The nucleotide sequence of the Tn5271 3-chlorobenzoate 3,4-dioxygenase genes (cbaAB) unites the class IA oxygenases in a single lineage

Cindy H. Nakatsu,† Neil A. Straus and R. Campbell Wyndham

Author for correspondence: R. Campbell Wyndham. Tel: +1 613 788 2600 ext. 3651. Fax: +1 613 788 3539. e-mail: cwyndham@ccs.carleton.ca

The Institute of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6

1 Institute of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6

2 Department of Botany, University of Toronto, Toronto, Ontario, Canada M5S 2B3

INTRODUCTION

Complex mixtures of organic compounds found in contaminated soils and surface waters are a powerful selective force for the evolution of bacterial catabolic pathways. Many isolates have been recovered from these environments on the basis of their ability to utilize selected contaminants as their sole source of carbon and energy (Sayler et al., 1990; van der Meer et al., 1992). While the aromatic catabolic determinants of some genera have been thoroughly studied, particularly the fluorescent pseudomonads and Acinetobacter (Gibson & Subramanian, 1984; Frantz & Chakrabarty, 1986; Reineke & Knackmuss, 1988; Harayama et al., 1992), other genera are not well characterized. As new catabolic gene sequences from diverse contaminant-degrading species are determined, new insights into evolutionary lineages and origins of these elements are revealed (Neidle et al., 1991).

The metabolism of chlorobenzoates illustrates the diversity of pathways for contaminant degradation that is likely to be encountered in natural communities or industrial consortia. There are now several known pathways for the metabolism of chlorobenzoates among the aerobic proteobacteria. The chlorobenzoate 1,2-dioxygenase pathway, wherein oxygen is introduced at the carboxyl and adjacent positions of the chlorobenzoate...
ring, forming chlorocatechols, has been thoroughly studied (Reineke & Knackmuss, 1988; Hickey & Focht, 1990; Harayama et al., 1992). This pathway is initiated either by benzoate (toluate) 1,2-dioxygenases with broad substrate specificity (Fetzner et al., 1992), or by ortho-halobenzoate 1,2-dioxygenases that are specific for ortho-substituted halobenzoates (Romanov & Hausinger, 1994). An operon representative of the former, ygiXYZ from the Pseudomonas putida plasmid pWW0, has been sequenced and the deduced amino acid sequences have been compared to the narrow-substrate-range benzoate 1,2-dioxygenase of Acinetobacter calcoaceticus (benABC) (Harayama et al., 1991; Neidle et al., 1991). Sequences of chlorocatechol ortho-ring-fission genes (clcABD) from P. putida (pAC27), required for complete metabolism of chlorobenzoates by this pathway, have also been determined. They reveal similarities to the corresponding catechol ortho-ring-fission genes found in Pseudomonas and Acinetobacter, and to other chloroaromatic ortho-ring-fission genes (Frantz & Chakraborty, 1987; Neidle et al., 1988; van der Meer et al., 1991, 1992; Harayama et al., 1992). Another known pathway involves the removal of a chlorine substituent by hydrolytic dehalogenation. This pathway functions in Pseudomonas sp. CBS-3 growing on 4-chlorobenzoate (Savard et al., 1986; Scholten et al., 1991). An early report on the metabolism of 3-chlorobenzoate in Pseudomonas attributed a similar transformation, to form 3-hydroxybenzoate, to a monoxygenase (Johnston et al., 1972). Very little is known concerning the biochemistry and genetics of the latter pathway. Recently we described a new pathway in Alcaligenes sp. strain BR60 that expresses a 3-chlorobenzoate 3,4-dioxygenase activity (Nakatsu & Wyndham, 1993). This dioxygenase introduces oxygen at the carbons of benzoate distal to the carboxyl group, yielding protocatechuate and chloroprotocatechuate metabolites. The genes for this pathway are carried on a 3-7 kb NotI–EcoRI fragment within the catabolic transposon Tn5271, originally found on the 85 kb conjugative plasmid pBRC60 (Wyndham et al., 1988, 1994; Nakatsu et al., 1991; Nakatsu & Wyndham, 1993).

Oxygogenases that are involved in the initial attack on aromatic substrates have been grouped according to the nature of the oxidation they carry out, and according to the properties of their electron transport proteins (Harayama et al., 1992; Mason & Cammack, 1992). The aromatic-ring-hydroxylating dioxygenases have been classified into two-component systems and three-component systems, based on the number of proteins involved in electron transport and hydroxylation (Fetzner et al., 1992). These broad classes have been further subdivided according to the detailed arrangements and cofactor requirements of their reductase and oxygenase components (Batie et al., 1991; reviewed by Mason & Cammack, 1992). Briefly, these classes and their distinguishing features are: class I, two-component oxygenases in which a flavin and iron–sulphur cluster are combined in the same reductase (class IA: FMN cofactor; class IB: FAD cofactor); class II, three-component dioxygenases in which the flavin (reductases) and iron-sulphur proteins (ferredoxins) are separate (class IIA: chloroplat-in-type ferredoxins; class IIB: Rieske-type ferredoxins); class III, three-component dioxygenase in which the reductase and ferredoxin are separate, but the reductase contains both a flavin and an iron–sulphur cluster. All of these classes have terminal oxygenase components with two common cofactors, a Rieske-type [2Fe–2S]R iron–sulphur centre and one mononuclear non-haem iron.

