseudomonas* sp. JS42 are very similar to the *dnt* genes, and the dioxygenase components show 67–84% amino acid identity to naphthalene dioxygenase of *P. putida* 9816-4 (Parales *et al.*, 1996). Recently a third locus, *nag*, which also belongs to the *dnt/ntd* group, was described for the naphthalene degrading strain *Pseudomonas* sp. U2 (Fuenmayor *et al.*, 1998). The *nagAa* gene for the ferredoxin component of this dioxygenase is separated from the reductase and ISP genes *nagAbAcAd* by two genes (*nagGH*) which apparently encode the ISP subunits for a monooxygenase which oxidises salicylate to gentisate, possibly using the electron transport functions encoded by the *nagAaAb* genes. *nagG* and *nagH* show high homology (>95%) to two ORFs of previously unknown function which are present in the same positions in both the *dnt* and *ntd* gene clusters, and which therefore probably have the same function as *nagG* and *nagH*. Although the *nag* genes showed lower homology to *nah*-like sequences, the region upstream of the structural genes was highly similar suggesting the *nag* system has diverged from a *nah*-like ancestor. Since the gene sequences of *nag*, *dnt*, and *ntd* are similar, and all loci possess two additional genes present in the same arrangement Fuenmayor *et al.* (1998) suggest both *dnt* and *ntd* systems have evolved from the divergent *nag* system.

S. yanoikuyae B1 and *S. paucimobilis* EPA505 exhibit a diverse catabolic potential attributable to a complex arrangement of degradative genes. Characterisation of the catabolic loci from *S. yanoikuyae* B1 revealed many different genes with homologies to catabolic genes from xylene, biphenyl, and naphthalene pathways, including genes for as many as three different ISPs (Kim & Zylstra, 1995; Zylstra *et al.*, 1997). It appears that this locus does not specify the degradation of a single compound or group of closely related compounds, but rather encodes all of the necessary enzymes to catabolise a wide range of aromatic substrates. The presence of such a gene cluster explains the wide catabolic ability of *S. yanoikuyae* B1 and of other *Sphingomonas* strains.

1.4.3 Regulation of the naphthalene catabolic genes of NAH7

Activation of the *nah* operons of NAH7 requires both an inducer and the product of a regulatory gene. Both operons are inducible by salicylate and its structural analogue

2-aminobenzoic acid (Barnsley, 1975b; 1976b). Although both operons are induced during growth in the presence of naphthalene, naphthalene is not an inducer of the NAH7-encoded naphthalene oxidation pathway. The apparent induction effect of naphthalene appears to be due to its conversion into an inducer, namely salicylate, through a low-level constitutive expression of the first *nah* operon.

NAH7 possesses a regulatory gene, *nahR*, which controls expression of the two *nah* operons. *nahR* is located upstream from the *nahG* gene (Yen & Gunsalus, 1985) and its product, a LysR-type transcriptional regulator, regulates both *nah* operons at the transcriptional level (Schell, 1985). The binding sites of the *nahR* product have been determined by cloning DNA regions containing the promoters of the two *nah* operons. The promoters were more precisely mapped by locating regions protected from S1 nuclease digestion by salicylate-induced mRNA from *P. putida* cells containing NAH7 (Schell, 1986). The nucleotide sequences of the promoter regions were subsequently determined, and the approximate transcriptional start sites of the two *nah* operons were located. Two interesting features are characteristic of the *nah* promoter regions. Each of the promoters contains nucleotide sequences in the −10 to −35 regions reminiscent of an *E. coli* promoter, but despite such structural similarity, studies of gene expression of cloned *nah* operon genes suggest that the promoters of the *nah* operon do not function efficiently in *E. coli* (Schell, 1985; Yen & Gunsalus, 1985). The two *nah* promoter regions share significant sequence homology with each other, particularly in three distinct regions.

The *nahR* gene has been mapped in the immediate vicinity of the *nahG* gene and is transcribed in the opposite direction to that of the other *nah* genes (Figure 1.14). As in the *nah* promoters, the *nahR* promoter region contains sequences in the −10 and −35 regions which are similar to those found in the same regions of a typical *E. coli* promoter (Schell, 1986).

In *P. putida* the *nahR* gene appears to be expressed constitutively (Schell, 1985; 1986). This is analogous to the model for the regulation of the TOL pathway as discussed in section 1.2.3. Conceivably the NahR protein, like XylR, can exist in two forms, an inactive form ($NahR_i$) and an active form ($NahR_a$), which are in equilibrium. In the absence of an inducer, $NahR_i$ is the predominant form, and presence of an inducer shifts the equilibrium to the formation of $NahR_a$ (Yen & Serdar, 1988).

1.5 Evolution of pathways for aromatic catabolism

The isolation of bacteria able to degrade diverse aromatic compounds, and the subsequent detailed characterisation of the genetic basis for the catabolic pathways, has led to a comprehensive understanding of the mechanisms for aromatic degradation. Also emerging is an understanding of how these pathways came about, and why catabolic pathways for different aromatic substrates, and from different bacterial genera, show such similarity.

1.5.1 Modular organisation of catabolic pathways

A common theme for aromatic degradative pathways is the arrangement of genes in a modular structure consisting of discrete operons with a defined catabolic function and under specific regulatory control. Such genetic organisation has been an important factor in the evolution and distribution of aromatic catabolic pathways since these operons seem to have been disseminated amongst bacteria and evolved as whole genetic units. In generalised terms, aromatic pathways consist of two groups of catabolic genes. An 'upper pathway' operon transforms the initial aromatic compound into a common intermediate such as benzoate or salicylate, usually with the release of pyruvate which may enter central metabolism via the tricarboxylic acid (TCA) cycle. A 'lower pathway' operon usually consists of upstream genes and a *meta* cleavage pathway. The upstream gene(s) convert the metabolic product of the upper pathway to catechol, which undergoes ring cleavage via the *meta* pathways yielding metabolites which enter the TCA cycle (Figure 1.16). Catechol may also be processed through the *ortho* cleavage pathway.

Organisation of a degradative pathway into discrete operons under regulatory control allows the efficient regulation of catabolic steps and the 'metabolic isolation' of catabolic metabolites, a concept introduced by Houghton & Shanley (1994). The formation of metabolic bottlenecks is avoided because the concentration of any given catabolic intermediate is governed by the 'specific and coincidental' induction of the enzymes that led to its formation as well as those that provide for its further degradation.

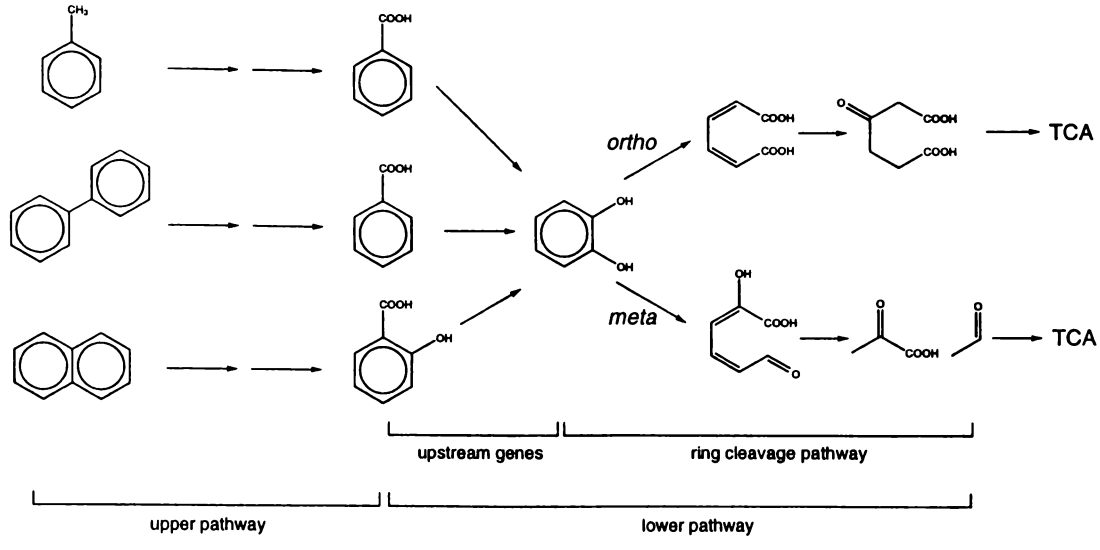


Figure 1.16 Modular organisation of aromatic catabolic pathways showing the role of the upper and lower pathways. These are generalised pathways for toluene (via the TOL pathway), biphenyl, and naphthalene via catechol and the *ortho* or *meta* pathways.

1.5.2 Accretion of genes to form a catabolic operon

Catabolic operons probably evolved by the recruitment of genes already part of the host genome which, for purposes of coordinate expression, became linked into a single regulon (Williams & Sayers, 1994). Biosynthetic pathways may have been a source of such genes as an anabolic enzyme may differ from a catabolic enzyme only by the position of equilibrium of the catalysed reaction.

The evolution of aromatic catabolic pathway genes into a single operon seems an unlikely event given the probable low availability of the chemically unstable intermediates as potential growth substrates in the environment. Prior to their acquisition into an aromatic pathway, individual genes could not have evolved to perform their catabolic functions due to the absence of the unstable substrates. Furthermore, a catabolic pathway would not endow any selective advantage to the host strain until it was fully functional and able to contribute metabolites such as pyruvate which may enter central metabolism and provide energy to the organism. Therefore the accretion of a single gene into an incipient pathway would have no selective advantage, and be of no benefit to the organism until the pathway was complete.

For example, catechol, the substrate for the *meta* cleavage pathway, is an unstable

compound and unlikely to persist in the environment for any length of time. Therefore, the ability to degrade catechol would be of little advantage to an organism unless catechol was being produced in the same cell, or it was close to a source of exogenous catechol. However, there is considerable selective advantage of acquiring an enzyme which generates catechol from a stable compound which exists in the environment, or from a catabolite of an upper pathway (eg. benzoate, salicylate, phenol). This may have been the selective pressure which caused fusions of genes coding for the generation of catechol to *meta* pathways.

Any regulatory functions of the operon were probably imposed after assembly of the catabolic unit since examples exist of apparently evolutionary related pathways which have different modes of regulation, for example the *nah* and *xyl* systems. Such regulatory functions were probably recruited from elements of regulatory systems present in the host genome (Williams & Sayers, 1994).

1.5.3 Evolutionary lineages of aromatic pathways

Aromatic compounds have probably been present in the environment for millions of years through events such as forest fires (which result in the pyrolysis of lignin) and natural oil seeps. Aromatic catabolic pathways have therefore had millions of years to evolve to their present form, and it has been suggested that the pathways present now evolved from a small number of common ancestors. The evolutionary lineages of the described 'modern' pathways can be speculated based on comparisons of nucleotide and predicted amino acid sequences of individual genes, and the order that genes are arranged within a catabolic operon, which reveals patterns between aromatic pathways. For example, the biphenyl, toluene (*tod*), and benzene dioxygenase genes are arranged in the same order, and the individual genes show high homology, which suggests these genetic units evolved from a common ancestor (Williams & Sayers, 1994). In comparison, the *nah* naphthalene dioxygenase genes are arranged differently and show lower sequence homology to this group, which suggests a more distant evolutionary relationship. As an indication of the time scale of such evolutionary processes it has been calculated that the level of nucleotide and amino acid sequence homology between different *meta* pathways suggests evolution from a common ancestral sequence, and that divergence from this common ancestor occurred about 50 million years ago (Harayama & Rekik, 1993).

Some substituted aromatic compounds are formed only through anthropogenic activities and have only become ubiquitously distributed since the industrial revolution, eg. PCBs, nitrotoluenes, and nitrogen-heterocycles. Pathways for the degradation of these 'novel' aromatic compounds have had less time to evolve, and in most cases have apparently diverged from a more deeply rooted catabolic pathway. Either a relaxation or change in substrate specificity of pathway enzymes allows the catabolism of aromatic substrates with a slightly different chemical structure. Situations where this has occurred are evident since in many cases the gene order and upstream regulatory sequences are conserved.

An example where a recent divergence has occurred which allows for the dissimilation of a 'novel' aromatic substrate is the metabolism of (di)nitrotoluene by the *dnt* and *ntd* systems (Parales *et al.*, 1996; Suen *et al.*, 1996). These loci apparently evolved from a *nag*-like precursor (for naphthalene degradation), which itself evolved from the *nah*-like group by the acquisition genes which allow degradation of salicylate through the gentisate pathway (Fuenmayor *et al.*, 1998). Pathways for the degradation of (di)nitrotoluenes may have arisen form *nag* rather than *nah* due to an advantage in metabolising these compounds through the gentisate pathway, or because the divergent *nag* initial dioxygenase is more adaptable to changes in substrate specificity than *nah*. Another example is the *dox* genes for dibenzothiophene (and naphthalene) degradation (Denome *et al.*, 1993) which are almost identical to the *nah* genes of *P. putida* G7. Presumably a small number of point mutations in the initial dioxygenase genes has extended the substrate range of the *dox* pathway to allow dissimilation of dibenzothiophene in addition to naphthalene.

1.5.4 Motility of catabolic pathways

It is likely the modular organisation of aromatic pathways has been a key factor in the dissemination of functional catabolic units between bacteria. The mechanisms by which these units may be passed from one strain to another are plasmid transfer and transposition.

Although regions of nucleotide sequence encoding catabolic genes have often been associated with plasmids (Conners & Barnsley, 1982; Saylor *et al.*, 1990), the importance of plasmids for aromatic degradation may have been overestimated. For

example, much has been made of the frequent observation that genes for the *meta* cleavage of catechol are usually plasmid-borne, whilst genes for the *ortho* pathway are found on the chromosome of the same organism. It now seems apparent that the location of catabolic genes is less significant since many catabolic operons are carried on transposons, and the site of integration is probably a random event. Furthermore, the perceived motility of catabolic operons which has resulted in the widespread dissemination of conserved genetic units for aromatic degradation need not be explained by their presence on conjugative plasmids, since transposition events would account for such motility.

Examples of catabolic genes which apparently reside on transposons are found for degradative pathways for diverse aromatic substrates. The TOL plasmid pWW0 contains two overlapping transposons of 56 kb and 70 kb which both carry a complete set of *xyl* genes (Assinder & Williams, 1990). The NAH7 plasmid of *P. putida* G7 for naphthalene degradation carries a defective 37.5 kb transposon which includes all of the *nah* genes (Tsuda & Iino, 1990) but does not encode a functional transposase. The genes for biphenyl degradation in *A. eutrophus* A5 (Springael *et al.*, 1993), for chlorobenzene degradation in *Pseudomonas* sp. P51 (van der Meer *et al.*, 1991) and *Alcaligenes* sp. BR60 (Nakatsu *et al.*, 1991), and for biphenyl degradation on the pWW100 plasmid of *P. putida* CB406 (Lloyd-Jones *et al.*, 1994) are also carried on transposons.

1.5.5 Mechanisms for evolution of catabolic pathways

A number of evolutionary mechanisms have played a part in moulding catabolic genes and operons into the forms in which they presently exist. The most basic evolutionary change is via point mutations and multiple nucleotide substitutions; these events have contributed to evolution of *meta* pathways (Harayama & Rekik, 1993). Deletions of DNA regions also occurs, for example the *nahD* isomerase gene of *P. putida* G7 has a 35 bp deletion near the end of the gene which does not effect the functionality of the transcribed enzyme, which is four residues longer than two other *nahD*-like isomerases since the deletion included the stop codon for this gene (Eaton, 1994). It is interesting to note that another *nahD*-like isomerase gene, *pahD* of *P. putida* OUS82 (Takizawa *et al.*, 1994), also has this 35 bp deletion which indicates these two *nah*-like pathways diverged from a common ancestor more

recently than the other *nah*-like pathways. DNA slippage and strand exchange is a mechanism which may account for the divergence of benzoate dioxygenase genes from *A. calcoaceticus* (*benABC*) and *P. putida* (*xylXYZ*) (Harayama *et al.*, 1991b). Recombination events resulting in the insertion or replacement of DNA sequences is another mechanism, for example the acquisition of genes encoding salicylate 5-hydroxylase resulting in the *nag* operon (Fuenmayor *et al.*, 1998).

1.6 Limits of current knowledge

It seems appropriate to conclude this literature review with a comment about the limits of current understanding of microbial degradation of aromatic compounds in the natural environment. As earlier discussed in section 1.2, the study of catabolic bacteria has invariably focused on those strains which grow best under laboratory culturing conditions, ie. the strains which grow fastest in enrichment cultures or during direct isolation techniques. This bias is readily apparent when one surveys the strains discussed in this review: the vast majority are *Pseudomonas* spp. or from related genera such as *Burkholderia* or *Sphingomonas*. Through the extensive study of cultured environmental strains has evolved a common belief that pseudomonads constitute the most versatile and prevalent microorganisms in soil and water ecosystems. Linked to this belief is the idea that since the catabolism of aromatics by pseudomonads is well understood, this equates to an excellent grasp of the predominant mechanisms for aromatic degradation in soil and water environments. However, recent studies of microbial community composition in contaminated soils by analysing 16S rRNA genes amplified using PCR from DNA extracted directly from soil, has revealed that *Pseudomonas* spp. make up only a small fraction of the microbial community, and that other types of bacteria are numerically more significant (Lloyd-Jones & Lau, 1998; Timmis, 1997). Furthermore, when *Pseudomonas* is introduced into a variety of habitats, it establishes a stable population which represents 0.01–1% of the total microbial community (Pipke *et al.*, 1992). Even at these population levels, pseudomonads probably do play an important role in the biodegradation of aromatic compounds, but there must be at least as equally important contributions made by the remainder of the microbial population. Options for studying microbes in this group are limited because they are difficult or

impossible to culture, and gene probes for detecting diverse catabolic genes will inevitibly be based on genes from well-characterised *Pseudomonas*-like strains.

The study of aromatic-degrading *Rhodococcus* and *Mycobacterium* strains is contributing to the understanding of aromatic catabolism by non-pseudomonads, but knowledge of these bacteria is limited at this stage.

CHAPTER 2

Experimental Methods

2.1 Growth and maintenance of bacterial strains

2.1.1 PAH degrading strains

Wild-type strains were routinely grown and maintained on half-strength Plate Count Agar (½PC) (Difco) and on solid selective media using a method similar to that of Bogardt & Hemmingsen, 1992. An agar base containing minimal media (MM) (4 g L^{-1} Na_2HPO_4, 2 g L^{-1} KH_2PO_4, 1 g L^{-1} $(NH_4)_2SO_4$ and 2 ml L^{-1} Herberts salts solution (Rosenberger & Elsden, 1960)) solidified with 1.5% purified agar (Difco) was overlaid with 4 ml MM containing 1% ultra pure LMP agarose (Gibco BRL) and 100 µl of a solution containing phenanthrene (Aldrich) dissolved in dimethylsulphoxide (DMSO) to a concentration of 1 g 20 ml^{-1}. The phenanthrene forms a fine white precipitate in this overlay medium which gives the overlay agar plate a cloudy appearance. When growing on this media, phenanthrene degrading strains produce zones of clearing surrounding the colonies as the phenanthrene in the vicinity is metabolised.

2.1.1.1 *Burkholderia* sp. strain RP007

Growth of RP007 on phenanthrene and naphthalene as sole sources of carbon was confirmed in liquid medium. For phenanthrene the strain was grown in MM amended with phenanthrene crystals which were added to 0.05% (wt vol^{-1}) in acetone (PTFE sterilised) to sterile flasks, the acetone allowed to evaporate before sterile minimal media was added to the flask. It was found that in this way the phenanthrene formed very fine crystals in the medium resulting in better growth compared to addition of the phenanthrene to the medium before autoclaving or in DMSO. Growth of RP007 on naphthalene required first incubation of the inoculated MM with naphthalene present in the headspace of the flask and therefore available only as a vapour. Once turbidity was observed (2–3 days) naphthalene crystals were added to the medium and subsequent growth occurred rapidly.

For all growth experiments inoculated flasks were incubated at 28°C with shaking at 200 rpm. Growth was measured by monitoring increase in A_{600} or by protein estimation using the Pierce BCA Protein Assay Reagent Kit (Pierce Chemical Company, Rockford, Illinois, USA) which uses the bicinchoninic acid reagent.

Bovine serum albumin (BSA) was used as a standard.

For DNA preparations RP007 was grown to A_{600}=1.0 in ½PC broth.

Burkholderia sp. strain RP007 has been deposited in the culture collection at Landcare Research (ICMP, c/o Landcare Research NZ Ltd, Private Bag 92170, Auckland) as strain number 13529.

2.1.1.2 PAH degrading culture collection

The naphthalene and phenanthrene degrading strains described in Chapter 6 were isolated by Rhonda Fraser and David Hunter from 11 different hydrocarbon-contaminated soils. 5 g of each soil was added to 45 ml of sterile 0.1% (wt vol^{-1}) *tetra* sodium pryophosphate containing 15 g glass beads in a 100 ml Schott bottle and shaken at 200 rpm for 1 hour. Serial dilutions made in 0.1% *tetra* sodium pryophosphate were directly plated onto MM agar plates. Isolates were cultured using either naphthalene or phenanthrene as sole carbon and energy source to increase the diversity of strains in this culture collection. Naphthalene was supplied as a vapour by incubation of unsealed petri dishes in a dessicator containing naphthalene crystals placed in its base, and phenanthrene vapour was provided by placing crystals within the lids of sealed petri dishes. Colonies were selected following up to 6 weeks of incubation at 28°C.

All naphthalene and phenanthrene degrading colonies were screened for growth at the expense of naphthalene, phenanthrene, biphenyl, phenol, and toluene. Colonies grown on ½PC were screened for ability to oxidise indole (resulting in accumulation of the blue product indigo) by placing indole crystals in the lid of a sealed petri dish, and catechol by spraying colonies with a 0.1 M aqueous solution of catechol.

2.1.2 *Escherichia coli* strains and recombinant constructs

Escherichia coli strains and recombinants were grown and maintained on Luria Bertani (LB) broth (10 g L^{-1} Bacto-tryptone (Difco), 5 g L^{-1} NaCl, 5 g L^{-1} Bacto-yeast extract (Difco) and LB agar (1.6%), amended if necessary with ampicillin (100 μg ml^{-1}). *E. coli* recombinant strains were also grown on Terrific Broth for use with some plasmid preparations (to a base broth (12 g Bacto-tryptone (Difco), 24 g Bacto-yeast extract (Difco), 4 ml glycerol; made up to 900 ml and autoclaved) 100 ml of a sterile solution of 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 was added). Growth of LB

broth and Terrific Broth cultures were at 37°C with shaking at 225 rpm.

2.1.3 Long term storage of strains

All bacterial strains and recombinant constructs used in this study were maintained at −70°C using the Protect system (Technical Service Consultants Ltd, Bury, England).

2.2 Nucleic acid manipulations

2.2.1 Routine molecular biology techniques and equipment

The molecular biology techniques and reagents used in this study were according to those described in Sambrook *et al.* (1989). Electrophoresis conditions used a 1× TBE or 1× TAE buffer system with 0.7–1.0% agarose gel, and were run at approximately 8 volts cm^{-1}. For size estimation of DNA bands a 1 kb DNA ladder standard (Gibco BRL) was routinely run alongside all samples. Gels were visualised at 302 nm using a UVP Model TM-15 ultraviolet transilluminator (UVP Inc., San Gabriel, California, USA). For microcentrifugation an Eppendorf Centrifuge 5410 was used, which had a single speed setting giving a maximum centrifugal force of 12800 ***g***. During DNA preparations ethanol precipitated DNA was dried using an Eppendorf Concentrator 5301 set at 30°C and 240 ***g***, attached to a Savant GP110 gel pump. For fixing Southern hybridisation membranes and drying SDS-PAGE gels a Bio-Rad Model 583 Gel Dryer was used, connected to the Savant GP110 gel pump. For Southern blotting procedures a Bio-Rad Model 785 vacuum blotter was used and the vacuum maintained at 5 in Hg using a Bio-Rad vacuum regulator. The ultracentrifuge used for caesium chloride gradients and enzyme purification procedures was a Sorvall RC M120EX (Dupont) using a RP 80AT fixed angle rotor. For centrifugation of large volumes of bacterial cultures a Sorvall RC-5B (Dupont) centrifuge was used with either an SS-34 or GSA rotor. For PCR and RT-PCR a Techne Cyclogene Dri-Block Cycler with HL-1 Heated Lid (Techne Ltd, Cambridge, England) was used.

All restriction enzymes, T4 DNA ligase, Proteinase K, bacterial alkaline phosphatase (BAP), *Taq* DNA polymerase, and dNTPs were supplied by Gibco BRL and were used according to the manufacturers instructions. RNase was supplied by Pharmacia Biotech. Ampicillin, X-gal (5-bromo-4-chloro-3-indoyl-β-D-galactoside) and IPTG

(isopropyl-β-D-thiogalactoside) were supplied by Boehringer Mannheim and were used as described by Sambrook *et al.* (1989).

For routine phenol extractions of DNA preparations phenol:chloroform:isoamyl alcohol (IAA) (25:24:1) (pH 8.0) prepared according to Sambrook *et al.* (1989) was used. Generally an equal volume of this reagent was added to the DNA solution, the tube was vortexed, then centrifuged to separate the phases, and the upper aqueous layer removed to a clean tube, avoiding any white residue at the interface.

Ethanol precipitation of DNA was by the addition of 0.1 volumes 3 M sodium acetate (pH 5.2) and 2 volumes 100% ethanol, incubation overnight at –20°C or on ice for 30 mins, followed by centrifugation (12800 ***g***, 15 mins) to pellet DNA. The pellet was washed with 70% ethanol to remove any salt associated with the DNA, then dried at 30°C in the vacuum concentrator and resuspended in an appropriate volume of TE buffer (pH 7.5) or sterile distilled water, and stored at –20°C.

For cloning PCR products into pUC18 for sequence analysis, a SureClone Ligation Kit (Pharmacia Biotech) was used. According to this method the single-base 3' overhangs of PCR products are removed by the Klenow fragment of DNA polymerase I, and the ends of the PCR fragment are concomitantly phosphorylated by T4 polynucleotide kinase. The PCR product is column purified and ligated to BAP-treated *Sma*I-pUC18.

For creating blunt-ended DNA fragments for subcloning restriction fragments with ends not compatible with pUC18 multiple cloning site restriction sites the S1 nuclease from a *double-stranded* Nested Deletion Kit (Pharmacia Biotech) was used. The manufacturers instructions were modified to account for the increased volume of the reaction compared to that for the deletion reaction.

2.2.2 Nested deletion experiments

To generate nested deletion derivatives of pUC18 constructs containing cloned DNA fragments for nucleotide sequencing, a *double-stranded* Nested Deletion Kit (Pharmacia Biotech) was used. This kit uses exonuclease III which is a 3'-exonuclease and is only active towards blunt and 5'-overhanging ends of double-stranded DNA. Because 3'-overhanging ends which are three or more bases in length are resistant to digestion by this enzyme, digestion of a recombinant clone with the appropriate restriction enzymes allows the deletion reaction to proceed in one

direction only, ie. into the cloned fragment. Aliquots of the deletion reaction are removed at time intervals, treated with S1 nuclease to remove single-stranded regions generated by the exo III, then ligated to recircularise the deleted clone. The result is a library of nested deletion subclones which contain inserts which have been deleted from one end only. The library is screened for deletion derivatives which have the appropriate size insert, and these are sequenced from a primer located on the vector adjacent to the site from where the deletion began. In this way the whole insert may be sequenced from the same primer as deletion of greater lengths of the insert allows sequencing further into the fragment.

For nested deletion of pUC18 constructs, restriction sites in the multiple cloning site (MCS) were chosen which were necessary to ensure deletion in one direction only. Sequencing of deletion derivatives was from the M13/pUC Forward or M13/pUC Reverse sequencing primers which are located either side of the pUC18 MCS. To sequence both strands of a cloned fragment, the fragment was cloned into pUC18 in both orientations which allowed deletion and subsequent nucleotide sequencing from both ends of the fragment using the same restriction enzymes and the same sequencing primer.

2.2.3 Nucleotide sequencing

Nucleotide sequencing was performed by the Waikato DNA Sequencing facility using a PRISM Ready Reaction DNA Terminator Cycle Sequencing Kit (Perkin-Elmer), and the reactions were resolved using an ABI model 377 automatic sequencer. The electronic sample files from the sequencing reaction were processed with the ABI Fractura computer programme to remove vector sequence and regions of low confidence (usually after 600 bp for each reaction), and the sequence data was assembled using the ABI Assembler computer programme.

2.2.4 Total DNA preparations from wild-type strains

A modification of the method of Ausubel *et al.* (1989) was used to prepare total genomic DNA extracts from the wild-type strains used in this study. This procedure uses hexadecyl trimethyl ammonium bromide (CTAB) to aid separation of the polysaccharide components of the lysed cells from the DNA fraction. Isolates were grown in 100 ml ½PC broth to late-log phase (48 hours) at 28°C with shaking (200

rpm). Cells were harvested by centrifugation (6000 ***g***, 10 mins) and resuspended in 4 ml TE buffer (pH 8.0). Lysozyme (Sigma) was added to 5 mg ml^{-1} and cells were incubated at 37°C for 1 hour, then Proteinase K was added to 100 μg ml^{-1} and SDS added to 0.5% and cells incubated at 65°C for 2.5 hours; the SDS concentration was then increased to 1% and incubation at 65°C continued for a further 30 mins. NaCl was added to 0.7 M and a 10% CTAB (Sigma)/0.7 M NaCl solution added to 1%, and incubation was continued at 60°C for 30 mins. The cell lysate was then extracted with an equal volume of phenol/chloroform/IAA, incubated on ice for 10 mins before centrifugation (10000 ***g***, 10 mins, 4°C) and removal of the top layer to a new tube. The phenol/chloroform/IAA extraction was repeated, then the DNA was precipitated by the addition of 0.6 volumes isopropylalcohol and incubation overnight at –20°C. The DNA was collected by centrifugation (10000 ***g***, 20 mins, –10°C) and the pellet washed in 70% ethanol, dried briefly, then resuspended in 4 ml sterile distilled water for 4 hours at 4°C with occasional mixing to aid solubilisation of the DNA. This solution was RNase treated by the addition of RNase to 50 μg ml^{-1} and incubation at 37°C for 1 hour, followed by addition of Proteinase K to 100 μg ml^{-1} and incubation at 65°C for 20 mins to inactivate RNase. The DNA solution was then extracted with an equal volume of phenol/chloroform/IAA as before and the DNA precipitated by the addition of 0.1 volumes 3 M sodium acetate and 2.5 volumes 100% ethanol and incubation overnight at –20°C. DNA was collected by centrifugation, washed with 70% ethanol and dried as before, and resuspended in 3 ml TE (pH 7.5) or sterile distilled water by mixing gently overnight at 4°C.

If necessary this DNA preparation was further purified using a caesium chloride gradient. CsCl (Boehringer Mannheim) was added to the DNA solution to give 1 g ml^{-1} and this solution was transferred to a sealable 4 ml Sorvall ultracentrifuge tube. Ethidium bromide was added to give a final concentration of 740 μg ml^{-1}, and the tube was topped up with a 1 g ml^{-1} CsCl/TE (pH 7.5) solution and sealed. To form the gradient, tubes were usually centrifuged at 70000 rpm for 16 hours at 19°C. To recover the DNA from the gradient the tube was clamped in a vertical position, the sealing screw removed, and a hypodermic needle attached to a 1 ml syringe was inserted through the tube wall underneath the DNA band, which was usually visible without illumination by UV. The DNA was carefully drawn into the syringe and transferred to a microcentrifuge tube. To remove the ethidium bromide the DNA

solution was extracted 3–4 times with an equal volume of CsCl- and water-saturated isopropylalcohol, and finally dialysed against 2–4 L deionised water. The DNA solution was stored at –20°C.

2.2.5 Plasmid isolation procedures

2.2.5.1 Alkaline lysis plasmid 'mini-prep'

For screening of large numbers of transformants for the routine characterisation of recombinant constructs the following plasmid 'mini-prep' was used. This method uses an alkaline lysis procedure to selectively purify supercoiled plasmid DNA, which is subsequently RNase treated and ethanol precipitated. This method is a modification of that described by Zagursky *et al.* (1985).

Cultures were grown overnight at 37°C with shaking (225 rpm) in 5–10 ml LB broth amended with ampicillin (100 μg ml^{-1}). 1.5–3 ml of each culture was pelleted in a microcentrifuge tube, then cells were resuspended in 200 μl cold buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris/HCl pH 8.0). Cells were lysed by the addition of 400 μl of a freshly made lysis solution (0.2 N NaOH, 1% (wt vol^{-1}) SDS) and the tubes inverted several times and placed on ice for 2–3 mins to ensure complete lysis. 300 μl cold 3 M potassium acetate was added to precipitate cellular debris and tubes were incubated on ice for 15 mins, then centrifuged for 3 mins. 750 μl of the supernatant was removed to a new tube and 0.45 ml cold isopropanol added to precipitate plasmid DNA. Tubes were centrifuged for 5 mins to pellet DNA, the pellet was washed in 70% ethanol before being resuspended in 100 μl TE buffer. After incubation with RNase (250 μg ml^{-1}) for 10 mins at 37°C, DNA was precipitated by the addition of 0.1 volumes 3 M sodium acetate (pH 5.2) and 2.5 volumes cold 100% ethanol. Tubes were centrifuged for 5 mins, the DNA pellet washed with 70% ethanol and dried before being resuspended in 50 μl TE buffer or sterile distilled water. Plasmid DNA was analysed by restriction digestion and stored at –20°C.

2.2.5.2 Bio-Rad Quantum Prep Plasmid Miniprep Kit

The Bio-Rad Quantum Prep provided a rapid and simple procedure for obtaining high quality DNA for direct nucleotide sequencing. The method was performed

following the manufacturers instructions and using the materials and reagents supplied. It was used for the preparation of known constructs for nucleotide sequencing which had first been identified using the manual plasmid mini-prep. The Quantum Prep kit uses the silicon dioxide exoskeleton of diatoms as the DNA binding matrix. *E. coli* cells (1–2 ml) grown in terrific broth are lysed using an alkaline lysis procedure and DNA is bound by the addition of the matrix suspension. The bound DNA/matrix solution is centrifuged through a spin filter where the matrix collects, and is then rinsed with a wash buffer. The DNA is eluted by applying sterile distilled water or TE buffer to the filter/matrix and centrifuging for 1 minute.

2.2.5.3 QIAGEN QIAprep Spin Miniprep Kit

The QIAGEN QIAprep Spin Miniprep Kit was also used for preparation of constructs for nucleotide sequencing. The method is similar to that of the Bio-Rad Quantum Prep but DNA is bound to a silica-gel membrane. The protocol provided by the manufacturers was followed using the materials and reagents supplied.

2.2.5.4 Promega Wizard Midiprep DNA purification system

The Promega Wizard Midiprep (Promega Corporation, Wisconsin, USA) was used for large scale plasmid preparations and yielded up to 200 µg of pUC18-based plasmid DNA from a 100 ml overnight LB culture. The procedure uses an alkaline lysis step followed by addition of a silica-based resin to bind DNA. The solution is drawn through a column attached to a vacuum manifold and the DNA-resin in the column is rinsed by the application of wash buffers. DNA is finally eluted in sterile distilled water or TE buffer. This method was used according to the manufacturers instructions and using the materials and reagents supplied.

2.2.5.5 Sucrose gradient procedure for wild-type strains

The method of Wheatcroft & Williams (1981) was used to prepare plasmid DNA from wild-type strains. This procedure partially denatures chromosomal DNA by the combined effects of mild shearing forces and SDS/NaOH treatment, and plasmid DNA is separated from the partially denatured chromosomal DNA and cellular debris by centrifugation through a preformed sucrose gradient.

Strains were grown to late-log phase (2 days) in 100 ml ½PC broth, and harvested by centrifugation (6000 ***g***, 10 mins). Cells were washed in fresh MM, then pelleted and

resuspended in 2 ml Reagent A (50 mM Tris base, 50 mM Na_2EDTA, 5% (vol vol^{-1}) Dow Corning Antifoam RD emulsion (BDH), 0.1 mg ml^{-1} xylene cyanol). 0.5 ml Reagent B (1 M NaOH saturated with SDS) was added and the solution gently mixed by hand for 2–3 mins until the xylene cyanol becomes green, indicating cell lysis. The solution is then vortexed vigorously for 3–5 mins until the viscosity decreases significantly. 2 ml of this preparation is layered onto a preformed 20% (wt vol^{-1}) sucrose gradient which is centrifuged (100,000 ***g***, 1 hour, 17°C) in a SW40 Ti swing-out rotor using a Beckman L8-55 centrifuge. 1 ml fractions were recovered from the gradient by siphoning from the bottom of the tube, and those containing DNA were pooled and centrifuged (300,000 ***g***, 1 hour). The lower 200 µl containing plasmid DNA was collected for subsequent analysis.

2.2.6 Frozen competent cell preparation

A 10 ml LB broth overnight culture of the desired *E. coli* strain was subcultured (1:100) into fresh LB broth and incubated at 37°C with shaking to A_{600}=0.3–0.35. The culture was chilled on ice for 5 mins, then cells were harvested by centrifugation (3000 ***g***, 5 mins, 4°C) and resuspended in $^2/_5$ of original volume in Buffer 1 (30 mM potassium acetate, 100 mM RbCl or KCl, 10 mM $CaCl_2.H_2O$, 50 mM $MnCl_2.4H_2O$, 15% (vol vol^{-1}) glycerol, pH adjusted to 5.8 with 0.2 M acetic acid, filter sterilised). Cells were left on ice for 5 mins, then harvested (3000 ***g***, 5 mins, 4°C) and resuspended in $^1/_{25}$ volume Buffer 2 (10 mM MOPS or PIPES, 75 mM $CaCl_2.2H_2O$, 10 mM RbCl, 15% (vol vol^{-1}) glycerol, pH adjusted to 6.5 with KOH, filter sterilised). Cells were left on ice for 15 mins, then 100 µl aliquots were dispensed into microcentrifuge tubes and snap-frozen in liquid nitrogen, then stored at –70°C.

2.2.7 Transformation procedure

Frozen competent cells were removed from –70°C and thawed at room temperature, then incubated on ice for 10 mins. 1–5 µl (50–100 µg) DNA was added to the cells and gently mixed, then cells were further incubated on ice for 30 mins. Cells were heat-shocked at 42°C for 45 seconds, then placed on ice for 2 mins. 0.9 ml LB broth or SOC medium (Sambrook *et al.*, 1989) was added to the cells which were then incubated at 37°C with shaking for 1 hour to allow expression. Cells were plated on to LB agar containing ampicillin (100 µg ml^{-1}).

2.2.8 Gel purification of DNA fragments

For subcloning of recombinant constructs or the preparation of labelled probes for hybridisation experiments it was necessary to prepare a DNA solution containing only the desired restriction fragment. DNA fragments were purified from a 0.7% 1× TAE agarose gel using three different techniques. The electroelution and glass wool methods were used during the early stages of this project, after which time the QIAGEN kit became available and was then used routinely. Excision of the gel slice containing the desired DNA band necessitated visualisation using UV, and to minimise nicking of the DNA this was always performed using the low intensity lamp on the UVP transilluminator.

2.2.8.1 Electroelution into dialysis bags

The desired DNA band was excised from the gel and placed in dialysis tubing prepared according to Sambrook *et al.* (1989) and prewashed in distilled water. 1 ml TE buffer was added to the bag before it was sealed at both ends and placed into the gel tank containing 1× TAE buffer so that the contents of the bag were completely submerged. The tank was run at 50 volts for 1 hour to elute the DNA out of the agarose, after which time the polarity was reversed for 45 seconds at the same voltage to desorb the DNA from the dialysis membrane. The buffer was then removed from the bag and phenol extracted, then ethanol precipitated and resuspended in 20 µl sterile distilled water.

2.2.8.2 Centrifugation through siliconised glass wool

Glass wool siliconised with Sigmacote (Sigma), a chlorinated organopolysiloxane compound, was packed into a 500 µl PCR tube with a pin hole in the bottom of the tube, which was placed into a 1.5 ml microcentrifuge tube for support. The excised gel fragment containing the DNA band to be purified was applied to the tube containing the glass wool and centrifuged for 2 mins. 100 µl TE was added to the glass wool and the tube centrifuged again for 2 mins. The eluted buffer containing the DNA was phenol extracted, then ethanol precipitated and resuspended in 20 µl TE buffer or sterile distilled water.

2.2.8.3 QIAquick Gel Extraction Kit (QIAGEN)

This kit provided a rapid and convenient method for the efficient recovery of DNA from agarose fragments. All steps were performed according to the manufacturers instructions. The slice of gel is first dissolved in buffer at 50°C, then centrifuged through a column containing a silica-gel membrane which adsorbs to DNA in the presence of high concentrations of salt. After washing steps the DNA is eluted in sterile distilled water. The eluted DNA often appeared contaminated with traces of ethanol, presumably carried over from the wash steps, and this was subsequently removed by evaporation or an ethanol precipitation step.

2.2.9 Southern blotting and hybridisation

The following method is the salt transfer protocol given by the manufacturers of the nylon membrane (NEN Life Science Products). After electrophoresis the 1× TBE gel containing the DNA to be probed was first depurinated by soaking in 0.25 N HCl for 10 mins, then rinsed twice with distilled water to remove excess HCl. The gel was then denatured in 0.4 N NaOH/0.6 M NaCl for 30 mins, then soaked in 1.5 M NaCl/0.5 M Tris-HCl pH 7.5 for a further 30 mins. The gel was blotted onto a GeneScreen nylon membrane (NEN Life Science Products), previously equilibrated in 10× SSC for 10 mins, using a Bio-Rad Model 785 vacuum blotter and a Bio-Rad vacuum pump and vacuum regulator to maintain the vacuum at 5 inches Hg. The gel was transferred in 10× SSC for 90 mins, then the membrane was agitated in 0.4 N NaOH for 1 minute to denature the DNA, then neutralised in 0.2 M Tris-HCl pH 7.5/1× SSC for 1 minute. The membrane was fixed by baking for 30 mins at 80°C under vacuum in the Bio-Rad gel dryer, then stored in a sealed plastic bag at room temperature until used for hybridisation.

Hybridisation experiments used either a non-radioactive method or a [^{32}P]-labelled probe. The Renaissance Random Primer Fluorescein Labelling Kit (NEN Life Science Products) is a non-radioactive DNA labelling and detection system which uses primer extension from random hexamers to incorporate Fluorescein-N6-dATP and other nucleotides to synthesise labelled DNA from denatured template DNA. The probe is denatured at 100°C before incubation with the membrane-bound target DNA in a buffer containing a blocking reagent. Hybridisation was overnight at 65°C in a

hybridisation oven using rolling cylinders and the stringency washes were: 2× SSC, 0.1% SDS for 15 mins at 65°C followed by 0.2× SSC, 0.1% SDS for 15 mins at 65°C. Following stringency washes the membrane is blocked with another buffer containing a blocking reagent to prevent non-specific protein binding, then incubated with Antifluorescein-HRP, a highly specific detector polyclonal antibody conjugate which reacts with the Fluorescein-labelled DNA. A chemiluminescent signal is generated by incubating the membrane with a chemiluminescent reagent, and the membrane is immediately exposed to Kodak X-OMAT AR autoradiography film (Eastman Kodak Co., New York, USA) for 30 mins and the film developed using Kodak reagents. In some cases exposure time was extended overnight to visualise very faint signals. The DNA labelling reaction for preparation of fluorescein-labelled probe, and all hybridisation procedures and stringency washes, were performed according to the manufacturers instructions and using the reagents and enzymes provided with the Renaissance kit.

[^{32}P]-labelled probes were prepared using a Prime-It II Random primer labelling kit (Stratagene) using [α-^{32}P]dCTP or [α-^{32}P]dATP (Sigma). Hybridisation and stringency wash conditions were as described for the non-radioactive method. Following the final post-hybridisation wash, wet membranes were placed between plastic sheets and exposed to autoradiography film at –70°C using an intensifying screen.

2.2.10 Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

Cell pellets from bacterial cultures were snap-frozen in liquid nitrogen and stored at –70°C until used for RNA purification. For isolation of RNA from cells the RNeasy Total RNA Kit (QIAGEN) was used. This procedure involves first lysing cells by incubation with lysozyme, followed by the addition of a lysis buffer containing guanidinium isothiocyanate, which creates highly denaturing conditions to inactivate RNases. The sample is applied to a silica-gel based membrane, washed with ethanol-based wash buffers, and RNA is eluted in 40 µl RNase-free water.

RNA samples were treated with DNase (Deoxyribonuclease I, Amplification Grade (Gibco BRL)) according to the manufacturers instructions. This step was necessary to ensure products amplified by PCR were derived from RNA and not contaminating DNA. Control PCR replicates not subjected to RT-PCR were included in all

experiments to show that no DNA template was present in the RNA preparations.
For RT-PCR the Titan One Tube RT-PCR System (Boehringer Mannheim) was used. This system uses AMV reverse transcriptase for first strand synthesis and the Expand High Fidelity enzyme blend, which consists of *Taq* DNA polymerase and *Pwo* DNA polymerase for the PCR part. The RT-PCR reaction buffer is optimised for all three enzymes so it is not necessary to open the reaction tube between steps. RNase-inhibitor (Boehringer Mannheim) was included in the RT-PCR reaction mixture to minimise the activity of any RNases present.
10 µl of the RNA preparation was used in the RT-PCR reaction mix. Primers were used at a final concentration of 0.1 µM. Thermal cycling conditions were as follows: for the reverse transcriptase step 50°C for 30 mins, then 94°C for 2 mins; for PCR: ten cycles of 94°C (30 s), 52°C (30 s), 68°C (60 s); then 15 cycles of 94°C (30 s), 52°C (30 s), 68°C (80 s); then 15 cycles of 94°C (30 s), 52°C (30 s), 68°C (2 mins), then one cycle of 94°C (30 s), 52°C (30 s), 68°C (10 mins), then 4°C until reaction tubes were removed from the thermal cycler. 10 µl of RT-PCR experiments were visualised by gel electrophoresis.

2.2.11 DNA extraction from soil

Direct DNA extraction from soil was performed according to the method described by Berthelet *et al.* (1996). To a 2 ml screw cap microcentrifuge tube was added 0.5 g soil, 2.5 g 0.1 mm diameter zirconia/silica beads, 0.5 ml 0.1 M sodium phosphate buffer (pH 8.0) and 0.25 ml lysis buffer (100 mM NaCl, 500 mM Tris-HCl pH 8.0, 10% SDS). The tube was first vortexed to mix, then shaken at maximum speed for 5 mins in a Mini Bead Beater (Biotech Products, Bartlesville, Oklahoma, USA). After centrifugation at 12800 ***g*** for 3 mins the supernatant was removed, mixed with 0.4 volume 7.5 M ammonium acetate, and incubated on ice for 10 mins. The tube was then centrifuged at 12800 ***g*** for 3 mins and 50–150 µl of the supernatant was removed and purified by centrifugation through a spin column of polyvinylpolypyrrolidone (PVPP) (Aldrich). The PVPP was previously acid-washed, then equilibrated and slurried in 20 mM potassium phosphate buffer (pH 7.4) before application to the spin column. For PCR experiments 1 µl of the DNA extract was used as a template for a 50 µl reaction.

2.3 Protein expression experiments

2.3.1 [^{35}S]methionine labelling of expressed proteins

Cultures of the *E. coli* JM105 strains carrying pKK223-3 (Pharmacia) constructs containing cloned hypothetical ORFs were grown overnight in 10 ml LB broth amended with ampicillin at 37°C with shaking. 500 µl was used to inoculate another 10 ml LB and grown to A_{600}=0.5 (mid-log phase). 100 µl of the culture was pelleted and resuspended in 500 µl methionine-free MEM amino acid media (Gibco BRL) with 2 mM IPTG and incubated with shaking for 2 hours to induce expression. 5 µCi [^{35}S]methionine (NEN Life Science Products) was added to each tube and incubated at 37°C in a waterbath for 1 hour with occasional mixing. Cultures were then chased for 10 mins with cold L-methionine (Gibco BRL) added to 0.05 µM. Cells were pelleted and resuspended in 30 µl SDS-reducing buffer (1.0 ml 0.5 M Tris-HCl pH 6.8, 0.8 ml glycerol, 1.6 ml 10% SDS, 0.4 ml 2-mercaptoethanol, 0.4% (wt vol^{-1}) bromophenol blue, deionised water to 8.0 ml). Samples were heated at 95°C for 5 mins to denature before loading on to an SDS-PAGE gel.

2.3.2 SDS-PAGE

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed in a Bio-Rad Mini-PROTEAN II Electrophoresis Cell according to the method described by Laemmli (1970). Samples were separated using a discontinuous gel system consisting of a 12% separating gel and a 4% stacking gel. The separating gel was prepared by mixing 3.35 ml deionised water, 2.5 ml 1.5 M Tris-HCl pH 8.8, 100 µl 10% SDS, 4.0 ml acrylamide/bis (30% stock); this solution was degassed for 15 mins, before adding 50 µl fresh 10% ammonium persulfate and 5 µl TEMED and pouring the gel. To prepare the stacking gel a solution of 6.1 ml deionised water, 2.5 ml 0.5 M Tris-HCl pH 6.8, 100 µl 10% SDS, 1.33 ml acrylamide/bis (30% stock) was prepared and degassed for 15 mins; 50 µl fresh 10% ammonium persulfate and 10 µl TEMED was added and the gel poured over the polymerised separating gel. The electrode (running) buffer consisted of 3 g l^{-1} Tris base, 14.4 g l^{-1} glycine, 0.1% SDS, pH 8.3. SDS-PAGE gels were run at 200 volts until the bromophenol blue front reached the bottom of the gel. The gel was then fixed by soaking in a 30% methanol/12% acetic

acid solution for 30 mins, then washed briefly in deionised water and soaked in Entensify (NEN Life Science Products), a universal autoradiography enhancer, for 30 mins in each of the two solutions. The gel was then transferred to Whatman 3MM chromatography filter paper and dried under vacuum at 80°C for 45 mins, then exposed to Kodak X-OMAT AR autoradiography film (Eastman Kodak Co., New York, USA) for an appropriate duration (generally 1 hour–overnight) and the film developed using Kodak reagents.

Bio-Rad Kaleidoscope prestained standards were included in all gels for size estimation of expressed proteins. These standards contained seven proteins with calibrated molecular weights ranging from 7.7 to 244 kilodaltons (kD).

2.4 [9-^{14}C]phenanthrene and [1-^{14}C]naphthalene mineralisation experiments

The following experiments were performed in the laboratory of Dr Carl Cerniglia at the National Centre for Toxicological Research (NCTR), Arkansas, USA during the months of May and June, 1997.

2.4.1 Mineralisation of [9-^{14}C]phenanthrene

300 µl of a phenanthrene-growing culture of RP007 in late-log phase was used to inoculate 30 ml MM in 250 ml biometer flasks with a sidearm containing 5 M NaOH CO_2 trap. The MM was supplemented with 50 µg ml^{-1} unlabelled phenanthrene and 0.5 µCi [9-^{14}C]phenanthrene (Amersham) added in acetone to sterile flasks and the solvent allowed to evaporate before addition of sterile media. The experiment was carried out in triplicate and included an uninoculated control.

Flasks were shaken (125 rpm) at 28°C and 0.5 ml samples of the CO_2 trap were removed at daily intervals, mixed with 15 ml LSC-cocktail (Ultima Gold, Packard, Downers Grove, Illinois, USA) and dpm determined by liquid scintillation counting (Packard Tri-Carb liquid scintillation analyser, model 2000A; Downers Grove, Illinois, USA).

At the completion of the experiment a dilution series of each flask was plated onto solid media to confirm purity of the culture and to estimate cell numbers. The culture

was centrifuged to pellet cells, and the supernatant was adjusted to pH 7.5 with 5 M NaOH and extracted three times with an equal volume (30 ml) ethyl acetate. The extract was dried with sodium sulfate and evaporated to dryness in a rotary evaporator, then resuspended in 3 ml ethyl acetate, evaporated to dryness under vacuum and finally resuspended in 200 µl methanol. This extract was used for high performance liquid chromatography (HPLC) analysis and for radioactivity estimation for mass balance calculations.

The remaining culture supernatant was adjusted to pH 2.5–3.5 with 5 M HCl and extracted with ethyl acetate as described above. In this way both neutral and acidified extracts were obtained for each culture replicate.

The cell pellet was resuspended in 10 ml distilled water and a 1 ml aliquot removed for scintillation counting. Radioactivity present in aliquots of the aqueous fraction of the supernatant after neutral and acid extractions, and in neutral and acidified extracts were measured to give a mass balance.

HPLC was conducted with a Hewlett Packard series 1050 pump system using a gradient of 50–95% methanol in water over 40 mins, followed by 10 mins at 95% methanol, at a flow rate of 1.0 ml min^{-1}. Analysis of acidified extracts used 1% acetic acid in the methanol and water. Detection of metabolites was achieved by a Hewlett Packard diode array detector model 1040A set at 254 nm and a Radiometric A-500 radio-chromatography detector connected in series, using Ultima-flo M LSC-cocktail (Packard, Downers Grove, Illinois, USA). The HPLC column was a SpherisorbODS-2 5-µm column, 46 × 250 mm (MetaChem Technologies, Torrence, California, USA).

2.4.2 High density cell suspension experiments

RP007 was grown in MM supplemented with 20 mM succinate to late-log phase. During the final 12 hours of growth phenanthrene or naphthalene (50 µg ml^{-1}) was added to induce cells. Cells were collected by centrifugation and washed in sterile MM, then centrifuged again and resuspended in 3 ml MM. This concentrated cell suspension (1 ml) was added to 30 ml MM in a 250 ml biometer flask with a sidearm containing the CO_2 trap (10 ml 5 M NaOH). The A_{600} of this suspension was approximately 1.7. Unlabelled phenanthrene or naphthalene was added to the medium to a concentration of 50 µg ml^{-1} with approximately 0.25 µCi [9-

^{14}C]phenanthrene (Amersham) or [1-^{14}C]naphthalene (Sigma). An uninoculated control was included in the experiment, and triplicate flasks were used for both phenanthrene and naphthalene incubations.

All flasks were incubated with shaking (125 rpm) at 28°C for 16 hours. After this time 0.5 ml was removed from the CO_2 trap for scintillation counting. The culture was centrifuged and the supernatant extracted as described above. The culture extracts were subjected to HPLC analysis and a mass balance was calculated for the recovery of the labelled PAH as described for the mineralisation experiments.

2.4.3 Time-course experiment for phenanthrene metabolism

This experiment was set up in an identical way to that described above for the metabolism of phenanthrene by a high density cell suspension of RP007, except that no [9-^{14}C]phenanthrene was added to the culture. Three identical culture flasks were set up, and flasks harvested after 4, 8, and 12 hours incubation time and extracted as described above. The area of the peak in the acidified extract corresponding to 1-hydroxy-2-naphthoic acid was measured and used as an estimate of the quantity of this metabolite accumulated in the culture.

2.5 Naphthalene transformation by *E. coli* clones

The ability of *E. coli* strains carrying recombinant *phn* genes to transform naphthalene was investigated by exposing a high density suspension of cells to naphthalene. *E. coli* strains were grown to late log phase in 100 ml M9 minimal media (Sambrook *et al.*, 1989) amended with 100 μg ml^{-1} ampicillin, and then harvested by centrifugation and washed twice in MM. Cells were resuspended in 20 ml fresh M9 media in a 50 ml stoppered conical flask and incubated with shaking at 28°C for 1 hour. Naphthalene was then added to the suspension to a concentration of 1 mg ml^{-1} from a stock solution in dimethylformamide, and the flasks were incubated at 28°C for 16 hours with shaking (200 rpm). The cell suspensions were harvested by centrifugation, and the supernatants removed and acidified to pH 3.0 with 1 M HCl and extracted three times with an equal volume (20 ml) ethyl acetate. Extracts were dried with sodium sulfate and then evaporated to dryness under vacuum at 37°C. The dried extract was resuspended in 2 ml methanol and for thin layer chromatography

(TLC) this was evaporated under methane to 100 µl.

TLC analysis used two solvent systems as described by Menn *et al.* (1993). Chloroform:methanol (98:2 vol vol^{-1}) was used to separate undegraded substrates and other less polar compounds from more polar metabolites; and a benzene:acetate: water (125:74:1 vol vol^{-1} vol^{-1}) system was used to resolve highly polar metabolites from substrates that remained at or near the origin of the TLC plates using the cloroform:methanol system. The TLC plates used were Merck Silica Gel 60 impregnated with a fluorescence indicator; plates were visualised under 302 nm UV.

2.6 Enzyme assays

2.6.1 Determination of naphthalene dioxygenase

Naphthalene dioxygenase was assayed according to the method of Shamsuzzaman & Barnsley (1974b). Cultures of RP007 grown on MM supplemented with acetate (20 mM), naphthalene (crystals) or phenanthrene (0.05% wt vol^{-1}) were harvested by centrifugation, washed, and resuspended in 0.1 M sodium phosphate buffer (pH 7.0) to A_{600}=0.15–0.2. Naphthalene in ethanol (10 µl, 10 mM) was added to 1 ml of suspension at 25°C, and the decrease in absorption due to naphthalene (A_{276}) was followed. The reference cuvette contained bacterial suspension (A_{600}=0.2). For the calculation of reaction rates the extinction coefficient of naphthalene at 276 nm of 4.51 mM^{-1} cm^{-1}, as determined experimentally by Shamsuzzaman & Barnsley (1974b), was used. BCA Protein Assay Reagent (Pierce) using bovine serum albumin as a standard was used to estimate protein concentrations in the cell suspensions.

2.6.2 Determination of catechol 2,3-dioxygenase

Catechol 2,3-dioxygenase activity was assayed according to the method of Nozaki (1970). Wild-type RP007 was grown in 1 L MM supplemented with acetate (10 mM), benzoate (5 mM), or phenanthrene (0.05% wt vol^{-1}) for up to 5 days to A_{600}=2–3, and dioxygenase clones in *E. coli* DH5α were grown overnight in 100 ml LB amended with ampicillin. Cells were harvested by centrifugation at 4°C (5000 ***g***, 10 mins) and washed in cold fresh MM, then resuspended in 0.1 M sodium phosphate buffer (pH 7.5) containing 10% (vol vol^{-1}) acetone to a concentration of

about 0.5 g ml^{-1} (wet wt). The cell suspension was disrupted by sonication (Ultrasonic Processor XL, Heat Systems-Ultrasonics Inc., New York, USA) for three 30 second periods, interspersed with cooling on ice for 1 min. The disrupted cell suspension was centrifuged at 4°C (30000 ***g***, 30 mins) to pellet cell debris and the supernatant carefully decanted and kept on ice.