In order to place the 3-chlorobenzoate 3,4-dioxygenase of Tn5271 into the context of the evolution of bacterial oxygenases, we determined the activity expressed by the genes designated cbaAB from Alcaligenes sp. BR60, and their nucleotide sequence. From this sequence we derived putative amino acid sequences and compared these with known sequences for other oxygenases. This comparison revealed a clearly defined lineage among the biochemically defined class IA oxygenases that shows several unique features.

**METHODS**

**Strains, plasmids and growth conditions.** The growth and maintenance of Alcaligenes sp. strain BR60 (pBRC60), strain BR6024 (chloramphenicol resistant, tryptophan auxotroph) and Escherichia coli strains have been described previously (Wyndham et al., 1988; Wyndham & Straus, 1988; Nakatsu et al., 1991; Nakatsu & Wyndham, 1993). E. coli strain DH5α was used as the host for the broad-host-range pBW13 constructs pBRCN5, pBRCN7, pBRCN11, pBRCN12, pBRCN15 and pBRCN16, containing the dioxygenase coding region of Tn5271 and deletion and point mutations of this region (Nakatsu & Wyndham, 1993; this study). E. coli strain JM109 was the host for the pUC18 plasmid construct pBRH2, containing the cloned HindIII restriction fragment H2 of Tn5271 (Wyndham et al., 1988). E. coli strain HB101 was used as the host for single-stranded DNA synthesis from clones in M13mp18 and M13mp19 (Yanisch-Perron et al., 1985).

**DNA cloning, sequencing and mutation.** DNA extraction and cloning was performed according to previously described methods (Nakatsu et al., 1991). The HindIII fragment H2 of Tn5271, originally cloned into the HindIII site of pUC18 to give pBRH2 (Wyndham et al., 1988), was the source of the DNA sequenced in this study. Plasmid pBRH2 was digested with EcoRI and the fragments E11 (2.9 kb), E13 (1.6 kb) and E14 (1.5 kb) were ligated into the replicative forms of phage M13mp18 and M13mp19 in both orientations. These constructs were transformed into E. coli HB101 to produce single-stranded DNA for sequencing in both directions using the commercial dideoxy chain-termination Sequenase kit (United States Biochemical). Oligonucleotides for sequencing were prepared as described by Nakatsu et al. (1991).

The cba gene region was mutagenized to determine the requirements for chlorobenzoate dioxygenase activity by creating a deletion and two site mutations, followed by the evaluation of expression in vivo. Unique Sva restriction sites within cbaA were used to create a deletion of the internal fragment of pBRCN5 in vitro, by restriction digestion and ligation, followed by transformation into E. coli DH5a and triparental mating into Alcaligenes sp. BR6024 (Nakatsu & Wyndham, 1993). The only known unique site within the second ORF (cbaB) was for EcoRV, but this site also existed within the vector pBW13. Therefore, the PstI–HindIII frag-
ment of pBRH2 was cloned into pUC18, then the EcoRV site was mutated by restriction digestion and treatment with exonuclease III for 1 s according to the manufacturer's directions (Boehringer Mannheim). The site was then filled, religated and the plasmid was transformed into E. coli DH5α as described (Sambrook et al., 1989). The modified fragment was then excised from pUC18, cloned into the broad-host-range vector pBW13 and conjugated into Alcaligenes sp. strain BR6024 as described previously (Nakatsu & Wyndham, 1993). A third mutation was created 3' to the cbaB ORF by digesting pBRCN5 with BstEI and treating with exonuclease III as described above. The presence of the specified modifications to all plasmids was confirmed by plasmid isolation and restriction mapping.