For the enzyme assay 100 µl of the cell extract was added to 900 µl of the assay substrate (10 mM in 0.1 M sodium phosphate buffer (pH 7.5) containing 10% acetone) in a 1 ml quartz cuvette. Extracts with high activities were diluted 1:10 for the assay. Extradiol dioxygenase activities were calculated for catechol by monitoring the increase of absorption at 375 nm (A_{max} for the *meta* cleavage product of catechol), and reaction rates were calculated using a molar extinction coefficient of 36000 $M^{-1}cm^{-1}$ (Nozaki, 1970); for 3-methylcatechol and 4-methylcatechol increase in absorbance and molar extinction coefficients were (388 nm and 15000 $M^{-1}cm^{-1}$) and (382 nm and 31500 $M^{-1}cm^{-1}$), respectively. Extradiol cleavage of 2,3-dihydroxybiphenyl was assayed with an initial substrate concentration of 0.1 mM 2,3-dihydroxybiphenyl, the increase of absorption at 434 nm was measured, and reaction rates were calculated using a molar extinction coefficient of 22000 $M^{-1}cm^{-1}$ (Taira *et al.*, 1988). All substrates were supplied by BDH (BDH Laboratory Supplies, Poole, Dorset, England) except 2,3-dihydroxybiphenyl which was supplied by Wako Chemicals (Wako Chemicals GmbH, Neuss, Germany). Substrates discoloured due to oxidation were purified by sublimation under vacuum at 60°C.

The protein concentration in the cell extracts was estimated using the Bio-Rad Protein Assay, which is based on the Bradford assay. The method was used according to the manufacturers instructions, and bovine serum albumin (BSA) was used to prepare a standard curve.

2.6.3 Assay of recombinant PAH extradiol dioxygenase

The PAH extradiol dioxygenase gene (*phnC*) was cloned into the vector pKK223-3 (Pharmacia) and expressed in *E. coli* JM105. Following growth of 100 ml cultures in LB with ampicillin to A_{600}=2 (overnight), IPTG was added to 500 µg ml^{-1} and the incubation continued for two hours to induce expression of *phnC*. *E. coli* cells were pelleted and washed in MM, then resuspended in 4 ml 0.1 M sodium phosphate buffer (pH 7.5) containing 10% acetone. Cells were disrupted by sonication and

cellular debris was removed by centrifugation to give a crude cell extract as described above. Activity against 1,2-dihydroxynaphthalene (1,2-DHN) (Tokyo Chemical Industry (TCI), Tokyo, Japan) was determined spectrophotometrically according to the method of Kuhm *et al.* (1991b). 100 µl cell extract was added to 900 µl 50 mM acetic acid-NaOH buffer (pH 5.5) and the reaction was started by the addition of 0.5 µmol 1,2-DHN in 10 µl dimethylformamide. The initial rate of decrease in absorbance at 331 nm was measured. 331 nm is an isobestic point for the oxidation of 1,2-DHN to 1,2-naphthoquinone which occurs rapidly in aqueous solution, but is minimised at pH 5.5 (Kuhm *et al.*, 1991b; Shamsuzzaman & Barnsley, 1974a). The molar reaction coefficient (ε) for 1,2-DHN at 331 nm is 2.6 $mM^{-1}cm^{-1}$ as calculated by Kuhm *et al.* (1991b), this value was used in the enzyme activity calculations. The PhnC extradiol dioxygenase was also assayed using catechol, 3-methylcatechol, 4-methylcatechol, and 2,3-dihydroxybiphenyl as described above for the catechol 2,3-dioxygenase assay. When using 3,4-dihydroxybiphenyl as a substrate, scans over the absorbance range 300–450 nm were used to follow any appearance of *meta* cleavage product. The anticipated *meta* cleavage product from 3,4-dihydroxybiphenyl has an absorbance maximum at 380 nm with a shoulder at 324 nm (pH 7) (Lloyd-Jones *et al.*, 1995).

2.7 Analysis of DNA sequence data

Sequence data was routinely managed using OMIGA (Version 1.0) (Oxford Molecular Group, Oxford, England). This programme translated DNA sequences to amino acid sequences, and searched for restriction sites and features such as inverse repeats, hairpin-loop structures and transmembrane regions. Comparison of nucleotide and amino acid sequences used the ClustalW programme (Thompson *et al.*, 1994) using default parameters, and 'msf' files generated by this programme were formatted using GeneDoc (Version 1.1.004). Identification of open reading frames (ORFs) within sequenced regions used an internet search tool provided by the Virtual Genome Centre located at http://alces.med.umn.edu/cuse.html. Searching of the GenBank and Swissprot databases using query nucleotide or amino acid sequences used the BLAST (Basic Local Alignment Search Tool) search engine provided by the National Centre for Biotechnology Information (NCBI) and located

at http://www.ncbi.nlm.nih.gov/BLAST/. For estimation of molecular weight of peptides predicted by hypothetical ORFs the Peptide Mass programme provided by the ExPASy World Wide Web Molecular Biology Server (Geneva) was used; this site is located at http://expasy.hcuge.ch/sprot/peptide-mass.html.

Dendrograms were constructed using the PHYLIP software package (Version 3.57c) (Felsenstein, 1989) using SEQBOOT, PROTDIST (or DNADIST), NEIGHBOR, and CONSENSE programmes. Bootstrap analyses used 100 data sets and distance measurements for protein sequences were computed using the Dayhoff PAM matrix, and for DNA sequences using the Jukes and Cantor matrix. From this data the 'neighbor joining method' was used to produce an unrooted tree for each data set, and the 'majority-rule consensus tree method' was used to find the strict consensus tree and bootstrap values for the major branch points. The final consensus tree was rooted through an outgroup which was specified after the analyses. The branch lengths were taken from the amino acid similarity values obtained from the ClustalW analysis, and each tree was redrawn to scale.

CHAPTER 3

Identification and growth of *Burkholderia* sp. strain RP007

3.1 Isolation of *Burkholderia* sp. strain RP007

RP007, the wild-type phenanthrene-degrading bacterium used in this study, was isolated in 1994 by M. Bruce from sludge in a contaminated pond at Rotowaro, near Huntly, New Zealand (Bruce, 1995). This pond contained sludge and waste water contaminated with PAHs, phenols, and aromatic heterocycles, which was effluent from a coal carbonisation plant located at this site. Phenolic waste from the plant, which was destroyed by fire in 1985, was shifted from holding ponds to an enclosed aerated treatment lagoon in an effort to encourage bioremediation of the waste (Aislabie *et al.*, 1993). In 1995, three years after treatment began, levels of contaminants in the sludge and water of the treatment pond were deemed to have reached acceptable levels and the sludge was disposed of by mixing with clay and spreading over 1 hectare of land on site. The pond sediment from which RP007 was isolated would therefore be expected to be enriched for bacteria capable of degrading aromatic compounds.

RP007 was isolated by a direct isolation technique for its ability to utilise phenanthrene as a sole carbon and energy source for growth. A dilution of the sludge was plated onto MM agar with a phenanthrene overlay, as described in Chapter 2, and RP007 was selected for its ability to form zones of clearing indicating metabolism of phenanthrene.

3.2 Identification of *Burkholderia* sp. strain RP007

3.2.1 Morphological, physiological and biochemical characteristics

RP007 is a Gram-negative rod-shaped bacterium, which forms light brown translucent colonies on ½PC agar. The results of API 20NE analysis indicate RP007 reduces nitrates to nitrites, produces catalase, and does not produce oxidase, urease, arginine dihydrolase, gelatinase, β-glactosidase, ferment glucose or produce indole from tryptophan. Glucose, arabinose, mannitol, N-acetyl-glucosamine and gluconate were assimilated by RP007. RP007 readily utilises acetate, succinate, benzoate, naphthalene, and phenanthrene when supplied as sole carbon sources.

3.2.2 Cellular fatty acid analysis

Fatty acid analysis (MIDI Microbial identification system (MIDI, Newark, Delaware, USA) showed the predominant long-chain fatty acids were $C_{16:0}$ saturated acid, $C_{17:0}$ cyclopropane, and $C_{19:0\ w8c}$ cyclopropane, and that 3-hydroxy $C_{16:0}$ was the only 3-hydroxy fatty acid present. This cellular fatty acid profile is very similar to those of the seven strains originally transferred to the *Burkholderia* genus by Yabuuchi *et al.* (1992); in particular the percentage compositions of $C_{16:0}$ (22%), $C_{17:0}$ (21%), and $C_{19:0\ w8c}$ (16%) are almost identical to that for *B. cepacia* EY645 (ATCC 25416), the type strain for the *Burkholderia* genus. The presence of 3-hydroxy $C_{16:0}$ is characteristic of members of the *Burkholderia* genus (Gillis *et al.*, 1995).

3.2.3 16S rRNA sequencing

Nucleotide sequencing of the 16S subunit of ribosomal DNA was used to further identify RP007. Using the eubacterial primers for the 16S gene (27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-TACGGGTACCTTGTTACGACTT-3')) (Lane, 1991) a 1.4 kb PCR product was amplified from RP007 genomic DNA. Numerous attempts to sequence this product failed to produce quality sequence, indicating multiple templates present in the PCR product. Using the same primers PCR amplification directly from a single colony failed to improve results, ruling out any possibility of contamination of the genomic DNA as the cause of the poor quality sequence data. Finally, the 16S product was amplified from RP007 using RT-PCR with the same primer set. To generate enough product for sequencing 1 µl of the RT-PCR product was used as a template for PCR, and using this product quality sequence data was generated using the 27F and 1492R primers for the sequencing reaction. An internal primer RP00716S (5'-CGCGTTGCATCGAATTAATC-3') was designed for sequencing of the internal region of the PCR product.

The presence of multiple variants of the 16S rRNA gene has been reported for several *Burkholderia* strains (Lessie *et al.*, 1996) which may explain the failure to obtain sequence data from the 16S PCR product. Apparently only one 16S gene is expressed which allowed sequencing of the RT-PCR product.

The nucleotide sequence of 1347 bp of the sequenced gene is shown in Figure 3.1. A BLAST search of this sequence revealed most similarity to the 16S rRNA genes of *Burkholderia* sp. LB400, a well characterised biphenyl degrading bacterium, and to

```
   1 cctggtggcg agtggcgaac gggtgagtaa tacatcggaa cgtgtcctgg
  51 agtgggggat agcccgggga aagccggatt aataccgcat acgctctacg
 101 gaggaaaggg ggggatctta ggacctttcg ctcaaggggc ggccgatggc
 151 agattagcta gttggtgggg taaaggccta ccaaggcgac gatctgtagc
 201 tggtctgaga ggacgaccag ccacactggg actgagacac ggcccagact
 251 cctacggaag gcagcagtgg ggaattttgg acaatgggcg caagcctgat
 301 ccagcaatgc cgcgtgtgtg aagaaggcct tcgggttgta aagcactttt
 351 gtccggaaag aaaacctcag ccctaatacg gcgggggat gacggtaccg
 401 gaagaataag caccggctaa ctacgtgcca gcagccgcgg taatacgtag
 451 ggtgcgagcg ttaatcggaa ttactgggcg taaagcgtgc gcaggcggtt
 501 cgttaagaca gatgtgaaat ccccgggctt aacctgggaa ctgcatttgt
 551 gactggcggg ctagagtatg gcagagcggg gtagaattcc acgtgtagca
 601 gtgaaatgcg tagagatgtg gaggaatacc gatggcgaag gcagcccct
 651 gggccaatac tgacgctcat gcacgaaagc gtggggagca aacaggatta
 701 gataccctgg tagtccacgc cctaaacgat gtcaactagt tgtcgggtct
 751 tcattgactt ggtaacgtag ctaacgcgtg aagttgaccg cctggggagt
 801 acggtcgcaa gattataact caaaggaatt gacggggacc cgcacaagcg
 851 gtggatgatg tggattaatt cgatgcaacg cgaaaaacct tacctaccct
 901 tgacatggac ggagcctggt tgaaaggtcg gggtgctcga aagagaaccg
 951 tcacacaggt gctgcatggc tgtcgtcagc tcgtgtcgtg agatgttggg
1001 ttaagtcccg caacgagcgc aacccttgtc cctagttgct acgcaagagc
1051 actctaggga gactgccggt gacaaaccgg aggaaggtgg ggatgacgtc
1101 aagtcctcat ggcccttatg ggtagggctt cacacgtcat acaatggtcg
1151 gaacagaggg ccgccaaccc gtgaggggga gccaatccca gaaaaccgat
1201 cgtagtccgg atcgcactct gcaactcgag tgcgtgaagc tggaatcgct
1251 agtaatcgcg gatcagcatg ccgcggtgaa tacgttcccg ggtcttgtac
1301 acaccgcccg tcacaccatg ggagtgggtt ttaccagaag tggctag
```

Figure 3.1 cds 1–1347 partial sequence of 16S rRNA gene of *Burkholderia* sp. RP007.

three phenanthrene degrading *Burkholderia* strains (N2P6, N2P5, N3P2) isolated from geographically diverse soils (Mueller *et al.*, 1997). The 16S rRNA sequence of RP007, together with that from LB400, N2P6, N2P5, and N3P2, was submitted to the Ribosomal Database Project (Maidek *et al.*, 1997) and a phylogenetic tree was produced showing these strains relative to representative *Burkholderia*, *Sphingomonas*, and *Pseudomonas* strains (Figure 3.2). This tree shows the PAH degrading *Burkholderia* strains (RP007, LB400, N2P6, N2P5, and N3P2) form a tight cluster and occupy a separate branch to the other *Burkholderia* strains. According to this tree, the RP007/LB400 group is most closely related to *B. caryophylli*.

3.2.4 Assignment of RP007 to *Burkholderia* genus

On the basis of the API20NE results, the cellular fatty acid profile, and the 16S rRNA nucleotide sequence, strain RP007 was assigned to the *Burkholderia* genus. Further definition to the species level was not attempted based on these results since no clear lineage seemed apparent. The 16S rRNA nucleotide sequence data suggests

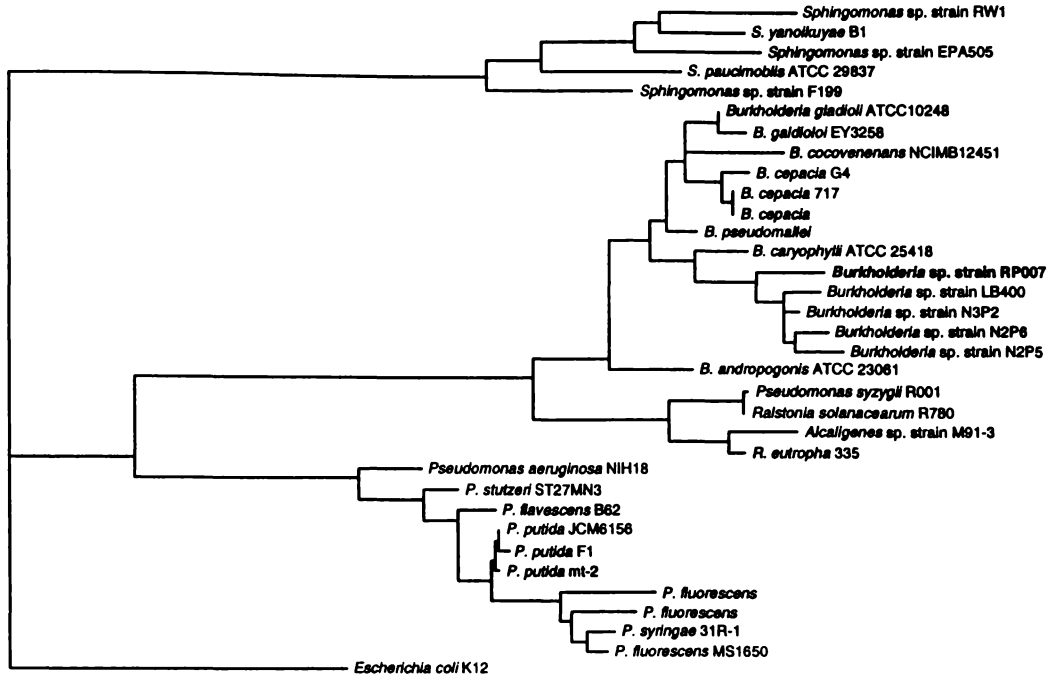


Figure 3.2 Phylogenetic tree based on 16S sequence generated by the Ribosomal Database Project (Maidak *et al.*, 1997) showing the position of RP007 and other PAH degrading *Burkholderia* strains (LB400, N3P2, N2P6, and N2P5) relative to representative *Burkholderia*, *Sphingomonas*, *Ralstonia*, and *Pseudomonas* species. *P. putida* F1, *P. putida* mt-2, *Sphingomonas* sp. RW1, *S. yanoikuyae* B1 and *Sphingomonas* sp. F199 are aromatic degrading strains.

RP007 is most similar to *B. caryophylli*, yet the fatty acid profile is more like that of the *B. cepacia* type strain. It is possible that the PAH degrading *Burkholderia* strains constitute a new species since the phylogenetic groupings based on 16S rRNA data show a separate cluster. Rigorous auxanographic studies and DNA:DNA hybridisation experiments would be necessary to further define the taxonomic classification of RP007, and this was not considered within the scope of this project.

3.3 RP007 growth at the expense of aromatic compounds

3.3.1 Phenanthrene

RP007 utilises phenanthrene as sole source of carbon and energy for growth on both solid and liquid media, without any requirement for supplementary vitamins. Zones of clearing on phenanthrene overlay plates are visible after 2–3 days, and evidence of colony growth is observable after 4–5 days.

Both non-induced (grown on ½PC media) and induced (grown on phenanthrene) RP007 cells inoculated into a MM broth containing phenanthrene crystals result in brownish discolouring of the media after 24 hours, indicative of phenanthrene transformation. With continued incubation an increase in cell density is observed, and the medium becomes increasingly discoloured due to metabolite formation. Due to this discoloration absorbance at 600 nm did not give an accurate measure of culture growth; estimation of protein concentration was therefore used to plot a phenanthrene growth curve for RP007 (Figure 3.3).

3.3.2 Naphthalene

RP007 also utilises naphthalene as sole source of carbon and energy for growth without any requirement for supplementary vitamins (Figure 3.4). Growth of RP007 on naphthalene is more difficult to initiate than on phenanthrene, but once growth has begun naphthalene serves as an excellent growth substrate. On solid MM with naphthalene supplied as crystals in the lid of the petri dish, small colonies are visible within 7–10 days. Growth of RP007 on naphthalene in liquid minimal medium requires either: (a) first supplying naphthalene as a vapour in the headspace of the flask until growth is observed (3–4 days), followed by addition of naphthalene crystals to the medium, after which time growth proceeds rapidly; or (b) inoculation of a large volume (1 ml) of mid–late log phase naphthalene-growing cells to fresh medium containing naphthalene crystals, rapid growth begins immediately. It appears the toxic effect of naphthalene prevents growth of small inocula of RP007 cells unless it is present in the growth medium at low concentrations; larger inocula of naphthalene-adapted cells are able to assimilate naphthalene without any toxic effect.

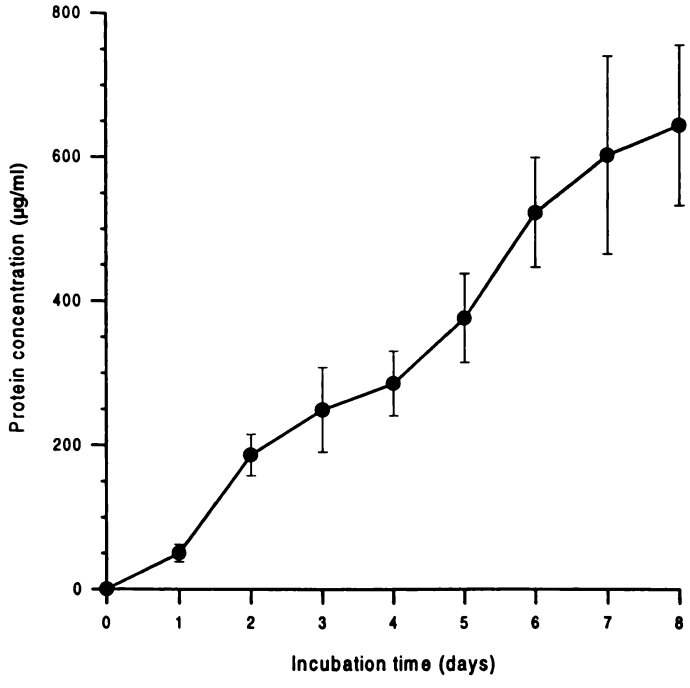


Figure 3.3 Growth of *Burkholderia* sp. RP007 using phenanthrene as sole source for carbon and energy. Data points are mean values for protein concentrations from three separate 100 ml cultures.

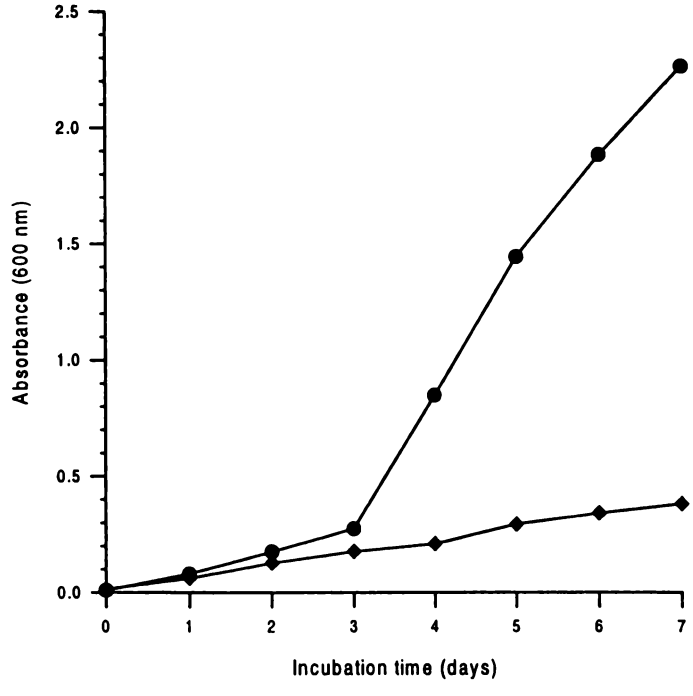


Figure 3.4 Representative plot showing growth of *Burkholderia* sp. RP007 at the expense of naphthalene. Closed circles (•) indicate growth on naphthalene supplied as crystals in the medium; diamonds (♦) indicate growth on naphthalene supplied as a vapour; both cultures were 100 ml volumes.

3.4 [^{14}C] mineralisation experiments

3.4.1 Transformation of [9-^{14}C]phenanthrene and [1-^{14}C] naphthalene by a high density cell suspension

A short incubation (16 hours) of a high density suspension of phenanthrene- or naphthalene-induced RP007 cells with [9-^{14}C]phenanthrene or [1-^{14}C]naphthalene, respectively, was used to confirm mineralisation of these substrates by RP007, and to identify accumulated metabolites.

During the period of this experiment RP007 mineralised approximately 30% of the added [9-^{14}C]phenanthrene or [1-^{14}C]naphthalene to $^{14}CO_2$. The mass balance (Table 3.1) shows significant portions of the added label for both substrates was associated with the cells or present in the acidified extract. The percentage recovery from the phenanthrene incubation (76.6%) was acceptable considering the volatility of this compound (C. E. Cerniglia, pers. comm.) and was comparable to the control flask (76.7%). The recovery from the naphthalene incubation was low (41.2%) but this was also considered acceptable due to even greater volatility of naphthalene. The presence of cells in the culture flasks appeared to prevent much volatilisation of the naphthalene relative to the control flask from which all [1-^{14}C]naphthalene disappeared during the incubation.

HPLC analysis of the extracts from the phenanthrene incubation revealed that virtually all of the label present in both the neutral and acidic extracts was accounted for by a peak with a retention time and absorbance spectra identical to that of the authentic 1-hydroxy-2-naphthoic acid standard. A peak identified as phenanthrene on the basis of retention time and its spectral profile accounted for a small quantity of the label present in the neutral extract.

Nominal label was present in the neutral extract from the naphthalene incubation, and of the ca. 5.9% present in the acidic extract all of it was attributable to a peak with an identical retention time and absorption spectra to the authentic salicylic acid standard.

Fraction	Phenanthrene	Phenanthrene control	Naphthalene	Naphthalene control
% CO_2	28.7±0.4	0	29.0±5.5	0.09
% Cells	12.7±0.5	0	5.1±0.9	0
% Aq phase	1.3±0.2	0.17	0.8±0.05	0.16
% Neutral extract	23.7±1.6	76.42	1.4±0.3	0.42
% Acid extract	10.3±0.9	0.11	5.9±1.1	0.13
% Accounted for	76.6±2.1	76.7	41.2±5.5	0.79

Table 3.1 Mass balance for [9-^{14}C]phenanthrene and [1-^{14}C]naphthalene high density cell suspension experiment for RP007. All values represent the percentage of dpm initially added as [9-^{14}C]phenanthrene or [1-^{14}C]naphthalene.

3.4.2 Phenanthrene time-course experiment

Since the 16 hour incubation of a high density cell suspension of RP007 with phenanthrene showed the presence of only 1-hydroxy-2-naphthoic acid as a metabolite, identical cultures (without the addition of [9-^{14}C]phenanthrene) were set up and harvested after 4, 8, and 12 hour incubations. As before, the only compounds detected by HPLC analysis in the neutral and acidified extracts were 1-hydroxy-2-naphthoic acid and phenanthrene. The area of the 1-hydroxy-2-naphthoic acid peak in the acidified extracts of each flask was seen to increase with longer incubation times (Figure 3.5).

This indicates RP007 rapidly metabolises phenanthrene to 1-hydroxy-2-naphthoic acid, with no intermediate metabolites accumulating. The metabolism of 1-hydroxy-2-naphthoic acid is apparently rate-limiting and hence this compound is seen to accumulate in the medium.

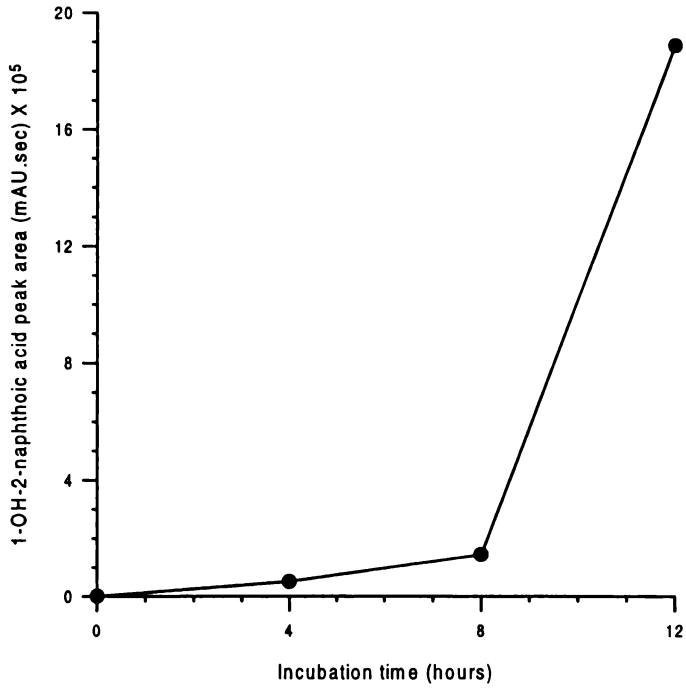


Figure 3.5 Accumulation of 1-hydroxy-2-naphthoic acid during growth of *Burkholderia* sp. RP007 at the expense of phenanthrene.

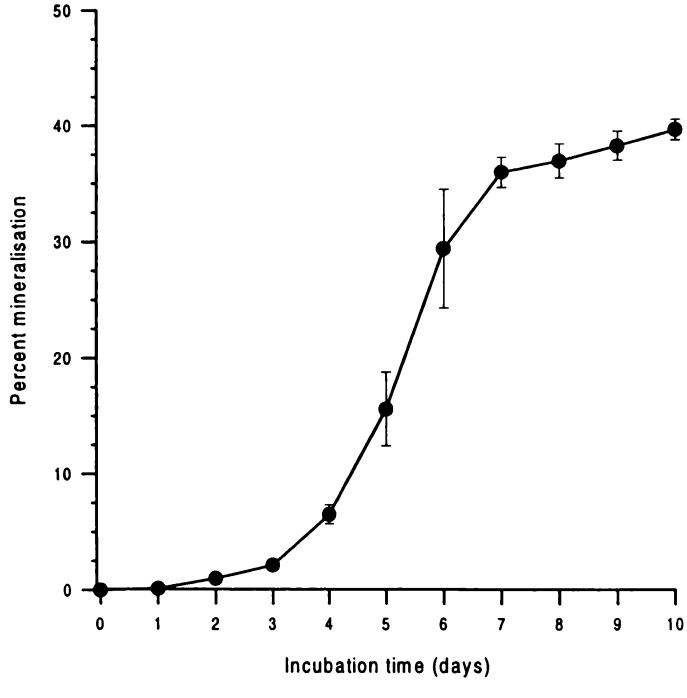


Figure 3.6 Mineralisation of [9-^{14}C]phenanthrene by *Burkholderia* sp. RP007. Y-axis values represent cumulative evolution of $^{14}CO_2$ as a percentage of dpm initially added as [9-^{14}C]phenanthrene.

Fraction	RP007	Control
% CO_2	39.7±0.9	0
% Cells	5.7±0.2	0
% Aq phase	2.2±0.3	0.2
% Neutral extract	6.0±0.3	38.8
% Acid extract	7.8±1.2	0.4
% Accounted for	61.5±2.5	39.4

Table 3.2 Mass balance for [9-^{14}C]phenanthrene mineralisation experiment by *Burkholderia* sp. RP007. All values represent the percentage of dpm initially added as [9-^{14}C]phenanthrene.

3.4.3 Mineralisation of [9-^{14}C]phenanthrene

Using phenanthrene as a sole source of carbon and energy for growth of RP007, the evolution of $^{14}CO_2$ was monitored. After an initial lag phase of 2 days RP007 rapidly mineralised [9-^{14}C]phenanthrene until approximately 40% of added label had been evolved as $^{14}CO_2$. Figure 3.6 shows the mineralisation curve from this experiment and Table 3.2 shows the mass balance data.

HPLC analysis of the neutral and acidic extracts of the culture supernatant at the conclusion of the experiment revealed the presence of 1-hydroxy-2-naphthoic acid and phenanthrene only. 1-hydroxy-2-naphthoic acid accounted for all of the label present in these extracts. In both the neutral and acidic extracts 1-hydroxy-2-naphthoic acid appeared as two peaks, both with a spectral profile identical to that of the authentic standard. In the neutral extract 1-hydroxy-2-naphthoic acid appeared as two peaks at RT 4.0 and 11.8 min, and in the acidic extract it appeared as peaks at RT 10.6 and 18.9 min. The RT for the 1-hydroxy-2-naphthoic acid standard under neutral and acidic conditions was 10.2 min and 18.6 min, respectively. It is possible that the compound appearing at RT 4.0 min and 10.6 min in the neutral and acidic extracts, respectively, is an acetylated derivative of 1-hydroxy-2-naphthoic acid (C. E. Cerniglia, pers. comm.). This derivatisation reaction could have occurred during the extraction with ethyl acetate or during the HPLC process, since acetic acid was present in the water and methanol reservoirs. However, as no such derivatisation was observed for the authentic standard, it seems more likely this compound was formed during the ethyl acetate extraction. This would explain the identical spectra yet different retention times observed for these compounds.

The 2 day lag phase evident in the RP007 [9-^{14}C]phenanthrene mineralisation curve

may be attributable to the position of the ^{14}C label; the difficulties associated with mineralisation studies using non-universally labelled compounds has been discussed by Shuttleworth & Cerniglia (1995). Since the ^{14}C label is present on the middle ring of phenanthrene no $^{14}CO_2$ would be evolved until this ring had been cleaved, which first requires that an outside ring be cleaved since bacterial attack on phenanthrene occurs at the 1,2 position. The time course experiment showed that RP007 rapidly accumulates 1-hydroxy-2-naphthoic acid, but degradation of phenanthrene to this point would not be indicated by evolution of $^{14}CO_2$. Thus, the initial lag phase seen in the mineralisation experiment may not indicate a lag in growth by RP007 on ^{14}C-phenanthrene, but rather during this time the strain was preferentially metabolising phenanthrene to 1,2-dihydroxynaphthalene. Indeed, the growth curve of RP007 on phenanthrene as measured by protein concentration shows no such lag phase (Figure 3.3).

CHAPTER 4

Cloning and characterisation of the *phn* operon

4.1 Probing RP007 with *nah*-like genes

Initial Southern hybridisation and PCR experiments revealed that the naphthalene- and phenanthrene-catabolic genes carried by *Burkholderia* sp. strain RP007 showed no homology to probes based on the archetypal *nah* genes under the stringency conditions used. A 5 kb probe made from a fragment of pDI1 containing the *pahAaAbAcAd* genes of *P. putida* OUS82 (Takizawa *et al.*, 1994) did not hybridise to plasmid or total genomic DNA from RP007 (not shown). Primers based on the *pahAc* gene of *P. putida* OUS82 (PAHTOP (5'-CCCGGCGACTATGTTACC-3') and PAHBOT (5'-GTTTTGCGATGCTGTTTC-3')) did not amplify a product from RP007 genomic DNA even using a low (45°C) annealing temperature.

These results suggested that the genes involved in naphthalene and phenanthrene degradation by RP007 were divergent to the well-characterised *nah*-like genes. This assumption provided the motivation to clone and characterise the naphthalene- and phenanthrene-catabolic genes from RP007.

4.2 Cloning of the pB1 fragment

In order to clone the genes responsible for the initial attack on naphthalene and phenanthrene by RP007, the ability of the aromatic initial dioxygenase to transform indole to indigo was exploited. This reaction, as described in Chapter 1, occurs due to the ability of the aromatic dioxygenase to catalyse the dihydroxylation of indole to indole dihydrodiol, which forms the blue product indigo (Ensley *et al.*, 1983). The wild-type RP007 is indole-positive and the strategy was to clone this phenotype into *E. coli*.

The cloning strategy involved the digestion of RP007 DNA (derived from a plasmid- or total genomic-DNA preparation) with a suitable restriction enzyme, then ligation of this DNA to pUC18 vector DNA digested with the same restriction enzyme, and treated with bacterial alkaline phosphatase to prevent reannealing of the vector ends. The ligation reaction was used to transform competent *E. coli* DH5α cells which were plated onto LB medium containing indole. This technique, referred to as 'shotgun cloning', yields a plasmid or genomic 'library' in pUC18. *E. coli* cells containing a pUC18 construct which includes an RP007 fragment encoding an aromatic initial dioxygenase will appear as blue colonies due to the production of

indigo in these cells. This technique has been successfully used by other workers to clone genes for aromatic catabolism; the *ntd* (Parales *et al.*, 1996), *nah* (Yang *et al.*, 1994), and *car* (Sato *et al.*, 1997a; b) genes were cloned using a similar strategy.

An indole-positive clone, pB1, was serendipitously discovered on transformation plates which had been stored on the bench for three weeks. The plates had been checked after overnight growth of the colonies at 37°C and again a week later, and no blue colonies were observed. Since transformation plates were usually discarded after one week, or sprayed with catechol to screen for clones carrying a catechol 2,3-dioxygenase gene, the possibility that it would take so long for the blue colour to develop was not considered. The pDI1 plasmid carrying the *pah* genes of *P. putida* OUS82 (Takizawa *et al.*, 1994) supplied by N. Takizawa yielded intense blue colonies after overnight growth on LB containing indole, and it was assumed indole-positive clones from RP007 would behave similarly. Indeed, after cultivation of the pB1 clone in pure culture the blue pigment developed after 2–3 days colony growth, but was not as intense as that for pDI1.

pB1 is an indole-positive clone obtained from a *Hin*dIII shotgun cloning experiment using a total genomic DNA preparation from *Burkholderia* sp. RP007. Restriction mapping showed that pB1 contained an 11.5 kb *Hin*dIII fragment inserted into the *Hin*dIII site of pUC18 (Figure 4.1A). To localise the region of the pB1 fragment coding for indigo production a number of subclones were derived from pB1, as shown in Figure 4.1B. This showed that the indole-positive region was located near the 3' end of the pB1 fragment adjacent to the M13pUC reverse sequencing primer of pUC18.

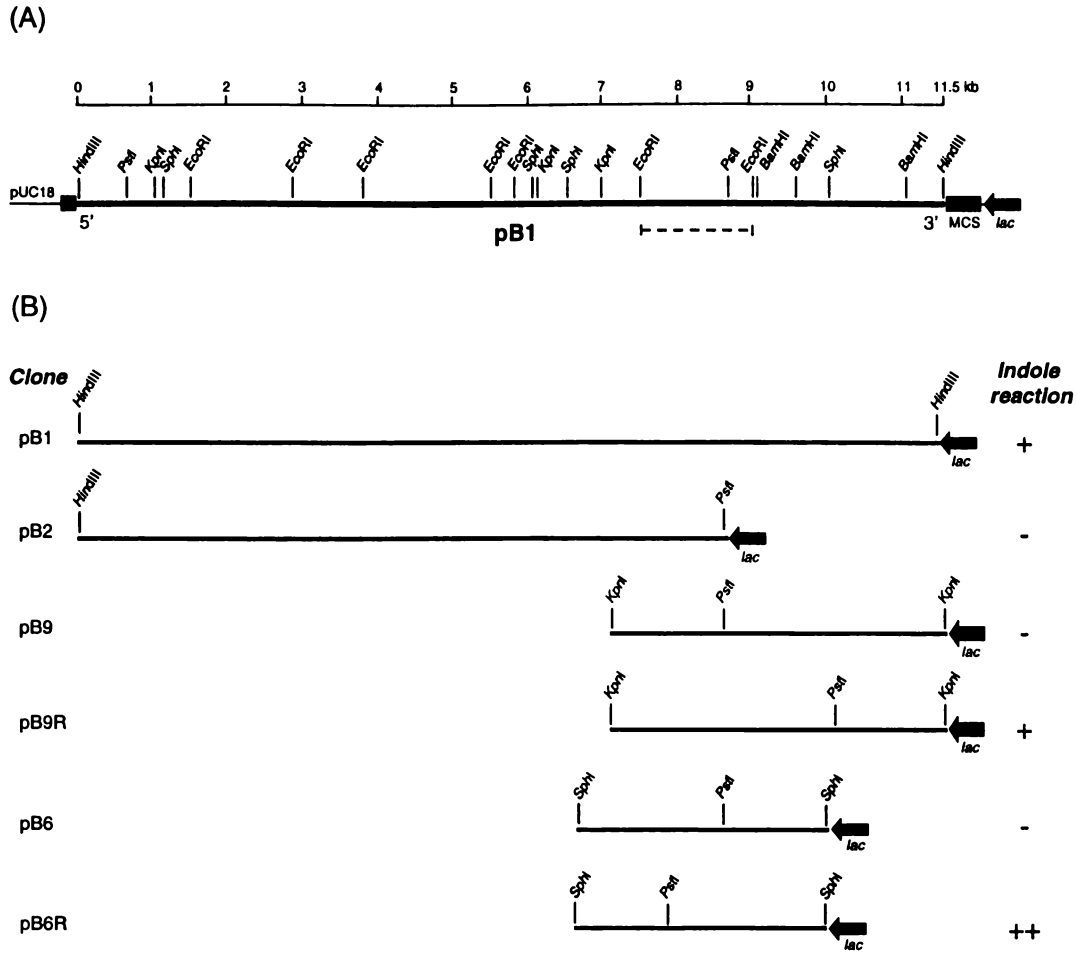


Figure 4.1 (A) The 11.5 kb pB1 insert shown relative to the pUC18 multiple cloning site (MCS) and *lac* promoter; the sites of common restriction enzymes are shown. The ends of the clone are labelled 5' and 3' for clarity. The dashed line represents the region used for the pB1 probe described in section 4.3. (B) Indole activity of subclones derived from pB1 indicating the position and orientation of dioxygenase genes which confer the indole-positive phenotype. The pB9 and pB6 fragments were cloned into pUC18 in both directions, as shown, and revealed dioxygenase genes were localised on the pB6 3.5 kb *Sph*I fragment and were transcribed in a 5'→3' direction with respect to the pB1 restriction map (4.1A). Details of these subclones are given in Table B.2 (Appendix B).

4.3 Southern hybridisation analysis

To confirm that the pB1 fragment was derived from RP007 DNA, and to attempt to localise these genes to either the plasmid or chromosomal DNA fraction of RP007, a Southern blot was probed with an internal 1.5 kb *Eco*RI fragment of pB1 known to include the genes for the aromatic initial dioxygenase (shown as a dashed line in Figure 4.1A). This experiment showed a strong hybridisation signal corresponding to a 1.5 kb *Eco*RI band of RP007 plasmid DNA, and no hybridisation to RP007 total genomic DNA, or to either genomic DNA from *P. putida* G7 (which carries the NAH7 plasmid) or the pDI1 plasmid carrying the *pah* genes (Takizawa *et al.*, 1994) (Figure 4.2). Detection of pB1 target DNA in a plasmid prep of RP007 indicates the pB1 fragment is plasmid-borne. Figure 4.2 also confirms that the initial dioxygenase genes of pB1 show no homology to the *nah*-type genes of *P. putida* G7 and *P. putida* OUS82.

The pB1 clone was derived from a cloning experiment using RP007 total genomic DNA which consists of both chromosomal and plasmid DNA. Failure to also detect target DNA in RP007 total genomic DNA in this experiment is probably due to the loading of insufficient DNA to these lanes, since previous experiments have shown hybridisation of a pB1 probe to RP007 total genomic DNA (not shown). The low relative representation of plasmid genes in a total DNA prep compared with their representation in a plasmid DNA prep makes it more difficult to detect these genes in total DNA.

Figure 4.3 shows a similar Southern blot probed with the *nahAc* gene of *P. putida* G7. This probe did not hybridise to RP007 plasmid or total genomic DNA and, as expected, did hybridise to *P. putida* G7 total DNA and to pDI1. This further confirms that the *phn* genes are not homologous to the *nah*-type genes, and validates the earlier Southern hybridisation and PCR results described in 4.1.

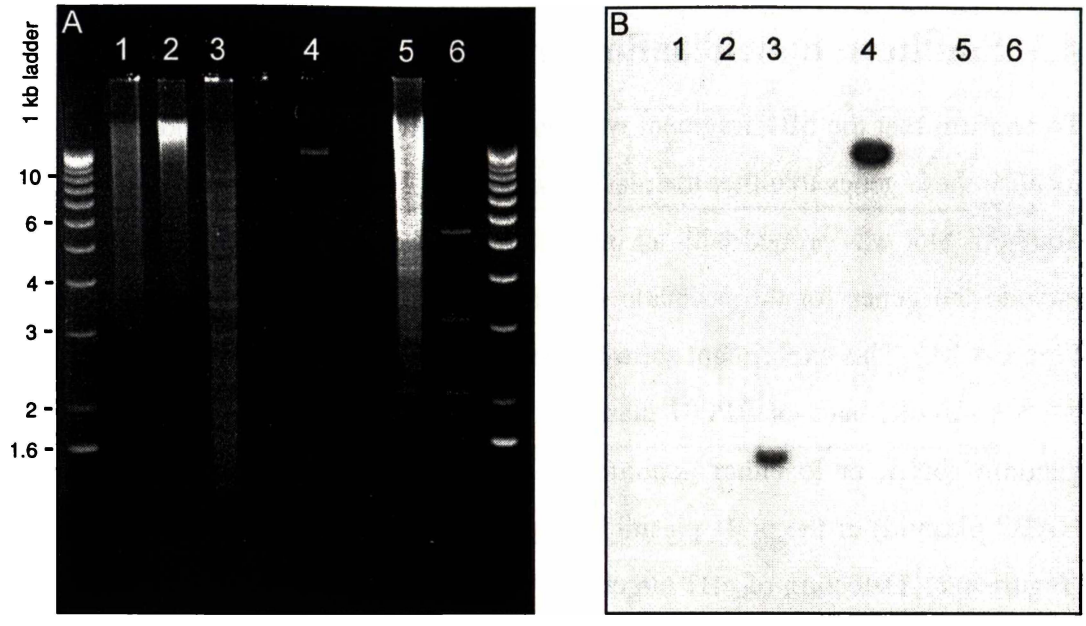


Figure 4.2 pB1 hybridisation experiment. (A) Agarose gel used for Southern blotting experiment; lane 1, *Eco*RI-digested RP007 genomic DNA; 2, *Hin*dIII- digested RP007 genomic DNA; 3, *Eco*RI-digested RP007 plasmid DNA; 4, *Hin*dIII-digested pB1; 5, *Eco*RI-digested *P. putida* G7 genomic DNA; 6, *Eco*RI-*Hin*dIII-digested pDI1. Both outside lanes are 1 kb ladder. (B) Autoradiogram showing hybridisation of [^{32}P]-labelled probe (1.5 kb *Eco*RI fragment of pB1) to RP007 plasmid DNA (lane 3) and the pB1 positive control (lane 4). The lanes are the same as for (A).

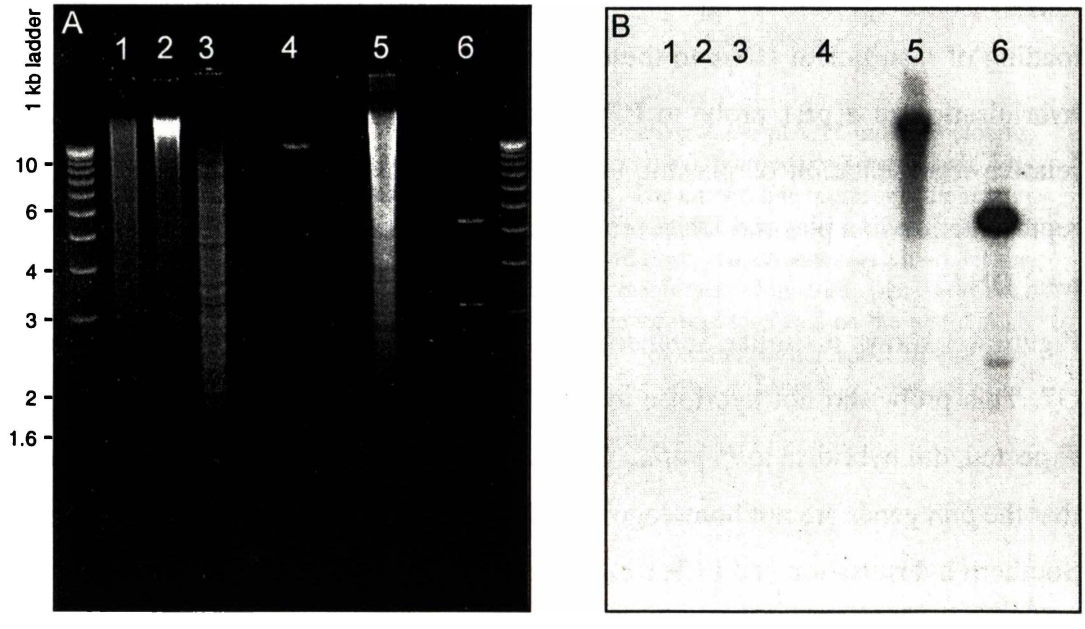


Figure 4.3 *nahAc* hybridisation experiment. (A) Agarose gel used for Southern blotting experiment; lane 1, *Eco*RI-digested RP007 genomic DNA; 2, *Hin*dIII- digested RP007 genomic DNA; 3, *Eco*RI-digested RP007 plasmid DNA; 4, *Hin*dIII-digested pB1; 5, *Eco*RI-digested *P. putida* G7 genomic DNA; 6, *Eco*RI-*Hin*dIII-digested pDI1. Both outside lanes are 1 kb ladder. (B) Autoradiogram showing hybridisation of [^{32}P]-labelled probe (1 kb region of *nahAc* gene) to *P. putida* G7 total genomic DNA (lane 5) and the pDI1 plasmid carrying the *pah* genes of *P. putida* OUS82 (lane 6). The lanes are the same as for (A).

Extract or standard	Solvent system	
	chloroform:methanol	benzene:acetate:water
[a]*E. coli* (pB1)	0.058	0.95
[a]wild-type RP007	0.058	0.95
salicylic acid std	0.058	0.95
naphthalene std	-	0.99

[a] Migration of purple fluorescent metabolite

Table 4.1 Thin layer chromatography (TLC) Rf values of purple fluorescent metabolites from *E. coli* (pB1) and RP007 incubations with naphthalene, and comparison with Rf values of authentic standards.

4.4 Naphthalene transformation by pB1

Incubation of a high density suspension of *E. coli* carrying the pB1 plasmid indicated expression of the pB1 genes in *E. coli* imparted the ability to transform naphthalene to salicylic acid. For this experiment *E. coli* DH5α was used as a negative control and the wild-type *Burkholderia* sp. RP007 included as a positive control, since it had been shown previously that a high density cell suspension of RP007 transformed [1-^{14}C]naphthalene to [^{14}C]salicylate (section 3.4.1).

After overnight incubation of cells with naphthalene, supernatants were extracted and analysed by thin layer chromatography (TLC) as described in Chapter 2. Using both solvent systems, *E. coli* (pB1) and wild-type RP007 gave a fluorescent blue metabolite under 302 nm which had the same Rf values as an authentic salicylic acid standard (Table 4.1). Shamsuzzaman & Barnsley (1974a) also detected salicylic acid on TLC plates by purple fluorescence under UV. The extract from the *E. coli* DH5α control did not show any fluorescent metabolites at similar Rf values.

Based on the similarity of the Rf values of the *E. coli* (pB1) blue fluorescent metabolite to those of the salicylic acid standard, and to the metabolite produced by the wild-type RP007, which has previously been shown to accumulate salicylic acid under these conditions, it appears *E. coli* (pB1) has transformed naphthalene to salicylic acid. This suggest the pB1 insert carries genes encoding a complete upper pathway for naphthalene/phenanthrene degradation.

4.5 Nucleotide sequencing of the pB1 fragment

The pB1 construct carries an 11.5 kb fragment of RP007 DNA which encodes an indole-positive phenotype which was localised to the 3' end of the fragment, and which imparts on *E. coli* the ability to transform naphthalene to salicylic acid. The pB1 fragment showed no detectable homology to previously characterised genes for naphthalene and phenanthrene degradation. Initial nucleotide sequencing from each end of the pB1 fragment using M13/pUC forward and reverse sequencing primers showed sequence homology indicative of an *ntrC*-like regulatory gene at the 5' end, and a naphthalene dihydrodiol dehydrogenase at the 3' end. This suggested that much of an operon for phenanthrene/naphthalene catabolism was present on the pB1 insert and on this basis it was decided to sequence the entire length of the pB1 fragment.

4.5.1 Subcloning and construction of nested deletion derivatives

The strategy for determining the nucleotide sequence of pB1 involved using the internal *Eco*RI restriction sites (and a *Pst*I site near the 3' end) to obtain subclones of the larger fragment. The smaller fragments were cloned into pUC18 in both directions, and from each of these was derived a nested deletion series which was used for the sequencing reactions (Figure 4.4). The exception to this method was the sequencing of pBE1 (from the 5' end of the pB1 insert) which could not be used for a nested deletion experiment since it was an *Eco*RI-*Hin*dIII clone and therefore did not have a complete multiple cloning site (MCS). In order to sequence this fragment in both directions a number of subclones were derived which utilised the internal restriction sites.

4.5.2 Assembly of the pB1 sequence

After complete nucleotide sequencing of both strands of the *Pst*I (pB4) and *Eco*RI subclones (pBE1–pBE6) of pB1, the fragments were assembled to give the complete pB1 sequence, cds 1–11451 (Figure 4.5). Assembly of pBE6 and pB4 was simplified since these fragments overlapped by 436 bp. The orientation of the pBE1 fragment was known based on the location of the *Hin*dIII site which marked the 5' end of the pB1 insert. The remaining fragments were assembled in accordance with their known position on the pB1 restriction map, and the orientations were determined by

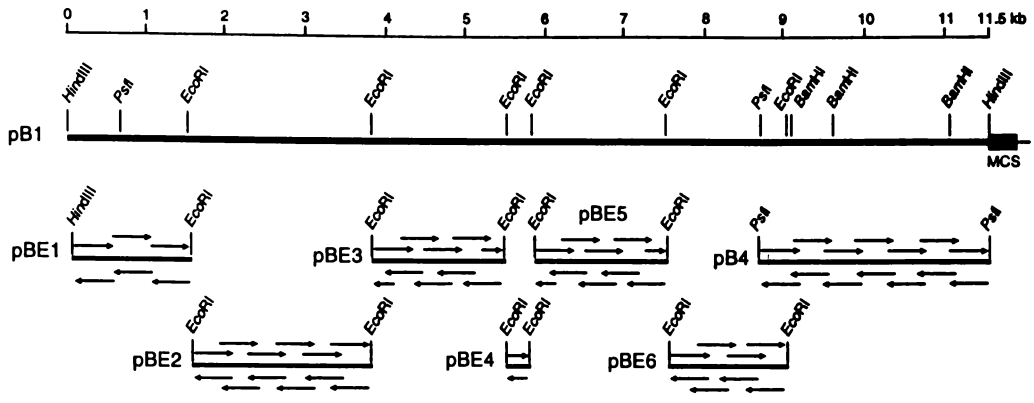
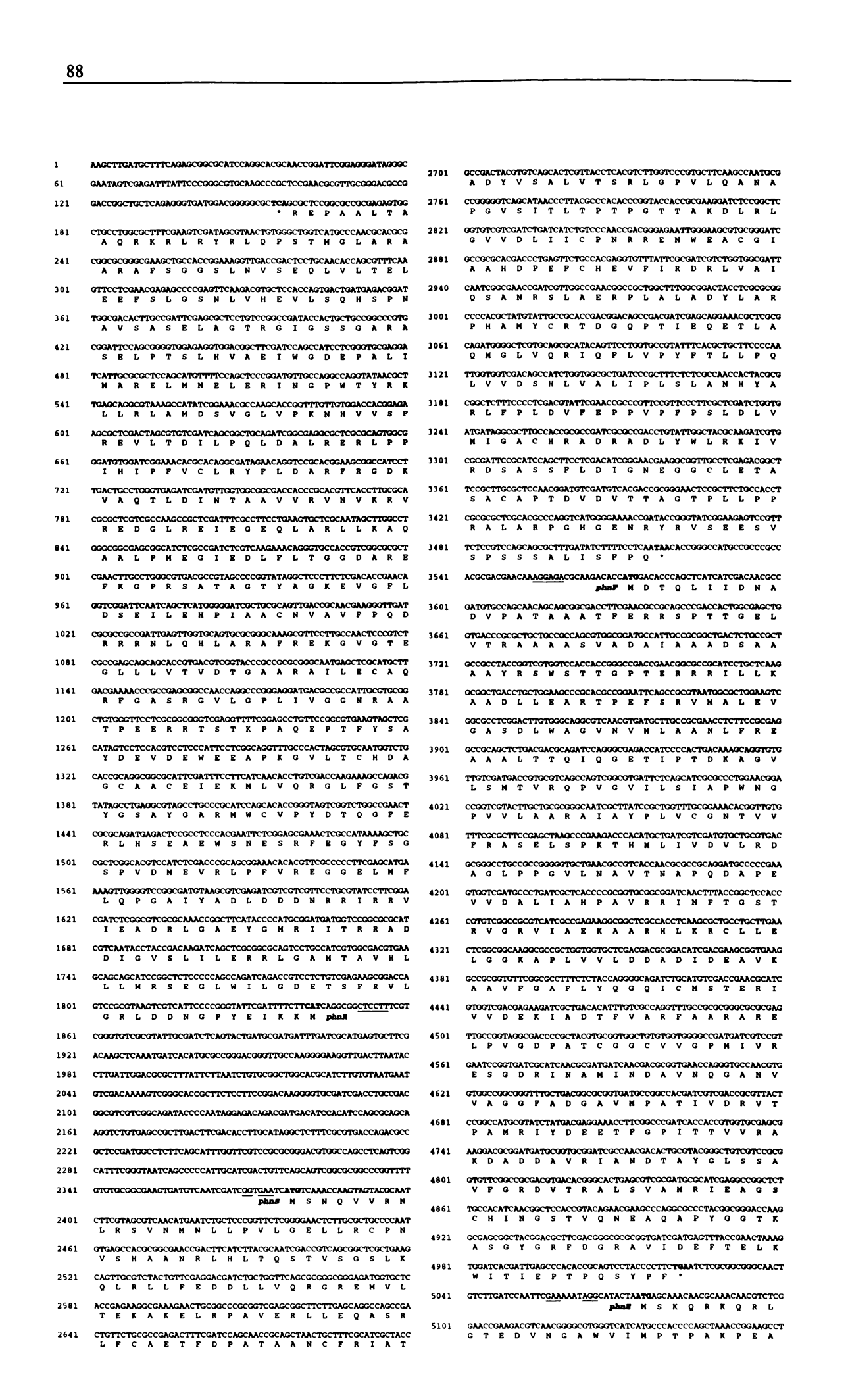


Figure 4.4 The strategy used to obtain nucleotide sequence data in both directions for the 11.5 kb pB1 fragment of RP007. The position of the subclones relative to the parent pB1 fragment is shown, and details of these subclones are given in Table B.2 (Appendix B). Small arrows represent the regions included in individual sequencing reactions from the subclones or nested deletion derivatives.

comparison of the restriction sites in the derived sequence with those of the restriction map. This assembly was verified by careful comparison of the nucleotide sequence across joined regions with the sequence of homologous genes, and of the transcribed amino acid sequence of these regions.


1 AAGCTTGATGCTTTCAGAGCGGCGCATCCAGGCACGCAACCGGATTCGGAGGGATAGGGC
61 GAATAGTCGAGATTTATTCCCGGGCGTGCAAGCCCGCTCCGAACGCGTTGCGGGACGCCG
121 GACCGGCTGCTCAGAGGGTGATGGACGGGGCGCTCAGCGCTCCGGCGCCGCGAGAGTGG
* R E P A A L T A
181 CTGCCTGGCGCTTTCGAAGTCGATAGCGTAACTGTGGGCTGGTCATGCCCAACGCACGCG
A Q R K R L R Y R L Q P S T M G L A R A
241 CGGCGCGGGCGAAGCTGCCACCGGAAAGGTTGACCGACTCCTGCAACACCAGCGTTTCAA
A R A F S G G S L N V S E Q L V L T E L
301 GTTCCTCGAACGAGAGCCCCGAGTTCAAGACGTGCTCCACCAGTGACTGATGAGACGGAT
E E F S L Q S N L V H E V L S Q H S P N
361 TGGCGACACTTGCCGATTCGAGCGCTCCTGTCCGGCCGATACCACTGCTGCCGGCCCGTG
A V S A S E L A G T R G I G S S G A R A
421 CGGATTCCAGCGGGGTGGAGAGGTGGACGGCTTCGATCCAGCCATCCTCGGGTGCGAGGA
S E L P T S L H V A E I W G D E P A L I
481 TCATTGCGCGCTCCAGCATGTTTTCCAGCTCCCGGATGTTGCCAGGCCAGGTATAACGCT
M A R E L M N E L E R I N G P W T Y R K
541 TGAGCAGGCGTAAAGCCATATCGGAAACGCCAAGCACCGGTTTGTTGTGGACCACGGAGA
L L R L A M D S V G L V P K M H V V S F
601 AGCGCTCGACTAGCGTGTCGATCAGCGGCTGCAGATCGGCGAGGCGCTCGCGCAGTGGCG
R E V L T D I L P Q L D A L R E R L P P
661 GGATGTGGATCGGAAACACGCACAGGCGATAGAACAGGTCCGCACGGAAGCGGCCATCCT
I H I P F V C L R Y F L D A R F R G D K
721 TGACTGCCTGGGTGAGATCGATGTTGGTGGCGGCGACCACCCGCACGTTCACCTTGCGCA
V A Q T L D I N T A A V V R V N V K R V
781 CGCGCTCGTCGCCAAGCCGCTCGATTTCGCCTTCCTGAAGTGCTCGCAATAGCTTGGCCT
R E D G L R E I E G E Q L A R L L K A Q
841 GGGCGGCGAGCGGCATCTCGCCGATCTCGTCAAGAAACAGGGTGCCACCGTCGGCGCGCT
A A L P M E G I E D L F L T G G D A R E
901 CGAACTTGCCTGGGCGTGACGCCGTAGCCCCGGTATAGGCTCCCTTCTCGACACCGAACA
F K G P R S A T A G T Y A G K E V G F L
961 GGTCGGATTCAATCAGCTCATGGGGGATCGCTGCGCAGTTGACCGCAACGAAGGGTTGAT
D S E I L E H P I A A C N V A V F P Q D
1021 CGCGCCGCCGATTAGTTGGTGCAGTGCGCGGGCAAAGCGTTCCTTGCCAACTCCCGTCT
R R R N L Q H L A R A F R E K G V G T E
1081 CGCCGAGCAGCAGCACCGTGACGTCGGTACCCGCCGCGCGGGCAATGAGCTCGCATGCTT
G L L L V T V D T G A A R A I L E C A Q
1141 GACGAAAACCCGCCGAGCGGCCAACCAGGCCCGGGAGGATGACGCCGCCATTGCGTGCGG
R F G A S R G V L G P L I V G G N R A A
1201 CTGTGGGTTCCTCGCGGCGGGTCGAGGTTTTCGGAGCCTGTTCCGGCGTGAAGTAGCTCG
T P E E R R T S T K P A Q E P T F Y S A
1261 CATAGTCCTCCACGTCCTCCCATTCCTCGGCAGGTTTGCCCACTAGCGTGCAATGGTCTG
Y D E V D E W E E A P K G V L T C H D A
1321 CACCGCAGGCGGCGCATTCGATTTCCTTCATCAACACCTGTCGACCAAGAAAGCCAGACG
G C A A C E I E K M L V Q R G L F G S T
1381 TATAGCCTGAGGCGTAGCCTGCCCGCATCCAGCACACCGGGTAGTCGGTCTGGCCGAACT
Y G S A Y G A R M W C V P Y D T Q G F E
1441 CGCGCAGATGAGACTCCGCCTCCCACGAATTCTCGGAGCGAAACTCGCCATAAAAGCTGC
R L H S E A E W S N E S R F E G Y F S G
1501 CGCTCGGCACGTCCATCTCGACCCGCAGCGGAAACACACGTTCGCCCCCTTCGAGCATGA
S P V D M E V R L P F V R E G G E L M F
1561 AAAGTTGGGGTCCGGCGATGTAAGCGTCGAGATCGTCGTTCCTGCGTATCCTTCGGA
L Q P G A I Y A D L D D D N R R I R R V
1621 CGATCTCGGCGTCGCGCAAACCGGCTTCATACCCCATGCGGATGATGGTCCGGCGCGCAT
I E A D R L G A E Y G M R I I T R R A D
1681 CGTCAATACCTACCGACAAGATCAGCTCGCGGCGCAGTCCTGCCATCGTGGCGACGTGAA
D I G V S L I L E R R L G A M T A V H L
1741 GCAGCAGCATCCGGCTCTCCCCCAGCCAGATCAGACCGTCCTCTGTCGAGAAGCGGACCA
L L M R S E G L W I L G D E T S F R V L
1801 GTCCGCGTAAGTCGTCATTCCCCGGGTATTCGATTTTCTTCATCAGGCGGCTCCTTTCGT
G R L D D N G P Y E I K K M phaR
1861 CGGGTGTCGCGTATTGCGATCTCAGTACTGATGCGATGATTTGATCGCATGAGTGCTTCG
1921 ACAAGCTCAAATGATCACATGCGCCGGGACGGGTTGCCAAGGGGAAGGTTGACTTAATAC
1981 CTTGATTGGACGCGCTTTATTCTTAATCTGTGCGGCTGGCACGCATCTTGTGTAATGAAT
2041 GTCGACAAAAGTCGGGCACCGCTTCTCCTTCCGGACAAGGGGTGCGATCGACCTGCCGAC
2101 GGCGTCGTCGGCAGATACCCCAATAGGAGACAGACGATGACATCCACATCCAGCGCAGCA
2161 AGGTCTGTGAGCCGCTTGACTTCGACACCTTGCATAGGCTCTTTCGCGTGACCAGACGCC
2221 GCTCCGATGGCCTCTTCAGCATTTGGTTCGTCCGCGCGGGACGTGGCCAGCCTCAGTCGG
2281 CATTTCGGGTAATCAGCCCCCATTGCATCGACTGTTCAGCAGTCGGCGCGGCCCGGTTTT
2341 GTGTGCGGCGAAGTGATGTCAATCGATCGGTGAATCATGTCAAACCAAGTAGTACGCAAT
phaB M S N Q V V R N
2401 CTTCGTAGCGTCAACATGAATCTGCTCCCGGTTCTCGGGGAACTCTTGCGCTGCCCCAAT
L R S V N M N L L P V L G E L L R C P N
2461 GTGAGCCACGCGGCGAACCGACTTCATCTTACGCAATCGACCGTCAGCGGCTCGCTGAAG
V S H A A N R L H L T Q S T V S G S L K
2521 CAGTTGCGTCTACTGTTCGAGGACGATCTGCTGGTTCAGCGCGGCGGGAGATGGTGCTC
Q L R L L F E D D L L V Q R G R E M V L
2581 ACCGAGAAGGCGAAAGAACTGCGGCCCGCGGTCGAGCGGCTTCTTGAGCAGGCCAGCGA
T E K A K E L R P A V E R L L E Q A S R
2641 CTGTTCTGCGCCGAGACTTTCGATCCAGCAACCGCAGCTAACTGCTTTCGCATCGCTACC
L F C A E T F D P A T A A N C F R I A T
2701 GCCGACTACGTGTCAGCACTCGTTACCTCACGTCTTGGTCCCGTGCTTCAAGCCAATGCG
A D Y V S A L V T S R L G P V L Q A N A
2761 CCGGGGGTCAGCATAACCCTTACGCCCACACCCGGTACCACCGCGAAGGATCTCCGGCTC
P G V S I T L T P T P G T T A K D L R L
2821 GGTGTCGTCGATCTGATCATCTGTCCCAACCGACGGGAGAATTGGGAAGCGTGCGGGATC
G V V D L I I C P N R R E N W E A C G I
2881 GCCGCGCACGACCCTGAGTTCTGCCACGAGGTGTTTATTCGCGATCGTCTGGTGGCGATT
A A H D P E F C H E V F I R D R L V A I
2940 CAATCGGCGAACCGATCGTTGGCCGAACGGCCGCTGGCTTTGGCGGACTACCTCGCGCGG
Q S A N R S L A E R P L A L A D Y L A R
3001 CCCCACGCTATGTATTGCCGCACCGACGGACAGCCGACGATCGAGCAGGAAACGCTCGCG
P H A M Y C R T D Q Q P T I E Q E T L A
3061 CAGATGGGGCTCGTGCAGCGCATACAGTTCCTGGTGCCGTATTTCACGCTGCTTCCCCAA
Q M G L V Q R I Q F L V P Y F T L L P Q
3121 TTGGTGGTCGACAGCCATCTGGTGGCGCTGATCCCGCTTTCTCTCGCCAACCACTACGCG
L V V D S H L V A L I P L S L A N H Y A
3181 CGGCTCTTTCCCCTCGACGTATTCGAACCGCCGTTCCGTTCCCTTCGCTCGATCTGGTG
R L F P L D V F E P P V P F P S L D L V
3241 ATGATAGGCGCTTGCCACCGCGCCGATCGCGCCGACCGTATTGGCTACGCAAGATCGTG
M I G A C H R A D R A D L Y W L R K I V
3301 CGCGATTCCGCATCCAGCTTCCTCGACATCGGGAACGAAGGCGGTTGCCTCGAGACGGCT
R D S A S S F L D I G N E G G C L E T A
3361 TCCGCTTGCGCTCCAACGGATGTCGATGTCACGACCGCGGGAACTCCGCTTCTGCCACCT
S A C A P T D V D V T T A G T P L L P P
3421 CGCGCGCTCGCACGCCCAGGTCATGGGGAAAACCGATACCGGGTATCGGAAGAGTCCGTT
R A L A R P G H G E N R Y R V S E E S V
3481 TCTCCGTCCAGCAGCGCTTTGATATCTTTTCCTCAATAACACCGGGCCATGCCGCCCGCC
S P S S S A L I S F P Q *
3541 ACGCGACGAACAAAGGAGACGCAAGACACCATGGACACCCAGCTCATCATCGACAACGCC
phaF M D T Q L I I D N A
3601 GATGTGCCAGCAACAGCAGCGGCGACCTTCGAACGCCGCAGCCCGACCACTGGCGAGCTG
D V P A T A A A T F E R R S P T T G E L
3661 GTGACCCGCGCTGCTGCCGCCAGCGTGGCGGATGCCATTGCCGCCGGCTGACTCTGCCGCT
V T R A A A A S V A D A I A A A D S A A
3721 GCCGCCTACCGGTCGTGGTCCACCACCGGGCCGACCGAACGGCGCCGCATCCTGCTCAAG
A A Y R S W S T T G P T E R R R I L L K
3781 GCGGCTGACCTGCTGGAAGCCCGCACGCCGGAATTCAGCCGCGTAATGGCGCTGGAAGTC
A A D L L E A R T P E F S R V M A L E V
3841 GGCGCCTCGGACTTGTGGGCAGGCGTCAACGTGATGCTTGCCGCGAACCTCTTCCGCGAG
G A S D L W A G V N V M L A A N L F R E
3901 GCCGCAGCTCTGACGACGCAGATCCAGGGCGAGACCATCCCCACTGACAAAGCAGGTGTG
A A A L T T Q I Q G E T I P T D K A G V
3961 TTGTCGATGACCGTGCGTCAGCCAGTCGGCGTGATTCTCAGCATCGCGCCCTGGAACGGA
L S M T V R Q P V G V I L S I A P W N G
4021 CCGGTCGTACTTGCTGCGCGGGCAATCGCTTATCCGCTGGTTTGCGGAAACACGGTTGTG
P V V L A A R A I A Y P L V C G N T V V
4081 TTTCGCGCTTCCGAGCTAAGCCCGAAGACCCACATGCTGATCGTCGATGTGCTGCGTGAC
F R A S E L S P K T H M L I V D V L R D
4141 GCGGGCCTGCCGCCGGGGGTGCTGAACGCCGTCACCAACGCGCCGCAGGATGCCCCCGAA
A G L P P G V L N A V T N A P Q D A P E
4201 GTGGTCGATGCCCTGATCGCTCACCCCGCGGTGCCGCCGATCAACTTTACCGGCTCCACC
V V D A L I A H P A V R R I N F T G S T
4261 CGTGTCGGCCGCGTCATCGCCGAGAAGGCGGCTCGCCACCTCAAGCGCTGCCTGCTTGAA
R V G R V I A E K A A R H L K R C L L E
4321 CTCGGCGGCAAGGCGCCGCTGGTGGTGCTCGACGACGCGGACATCGACGAAGCGGTGAAG
L G G K A P L V V L D D A D I D E A V K
4381 GCCGCGGTGTTCGGCGCCTTTCTCTACCAGGGGCAGATCTGCATGTCGACCGAACGCATC
A A V F G A F L Y Q G Q I C M S T E R I
4441 GTGGTCGACGAGAAGATCGCTGACACATTTGTCGCCAGGTTTGCCGCGCGGGCGCGCGAG
V V D E K I A D T F V A R F A A R A R E
4501 TTGCCGGTAGGCGACCCCGCTACGTGCGGTGGCTGTGTGGTGGGGCCGATGATCGTCCGT
L P V G D P A T C G G C V V G P M I V R
4561 GAATCCGGTGATCGCATCAACGCGATGATCAACGACGCGGTGAACCAGGGTGCCAACGTG
E S G D R I N A M I N D A V N Q G A N V
4621 GTGGCCGGCGGGTTTGCTGACGGCGCGGTGATGCCGGCCACGATCGTCGACCGCGTTACT
V A G G F A D G A V M P A T I V D R V T
4681 CCGGCCATGCGTATCTATGACGAGGAAACCTTCGGCCCGATCACCACCGTGGTGCGAGCG
P A M R I Y D E E T F G P I T T V V R A
4741 AAGGACGCGGATGATGCGGTGCGGATCGCCAACGACACTGCGTACGGGCTGTCGTCCGCG
K D A D D A V R I A N D T A Y G L S S A
4801 GTGTTCGGCCGCGACGTGACACGGGCACTGAGCGTCGCGATGCGCATCGAGGCCGGCTCT
V F G R D V T R A L S V A M R I E A G S
4861 TGCCACATCAACGGCTCCACCGTACAGAACGAAGCCCAGGCGCCCTACGGCGGGACCAAG
C H I N G S T V Q N E A Q A P Y G G T K
4921 GCGAGCGGCTACGGACGCTTCGACGGGCGCGCGGTGATCGATGAGTTTACCGAACTAAAG
A S G Y G R F D G R A V I D E F T E L K
4981 TGGATCACGATTGAGCCCACACCGCAGTCCTACCCCTTCTGAATCTCGCGGCGGGCAACT
W I T I E P T P Q S Y P F *
5041 GTCTTGATCCAATTCGAAAAATAGGCATACTAATGAGCAAACAACGCAAACAACGTCTCG
phaE M S K Q R K Q R L
5101 GAACCGAAGACGTCAACGGGCGTGGGTCATCATGCCCACCCCAGCTAAACCGGAAGCCT
G T E D V N G A W V I M P T P A K P E A

```
5161  CTGACTGGCGCGCCACCGACACGGTCGATCTGGATGAGACCGCCCGCATTGTCGAGGCGC
        S  D  W  R  A  T  D  T  V  D  L  D  E  T  A  R  I  V  E  A

5221  TGATCGATTCCGGCGTGAACGGCATTCTTAGCCTGGGCACTTTCGGCGAATGTGCCACGC
        L  I  D  S  G  V  N  G  I  L  S  L  G  T  F  G  E  C  A  T

5281  TGACCTGGGAGGAAAAGCAGGCTTTTATTGGCGCCGTCGTCGAGACTACGCGTGGCCGCG
        L  T  W  E  E  K  Q  A  F  I  G  A  V  V  E  T  T  R  G  R

5341  TGCCATTTTTCTGCGGTACCACTGCGCTCAACACGCGTGAGGTGGTCCGCCAGACGCGTG
        V  P  F  F  C  G  T  T  A  L  N  T  R  E  V  V  R  Q  T  R

5401  CCGCGCTGGACATCGGCGTTGACGGTACGATGCTCGGCGTACCGATGTGGTCCAGGATGG
        A  A  L  D  I  G  V  D  G  T  M  L  G  V  P  M  W  S  R  M

5461  AAGTGCCTGCCGCTGTGCAGTTCTACCGCGACGTCGCGGAGGCCTGCCCCGAAGCCGCGA
        E  V  P  A  A  V  Q  F  Y  R  D  V  A  E  A  C  P  E  A  A

5521  TCGCGGTCTATGCAAACGCGGATGCGTTCAAGTTCGAATTCCCGCGCGCCTTTTGGGCCC
        I  A  V  Y  A  N  A  D  A  F  K  F  E  F  P  R  A  F  W  A

5581  AGGTAGCCCAGATTCCGCAGGTCGTCACCGCCAAGTATCTGGGCATCGGAATGCTGGACC
        Q  V  A  Q  I  P  Q  V  V  T  A  K  Y  L  G  I  G  M  L  D

5641  TCGATCTCACGCTGGCACCGGGCATTCGCTTCCTGCCGCACGAGGACGACTACTATGCGG
        L  D  L  T  L  A  P  G  I  R  F  L  P  H  E  D  D  Y  Y  A

5701  CCGCGCGCGTGGCGCCAGAGCGCGTGACCGCGTTTTGGTCGAGTGGTGCGATGTGCGGGC
        A  A  R  V  A  P  E  R  V  T  A  F  W  S  S  G  A  M  C  G

5761  CGGCCACCGCTATTCGACTGCGCGACGAGGTGGCGAAGGCCAAGCAGACCGGGGATTGGC
        P  A  T  A  I  R  L  R  D  E  V  A  K  A  K  Q  T  G  D  W

5821  GTCTCGCCAAGGAACTTTCAGACGCGATGCGGCGTGCCGACGCCACGCTGTTTCCGCGCG
        R  L  A  K  E  L  S  D  A  M  R  R  A  D  A  T  L  F  P  R

5881  GGGATTTCGCCGAATTCTCCAAATACAACATCGCCATCGAGAAGGAGCGCATGAACGCCG
        G  D  F  A  E  F  S  K  Y  N  I  A  I  E  K  E  R  M  N  A

5941  CTGGCTGGCTCAGAGCGGGGCCATGCCGGCCCCCCTATCACATTGCGCCAGAAGAGTATC
        A  G  W  L  R  A  G  P  C  R  P  P  Y  H  I  A  P  E  E  Y

6001  TGGATGGTGCACGCCAGTCGGGCAGGGCATGGGCCGAATTGCACCAGCAATATTCAGATC
        L  D  G  A  R  Q  S  G  R  A  W  A  E  L  H  Q  Q  Y  S  D

6061  TGTAATGTAATTAGGTTAACTAAATTATGTGAAGGTTTTGAGAACAGCACCGACTGCCGA
        L  *

6121  TAACCGGCGGTACGGTGGCTAACTGGAAAAAATGGAGATGAGGATGAGCATCTACACCGT

6181  AGAACGCGTGCTGTGGGATCTGCATGCCCAGCCCGATCTCGCCAGCCGCTTCCGCGCGGA

6241  GCCGGAGGGGGTGCTCGACGGGTACCCGCTGGAAGGGGACGAGCGCGGGCTGGTTACGTC

6301  GTTGGATGTGCGTGCGATGGCTGATCGTGGCGCCAGCCAGATGCTGTTGTTCGTTGCCTG

6361  GACGGCGTTGCAGGGACCCGGCACTACGCCGGAATACATGCGGCGGATGAACACTCGTCT

6421  TTAAAGACGGGGATGAATAACCGGGTGGTCCAGAGAGACCATCAACAAGGAGACAAGA
                                                    phnC

6481  TGGCGAAAATCGTAGGGGGATTCATGATGCCCCACGATCCACTGATCCCCGCGACGCCTA
        M  A  K  I  V  G  G  F  M  M  P  H  D  P  L  I  P  A  T  P

6541  CTGCGCCACCCGCAGCGCAGCGCGAGATTTGCATGCACGCTTACGCAATCATTGTCGAGC
        T  A  P  P  A  A  Q  R  E  I  C  M  H  A  Y  A  I  I  V  E

6601  GCCTGCGTGCACTACAGGTCGACACTGTAGTGGTGATCGCGGATGACCATTACACCCTGA
        R  L  R  A  L  Q  V  D  T  V  V  V  I  A  D  D  H  Y  T  L

6661  ACGGGCCGTACTGCATCCCGATGGCGATGATCGGCATCGGTGACATCGAGGGTCCTTATG
        N  G  P  Y  C  I  P  M  A  M  I  G  I  G  D  I  E  G  P  Y

6721  AGCCGTGGCTGGGGATTCCGCGCGCCCAGATCGAAAACAACAGGCCACTGGCACATCACA
        E  P  W  L  G  I  P  R  A  Q  I  E  N  N  R  P  L  A  H  H

6781  TCATGCAGTACGGCCTCGAGTACGGCATCGATTGGGCGGTTTCGAAGTCGCTCGTTCTCG
        I  M  Q  Y  G  L  E  Y  G  I  D  W  A  V  S  K  S  L  V  L

6841  ATCACTCGGCGACCGTACCGATCCACTACGCGGTGCGTCCGGTGAAGGGTATGCGAGCGA
        D  H  S  A  T  V  P  I  H  Y  A  V  R  P  V  K  G  M  R  A

6901  TTCCGGTTTATCTCAACACTGGAATCGAACCGTTTATCACCAGCTGGCGCGCTCACGAGA
        I  P  V  Y  L  N  T  G  I  E  P  F  I  T  S  W  R  A  H  E

6961  TTGGACGTGTGATCGGCGACGCGGTCGCCAACTGGCAGGGCGACGAACGGGTGGCGATTT
        I  G  R  V  I  G  D  A  V  A  N  W  Q  G  D  E  R  V  A  I

7021  ACGGTACCGGCGGTCTCAGCCATTGGCCAGGCATGGCCGAAATGGGAGCAATCAATGAGG
        Y  G  T  G  G  L  S  H  W  P  G  M  A  E  M  G  A  I  N  E

7081  GGTGGGATCGCAAGATCATGAAACTGGTGGCGCAGGGCGACGTAGAGTCCCTGATCGCCC
        G  W  D  R  K  I  M  K  L  V  A  Q  G  D  V  E  S  L  I  A

7141  TCAGCGATGAAGAGATCTTACGCGACGGTGGCAATGGCGGAATTGAGATCAAGAACTGGA
        L  S  D  E  E  I  L  R  D  G  G  N  G  G  I  E  I  K  N  W

7201  TTTGCGCGATGGGGGCGCTCGGCGCCTGTCGGGGTGAAGTGATCGCGTACGAGCCGGTGG
        I  C  A  M  G  A  L  G  A  C  R  G  E  V  I  A  Y  E  P  V

7261  CCGAGTGGGTTTGCGGCTGCGGTTATATGGAAATGAAGGTGGCGTGAGGTCCTGTTGTGT
        A  E  W  V  C  G  C  G  Y  M  E  M  K  V  A  *

7321  CTCGATGGAGAGTGTGTCATGACGATCGATTTCTTCTTTGATTTTCTCAGCCCGTATGCC
                    phnD  M  T  I  D  F  F  F  D  F  L  S  P  Y  A

7381  TACCTGGCCCGCCATCGGTTGACCCAGGTCGCGGCGCTGCATGGTTGTGCGATCGCGTAT
        Y  L  A  R  H  R  L  T  Q  V  A  A  L  H  G  C  A  I  A  Y

7441  AAACCGATTGACCTGGCACGTGCCAAGCTGGCCATCGGCAACACCGGCCCAGCAAATCGC
        K  P  I  D  L  A  R  A  K  L  A  I  G  N  T  G  P  A  N  R

7501  GACATGCCAGTCAAGCTGGCATATGTCGTGGAAGATCTTAAGCGCTGGGCCGCGCGCTAT
        D  M  P  V  K  L  A  Y  V  V  E  D  L  K  R  W  A  A  R  Y

7561  CGGATTCCAATCGAATTCATCAAGAACTTCAACACAAAGCGAATGAACGTCGGCACGTTC
        R  I  P  I  E  F  I  K  N  F  N  T  K  R  M  N  V  G  T  F

7621  TATGCCGAAGCAAGGGGGCAGCAGGCCGACTACGTCCGACAGGCATACCACCTGGCGTGG
        Y  A  E  A  R  G  Q  Q  A  D  Y  V  R  Q  A  Y  H  L  A  W

7681  GGCGAAGGGGGCGCGCCGGACGACGATGCGGCGTTGCGCTCGATCGCCGTGTCAATGGGG
        G  E  G  G  A  P  D  D  D  A  A  L  R  S  I  A  V  S  M  G

7741  TGGGACGCAGCCGACTTCTTGCGATTTCTCGACTCGTCGGAAGCCGAGACGGCCTACAAC
        W  D  A  A  D  F  L  R  F  L  D  S  S  E  A  E  T  A  Y  N

7801  GAATCAACTTTAGAGGCGATCTCTGCCGGCGTCTTTGGCGTTCCGACGATGGCTGTCGGT
        E  S  T  L  E  A  I  S  A  G  V  F  G  V  P  T  M  A  V  G

7861  CGGGACATGTGGTGGGGTAACGACCGCATCGATTTTCTTGAAACCCATCTGGGTCGAATT
        R  D  M  W  W  G  N  D  R  I  D  F  L  E  T  H  L  G  R  I

7921  GCCGGGTAGACAATATAATTTTTTTGAAAGCAGGATCTACTTGATGGATCATAAACACAG
        A  G  *

7981  GAGGATGACAAATGGATACTACTCATACAATTAATTTCATCGACGTTGAGCATGGACGTC
                phnAc  M  D  T  T  H  T  I  N  F  I  D  V  E  H  G  R

8041  AAAGTCATAAAATATATCAGGACCAGACATTATTCGAGCTGGAATGTGAGCGCATCTTCG
        Q  S  H  K  I  Y  Q  D  Q  T  L  F  E  L  E  C  E  R  I  F

8101  CGCGCTGCTGGCTTTTTTTGACGCATGAATGCGCTATTCCCAATTACGGTGATTTCGTGA
        A  R  C  W  L  F  L  T  H  E  C  A  I  P  N  Y  G  D  F  V

8161  CCACGCGCATGGGAACGGATGAGGTGATTGTGATGCGGCAGAAGGATAACTCGATTAAAG
        T  T  R  M  G  T  D  E  V  I  V  M  R  Q  K  D  N  S  I  K

8221  CCTTTCTCAACGTTTGCAGGCATCGCGGTGCGCGGCTATGCGCAGTCGAAGCGGGTAACG
        A  F  L  N  V  C  R  H  R  G  A  R  L  C  A  V  E  A  G  N

8281  CGCGGGGCTTCGCCTGCAATTATCATGGCTGGGCGTATGGGGCAGATGGATCGCTTCTAA
        A  R  G  F  A  C  N  Y  H  G  W  A  Y  G  A  D  G  S  L  L

8341  GTGTTCCATTTGAGAATGAAATTTATCAAAAGCGTCTGGATAAATGCAAAAACGGTCTGC
        S  V  P  F  E  N  E  I  Y  Q  K  R  L  D  K  C  K  N  G  L

8401  ATGAAATTGGCAAGGTGGAGTCCTATCATGGATTCGTTTACGGCTGCTTCGATACAGATG
        H  E  I  G  K  V  E  S  Y  H  G  F  V  Y  G  C  F  D  T  D

8461  CGCCCAATCTAAAGGAATACCTGGGCGATTTTGCATGGTATCTCGATCTTTGGATGGAAG
        A  P  N  L  K  E  Y  L  G  D  F  A  W  Y  L  D  L  W  M  E

8521  CCGGCGGCGGCATCGAACTCGTCGGGCCGCCTGCCCGCTGCTTCATCGAAGCCAATTGGA
        A  G  G  G  I  E  L  V  G  P  P  A  R  C  F  I  E  A  N  W

8581  AAGCGCCGTCAGAGAATTTTGTCGGTGACGCCTATCATGTCGGCTGGACACATGCATCCG
        K  A  P  S  E  N  F  V  G  D  A  Y  H  V  G  W  T  H  A  S

8641  CGCTTAGATCCGGTCAGTCCGGGTTCGCCGGCATGGCAGGGAATAACGTGCTTCCGCCAG
        A  L  R  S  G  Q  S  G  F  A  G  M  A  G  N  N  V  L  P  P

8701  CAGGCGCCGGTCTGCAGGTTACCACGCGGCACGGTCACGGGATCGGTGCGCTATATGACG
        A  G  A  G  L  Q  V  T  T  R  H  G  H  G  I  G  A  L  Y  D

8761  TCTATGCGGGCGTCCACGACAACGAGTTATCCGAAGAGTTGATGGCATGGGGCCTGGCCA
        V  Y  A  G  V  H  D  N  E  L  S  E  E  L  M  A  W  G  L  A

8821  AAGAGCAGGTTCTTAAAGAAAAAATTGGCCCGATACGTGCCCGGCTTTATCGGTCACATT
        K  E  Q  V  L  K  E  K  I  G  P  I  R  A  R  L  Y  R  S  H

8881  TGAATGGAACCATCTTCCCCAATACCAGTTTCCTCACAGGATCGGGTGTATTTAAGGTCT
        L  N  G  T  I  F  P  N  T  S  F  L  T  G  S  G  V  F  K  V

8941  GGCAGCCACATGGGCCCAAAAAGACCGAGGTATTGACGTGGGCCATGGTCGAAAAAGACA
        W  Q  P  H  G  P  K  K  T  E  V  L  T  W  A  M  V  E  K  D

9001  TGTCTCCGGACATGAAAGAGAGAATTTCAAGGTCGGTGCAAAGGACGTTTGGAATCGCCG
        M  S  P  D  M  K  E  R  I  S  R  S  V  Q  R  T  F  G  I  A

9061  GTTATTGGGAAACTGAAGATAATGACAACATGGAGTCGGAAACCCACATGGGGCGAGGTT
        G  Y  W  E  T  E  D  N  D  N  M  E  S  E  T  H  M  G  R  G

9121  TCATGGCTGGCAAGAAAATGTTGAATTCCCAAATGGGGATCAGCGGGGACCGTCAGGATC
        F  M  A  G  K  K  M  L  N  S  Q  M  G  I  S  G  D  R  Q  D

9181  CGGATTATCCAGGTATCGTCGGTGATTCTGCTGTCGGAGAGACGTCGTATCGCGGCTACT
        P  D  Y  P  G  I  V  G  D  S  A  V  G  E  T  S  Y  R  G  Y

9241  ACCGCTTTTATAATGCCCTCATTCAATCATCAAATTGGAAAGAGGTTAATGCGCGTAGCA
        Y  R  F  Y  N  A  L  I  Q  S  S  N  W  K  E  V  N  A  R  S

9301  ATGACTGGGTGAATGAACTGATCACCGCGAGGCCCGAAAAATAAGGTCCGCTCCTTAGCG
        N  D  W  V  N  E  L  I  T  A  R  P  E  K  *

9361  GCCTTGCAATTCATCGAGATAAACAGACCCTGGAAATAAAGAGCGATGGAAACGAATCAA
                                        phnAd  M  E  T  N  Q

9421  GCAACTATGGAAACGACTCAAGCGCCGGTGAAAGCGGATCTGGCGAATGAGTATCCGGAA
        A  T  M  E  T  T  Q  A  P  V  K  A  D  L  A  N  E  Y  P  E

9481  CGTTATGTCAGTGAGGATCTTCGCAGAGAGGTCGAACGGTTTTTGTTCCGGGAAGCCCGA
        R  Y  V  S  E  D  L  R  R  E  V  E  R  F  L  F  R  E  A  R

9541  CTGATGGATACCGAGCAGTTTCGCGAATGGCATCAACAGATGCTGGATCCGGACATACGT
        L  M  D  T  E  Q  F  R  E  W  H  Q  Q  M  L  D  P  D  I  R

9601  TACGTTGTAACCAGTACCCAGCTTCGCTCCAGGAGAGAACGGCGGTATGCCCAACCTGAC
        Y  V  V  T  S  T  Q  L  R  S  R  R  E  R  R  Y  A  Q  P  D

9661  AAGGTTTTCATTTATGACGATGACCACAGACAGCTAGGCATTCGCGTGAACCAATATTAT
        K  V  F  I  Y  D  D  D  H  R  Q  L  G  I  R  V  N  Q  Y  Y

9721  GACCCCCAGCATTGGCGCATCGATCCCCCCGAAAAATACTGCCACACCGTTACCAACATC
        D  P  Q  H  W  R  I  D  P  P  E  K  Y  C  H  T  V  T  N  I

9781  GAGGCGTTCGAGAGCGACCGGGAGGACCAGATTAACGTCCGCAGCAATTGTTTGATTGTT
        E  A  F  E  S  D  R  E  D  Q  I  N  V  R  S  N  C  L  I  V

9841  CGTGCGCGGCGGTCCTATGAAGTGGATCAATTTTTCTACACACGAGAGGATGTGGTGCGA
        R  A  R  R  S  Y  E  V  D  Q  F  F  Y  T  R  E  D  V  V  R

9901  AGGAGCTCGGCGGGCACGTTGCGCTTGCTCAGTCGCGTAATTGACTATCCAGAGCGTTGT
        R  S  S  A  G  T  L  R  L  L  S  R  V  I  D  Y  P  E  R  C

9961  GTTCAAGGCCGTAATATGCTGATTTTTCTGTGAGGAGATGTTACGCGATCGGCGTGCAAC
        V  Q  G  R  N  M  L  I  F  L  *

10021 GCATGCGGCACGCCGCGAATAACATTAGTATACCGAATTATTGTGGGGGGATGTCATGGG
                                        phnB  M  G

10081 CTGGTTGAATAATACTGTCACTATCATTACCGGCGGCGGTTCGGGACTTGGAAAAGCGTT
        W  L  N  N  T  V  T  I  I  T  G  G  G  S  G  L  G  K  A  L

10141 GGTTGAGCGCTTCATCAACGAAGGGGGGCGGGTGGGCGTCCTTGAACGAAGTGGCGAGAG
        V  E  R  F  I  N  E  G  G  R  V  G  V  L  E  R  S  G  E  R

10201 GGCACGCGAGCTTGCGCGGGAGTTCGGAGACGCGGTCGAAGTTGTGGTCGGAGACGTGAC
        A  R  E  L  A  R  E  F  G  D  A  V  E  V  V  V  G  D  V  T

10261 CCTCTATGAGGACAACGTCCTGGTGGTGCAGAAAACTGTCGCGCGGTTTGGCCGGTTGGA
        L  Y  E  D  N  V  L  V  V  Q  K  T  V  A  R  F  G  R  L  D

10321 CAATTTCGTTGGTAACGCCGGGGTCTTTGATTTTTTCCAGACGTTGCCTCAAATGGACGC
        N  F  V  G  N  A  G  V  F  D  F  F  Q  T  L  P  Q  M  D  A
```

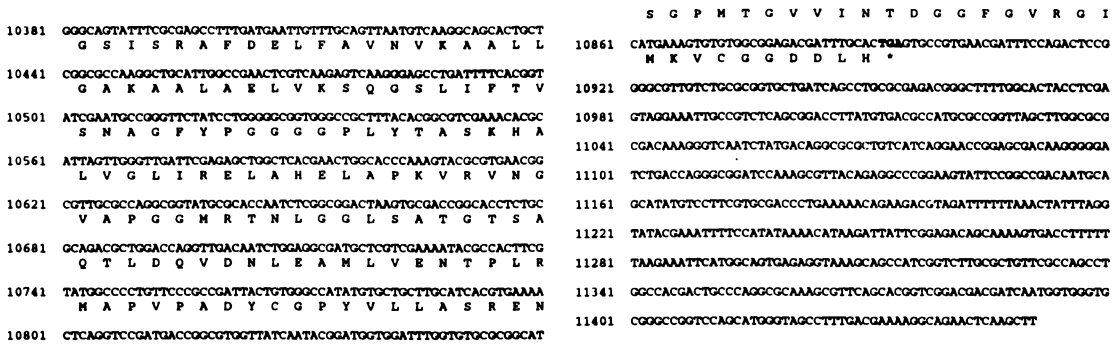


Figure 4.5 Complete nucleotide sequence of the pB1 fragment of *Burkholderia* sp. RP007, cds 1–11451, showing the location of the *phnRSFECDAcAdB* genes which were identified. Start and stop codons are in bold type, asterixes in the translated sequence indicate stop codons. Possible ribosomal binding sites (Shine-Dalgano sequences) are underlined. The sequence is shown 5'→3' as represented in Figure 4.1A; all genes are transcribed 5'→3' as shown here, except *phnR* which is transcribed in the opposite direction. A *Hin*dIII site (AAGCTT) is present at both ends of the fragment.

4.5.3 Identification of open reading frames (ORFs)

Open reading frames (ORFs) in the pB1 fragment were identified using an internet search tool provided by the Virtual Genome Centre. The programme was set to search submitted nucleotide sequences for ORFs in all six frames using the universal genetic code and to use ATG as a start codon. BLASTN and BLASTX searches of the nucleotide sequence provided approximate estimates of the location, identity and reading frame of genes on the pB1 fragment. BLASTP searches of the amino acid sequences of the predicted ORFs identified which of these were valid. Using the ClustalW and GeneDoc programmes the amino acid sequences of those ORFs which appeared to constitute valid genes were aligned with previously described analogous genes, usually of the *nah*-type, to confirm the validity of these ORFs.

Nine open reading frames were identified on the pB1 fragment, and based on the similarity of each ORF to previously described genes each was given a putative function; this is described further in section 4.7. The pB1 ORFs were designated the *phn* genes and their locations are shown in Figure 4.6A. Table 4.2 gives details of these ORFs including the amino acid similarity to the most homologous previously described gene, and Figure 4.8 shows the position where each of the catabolic genes acts in the degradative pathway of phenanthrene and naphthalene. Eight of the predicted pB1 ORFs (ORFs 2–9) are transcribed 5'→3' with respect to the sequence in Figure 4.5, and ORF 1 is transcribed in the opposite direction.

4.6 Expression of the *phn* genes in *E. coli*

4.6.1 Subcloning pB1 ORFs into pKK223-3

In order to express the hypothetical ORFs revealed by sequencing of pB1, and to confirm the predicted sizes of the peptide products, subclones containing the predicted ORFs were cloned into the expression vector pKK223-3. This vector contains the strong *tac* promoter which in *E. coli* JM105 is regulated by the *lac* repressor and induced by the addition of isopropyl-beta-D-thiogalactoside (IPTG) to the medium.

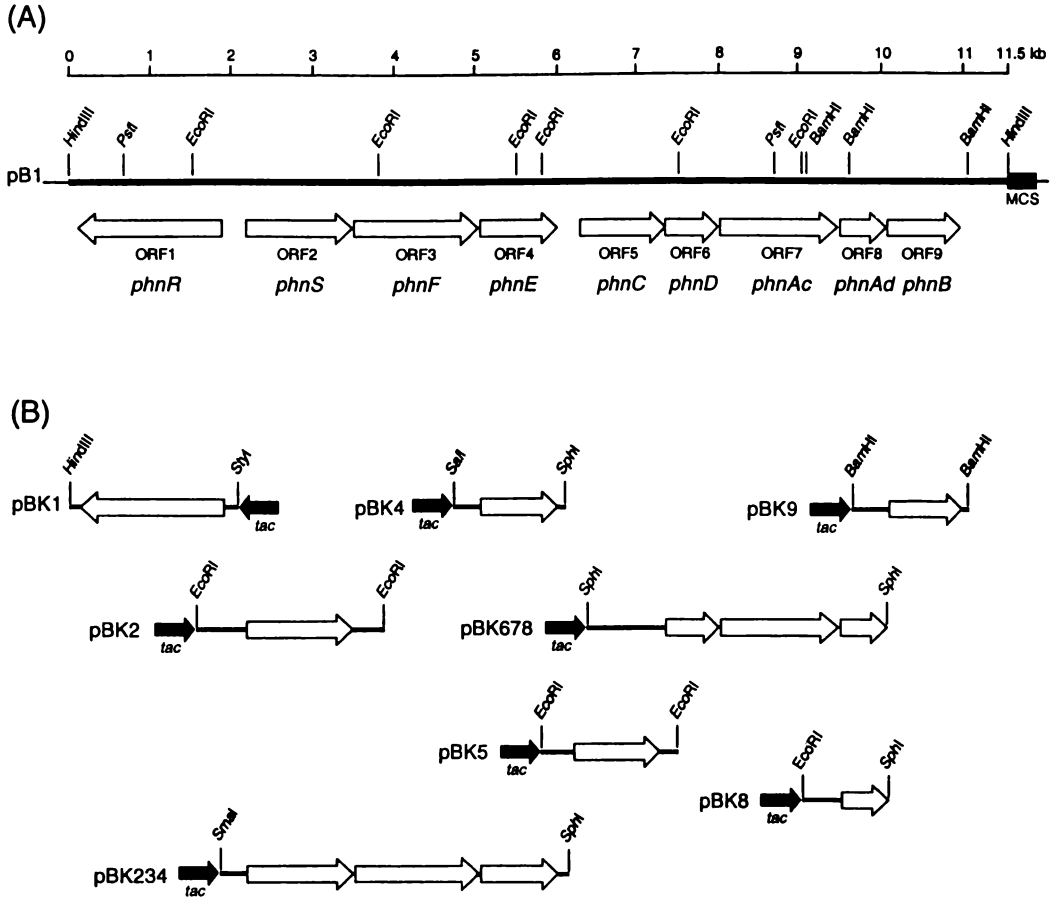


Figure 4.6 (A) The location and orientation of the nine open reading frames (ORFs) identified on the pB1 fragment of RP007. The gene designation of each ORF is also shown. (B) The subcloned derivatives of the pB1 fragment used for expression of the recombinant genes in *E. coli* JM105. These constructs were obtained by cloning regions of the pB1 fragment into the MCS of the expression vector pKK223-3. The orientation of the IPTG-inducible *tac* promoter is shown. Details of these subclones is given in Table B.3 (Appendix B).

pKK223-3 contains the pUC8 multiple cloning site, but since the *Bam*HI and *SalI*/*AccI*/*Hinc*II sites are not unique, only *Eco*RI, *SmaI*/*Xma*I, *Pst*I and *Hin*dIII sites are available for cloning inserts into pKK223-3. In addition, pKK223-3 does not contain the *lac*I'OPZ' genes for α-complementation, so X-gal screening of transformants to detect those with vectors containing inserts was not possible. These factors combined to make cloning of pB1 fragments into pKK223-3 a difficult and laborious procedure. Ligations were only successful when using two different restriction enzymes since the rate of religation of BAP-treated single-cut vector was unacceptably high. The orientation of the insert in pKK223-3 was critical for expression and it was usually necessary to first clone the desired fragment into

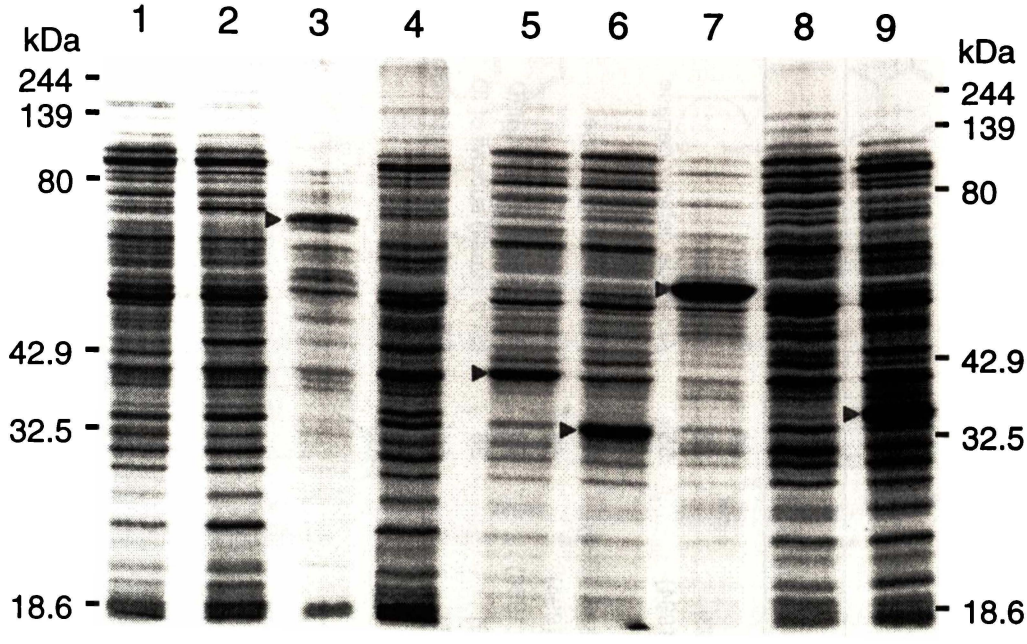


Figure 4.7 SDS-PAGE analysis of peptides encoded by *phn* genes expressed in *E. coli* JM105. The positions of molecular weight markers are shown, overexpressed proteins are marked with a small triangle. Samples were resolved using a 12% gel. Lanes are: 1, *E. coli*; 2, *E. coli* (pKK223-3); 3, *E. coli* (pBK1 (*phnR*)); 4, *E. coli* (pBK234 (*phnSFE*)); 5, *E. coli* (pBK4 (*phnE*)); 6, *E. coli* (pBK5 (*phnC*)); 7, *E. coli* (pBK678 (*phnDAcAd*)); 8, *E. coli* (pBK8 (*phnAd*)); 9, *E. coli* (pBK9 (*phnB*)).

pUC18, then transfer the insert from that construct, excised using a different set of restriction enzymes, into pKK223-3. In some cases the individual ORF could not be cloned into pKK223-3 alone so was present on a fragment which included adjacent ORFs. Table B.3 (Appendix B) and Figure 4.6B give full details of the pKK223-3 constructs used for the expression experiments, and the pUC18 intermediate clones used to obtain the final pKK223-3 constuct.

4.6.2 Analysis of expressed proteins by denaturing gel electrophoresis

Five recombinant peptides were successfully expressed in *E. coli* JM105 and visualised by SDS-PAGE analysis (Figure 4.7). The sizes as estimated by SDS-PAGE agree well with the predicted sizes (Table 4.2), although the experimentally determined molecular weight for *phnB* at 35 kDa was larger than the predicted size of 28.4 kDa. The remaining ORFs were not visualised, which may reflect a failure of *E. coli* to express the recombinant peptides from these subcloned derivatives.

ORF	Gene	Nucleotide position	No. of aa	Molecular mass (kDa) Predicted	Exptl[a]	Closest relative[b]	% aa identity	Protein feature[c]
1	*phnR*	155-1843	562	62.2	67	*phhR* (Ng *et al.*, 1995)	44	Regulatory
2	*phnS*	2377-3519	380	41.7	nd	*nodD* (Sousa *et al.*, 1993)	22	Regulatory
3	*phnF*	3529-5022	497	52.6	nd	*pahF* (Takizawa *et al.*, 1994)	65	Aldehyde dehydrogenase
4	*phnE*	5073-6065	330	36.5	39	*nahE* (Eaton, 1994)	73	Hydratase/Aldolase
5	*phnC*	6480-7307	275	30.0	32	*mpcI* (Kabisch & Fortnagel, 1990)	24	Extradiol dioxygenase
6	*phnD*	7339-7929	196	21.9	nd	*pahD* (Takizawa *et al.*, 1994)	48	Isomerase
7	*phnAc*	7992-9344	450	50.5	54	*pahAc* (Takizawa *et al.*, 1996)	56	ISPα (large) subunit of ID
8	*phnAd*	9406-9993	195	23.7	nd	*dntAd* (Suen *et al.*, 1996)	36	ISPβ (small) subunit of ID
9	*phnB*	10076-10894	272	28.4	35	*bphB* (Masai *et al.*, 1995)	55	Dihydrodiol dehydrogenase

[a] nd, over-expressed protein was not detected in *E. coli.*

[b] GenBank sequence showing most homology to each ORF, the reference is shown in parentheses.

[c] ID – PAH initial dioxygenase; ISP – iron sulfur-protein.

Table 4.2 Properties of the genes identified on the pB1 fragment.

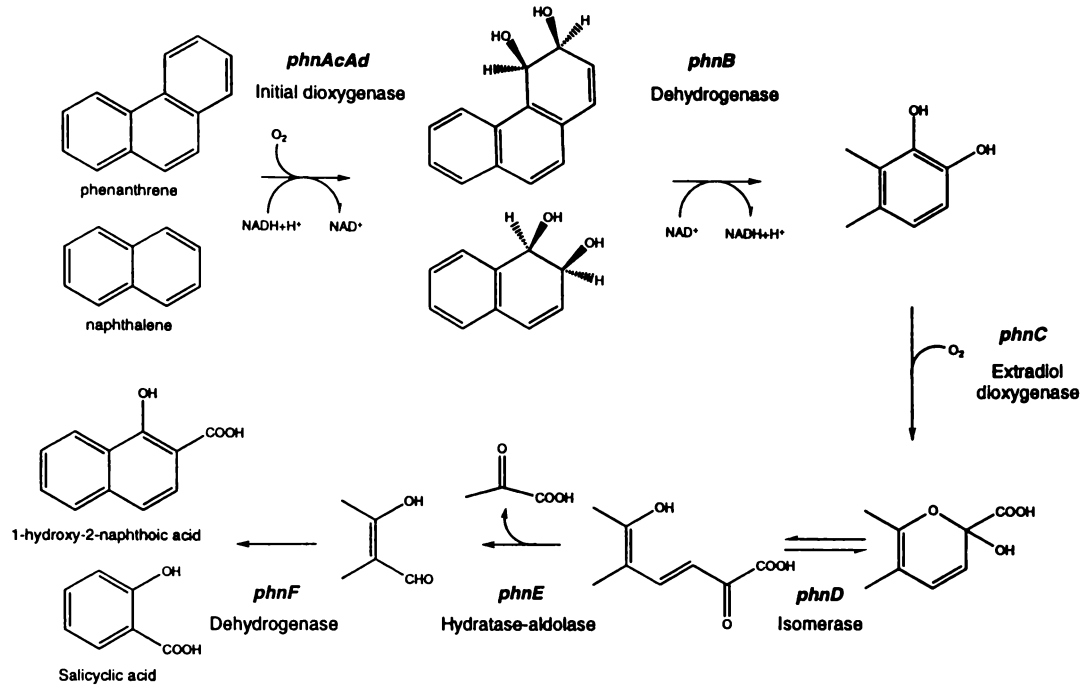


Figure 4.8 The putative functions of the *phn* catabolic genes identified from the nucleotide sequence of the pB1 fragment of *Burkholderia* sp. RP007.

4.7 Phylogenetic analysis of the *phn* genes

All nine genes identified on the pB1 fragment of *Burkholderia* sp. RP007 were designated the *phn* genes (for phenanthrene degradation). Although the genes from the phenanthrene degrading strains *P. aeruginosa* PaK1 (Takizawa *et al.*, 1996) and *P. putida* OUS82 (Takizawa *et al.*, 1994) were given the *pah* designation (for PAH degradation), it was felt the RP007 genes deserved a separate designation since they are significantly different to the *nah/pah/dox* genes. The individual gene designations for the *phn* genes, which denote the function of the transcribed peptide, is consistent with those of the *nah*-type genes.

The pB1 ORFs are discussed below in the order that the transcribed peptides act in the catabolic reactions against phenanthrene (Figure 4.8). In these descriptions of the individual ORFs the *phn* genes are compared to the isofunctional '*nah*-like' gene. The *nah*-like group includes members of the *nah* (Simon *et al.*, 1993; Eaton, 1994), *ndo* (Kurkela *et al.*, 1988), *dox* (Denome *et al.*, 1993) and *pah* (Kiyohara *et al.*, 1994; Takizawa *et al.*, 1994; Takizawa *et al.*, 1996) gene families, which share greater than 90% amino acid similarity. All identified ORFs were preceded by putative ribosomal binding sites (Shine & Dalgarno, 1975); these are shown in Figure 4.5.

4.7.1 *phnAc* – PAH initial dioxygenase ISPα subunit

ORF 7 encodes a 450 amino acid peptide which shows closest homology to the iron-sulfur protein large (α) subunit (ISPα) of naphthalene dioxygenase. This gene shows 54–56% amino acid sequence homology and 56–57% nucleotide sequence homology to the *nahAc*-like gene.

Iron-sulfur proteins are characterised by the presence of Rieske-type iron-sulfur centres which are [2Fe-2S] binding sites. These sites consist of cysteine and histidine residues arranged in a conserved motif ($CXHX_{16-17}CXXH$) (Rieske *et al.*, 1964). This motif is found in the PhnAc sequence between residues 81 and 105 and is shown in the ISPα amino acid alignment in Figure C.1 (Appendix C).

On the basis of the similarity in size and regions of amino acid sequence homology, particularly with respect to the iron-sulfur binding regions, ORF 7 is designated the *phnAc* gene encoding the ISPα subunit of a PAH initial dioxygenase.

Figure 4.9 shows the phylogenetic relationship of the PhnAc amino acid sequence to other described ISPα subunits for aromatic degradation. It is clear that PhnAc is part of the NahAc group, which also includes DntAc and NtdAc, but homology to the members of this group is significantly less (ca. 55%) than that shown between these members (>80%). Both PhnAc and NahAc-like sequences show ca. 30% amino acid homology to analogous biphenyl dioxygenase genes. The ISPα peptides from biphenyl/PCB pathways (Bph or Bpd) are localised within two phylogenetic groups which both include ISPα peptides from pathways for the degradation of single-ring aromatic compounds. A fourth less tightly clustered group consists of ISPα peptides from four other pathways for single-ring aromatic compound degradation. CarA1 (*Sphingomonas* sp. CB3) is the ISPα from a carbazole catabolic pathway and loosely groups with the Bph-type ISPαs (Shepherd & Lloyd-Jones, 1998). The *Pseudomonas* sp. CA10 CarAa peptide is also from a carbazole catabolic pathway and was selected as an outgroup due to its low amino acid homology (<15%) to all other ISPαs in the alignment.

Experiments using hybrid dioxygenases constructed with the genes encoding the NtdA (*Pseudomonas* sp. JS42) and DntA (*Burkholderia* sp. DNT) dioxygenases expressed in *E. coli*, demonstrated that the C-terminal region of the ISPα peptide was responsible for the substrate specificity of the dioxygenase (Parales *et al.*, 1998a; b).

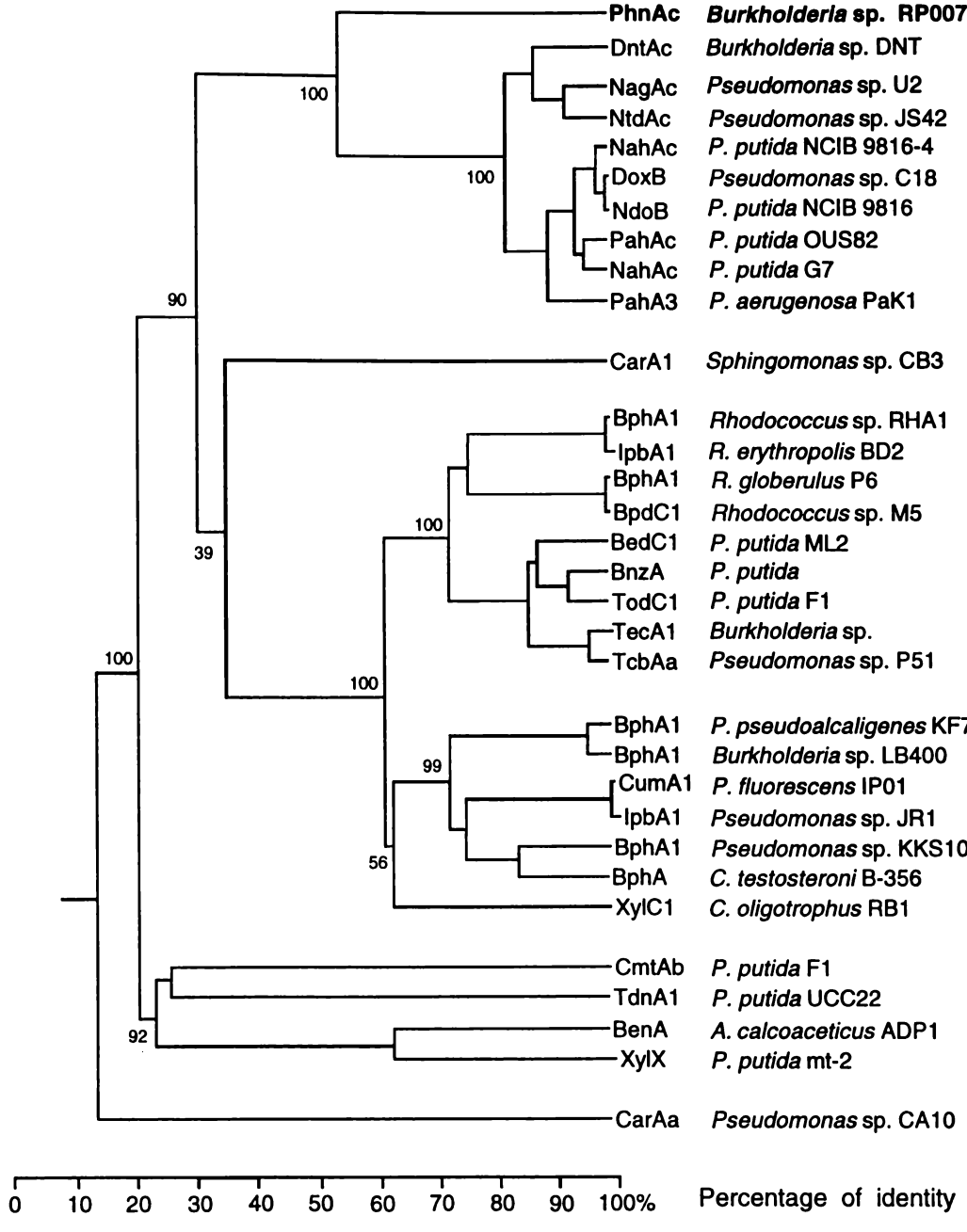


Figure 4.9 Dendrogram of predicted amino acid sequences of iron-sulfur protein large (α) subunits of aromatic compound dioxygenases which show homology to PhnAc of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using CarAa (*Pseudomonas* sp. CA10) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Figure A.1 (Appendix A).

An amino acid sequence alignment of ISPαs from naphthalene/phenanthrene, biphenyl/PCB and single-ring aromatic pathways reveals greater similarity in the N-terminal half of the peptide; this is the region where the Rieske-type iron-sulfur binding sites are located. The similarity in the C-terminal half of the ISPα sequences

across this group is lower, but amongst ISPαs that attack the same substrate the similarity is higher. Amino acid alignments of PhnAc and NahAc of *P. putida* G7 revealed similar homologies for the N-terminal and C-terminal halves of the peptide.

4.7.2 *phnAd* – PAH initial dioxygenase ISPβ subunit

ORF 8 encodes a 195 amino acid peptide and is located 62 bp downstream of the *phnAc* gene (ORF 7) and is transcribed in the same direction. This gene shows closest homology to *dntAd* encoding the iron-sulfur protein small (β) subunit (ISPβ) of dinitrotoluene dioxygenase (Suen *et al.*, 1996) and has been designated *phnAd*, encoding the ISPβ of PAH initial dioxygenase. PhnAd shows 28–33% amino acid homology to *nahAd*-like genes. Close analysis of the alignment of the PhnAd amino acid sequence with the NahAd-like sequences reveals a nonhomologous region for the first 40 amino acid residues; this region of PhnAd also shows little or no similarity to the corresponding regions of ISPβs from biphenyl/PCB pathways. The translation of the corresponding nucleotide sequence in the other two reading frames revealed no homology to the NahAd-type sequence ruling out any possibility of a translation frameshift due to a mistake in the nucleotide sequence.

The amino acid homology of PhnAd to NahAd-type is low, but nine amino acid residues which are invariant in the ISPβs of many other oxygenases (Neidle *et al.*, 1991) are conserved in the PhnAd sequence. The location of this gene directly downstream of the *phnAc* gene, the presence of a putative ribosomal binding site directly upstream of the ATG start codon, and the presence of residues which are highly conserved in other ISPβs, supports the designation of ORF8 as *phnAd* despite the low (ca. 30%) predicted amino acid similarity to *nahAd*-like sequences.