Sequence analyses. The National Center for Biotechnology Information (NCBI) BLAST and BLASTX programs were used to search for similarities to the cbaAB nucleotide sequences and deduced amino acid sequences (Altschul et al., 1990). The non-redundant components of peptide sequence databases, including SwissProt, PIR, GenPept (translations from GenBank) and GUpdate (cumulative daily updates to the major release), were searched. The GCG sequence analysis programs of the University of Wisconsin (Devereux et al., 1984), provided through the Molecular Biology Database System of the National Research Council of Canada, were used for all analyses. Amino acid sequence alignments were carried out using the GAP, BESTFIT and PILEUP programs of the GCG7 package. For these alignments a gap weight of 3-0 and a gap length weight of 0-1 were used. Complete sequences were used for alignments within the class IA oxygenase group containing Cba, Van and Phe protein sequences. Complete sequences and partial sequences were used for alignments between amino acid sequences of class IA and the other classes of oxygenases as defined by Batie et al. (1991). The sequences of the 200 amino acids spanning the Rieske-type [2Fe-2S]_{R} domains of the terminal oxygenase components were aligned to show similarities across oxygenase classes. Similarly, 200 amino acids containing the contiguous FMN- (FAD-) isoxaloazine binding and NAD-ribose-binding domains of the reductase components, and 110 amino acids spanning the plant ferredoxin-like [2Fe-2S] FMF domains of these same components, were aligned individually. The PILEUP program uses a progressive alignment method the order of which is based on pairwise similarity scores for each pair of sequences. The pairwise similarity scores were used to cluster the sequences by similarity. Relative evolutionary distances were presented graphically as dendrograms using these pairwise similarity scores.

RESULTS

Localization and nucleotide sequence of the cbaAB genes

A restriction digest map of Tn5271, showing the flanking IS1071 elements and the non-repeated region carrying the cba-AB (3-chlorobenzoate 3,4-dioxygenase) genes is presented in Fig. 1. This figure illustrates the subclones constructed to determine the function of the cbaAB genes. The nucleotide sequence from position 4600 to 6938 is presented in Fig. 2. The nucleotide numbering system begins with nucleotide 1 as the first G in the left inverted repeat of the IS1071 L element of Tn5271 (Fig. 1; GenBank accession number M65135; Nakatsu et al., 1991). Previous studies (Nakatsu & Wyndham, 1993) demonstrated that unique protein products were produced only in the left to right (PstI to HindIII) orientation, the upper strand in this study, therefore only the deduced amino acid sequences of the ORFs of the upper strand are included in Fig. 2. The ORFs were designated cbaA and cbaB.

Conjugation of the pBRCN5 construct containing these ORFs plus flanking DNA into Alcaligenes sp. BR6024 restored the ability of this mutant to grow on 3-chlorobenzoate. Elimination of the Scal fragment within the cbaA ORF to give the pBRCN7 plasmid resulted in a failure to restore the degradative phenotype.

The clone pBRCN15 was created by eliminating the EcoRV restriction enzyme site within cbaB of pBRCN5 by exonuclease digestion. The introduction of this plasmid into Alcaligenes sp. strain BR6024 did not initially produce 3-chlorobenzoate-degrading transconjugants. After prolonged incubation (3 weeks), colonies did appear on the 3-chlorobenzoate selection plates. When these putative transconjugants were patched to fresh 3-chlorobenzoate selection plates, growth of 10% (5/50) of the patched colonies occurred. If the tetracycline resistance marker of pBRCN15 was used for selection of transconjugants on succinate plates, followed by patching of transconjugants onto 3-chlorobenzoate, no growth occurred even after prolonged incubation. These observations suggest CbaB is required for 3-chlorobenzoate 3,4-dioxygenase activity, and that we observed mutational recovery of this activity due to repair of the deletion around the EcoRV site in vivo.

The clone pBRCN16 was created by eliminating the BstEI restriction enzyme site that lies approximately 0.3 kb beyond the stop codon of cbaB. The clones pBRCN11 and pBRCN12 were created by nested deletions in the same region of DNA, 3' to the cbaB gene (Nakatsu & Wyndham, 1993). The introduction of these plasmids into Alcaligenes sp. strain BR6024 did not produce 3-chlorobenzoate-degrading transconjugants. However, the pBRCN16 and pBRCN11 transconjugants metabolized 3-chlorobenzoate to dihydrodiol intermediates that were unstable (Nakatsu & Wyndham, 1993). The major product accumulating in the medium of these transconjugants was a chlorohydroxybenzoate, presumed to be the product of spontaneous dehydrogenation of a chlorohydroxybenzoate intermediate. The pBRCN12 transconjugants failed to metabolize 3-chlorobenzoate, indicating that the nested deletion had eliminated an essential sequence required for dioxygenase expression. Restriction mapping revealed that this deletion extended into the cbaB gene.

Sequence characterization

The first ORF, cbaA, has potential initiation codons at nucleotides 4631, 4685 and 4697 of Tn5271 (Fig. 2). The potential ribosome-binding site preceding nucleotide 4631 is AGGAG, very close to the ribosomal-binding site consensus sequence for E. coli, therefore this was chosen as a tentative initiation codon. The cbaA gene potentially codes for a protein of 432 amino acids with a deduced molecular mass of 48932 Da. The second ORF, cbaB, has
C. H. NAKATSU, N. A. STRAUS and R. C. WYNDHAM

![Diagram of Tn5271 restriction enzyme map](image)

**Fig. 1.** Restriction enzyme digest map of Tn5271 showing the location of the *cbaA* and *cbaB* genes. (a) Map of the entire Tn5271 element. Restriction enzymes used were: Bs, BstEII; E, EcoRI; Ev, EcoRV; H, HindIII; N, NotI; P, PstI; Sc, Scal. (b) The region of Tn5271 cloned into the IncPβ broad-host-range vector pBW13 to form pBRCN5. The *cbaA* and *cbaB* genes are designated by arrows (encoded on the upper strand as drawn, with the 5' end to the left). Downstream of *cbaB* there is another ORF (ORF3) not fully characterized. Distances are marked in kb relative to the numbering system for Tn5271, with nucleotide 1 being the first G in the left inverted repeat of IS1071 L (Nakatsu et al., 1991). The mutations of pBRCN5 used to determine functions of the ORFs are shown below. The construction of pBRCN7, pBRCN11 and pBRCN12 was previously described (Nakatsu & Wyndham, 1993). The pBRCN15 and 16 mutations were created by deletion of nucleotides at restriction enzyme sites for EcoRV and BstEII, respectively.

an initiation codon at nucleotide 5983 with a potential ribosome-binding site with the sequence GGAGG. The *cbaB* gene potentially codes for a 288 amino acid protein with a deduced molecular mass of 31 704 Da.

The G+C composition of *cba* is 58.7% and of *cbaB* is 59.2%, in the expected range for the *Alcaligenes* sp. BR60 family *Comamonadaceae* (57–70% G+C) (Willems et al., 1991). This G+C composition is similar to the 58.8% G+C content for the IS1071 transposase sequence flanking *cbaAB* within Tn5271 (Nakatsu et al., 1991). There is very little information on codon usage by species within the *Comamonadaceae*, however we can comment on the differences in codon usage when compared to codons used by fluorescent *Pseudomonas* species (West & Iglewski, 1988). For example, arginine AGA and AGG are used nine times in the *cbaAB* genes, which is unlike codon usage in the fluorescent pseudomonads. Hayayama et al. (1991) found that there was an exceptional use of these same codons in the *xylX* gene of *P. putida*, prompting them to speculate on alternative origins for this gene. Codons considered rare in *Pseudomonas aeruginosa*, but used frequently in *cbaAB* are: arginine-CGA, alanine-GCA, isoleucine-ATT, leucine-TTG, glycine-GGG and GGA, and valine-GTT. In *cbaA* alanine-GCT and arginine-CGT and AGG, and in *cbaB* asparagine-AAT, cysteine-TGT, isoleucine-ATA and serine-AGT, are all rare codons for *Pseudomonas* that are used in the *cbaAB* genes. These codon usage patterns suggest the *cbaAB* genes do not derive from fluorescent *Pseudomonas* species.

### Alignments of oxygenase amino acid sequences

The **BLASTX** search of peptide sequence databases for similarities to the *cbaA* gene product listed the products of *phb* and *vanA* as the closest matching full sequences (Fig. 3). The *phb12345* genes were described by Nomura et al. (1992). They are found on a 7 kb *EcoRI* fragment of the PHT plasmid of *P. putida* strain NMH102-2 and their expression allows this strain to grow with phthalate as sole source of carbon. The chromosomal *vanAB* genes encode vanillate demethylase activity in a *Pseudomonas* species (Brunel & Davison, 1988). The latter activity has been described as a monooxygenase-catalysed oxidative demethylation. In addition to the above matches, similarity was also detected to a partial amino acid sequence of *C. testosteroni* (Davison et al., 1988). The latter activity has been described as a monooxygenase-catalysed oxidative demethylation. In addition to the above matches, similarity was also detected to a partial amino acid sequence of the NH₂-terminal 35 residues sequenced from the purified component A of the 4-sulphobenzoate 3,4-dioxygenase of *C. testosteroni* strain T2 (Locher et al., 1991). The NH₂-terminal sequence of the 4-sulphobenzoate 3,4-dioxygenase showed 48% identity to CbaA over the available 35 amino acid sequence. This was a much greater similarity than for the CbaA/Pht3 comparison in the same region of the sequence (37%). This suggests that when the complete 4-sulphobenzoate 3,4-dioxygenase sequence
ponents have been shown to bind mononuclear, non-

Within this region there is 58% identity between CbaA with similar regions of the terminal oxygenase and Pht3, and 35% identity between CbaA and V-anA.

... proposed to coordinate this Fe(II) (Neidle... 1982). This cofactor... defined as the first G in the sequence of the nucleotides of an ORF... 1991) is shown. Potential ribosomal-binding sites are underlined and initiation codons are in bold letters. Stop codons are indicated by an asterisk. The first 87 nucleotides of an ORF downstream of cbaB that has been partly characterized are also shown.

Fig. 2. Nucleotide and deduced amino acid sequence of the cbaA and cbaB genes. The nucleotide sequence of the upper strand of Tn5271 between nucleotide positions 4600 and 6938 (nucleotide 1 defined as the first G in the sequence of the left IS1071 element of Tn5271; Nakatsu et al., 1991) is shown. Potential ribosomal-binding sites are underlined and initiation codons are in bold letters. Stop codons are indicated by an asterisk. The first 87 nucleotides of an ORF downstream of cbaB that has been partly characterized are also shown.

is eventually determined, the alignment with CbaA will likely be very strong.