The dendrogram (Figure 4.10) of ISPβs from a number of aromatic compound initial dioxygenases shows a similar phylogenetic arrangement to the ISPα dendrogram. PhnAd groups with the NahAd-type peptides, but the similarity to members of this group is lower than that for PhnAc to the NahAc homologues. The biphenyl/PCB pathway ISPβs and the single aromatic ring ISPβs are grouped in a similar way to the corresponding ISPαs.

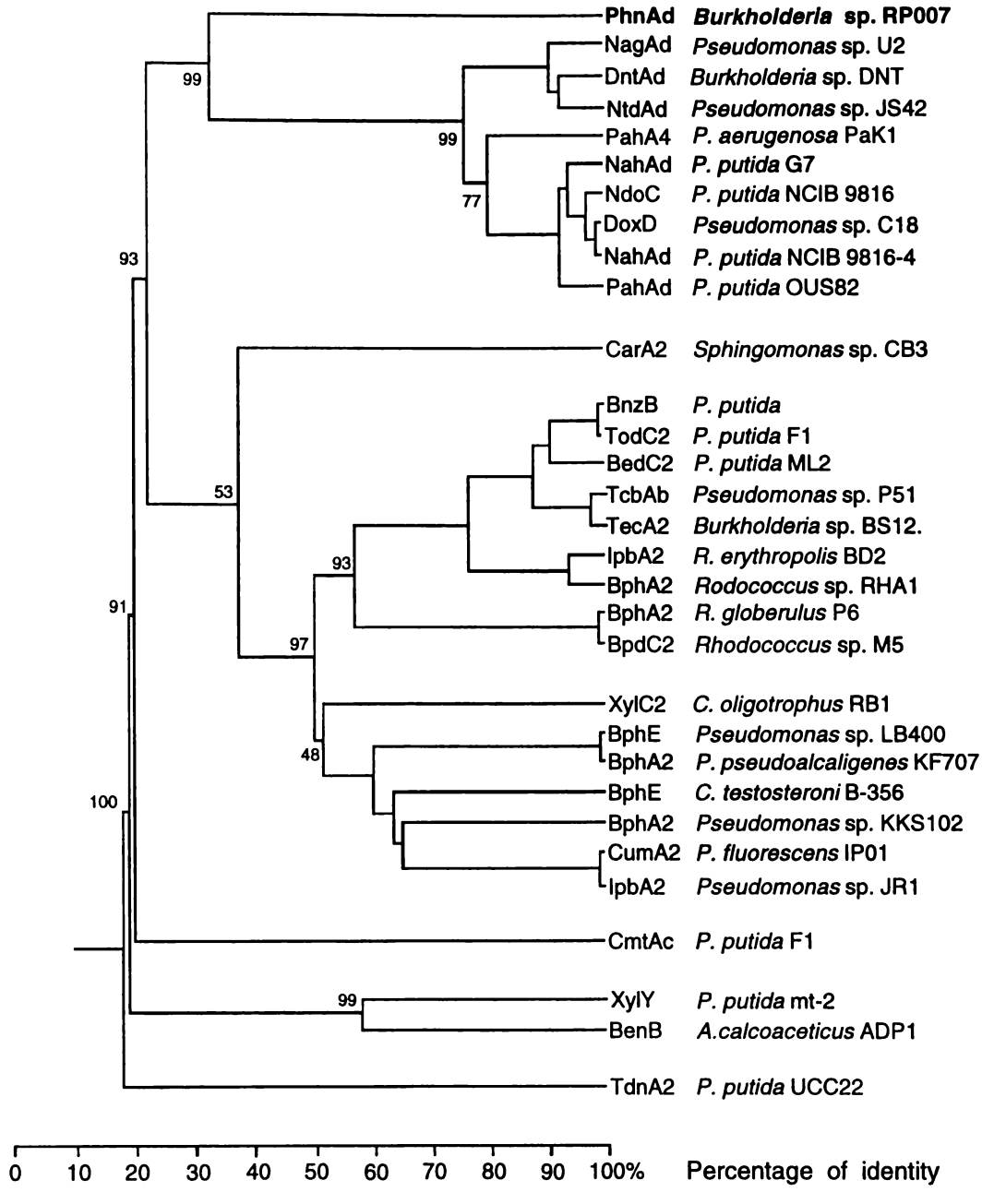


Figure 4.10 Dendrogram of predicted amino acid sequences of iron-sulfur protein small (β) subunits of aromatic compound initial dioxygenases which show homology to PhnAd of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using TdnA2 (*P. putida* UCC22) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

4.7.3 *phnB* – PAH dihydrodiol dehydrogenase

ORF 9 begins 82 nucleotides downstream from the stop codon of ORF 8 (*phnAd*) and encodes a putative PAH *cis*-dihydrodiol dehydrogenase; this gene is designated *phnB*. Only three dehydrogenase genes have been sequenced from *nah*-like operons to date: *pahB* (Takizawa *et al.*, 1994), *pahB* (Takizawa *et al.*, 1996), and *doxE* (Denome *et al.*, 1993). This group shares >85% amino acid homology and is referred to here as *nahB*-like. *phnB* shows greater similarity to the analogous dehydrogenase genes from biphenyl degrading strains than to the *nahB*-like sequence. *phnB* shows most homology to the *bphB* gene from *Rhodococcus* sp. RHA1 (Masai *et al.*, 1995) (56% amino acid homology) and shows 44–47% amino acid similarity to other *bphB*-type genes. The similarity to the *nahB*-like dehydrogenases is 38–40% at the amino acid level.

This phylogenetic arrangement is illustrated in the dendrogram in Figure 4.11. PhnB groups with the BphB-type dehydrogenases and the NahB-types form a separate and divergent cluster. Many of the relationships shown between biphenyl/PCB and single ring pathway ISPαs and ISPβs are conserved in the dehydrogenase arrangement.

A comparison of BphB dehydrogenase from *R. globerulus* P6 with 20 members of the family of short chain alcohol dehydrogenases, of which the P6 BphB is a member, revealed five strictly conserved residues (Asturias *et al.*, 1994). Both the PhnB and NahB-like enzymes also shared these residues; these are shown in the amino acid alignment of dehydrogenases in Figure C.3 (Appendix C). Furthermore, residues 8–36 in the P6 BphB enzyme are highly conserved and thought to be involved in binding of $NAD(P)^+$ cofactors; this region showed significant similarity to the corresponding PhnB region but was less similar in the NahB-like sequences.

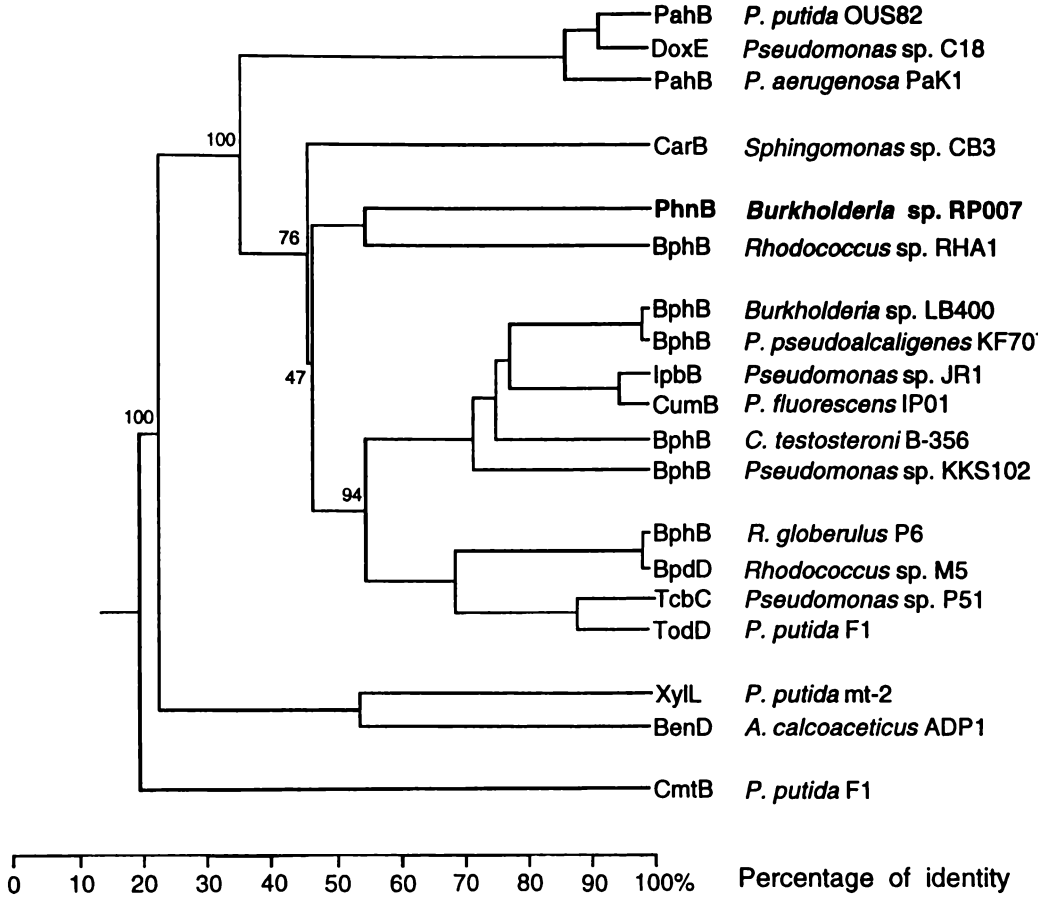


Figure 4.11 Dendrogram of predicted amino acid sequences of aromatic compound dehydrogenases which show homology to PhnB of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using CmtB (*P. putida* F1) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

4.7.4 *phnC* – PAH extradiol dioxygenase

ORF 5 encodes a 275 amino acid peptide and has been designated *phnC*, the extradiol dioxygenase of this PAH catabolic pathway. The predicted amino acid sequence of this gene shows no homology to NahC-like 1,2-dihydroxynaphthalene dioxygenases, which share >96% homology. PhnC also shows no homology to 2,3-dihydroxybiphenyl 1,2-dioxygenase typified by BphC of *Burkholderia* sp. LB400 (Erickson & Mondello, 1992; Han *et al.*, 1995). Six amino acid residues are strictly conserved amongst both 1,2-dihydroxybiphenyl dioxygenases and 1,2-dihydroxynaphthalene dioxygenases (His^{146}, His^{195}, His^{210}, His^{241}, Tyr^{250} and Glu^{260}, numbered according to the *R. globerulus* P6 BphCI enzyme) and are proposed to be among the ligands of the ferrous iron (Asturias *et al.*, 1994); these residues were not conserved in the PhnC peptide sequence (Figure C.4, Appendix C). PhnC also showed no amino acid similarity to the BphC2 and BphC3 enzymes from *R. globerulus* P6, which show significant homology to each other but no similarity to other BphC- or catechol 2,3-dioxygenase-types (Asturias *et al.*, 1994).

A BLASTP search of PhnC revealed homology to MhpB, a 2,3-dihydroxyphenylproprionate 1,2-dioxygenase from *E. coli* K12, MpcI, a 2,3-dihydroxyphenylproprionate 1,2-dioxygenase from *Alcaligenes eutrophus* (Kabisch & Fortnagel, 1990) and LigB, a protocatechuate 4,5-dioxygenase from *S. paucimobilis* (Noda *et al.*, 1990). Spence *et al.* (1996) suggested that this group of extradiol dioxygenases, characterised by MhpB and MpcI, and including LigB, represented a new class (Class III) of aromatic ring-cleavage dioxygenases. This group of dioxygenases shows no sequence similarity with the major family of ring cleavage enzymes.

The model proposed by Spence *et al.* (1996) suggests Class I extradiol dioxygenases, as typified by the small (21 kDa) *R. globerulus* P6 BphC2, is a single domain enzyme. The *Burkholderia* sp. LB400 BphC enzyme typifies the two-domain Class II extradiol dioxygenases in which a Class I enzyme has been duplicated followed by the loss of functionality of the N-terminal domain, and the C-terminal domain binds iron(II) and remains catalytically active. Extradiol dioxygenases of Class III are two domain enzymes typified by the *E. coli* MhpB enzyme, and evolved by the duplication of a Class I enzyme, followed by loss of function of the C-terminal

domain and the N-terminal domain remains catalytically active.

An alignment of the PhnC amino acid sequence with that of MpcI, MhpB and LigB shows 20–25% similarity. PhnC and MhpB share 59 (21%) identical residues. The PhnC enzyme is a 275 residue peptide which is smaller than MpcI (313 residues), MhpB (314) and LigB (302). Spence *et al.* (1996) noted four highly conserved histidine residues amongst five extradiol dioxygenases of this class and suggested these were involved in coordination of the non-heme iron cofactor. Experimental evidence confirmed that histidine residues were essential for functionality of MhpB and MpcI. The conserved histidine residues (His^{10}, His^{53}, His^{115}, and His^{179} positions in MhpB sequence) are all conserved in PhnC which supports grouping of PhnC with this class of enzymes; these are marked in the amino acid alignment in Figure C.5 (Appendix C).

The dendrogram in Figure 4.12 shows the position of the cluster of Class III extradiol dioxygenases, which includes PhnC, relative to the main group of Class II dioxygenases. The low homology between members of Class II is illustrated by the branch lengths within this subfamily. The NahC-like 1,2 dihydroxynaphthalene 1,2 dioxygenases occupy a separate cluster within the Class II group, which also includes BphC from *S. yanoikuyae* Q1 and *S. yanoikuyae* B1. The Q1 BphC dioxygenase has a broad substrate specificity with 70-fold more activity towards 1,2 dihydroxynaphthalene than 2,3-dihydroxybiphenyl (Kuhm *et al.*, 1991a). The activity of the B1 BphC dioxygenase towards naphthalene was not measured (Kim & Zylstra, 1995) but since this strain degrades phenanthrene and anthracene it probably has a similarly broad substrate specificity to the Q1 BphC. This may explain why the B1 and Q1 BphC dioxygenases are more closely related to NahC-like dioxygenases than to other BphC-types.

Spence *et al.* (1996) also noted some similarity between N-terminal and C-terminal domains of MhpB and MpcI which supports the proposition that these enzymes evolved through the duplication of a single-domain enzyme. A similar alignment of the N- and C-terminal domains of PhnC showed some similarity but significantly less than that observed for MhpB and MpcI (not shown).

Phylogenetic analysis of known extradiol dioxygenases reveals that enzymes for cleavage of aromatic compounds containing two rings show significantly less similarity to catechol 2,3-dioxygenases and similar dioxygenases which cleave

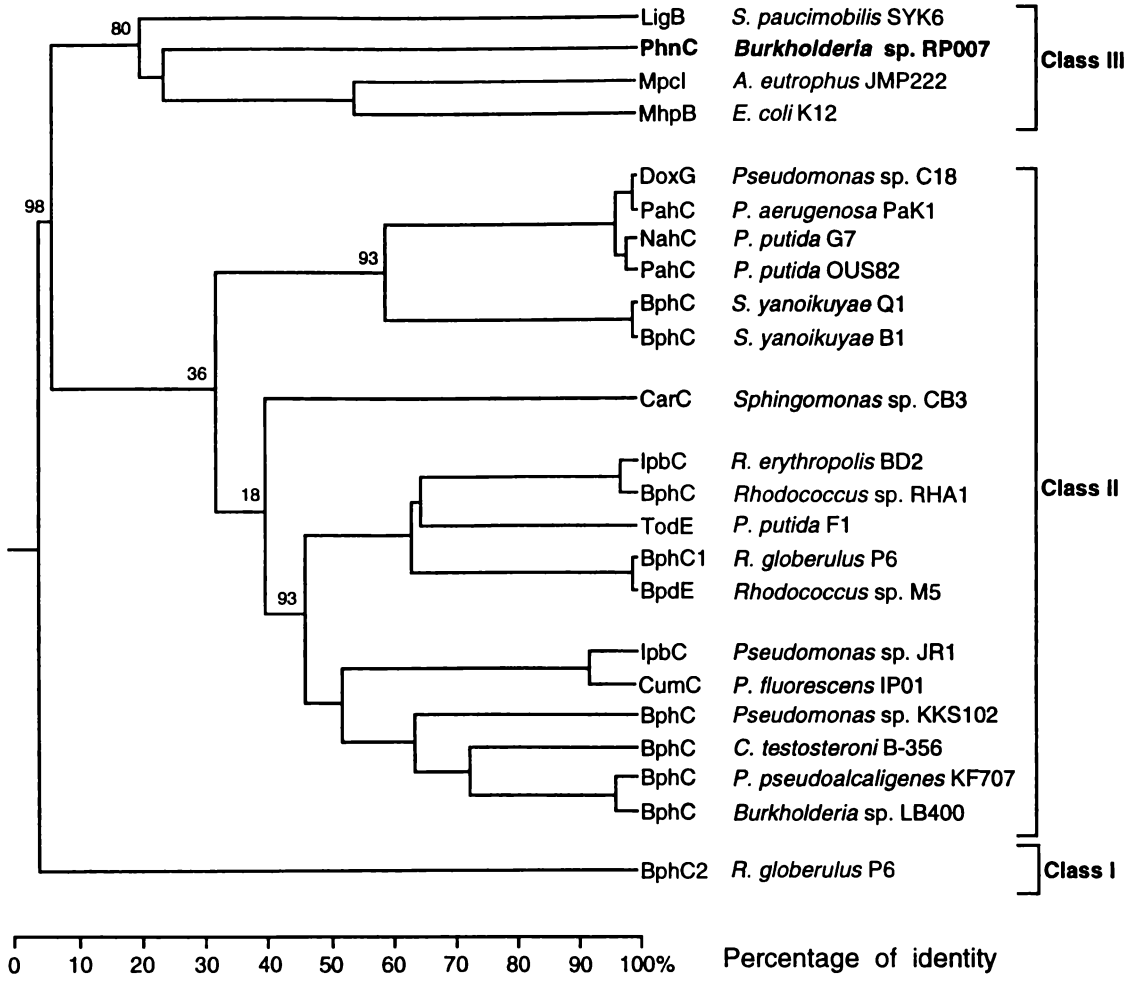


Figure 4.12 Dendrogram of predicted amino acid sequences of extradiol dioxygenases showing PhnC is part of the Class III extradiol dioxygenases. This dendrogram is an unrooted tree rerooted using BphC2 (*R. globerulus* P6) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

single-ring aromatic substrates. Harayama & Rekik (1989) thus grouped the extradiol dioxygenases characterised at that time into two different families comprising dioxygenases for monocyclic or bicyclic compounds. A representation of these two families of extradiol dioxygenases is shown in a dendrogram by Kim & Zylstra (1995) which includes many other BphC- and C23O-type enzymes since characterised.

The classification of PhnC is not consistent with the grouping of extradiol dioxygenases as postulated by Harayama & Rekik (1989) since it forms part of an operon which encodes the degradation of naphthalene and phenanthrene, yet shows closest similarity to dioxygenases for the cleavage of substituted catechol compounds.

Substrate[a]	Specific activity[b] (μmol/min/mg protein)	Relative activity[c] (%)
Catechol	6.9±2.2	1.2
3-Methylcatechol	31±6.8	5.5
4-Methylcatechol	41±12	7.3
2,3-Dihydroxybiphenyl	114±17	20.2
1,2-Dihydroxynaphthalene	564±158	100

[a] Activity toward 3,4-dihydroxybiphenyl was also observed but the reaction rate was not determined since no molar extinction coefficient (ε) for this substrate was available.

[b] Each value is the mean from three different cultures, showing standard deviation.

[c] Relative activities are expressed as percentages of the activity with 1,2-dihydroxynaphthalene.

Table 4.3 Enzyme activities of the PhnC extradiol dioxygenase with different substrates. Values represent activities of a cell extract of *E. coli* (pBK5).

4.7.4.1 Extradiol activity of PhnC

Since the identification of ORF5 as the *phnC* gene encoding the extradiol dioxygenase of the *phn* PAH catabolic pathway was based on low amino acid sequence homology (ca. 20%) to the MpcI/MphB/LigB (ClassIII) extradiol dioxygenases, it was necessary to show empirically that the *phnC* gene did code for this enzyme. This was achieved by expressing the recombinant *phnC* gene in *E. coli* and demonstrating *meta* cleavage activity in the cell extract towards 1,2-dihydroxynaphthalene.

The recombinant construct pBK5 is a 1.7 kb *Eco*RI fragment of pB1, which includes the complete *phnC* gene (ORF5), cloned into the expression vector pKK223-3 (Figure 4.6B). The specific and relative activities of the *phnC* extradiol dioxygenase are shown in Table 4.3. The clear preference for 1,2-dihydroxynaphthalene indicates PhnC is a PAH extradiol dioxygenase.

4.7.5 *phnD* – isomerase

ORF 6 is identified as the *phnD* gene which encodes a putative isomerase. In the naphthalene catabolic pathway this enzyme catalyses the conversion of 2-hydroxychromene-2-carboxylate to *trans-o*-hydroxybenzylidenepyruvate, which is the substrate for cleavage by a hydratase-aldolase to yield pyruvate and salicylaldehyde (Eaton & Chapman, 1992). Since these steps are unique to the oxidation of aromatic compounds of two or more rings (not including biphenyl), there exist fewer previously characterised genes to which new sequences genes can be compared. Previously described 2-hydroxychromene-2-carboxylate isomerase genes are *nahD* from the NAH7 plasmid of *P. putida* G7 (Eaton, 1994), *doxJ* from *Pseudomonas* sp. C18 (Denome *et al.*, 1993), *pahD* from *P. putida* OUS82, and *pahD* from *P. aeruginosa* PaK1 (Takizawa *et al.*, 1996). Members of this NahD-like group are highly similar; *doxJ* and *pahD* (PaK1) are 100% identical at the nucleotide level (the *doxJ* sequence as given by the GenBank file was adjusted since the authors had not recognised a GTG start codon which had been used for the *nahD* and *pahD* genes). Homology amongst all four members of this group is >80%. *doxJ* and *nahD* show 91% amino acid homology up to a 35 bp deletion in *nahD* which occurs at a position corresponding to 25 bp before the end of the *doxJ* gene; consequently there was no homology to *doxJ* after this point and *nahD* is four amino acid residues longer than *doxJ* since the stop codon was part of the deletion (Eaton, 1994). *pahD* (OUS82) also has this 35 bp deletion.

As shown in Figure 4.13 the predicted amino acid sequence of PhnD from RP007 is 47% similar to that of the NahD-like isomerases. The high similarity of the PahD(PaK1)/DoxJ and PahD(OUS82)/NahD pairs is illustrated, and the reduced homology between these pairs which is a consequence of the 35 bp deletion. That both *pahD*(OUS82) and *nahD* share this deletion suggests these genes have recently evolved from a common ancestor.

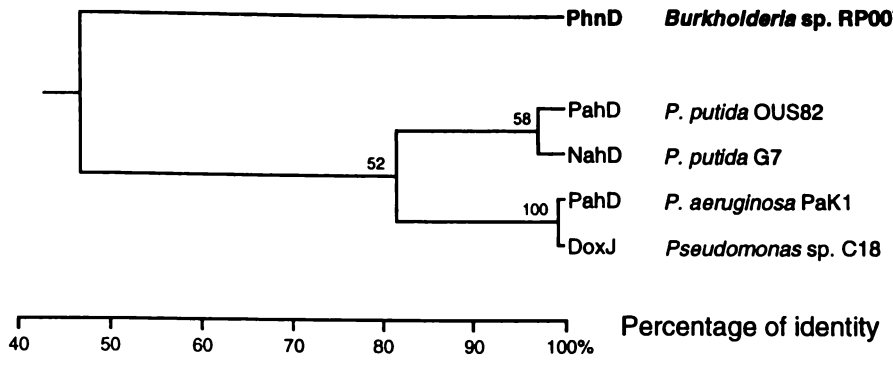


Figure 4.13 Dendrogram of predicted amino acid sequences of 2-hydroxychromene-2-carboxylate isomerases which show homology to PhnD of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using PhnD as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

4.7.6 *phnE* – hydratase-aldolase

ORF 4 is a putative hydratase-aldolase gene designated *phnE*. The predicted amino acid sequence of *phnE* is 73% similar to the four highly homologous NahE-like *trans-o*-hydroxy-benzylidenepyruvate hydratase-aldolase peptides (NahE (Eaton, 1994), PahE (OUS82) (Takizawa *et al.*, 1994), PahE (PaK1) (Takizawa *et al.*, 1996) and DoxI (Denome *et al.*, 1993). These NahE-like enzymes share >94% amino acid homology, and PahE (PaK1) and DoxI are identical. Of all the *phn* genes present on the pB1 fragment, *phnE* is the most similar to the corresponding *nah*-like gene.

The region of PhnE which is least similar to the NahE-like enzyme is in the eight amino acids at the N-terminal of the peptide; Eaton (1994) suggested these residues were not required for enzyme activity since other workers showed hydratase-aldolase activity by an *E. coli* clone containing a truncated *nahE* gene which lacked the eight N-terminal amino acids.

A dendrogram showing the phylogenetic position of PhnE relative to NahE-like peptides is not included here as the theme is very similar to the PhnD tree; that is, a tight cluster of Nah-type peptides with the Phn enzyme forming a separate branch. An alignment of the predicted amino acid sequence PhnE and the NahE-like peptides is shown in Figure C.7 (Appendix C).

4.7.7 *phnF* – dehydrogenase

ORF 3 is a putative dehydrogenase gene, *phnF*, which has a predicted amino acid sequence that is 65% similar to NahF-like salicylate dehydrogenases (PahF (OUS82) (Takizawa *et al.*, 1994), PahF (PaK1) (Takizawa *et al.*, 1996), and DoxF (Denome *et al.*, 1993). The NahF-like dehydrogenases share >91% amino acid homology. The predicted amino acid alignment for PhnE and NahE-like peptides is shown in Figure C.8 (Appendix C).

4.7.8 *phnS* – LysR-type transcriptional regulator

The predicted amino acid sequence of ORF 2 shows low similarity to a number of LysR-type transcriptional regulators (LTTRs). ORF2 is designated *phnS* and is a putative transcriptional activator involved in the regulation of the *phn* pathway or lower pathway genes. As shown in the dendrogram in Figure 4.14 the amino acid sequence of PhnS shows most homology to a number of LTTRs, designated NodD, involved in regulation of nodulation proteins in rhizosphere-dwelling bacteria. PhnS is also 21% similar to NahR, the LTTR involved in regulation of the *nah* and *sal* promoters of the NAH7 plasmid of *P. putida* G7 (Schell & Wender, 1986). The dendrogram in Figure 4.14 also shows other LTTRs involved in the regulation of aromatic degradation pathways. LTTRs are a loosely homologous group as illustrated by the branch lengths in this tree. Indeed, the criteria for identifying an LTTR is the presence of >20% amino acid homology with another LTTR family member, or high identity to a consensus sequence for the highly conserved amino terminus (Schell, 1993).

The most highly conserved region of all LTTRs is at residues 1–65, numbered according to the NahR amino acid sequence (Schell, 1993). The central portion of this region (residues 23–42) is a helix-turn-helix (HTH) motif which is involved in DNA binding and is 40% conserved in all LTTRs. PhnS shows high identity in this region, except at positions 31 and 35 where Gly and Pro, respectively, differ from classic HTHs (Figures 4.15 and C.9 (Appendix C); divergence at these positions is common among LTTRs (Schell, 1993). Of the seven most highly conserved residues in this region which are identical in 70% of LTTRs, (Ala^{27}, $Thr(Ser)^{33}$, Gln^{34}, Pro^{35}, $Ser(Thr)^{38}$, Leu^{44}, and Glu^{45}) all are conserved in PhnS except Pro^{35} (Ser). The amino acid homology between LTTRs within this region was considerably higher than that shown across the whole sequence, and revealed a different phylogenetic relationship than that shown in Figure 4.14. Across this region PhnS is more similar to NahR (40%) than the NodD LTTRs (30–35%).

A conserved C-terminal domain (residues 227–253) is also important for DNA interactions and transcriptional activation and may play a role in multimerisation of LTTR units or participate in the DNA-protein binding (Schell, 1993). In this region 13 residues were conserved with respect to NahR.

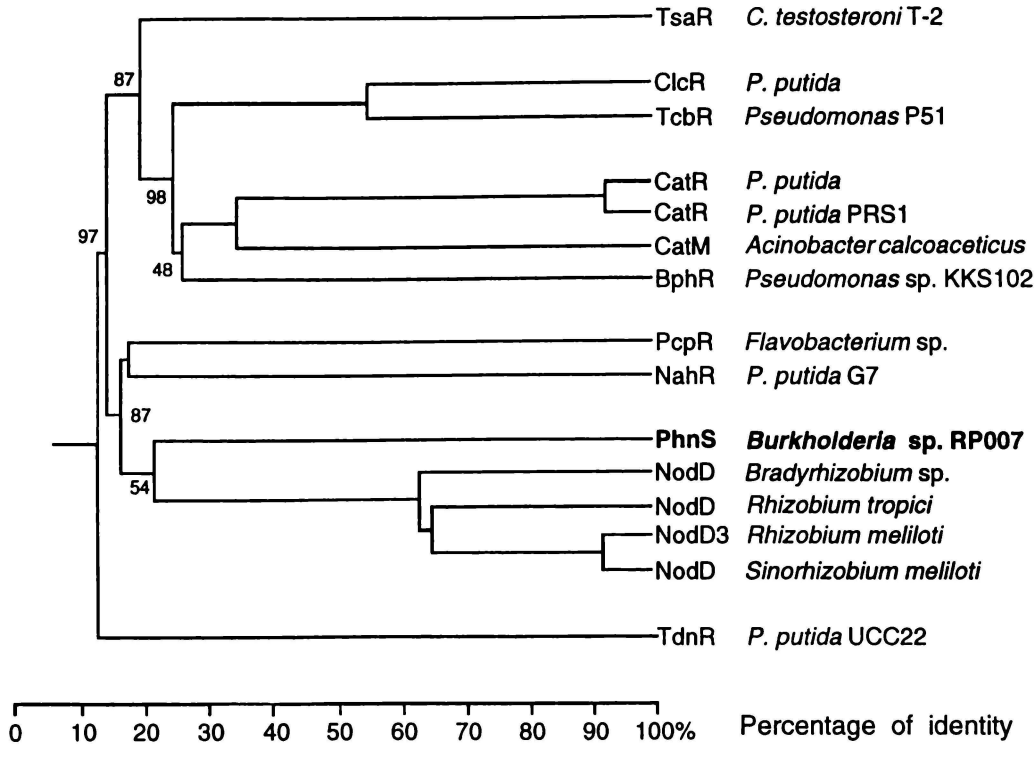


Figure 4.14 Dendrogram of predicted amino acid sequences of LysR-type transcriptional regulators (LTTRs) which show homology to PhnS of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using TdnR (*P. putida* UCC22) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

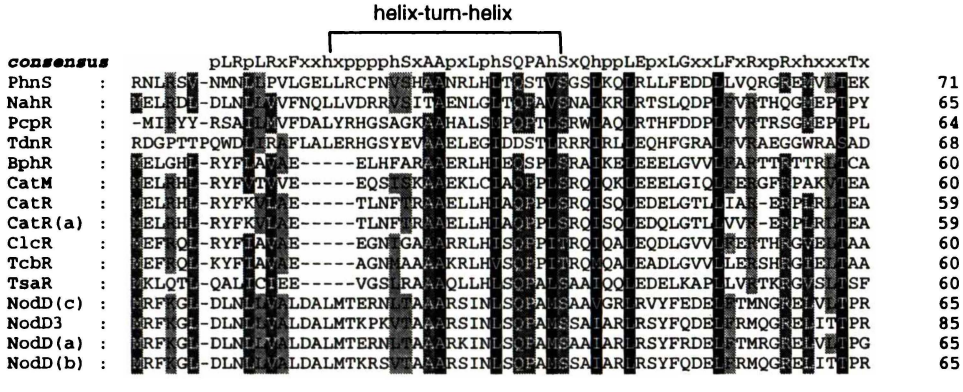


Figure 4.15 Alignment of LysR-type transcriptional regulators (LTTRs) of the region corresponding to residues 1–65 of the NahR amino acid sequence. The 20 residues which form the helix-turn-helix motif are shown. Shading indicates positions where amino acid identity is conserved (black – 100% , grey – 80%, light grey – 60%). The consensus sequence shows residues that are conserved in 70% of characterised LTTRs (Schell, 1993); lower case letters are: x, any residue; h, hydrophobic residue (V, I, L, M); p, hydrophilic residue (T, S, N, Q, D, E, K, R, H).

Amino acid homology across the region corresponding to 95–173 residues of the NahR peptide, which is the domain involved in coinducer recognition/response, was not higher between PhnS and NahR (18%) than the overall homology. Greater similarity across this region might be expected since it is possible both peptides recognise a similar coinducer. In this region there was greater homology between the NodD LTTRs (75–92%) than the overall homology (62–91%) and this may reflect similar coinducers for these peptides. Increased homology in this region was also seen amongst some of the LTTRs implicated in catechol and chlorocatechol metabolism.

4.7.9 *phnR* – NtrC-type transcriptional activator

ORF 1 is a large peptide of 562 amino acids and is divergently transcribed with respect to the other eight *phn* genes. The predicted amino acid sequence of this gene, designated *phnR*, shows high homology to a group of positive transcriptional regulators which form the NtrC family, so called because of similarity to the *Klebsiella pneumoniae* NtrC and NifA proteins, both of which are transcriptional regulators for the *ntr* and *nif* genes for nitrogen metabolism (Buikema *et al.*, 1985). Members of the NtrC family of transcriptional regulators control a variety of physiological processes in response to environmental signals, and are characterised by the dependence on RNA polymerase that utilises the alternative σ^{54} cofactor; they are therefore referred to as σ^{54}-dependent regulators (Shingler, 1996).

PhnR shows >40% amino acid similarity to a number of σ^{54}-dependent regulators involved in regulation of aromatic degradative pathways. The dendrogram in Figure 4.16 shows PhnR is related to a group of six σ^{54}-dependent regulators which share >65% homology. σ^{54}-dependent regulators are modular proteins consisting of four functional domains. The central C-domain possesses an NTP binding site which mediates the ATPase activity of this domain. This region is thought to be involved in the association with the promoter-bound σ^{54}-RNA polymerase holoenzyme complex and is highly conserved amongst regulators of this class. Across this region PhnR shows high similarity to other σ^{54}-dependent regulators and the motifs (GXXGXK and QXXLLRVL) implicated in ATP binding and hydrolysis (Shingler, 1996) are present. The C-terminal D-domain of σ^{54}-dependent regulators contains a helix-turn-helix motif (AXXXXXXXXXXAAXXLG) involved in DNA binding which is also conserved in the PhnR sequence. These features are marked in the amino acid sequence alignment of PhnR with other σ^{54}-dependent regulators (Figure C.10, Appendix C).

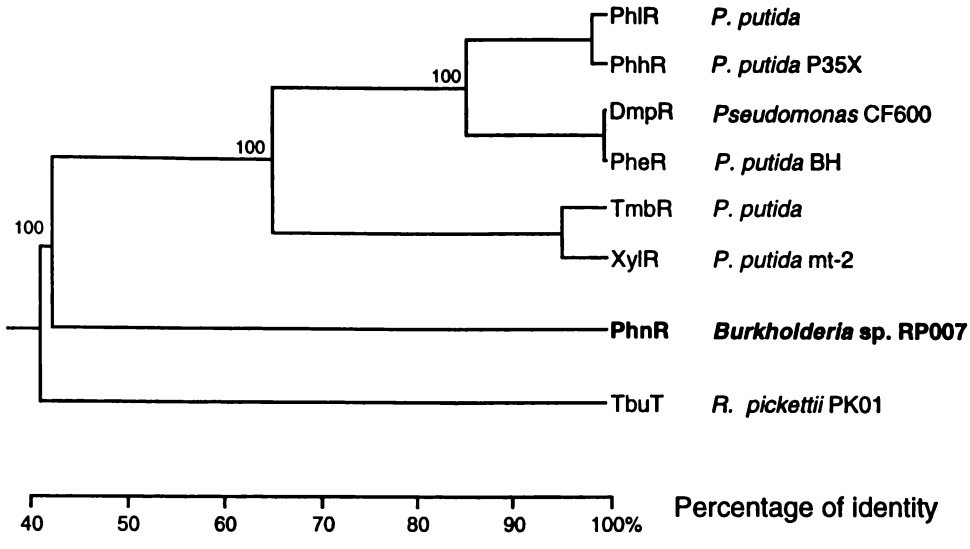


Figure 4.16 Dendrogram of predicted amino acid sequences of σ^{54}-dependent transcriptional regulators involved in regulation of aromatic catabolic pathways which show homology to PhnR of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using TbuT (*Ralstonia pickettii* PK01) as an outgroup. The scale shows the percentage amino acid similarity between the sequences, and the bootstrap values of the major branch points are shown. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Table A.1 (Appendix A).

4.8 Detection of *phn* mRNA by RT-PCR

4.8.1 *phn* genes are not constitutively expressed

To show that expression of the *phn* genes is associated with catabolism of phenanthrene and naphthalene by RP007, evidence was sought for induction of the *phn* enzymes during growth on these substrates. Reverse transcriptase-polymerase chain reaction (RT-PCR) was chosen as a means of detecting the mRNA transcript after lack of success using Northern blotting techniques.

RP007 was grown to late-log phase in liquid minimal media supplemented with acetate (5 mM), phenanthrene (0.05% wt vol^{-1}), or naphthalene (0.05% wt vol^{-1}). Cell pellets from these cultures were used for the RT-PCR experiment as described in Chapter 2.

For detection of *phn* genes two sets of primers were designed (Table 4.4, Figure 4.17). The P6897/P8420 primer set spans three ORFs (*phnCDAc*) and amplifies a 1524 bp fragment. The P7525/P8197 primer set is internal to the P6897/P8420 set and amplifies a 673 bp fragment spanning *phnDAc*. The internal primers were designed due to initial lack of success using the P6897/P8420 set and it was considered a primer set amplifying a smaller product might be more successful. Also the internal primers could be used to confirm the identity of an RT-PCR product derived using the P6897/P8420 primer pair.

The ability of both sets of primers to amplify products of the expected size from RP007 genomic DNA was first confirmed by PCR. Use of a range of $MgCl_2$ concentrations showed both primer sets gave a most intense product at 1.5 mM, therefore the RT-PCR protocol did not require adjustment.

Total RNA extracted from cell pellets derived from phenanthrene and naphthalene cultures consistently yielded RT-PCR products of the expected size from both primer sets (Figure 4.18). Both positive and negative control tubes for each set of primers and template RNA gave expected results. PCR using the internal primers against the RT-PCR product obtained using the P6897/P8420 primer set gave a product of the expected size confirming the identity of the 1.5 kb product.

RNA extracted from acetate-grown RP007 cells consistently failed to yield an RT-PCR product using either of the primer sets. All positive and negative controls using

Primer	Primer sequence	Gene (nucleotide position)[a]
P7525	5'-GTCGTGGAGGATCTTAAGCG-3'	*phnD* (7525-7544)
P8197	5'-CGCATCACAATCACCTCATC-3'	*phnAc* (8178-8197)
P6897	5'-GCGATTCCGGTTTATCTCAA-3'	*phnC* (6897-6916)
P8420	5'-CTCCACCTTGCCAATTTCAT-3'	*phnAc* (8401-8420)
P4940	5'-CCGAACTAAAGTGGATCACGA-3'	*phnF* (4970-4990)
P6510	5'-GATCAGTGGATCGTGGGG-3'	*phnC* (6510-6527)
P3076	5'-CAGCGCATACAGTTCCTGGT-3'	*phnS* (3076-3095)
P3953	5'-GGTCATCGACAACACACCTG-3'	*phnF* (3953-3972)

[a] Nucleotide position shows the location of the primer site with respect to the pB1 sequence (Figure 4.5).

Table 4.4 Primers used for *Burkholderia* sp. RP007 RT-PCR experiments described in the text.

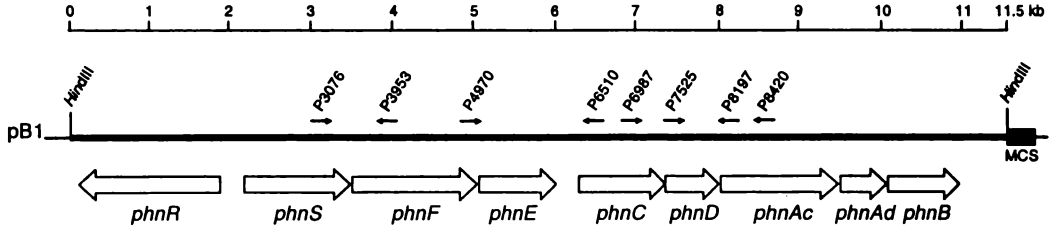


Figure 4.17 Location of the primers described in Table 4.4 relative to the pB1 fragment. The small arrows (not to scale) show the direction and location of each primer.

this template gave expected results confirming the validity of the negative result. These results indicate that the *phn* genes of RP007 are expressed only during growth on aromatic substrates and are therefore regulated at the transcriptional level. This is expected since transcriptional control of other aromatic catabolic operons has been shown (Schell, 1985; Ng *et al.*, 1995; Gallegos *et al.*, 1997; Delgado & Ramos, 1994). Since these results show transcription of the *phn* genes is under regulatory control, they support the hypothesis that the products of the *phnS* and/or *phnR* genes are involved in regulation of the *phn* genes.

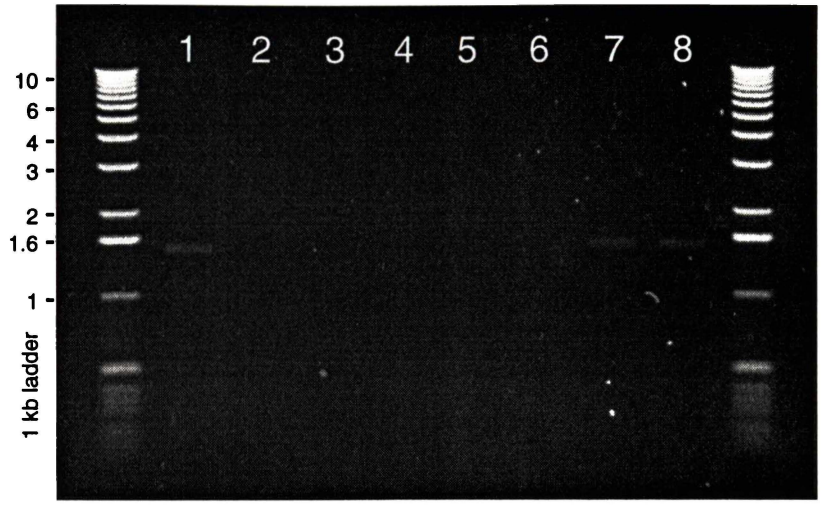


Figure 4.18 RT-PCR amplification of a 1.5 kb fragment from total RNA of *Burkholderia* sp. RP007 using the P6897/P8420 primers. Lanes 6, 7 and 8 show the RT-PCR experiment for acetate-, naphthalene-, and phenanthrene-grown RP007 cells, respectively; a product is only observed in lanes 7 and 8. Other lanes show PCR controls with the following templates: 1, RP007 genomic DNA; 2, no template DNA; 3, total RNA from acetate-grown cells; 4, total RNA from naphthalene-grown cells; 5, total RNA from phenanthrene-grown cells.

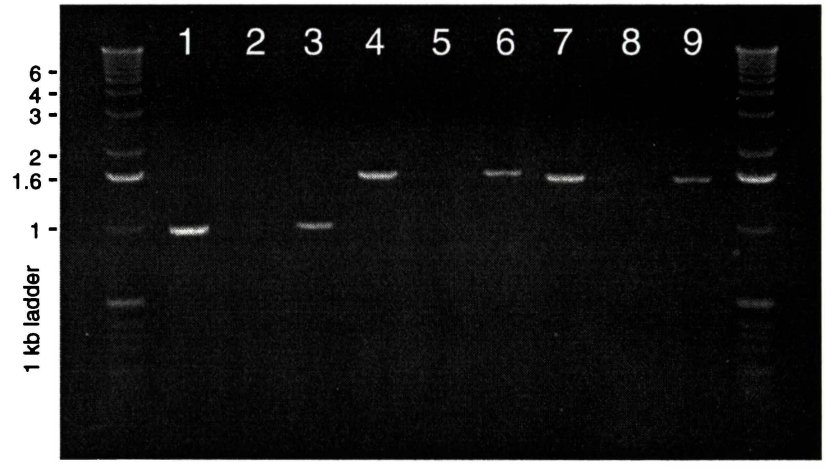


Figure 4.19 RT-PCR amplification of regions of the *phn* locus of *Burkholderia* sp. RP007 showing the *phnSFECDAc* genes are expressed on a single mRNA transcript. Lanes 3, 6, and 9 show RT-PCR products using the P3078/P3953, P4970/P6510, and P6987/P8420 primer sets, respectively. Other lanes: 1, P3078/P3953 positive PCR control; 2, P3078/P3953 negative PCR control; P4970/P6510 postive PCR control; P4970/P6510 negative PCR control; P6987/P8420 postive PCR control; P6987/P8420 negative PCR control (positive PCR controls contained RP007 genomic DNA, negative PCR controls contained the RNA prep from naphthalene-grown RP007 as used for the RT-PCR experiments).

4.8.2 The *phnSFECDAc* genes are expressed on a single transcript

Arrangement of the *phnSFECDAcAdB* genes on the pB1 fragment suggests they constitute an operon and are therefore expressed on a single mRNA transcript. The experiment described in section 4.7.1 demonstrated that the *phnCDAc* genes are expressed on the same transcript since RT-PCR using the P6897/P8420 primer set amplified a 1.5 kb fragment spanning across these genes.

The close physical arrangement of *phnSFE* suggests these genes are transcribed together; *phnCDAcAbB* are also closely arranged suggesting they too are transcribed on a single transcript. However, the large (414 bp) intergenic space between *phnE* and *phnC* suggests that *phnSFE* and *phnCDAcAbB* may be transcribed separately.

Primers were designed which spanned *phnFEC* including the intergenic space between *phnE* and *phnC* (P4970/P6510), and which spanned *phnSF* (P3076/P3953) (Table 4.4, Figure 4.17). Using both primer sets products of the expected size were amplified from total RNA extracts of naphthalene-grown RP007 cells using RT-PCR (Figure 4.19). All positive and negative controls for the RT-PCR experiments gave expected results.

The results of this experiment, in concert with the results described in section 4.7.1 for the P6897/P8420 primer set, demonstrate that the *phnSFECDAc* genes are transcribed on a single mRNA transcript. The regions amplified by these primers did not include *phnAd* and *phnB*, but since both *phnAc* and *phnAd* encode subunits of the same enzyme (ISP) they must be cotranscribed. The intergenic space between *phnAd* and *phnB* is 82 bp, and since *phnB* appears to be the most downstream gene of this operon, it is probable that it is also transcribed on the same fragment as *phnSFECDAc*.

Therefore, on the basis of this evidence the *phnSFECDAcAdB* genes constitute a cotranscribed, transcriptionally regulated operon.

CHAPTER 5

Cloning and characterisation of lower pathway genes

5.1 Catechol 2,3-dioxygenase activity of RP007

Catechol is a common intermediate in the catabolism of aromatic compounds and may be metabolised by extradiol cleavage by a catechol 2,3-dioxygenase (C23O) and enter the *meta* cleavage pathway, or by intradiol cleavage by a catechol 1,2-dioxygenase whereafter it is processed via the β-ketoadipate pathway. *Burkholderia* sp. RP007 possesses C23O activity that is induced by growth on aromatic compounds, which suggests RP007 metabolises catechol via the *meta* cleavage pathway.

RP007 was grown in 1 L cultures of MM supplemented with either acetate (10 mM), benzoate (5 mM), or phenanthrene (0.05% wt vol^{-1}) to a cell density of approximately 5×10^{10} cfu ml^{-1}. Cell pellets from these cultures were disrupted by sonication to yield a crude cell extract which was used for the *meta* cleavage enzyme assay as described in Chapter 2. The specific activities are shown in Table 5.1.

These specific activities represent a combination of the individual activities of possibly several enzymes in RP007 which have *meta* cleavage activity. Therefore no conclusions can be made about substrate specificity of any RP007 C23O since any C23O activity will not be distinguishable from the background of other extradiol dioxygenase activity. However, the data does show that specific activity towards catechol is 5-fold greater when RP007 is grown on aromatic substrates than on acetate, and that the RP007 cell extracts show greater activity towards catechol and 3-methylcatechol than towards 4-methylcatechol.

5.2 Cloning of C23O genes from RP007

The gene for C23O may be readily screened for using a colorimetric reaction which exploits the ability of C23O to cleave catechol yielding the intensely yellow product 2-hydroxymuconic semialdehyde. *E. coli* clones carrying a C23O gene will appear as yellow colonies when sprayed with a dilute aqueous solution of catechol. This technique was used to detect C23O genes cloned from RP007.

RP007 genomic DNA was used to generate a library in pUC18, as described in Chapter 4 for cloning upper pathway genes. Competent *E. coli* cells transformed with this library were plated onto solid LB media and sprayed with a catechol solution

Growth substrate	Specific activity (μmol/min/mg protein)		
	Catechol	3-Methylcatechol	4-Methylcatechol
Acetate	49	34	12
Benzoate	251	175	64
Phenanthrene	251	62	76

Table 5.1 The effect of growth substrate on the *meta* cleavage activity of crude cell extracts of *Burkholderia* sp. RP007 towards catechol and methylcatechols.

after overnight growth of colonies. Two clones were obtained which showed a 'catechol-positive' phenotype; these were designated pH1 and pH2. Restriction analysis of these constructs revealed pH1 and pH2 are *Hin*dIII fragments of 10.5 kb and 11.2 kb cloned into the *Hin*dIII site of the pUC18 MCS (Figures 5.3A, 5.4A).

5.3 Hybridisation of pH1 and pH2 C23O genes

A 1.5 kb *Eco*RI-*Bam*HI fragment of pH1, known to include part of the C23O gene, hybridised to a 2.5 kb band of *Eco*RI-digested RP007 plasmid DNA (Figure 5.1). This size corresponds to the pH5 *Eco*RI fragment of pH1 which includes the C23O gene (Figure 5.3B). It did not hybridise to RP007 total genomic DNA or to the pH2 clone.

A 1.2 kb fragment of pH2 including the entire C23O gene was amplified by PCR using custom primers. This fragment hybridised to bands of 10–12 kb in *Eco*RI-digested RP007 plasmid DNA, *Eco*RI- and *Hin*dIII-digested RP007 total genomic DNA (Figure 5.2). The hybridised band of *Hin*dIII-genomic DNA represents the 11.2 kb *Hin*dIII pH2 fragment. The expected size of the hybridised band of the *Eco*RI plasmid and genomic digests is unknown since the C23O gene is not located on a discrete *Eco*RI fragment in pH2, ie. the adjacent *Eco*RI site falls outside the pH2 fragment. The pH2 C23O did not hybridise to the pH1 clone.

Hybridisation of pH1 and pH2 probes to plasmid preparations of RP007 indicates both C23O genes are plasmid-borne. The pH2 C23O probe also hybridised to RP007 total genomic DNA, whereas the pH1 probe did not. The ability to detect plasmid DNA in total genomic DNA preparations of RP007 is variable, as discussed in section 4.3, and may be related to the relative concentrations of the target gene in these preparations.

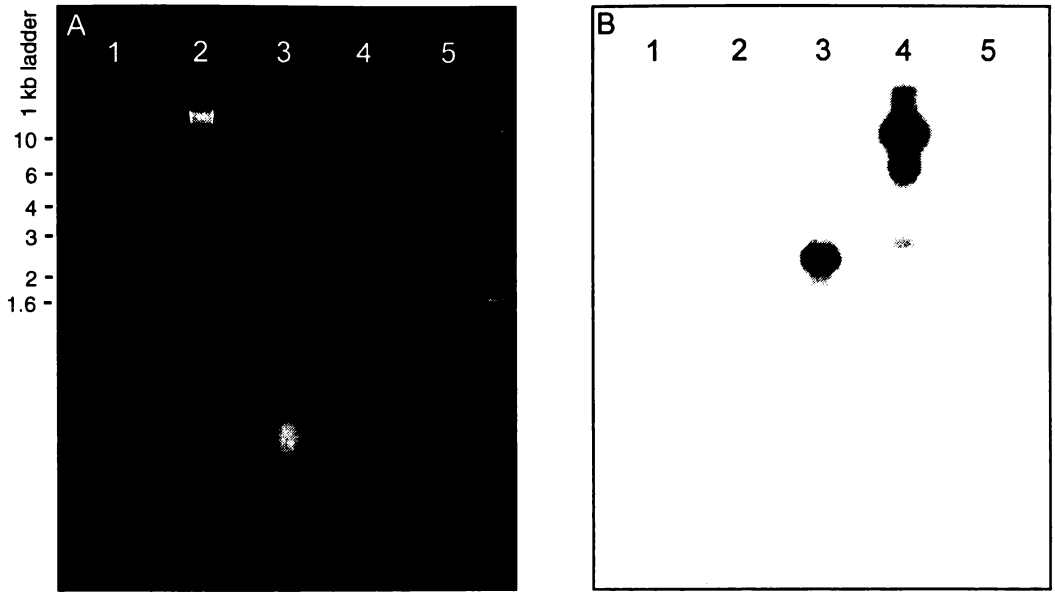


Figure 5.1 pH1 C23O-gene hybridisation experiment. (A) Agarose gel used for Southern blotting experiment; lane 1, *Eco*RI-digested RP007 genomic DNA; 2, *Hin*dIII-digested RP007 genomic DNA; 3, *Eco*RI-digested RP007 plasmid DNA; 4, *Hin*dIII-digested pH1; 5, *Hin*dIII-digested pH2. Both outside lanes are 1 kb ladder. (B) Autoradiogram showing hybridisation of [^{32}P]-labelled probe (1 kb of pH1 C23O gene) to RP007 plasmid DNA and the pH1 positive control; there is no hybridisation to pH2. The lanes are the same as for (A).

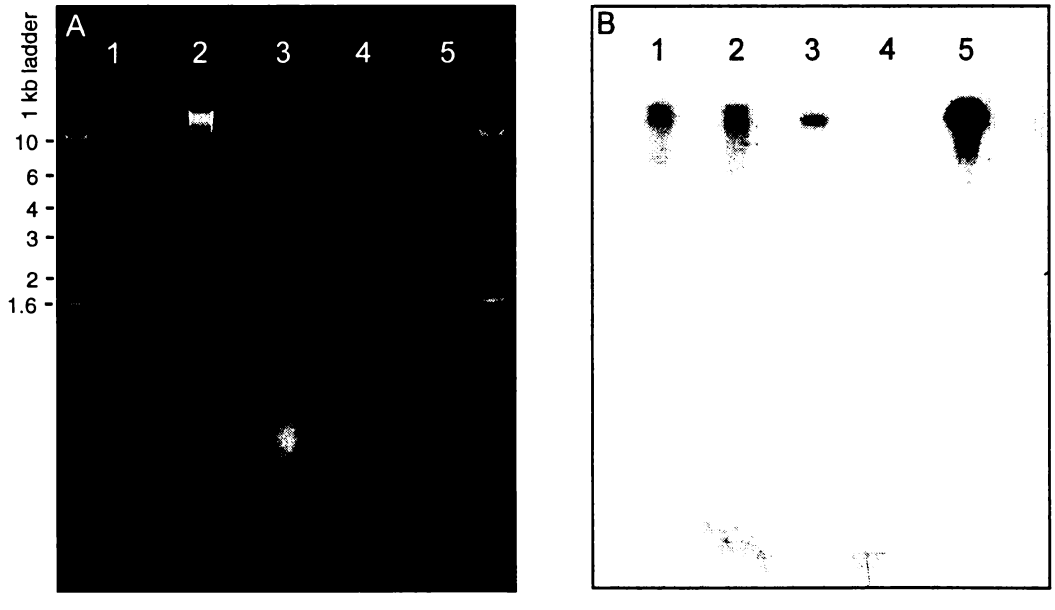


Figure 5.2 pH2 C23O-gene hybridisation experiment. (A) Agarose gel used for Southern blotting experiment; lane 1, *Eco*RI-digested RP007 genomic DNA; 2, *Hin*dIII-digested RP007 genomic DNA; 3, *Eco*RI-digested RP007 plasmid DNA; 4, *Hin*dIII-digested pH1; 5, *Hin*dIII-digested pH2. Both outside lanes are 1 kb ladder. (B) Autoradiogram showing hybridisation of [^{32}P]-labelled probe (1.2 kb of pH2 C23O gene) to both *Eco*RI- and *Hin*dIII-digested RP007 genomic DNA, RP007 plasmid DNA, and the pH2 positive control; there is no hybridisation to pH1. The lanes are the same as for (A).

Substrate	pH1		pH2	
	Spec.[a]	Rel.[b]	Spec.[a]	Rel.[b]
Catechol	1735±294	100	781±78	89
3-Methylcatechol	257±25	15	873±102	100
4-Methylcatechol	529±105	30	416±47	48
2,3-dihydroxybiphenyl	0	0	64±16	7

[a] Specific activity (μmol min^{-1} mg^{-1} protein). Each value is the mean from three different cultures, showing standard deviation.

[b] Relative activity (%)

Table 5.2 *meta* cleavage activities of the pH1 and pH2 catechol 2,3-dioxygenases.

5.4 Substrate specificities of pH1 and pH2

The specific activity of the pH1 and pH2 C23Os expressed in *E. coli* against different substrates was measured (Table 5.2). The C23Os show slightly different substrate specificity profiles. pH1 has an obvious preference for catechol and no activity towards 2,3-dihydroxybiphenyl, whilst pH2 has similar activities to catechol and 3-methylcatechol, and some activity towards 2,3-dihydroxybiphenyl.

5.5 Nucleotide sequencing of pH1 and pH2

The *meta* cleavage activity of subcloned derivatives of pH1 and pH2 allowed the localisation of the C23O genes within these fragments. Figure 5.3B shows the pH1 C23O gene was localised to a central *Eco*RI fragment (pH5) and flanked the *Bam*HI site. Nested deletion derivatives of pH5 and pH5R (the fragment inserted in the pUC18 MCS in the opposite orientation) were used to obtain the complete nucleotide sequence of the pH5 *Eco*RI fragment (Figure 5.5).

The pH2 C23O activity was localised to a 3.5 kb *Eco*RI-*Eco*RV fragment (pH2E3) as shown in Figure 5.4B. Nested deletion derivatives of pH2E3 and sequencing from custom primers were used to obtain the complete nucleotide sequence for both strands of the pH2E3 fragment (Figure 5.6).