Fig. 4(a) shows an alignment of amino acids 50–108 of CbaA with similar regions of the terminal oxygenase components representing several other oxygenase classes. Within this region there is 58% identity between CbaA and Phth, and 35% identity between CbaA and V-anA. Alignments in this region are similar to previously determined conserved regions containing the proposed cysteine and histidine ligands of Rieske-type [2Fe–2S] centers of several dioxygenases (Neidle et al., 1991; Mason & Cammack, 1992). In addition to the Rieske-type [2Fe–2S]R cofactor, class IB terminal oxygenase components have been shown to bind mononuclear, non-haem Fe(II) (Yamaguchi & Fujisawa, 1982). This cofactor is thought to be involved in oxygen binding (Mason & Cammack, 1992). Histidine and tyrosine residues conserved in the class IB, IIB and III oxygenases have been proposed to coordinate this Fe(II) (Neidle et al., 1991). In the class IA alignment of Fig. 3, an aspartate and two histidine residues are conserved in a D-x-x-H-x-x-x-H motif in about the same position relative to the Rieske domain (separated by 90–100 amino acids) as in the other oxygenases. No tyrosine residues are conserved in this region.

The relative evolutionary distances of the terminal oxygenase components from different classes were determined by aligning the NH2-terminal 250 amino acids of
Fig. 3. Alignment of CbaA, Ph3 (Nomura et al., 1992) and Vana (Brunel & Davison, 1988) amino acid sequences. The three complete sequences are shown, along with the first 35 amino acids of component A of the 4-sulphobenzoate 3,4-dioxogenase (SbaA; Locher et al., 1991). Amino acids conserved in two or more sequences are shown in upper case, and the identity consensus sequence is shown on the lower line (Cons). * -- Gap introduced into the alignment. Amino acid numbers refer to the CbaA sequence only. Cysteine and histidine residues conserved in the Rieske [2Fe-2S] centre are underlined at CbaA amino acids 69 (see also Fig. 4). Aspartate and histidine residues conserved at the proposed mono-nuclear, non-haem Fe(II)-binding domain are underlined beginning at CbaA amino acid 177. Arginine residues potentially involved in the active site binding of the carboxyl group of the different substrates are underlined at Cba amino acids 329, 377 and 378.

Fig. 4. (a) Alignment of the conserved Rieske [2Fe-2S] centres of the oxygenase components of cis-diol-forming oxygenases. The class designation for oxygenases is based on the system of Batie et al., 1991). The two cysteine and two histidine residues are known to be the ligands of the [2Fe-2S] centre. *, Amino acids above this position are conserved in all of the aligned sequences; --, gap introduced into the alignment. Spaces flanking conserved residues are used for emphasis only. CbaA, chlorobenzoate dioxygenase (Nomura et al., 1992); VanA, vanillate demethylase (Nomura et al., 1992); BphA, biphenyl dioxygenase (Eckardt et al., 1991); BnzA, benzene dioxygenase (Irie et al., 1987). (b) Relative evolutionary distances of terminal oxygenase components determined using the NH2-terminal 250 amino acids of all sequences and drawn as a dendrogram based on pairwise similarity scores.

![Fig. 3. Alignment of CbaA, Ph3 (Nomura et al., 1992) and Vana (Brunel & Davison, 1988) amino acid sequences.](image1)

![Fig. 4. (a) Alignment of the conserved Rieske [2Fe-2S] centres of the oxygenase components of cis-diol-forming oxygenases. The class designation for oxygenases is based on the system of Batie et al., 1991). The two cysteine and two histidine residues are known to be the ligands of the [2Fe-2S] centre. *, Amino acids above this position are conserved in all of the aligned sequences; --, gap introduced into the alignment. Spaces flanking conserved residues are used for emphasis only. CbaA, chlorobenzoate dioxygenase (Nomura et al., 1992); VanA, vanillate demethylase (Brunel & Davison, 1988); BphA, biphenyl dioxygenase (Eckardt et al., 1991); BnzA, benzene dioxygenase (Irie et al., 1987). (b) Relative evolutionary distances of terminal oxygenase components determined using the NH2-terminal 250 amino acids of all sequences and drawn as a dendrogram based on pairwise similarity scores.](image2)
residues characteristic of the chloroplast-type ferredoxin [2Fe–2S]_{\text{Fd}}-binding domain (Otaka & Ooi, 1989). Fig. 5(c) shows an alignment of the class IA and IB sequences in this region with the xylene monoxygenase reductase and plant ferredoxin sequences.

In order to evaluate relative evolutionary distances of the reductase components, the protein sequences were split into two separate parts. This was necessary because the electron transfer components of different reductases have been found to be organized in modular structures that are fused in different ways to form the primary structure (Correll et al., 1992). The first part examined included the FMN/FAD-isoalloxazine- and NAD-ribose-binding domains together in sequences averaging 200 amino acids in length. The second part contained the [2Fe–2S]_{\text{Fd}}-binding domains of the reductases and two plant ferredoxins and averaged 110 amino acids in length. Fig. 6(a) presents a scheme of alignments between two class IA reductases (VanB and CbaB) and a class IB reductase (BenC) to illustrate the alternative arrangements of the conserved regions containing FMN/FAD-isoalloxazine-, NAD-ribose-, and ferredoxin-like ([2Fe–2S]_{\text{Fd}})-binding domains (Neidle et al., 1991; Correll et al., 1992). Fig. 6(b) illustrates the relative evolutionary distances of these and other class IA and IB dioxygenase reductases as dendrograms based on pairwise similarity scores for the FMN/FAD-isoalloxazine- and NAD-ribose domain (left), and ferredoxin-like ([2Fe–2S]_{\text{Fd}}) domain (right). Xylene monoxygenase reductase (XylA) was included in these alignments as an outlier as it has previously been shown to contain the three conserved domains shown in Fig. 5, in the same orientation as the class IB reductases (Neidle et al., 1991).

Class IIB reductase components from the biphenyl and toluene catabolic pathways (BphG and TodA, respectively) and their associated ferredoxin sequences (BphF and TodB) (Furukawa et al., 1989; Zylstra & Gibson, 1989; Erickson & Mondello, 1992) showed very little similarity to the reductase components of the class IA dioxygenases (data not shown). No similarity was found in the alignments in the region expected for the FMN/FAD-isoalloxazine consensus. Some similarity was detected between the potential NAD-ribose-binding domain shown in Fig. 5(b), and the second dinucleotide-binding $\beta\alpha\beta$-fold described for the TodA sequence (Neidle et al., 1991; Mason & Cammack, 1992). The ferredoxin domains were not similar. This was expected because the class IIB ferredoxins have spectroscopic properties and consensus sequences characteristic of Rieske-type [2Fe–2S]$_{\text{R}}$ proteins as opposed to chloroplast-type ferredoxin [2Fe–2S]$_{\text{Fd}}$ proteins (Morrice et al., 1988; Mason & Cammack, 1992).

Another ORF, tentatively designated $cbaC$, has been identified from DNA sequence data extending beyond the $cbaB$ gene (data not shown). Mutation of this region of the cloned chlorobenzoate dioxygenase genes of pBRCN5 resulted in metabolism of 3-chlorobenzoate to dihydrodiol intermediates (results presented above). Expression studies using the thermally regulated T7 polymerase-promoter of pGEM3ZF indicated that the region down-

---

**Fig. 5.** Conserved regions in the reductase component of dioxygenases. * Asterisks indicate amino acids above this position which are conserved in all of the aligned sequences above; --- gap introduced into the alignment. Spaces flanking conserved residues are used for emphasis only. CbaB, chlorobenzoate dioxygenase reductase (this study); PhB2, phtahalate 4,5-dioxygenase reductase (nomura et al., 1992); PDR, phtahalate dioxygenase reductase (correll et al., 1992) (note that the sequences reported for the PDR protein were determined by peptide sequencing or were inferred from electron density maps). VanB, vanillate demethylease reductase (Brunel & Davison, 1988); BenC, chlorobenzoate dioxygenase reductase (Neidle et al., 1991); XylZ, toluate dioxygenase reductase (Harayama et al., 1991); XylA, xylene monoxygenase reductase (Suzuki et al., 1991); Fer1, ferredoxin, Fischerella sp. (Hase et al., 1978; GenBank accession number A00252); Fer2, ferredoxin, Anaebadusa variabilis (chan et al., 1983).
stream of chaB encodes a 42 kDa protein (Nakatsu & Wyndham, 1993). These preliminary results suggest this region of Tn5271 encodes a 3-chlorobenzoate 3,4-dihydriodiol dehydrogenase. Sequencing, database comparisons, and expression studies are continuing on this ORF in an effort to identify the product and determine its similarity to known dehydrogenases.

**DISCUSSION**

Earlier expression studies (Nakatsu & Wyndham, 1993), using constructs with various fragments of Tn5271 placed under the control of either the hybrid tac promoter of the broad-host-range plasmid pMMB66HE or the T7 promoter of pGEM3Zf, showed that two protein products (51 and 42 kDa) were formed from clones capable of restoring 3-chlorobenzoate degradation. The pattern of metabolites formed by clones that did not express the dioxygenase gene and the ORF downstream of chaB remain to be completed. Expression of the entire operon in natural hosts of Tn5271 probably depends on the activity of an upstream promoter contained within an A + T-rich region of 142 bp (60% A + T), located between the Not site (Fig. 1) and the putative ribosome-binding site of chaA (complete sequence not shown). There are no obvious similarities between this sequence and the known promoter sequences of other multi-component oxygenases, including the Pm or Pu promoters of the TOL plasmid and the putative promoter regions of the pht12345 and vanAB operons. Characterization of the chaAB promoter is in progress.