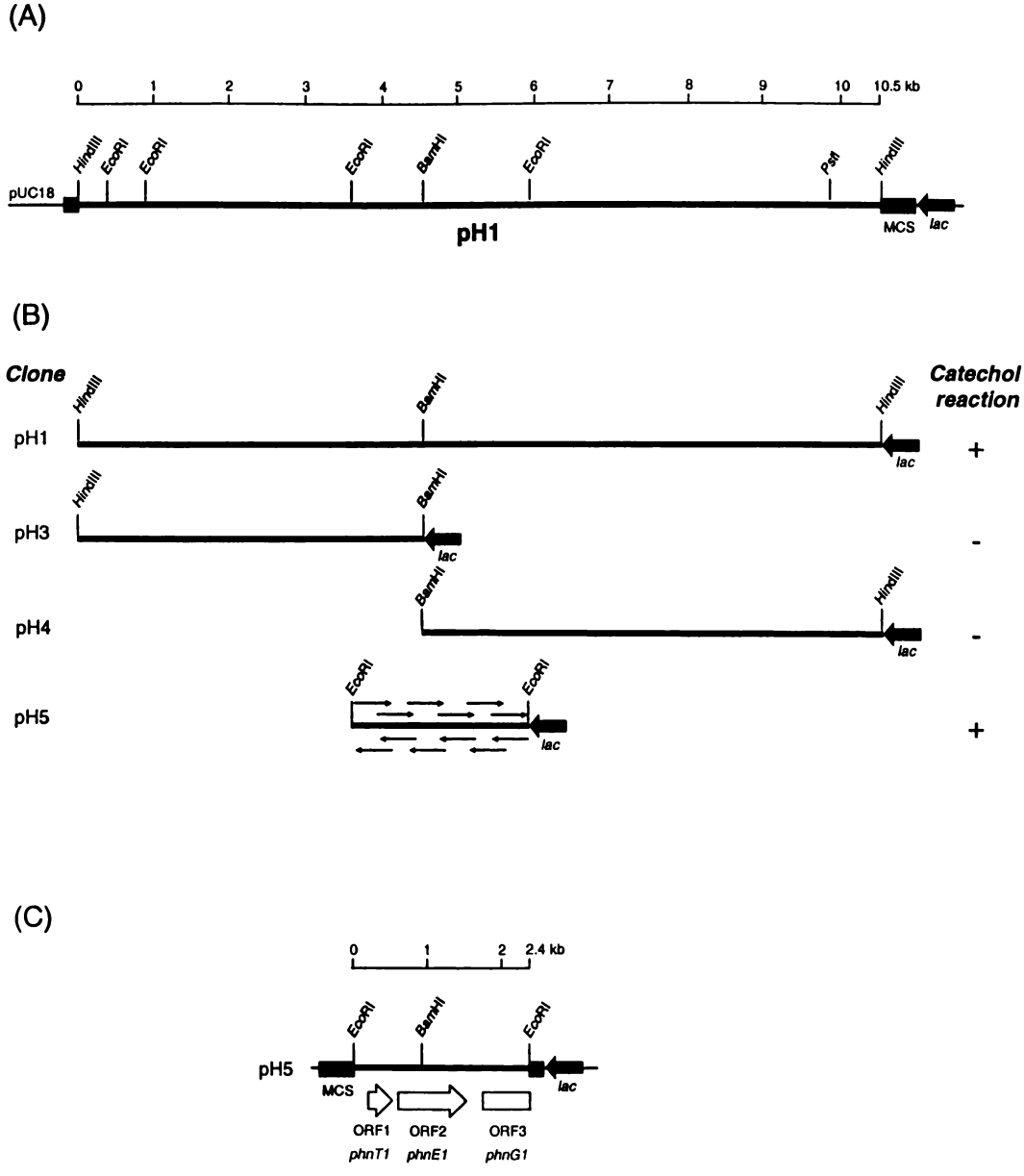


Figure 5.3 (A) Physical map of the 10.5 kb *Hin*dIII insert of pH1 showing sites of common restriction enzymes, and the orientation of the *lac* promoter. (B) Catechol *meta* cleavage activities of subclones of pH1 showing the localisation of the catechol 2,3-dioxygenase activity to the central 2.4 kb *Eco*RI fragment. The small arrows show the sequencing strategy for pH5. (C) Physical map of the pH5 fragment showing the position and orientation of the two complete and one partial ORFs identified on this fragment.

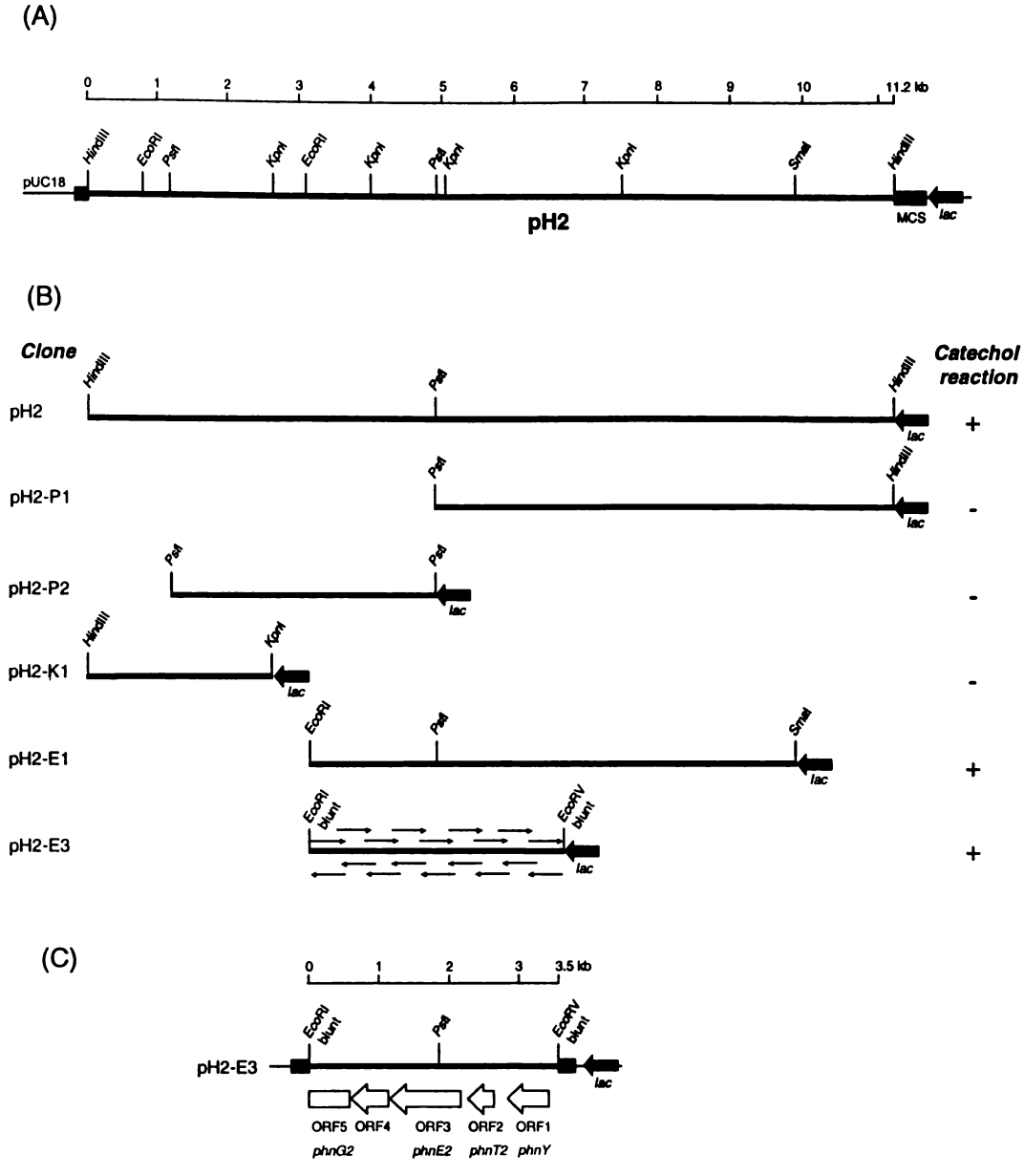


Figure 5.4 (A) Physical map of the 11.2 kb *Hin*dIII insert of pH2 showing sites of common restriction enzymes, and the orientation of the *lac* promoter. (B) Catechol *meta* cleavage activities of subclones of pH2 showing the localisation of the catechol 2,3-dioxygenase activity to the central 3.5 kb *Eco*RI-*Eco*RV fragment. The small arrows show the sequencing strategy for pH2E3. (C) Physical map of the pH2E3 fragment showing the position and orientation of the four complete and one partial ORFs identified on this fragment.

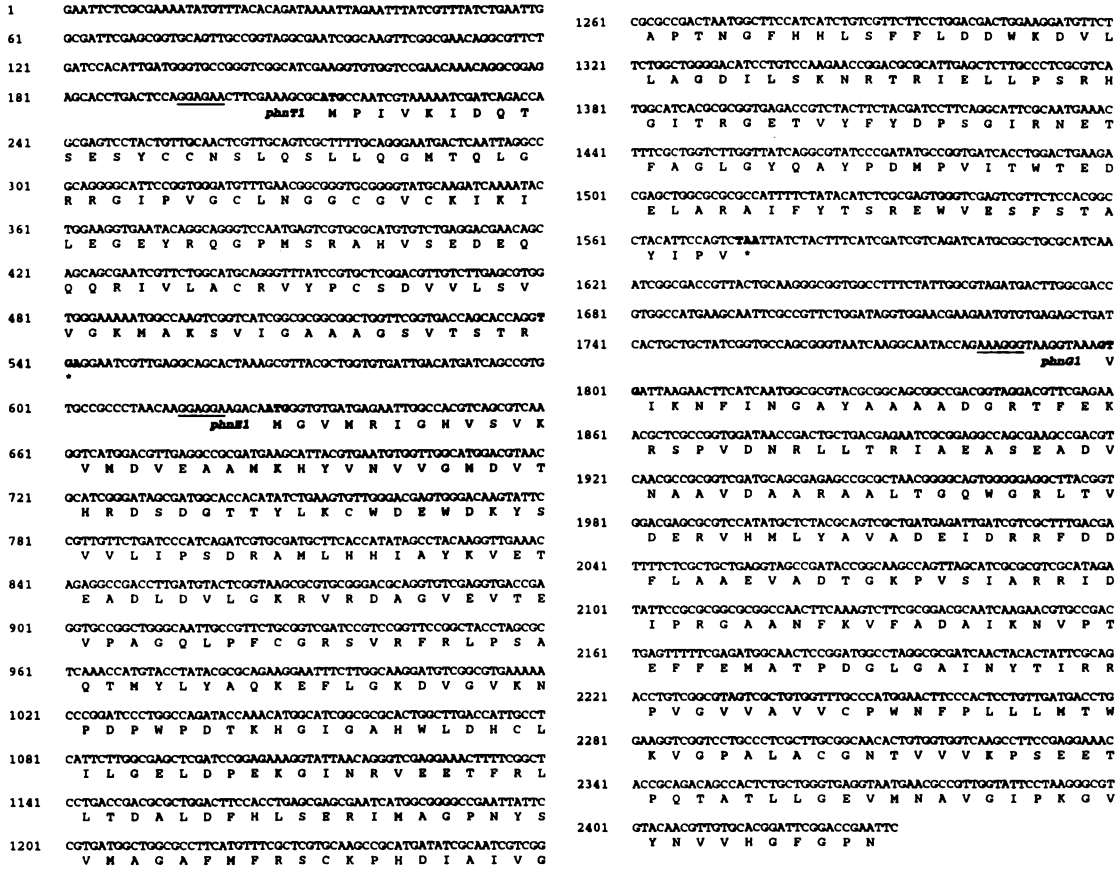


Figure 5.5 Complete nucleotide sequence of the pH5 subcloned fragment of pH1 from *Burkholderia* sp. RP007, cds 1–2433. Start and stop codons are in bold type, asterixes in the translated sequence indicate stop codons. Possible ribosomal binding sites (Shine-Dalgano sequences) are underlined. The sequence is shown 5'→3'; all genes are transcribed 5'→3' as shown here.

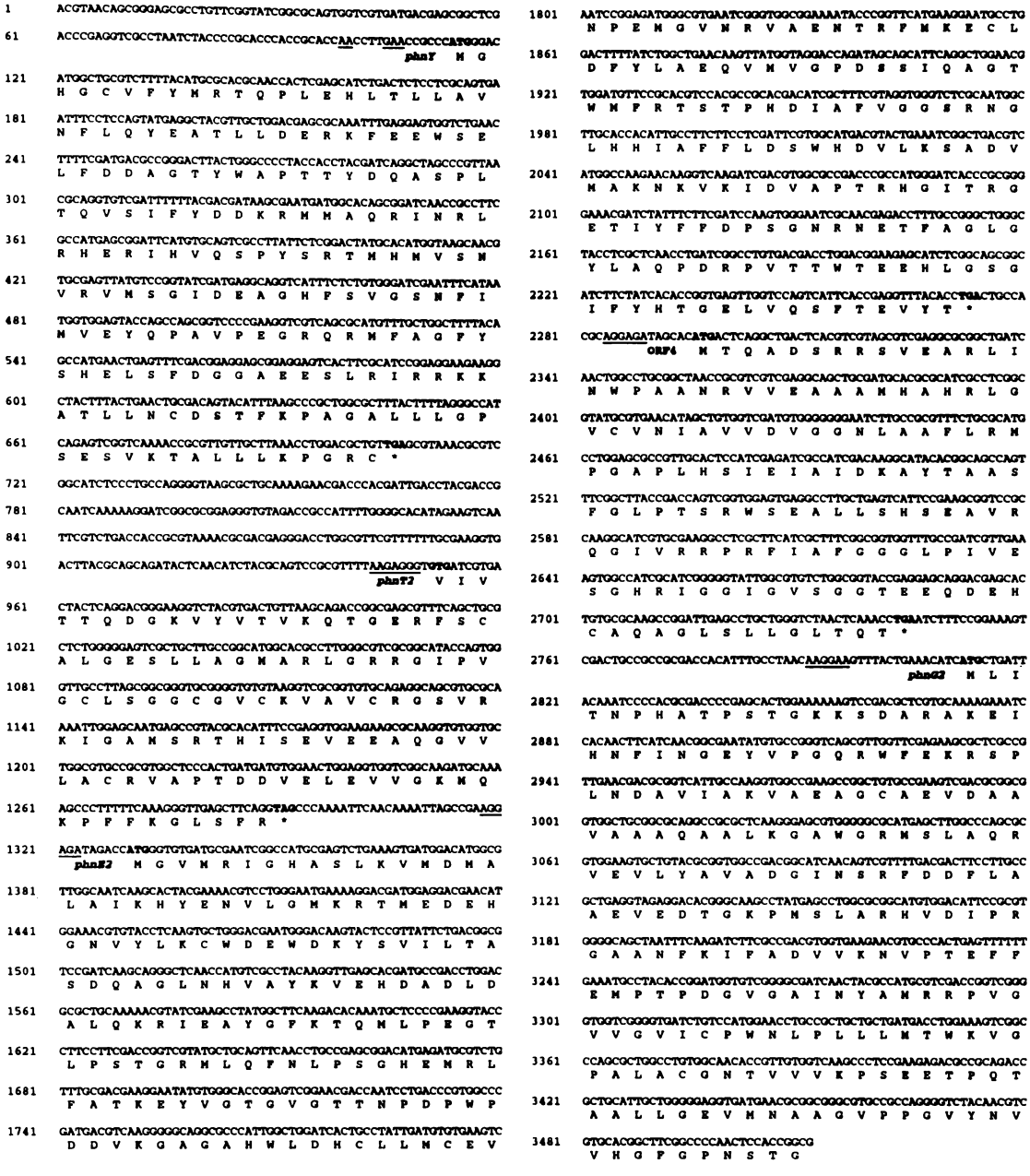


Figure 5.6 Complete nucleotide sequence of the 3511 bp *Eco*RI-*Eco*RV pH2E3 subcloned fragment of pH2 from *Burkholderia* sp. RP007. Start and stop codons are in bold type, asterixes in the translated sequence indicate stop codons. Possible ribosomal binding sites (Shine-Dalgano sequences) are underlined. The sequence is shown 5'→3'; all genes are transcribed 5'→3' as shown here.

5.6 Analysis of the pH1 and pH2 sequence

5.6.1 pH1

Analysis of the nucleotide sequence of the 2433 bp central *Eco*RI fragment of pH1 (Figure 5.5) revealed two complete ORFs of 109 and 316 amino acid residues and a partial ORF for which the sequenced region coded 211 residues (Figure 5.3C). ORF 2 is the C23O gene, designated *phnE1*, and is most similar to the C23O of *B. cepacia* AA1 (Ma & Herson, 1996) to which the amino acid homology is 59%. The dendrogram in Figure 5.7 shows the C23Os so far described are grouped into three classes. PhnE1 is part of the class which includes TdnC of *P. putida* UCC2 and C23OII of *P. putida* mt-15.

Upstream of *phnE1* is ORF 1 which encodes a putative chloroplast-type ferredoxin and is designated *phnT1*. This peptide shows ca. 30% amino acid similarity to chloroplast-type ferredoxins from other *meta* pathways including XylT and NahT from the TOL and NAH *meta* pathways (Harayama *et al.*, 1991a). Like other ferredoxins, all chloroplast-type ferredoxins contain a [2Fe-2S] cluster which is formed by four sulfide ligands contributed by four cysteine residues arranged in a $CX_4CX_2CX_{31-32}$ motif (Mason & Cammack, 1992; Neidle *et al.*, 1991). In many *meta* pathways a chloroplast-type ferredoxin is found immediately upstream of the C23O gene. The function of this gene was unknown until Polissi & Harayama (1993) demonstrated that XylT was necessary to reactivate the C23O after non-competitive inhibition by the *meta* cleavage product of catechol or methyl-substituted catechols. The reduced form of XylT reactivates the inactivated C23O; however, it was not demonstrated how the oxidised XylT is subsequently reduced.

The incomplete ORF (ORF 3) is designated *phnG1* and is initiated by a GTG start codon. The sequence of the 211 amino acids encoded by this region shows high homology to the N-terminal portion of several 2-hydroxymuconic semialdehyde dehydrogenases (HMSD) of other lower pathways for aromatic degradation, including 68% to DmpC of *Pseudomonas* sp. CF600 (Nordlund & Shingler, 1990) and 58% to XylG from *P. putida* mt-2 (pWW0) plasmid (Harayama *et al.*, 1991b). The similarity of the partial amino acid sequence of this gene and its location directly downstream of the C23O gene, which is highly conserved for other HMSD, supports

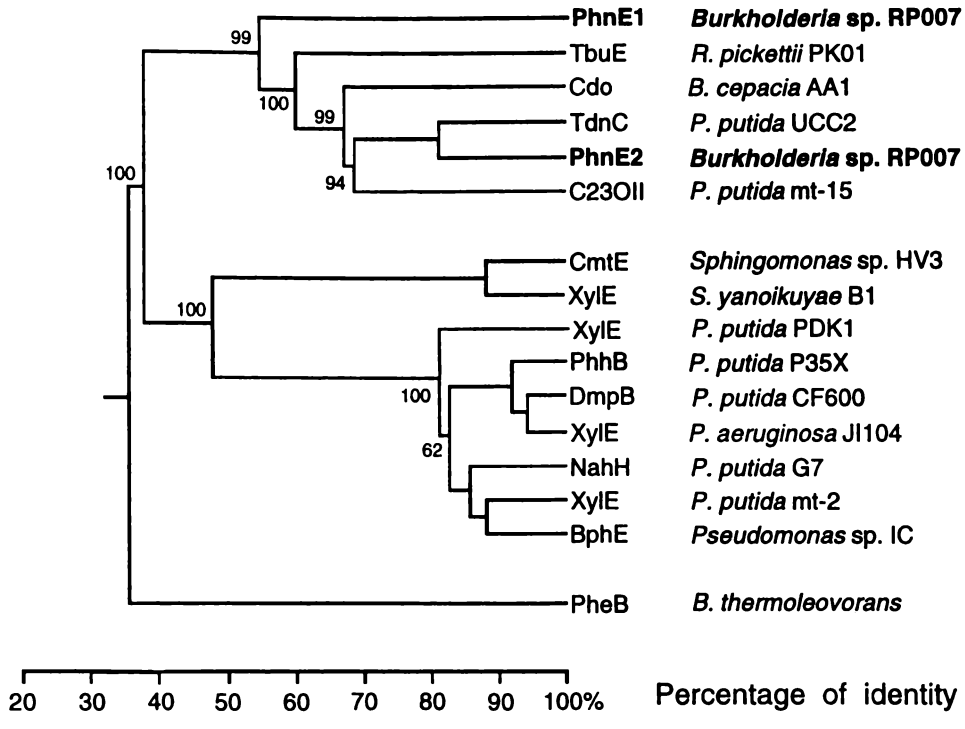


Figure 5.7 Dendrogram of predicted amino acid sequences of catechol 2,3-dioxygenases (C23Os) which show homology to PhnE1 and PhnE2 of *Burkholderia* sp. RP007. This dendrogram is an unrooted tree rerooted using PhnB (*B. thermoleovorans*) as an outgroup. The scale shows the percentage amino acid similarity between the sequences. The GenBank accession numbers and/or references for each protein and strain, including the substrate on which that strain was isolated, can be found in Figure A.1 (Appendix A).

the designation of this partial gene as *phnG1* encoding a putative HMSD.

The order of the three genes identified on the sequenced region of pH1 (*phnTEG*) is conserved in many *meta* cleavage pathways for aromatic degradation including the *xyl* operon of the TOL pathway (Harayama & Rekik, 1993), the *nah* genes of the NAH7 *sal* operon (Yen & Gunsalus, 1985) and the *dmp* operon (Shingler *et al.*, 1992) (Figure 5.8). The conservation of the gene order of these *phn meta* pathway genes suggests the remainder of the operon also has a similar order to the *xyl/nah/dmp* operons (*TEGFJKIH*). Some of these genes are likely present on the pH1 fragment, but no attempt was made to characterise them.

5.6.2 pH2

The nucleotide sequence was determined for 3511 bp of an internal *Eco*RI-*Eco*RV fragment of pH2 which includes the C23O gene (Figure 5.6). Analysis of this sequence identified four complete ORFs and one partial ORF (Figure 5.4C).

ORF1 encodes a 197 amino acid peptide which shows low homology to the β-

subunits of a group of aromatic 1,2-dioxygenases. ORF 1 shows 18–20% amino acid homology to BenB for benzoate oxidation by *A. calcoaceticus* (Neidle *et al.*, 1991), XylY for toluene oxidation by *P. putida* mt-2 (pWW0) (Harayama *et al.*, 1991b), BnxB and Bed2 for benzene oxidation by *P. putida* (Irie *et al.*, 1987; Tan *et al.*, 1993), and BphE for biphenyl oxidation by *Burkholderia* sp. LB400 (Erickson & Mondello, 1992). Although this similarity is low it is similar to the level of homology shown between some peptides of this group; BenB is 58% similar to XylY but only 14–17% similar to BnxB, Bed2 and BphE; and BnxB is 15–20% similar to all of these peptides. The homology to these subunits suggests ORF 1, which is designated *phnY*, is part of a three component aromatic 1,2-dioxygenase analogous to the *xylXYZ* and *benABC* (Neidle *et al.*, 1991) dioxygenases, and its function is to generate catechol (or substituted catechol) which serves as the substrate for the downstream *meta* pathway enzymes.

ORF 2 is located 243 bp downstream from the stop codon of ORF 1 and encodes a putative chloroplast-type ferredoxin of 113 amino acids, designated *phnT2*. PhnT2 is 39% similar to the PhnT1 ferredoxin of pH1 and ca. 25% similar to other chloroplast-type ferredoxins including XylT and NahT. The *phnT2* gene is initiated by a GTG start codon and its length is close to that of PhnE1 (109 residues), XylT (112) and NahT (108).

ORF 3 is the C23O gene which is designated *phnE2*. The predicted amino acid sequence of PhnE2 is most similar to the 3-methylcatechol 2,3-dioxygnease, TdnC, of *P. putida* UCC2 (McClure *et al.*, 1991). The dendrogram in Figure 5.7 shows PhnE2 is part of the same class of C23Os as PhnE1, but shows higher similarity to members of this group than PhnE1, which is the most outlying member. The amino acid homology between PhnE1 and PhnE2 is 57%.

ORF 4 begins 22 bp downstream from the stop codon of *phnE2* and codes for a 149 amino acid peptide. The predicted amino acid sequence of ORF 4 is 41% similar to CmpX of *Sphingomonas* sp. HV3 (Yrjälä *et al.*, 1997) for which the authors could not assign a function. Both *cmpX* and ORF 4 are located between the C23O gene and the HMSD gene of their respective *meta* pathways, and a similar sized ORF of unidentified function has been found in the same location for other *meta* pathways, including the *phn* operon of *Pseudomonas* sp. DJ77 (Kim *et al.*, 1997a).

67 bp downstream of ORF 4 begins ORF 5 which extends to the end of the

sequenced region. The 235 amino acids encoded by this partial ORF are 67% similar to the N-terminal partial sequence of PhnG1 and ca. 60% similar to 2-hydroxymuconic semialdehyde dehydrogenases (HMSD) of the *xyl*, *dmp*, and *cmp meta* operons. On this basis ORF 5 was designated *phnG2* encoding the HMSD of this pathway.

5.6.3 Gene order of pH1 and pH2 *meta* pathways

The order of the genes identified on the pH1 *meta* pathway is: *phnT1* (chloroplast-type ferredoxin) – *phnE1* (C23O) – *phnG1* (HMSD). The order of the pH2 *meta* pathway genes is: *phnY* (β-subunit of aromatic 1,2-dioxygenase) – *phnT2* (chloroplast-like ferredoxin) – *phnE2* (C23O) – ORF 4 (unidentified) – *phnG2* (HMSD).

Much significance has been attributed to the gene order of *meta* pathways, particularly to the genes adjacent the C23O gene. The high degree of conservation of gene order between several well characterised *meta* pathways (*xyl*, *nah*, *dmp*, *bph*) suggests that these operons evolved from a common ancestor (Williams & Sayers, 1994). A difference in the order of two genes (HMSD and 2-hydroxymuconic semialdehyde hydrolase (HMSH)) which follow the C23O gene distinguishes the *tbu* operon from this group (Kukor & Olsen, 1996), and the *tod* genes of *P. putida* F1 (Zylstra & Gibson, 1989) also have a divergent gene order. Two recently characterised *meta* operons show a further difference in gene order. The *cmp* genes of *Sphingomonas* sp. HV3 (Yrjälä *et al.*, 1997) and the *phn* genes of *Pseudomonas* sp. DJ77 (Kim *et al.*, 1997a) are arranged HMSH–C23O–U–HMSD (where U is an unidentified ORF (*cmpX*)).

The pH1 *meta* operon shows the same order for the ferredoxin–C23O–HMSD genes as the *xyl/nah/dmp/bph* pathways which suggests it belongs to this group (Figure 5.8). The pH2 *meta* operon shows a similar order to the *cmp/phn* group, except that upstream of the C23O gene is a chloroplast-type ferredoxin and apparently the genes for an aromatic 1,2-dioxygenase. It is noteworthy that the pH2 operon also includes the gene of unknown function found in the *cmp* and *phn* operons.

It is unlikely that the pH2 *meta* pathway is involved in salicylate degradation since it appears a multicomponent dioxygenase generates the catechol substrate for the C23O and subsequent steps analogous to the *xyl* and *ben* pathways. In the *nah meta*

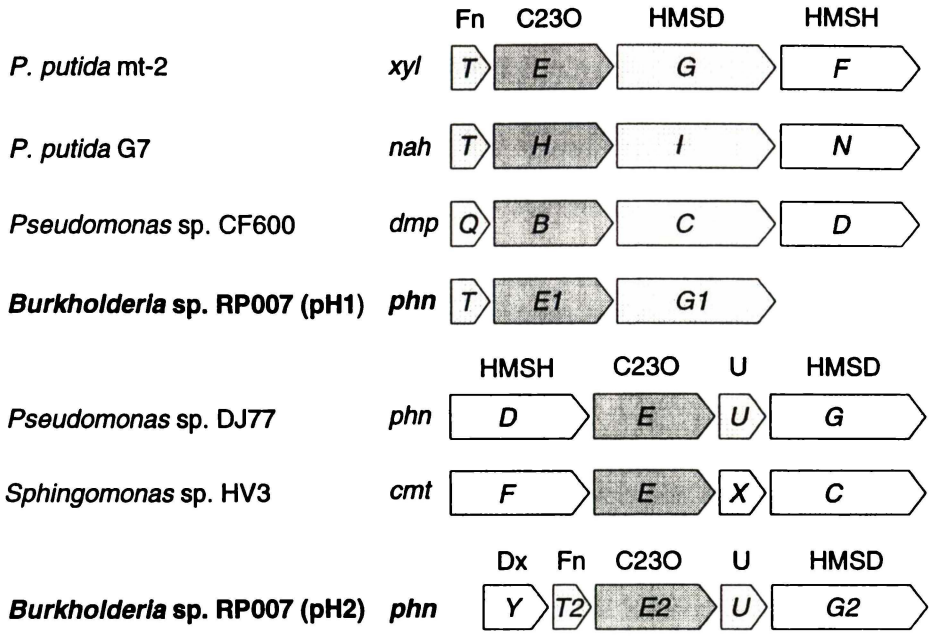


Figure 5.8 Gene organisation of different *meta* pathway operons relative to the *Burkholderia* sp. RP007 *meta* operons encoded on pH1 and pH2. Gene abbreviations are: Fn, chloroplast-like ferredoxin; C23O, catechol 2,3-dioxygenase; HMSD, 2-hydroxymuconic semialdehyde dehydrogenase; HMSD, 2-hydroxymuconic semialdehyde hydrolase; U, gene of unknown function; Dx, β-subunit of aromatic dioxygenase.

and subsequent steps analogous to the *xyl* and *ben* pathways. In the *nah meta* pathway salicylate is converted to catechol by a single monooxygenase (salicylate hydroxylase) and the presence of a putative β-subunit of a multicomponent dioxygenase on pH2 is not consistent with this reaction scheme.

One can therefore speculate the roles of the two *Burkholderia* sp. RP007 *meta* pathway operons. The three genes characterised from the pH1 pathway have the same gene order as the *nah meta* pathway genes for salicylate degradation, and in the absence of any evidence suggesting this operon does not include a gene for salicylate hydroxylase, it is possible the pH1 operon encodes a pathway for salicylate degradation. This operon is therefore the lower pathway for naphthalene and phenanthrene catabolism. The presence of a putative multicomponent dioxygenase in the pH2 operon suggests this pathway is not involved in the dissimilation of salicylate but may be an independently acquired operon for the catabolism of other single ring aromatic substrates such as benzoate, which is readily utilised by RP007. The gene order and the presence of an unidentified ORF with homology to *cmtX* suggests this operon is of a similar class to the *meta* pathways of *Sphingomonas* sp. HV3 and *Pseudomonas* DJ77.

The roles of the two RP007 *meta* pathway operons as suggested by this hypothesis are supported by the enzyme activity data for the pH1 and pH2 clones (Table 5.2) and the RP007 wild-type (Table 5.1). The pH2 C23O has a high activity towards both catechol and 3-methylcatechol, whereas the pH1 C23O has maximum activity toward catechol and low activity toward 3-methylcatechol. When RP007 was grown at the expense of benzoate, cell extracts showed high activity towards both catechol and 3-methylcatechol, indicating induction of the pH2 C23O under these conditions. When RP007 was grown at the expense of phenanthrene, cell extracts showed high activity toward catechol only, and only a small increase in activity towards 3-methylcatechol relative to the cell extract from RP007 grown on acetate. These results suggest that growth with benzoate induces the pH2 C23O which has a broad substrate specificity, and that growth with phenanthrene induces the narrow substrate range pH1 C23O. Since phenanthrene is presumably degraded via salicylate in RP007, this supports the involvement of the pH1 *meta* pathway in the lower pathway degradation of phenanthrene (and naphthalene).

CHAPTER 6

Molecular ecology of *phn* and *nah* genes

6.1 PAH catabolic genes in the environment

It is known that the *nah*-like genes for naphthalene and phenanthrene degradation are highly conserved and are ubiquitously distributed throughout the globe. Characterisation of the catabolic genes of PAH degrading isolates derived from diverse geographical locations has invariably revealed the presence of genes highly homologous to the archetypal *nah* genes of *P. putida* G7 (reviewed in Chapter 1). Furthermore, several studies have shown that *nah*-like genes can be detected in DNA extracted directly from hydrocarbon-contaminated soil, indicating the presence of strains harbouring *nah*-like operons in these environments (Berthelet *et al.*, 1996; Guo *et al.*, 1997; Herrick *et al.*, 1993).

Despite the failure of any workers to characterise PAH degradative genes that are not homologous to *nah*-like genes, it is commonly accepted that one highly conserved group of genes for PAH catabolism cannot reflect the true diversity of such determinants in the environment (Goyal & Zylstra, 1996; Zylstra *et al.*, 1997). However, the options remain limited for attempts to characterise or detect genes divergent to *nah*-like genes since the only probes available are those based on *nah*-like sequences.

The characterisation of the *phn* genes of *Burkholderia* sp. RP007 has revealed an operon that is genotypically divergent to *nah*-like loci and which therefore represents a new group of naphthalene/phenanthrene catabolic genes. The degree of divergence of the *phn* genes to the corresponding *nah*-like genes means that probes designed based on *nah* sequences will not detect *phn* genes. It was therefore of interest to investigate the distribution of this new group of PAH catabolic genes to determine its prevalence in the environment with respect to *nah*-like genes. Are they represented in the environment as much as *nah* (with respect to copy number and distribution) or is this group present in a much smaller population of PAH degraders than *nah*?

Two approaches were used to investigate the prevalence and distribution of *phn*-like genes in the environment, relative to *nah*-like genes. Firstly, isolates were cultured from hydrocarbon-contaminated soil samples for their ability to utilise naphthalene or phenanthrene as sole carbon source, and screened for hybridisation to probes based on *phn* or *nah* genes. The second approach was to use *phn* or *nah* probes to detect like sequences in DNA extracted directly from soil samples. In this way the presence

of *phn*- or *nah*-like genes in both the culturable and nonculturable microbial communities could be estimated.

6.2 Design of *phn* and *nah* probes

The gene encoding the iron-sulpur protein α subunit (ISPα) of the initial dioxygenase (*phnAc*/*nahAc*) was selected as a probe indicative of either *phn*- or *nah*-like genes. This gene was chosen because: (a) it has a role in determining the substrate specificity of the initial dioxygenase (Parales *et al.*, 1998) and therefore determines the substrate specificity of the whole pathway; (b) the ISPα gene has been characterised for many aromatic degrading strains meaning highly conserved regions could be identified; and (c) the low homology between *phnAc* and *nahAc*-like genes (55% amino acid homology) means probes will not hybridise to target DNA of the other group.

The P8073/P9047 primer set was designed to amplify a 994 bp product from the *phnAc* gene. The nahAcfor/nahAcrev primer set was designed based on *nahAc*-like sequences and amplified a 993 bp product from *nahAc*-like genes. These primers were designed so that each set amplified a product from the same region of each gene of an almost identical size. The ability of each primer set to amplify a product of the expected size from the corresponding wild-type strains was verified. That the *phn* primer set did not amplify a product from *nah* strains, and the *nah* primers did not amplify a product from RP007, even under relaxed PCR conditions, was also confirmed. Table 6.1 and Figure 6.1 show the sequences and locations of these primers in the ISPα gene.

PCR conditions for detecting *phn* or *nah* genes in cultured strains consisted of 25 cycles with a 55°C annealing temperature. Under these conditions the nahAcfor/ nahAcrev primer set amplified a product from *P. putida* OUS82, indicating that 2–3 mismatches are tolerated.

Primer	Sequence
P8073	5'-TTCGAGCTGGAATGTGAGC-3'
P9047	5'-AATAACCGGCGATTCCAAAC-3'
nahAcfor	5'-TGGCGATGAAGAACTTTTCC-3'
nahAcrev	5'-AACGTACGCTGAACCGAGTC-3'

Table 6.1 Primers used to amplify *phnAc*-like (P8073/P9047) or *nahAc*-like (nahAcfor/ nahAcrev) genes.

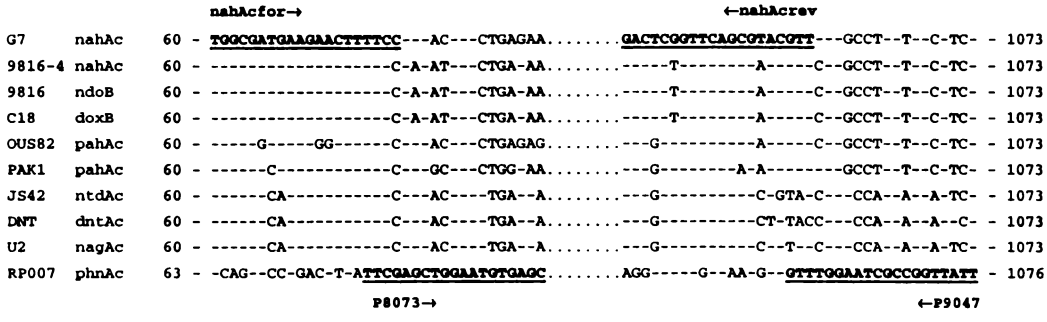


Figure 6.1 Alignment of nucleotide sequences of *phnAc* and *nahAc*-like genes showing the location of the P8073/P9047 and nahAcfor/nahAcrev primers (underlined). Nucleotides which conform to the primer sequence are shown as dashes, the nucleotides which differ are shown. For clarity the downstream primers nahAcrev and P9047 are shown as their complimentary sequence.

6.3 Isolation and characterisation of PAH catabolic strains

PAH degrading bacteria were isolated from 11 different contaminated soil samples to reduce the possibilty of duplication of isolates, and to ensure the culture collection represented diverse environmental samples. Bacteria were cultured based on their ability to utilise either naphthalene or phenanthrene as sole carbon and energy source, as described in Chapter 2. Each strain was further tested for growth on naphthalene, phenanthrene, toluene, and biphenyl, and for its ability to oxidise indole and catechol. The Gram-type of each strain was determined using the KOH test.

PCR and dotblot hybridisation was used to assess the genotype of the PAH catabolic determinant for each strain. A genomic DNA preparation was obtained for each strain and used as a template for PCR using either the P8073/P9047 (*phn*) or nahAcfor/nahAcrev (*nah*) primers. For dotblot hybridisation approximately 3 μg of each DNA prep was fixed to a nylon membrane which was hybridised to *phn* or *nah* probes. [^{32}P]-labelled probes were prepared from PCR products amplified by the P8073/P9047 and nahAcfor/nahAcrev primers from RP007 and *P. putida* G7

genomic DNA, respectively.

Analysis of the growth and hybridisation data (Table 6.2) reveals that 20 of the 44 naphthalene isolates (45%), and 8 of the 33 phenanthrene isolates (24%) had *nah*-like genes. All hybridising naphthalene isolates were Gram-negative, only one of which was also able to degrade phenanthrene. Positive indole (ind^+) and catechol (cat^+) reactions correlate well with the *nah* genotype being present in 17 of 20 hybridising isolates. A *nahAc* PCR product was amplified from only 6 of 33 phenanthrene isolates, three of which conformed to the naphthalene and phenanthrene degrading/ind^+cat^+ phenotype. No homology to the *phn* probe was detected by PCR or hybridisation for any of the isolates. None of the Gram-positive isolates hybridised to the *nah* probe, and the ability of strains to utilise biphenyl or toluene did not appear to be related to indole or catechol phenotypes or to hybridisation to *nah*.

There was excellent correlation between strains from which the *nahAc* PCR primers amplified a product, and those which hybridised to the *nahAc* probe. Only 1 naphthalene and 2 phenanthrene isolates which did not give a *nahAc* PCR product hybridised to the *nahAc* probe, and 2 naphthalene isolates which gave a *nahAc* PCR product did not hybridise to the *nahAc* probe. It was expected the dotblot hybridisation would show more strains with *nahAc* homology than the PCR screening since the hybridisation experiment would detect homology over a 1 kb region and would be more tolerant of short regions of low homology. Conversely the PCR experiment effectively relies on high homology to two 20 bp regions where the primers bind, and greater than 3 mismatches in each of these regions will prevent annealing of the primers and subsequent amplification, regardless of whether the region internal to the primer-binding sites shows high homology to *nahAc*. The similar results for these two screening methods suggests the *nahAc* primer sites correspond to highly conserved regions of the *nahAc* gene, and are therefore excellent primers for detecting *nahAc*-like sequences.

The complete lack of homology of any of the cultured strains to the *phnAc* probe was a surprising result. It suggests that the *phn* genotype is very rare amongst culturable bacteria, even amongst those strains isolated from Rotowaro, which was where *Burkholderia* sp. RP007 was isolated.

The results show that nearly all strains which are ind^+ (20 of 24) exhibit homology to the *nahAc* gene. Oxidation of indole resulting in the blue product indigo has been

routinely used as a marker for cloning aromatic dioxygenase genes from PAH degrading strains. The *ndo*, *nah*, *dox*, *nag*, and *pah* loci were all cloned into recombinant vectors based on the indole colorimetric reaction. According to these results, ind^+ phenotypes usually have a *nah* genotype, and it is therefore not surprising that all of these catabolic loci are very similar to the achetypal *nah* genes. However, the *phn* genes of RP007 were also cloned using the indole reaction as a marker and these genes are divergent to *nah*-like genes.

The *nahAc* PCR products generated from the naphthalene and phenanthrene isolates were screened for restriction fragment length polymorphisms (RFLP). The ISPα genes from *nah*-like sequences possess a conserved *Pst*I site which cleaves the fragment amplified by the nahAcfor/nahAcrev primers into 270 bp and 722 bp fragments. 21 of 26 *nahAc* PCR products had this internal *Pst*I site indicating these sequences are very similar *nahAc*. The five sequences which lacked the *Pst*I site showed an RFLP pattern that was identical to that of *ntdAc* of *Pseudomonas* sp. JS42 (Parales *et al.*, 1996) and *dntAc* of *Burkholderia* sp. DNT (Suen *et al.*, 1996).

Ten *nahAc* PCR products comprising representatives of each group were partially sequenced to evaluate the diversity of the *nahAc* gene in these isolates. Phylogenetic analysis of this data revealed amplified sequences belonged to one of two groups. Four sequences were >90% similar to *nahAc*-like genes and were grouped with *nahAc*, *ndoB*, *doxB*, and *pahAc*, whilst the other six sequences were >90% similar to *ntdAc*, *dntAc*, and *nagAc* and were grouped with these genes. The homology between these two groups is ca. 80% which means members of both groups will hybridise to the probe derived from *nahAc* of *P. putida* G7. The exception to this was 4N2b and 4N2c which are 78% similar to *nahAc* of *P. putida* G7 and did not hybridise to this probe, but were amplified by the nahAcfor/nahAcrev primers. This result is a useful indicator of the stringency of the hybridisation conditions and shows that sequences with above 80% homology to the probe are detected.

The phylogenies of the sequenced *nahAc* regions form these cultured strains is shown in the dendrogram in Figure 6.4.

		Growth on[c]							PCR		Dotblot[g]	
Isolate[a]	Substrate[b]	naphthalene	phenanthrene	biphenyl	toluene	Indole m[d]	Catechol m[e]	Gram (KOH)[f]	*nah*	*phn*	*nah*	*phn*
LSN1	nap	+	+	-	+	-	-	-	-	-	-	-
LSN2	nap	+	+	-	+	-	-	-	-	-	-	-
LSN3	nap	+	-	-	-	+	+	-	+	-	-	-
LSN4	nap	+	+	-	+	+	+	-	+	-	+	-
LSN5	nap	+	+	-	+	-	-	+	-	-	-	-
LSN6	nap	+	-	-	-	-	-	-	-	-	-	-
LSN7	nap	+	-	-	-	+	+	-	+	-	+	-
LSN8	nap	+	-	-	+	-	-	+	-	-	-	-
LSN10	nap	+	-	-	-	+	+	-	+	-	+	-
LSN11	nap	+	+	-	-	+	-	+	-	-	-	-
LSN12	nap	+	+	-	-	-	-	+	-	-	-	-
LSN13	nap	+	+	-	-	+	+	+	-	-	-	-
LSN14	nap	+	-	-	-	+	+	-	+	-	+	-
LSN15	nap	+	+	-	-	-	-	-	+	-	+	-
LSNa	nap	+	-	-	-	+	+	-	+	-	+	-
LSNb	nap	+	-	-	-	+	+	-	+	-	+	-
LSNc	nap	+	-	-	-	+	+	-	+	-	+	-
LSNd	nap	+	-	-	-	+	+	-	+	-	-	-
LSNe	nap	+	-	-	-	+	+	-	+	-	+	-
LSNf	nap	+	-	-	-	+	+	-	+	-	+	-
HHSN4	nap	+	-	-	-	-	-	-	+	-	+	-
1N1	nap	+	-	-	-	-	-	-	-	-	-	-
1N2	nap	+	-	-	-	-	-	-	-	-	-	-
1N3	nap	+	-	-	+	+	-	-	+	-	+	-
1N4	nap	+	-	-	-	-	-	-	-	-	-	-
1N5	nap	+	-	-	-	-	-	-	-	-	-	-
1N6	nap	+	-	-	+	+	-	-	+	-	+	-
1N7	nap	+	-	-	+	-	-	-	+	-	+	-
2N1	nap	+	-	-	-	+	+	-	+	-	+	-
2Na	nap	+	+	-	+	+	+	-	-	-	-	-
3N1	nap	+	-	-	+	-	-	+	-	-	-	-
3N3a	nap	+	-	+	+	-	-	+	-	-	-	-
3N3b	nap	+	+	+	+	+	-	+	-	-	-	-
4N2a	nap	+	+	-	+	-	-	-	-	-	-	-
4N2b	nap	+	-	-	-	+	-	-	+	-	-	-
4N2c	nap	+	-	-	-	+	-	-	+	-	-	-
4Na	nap	+	-	-	-	-	-	-	+	-	+	-
5N1	nap	+	-	-	+	+	-	-	-	-	-	-
6N1	nap	+	-	-	+	-	-	-	-	-	-	-
7N1	nap	+	-	+	+	-	-	+	-	-	-	-
8N1	nap	+	-	+	+	-	-	+	-	-	-	-
8N2	nap	+	-	+	+	-	-	+	-	-	-	-
8N2a	nap	+	-	-	+	-	-	-	-	-	-	-
8N3a	nap	+	-	-	+	-	-	-	-	-	-	-
LSP1	phe	-	+	-	-	-	-	-	-	-	-	-
LSP2	phe	-	+	-	-	-	-	+	-	-	-	-
LSP3	phe	-	+	-	-	-	-	+	-	-	-	-
LSP5	phe	-	+	-	-	-	-	-	-	-	-	-
HHSP1	phe	-	+	-	-	-	-	-	-	-	-	-
HHSP2	phe	-	+	-	-	-	-	-	-	-	-	-

		Growth on[c]							PCR		Dotblot[g]	
Isolate[a]	Substrate[b]	naphthalene	phenanthrene	biphenyl	toluene	Indole m[d]	Catechol m[e]	Gram (KOH)[f]	*nah*	*phn*	*nah*	*phn*
HHSP3	phe	+	+	-	-	+	-	-	-	-	+	-
HSP2	phe	+	+	-	-	+	+	-	+	-	+	-
HSP4	phe	+	+	-	-	+	-	-	+	-	+	-
1P1	phe	-	+	-	-	-	-	+	-	-	-	-
1P2	phe	-	+	-	-	-	-	-	-	-	-	-
1P3	phe	+	+	-	-	-	-	-	-	-	-	-
1P4	phe	+	+	-	-	-	-	-	-	-	-	-
2Pa	phe	+	+	-	+	-	+	-	-	-	-	-
2Pb	phe	+	+	-	-	-	-	-	-	-	-	-
2Pc	phe	-	+	-	-	-	-	-	-	-	-	-
2Pe	phe	-	+	-	-	-	-	-	-	-	-	-
2Pf	phe	+	+	-	-	-	-	-	-	-	-	-
2Pg	phe	+	+	-	-	-	-	-	-	-	-	-
2Ph	phe	-	+	-	-	-	-	-	-	-	-	-
3Pa	phe	+	+	-	+	-	-	-	-	-	-	-
3Pb	phe	+	+	-	+	-	-	-	-	-	-	-
3Pc	phe	+	+	-	-	-	-	+	-	-	-	-
3Pd	phe	-	+	-	-	-	-	+	-	-	-	-
5Pa	phe	+	+	-	-	-	-	-	-	-	+	-
6Pd	phe	-	+	-	-	-	-	+	-	-	-	-
7Pa	phe	+	+	-	-	-	-	-	-	-	-	-
8Pa	phe	+	+	-	-	-	-	-	+	-	+	-
RP001	phe	-	+	-	-	-	-	-	-	-	-	-
RP002	phe	-	+	-	-	-	-	-	+	-	+	-
RP003	phe	-	+	-	-	-	-	-	-	-	-	-
RP004	phe	-	+	-	-	-	-	-	+	-	+	-
RP005	phe	-	+	-	-	-	-	-	+	-	+	-

[a] The first number or letter combination (1–8, LS, HS, HHS) represents the soil sample from which that strain was isolated (see Table 6.3). RP- strains were isolated from contaminated sludge at Rotowaro, Huntly.

[b] The substrate with which the strain was isolated; nap – naphthalene, phe – phenanthrene.

[c] The substrate was supplied as crystals in the lid of a petri dish of solid MM. Phenanthrene growth was also tested using overlay plates.

[d] Strains were grown on ½PC agar until visible colonies formed. Indole crystals were placed in the lid of the petri dish and after 2–3 days further incubation strains were scored positive if a dark blue pigment accumulated within colonies.

[e] Colonies grown on ½PC agar were sprayed with a dilute (100 mM) aqueous solution of catechol. Strains were marked positive if a yellow *meta* cleavage product of catechol accumulated within colonies after 15 mins.

[f] The KOH test was used as a rapid indicator of the Gram-type of each strain.

[g] To each well was added 100 µl of genomic DNA diluted to 30 µg ml^{-1}.

Table 6.2 Properties of 77 naphthalene or phenanthrene degrading strains isolated from contaminated soil from the Waikato/South Auckland region. Rhonda Fraser and David Hunter are acknowledged for the isolation of all strains except RP001–5, and the substrate utilisation screening and biochemical tests (catechol, indole, KOH) for all strains in this table.

Soil sample	Description
1	Contaminated soil near leaking oil barrels, Leslie Coaches, Gordonton.
2	Diesel-contaminated soil in gravel carpark, Caltex Service Station, Clyde St, Hamilton.
3	Oil- or petrol-contaminated soil at car crash/dumping site, Gordonton Rd.
4	Oil-contaminated soil adjacent State Highway 1, Ramarama.
5	Oil-contaminated soil at roadworks site near railway tracks, Mercer.
6	Soil adjacent NZ Vanlines driveway, Mangere.
7	Sample from an oily sludgy puddle, adjacent State Highway 1, Ramarama.
8	Garden soil adjacent Papakura train/bus terminal carpark.
LS	Sample from Hamilton town gas site containing low levels PAHs and BTEX.
HS	Sample from Hamilton town gas site containing high levels PAHs and BTEX.
HHS	Sample from Hamilton town gas site containing medium levels PAHs and BTEX.
A	Agricultural soil (maize) from Horotiu, low humic/organic content
B	Pristine soil from Garrett's Bush, high humic/organic content.

Table 6.3 Description of soil samples from the Waikato/South Auckland region from which PAH degrading bacteria were isolated. DNA extracts from these soils (except Soil 8) were screened for the presence of *phnAc*- and *nahAc*-like genes.

6.4 Detection of *phn* and *nah* genes in soil

Screening of the 77 strains isolated from local contaminated soils based on naphthalene or phenanthrene utilisation revealed the *nah* genotype could account for 28 (36%) of the isolates, but the *phn* genotype was not present at all in the culture collection. Does the absence of *phn*-like genes amongst cultivatible bacteria indicate this genotype is extremely rare in the environment, or is it merely a reflection of the bias associated with culturing techniques?

To investigate the presence of the *phn* genotype in the nonculturable microbial community, PCR was used to amplify the *phnAc* gene from DNA extracted directly from soil. Such a DNA preparation will ideally be representative of all organisms present in a given soil sample, and allows detection of specific genes without prior culturing. Soils were also screened for the presence of the *nahAc* gene so that the presence of *phn* relative to *nah* could be determined.

DNA preparations from soil were obtained using the method of Berthelet *et al.* (1996) as described in Chapter 2. For each PCR experiment 1 µl of a DNA prep was used as the template in a 50 µl reaction volume. PCR cycling conditions consisted of 40 cycles with an annealing temperature of 55°C. Several cycling regimes were tested and 40 cycles was found to be both necessary and sufficient to detect the *phnAc* or *nahAc* gene in soil-extracted DNA. To increase sensitivty and to confirm

the identity of amplified products, PCR experiments were Southern blotted and hybridised with [^{32}P]-labelled *phnAc* or *nahAc* probes. Due to the sensitivity of the PCR method contamination problems were experienced which resulted in amplification products present in negative control tubes. The use of barrier pipette tips and the replacement of all stocks and buffers resolved the contamination problem and reliable results were obtained thereafter.

DNA extracts were prepared for all soil samples described in Table 6.3, except Soil 8. PCR experiments used either the P8073/P9047 primers to amplify DNA homologous to *phnAc*, or the nahAcfor/nahAcrev primers to amplify *nahAc*-like sequence. The results are shown in Figures 6.2 and 6.3, and reveal *phnAc* was detected in 8 of 12 soil samples, and *nahAc* was detected in 7 samples. Most soils showed the presence of both genotypes, but in Soil 1 only *nahAc* was present and in Soils 4 and 7 only *phnAc* was detected. Neither *phnAc* or *nahAc* was present in the pristine (non-contaminated) soils A and B.

These results show that despite the absence of *phn* genes in cultured PAH degrading bacteria, the *phn* genotype is prevalent in the microbial communities of contaminated soils. Furthermore, in two of the soils *phnAc* was detected, but not *nahAc*, suggesting that in these communities bacteria of the *phn* genotype represent a significantly greater proportion of the PAH degrading community than those with a *nah* genotype.

Although the intensities of the *phn* and *nah* PCR products were similar, no inference can be made of the relative initial amounts of *phn* and *nah* template DNA due to the high number of cycles used in the amplification. A method based on competitive PCR would be necessary to estimate initial amounts of *phn* and *nah* template DNA in a soil DNA extract, and so estimate the relative sizes of *phn*- and *nah*-genotype populations in these microbial communities.

As an indication of the ability of the *phn* primers to amplify sequences homologous to *nah*, and *vice versa*, the membranes shown in Figures 6.2 and 6.3 were stripped and hybridised with the opposite probe. None of the *phn* or *nah* PCR products showed homology to the other probe, showing that the *phn* and *nah* primers were specific for the genotype for which they were designed.

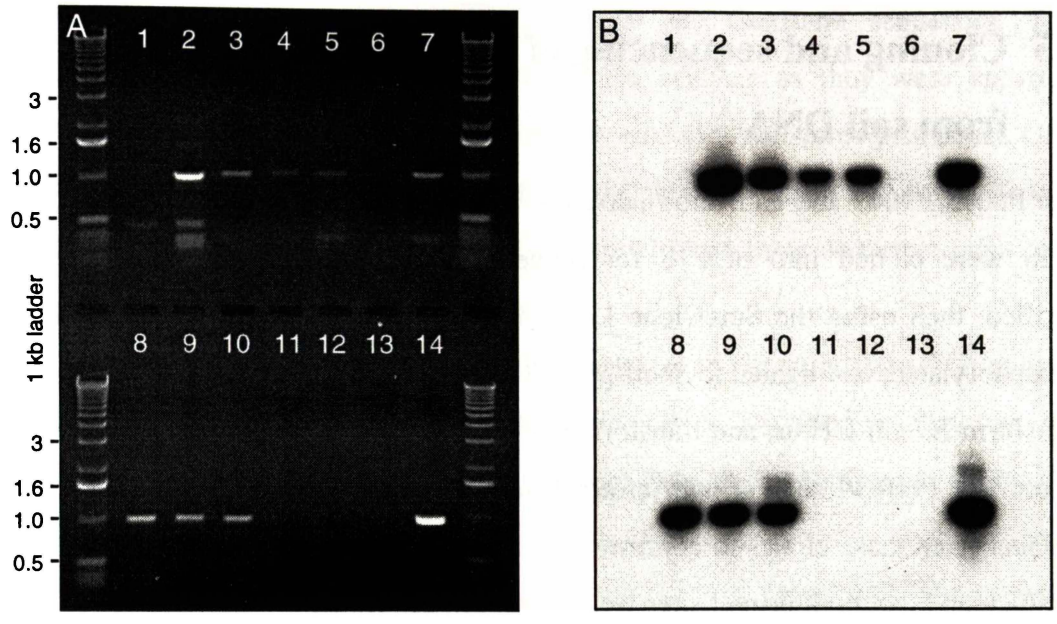


Figure 6.2 (A) Visualisation by electrophoresis of PCR products amplified from DNA preparations from various soil samples using the *phnAc* primers (P8073/P9047). Lanes 1–12 show PCR from soil samples described in Table 6.3: lanes 1–7, Soils 1–7; 8, Soil LS; 9, Soil HS; 10, Soil HHS; 11, Soil A; 12, Soil B. Lane 13 is a negative PCR control (no DNA template added), lane 14 is a positive PCR control (1 µl RP007 genomic DNA). (B) Southern blot of the gel shown in (A) hybridised to a [^{32}P]-labelled *phnAc* probe. The lanes are the same as for (A).

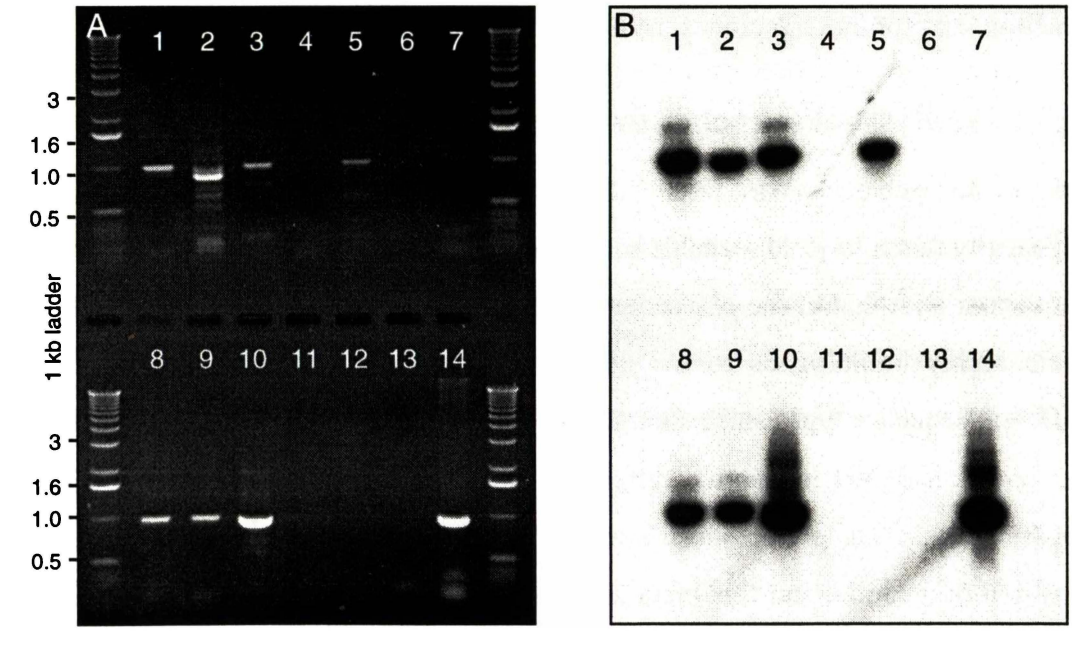


Figure 6.3 (A) Visualisation by electrophoresis of PCR products amplified from DNA preparations from various soil samples using the *nahAc* primers (nahAcfor/nahAcrev). Lanes 1–12 show PCR from soil samples described in Table 6.3: lanes 1–7, Soils 1–7; 8, Soil LS; 9, Soil HS; 10, Soil HHS; 11, Soil A; 12, Soil B. Lane 13 is a negative PCR control (no DNA template added), lane 14 is a positive PCR control (1 µl *P. putida* G7 genomic DNA). (B) Southern blot of the gel shown in (A) hybridised to a [^{32}P]-labelled *nahAc* probe. The lanes are the same as for (A).

6.5 Cloning and sequencing of *phn* and *nah* PCR products from soil DNA

The P8073/P9047 and nahAcfor/nahAcrev PCR products amplified from Soil 2 and HHS were cloned into pUC18 for sequence analysis. PCR products were gel-purified, then using the SureClone Ligation Kit (Pharmacia) were blunt-ended and phosphorylated, and ligated to *Sma*I-pUC18 vector. This ligation reaction was used to transform *E. coli* DH5α, and transformants were screened for the presence of a 1 kb insert. The P8073/P9047 or nahAcfor/nahAcrev primers were used to amplify a PCR product from these clones to confirm the insert was derived from the origninal PCR product, and that both primer sites were present in the clone.

Five clones with desired inserts were selected from each cloning experiment, ie. ten *phn* clones and ten *nah* clones in total. The *phn* clones from Soil 2 were designated P2-1–P2-5, and from Soil HHS were designated PHS-1–PHS-5. The *nah* clones were designated N2-1–N2-5 and NHS-1–NHS-5 in the same way. Nucleotide sequence from the *phn* clones was determined using the P8073 primer for the sequencing reaction, and for the *nah* clones the nahAcfor primer was used.

6.5.1 Analysis of sequence data

Successful sequence was obtained from all *phn* clones except PHS-1, which repeatedly failed to yield readable sequence data. An alignment of 600 bp of the nine sequences with that of the *phn* sequence of *Burkholderia* sp. RP007 revealed that all were highly homologous to the *phn* sequence. The nine clones represented six different sequence types since a number of the cloned sequences were identical; these sequences had 0–9 nucleotide differences relative to the RP007 *phn* sequence. In most cases the nucleotide substitutions relative to RP007 *phn* occurred at the third position of a codon, and this mutation was therefore silent resulting in no change of the amino acid sequence. The most divergent sequence was PHS-5, which showed 9 nucleotide differences relative to RP007 *phn*, and all of these were silent mutations.

All clones of *nah* PCR products gave successful nucleotide sequence results except NHS-4. The sequence of four *nah* clones from Soil 2 (N2-1–N2-3, N2-5) were very similar but showed no homology to the corresponding region of the *P. putida* G7

nahAc gene, and no significant homology to any GenBank sequences. These sequences were therefore disregarded from the analysis as they were apparently unrelated DNA coincidentally amplified by the nahAcfor/nahAcrev primer set. Of the other five clones, two sequences were identical (NHS-2, NHS-3) and two differed by only one base (NHS-1, NHS-5). The homology of these three sequence types to the *P. putida* G7 *nah* sequence was 80–94%.

A dendrogram (Figure 6.4) was constructed based on an alignment of a 577 bp region of the *phn* and *nah* PCR products from soil DNA. Included in this tree were sequences from ten *nah*-type strains isolated for their ability to degrade naphthalene and phenanthrene (described in section 6.3), and the corresponding sequence region from characterised PAH degrading strains (available from GenBank). The sequences in this analysis are arranged into three distinct phylogenetic groups. The highly homologous *phn* group includes the RP007 *phn* sequence and the six amplified *phn* PCR products, which share >98.5% nucleotide similarity. Collectively this group shows 58–60% homology to the *nah*-like sequences.

Sequences similar to *nah* are organised into two groups, designated *nah* and *dnt/ntd*, which are 78–80% similar. The *nah* group includes the *nah*, *ndo*, *pah*, and *dox* genes, and the *dnt/ntd* group consists of the divergent *dnt*, *ntd*, and *nag* genes. The sequences from the cultured isolates and from directly amplified *nah* PCR products fall into both *nah* and *dnt/ntd* groups. Amongst the *nah* group the new sequences do not show greater diversity than the characterised genes of that group, but rather cluster between the *nah*, *ndo*, *pah*, and *dox* sequences. In the *dnt/ntd* group, the sequences from both the cultured strains and from the PCR products show a greater diversity than that between the *dnt*, *ntd*, and *nag* sequences, and the inclusion of these sequences increases the diversity of this group.

The phylogenetic distribution of the sequences amplified by PCR using the *phn* and *nah* primer sets may reflect the design of these primers. The nahAcfor/nahAcrev primer binding sites correspond to highly conserved regions in seven *nah*-like sequences, including *ntd* and *dnt*. The sequences amplified by these primers would therefore be expected to reflect the diversity of the *nah*-like group, although it is worth noting that the nahAcfor/nahAcrev primers would probably not amplify a product from *dnt* and *ntd* sequences due to five mismatches at the nahAcrev primer binding site. In contrast, the *phn* primers were designed based on the RP007 *phn*

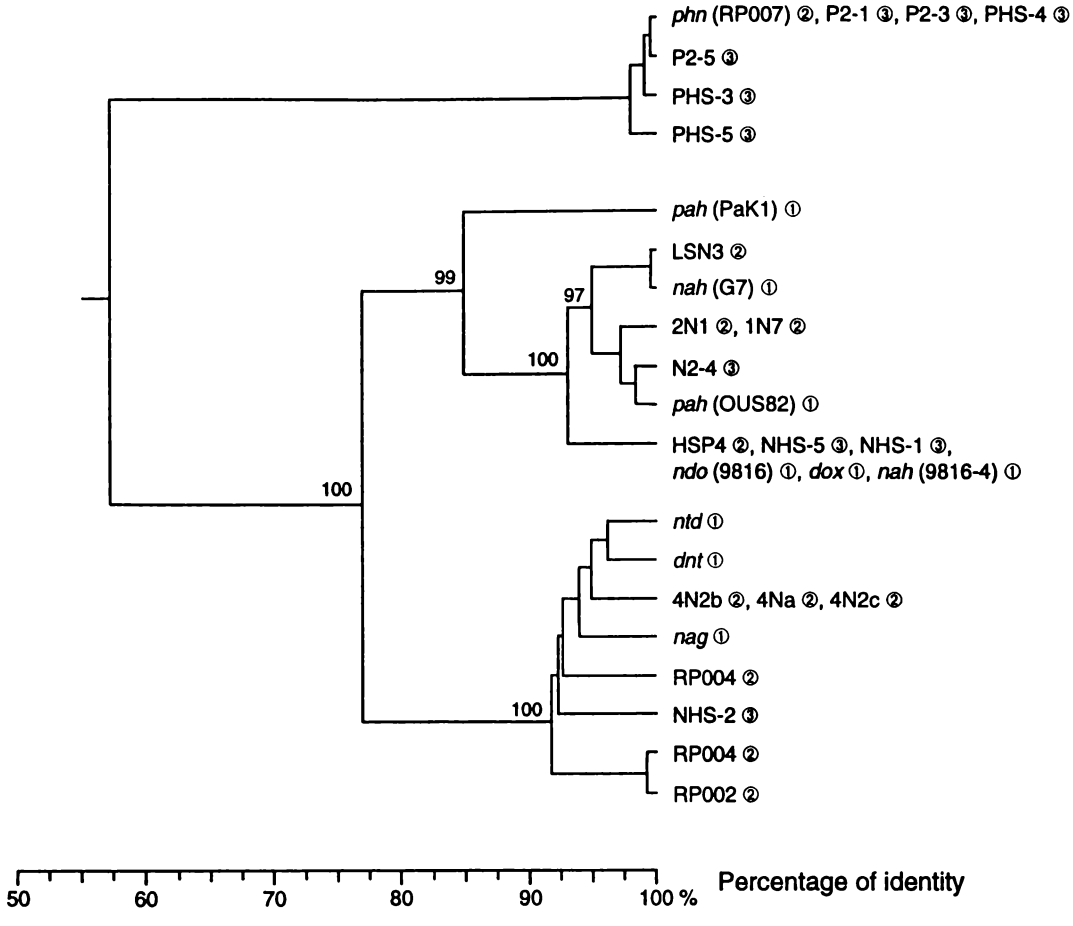


Figure 6.4 Phylogeny of the partial initial dioxygnase iron-sulfur protein α subunit gene from cultured PAH degrading strains and directly amplified from soil DNA. Numbers indicate the source of the nucleotide sequence: ①, characterised PAH degrading strains (sequences obtained from GenBank); ②, isolates from Waikato/South Auckland soils cultured for their ability to degrade PAHs (described in section 6.3); ③, PCR products amplified directly from soil DNA using primers specific for *phnAc* or *nahAc*-like genes. The dendrogram is an unrooted tree constructed using the *phn* branch as an outgroup, and was based on an alignment of 577 bp of nucleotide sequence. Sequences with 99–100% homology are shown on the same branch, and the bootstrap values of major branch points are included.

sequence only, and it is therefore unknown whether these primer sites correspond to conserved or variable regions of the *phn*-type sequence. The high homology of the amplified *phn* sequences indicates the *phn* primers are specific for sequences highly similar to RP007 *phn*.

The lack of cross-hybridisation of the *phn* PCR products from soil DNA to a *nah* probe, and *nah* PCR products to a *phn* probe, suggested that the *phn* and *nah* genotypes are distinct groups which do not overlap. Since the detection limits of the hybridisation protocol are known to be 78–80% nucleotide homology, the mixed

population of *phn* and *nah* PCR products are therefore <80% similar. This is supported by the phylogenies of sequence from cloned *phn* and *nah* PCR products which conform to these distinct groups.

There are a number of methodological limitations which effected the results obtained from this experiment. Despite selecting ten clones for each primer set, the number of different sequences which could be used for the phylogenetic analyses was considerably fewer. This was due in part to the presence of clones with identical insert sequences, which presumably arose either from the cloning of identical PCR fragments, or the replication of a transformant cell during the transformation procedure. The cloning of PCR products unrelated to the target sequence which were coincidentally amplified by the nahAcfor/nahAcrev primers also reduced the number of usable sequences. Neither of these problems can be easily avoided since this information only becomes apparent after analysis of the sequence data, but to decrease the chance of cloning identical sequences all clones could be derived from PCR products from different soils.

The high homology and in some cases identical sequence of the cloned *phn* PCR products to the RP007 *phn* sequence suggested the possibility that these PCR products were the result of contamination by RP007 *phn* DNA. A number of factors suggest this is not the case and that these sequences do represent real genes amplified from soil DNA. Firstly, the PCR experiments from which the *phn* products were cloned included negative controls which were analysed by Southern hybridisation and showed no contamination of the PCR reagents. Secondly, most mutations relative to RP007 *phn* were silent, in particular the nine mutations of PHS-5 all occurred in the third base of the codon, and if these substitutions were random errors introduced by *Taq* polymerase they would be expected to occur equally at all three codon positions. A further consideration is that the parallel experiment in which *nah*-like sequences were amplified showed high diversity relative to the *P. putida* G7 *nah* sequence, and this validates the *phn* result.

To show that the *phn* PCR results are real it would be necessary to empirically demonstrate the fidelity of *Taq* under similar PCR conditions. This could be achieved by PCR from a low quantity of RP007 genomic DNA using the same conditions (40 cycles) and cloning and sequencing a number of PCR products obtained in this way. To confirm that the cloned PCR products represented the diversity of the total PCR

product, ie. that the *phn* PCR product is highly conserved, the total PCR product could be analysed by denaturing gradient gel electrophoresis (DGGE) or temperature gradient gel electrophoresis (TGGE) which would resolve individual variants of the total PCR mix. This could also be used to show the characterised *nah* PCR products are representative of the diversity of the total PCR mix.

CHAPTER 7

Discussion

7.1 The *Burkholderia* genus

Strain RP007 was assigned to the *Burkholderia* genus based on the results of biochemical tests, cellular fatty acids analysis, and nucleotide sequencing of 1.35 kb of the 16S rRNA gene. *Burkholderia* is a recently described genus to which a number of strains previously designated *Pseudomonas* have been reassigned. It had earlier been recognised that the type species for the genus, *B. cepacia* (initially designated *Pseudomonas multivorans*), constituted a separate group within the aerobic pseudomonads due to its extraordinary nutritional versatility (Stanier *et al.*, 1966), a trait that has become characteristic of members of the *Burkholderia* genus. The results of rRNA/DNA hybridisation experiments indicated *B. cepacia* was not closely related to the fluorescent pseudomonads; this was later confirmed by 16S rRNA nucleotide sequence analysis which confirmed *B. cepacia* belongs to the β-subclass of the *Proteobacter* rather than the γ-subclass which includes *P. aeruginosa*, the type species for *Pseudomonas* (Palleroni & Holmes, 1981).

Subsequently Yabuuchi *et al.* (1992) proposed the transfer of seven species of the genus *Pseudomonas* homology group II to a new genus *Burkholderia*, named after the American bacteriologist W. H. Burkholder, who first discovered the etiological agent of rotten onion. Based on the results of a polyphasic study (phenotypic characteristics, assimilation profile, cellular lipids profile, cellular fatty acids profile, and 16S rRNA sequence) of the type or reference strains the following species were transferred to the new genus: *Pseudomonas mallei*, *P. pseudomallei*, *P. cepacia*, *P. solanacearum*, *P. caryophylli*, *P. gladioli* and *P. pickettii*. The genus was expanded when Urakami *et al.* (1994) recharacterised a selection of plant-associated bacteria and proposed the transfer of *Pseudomonas planterii* and *P. glumae* to the genus *Burkholderia*, and established a new species, *Burkholderia vandii*, for a strain isolated from the roots of *Vanda* species. Gillis *et al.* (1995) created a new *Burkholderia* species, *B. vietnamiensis*, for a collection of nitrogen-fixing strains isolated from rice rhizosphere soil in Vietnam. On the basis of polyphasic taxonomic results and the previous studies of other workers including Yabuuchi *et al.* (1992) and Urakami *et al.* (1994), Gillis *et al.* (1995) provided an emended description the genus *Burkholderia*:

'cells are Gram-negative, nonfermentative, straight rods that have a single polar flagellum or a tuft of polar flagella.... Catalase is produced and oxidase activity varies between species. The cellular fatty acids are characterised by the presence of 3-hydroxy $C_{16:0}$. The type strains of several species are characterised by the presence of two types of ornithine lipids. Most species grow at 40°C. All species can grow with the following substrates as sole carbon sources: glucose, glycerol, inositol, galactose, sorbitol, and mannitol. Some species are pathogenic for humans, animals, or plants. Isolated from plant material, soil, or clinical samples. Can be recognised on the basis of 16S rRNA characteristics (sequence and/or DNA-rRNA hybridisation data). Most strains accumulate polyhydroxybutyrate as carbon reserve material and are capable of *ortho* cleavage of protocatechuate. The G+C content is 59.0–69.5 mol%. The type species is *B. cepacia* (Yabuuchi *et al.*, 1992).'

Gillis *et al.* (1995) removed [*B.*] *solanacea* and [*B.*] *pickettii* from the *Burkholderia* genus since in these species the 3-hydroxy $C_{16:0}$ fatty acid is absent or present in small amounts and so constitutes a separate fatty acid cluster.

B. cepacia has an interesting and unusual genomic organisation. Several *B. cepacia* isolates studied contained between two and four chromosomes with overall genome sizes in the range of 5 to 9 Mb (Lessie *et al.*, 1996). *B. cepacia* is also notable in harbouring an extensive array of insertion sequences (transposons) which are thought to contribute to the genomic plasticity of *B. cepacia* and play a role in evolution of novel catabolic functions (Lessie *et al.*, 1996).

These observations are borne out by the study of *B. cepacia* AC1100 which degrades 2,4,5-trichlorophenoxyacetic acid (2,4,5-T). This strain is comprised of five replicons of 0.15–4.0 Mb, and the two *tft* gene clusters for 2,4,5-T degradation are carried on different replicons and are located adjacent to three known insertion sequences (Hübner *et al.*, 1998). *B. cepacia* AC1100 acquired its ability to degrade 2,4,5-T under strong selective pressure in chemostat culture, presumably through the recruitment of genes present in the chemostat consortium and their integration into its genome. Growth of AC1100 in the absence of 2,4,5-T leads to mutants which have lost the ability to catabolise this substrate, and analysis of 2,4,5-T-negative mutants

shows the catabolic genes are deleted and rearranged at high frequencies.

This type of genome organisation may explain the variable results obtained when probing RP007 plasmid or total genomic DNA preparations with the *phn* catabolic genes. The distinction between chromosomal- and plasmid-borne genes may be blurred given that a small chromosome may behave like a large plasmid during isolation procedures. The location of the three *phn* operons may ultimately be less significant if they are carried on transposon elements, and given that the genome of *Burkholderia* spp. is such a dynamic system.

Many *Burkholderia* spp. have been implicated in the biodegradation of diverse aromatic compounds, which has led to a reputation of this genus for catabolic versatility. The first report of a member of this genus which degrades PAHs was by Grifoll *et al.* (1995b). *B. cepacia* F297 is a fluorene degrading isolate which also degrades two- and three-ring PAHs and dibenzothiophene. Mueller *et al.* (1997) isolated degradative strains from geographically diverse soils using PAH-enrichment cultures, and found phenanthrene enrichments from American and Norwegian soils yielded a number of *B. cepacia*-like isolates. All of these strains were only able to grow on phenanthrene and lower molecular weight PAHs. The authors suggest some factor of *B. cepacia* may explain the frequent observation that this species degrades phenanthrene but no larger PAHs. Such factors could include a competitive growth advantage of *B. cepacia* under these conditions, cell surface characteristics that allow better association of phenanthrene with the degradative enzymes, or a cultural bias such as growth medium or incubation temperature that leads to enrichment of the same genera. Members of the *Burkholderia* genus are able to degrade PAHs larger than phenanthrene; Juhasz *et al.* (1997a; b) described a *B. cepacia* strain (VUN10001) with the ability to degrade a wide range of aromatic compounds including high molecular weight PAHs. This strain could grow on PAHs of up to five rings, *n*-alkanes, chlorinated- and nitro-aromatic compounds. Several other *Burkholderia* strains have been described which degrade aromatic compounds (Beil *et al.*, 1997; Erickson & Mondello, 1992; Haak *et al.*, 1995; Shields *et al.*, 1995; Suen *et al.*, 1996).

Phylogenetic analysis of 16S rRNA nucleotide sequence of *Burkholderia* strains which degrade aromatic compounds reveals these strains form a cluster distinct from other *Burkholderia* species (Figure 3.1). *B. cepacia* AC1100 also groups with this

cluster, and Lessie *et al.* (1996) suggested this line of aromatic degraders differs sufficiently from other *Burkholderia* species to constitute a distinct species.

Burkholderia has been targeted as an indicator of environmental perturbance. Using a DNA probe specific for the 16S rRNA gene of *B. cepacia*, Lemke *et al.* (1997) found this species was present at a range of sites indicating adaptation to a wide variety of environmental conditions, but that populations of *B. cepacia* did not increase in areas contaminated with aliphatic and aromatic hydrocarbons.

The G+C content of the sequenced regions of *Burkholderia* sp. RP007 is: pB1 – 58.9%, pH1 – 57.6%, and pH2 – 55.4%. These values fall outside the G+C range given in the definition for *Burkholderia* of 59–69.5% (Gillis *et al.*, 1995), and may reflect the origin of these catabolic operons was from bacteria characterised by a lower G+C content.

7.2 The PAH upper pathway *phn* locus

The upper pathway *phn* genes cloned on pB1 constitute a novel operon for PAH catabolism. Although five of the *phn* catabolic genes are phylogenetically grouped with the corresponding *nah*-like genes, in each case the *phn* gene is consistently an out group showing only 36–73% amino acid homology to the highly homologous *nah*-like genes. The dihydrodiol dehydrogenase, *phnB*, shows 15% greater amino acid identity to the corresponding gene from biphenyl degradation pathways than to the *nah*-like dehydrogenases. The extradiol dioxygenase of the *phn* pathway, *phnC*, shows an even more interesting phylogeny. This gene has no homology to any extradiol dioxygenases from naphthalene or biphenyl pathways, but is a member of the recently described Class III extradiol dioxygenases. Previously characterised members of this class are from pathways for the degradation of substituted single-ring aromatic compounds such as phenylproprionate, and *phnC* is the first report of a Class III extradiol dioxygenase which cleaves a multi-ring aromatic substrate.

The gene order of the catabolic genes of the pB1 *phn* locus differs from the highly conserved order found in *nah*-like operons. The *phn* catabolic genes are arranged *FECDAcAdB* compared with the conserved order for *nah*-like loci of *AaAbAcAdBFCQED* (Zylstra *et al.*, 1997).

A further feature which distinguishes the *phn* locus from *nah*-like systems is the

genes encoding the *phn* initial dioxygenase. The *AaAb* genes which encode the ferredoxin and reductase components of the multicomponent PhnA dioxygenase are not present on the *phn* locus. It is expected that the PhnA dioxygenase requires these components since there was no evidence of electron transport functions apparent in the *phnAcAd* predicted amino acid sequence. It is possible the *AaAb* genes lie downstream of *phnB*, as shown for the catabolic genes of a *Comamonas testosteroni* strain (Goyal & Zylstra, 1996), but no homology could be found to ferredoxin or reductase genes in the 560 bp downstream of the *phnB* stop codon to the end of the pB1 fragment. The ferredoxin and reductase genes for electron transport functions for PhnA may be located elsewhere on the RP007 chromosome or may be supplied by cellular housekeeping genes. That the PhnA dioxygenase was fully functional in *E. coli* was demonstrated by the indole-positive phenotype of pB1 and the ability of pB1 to transform naphthalene to salicylate. Ensley *et al.* (1987) found the cloned *nahAcAd* genes (ISPα and ISPβ) of *P. putida* G7 could oxidise indole to indigo and Simon *et al.*, (1993) suggested *E. coli* contains a nonspecific reductase which provides the electron transport functions for recombinant dioxygenases. Expression of an active PhnA suggests *E. coli* must also contain a nonspecific reductase component which complements PhnA ISPα and ISPβ.

Three experimental approaches were used to provide evidence to support the nucleotide sequence data which suggested the pB1 fragment contained an operon encoding the upper pathway for PAH catabolism. *E. coli* carrying the pB1 plasmid was shown to convert naphthalene to salicylic acid, which was predicted based on the catabolic genes present on pB1. The enzyme activity of the *phnC* extradiol dioxygenase showed a clear substrate preference for 1,2-dihydroxynaphthalene over catechol, methylcatechols, and 2,3-dihydroxybiphenyl. Finally, it was shown by RT-PCR experiments that expression of the *phn* genes was linked to growth on naphthalene or phenanthrene. RT-PCR was also used to show that the *phn* catabolic genes are expressed on a single mRNA transcript which justifies referring to the *phn* locus as an operon.

The phylogeny of the pB1 *phn* genes shows homology of a patchwork nature. Although five of the *phn* catabolic genes bear most similarity to *nah*-like genes, *phnB* and *phnC* are more distantly related to their *nah*-like counterparts. The *phn* gene order is quite different also, and these factors suggest the *phn* and *nah* genotypes

have a distant evolutionary relationship. If the *phn* and *nah* genotypes have evolved from a common ancestor, such differences suggest the divergence of these groups occurred a long time ago, so that subsequent gene evolution and reorganisation, possibly through the recruitment of other genes, has lead to the differences observed now. However, evolutionary processes do not adequately explain the level of divergence of *phn* to *nah*. Catabolic operons tend to move between bacteria and evolve as entire units, and although sequence divergence of individual genes may occur to reflect changes in substrate specificity, the same general homology patterns are seen between related loci (Williams & Sayers, 1994). Recombination events may result in insertions of genes into groups of genes, but this would not explain the differences between *phn* and *nah*. An alternative scenario is that the *phn* locus was compiled by the recruitment of individual genes from a variety of catabolic pathways, possibly including *nah* and *bph* loci, and subsequently evolved to specifically catabolise PAHs. Similarities to *nah* and *bph* genes therefore reflect the origin of these genes, but the gene arrangement and composition is novel.

7.3 The *phn* lower pathway genes

Burkholderia sp. RP007 possesses at least two nonhomologous *meta* pathway operons which belong to different evolutionary groups. One of these, cloned on pH1, may be the lower pathway for PAH metabolism. The other was cloned on pH2 and appears to be a separate pathway possibly involved in degradation of benzoate due to the presence of a multicomponent dioxygenase upstream of the *meta* pathway genes. Two nonhomologous independently regulated catechol 2,3-dioxygenase genes were reported to reside on the same TOL plasmid (pWW15) of *P. putida* mt-15 (Keil *et al.*, 1985), presumably indicating the presence of two different *meta* pathways. This suggests it is not unusual for a single strain to possess multiple divergent pathways for the same or similar aromatic substrate, and supports the theory that such operons are acquired by bacteria as discrete units (Williams & Sayers, 1994).

7.4 Transcriptional regulators for aromatic catabolism

Predicted amino acid sequence data suggest *phnR* and *phnS* encode a σ^{54}-dependent regulator (NtrC-like) and a LysR-type transcriptional regulator (LTTR), respectively. Both classes of regulators are positive transcriptional activators which, in the presence of specific coinducers, activate transcription from specific promoters.

Many pathways for aromatic catabolism have been shown to be transcriptionally regulated, and in many cases the regulatory genes/proteins have been described. The transcriptional activators for aromatic pathways are usually classed as σ^{54}-dependent transcriptional activators or LysR-type transcriptional regulators. Exceptions are XylS of the TOL pathway of *P. putida* mt-2 which is a member of the XylS/AraC family of regulatory proteins (Gallegos *et al.*, 1997), and members of the two-component signal transduction family (eg. TodS-TodT of *P. putida* F1) (Lau *et al.*, 1996). Regulators of both the σ^{54}-dependent and LysR-type families have been extensively studied to elucidate the mechanisms by which these proteins exert transcriptional control over the operons they regulate.

7.4.1 σ^{54}-dependent transcriptional activators

The σ^{54}-dependant family of transcriptional activators controls the expression of a variety of cellular processes in many different eubacteria. NtrC and NifA from *Kleibsiella pneumoniae* are two archetypal members of this family which regulate genes for nitrogen assimilation and fixation and have been well studied, contributing to much of the current understanding of σ^{54}-dependent regulators (Buikema *et al.*, 1985). This class of transcriptional activators is also referred to as 'NtrC-type' as a reflection of this. This family is now known to include regulators involved in a range of cellular processes including specific transport systems, utilisation of alternative carbon and energy sources, production of extracellular structures, and virulence determinants (Shingler, 1996).

σ^{54}-dependent regulators require the alternative sigma factor σ^{54} encoded by *rpoN* (*ntrA*) and its homologues. σ^{54} confers unusual properties on the holoenzyme RNA polymerase utilising this factor ($E\sigma^{54}$), primarily the ability to recognise and initiate transcription from a distinct class of –24/–12 bacterial promoters that differ from the usual –35/–10 -type promoters recognised by RNA polymerase utilising the σ^{70}

'housekeeping' sigma factor ($E\sigma^{70}$) (Merrick, 1993).

σ^{54}-dependent regulators are modular in structure and are composed of three functionally distinct domains: a C-terminal domain (domain D) containing a helix-turn-helix motif (AXXXXXXXXXXAAXXLG) for DNA binding, a central domain (domain C) responsible for activation of transcription by an ATPase activity mediated by an NTP binding site (GXXGXGK), and an N-terminal domain (domain A) which is thought to be the target for specific regulatory signals. Domains A and C are joined by a flexible minor domain, domain B.

DmpR, from *Pseudomonas* sp. CF600, and XylR, from the TOL plasmid of *P. putida* mt-2, are σ^{54}-dependent regulators which regulate the expression of phenol-cresol and toluene-xylene catabolism, respectively (Abril *et al.*, 1989; Shingler *et al.*, 1993). These peptides are part of a class of σ^{54}-dependent regulators which respond to the presence of aromatic substrates (or metabolites) of the pathways they control.

Transcriptional activation by DmpR/XylR-type σ^{54}-dependent regulators is by an elaborate process based on derepression as a control mechanism. The $E\sigma^{54}$ RNA polymerase binds to the promoter region and forms a stable closed promoter complex. Unlike $E\sigma^{70}$, $E\sigma^{54}$ is isomerisation incompetent, and is unable to proceed to form an open transcriptional complex which allows transcription, since this process is thermodynamically unfavourable. ATP hydrolysis by the ATPase activity present on the C-domain of the σ^{54}-dependent regulator provides the necessary free energy for the isomerisation process (Wedel & Kustu, 1995). However, in the absence of effector molecules for the σ^{54}-dependent regulator, domain A interacts with domain C and represses the ATPase activity, thereby preventing formation of the open transcriptional complex at the promoter (Ng *et al.*, 1996). The binding of an effector molecule to domain A imposes a conformational change on domain A which prevents interaction with domain C and allows ATPase activity.

Other classes of σ^{54}-dependent regulators have a similar mechanism of action but do not interact directly with an effector molecule. Members of the NtrC subgroup are the response proteins of two-component regulatory systems and are activated by phosphorylation. A sensory histidine kinase senses the environmental signal (eg. nitrogen-limiting conditions) and activates the response protein by phosphorylation at a conserved aspartate residue located in the A-domain. Phosphorylation at this

residue is necessary for ATPase activity. Another class of σ^{54}-dependent regulators typified by NifA are regulated by an antagonistic protein-protein interaction which, under certain conditions, interferes with the activation properties of NifA (Shingler, 1996).

σ^{54}-dependent regulators are unusual in that they bind to specific sequences, often inverted repeats, located 100–200 bp upstream from the promoters they regulate; these are called upstream activating sequences (UASs) or enhancer-like elements (ELEs) (Morett & Segovia, 1993). The orientation of these upstream sequences on the DNA helix relative to that of the promoter is critical. Close physical proximity and correct protein-protein interaction between the UAS-bound activator and the promoter/RNA polymerase closed transcriptional complex requires precise looping out of the intervening DNA.

7.4.2 LysR-type transcriptional regulators (LTTRs)

The LysR-R family of transcriptional regulators is a rapidly expanding group and thought to be the second most common type of transcriptional regulators in prokaryotes, after the two-component systems (Schell, 1993). LysR-type transcriptional regulators (LTTRs) are widely distributed in diverse genera of prokaryotes but have not been found in the arhaea or eukaryotes. LTTRs activate divergent transcription of linked target genes or unlinked regulons encoding extremely diverse functions. Much of the understanding of LTTRs has come from the study of NahR which is the regulatory protein for the *nah* and *sal* operons on the NAH7 plasmid of *P. putida* G7 (Schell, 1985; 1986; Schell & Sukordhaman, 1989; Schell & Wender, 1986).

Three functional regions are identified in LTTRs: an N-terminal region DNA binding domain (residues 1–65 numbered according to NahR) which employs a helix-turn-helix motif, domains involved in coinducer recognition and/or response (residues 100–173 and 196–296), and a domain required for both DNA binding and coinducer response (residues 227–253).

The mode of action of LTTRs is simpler than that of the σ^{54}-dependent regulators. In the absence of a coinducer molecule, the LTTR binds to the promoter(s) it regulates via a 15 bp dyadic sequence near the –65 promoter region. Binding of the coinducer

to the LTTR results in a conformational change of the LTTR causing extension of the binding region towards the –35 promoter region, and increased affinity for RNA that causes increased transcriptional initiation.

LTTRs are usually divergently transcribed from a promoter that is close to the regulated gene, although some LTTRs lack this divergent structure. The *nahR* gene is transcribed divergently and directly upstream of the *sal* operon, such that the divergent *nahR* and *sal* promoters are very close. NahR binds to both the *sal* and *nahR* promoters causing a three-fold negative autoregulation of *nahR*, and in the presence of a coinducer (salicylate) only the *sal* promoter is induced (Huang & Schell, 1991). Most LTTRs are multimeric *in vivo*; NahR is a tetramer, others are dimers. NahR appears to be membrane associated but most others are soluble cytoplasmic proteins.

7.4.3 A model for the role of PhnR/PhnS

The location and direction of transcription of *phnR* and the large intercistronic region between *phnR* and the other *phn* genes represents an arrangement similar to that of other aromatic catabolic operons regulated by a σ^{54}-dependent regulator (Ng *et al.*, 1995). *phnS* was shown to be cotranscribed with the *phn* catabolic genes, and although no supplementary evidence was obtained to confirm the role of *phnR* and *phnS*, a possible model for the regulatory functions of these genes can be formulated. According to this hypothesis, PhnR is expressed constitutively and in the presence of an inducing substrate stimulates transcription of the entire *phn* operon at a promoter situated upstream of *phnS*. PhnS is coordinately expressed with the other *phn* genes and in the presence of another inducing substrate acts to induce the lower pathway genes located distal to the upper pathway *phn* genes (possibly the *meta* pathway cloned on pH1). This model is similar to the proposed XylR/XylS regulatory system of the pWW0 TOL plasmid (Ramos *et al.*, 1997).

The PhnR/PhnS regulatory system of the *Burkholderia* sp. RP007 *phn* locus may provide the first observation of a σ^{54}-dependent regulator and a LysR-type regulator acting in concert to control transcription of the same overall pathway.

7.5 Detection of PAH catabolic genes in soil

The use of molecular biology techniques to detect specific nucleotide sequences in water or terrestrial environments is emerging as a powerful experimental tool. This approach side-steps the inherent deficiencies associated with classical microbiological techniques which are time consuming and, more importantly, are subject to limitations and biases since some types of bacteria are more amenable to culturing under laboratory conditions than others. Estimates of the proportion of all bacterial species that are culturable are 0.01–10% (Ferguson *et al.*, 1984; Torsvik *et al.*, 1990). DNA extracted directly from water, soil, or sediment samples would ideally be qualitatively and quantitatively representative of the microorganisms present in that sample, and therefore allows assessment of the presence and relative numbers of a given genotype, and the amplification of sequences from unculturable species.

The study of mercury resistance genes (*mer*) successfully applied molecular biology techniques to detect and amplify specific *mer* sequences from aquatic and soil environments (Barkay *et al.*, 1989; Bruce *et al.*, 1992). The extraction of DNA from soils for PCR amplification is inherently more difficult than from water samples due to substances such as humic acids which are coextracted with DNA and which are inhibitors of *Taq* DNA polymerase. Bruce *et al.* (1992) extracted DNA from soil by incubation of the sample with a buffer containing 1% SDS at 70°C, followed by precipitation of DNA and purification through a caesium chloride gradient, then extraction with phenol-chloroform, ethanol precipitation and finally repeated washing with 70% ethanol. This method was used in a study of genetic diversity of *mer* genes amplified from DNA directly isolated from soil and sediments, which showed that *mer* diversity of directly amplified sequences was significantly greater than between cultivated bacteria (Bruce *et al.*, 1995).

Regions of DNA involved in degradation of PAHs have been successfully amplified from DNA extracted directly from soils or sediments. Herrick *et al.* (1993) used primers based on the *nahAc* gene of *P. putida* G7 to amplify a product from two sediments with a demonstrable naphthalene mineralisation potential. Restriction fragment length polymorphisms (RFLP) of these products indicated some diversity of the amplified products relative to the *P. putida* G7 *nahAc* gene, and the inability to

amplify a product from a third sediment which was shown to have a naphthalene degrading population suggested divergence of *nahAc* sequence in this sample. Herrick *et al.* (1993) extracted and purified DNA from the sediment using a lysozyme/SDS/freeze-thaw protocol, followed by electrophoresis through an agarose gel containing polyvinylpyrrolidone (PVPP). PVPP is thought to bind humic compounds. After separation from the contaminating components of the sediment by the electrophoresis process, the DNA was excised from the gel and a portion added directly to a tube for PCR.

Berthelet *et al.* (1996) described a simpler and more rapid procedure for the extraction of DNA from soil which was used as a template for amplifying naphthalene catabolic genes using primers based on *ndoB* of *P. putida* 9816. This method also uses PVPP to remove humic compounds inhibitory to PCR, and was the method used for the experiments described in Chapter 6. Soil is shaken at high speed in the presence of zirconia/silicone beads and a NaCl/Tris/SDS lysis buffer. The crude lysate is passed through a spin column containing acid-washed PVPP. This DNA extraction technique was used to assess the degradative potential of petroleum hydrocarbon contaminated soils from the high Arctic. Primers based on the *ndoB* and *alkB* genes from the naphthalene degradative pathway of *P. putida* 9816 and alkane degradative pathway of *P. oleovorans* amplified products from some of the soils tested, indicating the presence of genes similar to *ndoB* and *alkB* in these soils.

A 2.1 kb probe including the *nahAcAd* genes of *P. putida* G7 was used to assess whether soils or sediments contaminated with aromatic or straight chain hydrocarbons were enriched for bacterial strains capable of degrading such compounds (Guo *et al.*, 1997). A *xylE* probe from *P. putida* mt-2 was used to represent degradative pathways for BTEX compounds (benzene, toluene, ethylbenzene, xylene) and a 23S rRNA probe was used to assess the presence of *Pseudomonas* strains. Probes were hybridised to DNA extracted from soil cores near the site of a leaking underground storage tank, and from soil microcosms enriched with gasoline at different concentrations. Levels of hybridisation were normalised and expressed as 'gene mass-equivalents' with respect to grams of soil or total DNA extracted. *nahAcAd* and *xylE* probes showed contaminated soils were enriched for these genes relative to the noncontaminated soil, but the 23S probe showed no significant difference. There was no linear correlation observed between the

concentrations of the hydrocarbons in the soil and the genes tested. The *xylE* gene was a more sensitive probe than *nahAcAd* for indicating contamination.

Although the presence of PAH catabolic genes can be readily confirmed in contaminated soils by hybridisation and PCR techniques, these results give no information about the rates of *in situ* expression of degradative genes necessary to predict rates of pollutant biodegradation. The *nah* genes for naphthalene degradation have been used as a model system to relate expression of *nah* genes to naphthalene mineralisation rates (Fleming *et al.*, 1993; Sanservino *et al.*, 1993b). Total RNA was purified from contaminated soils and mRNA transcripts of the *nahA* genes were quantified by a ribonuclease protection assay. *nahA* transcription levels correlated positively with [^{14}C]naphthalene mineralisation rates, soil naphthalene concentration, and *nahA* gene frequency. Induction of *nahA* transcription by the addition salicylate resulted in an increase in *nahA* mRNA in only one of four soils tested, and two of the soils were apparently already fully induced. In addition, a naphthalene-*lux* reporter system was used to determine the bioavailability of naphthalene within these soils, which showed that in most of the soils tested naphthalene was bioavailable, but in some soils in which naphthalene was present, it was apparently not bioavailable.

The results described in Chapter 6 pertaining to the environmental distribution of the *nah* genotype concur with these previously published studies. PAH degrading bacteria were readily isolated from hydrocarbon contaminated soils, and a third of isolates hybridised to a *nah* probe. *nahAc*-like sequences could be amplified from DNA extracted from most soils, but no *nahAc* products were observed in pristine (non-contaminated) soils.

The sequenced region of the initial dioxygenase ISPα gene from *nah*-type cultured isolates and from *nah* PCR products from soil DNA fell into two phylogenetic groups: *nah* and *dnt/ntd* (Figure 6.4). The *nah*-like sequences were similar to previously described *nah*, *ndo*, *pah*, and *dox* sequences, and their inclusion in this group did not contribute to greater diversity within this cluster. The sequences which aligned to the *dnt/ntd* group were less similar to the *dnt*, *ntd*, and *nag* sequences, and these new members expand the diversity of this group.

Relative to the sequences from strains cultured for their ability to degrade naphthalene or phenanthrene, the *nah*-like sequences amplified directly from soil DNA did not show an increased diversity. The use of this technique has not revealed

any greater diversity amongst the targeted gene than was detected using classical culturing methods. This finding was not similar to that of Bruce *et al.* (1995) who found greater diversity in directly amplified *mer* genes than in cultured strains. This may reflect that the *nahAc* primers were based on a consensus of *nah*-like sequences from cultured strains which would presumably tend to amplify similar sequences. It also confirms the *nah* group is highly conserved in the environment as well as amongst cultivated strains. This is the first case of the sequence characterisation of *nah*-like genes amplified directly from soil DNA.

None of the cultured isolates described in Chapter 6 showed homology to *phnAc* of *Burkholderia* sp. RP007, but homologous sequences were detected in DNA extracted from eight of the soils studied, including two soils in which no homology to the *nahAc* probe was found. Despite the apparent abundance of strains harbouring *phn*-type genes they are obviously difficult to cultivate, which may explain the reason there has been no previous report of this PAH catabolic genotype. The original culturing of *Burkholderia* sp. RP007 was therefore fortuitous, and may reflect the acquisition of the *phn* genes by a readily culturable strain (RP007) from other non-culturable bacteria in which this genotype normally resides.

Nucleotide sequencing of *phnAc* PCR products amplified from soil DNA revealed highly homologous (98.5–100%) sequence to the RP007 *phnAc* gene. The possibility that these PCR products were derived from a RP007 *phnAc* contaminant was considered, but discounted since negative PCR controls did not suggest contamination, and most nucleotide substitutions relative to RP007 *phn* occurred at the third position of the codon. However, a number of further experiments to investigate the fidelity of *Taq* polymerase and assess the true diversity of the total *phn* PCR product would be necessary to validate these results. If these results are accepted as a true representation of the diversity of the *phnAc* gene present in soil communities, they suggest the presence of a highly conserved genotype amongst different microbial populations. Such high homology could be a result of the primers which were designed based on the RP007 *phnAc* sequence only, and therefore may be very specific for sequences almost identical to *phnAc*.

7.6 Future research

This study of *Burkholderia* sp. RP007 has comprehensively described a novel operon which encodes a common pathway for PAH degradation, has partially characterised two divergent lower pathways for aromatic degradation, and investigated the molecular ecology and diversity of these catabolic genes. Several avenues of research which would expand the results of this study seem apparent.

The first of these pertains to the taxonomy of *Burkholderia* sp. RP007 as indicated by the 16S rRNA nucleotide sequence. The cluster of aromatic degrading *Burkholderia* strains based on the 16S phylogeny was distinct from other *Burkholderia* species, and Lessie *et al.* (1996) suggested these strains probably constituted a separate species. It would be interesting to obtain many of the described *Burkholderia* strains which degrade aromatic compounds and undertake a polyphasic taxonomic analysis, with a view to classifying these strains into a new species (eg. *Burkholderia aromatovorans* or *B. degradens*). The 16S sequence data could also be used to design a probe specific for aromatic degrading *Burkholderia* strains if it could be demonstrated there was significant difference between these strains and other *Burkholderia* spp.

The second area for further investigation would be the regulatory system of the *phn* upper pathway. A *phnR* insertional mutant of RP007 would show whether *phnR* was essential for expression of the *phn* operon. If this were the case, gel mobility shift experiments using purified PhnR and the amplified intergenic region between *phnS* and *phnR* might resolve whether salicylate is the required coinducer for expression of *phn*, as is the case for *nah*, or some other metabolite. The hypothesised role of *phnS* could be tested by studying growth of a *phnS* insertional mutant; if PhnS induced the lower pathway for PAH degradation, a *phnS* mutant would be expected to show poorer growth on naphthalene relative to the wild-type since only one molecule of pyruvate would be available for growth per naphthalene molecule.

Since the PhnC extradiol dioxygenase belongs to a new class of extradiol dioxygenases, and is very different to the NahC- and BphC-type dioxygenases, it is worthy of further study. PhnC is novel in the sense that no other members of Class III extradiol dioxygenases preferentially cleave substrates with more than one ring, and therefore warrants purification, biochemical analysis, and structure determination by X-ray crystallography.

Finally, it would be interesting to accurately determine the relative populations of bacteria with *phn* and *nah* genotypes in the environment. The intensities of *phn* and *nah* PCR products in Figures 6.2 and 6.3 were similar, but no inferences can be made about the relative populations of bacteria with these genotypes due to the high number of cycles used in the amplification. Competitive PCR experiments would allow the quantitation of a given DNA template relative to a standard template of a known concentration which is amplified by the same primer set. Estimation of *phn/nah* template proportions in different contaminated soils would indicate the relative ecological significance of these genotypes.

APPENDIX A Details of bacterial strains and genes referred to in the text

Strain	Gene[a]	Substrate[b]	GB Acc. no.[c]	Reference
P. putida ML2	*bed*	benzene	L04642	Tan *et al.*, 1993
Acinetobacter calcoaceticus ADP1	*ben*	benzoate	M76990	Neidle *et al.*, 1991
P. putida	*bnz*	benzene	M17904	Irie *et al.*, 1987
Rhodococcus sp. M5	*bpd*	biphenyl/PCB	U27591	Wang *et al.*, 1995
Rhodococcus sp. RHA1	*bph*	PCB	D31242	Masai *et al.*, 1995
R. globerulus P6	*bph*	biphenyl	X80041 (A1-A4), X75633 (C1), X75634 (C2), X75635 (C3)	Asturias *et al.*, 1994 Asturias *et al.*, 1995
P. alcaligenes KF707	*bph*	biphenyl	M83673	Taira *et al.*, 1992
Pseudomonas sp. IC	*bph*	biphenyl	U01825	Carrington *et al.*, 1994
Pseudomonas sp. LB400	*bph*	biphenyl/PCB	M86348	Erickson & Mondello, 1992
Pseudomonas sp. KKS102	*bph*	PCB	D17319, D38633(R)	Fukada *et al.*, 1994
Sphingomonas yanoikuyae Q1	*bph*	PCB	M20640	Taira *et al.*, 1988
S. yanoikuyae B1	*bph*	biphenyl/PAHs	U23374	Kim & Zylstra, 1995
Comamonas testosteroni B-356	*bph*	biphenyl/PCB	U47637	Sylvestre *et al.*, 1996
Sphingomonas sp. CB3	*car*	carbazole	AF060489	Shepherd & Lloyd-Jones, 1998
Pseudomonas sp. CA10	*car*	carbazole	D89064, D89065	Sato *et al.*, 1997a; b
A. calcoaceticus	*cat*	catechol	M76991	Romero-Arroyo *et al.*, 1995
P. putida	*cat*	catechol	M33817	Rothmel *et al.*, 1990
P. putida PRS1	*cat*	catechol	U12557	Houghton *et al.*, 1995
B cepacia AA1	–	–	U47111	unpublished
P. putida mt-15	–	toluene	U01826	Keil *et al.*, 1985
P. putida	*clc*	chlorocatechol	L06464	Coco *et al.*, 1993
Sphingomonas sp. HV3	*cmp*	chloroaromatics	L10655	Yrjälä *et al.*, 1994
P. putida F1	*cmt*	*p*-cumate	U24215	Eaton, 1996
P. fluorescens IP01	*cum*	cumene	D37828	Aoki *et al.*, 1996
Pseudomonas CF600	*dmp*	(methyl)phenol	X68033 P17262(B)	Shingler *et al.*, 1993 Bartilson & Shingler, 1989
Burkholderia sp. DNT	*dnt*	2,4-dinitrotoluene	U62430	Suen *et al.*, 1996
Pseudomonas sp. C18	*dox*	dibenzothiophene	M60405	Denome *et al.*, 1993
R. erythropolis BD2	*ipb*	isopropylbenzene	U24277	Kesselar *et al.*, 1996
Pseudomonas sp. JR1	*ipb*	isopropylbenzene	U53507	Pflugmacher, *et al.*, 1996
S. paucimobilis SYK6	*lig*	protocatechuate	M34835	Noda *et al.*, 1990
Alcaligenes eutrophus JMP222	*mpc*	catechol	X52414	Kabisch & Fortnagel, 1990
E. coli K12	*mhp*	3-phenyl-proprionic acid	D86239	Spence *et al.*, 1996
Pseudomonas sp. U2	*nag*	naphthalene	AF036940	Fuenmayor *et al.*, 1998
P. putida G7	*nah*	naphthalene	M83949 (Aa-Ad) J04994 (C) U09057 (QED) J04233 (R) P08127(H)	Simon *et al.*, 1993 Harayama & Rekik, 1989 Eaton, 1994 Schell & Sukordhaman, 1989 Ghosal *et al.*, 1987

Strain	Gene[a]	Substrate[b]	GB Acc. no.[c]	Reference
P. putida NCIB 9816-4	*nah*	naphthalene	M83950, U49496	Simon *et al.*, 1993
P. putida NCIB 9816	*ndo*	naphthalene	M23914	Kurkela *et al.*, 1988
Bradyrhizobium sp.	*nod*	na	P04682	Scott, 1986
Rhizobium meliloti	*nod*	na	X53820	Rushing *et al.*, 1991
R. meliloti	*nod*	na	S63823	unpublished
Rhizobium tropici	*nod*	na	L04660	Sousa *et al.*, 1993
Pseudomonas sp. JS42	*ntd*	2-nitrotoluene	U49504	Parales *et al.*, 1996
P. putida OUS82	*pah*	phenanthrene	D16629	Takizawa *et al.*, 1994
P. aeruginosa PaK1	*pah*	phenanthrene	D84146	unpublished
Flavobacterium sp. ATCC 39723	*pcp*	pentchlorophenol	U12290	unpublished
Bacillus thermoleovorans	*phe*	phenol	AF031325	Duffner & Mueller, 1998
P. putida BH	*phe*	phenol	D63814	unpublished
P. putida P35X	*phh*	(methyl)phenol	X79599(R) X77856(B)	Ng *et al.*, 1994
P. putida	*phl*	phenol	X91145	Muller *et al.*, 1996
Ralstonia picketti PK01	*tbu*	benzene, toluene	U72645	Byrne & Olsen, 1996
			U20258(E)	Kukor & Olsen, 1996
Pseudomonas sp. P51	*tcb*	chlorobenzene	U15298	Werlen *et al.*, 1996
				van der Meer *et al.*, 1991
Burkholderia sp.	*tec*	1,2,4,5-tetrachlorobenzene	U78099	Beil *et al.*, 1997
P. putida UCC22 (UCC2)	*tdn*	aniline	D85415 X59790(C)	Fukumori & Saint, 1997
P. putida	*tmb*	1,2,4-trimethylbenzene	U41301	unpublished
P. putida F1	*tod*	toluene	J04996	Zylstra & Gibson, 1989
C. testosteroni T-2	*tsa*	*p*-toluenesulfonate	U32622	Junker *et al.*, 1997
Cycloclasticus oligotrophus RB1	*xyl*	toluene, xylene	U51165	Wang *et al.*, 1996
P. aeruginosa JI104	*xyl*	toluene, xylene	D83057	Kitayama *et al.*, 1996
P. putida mt-2	*xyl*	toluene, xylene	M64747	Harayama *et al.*, 1991
			M10143(R)	Inouye *et al.*, 1988
				Nakai *et al.*, 1983
P. putida PDK1	*xyl*	toluene	Q04285	Benjamin *et al.*, 1991
S. yanoikuyae B1	*xyl*	biphenyl	U23375	Kim & Zylstra, 1995

[a] The designation of genes characterised from this strain, (–) indicates no gene designation was given.

[b] The substrate on which catabolic strains were isolated, na – not applicable (for rhizospere-dwelling strains), (–) indicates the catabolised substrate was not given.

[c] GenBank accession number pertaining to these genes, letters in parenthesis indicate these genes have different accession numbers.

Table A.1 Details of bacterial strains and characterised genes referred to in the text.

APPENDIX B Details of subclones derived from pB1

Clone	Size[a]	Description	Orientn[b]
pB1	11.5	*Hin*dIII fragment derived from a total genomic DNA preparation of *Burkholderia* sp. RP007, cloned into *Hin*dIII site of pUC18; indole-positive	→
pB2	9.5	*Bam*HI deletion of pB1. Contains 9.5 kb *Hin*dIII-*Bam*HI fragment of pB1, indole negative	→
pB3	8.5	8.5 kb *Pst*I fragment of pB1 cloned into *Pst*I site of pUC18, indole negative	→
pB5	5.2	5.2 kb *Sph*I fragment of pB1 cloned into *Sph*I site of pUC18, indole negative	←
pB6	3.5	3.5 kb *Sph*I fragment of pB1 cloned into *Sph*I site of pUC18, indole negative	→
pB6R	3.5	As for pB6, insert in reverse orientation, strongly indole-positive	←
pB9	4.5	4.5 kb *Kpn*I fragment at 3' end of pB1 cloned into *Kpn*I site of pUC18, weakly indole positive	→
pB9R	4.5	As for pB9, insert in reverse direction, indole positive	←

[a] Size of insert in kb

[b] Orientation of insert in pUC18 MCS with respect to the orientation of the pB1 insert; → indicates orientation is 5'→3' as for pB1, ← indicates orientation is opposite to pB1.

Table B.1 Details of the subclones initially derived from pB1 to establish the location of the dioxygenases genes responsible for the indole-positive phenotype.

Clone	Size[a]	Description	Orientn[b]	Site A[c]	Site B[d]
pBE1	1.5	*Eco*RI deletion of pB1, contains 1.5 kb fragment from MCS-*Hin*dIII end of pB1	→	NA[e]	NA
pBE2	2.4	2.4 kb *Eco*RI fragment of pB1 cloned into *Eco*RI site of pUC18	→	*Xba*I	*Pst*I
pBE2R	2.4	As for pBE2, insert in reverse orientation	←	*Xba*I	*Pst*I
pBE3	1.7	1.7 kb *Eco*RI fragment of pB1 cloned into *Eco*RI site of pUC18	→	*Xba*I	*Pst*I
pBE3R	1.7	As for pBE3, insert in reverse orientation	←	*Xba*I	*Pst*I
pBE4	0.3	0.3 kb *Eco*RI fragment of pB1 cloned into *Eco*RI site of pUC18	→	NA	NA
pBE5	1.7	1.7 kb *Eco*RI fragment of pB1 cloned into *Eco*RI site of pUC18	→	*Xba*I	*Pst*I
pBE5R	1.7	As for pBE5, insert in reverse orientation	←	*Xba*I	*Pst*I
pBE6	1.6	1.6 kb *Eco*RI fragment of pB1 cloned into *Eco*RI site of pUC18	→	*Xba*I	*Sph*I
pBE6R	1.6	As for pBE6, insert in reverse orientation	←	*Xba*I	*Sph*I
pB4	2.7	2.7 kb *Pst*I fragment from MCS-*Eco*RI end of pB1 (includes MCS *Pst*I site) cloned into *Pst1* site of pUC18	→	*Xba*I	*Kpn*I
pB4R	2.7	As for pB4, insert in reverse orientation	←	*Xba*I	*Kpn*I

[a] Size of insert in kb

[b] Orientation of insert in pUC18 MCS with respect to the orientation of the pB1 insert; → indicates orientation is is 5'→3' as for pB1, ← indicates orientation is opposite to pB1.

[c] Restriction enzyme used for nested deletion experiment to create Site A. Site A requires a 5'-overhang which is susceptible to exonuclease III digestion

[d] Restriction enzyme used for nested deletion experiment to create Site B. Site B requires a 3'-overhang of 3 or more bases and is resistant to exonuclease III digestion

[e] NA – not applicable, this construct was not used for nested deletion experiment

Table B.2 Subcloned derivatives of pB1 cloned using internal *Eco*RI sites and used to obtain nested deletion derivatives for nucleotide sequencing. Constructs in this table are shown in Figure 4.4.

Clone	Vector	Size[a]	Description	ORFs[b]	Orientn[c]	Confirmed by[d]
pBX1B	pUC18	3.0	*Sty*I deletion of pB1. Contains ORF1 and part of ORF9 present on 1.5 kb of MCS end of pB1	1	←	*Eco*R1-*Hin*dIII *Hin*dIII-*Bam*HI
pBK1	pKK223-3	2.0	*Sty*I (blunt)-*Hin*dIII fragment of pBX1B cloned into *Hin*dIII-*Sma*I sites of pKK223-3	1	→	*Eco*R1-*Hin*dIII *Pst*I
pBE2H	pUC18	2.4	Blunt-ended 2.4 kb *Eco*RI insert of pBE2 cloned into *Hin*cII site of pUC18	2	←	*Eco*R1-*Hin*dIII *Sma*I
pBK2	pKK223-3	2.4	*Eco*RI-*Hin*dIII insert of pBE2H cloned into *Eco*RI-*Hin*dIII sites of pKK223-3	2	→	*Eco*RI-*Hin*dIII *Kpn*I, *Sal*I, *Sma*I
pBX4	pUC18	1.5	1.5 kb *Sph*I-*Sal*I fragment of pB5 cloned into *Sph*I-*Sal*I sites of pUC18	4	←	*Sph*I-*Sal*I *Eco*RI, *Kpn*I
pBK4	pKK223-3	1.5	*Hin*dIII-*Sma*I insert of pBX4 cloned into *Hin*dIII-*Sma*I sites of pKK223-3	4	→	*Hin*dIII-*Sma*I *Hin*dIII-*Kpn*I *Eco*RI
pBE5H	pUC18	1.7	Blunt-ended 1.7 kb *Eco*RI insert of pBE5R cloned into *Hin*cII site of pUC18	5	←	*Eco*RI-*Hin*dIII *Sph*I
pBK5	pKK223-3	1.7	*Eco*RI-*Hin*dIII insert of pBE5H cloned into *Eco*RI-*Hin*dIII sites of pKK223-3	5	→	*Eco*RI-*Hin*dIII *Sph*I-*Hin*dIII
pBK678	pKK223-3	3.5	3.5 kb insert of pB6R excised with *Hin*dIII-*Sma*I and cloned into *Hin*dIII-*Sma*I sites of pKK223-3	6,7,8	→	*Eco*RI-*Hin*dIII *Pst*I
pBX8	pUC18	1.2	1.2 kb *Sph*I-*Pst*I fragment of pB6 cloned into *Sph*I-*Pst*I sites of pUC18	8	←	*Sph*I-*Pst*I *Eco*RI, *Bam*HI
pBK8	pKK223-3	0.75	0.75 kb *Eco*RI-*Hin*dIII fragment of pBX8 cloned into *Eco*RI-*Hin*dIII sites of pKK223-3	8	→	*Eco*RI-*Hin*dIII *Bam*HI
pBX9R	pUC18	1.5	1.5 kb *Bam*HI fragment of pB1 cloned into *Bam*HI site of pUC18	9	←	*Bam*HI, *Sph*I
pBK9	pKK223-3	1.5	1.5 kb insert of pBX9R excised with *Eco*RI-*Hin*dIII and cloned into *Eco*RI-*Hin*dIII sites of pKK223-3	9	→	*Eco*RI-*Hin*dIII *Sph*I

[a] Size of insert in kb.