The biochemical classification scheme for bacterial oxygenases described by Batie et al. (1991) is continually being reinforced as new aromatic ring oxygenases are characterized (Bünz & Cook, 1993; Romanov & Hausinger, 1994). As the amino acid sequences of these oxygenases become known, evidence is accumulating that the biochemical classification has a strong evolutionary basis (Neidle et al., 1991; Harayama et al., 1992). In this study the class IA oxygenases have been united in a single lineage the basis of similarities between their component amino acid sequences. Relative evolutionary
distances determined from pairwise amino acid sequence comparisons correspond well with the biochemical classification system of Batie et al. (1991). The terminal oxygenase components of class IA, in the region containing the [2Fe–2S]$_{18}$-binding domains, are more closely related to each other than to the terminal oxygenases of the other classes (Figs 3 and 4). The reductase components of the class IA oxygenases (Fig. 5) show similar cofactor- and iron-binding domains to the class IB reductases and xylene monooxygenase reductase. However, the orientation of the chloroplast-type [2Fe–2S]$_{18}$-binding domain and the FMN/FAD-isoxazoloxidase- and NAD-ribose-binding domains are reversed in all of the known class IA reductases, including the 3-chlorobenzoate dioxygenase reductase described here, compared to other dioxygenase reductases (Fig. 6). This reversal has been attributed to alternative fusions of the three distinct domains of these reductases (Correll et al., 1992). Determination of the crystal structure of phthalate dioxygenase reductase has shown that alternative fusions of the cofactor- and iron-binding domains of the reductases can be achieved by short peptide linkers that cause minimal disturbance to the structure of the reductase (Correll et al., 1992). This would suggest that alternative fusions may arise frequently during the evolution of these electron transport systems. However, the sequence comparisons and relative evolutionary distance diagrams presented in Figs 4 and 6 indicate that the class IA oxygenases form a unified lineage that has conserved the unique fusion arrangement of the reductase domains despite considerable divergence in sequence and function. This suggests that the observed fusion of the cofactor- and iron-binding domains of the class IA dioxygenase reductases probably occurred early in the evolution of the electron transport systems of bacterial oxygenases.

There is good evidence that operons coding for pollutant biodegradation are assembled in a stepwise manner from existing catabolic genes (van der Meer et al., 1992). It is possible that reductase genes for one class of oxygenase may be recruited into another class. Recently a reductase gene (pcpD) has been described that encodes an amino acid sequence that is >55% similar to the VanB and Pht2 class IA reductases (Lange et al., 1994). This gene is located 13 nucleotides downstream from the pentachlorophenol 4-monoxygenase gene pcpB of Flavobacterium sp. strain ATCC 39723, suggesting it is co-transcribed. Unlike the class IA oxygenases, however, pentachlorophenol 4-monoxygenase (PcpB) is a NADPH-dependent flavoprotein monoxygenase (Orser et al., 1993). This observation suggests that the class IA reductases may have been recruited for a variety of aromatic ring oxidation reactions.

The unique features of the class IA dioxygenases may be used in the future to recruit new oxygenases to this class. For example, the 4-sulphobenzoate 3,4-dioxygenase of C. testosteroni T2 belongs to this class on the basis of biochemical studies (Locher et al., 1991). Based on the N-terminal 35 amino acids sequenced from the terminal oxygenase subunit, compared to the corresponding sequences of the CbaA and Pht3 proteins, this dioxygenase belongs to the class IA lineage defined here. It remains for complete sequencing of this dioxygenase and the associated reductase to establish its evolutionary similarity. Isolation and sequencing of dehalogenating class IA dioxygenases other than CbaAB, for example the 4-chlorophenylacetate 3,4-dioxygenase of Pseudomonas sp. CBS3 (Markus et al., 1986; Schweizer et al., 1987) is also expected to strengthen the conclusions presented here.

Both monooxygenase (vanillate demethylase) and dioxygenase enzymes belong to the class IA oxygenases. The terminal oxygenase components of these enzymes are more closely related to one another than to dioxygenases from different classes, suggesting that biochemical variation may arise rapidly within a single lineage of oxygenase genes. Sequence comparisons have not shed light on the important amino acid domain(s) determining substrate binding and monooxygenase versus dioxygenase activity. Biochemical studies have shown that the aromatic hydrocarbon dioxygenases may catalyse...
monooxygenase reactions (Wackett et al., 1988), and the 4-methoxybenzoate monooxygenase of P. putida can dihydroxylate the alkyl side chain of 4-methoxystyrene (Wende et al., 1989). Therefore, the fate of the oxygen atoms at the active site of these enzymes may not be as important an evolutionary character as the orientation of attack on the substituted aromatic ring.