[b] Complete open reading frames (ORFs) of pB1 which were present on this subcloned fragment.

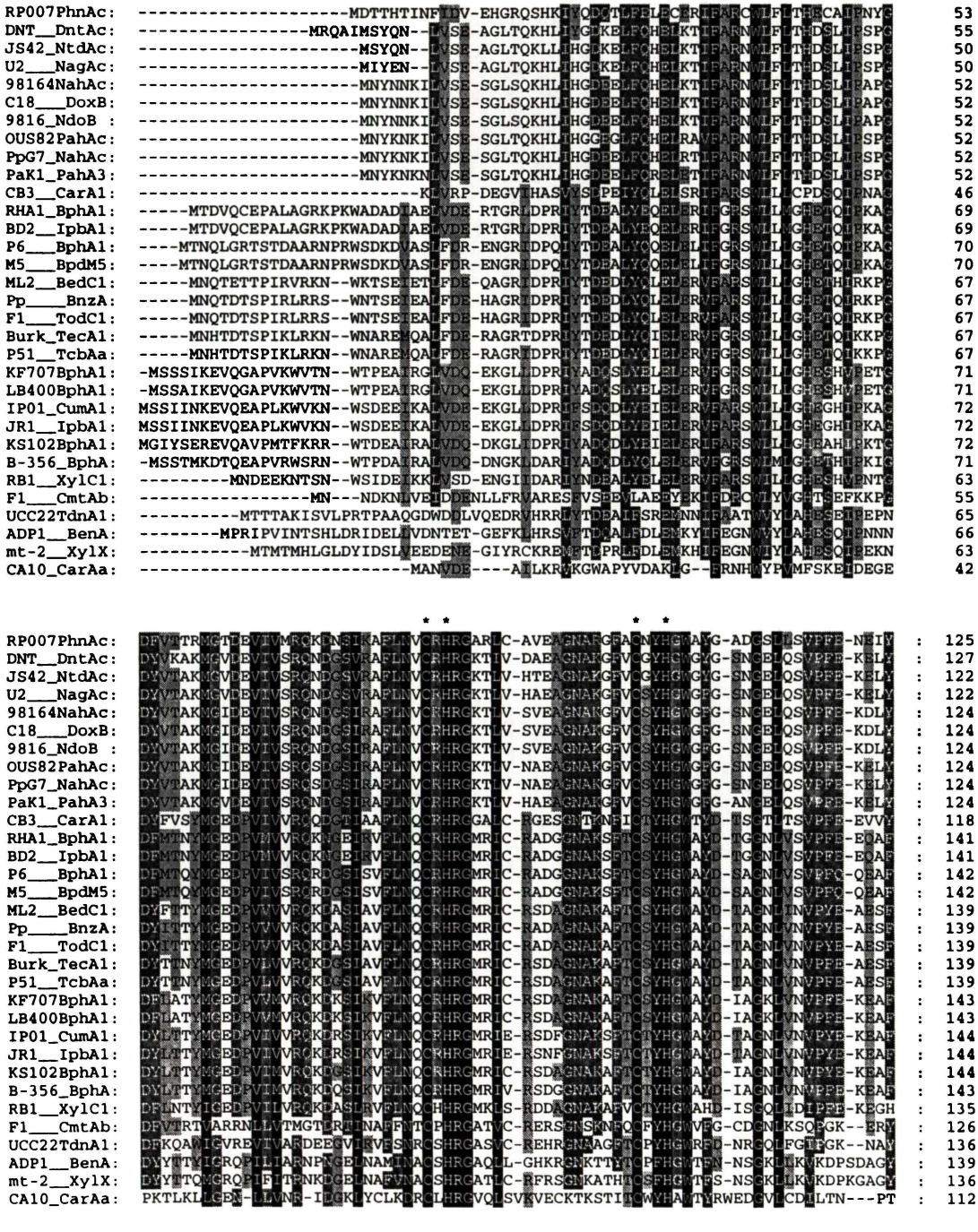
[c] For pUC18 constructs, → indicates the direction of transcription of the ORF(s) of that insert is *Hin*dIII→*Eco*RI relative to the pUC18 MCS, ← indicates transcription in the opposite direction. For pKK223-3 constructs, → indicates the direction of transcription is *Eco*RI→*Hin*dIII relative to the MCS of this vector, and is the required direction for expression from the *tac* promoter.

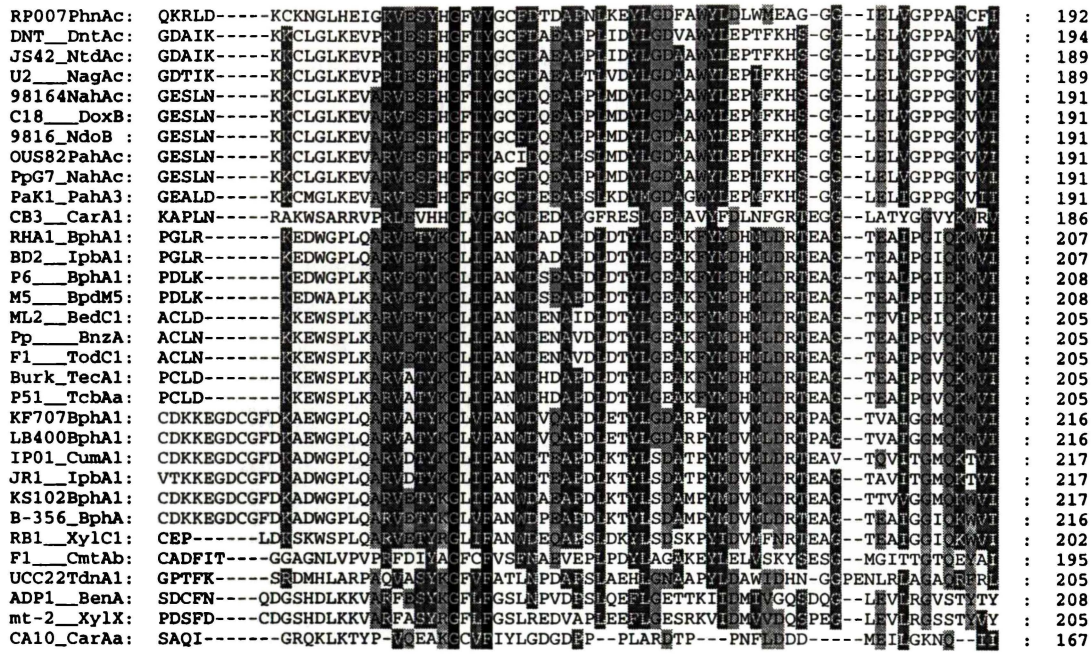
[d] Digestion of the construct by these restriction enzymes gave fragments of the expected sizes and confirmed the identity of the clone.

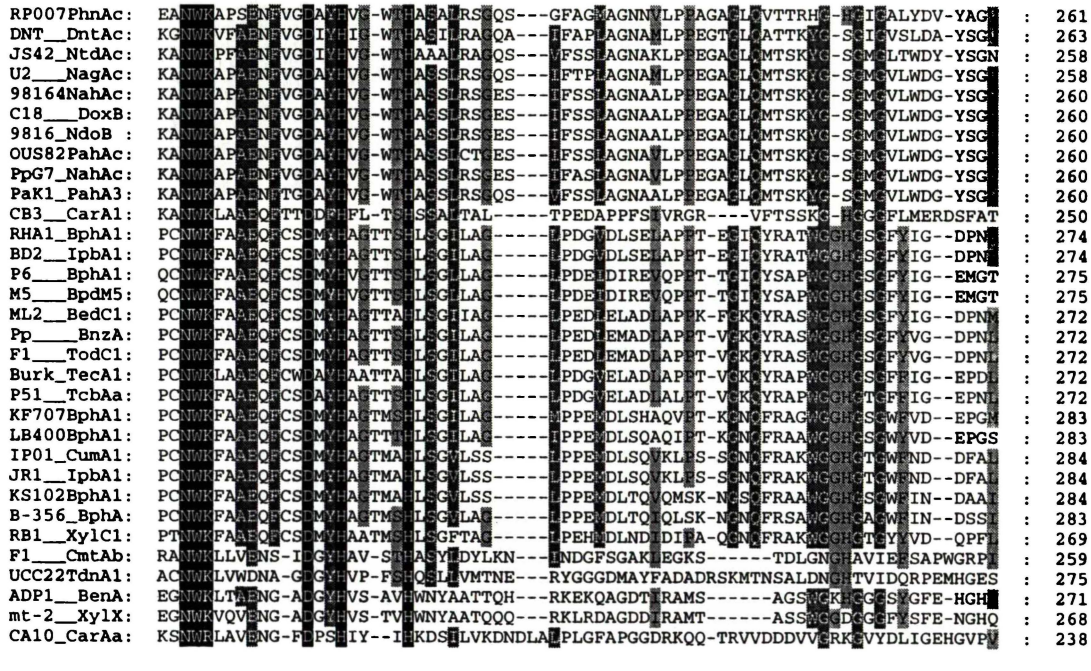
Figure B.3 Details of the subclones derived from pB1 constructed for the expression of the predicted ORFs of pB1. Fragments were first cloned into pUC18 to enable subsequent excision of the insert with the desired restriction enzymes for ligation into the pKK223-3 MCS, which has a limited number of retriction sites available for cloning. The pKK223-3 constructs described here are shown in Figure 4.6

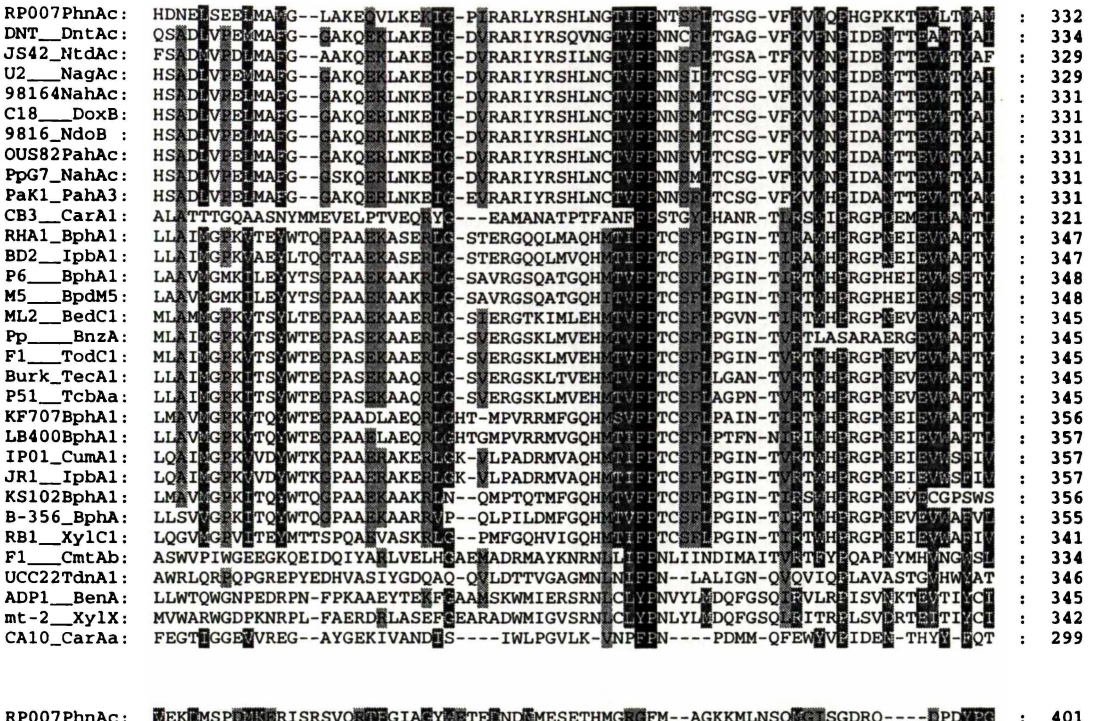
APPENDIX C Amino acid sequence alignments

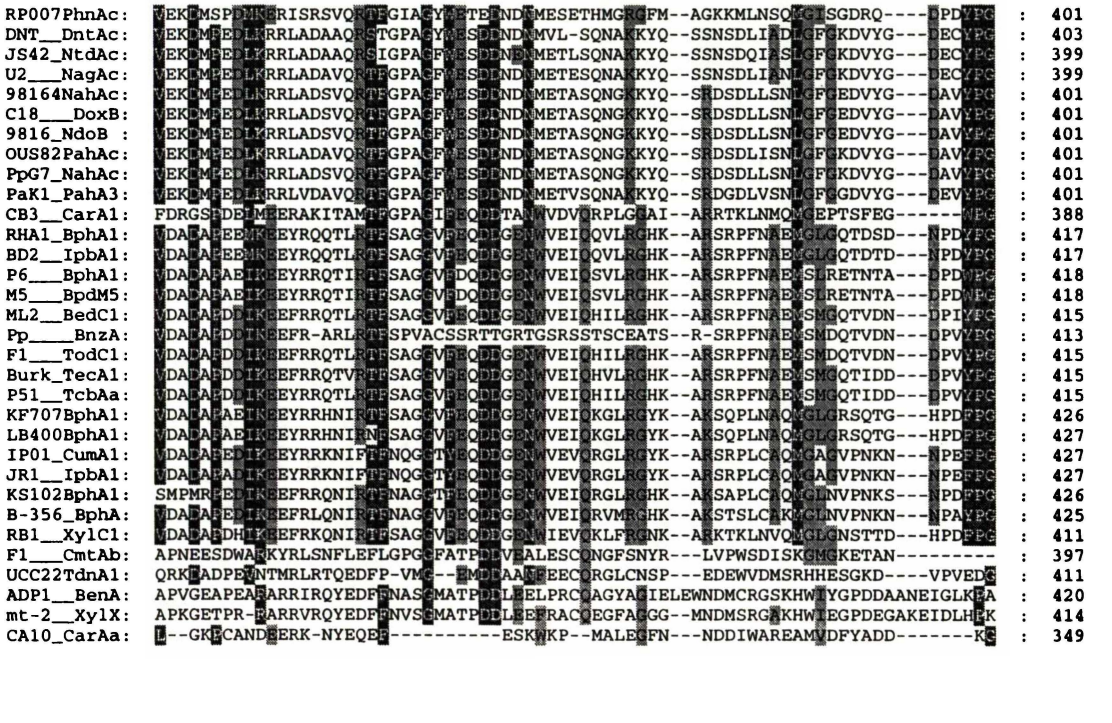
Shading for amino acid alignments represents the degree of conservation amongst aligned residues; black – 100%, grey – 80%, light grey – 60%.




RP007PhnAc:
DNT__DntAc:
JS42_NtdAc:
U2___NagAc:
98164NahAc:
C18___DoxB:
9816_NdoB :
OUS82PahAc:
PpG7_NahAc:
PaK1_PahA3:
CB3__CarA1:
RHA1_BphA1:
BD2__IpbA1:
P6___BphA1:
M5___BpdM5:
ML2__BedC1:
Pp____BnzA:
F1___TodC1:
Burk_TecA1:
P51__TcbAa:
KF707BphA1:
LB400BphA1:
IP01_CumA1:
JR1__IpbA1:
KS102BphA1:
B-356_BphA:
RB1__XylC1:
F1___CmtAb:
UCC22TdnA1:
ADP1__BenA:
mt-2__XylX:
CA10_CarAa:


RP007PhnAc:
DNT__DntAc:
JS42_NtdAc:
U2___NagAc:
98164NahAc:
C18___DoxB:
9816_NdoB :
OUS82PahAc:
PpG7_NahAc:
PaK1_PahA3:
CB3__CarA1:
RHA1_BphA1:
BD2__IpbA1:
P6___BphA1:
M5___BpdM5:
ML2__BedC1:
Pp____BnzA:
F1___TodC1:
Burk_TecA1:
P51__TcbAa:
KF707BphA1:
LB400BphA1:
IP01_CumA1:
JR1__IpbA1:
KS102BphA1:
B-356_BphA:
RB1__XylC1:
F1___CmtAb:
UCC22TdnA1:
ADP1__BenA:
mt-2__XylX:
CA10_CarAa:


RP007PhnAc:
DNT__DntAc:
JS42_NtdAc:
U2___NagAc:
98164NahAc:
C18___DoxB:
9816_NdoB :
OUS82PahAc:
PpG7_NahAc:
PaK1_PahA3:
CB3__CarA1:
RHA1_BphA1:
BD2__IpbA1:
P6___BphA1:
M5___BpdM5:
ML2__BedC1:
Pp____BnzA:
F1___TodC1:
Burk_TecA1:
P51__TcbAa:
KF707BphA1:
LB400BphA1:
IP01_CumA1:
JR1__IpbA1:
KS102BphA1:
B-356_BphA:
RB1__XylC1:
F1___CmtAb:
UCC22TdnA1:
ADP1__BenA:
mt-2__XylX:
CA10_CarAa:


RP007PhnAc:
DNT__DntAc:
JS42_NtdAc:
U2___NagAc:
98164NahAc:
C18___DoxB:
9816_NdoB :
OUS82PahAc:
PpG7_NahAc:
PaK1_PahA3:
CB3__CarA1:
RHA1_BphA1:
BD2__IpbA1:
P6___BphA1:
M5___BpdM5:
ML2__BedC1:
Pp____BnzA:
F1___TodC1:
Burk_TecA1:
P51__TcbAa:
KF707BphA1:
LB400BphA1:
IP01_CumA1:
JR1__IpbA1:
KS102BphA1:
B-356_BphA:
RB1__XylC1:
F1___CmtAb:
UCC22TdnA1:
ADP1__BenA:
mt-2__XylX:
CA10_CarAa:

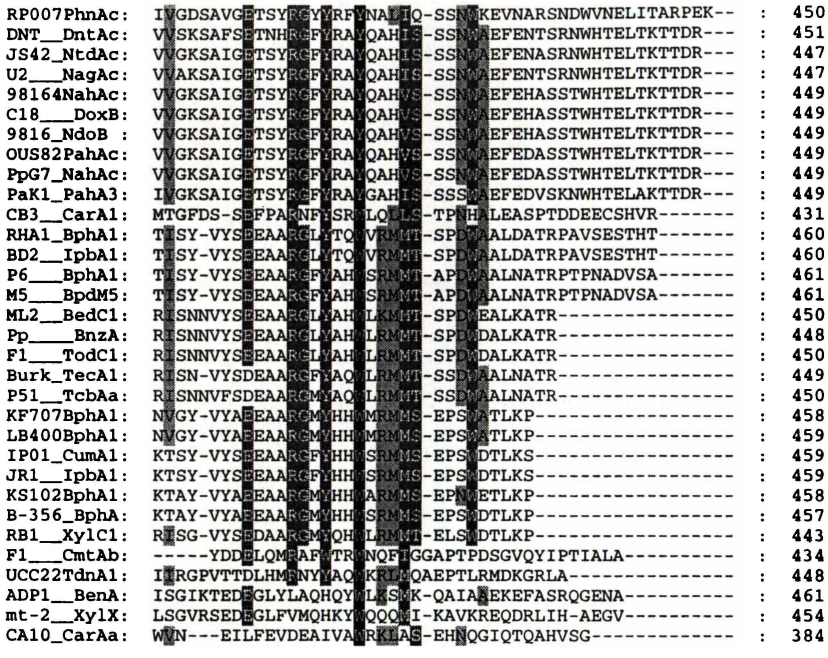
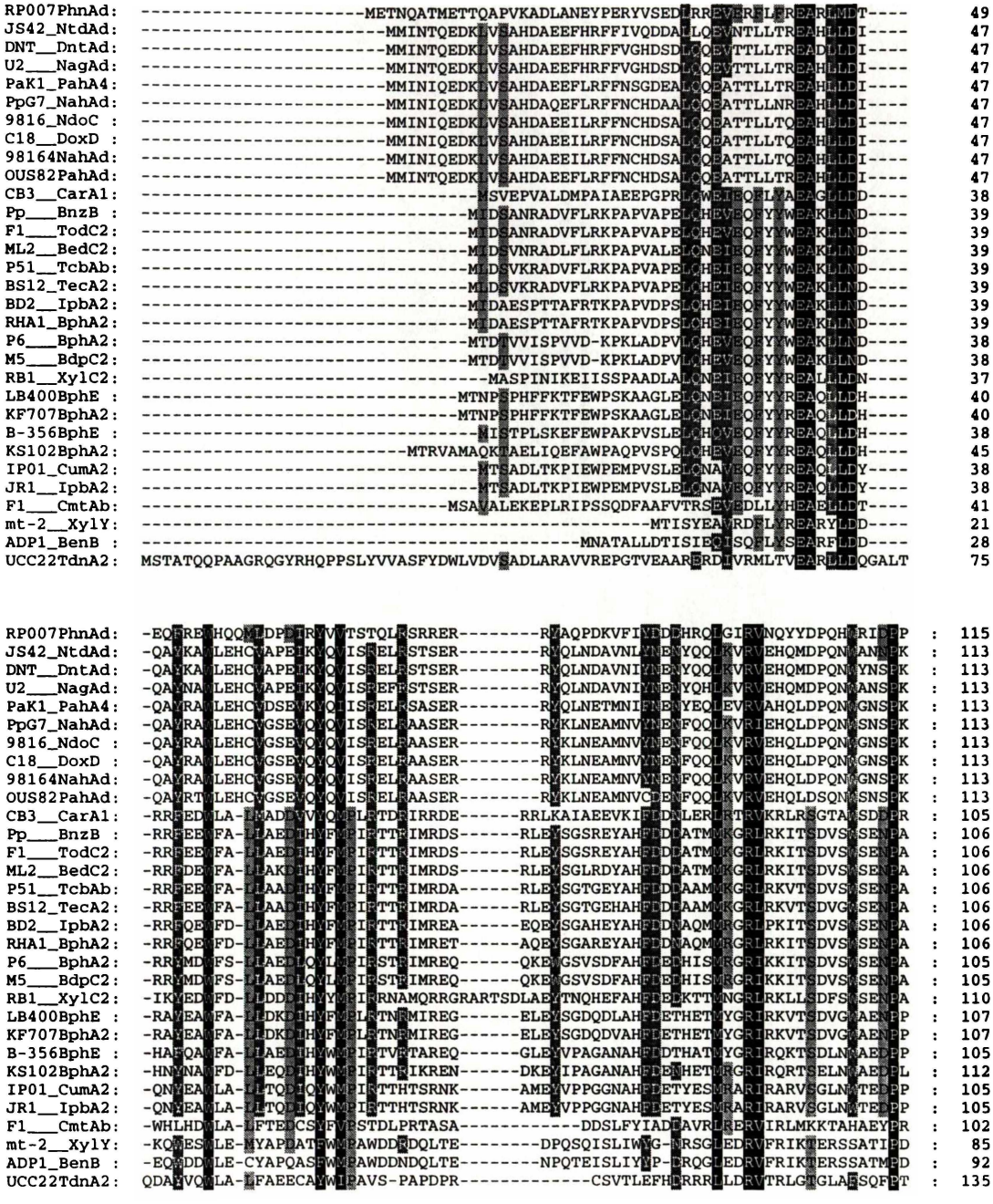


Figure C.1 Alignment of predicted amino acid sequences of different iron-sulfur protein α subunits (ISPα) of aromatic initial dioxygenases which show homology to PhnAc of *Burkholderia* sp. RP007. The conserved cysteine and histidine residues which coordinate the Rieske-type [2Fe-2S] binding sites are marked with asterixes. Details of the strains from which these genes were characterised can be found in Table A.1.



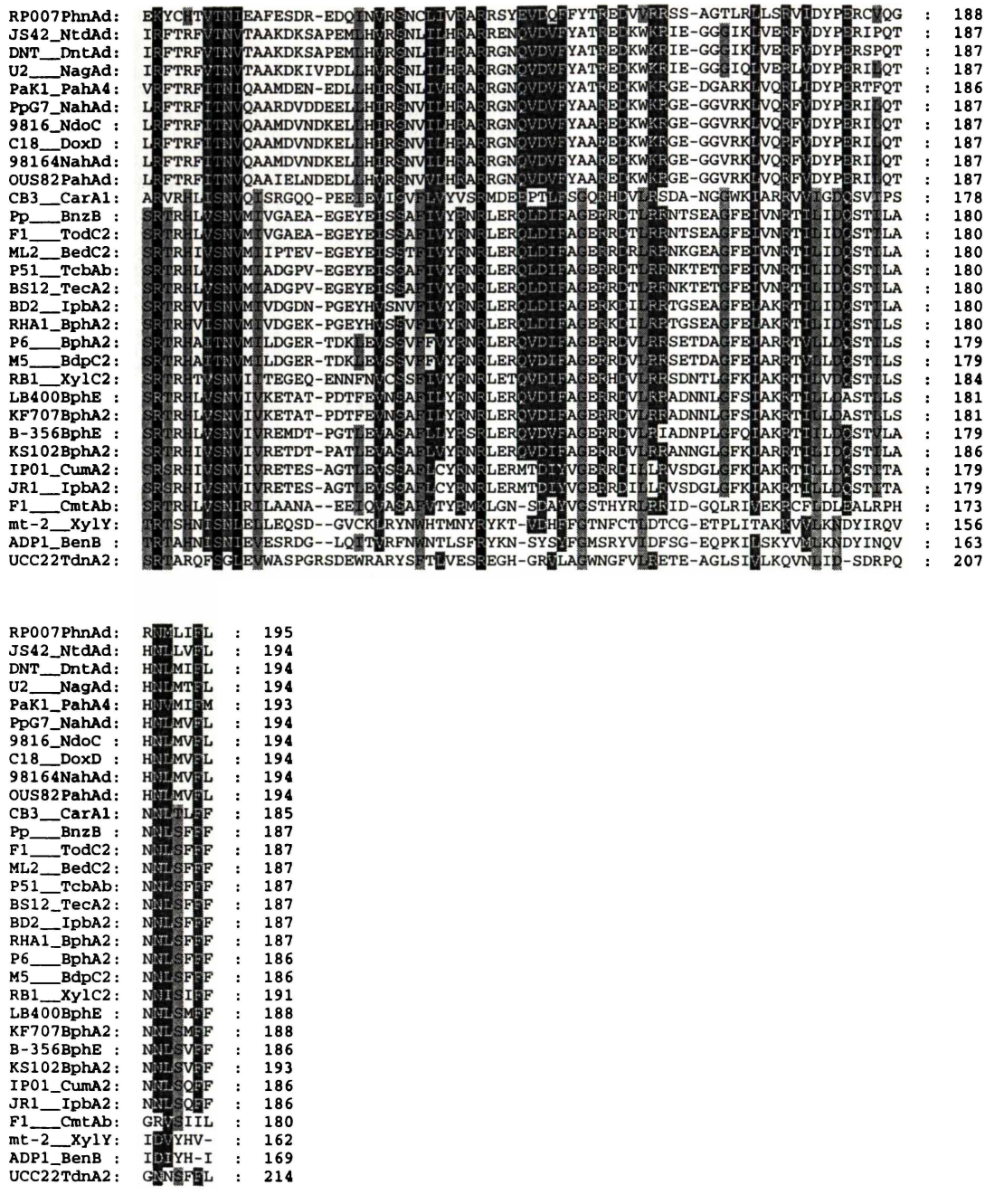
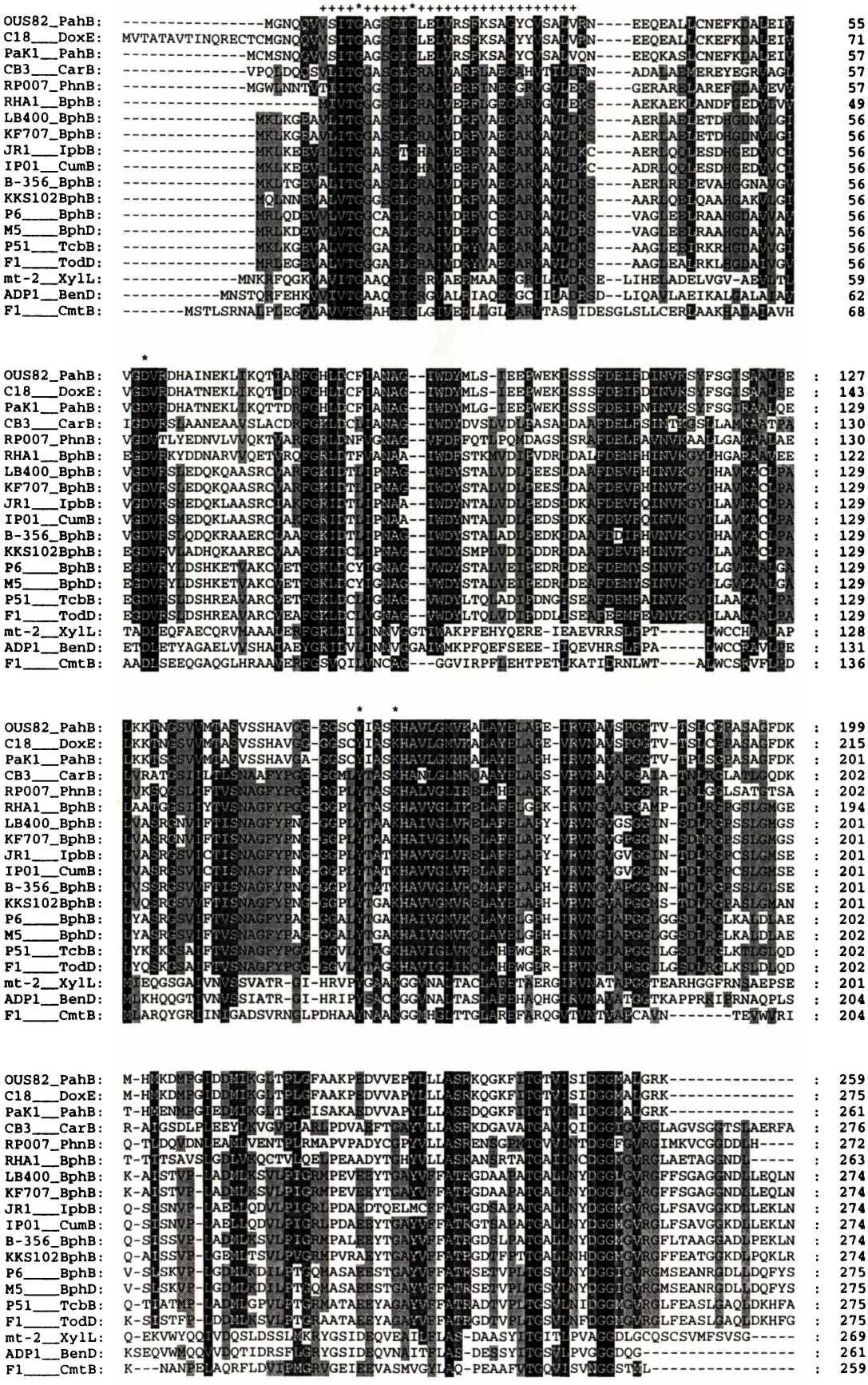
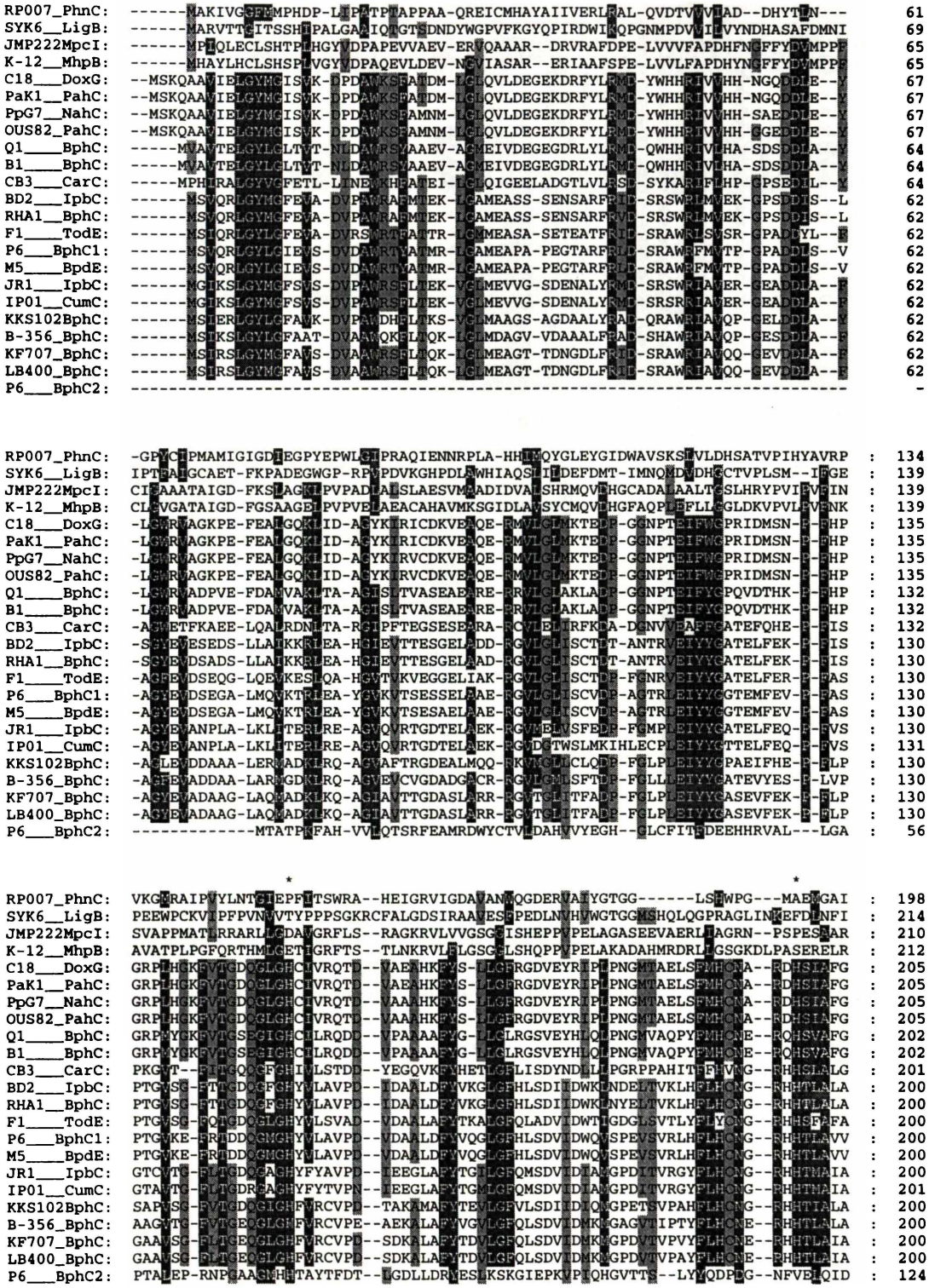


Figure C.2 Alignment of predicted amino acid sequences of different iron-sulfur protein β subunits (ISPβ) of aromatic initial dioxygenases which show homology to PhnAd of *Burkholderia* sp. RP007. Details of the strains from which these genes were characterised can be found in Table A.1.



```
OUS82_PahB:  -------  :    -
C18___DoxE:  -------  :    -
PaK1__PahB:  -------  :    -
CB3___CarB:  E------  :  277
RP007_PhnB:  -------  :    -
RHA1__BphB:  -------  :    -
LB400_BphB:  IHP----  :  277
KF707_BphB:  IHP----  :  277
JR1___IpbB:  ID-----  :  276
IP01__CumB:  ID-----  :  276
B-356_BphB:  INREGQE  :  281
KKS102BphB:  LS-----  :  276
P6____BphB:  KGALV--  :  280
M5____BphD:  KGALV--  :  280
P51___TcbB:  -------  :    -
F1____TodD:  -------  :    -
mt-2__XylL:  -------  :    -
ADP1__BenD:  -------  :    -
F1____CmtB:  -------  :    -
```

Figure C.3 Alignment of predicted amino acid sequences of different aromatic dihydrodiol dehydrogenases which show homology to PhnB of *Burkholderia* sp. RP007. The five residues strictly conserved amongst short chain alcohol dehydrogenases (Gly^{12}, Gly^{18}, Asp^{59}, Tyr^{154}, Lys^{159} numbered according to BphB of *R. globerulus* P6) (Asturias *et al.*, 1994) are marked with asterixes, and the region involved in binding $NAD(P)^+$ cofactors is marked with (+). Details of the strains from which these genes were characterised can be found in Table A.1.



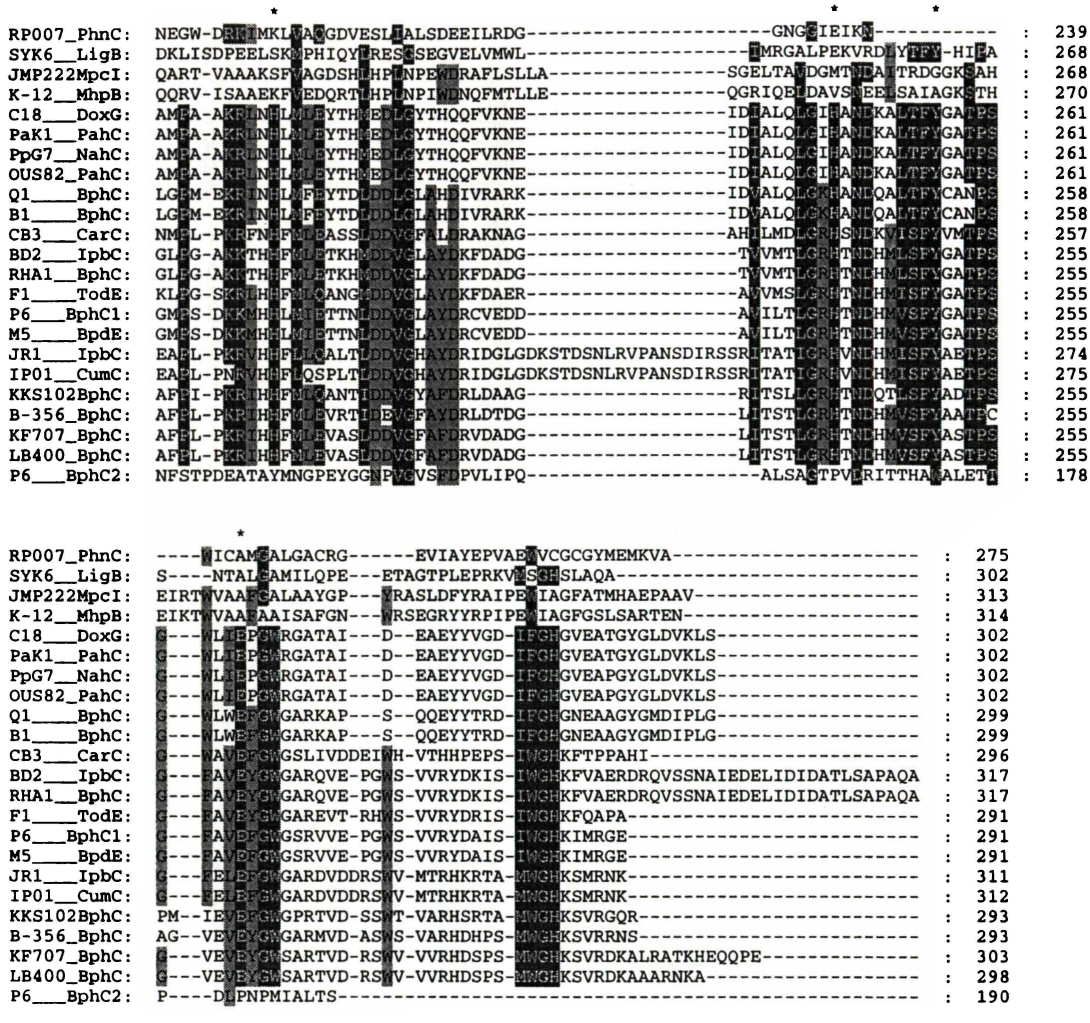


Figure C.4 Alignment of predicted amino acid sequences of different aromatic extradiol dioxygenases with PhnC of *Burkholderia* sp. RP007. Six residues which are strictly conserved amongst both 1,2-dihydroxybiphenyl dioxygenases and 1,2-dihydroxynaphthalene dioxygenases (His^{146}, His^{195}, His^{210}, His^{241}, Tyr^{250}, Glu^{260} numbered according to BphCI of *R. globerulus* P6), and are proposed to be ligands for the ferrous iron (Asturias *et al.*, 1994), are marked with asterixes; these are not conserved in PhnC. Details of the strains from which these genes were characterised can be found in Table A.1.

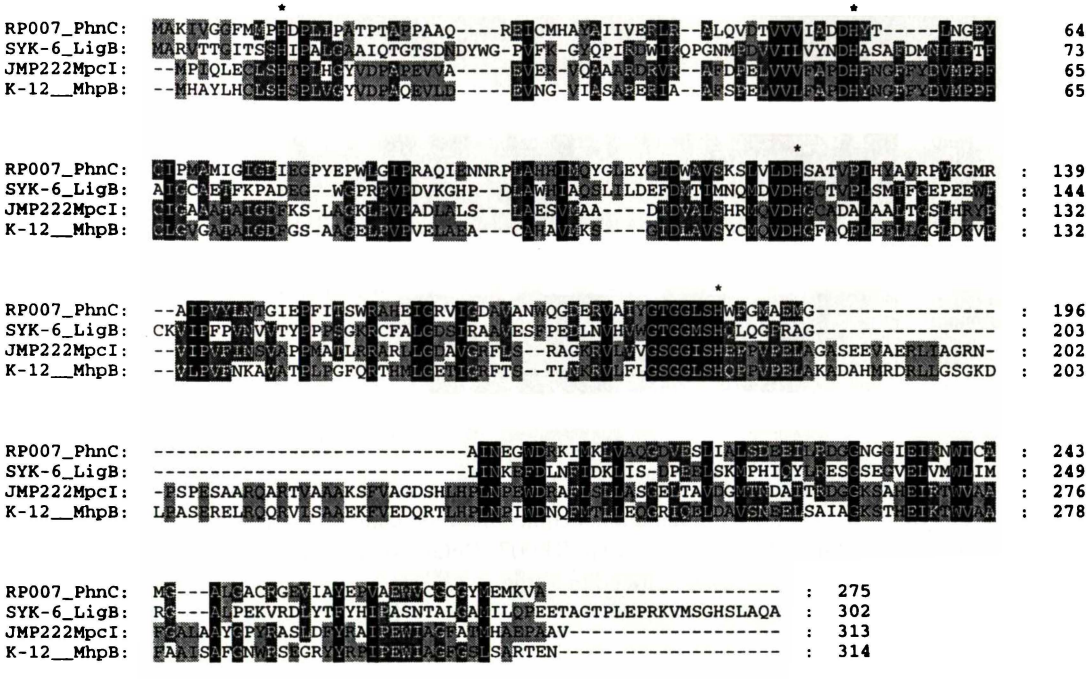


Figure C.5 Alignment of predicted amino acid sequences of Class III aromatic extradiol dioxygenases with PhnC of *Burkholderia* sp. RP007, showing the locations of the four histidine residues (His^{10}, His^{53}, His^{115}, His^{179} numbered according to MhpB of *E. coli*), which may be involved in coordination of the non-heme cofactor (Spence *et al.*, 1996), marked with asterixes.

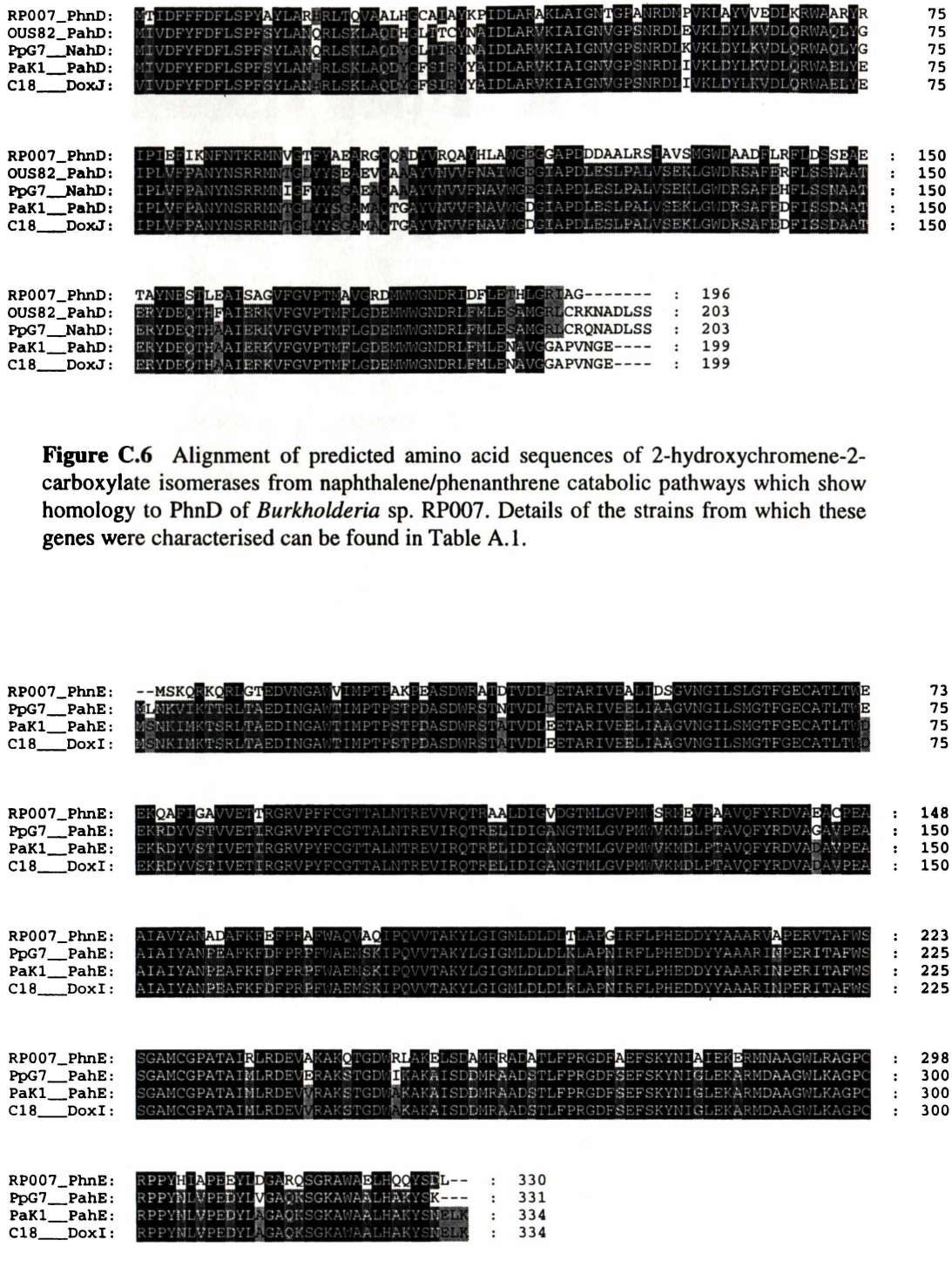


Figure C.6 Alignment of predicted amino acid sequences of 2-hydroxychromene-2-carboxylate isomerases from naphthalene/phenanthrene catabolic pathways which show homology to PhnD of *Burkholderia* sp. RP007. Details of the strains from which these genes were characterised can be found in Table A.1.

Figure C.7 Alignment of predicted amino acid sequences of the hydratase-aldolase from naphthalene/phenanthrene catabolic pathways which show homology to PhnE of *Burkholderia* sp. RP007. Details of the strains from which these genes were characterised can be found in Table A.1.

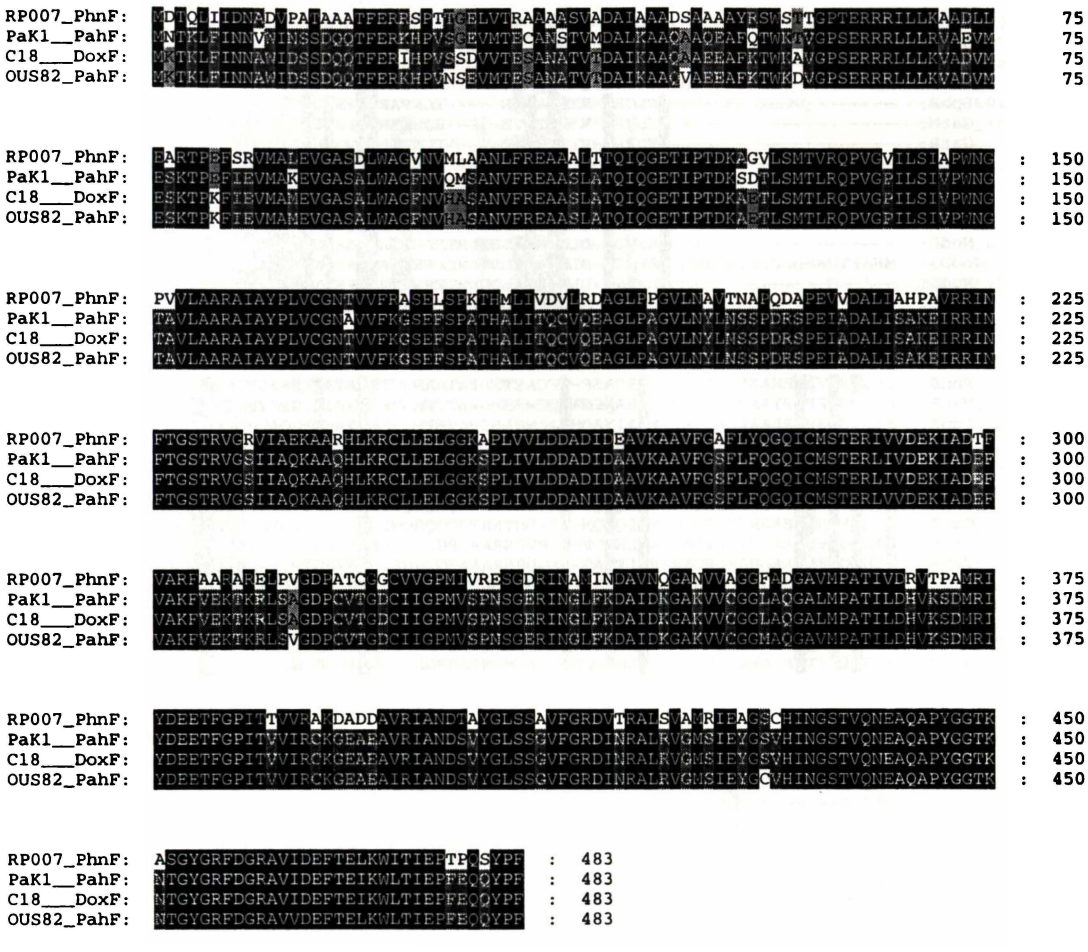
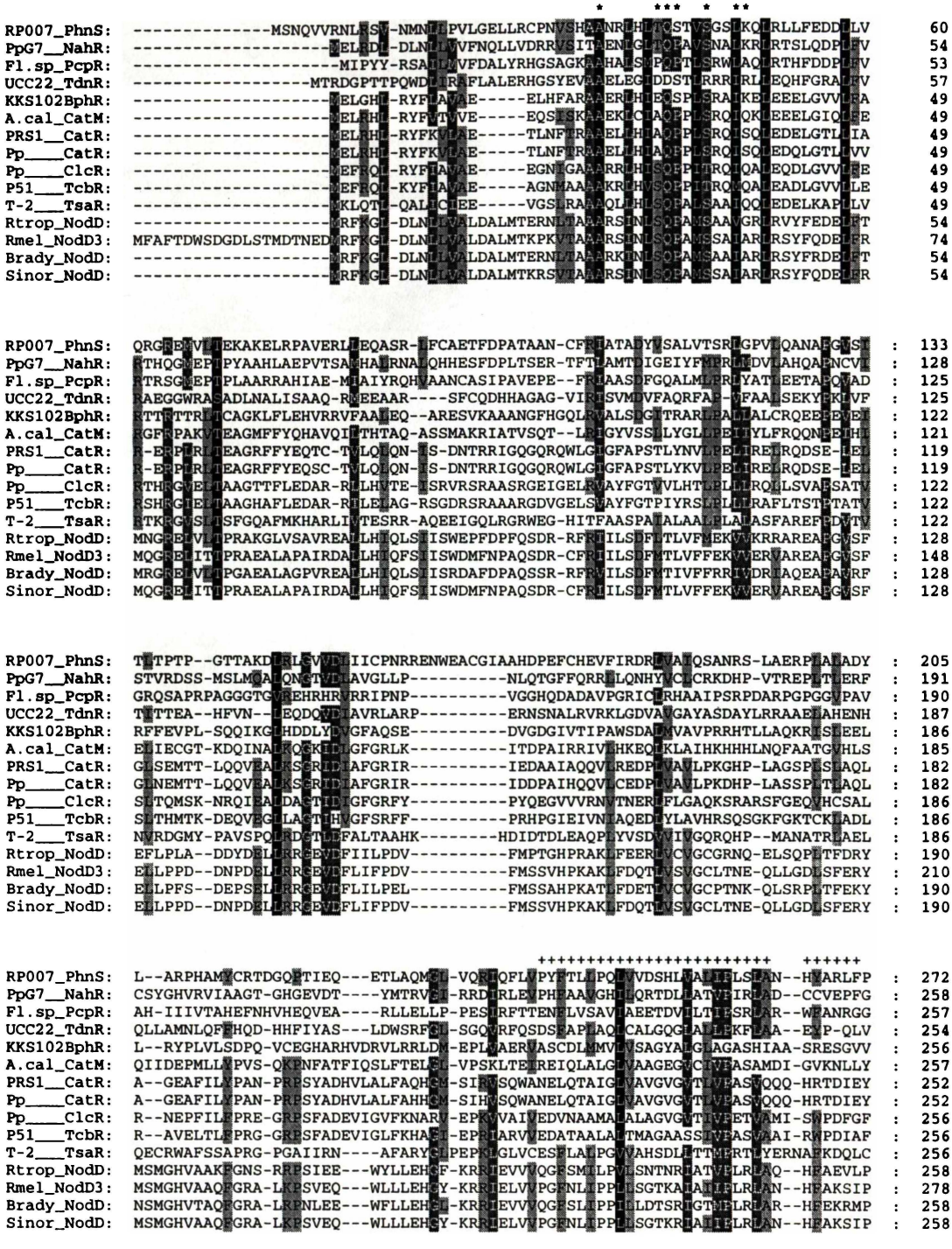


Figure C.8 Alignment of predicted amino acid sequences of salicylaldehyde dehydrogenases from naphthalene/phenanthrene catabolic pathways which show homology to PhnF of *Burkholderia* sp. RP007. Details of the strains from which these genes were characterised can be found in Table A.1.



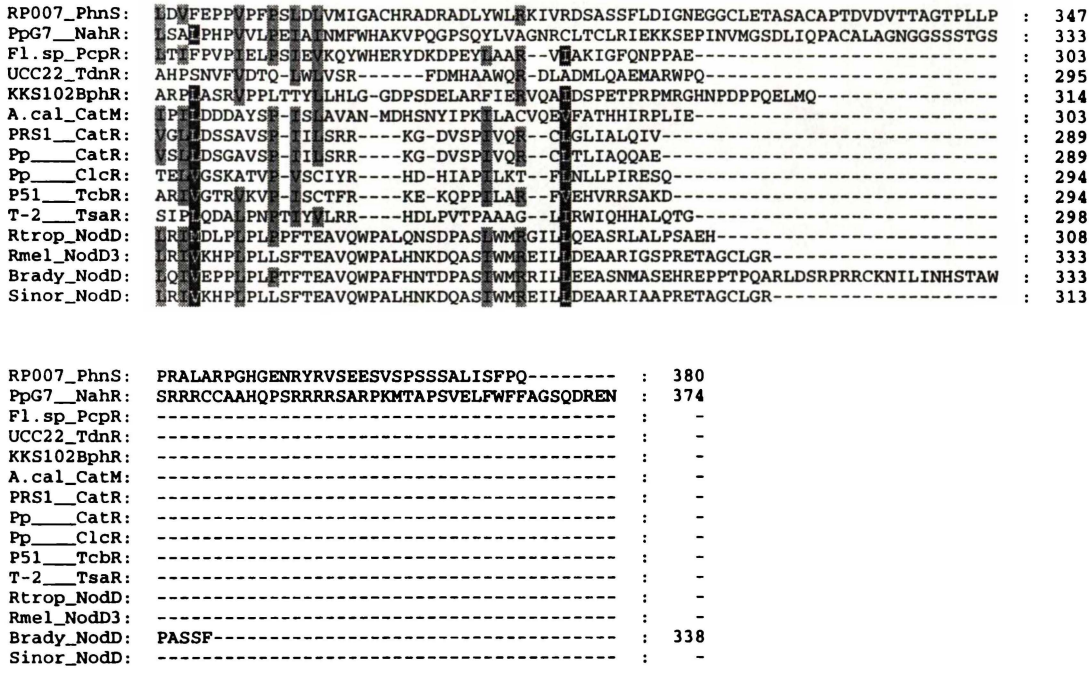


Figure C.9 Alignment of predicted amino acid sequences of LysR-type transcriptional regulators from aromatic catabolic pathways and from rhizosphere-dwelling bacteria which show homology to PhnS of *Burkholderia* sp. RP007. The helix-turn-helix motif is located in the region corresponding to residues 23–42 of NahR of *P. putida* G7, and the seven highly conserved residues in this region (Ala[27], Thr(Ser)[33], Gln[34], Pro[35], Ser(Thr)[38], Leu[44], Glu[45] numbered according to NahR) (Schell, 1993) are marked with asterixes. A conserved C-terminal domain (NahB residues 227–253) important for DNA interactions and transcriptional activation is also shown (+). Details of the strains from which these genes were characterised can be found in Table A.1.

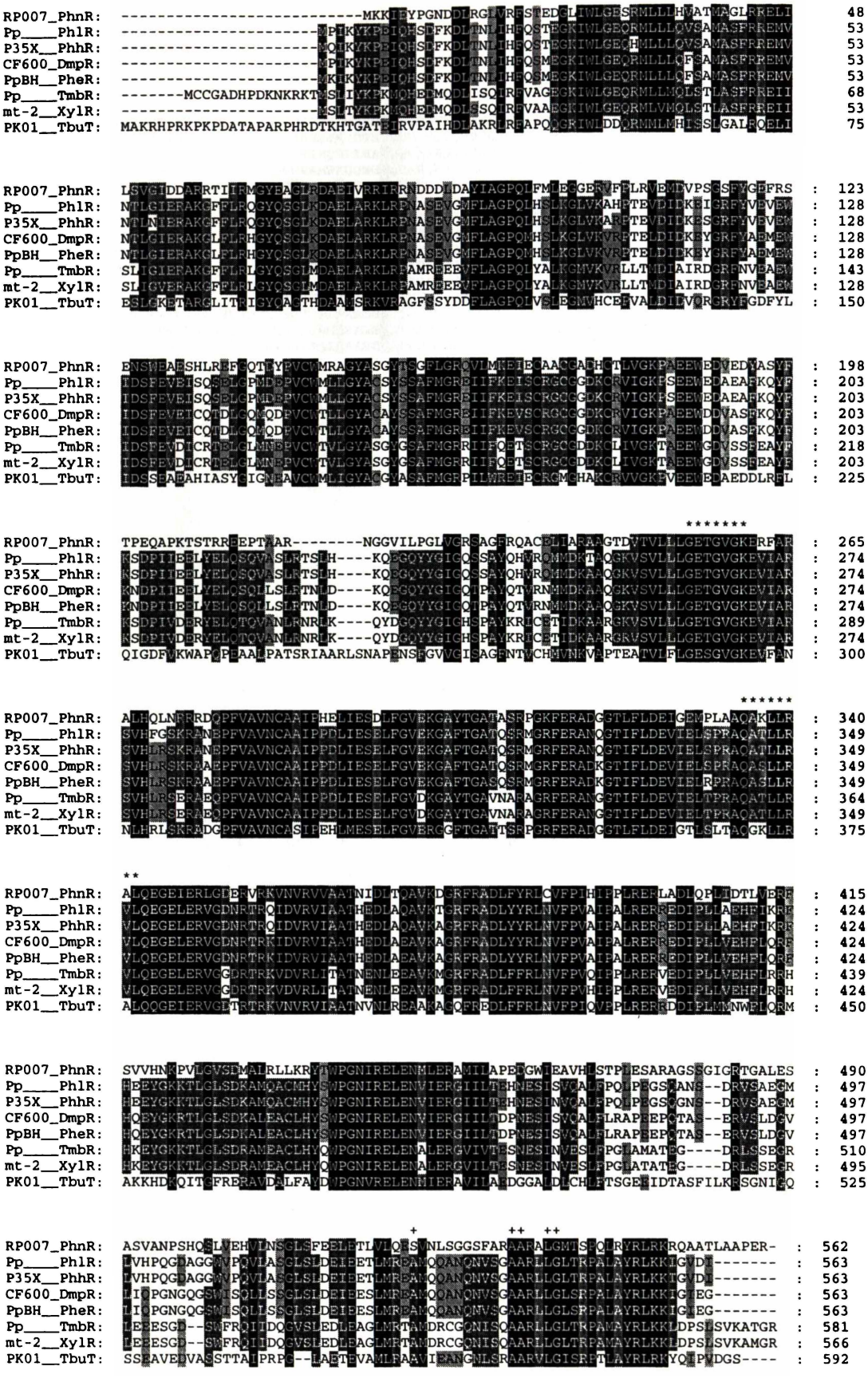
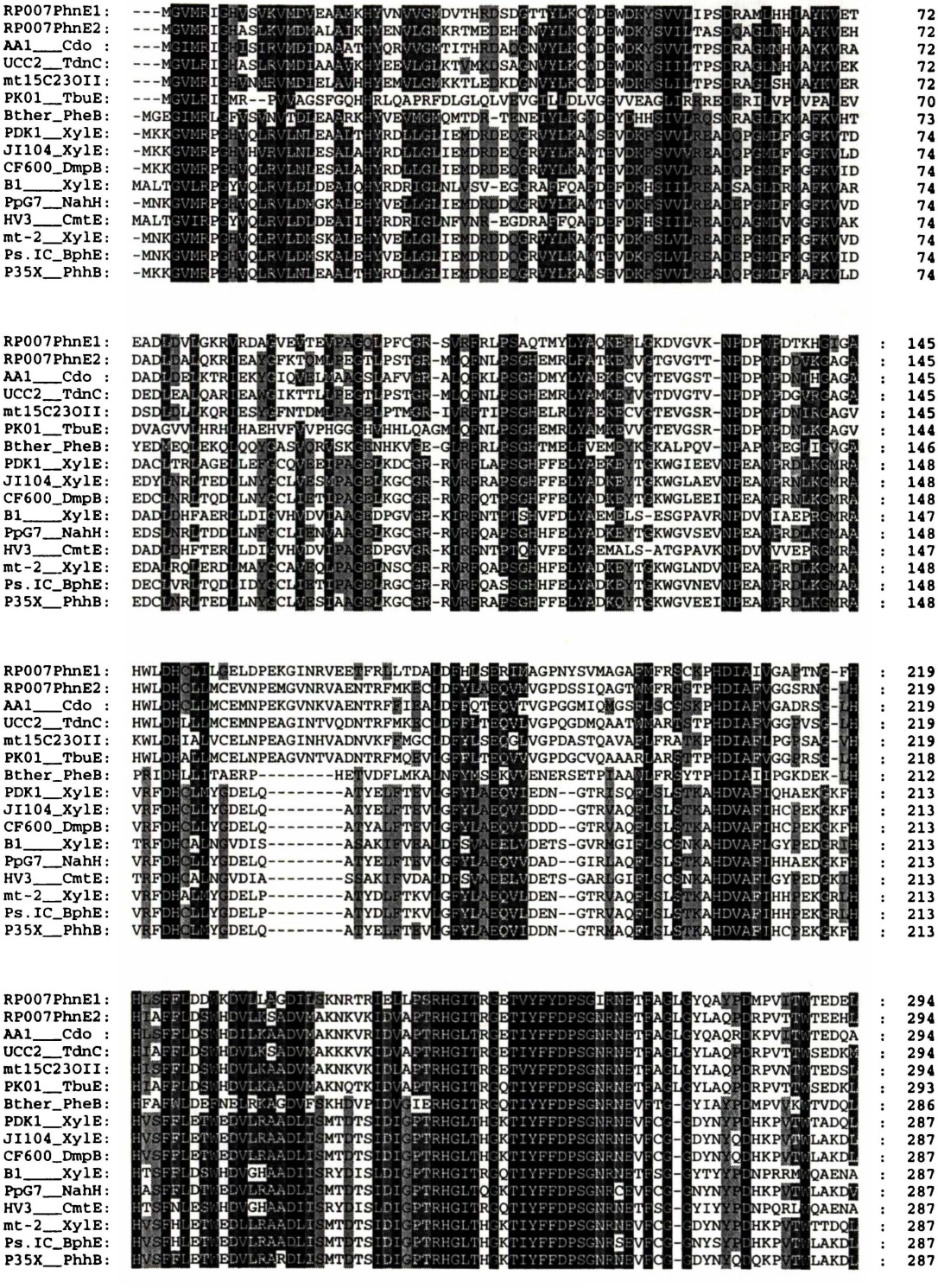


Figure C.10 Alignment of predicted amino acid sequences of σ^{54}-dependent (NtrC-like) transcriptional regulators from aromatic catabolic pathways which show homology to PhnR of *Burkholderia* sp. RP007, showing conserved C-domain motifs (*) and residues involved in the D-domain helix-turn-helix (+). Details of the strains from which these genes were characterised can be found in Table A.1.



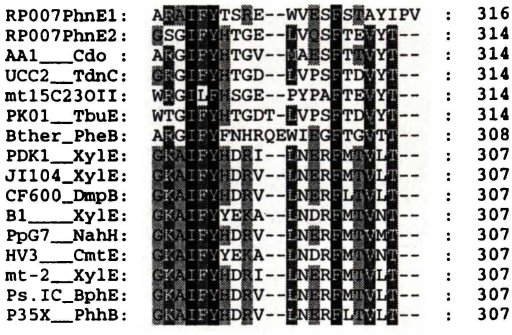


Figure C.11 Alignment of predicted amino acid sequences of different catechol 2,3-dioxygenases from aromatic catabolic pathways which show homology to PhnE1 and PhnE2 of *Burkholderia* sp. RP007. Details of the strains from which these genes were characterised can be found in Table A.1.

References

Abril, M. A., Michan, C., Timmis, K. N., Ramos, J. L. (1989). Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J Bacteriol* **171**, 6782–6790.

Aihara, J. (1992). Why aromatic compounds are stable. *Sci Am* (March), 44–50.

Aislabie, J., Riddell, P., Huser, B. (1993). Bioremediation – theory and application: The use of microorganisms in the cleanup of a former carbonisation plant. *Water and Wastes in NZ* **77**, 54–59.

Aoki, H., Kimura, T., Habe, H., Yamane, H., Kodama, T., Omori, T. (1996). Cloning, nucleotide sequence, and characterisation of the genes encoding enzymes involved in the degradation of cumene to 2-hydroxy-6-oxo-7-methylocta-2,4-dienoic acid in *Pseudomonas fluorescens* IP01. *J Ferment Bioeng* **81**, 187–196.

Assinder, S. J. & Williams, P. A. (1990). The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Micro Physiol* **31**, 1–69.

Asturias, J. A., Diaz, E., Timmis, K. N. (1995). The evolutionary relationship of biphenyl dioxygenase from Gram-positive *Rhodococcus globerulus* P6 to multicomponent dioxygenases from Gram-negative bacteria. *Gene* **156**, 11–18.

Asturias, J. A., Eltis, L. D., Prucha, M., Timmis, K. N. (1994). Analysis of three 2,3-dihydroxybiphenyl 1,2-dioxygenases found in *Rhodococcus globerulus* P6. Identification of a new family of extradiol dioxygenases. *J Biol Chem* **269**, 7807–7815.

Ausubel, F. M., Brent, M., Kingstone, R. E., Moore, D. E., Seidman, J. G., Smith, J. A., Struhl. K. (1989). *Current protocols in molecular biology.* New York: John Wiley & Sons.

Barkay, T., Liebert, C., Gillman, M. (1989). Hybridisation of DNA probes with whole-community genome for detection of genes that encode microbial responses to pollutants: *mer* genes and Hg^{2+} resistance. *Appl Environ Microbiol* **55**, 1574–1577.

Barnsley, E. A. (1975a). Bacterial degradation of fluoranthene and benz[*a*]pyrene. *Can J Microbiol* **146,** 173–183.

Barnsley, E. A. (1975b). The induction of the enzymes of naphthalene catabolism in Pseudomonads by salicylate and 2-aminobenzoate. *J Gen Microbiol* **88**, 193–197.

Barnsley, E. A. (1976a). Naphthalene metabolism by Pseudomonads: the oxidation of 1,2-dihydroxynaphthalene to 2-hydroxychromene-2-carboxylic acid and the formation of 2-hydroxybenzalpyruvate. *Biochem Biophys Res Commun* **72**, 1116–1121.

Barnsley, E. A. (1976b). Role and regulation of the *ortho* and *meta* pathways of catechol metabolism in Pseudomonads metabolising naphthalene and salicylate. *J Bacteriol* **125**, 404–410.

Bartilson, M. & Shingler, V. (1989). Nucleotide sequence and expression of the catechol 2,3-dioxygenase-encoding gene of phenol-catabolising *Pseudomonas* CF600. *Gene* **85**, 233–238.

Beil, S., Happe, B., Timmis, K. N., Pieper, D. H. (1997). Genetic and biochemical characterisation of the broad spectrum chlorobenzene dioxygenase from *Burkholderia* sp. strain PS12 – dechlorination of 1,2,4,5-tetrachlorobenzene. *Eur J Biochem* **247**, 190–199.

Benjamin, R. C., Voss, J. A., Kunz, D. A. (1991). Nucleotide sequence of *xylE* from the TOL pDK1 plasmid and structural comparison with isofunctional catechol 2,3-dioxygenase genes from TOL pWWO and NAH7. *J Bacteriol* **173**, 2724–2728.

Benner, B. A., Bryner, N. P., Wise, S. A., Mulholland, G. W., Lao, R. C., Fingas, M. F. (1990). Polycyclic aromatic hydrocarbon emissions from the combustion of crude oil on water. *Environ Sci Tech* **24**, 1418–1427.

Berthelet, M., Whyte, L. G., Greer, C. W. (1996). Rapid, direct extraction of DNA from soils for PCR analysis using polyvinylpolypyrrolidone spin columns. *FEMS Microbiol Lett* **138**, 17–22.

Blumer, M. (1976). Polycyclic aromatic compounds in nature. *Sci Am* **234**, 34–45.

Bogardt, A. H. & Hemmingsen, B. B. (1992). Enumeration of phenanthrene-degrading bacteria by an overlayer technique and its use in evaluation of petroleum-contaminated sites. *Appl Environ Microbiol* **58**, 2579–2582.

Boldrin, B., Tiehm, A., Fritzsche, C. (1993). Degradation of phenanthrene, fluorene, fluoranthene, and pyrene by a *Mycobacterium* sp. *Appl Environ Microbiol* **59**, 1927–1930.

Bos, R. P., Theuws, J. L. G., Leijdekkers, Ch-M., Henderson, P. Th. (1984). The presence of the mutagenic polycyclic aromatic hydrocarbons benzo[*a*]pyrene and benz[*a*]anthracene in creosote. *Mutation Res* **130**, 153–158.

Bruce, M. (1995). Isolation and characterisation of high molecular weight polycyclic aromatic hydrocarbon degrading bacteria. *Landcare Research NZ Ltd internal report.*

Bruce, K. D., Hiorns, W. D., Hobman, J. L., Osborn, A. M., Strike, P., Ritchie, D. A. (1992). Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. *Appl Environ Microbiol* **58**, 3413–3416.

Bruce, K. D., Osborn, A. M., Pearson, A. J., Strike, P., Ritchie, D. A. (1995). Genetic diversity within *mer* genes directly amplified from communities of noncultivated soil and sediment bacteria. *Mol Ecol* **4**, 605–612.

Buikema, W. J., Szeto, W. W., Lemley, P. V., Orme-Johnson, W. H., Ausubel, F. M. (1985). Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. *Nucleic Acids Res* **13**, 4539–4555.

Burlage, R. S., Hooper, S. W., Saylor, G. S. (1989). The TOL (pWW0) catabolic plasmid. *Appl Environ Microbiol* **55**, 1323–1328.

Byrne, A. M. & Olsen, R. H. (1996). Cascade regulation of the toluene-3-monooxygenase operon (*tbuA1UBVA2C*) of *Burkholderia pickettii* PKO1: role of the *tbuA1* promoter (*PtbuA1*) in the expression of its cognate activator, Tbu1. *J Bacteriol* **178**, 6327–6337.

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.

Carrington, B., Lowe, A., Williams, P. A. (1994). The lower pathway operon for benzoate catabolism in biphenyl-utilising biphenyl *Pseudomonas* sp. strain IC. *Microbiology* **140**, 499–508.

Cerniglia, C. E. (1992). Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* **3**, 351–368.

Cerniglia, C. E. & Heitkamp, M. A. (1989). Microbial degradation of polycyclic aromatic hydrocarbons (PAH) in the aquatic environment. In *Metabolism of polycyclic aromatic hydrocarbons in the aquatic environment,* pp. 41–67. Edited by U. Varanasi. Boca Rato, Florida: CRC Press, Inc.

Chakrabarty, A. M. (1972). Genetic basis of the biodegradation of salicylate in *Pseudomonas*. *J Bacteriol* **112**, 815–821.

Coco, W. M., Rothmel, R. K., Henikoff, S., Chakrabarty, A. M. (1993). Nucleotide sequence and initial functional characterisation of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. *J Bacteriol* **175**, 417–427.

Conners, M. A. & Barnsley, E. A. (1982). Naphthalene plasmids in *Pseudomonads*. *J Bacteriol* **149**, 1096.

Davies, J. I. & Evans, W. C. (1964). Oxidative metabolism of naphthalene by soil pseudomonads. *Biochem J* **91**, 251–261.

Delgado, D., & Ramos, J. L. (1994). Genetic evidence for the inactivation of the positive transcriptional regulator, XylR, a member of the NtrC family of regulators, by effector binding. *J Biol Chem* **269**, 8059–8062.

Denome, S. A., Stanley, D. C., Olson, E. S., Young, K. D. (1993). Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. *J Bacteriol* **176**, 2158–2164.

Dipple, A., Moschel, R. C., Bigger, C. A. H. (1984). Polynuclear aromatic carcinogens. In *Chemical Carcinogens,* Second Ed. Vol. 1, pp. 41–163. Edited by C. E. Searle. Washington DC: American Chemical Society Monograph 182.

Duffner,F. D. & Mueller, R. (1998). Novel phenol hydroxylase and catechol 2,3-dioxygenase from the thermophilic *Bacillus thermoleovorans* strain A2: Nucleotide sequence and analysis of the genes. *FEMS Microbiol Lett* **161**, 37-45.

Dunn, D. W. & Gunsalus, I. C. (1973). Transmissable plasmid encoding early enzymes of naphthalene oxidation in *Pseudomonas putida. J Bacteriol* **114**, 974–979.

Eaton, R. W. (1994). Organisation and evolution of naphthalene catabolic pathways: sequence of the DNA encoding 2-hydroxychromene-2-carboxylate isomerase and *trans-o*-hydroxybenzylidenepyruvate hydratase-aldolase from the NAH7 plasmid. *J Bacteriol* **176**, 7757–7762.

Eaton, R. W. (1996). *p*-Cumate catabolic pathway in *Pseudomonas putida* Fl: cloning and characterisation of DNA carrying the *cmt* operon. *J Bacteriol* **178**, 1351–1362.

Eaton, R. W. & Chapman, P. J. (1992). Bacterial metabolism of naphthalene: construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions. *J Bacteriol* **174**, 7542–7554.

Ensley, B. D., Gibson, D. T., Laborde, A. L. (1982). Oxidation of naphthalene by a multicomponent enzyme system from *Pseudomonas* sp. strain NCIB 9816. *J Bacteriol* **149**, 948–952.

Ensley, B. D., Osslund, T. D., Joyce, M., Simon, M. J. (1987). Expression and complementation of naphthalene dioxygenase activity in *Escherichia coli*, In *Microbial metabolism and the carbon cycle*, pp. 437–455. Edited by S. R. Hagedorn, R. S. Hanson & D. A. Kunz. New York: Harwood Academic Publishers.

Ensley, B. D., Ratzkin, B. J., Osslund, T. D., Simon, M. J., Wackett, L. P., Gibson, D. T. (1983). Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* **222**, 167–169.

Erickson, B. D. & Mondello, F. J. (1992). Nucleotide sequencing and transcriptional mapping of genes encoding biphenyl dioxygenase, a multicomponent PCB-degrading enzyme in *Pseudomonas* strain LB400. *J Bacteriol* **174**, 2903–2912.

Evans, W. C., Fernley, H. N., Griffiths, E. (1965). Oxidative metabolism of phenanthrene and anthracene by soil pseudomonads; the ring fission mechanism. *Biochem J* **95**, 819–825.

Felsenstein, J. (1989). PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166.

Ferguson, R. L., Buckley, E. N., Palumbo, A. V. (1984). Response of marine bacterioplankton to differential filtration and confinement. *Appl Environ Microbiol* **47**, 49–55.

Fleming, S. T., Sanseverino, J., Sayler, G. S. (1993). Quantitative relationship between naphthalene catabolic gene-frequency and expression in predicting PAH degradation in soils at town gas manufacturing sites. *Environ Sci Technol* **27**, 1068-1074.

Fredrickson, J. K., Balkwill, D. L., Drake, G. R., Romine, M. F., Ringleberg, D. B., White, D. C. (1995). Aromatic-degrading *Sphingomonas* isolates from the deep subsurface. *Appl Environ Microbiol* **61**, 1917–1922.

Freeman, D. J. & Cattell, F. C. R. (1990). Woodburning as a source of atmospheric polycyclic aromatic hydrocarbons. *Environ Sci Tech* **24**, 1581–1585.

Fritzsche, C. (1994). Degradation of pyrene at low defined oxygen concentrations by a *Mycobacterium* sp. *Appl Environ Microbiol* **60**, 1687–1689.

Fuenmayor, S. L., Wild, M., Boyes, A. L., Williams, P. A. (1998). A gene cluster encoding steps in conversion of naphthalene to gentisate in *Pseudomonas* sp. strain U2. *J Bacteriol* **180**, 2522–2530.

Fukuda, M., Yasukochi, Y., Kikuchi, Y., Nagata, Y., Kimbara, K., Horiuchi, H., Takagi, M., Yano, K. (1994). Identification of the *bphA* and *bphB* genes of *Pseudomonas* sp. strain KKS102 involved in degradation of biphenyl and polychlorinated biphenyls. *Biochem Biophys Res Commun* **202**, 850–856.

Fukumori, F. & Saint, C. P. (1997). Nucleotide sequences and regulational analysis of genes involved in conversion of aniline to catechol in *Pseudomonas putida* UCC22 (pTDN1). *J Bacteriol* **179**, 399–408.

Furukawa, K., Hayase, N., Taira, K., Tomizuka, N. (1989). Molecular relationship of chromosomal genes encoding biphenyl/polychlorinated biphenyl catabolism: some bacteria possess a highly conserved *bph* operon. *J Bacteriol* **171**, 5467–5472.

Furukawa, K., Hayase, N., Taira, K. (1990). Biphenyl/polychlorinated biphenyl catabolic gene (*bph* operon): organisation, function and molecular relationships in various pseudomonads. In Pseudomonas*: biotransformations, pathogenesis, and evolving biotechnology,* pp. 111–132. Edited by S. Silver, A. M. Chakrabarty, B. Iglewski & S. Kaplan. Washington, DC: American Society for Microbiology.

Furukawa, K., Hayashida, S., Taira, K. (1992). Biochemical and genetic basis for the degradation of polychlorinated biphenyls in soil bacteria. In Pseudomonas*: molecular biology and biotechnology*, pp. 259–267. Edited by E. Galli, S. Silver, & B. Witholt. Washington, DC: American Society for Microbiology.

Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., Ramos, J. L. (1997). AraC/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61**, 393–410.

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

Gibson, D. T., Mahadevan, V., Jerina, D. M., Yagi, H., Yeh, H. J. C. (1975). Oxidation of the carcinogens benzo[*a*]pyrene and benz[*a*]anthracene to dihydrodiols by a bacterium. *Science* **189**, 295–296.

Gibson, D. T. & Subramanian, V. (1984). Microbial degradation of aromatic hydrocarbons. In *Microbial degradation of organic compounds,* pp. 181–252. Edited by D. T. Gibson. New York: Marcel Dekker.

Gillis, M., van Van, T., Bardin, R., Goor, M., Hebbar, P., Willems, A., Segers, P., Kersters, K., Heulin, T., Fernandez, M. P. (1995). Polyphasic taxonomy in the genus *Burkholderia* leading to an emended description of the genus and proposition of *Burkholderia vietnamiensis* sp. nov. for N_2-fixing isolates from rice in Vietnam. *Int J Syst Bacteriol* **45**, 274–289.

Goyal, A. K. & Zylstra, G. J. (1996). Molecular cloning of novel genes for polycyclic aromatic hydrocarbon degradation from *Comamonas testosteroni* GZ39. *Appl Environ Microbiol* **62**, 230–236.

Grifoll, M., Casellas, M., Bayona, J. M., Solanas, A. M. (1992). Isolation and characterisation of a fluorene-degrading bacterium: identification of ring oxidation and ring fission products. *Appl Environ Microbiol* **58**, 2910–2917.

Grifoll, M., Selifonov, S. A., Gatlin, C. V., Chapman, P. J. (1995). Actions of a versatile fluorene-degrading bacterial islolate on polycyclic aromatic compounds. *Appl Environ Microbiol* **61**, 3711–3723.

Grosser, R. J., Warshawsky, D., Vestal, J. R. (1991). Indigenous and enhanced mineralisation of pyrene, benzo[a]pyrene, and carbazole in soils. *Appl Environ Microbiol* **57**, 3462–3469.

Guerin, W. F. & Jones, G. E. (1988). Mineralisation of phenanthrene and anthracene by a *Mycobacterium* sp. *Appl Environ Microbiol* **54**, 937–944.

Guo, C., Sun, W., Harsh, J. B., Ogram, A. (1997). Hybridisation analysis of microbial DNA from fuel oil-contaminated and noncontaminated soil. *Microb Ecol* **34**, 178–187.

Haak, B., Fetzner, S., Lingens, F. (1995). Cloning, nucleotide sequence, and expression of the plasmid-encoded genes for the two-component 2-halobenzoate 1,2-dioxygenase from *Pseudomonas cepacia* 2CBS. *J Bacteriol* **177**, 667-675.

Han, S., Eltis, L. D., Timmis, K. N., Muchmore, S. W., Bolin, J. T. (1995). Crystal structure of the biphenyl-cleaving extradiol dioxygenase from a PCB-degrading Pseudomonad. *Science* **270**, 976–980.

Harayama, S., Kok, M., Neidle, E. L. (1992). Functional and evolutionary relationships among diverse oxygenases. *Ann Rev Micobiol* **46**, 565–601.

Harayama, S., Polissi, A., Rekik, M. (1991a). Divergent evolution of chloroplast-type ferredoxins. *FEBS Lett* **285**, 85–88.

Harayama, S. & Rekik, M. (1989). Bacterial aromatic ring-cleavage enzymes are classified into two different gene families. *J Biol Chem* **264**, 15328–15333.

Harayama, S. & Rekik, M. (1990). The *meta*-cleavage operon of TOL degradative plasmid pWWO 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 the TOL plasmid pWWO 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., Bairoch, A., Neidle, E. L., Ornston, L. N. (1991b). Potential DNA slippage structures acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal *benABC* and *Pseudomonas putida* TOL pWWO plasmid *xylXYZ* genes encoding benzoate dioxygenases. *J Bacteriol* **173**, 7540–7548.

Harayama, S., Rekik, M., Wubbolts, M., Rose, K., Leppik, R. A., Timmis, K. N. (1989). Characterisation of five genes from the upper pathway operon of TOL plasmid pWWO from *Pseudomonas putida* and identification of the gene products. *J Bacteriol* **171**, 5048–5055.

Heitkamp, M. A. & Cerniglia, C. E. (1988). Mineralisation of polycyclic aromatic hydrocarbons by a bacterium isolated from sediment below an oil field. *Appl Environ Microbiol* **54**, 1612–1614.

Heitkamp, M. A., Franklin, W., Cerniglia, C. E. (1988a). Microbial metabolism of polycyclic aromatic hydrocarbons: isolation and characterisation of a pyrene-degrading bacterium. *Appl Environ Microbiol* **54**, 2549–2555.

Heitkamp, M. A., Freeman, J. P., Miller, D. W. (1988b). Pyrene degradation by a *Mycobacterium* sp.: isolation and identification of ring oxidation and ring fission products. *Appl Environ Microbiol* **54**, 2556–2565.

Herbes, S. E. & Schwall, L. R. (1978). Microbial transformation of polyaromatic hydrocarbons in pristine and petroleum contaminated sediments. *Appl Environ Microbiol* **35**, 306–316.

Herrick, J. B., Madsen, E. L., Batt, C. A., Ghiorse, W. C. (1993). Polymerase chain reaction amplification of naphthalene-catabolic and 16S rRNA gene sequences from indigenous sediment bacteria. *Appl Environ Microbiol* **59**, 687–694.

Houghton, J. E., Brown, T. M., Appel, A. J., Hughes, E. J., Ornston, L. N. (1995). Discontinuities in the evolution of *Pseudomonas putida cat* genes. *J Bacteriol* **177**, 401–412.

Houghton, J. E. & Shanley, M. S. (1994). Catabolic potential of pseudomonads: a regulatory perspective. In *Biological degradation and bioremediation of toxic chemicals,* pp. 11–31. Edited by R. G. Chaudhry. Oregon: Dioscorides Press.

Huang, J. & Schell, M. A. (1991). *In vivo* interactions of the NahR transcriptional activator with its target sequences: inducer mediated changes resulting in transcription activation. *J Biol Chem* **266**, 10830–10838.

Hübner, A., Danganan, C. E., Xun, L., Chakrabarty, A. M., Hendrickson, W. (1998). Genes for 2,4,5-trichlorophenoxyacetic acid metabolism in *Burkholderia cepacia* AC1100: characterisation of the *tftC* and *tftD* genes and locations of the *tft* operons on multiple replicons. *Appl Environ Microbiol* **64**, 2086–2093.

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 characterisation of the genes encoding benzene oxidation enzymes of *Pseudomonas putida. J Bacteriol* **169**, 5174–5179.

Iwabuchi, T. & Harayama, S. (1997). Biochemical and genetic characterisation of 2-carboxybenzaldehyde dehydrogenase, an enzyme involved in phenanthrene degradation by *Nocardioides* sp. strain KP7. *J Bacteriol* **179**, 6488–6494.

Jerina, D. M., Selander, H., Yagi, H., Wells, M. C., Davey, J. F., Mahadevan, V., Gibson, D. T. (1976). Dihydrodiols from anthracene and phenanthrene. *J Am Chem Soc* **98**, 5988–5996.

Jerina, D. M., van Bladeren, P. J., Yagi, H., Gibson, D. T., Mahadevan, V., Neese, A. S., Koreeda, M., Sharma, N. D., Boyd, D. (1984). Synthesis and absolute configuration of *cis*-1,2-, 8,9- and 10,11-dihydrodiol metabolites of benz[*a*]-anthracene formed by a strain of *Beijerinckia*. *J Org Chem* **49**, 1075–1082.

Jones, K. C., Stratford, J. A., Tidridge, P., Waterhouse, K. S. (1989a). Polynuclear aromatic hydrocarbons in an agricultural soil: long-term changes in profile distribution. *Environ Pollut* **56**, 337–351.

Jones, K. C., Stratford, J. A., Waterhouse, K. S., Furlong, E. T., Giger, W., Hites, R. A., Schaffner, C., Johnston A. E. (1989b). Increases in the polynuclear aromatic hydrocarbon content of an agricultural soil over the last century. *Environ Sci Technol* **23**, 95–101.

Juhasz, A. L., Britz, M. L., Stanley, G. A. (1996). Degradation of benzo[*a*]pyrene, dibenz[*a*,*h*]anthracene and coronene by *Burkholderia cepacia*. *Wat Sci Tech* **36**, 45–51.

Juhasz, A. L., Britz, M. L., Stanley, G. A. (1996). Degradation of fluoranthene, pyrene, benz[*a*]anthracene and dibenz[*a*,*h*]anthracene by *Burkholderia cepacia*. *J Appl Microbiol* **83**, 189–198.

Junker, F., Kiewitz, R., Cook, A. M. (1997). Characterization of the *p*-toluene-sulfonate operon *tsaMBCD* and *tsaR* in *Comamonas testosteroni* T-2. *J Bacteriol* **179**, 919-927.

Kabisch, M. & Fortnagel, P. (1990). Nucleotide sequence of metapyrocatechase I (catechol 2,3-oxygenase I) gene *mpcI* from *Alcaligenes eutrophus* JMP222. *Nucleic Acids Res* **18**, 3405–3406.

Keil, H., Lebens, M. R. & Williams, P. A. (1985). TOL plasmid pWW15 contains two nonhomologous, independently regulated catechol 2,3-dioxygenase genes. *J Bacteriol* **163**, 248–255.

Kelley, I. & Cerniglia, C. E. (1991). The metabolism of fluoranthene by a species of *Mycobacterium*. *J Indust Microbiol* **7**, 19–26.

Kelley, I., Freeman, J. P., Evans, F. E., Cerniglia, C. E. (1993). Identification of metabolites from the degradation of fluoranthene by *Mycobacterium* sp. strain PYR-1. *Appl Environ Microbiol* **59**, 800–806.

Kesseler, M., Dabbs, E. R., Averhoff, B., Gottschalk, G. (1996). Studies on the isopropylbenzene 2,3-dioxygenase and the 3-iso-propylcatechol 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. *Microbiology* **142**, 3241–3251.

Kim, S., Shin, H-J., Kim, Y., Kim, S. J., Kim, Y-C. (1997a). Nucleotide sequence of the *Pseudomonas* sp. DJ77 *phnG* gene encoding 2-hydroxymuconic semialdehyde dehydrogenase. *Biochem Biophys Res Commun* **240**, 41–45.

Kim, E. & Zylstra, G. J. (1995). Molecular and biochemical characterisation of two *meta*-cleavage dioxygenases involved in biphenyl and *m*-xylene degradation by *Beijerinckia* sp. strain B1. *J Bacteriol* **177**, 3095–3103.

Kim, E., Zylstra, G. J., Freeman, J. P., Heinze, T. M., Deck, J., Cerniglia, C. E. (1997b). Evidence for the role of 2-hydroxychromene-2-carboxylate isomerase in the degradation of anthracene by *Sphingomonas yanoikuyae* B1. *FEMS Microbiol Lett* **153**, 479–484.

Kitayama, A., Achioku, T., Yanagawa, T., Kanou, K., Kikuchi, M., Ueda, H., Suzuki, E., Nishimura, H., Nagamune, T., Kawakami, Y. (1996). Cloning and characterisation of extradiol aromatic ring-cleavage dioxygenases of *Pseudomonas aeruginosa* JI104. *J Ferment Bioeng* **82**, 217–223.

Kiyohara, H., Torigoe, S., Kaida, N., Asaki, T., Iida, T., Hayashi, H., Takizawa, N. (1994). Cloning and characterisation of a chromosomal gene cluster, *pah*, that encodes the upper pathway for phenanthrene and naphthalene utilisation by *Pseudomonas putida* OUS82. *J Bacteriol* **176**, 2439–2443.

Koreeda, M., Akhtar, M. N., Boyd, D. R., Neill, J. D., Gibson, D. T., Jerina, D. M. (1978). Absolute stereochemistry of *cis*-1,2-, *trans*-1,2 and *cis*-3,4-dihydrodiol metabolites of phenanthrene. *J Org Chem* **43**, 1023–1025.

Kukor, J. J. & Olsen, R. H. (1996). Catechol 2,3-dioxygenases functional in oxygen-limited (hypoxic) environments. *Appl Environ Microbiol* **62**, 1728–1740.

Kuhm, A. E., Stolz, A., Knackmuss, H. J. (1991a). Metabolism of naphthalene by the biphenyl-degrading bacterium *Pseudomonas paucimoblis* Q1. *Biodegradation* **2**, 115–120.

Kuhm, A. E., Stolz, A., Ngai, K-L., Knackmuss, H. J. (1991b). Purification and characterisation of a 1,2-dihydroxynaphthalene dioxygenase from a bacterium that degrades naphthalenesulfonic acids. *J Bacteriol* **173**, 3795–3802.

Kurkela, S., Lehvaslaiho, H., Oalva, E. T., Teeri, T. H. (1988). Cloning, nucleotide sequencing and characterisation of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB 9816. *Gene* **73**, 355–362.

Laemmli, U. K. (1970). Cleavage of strutural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lane, D. J. (1991). 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp. 115–175. Edited by E. Stackebrandt & M. Goodfellow. Chichester, UK: John Wiley & Sons.

Lau, P. C. K., Wang, Y., Labbé, D., Bergeron, H., Garnon, J. (1996). Two-component signal transduction systems regulating toluene and biphenyl-polychlorinated biphenyl degradations in a soil pseudomonad and an actinomycete. In *Molecular Biology of Pseudomonads*, pp. 176-187. Edited by T. Nakazawa *et al.* Washington, DC: American Society for Microbiology.

Lemke, M. J., Brown, B. J., Leff, L. G. (1997). The response of three bacterial populations to pollution in a stream. *Microb Ecol* **34**, 224-231.

Lessie, T. G., Hendrickson, W., Manning, B. D., Devereux, R. (1996). Genomic complexity and plasticity of *Burkholderia cepacia. FEMS Microbiol Lett* **144**, 117–128.

Lloyd-Jones, G., DeJong, C., Ogden, R. C., Duetz, W. A., Williams, P. A. (1994). Recombination of the *bph* (biphenyl) catabolic genes from the plasmid pWW100 and their deletion during growth on benzoate. *Appl Environ Microbiol* **60**, 691–696.

Lloyd-Jones, G., Ogden, R. C., Williams P. A. (1995). Inactivation of 2,3-dihydroxybiphenyl 1,2-dioxygenase from *Pseudomonas* sp. strain CB406 by 3,4-dihydroxybiphenyl. *Biodegradation* **6**, 11–17.

Ma, Y. & Herson, D. S. (1996). GenBank accession no. U47111.

Mahaffey, W. R., Gibson, D. T., Cerniglia, C. E. (1988). Bacterial oxidation of chemical carcinogens: formation of polycyclic aromatic acids from benz[*a*]anthracene. *Appl Environ Microbiol* **54**, 2415–2423.

Maidak, B. L., Olsen, G. J., Larsen, N., Overbeek, R., Mccaughey, M. J., Woese, C. R. (1997). The RDP (Ribosomal Database Project). *Nucleic Acids Res* **25**, 109–110.

Masai E., Yamada, A., Healy, J. M., Hatta, T., Kimbara, K., Fukuda, M., Keiji, Y. (1995). Characterisation of biphenyl catabolic genes of Gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. *Appl Environ Microbiol* **61**, 2079–2085.

Mason J. R. & Cammack, R. (1992). The electron transport proteins of hydroxylating bacterial dioxygenases. *Ann Rev Microbiol* **46**, 277–305.

McClure, N. C., Saint, C. P., Weightman, A. S. (1991). GenBank accession no. X59790.

Means, J. C., Wood, S. G., Hassett, J. J., Banwort, W. L. (1980). Sorption of polynuclear aromatic hydrocarbons by sediments and soils. *Environ Sci Technol* **14**, 1524–1528.

Menn, F-M., Applegate, B. M., Sayler, G. S. (1993). NAH plasmid-mediated catabolism of anthracene and phenanthrene to napthoic acids. *Appl Environ Microbiol* **59**, 1938-1942.

Menzie, C. A., Potocki, B. B., Santodonato, J. (1992). Exposure to carcinogenic PAHs in the environment. *Environ Sci Technol* **26**, 1278–1283.

Merrick, M. J. (1993). In a class of its own – the RNA polymerase factor σ^{54} (σ^{N}). *Mol Microbiol* **10**, 903–909.

Monna, L., Omori, T., Kodama, T. (1993). Microbial degradation of dibenzofuran, fluorene, and dibenzo-*p*-dioxin by *Staphylococcus auriculans* DBF63. *Appl Environ Microbiol* **59**, 285–289.

Morett, E. & Segovia, L. (1993). The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol* **175**, 6067–6074.

Mueller, J. G., Chapman, P. J., Blattmann, B. O., Pritchard, P. H. (1990). Isolation and characterisation of a fluoranthene-utilizing strain *of Pseudomonas paucimobilis. Appl Environ Microbiol* **56**, 1079–1086.

Mueller, J. G., Devereux, R., Santavy, D. L., Lantz, S. E., Willis, S. G., Pritchard, P. H. (1997). Phylogenetic and physiological comparisons of PAH-degrading bacteria from geographically diverse soils. *Antonie Leeuwenhoek* **71**, 329–343.

Mueller, J. G., Lantz, S. E., Devereux, R., Berg, J. D., Pritchard, P. H. (1993). Studies on the microbial ecology of PAH biodegradation. In *Proceedings, Battelle's* 'in situ *and on-site bioreclamation: the second international symposium'* April 5–8, 1993, San Diego, California.

Muller, C., Petruschka, L., Cuypers, H., Burchhardt, G., Herrmann, H. (1996). Carbon catabolite repression of phenol degradation in *Pseudomonas putida* is mediated by the inhibition of the activator protein PhlR. *J Bacteriol* **178**, 2030–2036.

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.

Nakatsu, C., Ng, J., Singh, R., Straus, N., Wyndham, C. (1991). Chlorobenzoate catabolic transposon Tn*5271* is a composite class I element with flanking class II insertion sequences. *Proc Natl Acad Sci USA* **88**, 8312–8316.

Nakazawa, T., Inouye, S., Nakazawa, A. (1990). Regulatory systems for the expression of *xyl* genes on the TOL plasmid. In Pseudomonas*: biotransformations, pathogenesis, and evolving biotechnology,* pp. 133–140. Edited by S. Silver, A. M. Chakrabarty, B. Iglewski & S. Kaplan. Washington, DC: American Society for Microbiology.

Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M., Harayama, S. (1991). Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. *J Bacteriol* **173**, 5385–5395.

Ng, L. C., Oneill, E., Shingler, V. (1996). Genetic evidence for interdomain regulation of the phenol-responsive σ^{54}-dependent activator DmpR. *J Biol Chem* **271**, 17281–17286.

Ng, L. C., Poh, C. L., Shingler, V. (1995). Aromatic effector activation of the NtrC-like transcriptional regulator PhhR limits the catabolic potential of the (methyl)phenol degradative pathway it controls. *J. Bacteriol* **177**, 1485–1490.

Ng, L. C., Shingler, V., Sze, C. C., Poh, C. L. (1994). Cloning and sequences of the first eight genes of the chromosomally encoded (methyl)phenol degradation pathway from *Pseudomonas putida* P35X. *Gene* **151**, 29–36.

Nishioka, M., Chang, H-C., Lee, M. L. (1986). Structural characteristics of polycyclic aromatic hydrocarbons in coal tars and combustion products. *Environ Sci Technol* **20**, 1023–1027.

Noda, Y., Nishikawa, S., Shiozuka, K., Kadokura, H., Nakajima, H., Yoda, K., Katayama, Y., Morohoshi, N., Haraguchi, T., Yamasaki, M. (1990). Molecular cloning of the protocatechuate 4,5-dioxygenase genes of *Pseudomonas paucimobilis. J Bacteriol* **172**, 2704–2709.

Nordlund, I. & Shingler, V. (1990). Nucleotide sequences of the *meta*-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxy-muconic semialdehyde hydrolase from *Pseudomonas* CF600. *Biochem Biophys Acta* **1049**, 227–230.

Nozaki, M. (1970). Metapyrocatechase (*Pseudomonas*). *Methods Enzymol* **17A**, 522–525.

Palleroni, N. J. & Holmes, B. (1981). *Pseudomonas cepacia* sp. nov., nom. rev. *Int J Syst Bacteriol* **31**, 479–481.

Parales, R. E., Emig, M. D., Lynch, N. A., Gibson, D. T. (1998a). Substrate specificity of hybrid naphthalene and 2,4-dinitrotoluene dioxygenase enzyme systems. *J Bacteriol* **180**, 2337–2344.

Parales, J. V., Kumar, A., Parales, R. E., Gibson, D. T. (1996). Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from *Pseudomonas* sp. JS42. *Gene* **181**, 57–61.

Parales, J. V., Parales, R. E., Resnick, S. M., Gibson, D. T. (1998b). Enzyme specificity of 2-nitrotoluene 2,3-dioxygenase from *Pseudomonas* sp. strain JS42 is determined by the C-terminal region of the α subunit of the oxygenase component. *J Bacteriol* **180**, 1194–1199.

Park, K. S., Sims, R. C., Dupont, R. R., Doucette, W. J., Matthews, J. E. (1990). Fate of PAH compounds in two soil types: infuence of volatilisation, abiotic loss and biological activity. *Environ Toxicol Chem* **9**, 187–195.

Patel, T. R. & Barnsley, E. A. (1980) Naphthalene metabolism by Pseudomonads: purification and properties of 1,2-dihydroxynaphthalene oxygenase. *J Bacteriol* **143**, 668–673.

Patel, T. R. & Gibson, D. T. (1974). Purification and properties of (+)-*cis*-naphthalene dihydrodiol dehydrogenase of *Pseudomonas putida*. *J Bacteriol* **119**, 879–888.

Pflugmacher, U., Averhoff, B., Gottschalk, G. (1996). Cloning, sequencing, and expression of isopropylbenzene degradation genes from *Pseudomonas* sp. strain JR1: identification of isopropylbenzene dioxygenase that mediates trichloroethene oxidation. *Appl Environ Microbiol* **62**, 3967–3977.

Pipke, R., Wagner-Döbler, I., Timmis, K. N., Dwyer, D. F. (1992). Survival and function of a genetically engineered Pseudomonad in aquatic sediment microcosms. *Appl Environ Microbiol* **58**, 1259–1265.

Polissi, A. & Harayama, S. (1993). *In vivo* reactivation of catechol 2,3-dioxygenase mediated by a chloroplast-type ferredoxin: a bacterial strategy to expand the substrate specificity of aromatic degradative pathways. *EMBO J* **12**, 3339–3347.

Pothuluri, J. V. & Cerniglia, C. E. (1994). Microbial metabolism of polycyclic aromatic hydrocarbons. In *Biological degradation and bioremediation of toxic chemicals*, pp. 92–124. Edited by R. G. Chaudhry. Oregon: Dioscorides Press.

Ramos, J. L., Marques, S., Timmis, K. N. (1997). Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Ann Rev Microbiol* **51**, 341–373.

Rieske, J. S., Maclennan, D. H., Coleman, R. (1964). Isolation and properties of an iron protein from the (reduced coenzyme Q)-cytochrome C reductase complex of the respiratory chain. *Biochem Biophys Res Commun* **15**, 338–344.

Romero-Arroyo, C. E., Schell, M. A., Gaines III, G. L., Neidle, E. L. (1995). *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. *J Bacteriol* **177**, 5891–5898.

Rosenberger, R. F. & Elsden, S. R. (1960). The yields of *Streptococcus faecalis* grown in continuous culture. *J Gen Microbiol* **22**, 726–739.

Rothmel, R. K., Aldrich, T. L., Houghton, J. E., Coco, W. M., Ornston, L. N., Chakrabarty, A. M. (1990). Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the *lysR* family. *J Bacteriol* **172**, 922–931.

Rushing, B. G., Yelton, M. M., Long, S. R. (1991). Genetic and physical analysis of the *nodD3* region of *Rhizobium meliloti*. *Nucleic Acids Res* **19**, 921–927.

Sambrook, J., Fritsch, E. F., Maniatis T. (1989). *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbour Laboratory Press.

Sanseverino, J., Applegate, B. M., Henry King, J. M., Sayler, G. S. (1993a). Plasmid-mediated mineralisation of naphthalene, phenanthrene, and anthracene. *Appl Environ Microbiol* **59**, 1931–1937.

Sanseverino, J., Werner, C., Fleming, J., Applegate, B., King, J. M. H., Saylor, G. S. (1993b). Molecular diagnostics of polycyclic aromatic hydrocarbon biodegradation in manufactured gas plant soils. *Biodegradation* **4**, 303-321.

Sato, S. I., Nam, J. W., Kasuga, K., Nojiri, H., Yamane, H., Omori,T. (1997a). Identification and characterisation of genes encoding carbazole 1,9a-dioxygenase in *Pseudomonas* sp. strain CA10. *J Bacteriol* **179**, 4850–4858.

Sato, S. I., Ouchiyama, N., Kimura, T., Nojiri, H., Yamane, H., Omori, T. (1997b). Cloning of genes involved in carbazole degradation of *Pseudomonas* sp. strain CA10: nucleotide sequences of genes and characterisation of *meta*-cleavage enzymes and hydrolase. *J Bacteriol* **179**, 4841–4849.

Saylor, G. S., Hooper, S. W., Layton, A. C., Henry King, J. M. (1990). Catabolic plasmids of environmental and ecological significance. *Microb Ecol* **19**, 1–20.

Schell, M. A. (1983). Cloning and expression in *Escherichia coli* of the naphthalene degradation genes from plasmid NAH7. *J Bacteriol* **153**, 822–829.

Schell, M. A. (1985). Transcriptional control of the *nah* and *sal* hydrocarbon-degrading operons by the *nahR* gene product. *Gene* **36**, 301–304.

Schell, M. A. (1986). Homology between nucleotide sequences of promoter regions of *nah* and *sal* operons of NAH7 plasmid of *Pseudomonas putida*. *Proc Natl Acad Sci USA* **83**, 369–372.

Schell, M. A. (1993). Molecular biology of LysR family of transcriptional regulators. *Annu Rev Microbiol* **47**, 597–626.

Schell, M. A. & Sukordhaman, M. (1989). Evidence that the transcription activator encoded by *nahR* gene of *Pseudomonas putida* is evolutionarily related to transcriptional activators encoded by the *Rhizobium nodD* genes. *J Bacteriol* **171**, 1952–1959.

Schell, M. A. & Wender, P. E. (1986). Identification of the *nahR* gene product and nucleotide sequences required for its activation of the *sal* operon. *J Bacteriol* **166**, 9–14.

Scott, K. F. (1986). Conserved nodulation genes from the non-legume symbiont *Bradyrhizobium* sp. (Parasponia). *Nucleic Acids Res* **14**, 2905–2919.

Shamsuzzaman, K. M. & Barnsley, E. A. (1974a). The regulation of naphthalene metabolism in pseudomonads. *Biochem Biophys Res Commun.* **60**, 582–589.

Shamsuzzaman, K. M. & Barnsley, E. A. (1974b). The regulation of naphthalene oxygenase in pseudomonads. *J Gen Microbiol* **83**, 165–170.

Shepherd, J. M. & Lloyd-Jones, G. (1998). Novel carbazole degradation genes of *Sphingomonas* CB3: sequence analysis, transcription, and molecular ecology. *Biochem Biophys Res Commun* **247**, 129–135.

Shields, M. S., Reagin, M. J., Gerger, R. R., Campbell, R., Somerville, C. (1995). TOM, a new aromatic degradative plasmid from *Burkholderia* (*Pseudomonas*) *cepacia* G4. *Appl Environ Microbiol* **61**, 1352–1356.

Shine, J. & Dalgarno, L. (1975). Determinant of cistron specificity in bacterial ribosomes. *Nature* **254**, 34–38.

Shingler, V. (1996). Signal sensing by sigma(54)-dependent regulators – derepression as a control mechanism. *Mol Microbiol* **19**, 409–416.

Shingler, V., Bartilson, M., Moore, T. (1993). Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators. *J Bacteriol* **175**, 1596–1604.

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.

Shuttleworth, K. W. & Cerniglia, C. E. (1995). Environmental aspects of PAH degradation. *Appl Biochem Biotechnol* **54**, 291–302.

Simon, M. J., Osslund, T. D., Saunders, R. (1993). Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. *Gene* **127**, 31–37.

Sims, R. C. & Overcash, M. R. (1983). Fate of polynuclear aromatic hydrocarbons (PNAs) in soil-plant systems. *Residue Reviews* **88**, 1–68.

Sousa, C., Folch, J. L., Boloix, P., Megias, M., Nava, N., Quinto, C. (1993). A *Rhizobium tropici* DNA region carrying the amino-terminal half of a *nodD* gene and a *nod*-box-like sequence confers host-range extension. *Mol Microbiol* **9**, 1157–1168.

Smith, M. R. (1990). The biodegradation of aromatic hydrocarbons by bacteria. *Biodegradation* **1**, 191–206.

Spence, E. L., Kawamukai, M., Sanvoison, J., Braven, H., Bugg, T. D. H. (1996). Catechol dioxygenases from *Escherichia coli* (MhpB) and *Alcaligenes eutrophus* (MpcI): sequence analysis and biochemical properties of a third family of extradiol dioxygenases. *J Bacteriol* **178**, 5249–5256.

Springael, D., Kreps, S., Mergeay, M. (1993). Identification of a catabolic transposon Tn*4371* carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. *J Bacteriol* **175**, 1674–1681.

Stanier, R. Y., Palleroni, N. J., Doudoroff, M. (1966). The aerobic pseudomonads: a taxonomic study. *J Gen Microbiol* **43**, 159–271.

Suen, W-C., Haigler, B. E., Spain, J. C. (1996). 2,4-dinitrotoluene dioxygenase from *Burkholderia* sp. strain DNT: similarity to naphthalene dioxygenase. *J Bacteriol* **178**, 4926–4934.

Sylvestre, M., Sirois, M., Hurtubise, Y., Bergeron, J., Ahmad, D., Shareck, F., Barriault, D., Guillemette, I., Juteau, J. M. (1996). Sequencing of *Comamonas testosteroni* strain B-356 biphenyl/chlorobiphenyl dioxygenase genes: evolutionary relationships among Gram-negative bacterial biphenyl dioxygenases. *Gene* **174**, 195–202.

Taira, K., Hayase, N., Arimura, N., Yamashita, S., Miyazaki, T., Furukawa, K. (1988). Cloning and nucleotide sequence of the 2,3-dihydroxybiphenyl dioxygenase gene from the PCB-degrading strain of *Pseudomonas paucimobilis* Q1. *Biochemistry* **27**, 3990–3996.

Taira, K., Hirose, J., Hayashida, S., Furukawa, K. (1992). Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. *J Biol Chem* **267**, 4844–4853.

Takizawa, N., Iida, T., Yamauchi, K., Satoh, S., Wang, Y., Fukuda, M., Kiyohara, H. (1996). GenBank accession no. D84146.

Takizawa, N., Kaida, N., Torigoe, S., Moritani, T., Sawada, T., Satoh, S., Kiyohara, H. (1994). Identification and characterisation of genes encoding polycyclic aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase in *Pseudomonas putida* OUS82. *J Bacteriol* **176**, 2444–2449.

Tan, H. M., Tang, H. Y., Joannou, C. L., Abdel-Wahab, N. H., Mason, J. R. (1993). The *Pseudomonas putida* ML2 plasmid-encoded genes for benzene dioxygenase are unusual in codon usage and low in G+C content. *Gene* **130**, 33–39.

Thompson, J. D., Higgins, D. G., Gibson T. J. (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Timmis, K. N. (1997). Abstract I. In: VI International Congress on *Pseudomonas*: Molecular Biology and Biotechnology, 4–8 September, 1997, Madrid, Spain.

Torsvik, V., Goksoyr, J., Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Appl Environ Microbiol* **56**, 782–787.

Tsuda, M. & Iino, T. (1990). Naphthalene degrading genes on plasmid NAH7 are on a defective transposon. *Mol Gen Genet* **223**, 33–39.

Urakami, T., Ito-Yoshida, C., Araki, H., Kijima, T., Suzuki, K-I., Komagata, K. (1994). Transfer of *Pseudomonas plantarii* and *Pseudomonas glumae* to *Burholderia* as *Burkholderia* spp. and description of *Burkholderia vandii* sp. nov. *Int J Syst Bacteriol* **44**, 235–245.

van der Meer, J. R., Zehnder, A. J. B., de Vos, W. M. (1991). Identification of a novel composite transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. *J Bacteriol* **173**, 7077–7083.

Walter, U., Beyer, M., Klein, J., Rehm, H-J. (1991). Degradation of pyrene by *Rhodococcus* sp. UW1. *Appl Microbiol Biotechnol* **34**, 671–676.

Wang, Y., Garnon, J., Labbé, D., Bergeron, H., Lau, P. C. K. (1995). Sequence and expression of the *bpdC1C2BADE* genes involved in the initial steps of biphenyl/chlorobiphenyl degradation by *Rhodococcus* sp. M5. *Gene* **164**, 117–122.

Wang, Y., Lau, P. C. K., Button, D. K. (1996). A marine oligobacterium harboring genes known to be part of aromatic hydrocarbon degradation pathways of soil pseudomonads. *Appl Environ Microbiol* **62**, 2169–2173.

Wedel, A. & Kustu, S. (1995). The bacterial enhancer-binding protein NtrC is a molecular machine – ATP hydrolysis is coupled to transcriptional activation. *Genes & Dev* **9**, 2042–2052.

Weissenfels, W. D., Beyer, M., Klein, J., Rehm, H. J. (1991). Microbial metabolism of fluoranthene: identification of ring fission products. *Appl Microbiol Biotechnol* **34**, 528–535.

Weissenfels, W. D., Beyer, M., Klein, J. (1990). Degradation of phenanthrene, fluorene and fluoranthene by pure bacterial cultures. *Appl Microbiol Biotechnol* **32**, 479–484.

Werlen, C., Kohler, H. P., van der Meer, J. R. (1996). The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. *J Biol Chem* **271**, 4009–4016.

Wheatcroft, R. & Williams, P. A. (1981). Rapid methods for the study of both stable and unstable plasmids in *Pseudomonas*. *J Gen Microbiol* **124**, 433–437.

Williams, P. A. & Sayers, J. R. (1994). The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. *Biodegradation* **5**, 195–217.

Wilson, S. C. & Jones, K. C. (1993). Bioremediation of soil contaminated with polynuclear aromatic hydrocarbons (PAHs): a review. *Environ Pollut* **81**, 229–249.

Yabuuchi, E., Kosako, Y., Oyaizu, H., Yanu, I., Hotta, H., Hashimoto, Y., Ezaki, T., Arakawa, M. (1992). Proposal of *Burholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiol Immunol* **36**, 1251–1275.

Yang, Y., Chen, R. F., Shiaris, M. P. (1994). Metabolism of naphthalene, fluorene, and phenanthrene: preliminary characterisation of a cloned gene cluster from *Pseudomonas putida* NCIB 9816. *J Bacteriol* **176**, 2158–2164.

Yen, K-M. & Gunsalus, I. C. (1982). Plasmid gene organisation: naphthalene/ salicylate oxidation. *Proc Natl Acad Sci USA* **79**, 874–878.

Yen, K-M. & Gunsalus, I. C. (1985). Regulation of naphthalene catabolic genes of plasmid NAH7. *J Bacteriol* **162**, 1008–1012.

Yen, K-M. & Serdar, C. M. (1988). Genetics of naphthalene catabolism in pseudomonads. *CRC Crit Rev Microbiol* **15**, 247–268.

Yrjälä, K., Paulin, L., Kilpi, S., Romantschuk, M. (1994). Cloning of *cmpE*, a plasmid-borne catechol 2,3-dioxygenase-encoding gene from the aromatic- and chloroaromatic-degrading *Pseudomonas* sp. HV3. *Gene* **138**, 119–121.

Yrjälä, K., Paulin, L., Kilpi, S., Romantschuk, M. (1997). Novel organisation of catechol *meta*-pathway genes in *Sphingomonas* sp. HV3 pSKY4 plasmid. *FEMS Microbiol Lett* **154**, 403–408.

Zagursky, R. J., Baumeister, K., Lomax, N., Berman, M. L. (1985). Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. *Gene Anal Tech* **2**, 89–94.

Zylstra, G. J. & Gibson, D. T. (1989). Toluene degradation by *Pseudomonas putida* F1. Nucleotide sequence of the *todC1C2BADE* genes and their expression in *E. coli*. *J Biol Chem* **264**, 14940–14946.

Zylstra, G. J., Kim, E., Goyal, A. K. (1997). Comparative molecular analysis of genes for polycyclic aromatic hydrocarbon degradation. *Genet Eng* **19**, 257–269.

AMENDMENTS December 1998

List of Abbreviations

bp	base pair
BSA	bovine serum albumin
C23O	catechol 2,3-dioxygenase
CTAB	hexadecyl trimethyl ammonium bromide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
dpm	disintegrations per minute
FAD	flavin adenine dinucleotide
GC	Gas chromatography
HPLC	high performance liquid chromatography
ISP	iron-sulphur protein
kb	kilobase
LB	Luria Bertani medium
LSC	liquid scintillation counting
MCS	multiple cloning site
MM	minimal media
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
ORF	open reading frame
PAH	polycyclic aromatic hydrocarbon
PCB	polychlorinated biphenyl
PCR	polymerase chain reaction
PVPP	polyvinylpolypyrrolidone
RNA	ribonucleic acid
RT-PCR	reverse transcriptase-polymerase chain reaction
SDS	sodiumdodecylsulphate
SDS-PAGE	sodiumdodecylsulphate-polyacrylamide gel electrophoresis
UV	ultraviolet
½PC	half-strength plate count medium

p70, line 7 *P. pseudoalcaligenes* LB400 is normally regarded as a PCB degrader.

Reference omitted:

Lloyd-Jones, G. & Lau, P. C. K. (1998). A molecular view of microbial diversity in a dynamic landfill in Quebec. *FEMS Microbiol Lett* **162**, 219–226.