We speculated earlier that there may be an evolutionary relationship between the 4-sulphobenzoate 3,4-dioxygenase and the 3-chlorobenzoate 3,4-dioxygenase because both host organisms belong to the Comamonadaceae, and because the relative stereospecificity of the reactions was similar (Nakatsu & Wyndham, 1993). An examination of Fig. 7, showing the orientations of oxygen insertion into the four different substrates of the class IA oxygenases, supports this conclusion. The over-riding evolutionary constraint acting on the divergence of the class IA oxygenases would appear to be the requirement for a carboxyl group para to the site of oxygen insertion. The amino acid(s) of the dioxygenase component that orient the substrate within the active site and that presumably recognize the carboxyl group are not known, however they should be conserved in all class IA dioxygenase sequences. Conserved amino acids that may serve this function in the class IA dioxygenases are the three arginines at positions 329, 377 and 378 (Fig. 3, CbaA numbering). Arginines also make up three of the nine conserved residues in an alignment of the P-subunits of the class IB, IIB and III oxygenases that are required for substrate binding (Neidle et al., 1991). Further sequencing, mutagenesis and expression studies are required to establish the nature of active sites and the mechanistic and evolutionary relationships between dioxygenase and monooxygenase reactions in these enzymes.

ACKNOWLEDGEMENTS

The authors thank Dr Rama Singh, Bess Wong, Barbara Holland and Marysoon Salih for their assistance. The work reported here was supported by a Research Grant to R.C.W. from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

phanotic two-component enzyme system from Comamonas testo-


properties of component A of the 4-chlorophenylacetate-3,4-
dioxygenase from Pseudomonas species strain CBS. J Biol Chem 261,

12883–12888.


hydroxylation bacterial dioxygenases. Annu Rev Microbiol 46,

277–305.

van der Meer, J. R., Eggen, R. I. L., Zehnder, A. J. B. & de Vos,


gene cluster, which encodes metabolism of chlorinated catechols: evi-

dence for specialization of catechol-1,2-dioxygenases for chlorinated


van der Meer, J. R., de Vos, W. M., Harayama, S. & Zehnder,


Morrice, N., Geary, P., Cammack, R., Harris, A., Beg, F. & Aitken,

A. (1998). Primary structure of protein B from Pseudomonas putida,

member of a new class of 2Fe-2S ferredoxins. FEBS Lett 231,


the transposable chlorobenzoate-3,4-dioxygenase genes of


Chlorobenzoate catabolic transposon Tn5271 is a composite class I

element with flanking class II insertion sequences. Proc Natl Acad

Sci USA 88, 8312–8316.


DNA sequence of the Acinetobacter salaoacetius catechol 1,2-
dioxygenase I structural gene catA: evidence for evolutionary divergence

of intradiol dioxygenases by acquisition of DNA sequence repetitions.

J Bacterial 170, 4874–4880.

Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M. &


salaoacetius benABC genes for benzoate 1,2-dioxygenase reveal evolu-

tional relationships among multicomponent oxygenases. J Bacterial

173, 5385–5395.

Nomura, Y., Nakagawa, M., Ogawa, N., Harashima, S. & Oshima,

Y. (1992). Genes in PHT plasmid encoding the initial degradation

pathway of phthalate in Pseudomonas putida. J Ferment Biog 74,

333–344.

Orser, C. S., Lange, C. C., Xun, L., Zahrt, T. C. & Schneider, B. J.

(1993). Cloning, sequence analysis, and expression of the Flavo-

bacterium pentachlorophenol-4-monoxygenase gene in Escherichia


homologies. V. New perspectives on evolution between bacterial and

chloroplast-type ferredoxins inferred from sequence evidence.


Reineke, W. & Knackmuss, H.-J. (1988). Microbial degradation of


uses a three-component ortho-halobenzoate-1,2-dioxygenase for

metabolism of 2,4-dichloro- and 2-chlorobenzoate. J Bacterial

176, 3368–3374.


Laboratory.


Pseudomonas sp. strain CBS3 genes specifying dehalogenation of


Sayler, G. S., Hooper, S. W., Layton, A. C. & Henry King, J. M.

(1990). Catabolic plasmids of environmental and ecological signifi-


Scholten, J. D., Chang, K. H., Babbett, P. C., Charest, H., Sylvèstre,


Schweizer, D., Markus, A., Seez, M., Ruf, H. H. & Lingens, F.

(1987). Purification and some properties of component B of the

4-chlorophenylacetate-3,4-dioxygenase from Pseudomonas species


Suzuki, M., Hayakawa, T., Shaw, J. P., Rekik, M. & Harayama, S.

(1991). Primary structure of xylene monooxygenase: similarities to

and differences from the alkane hydroxylation system. Eur J Bacterial

173, 1690–1695.


monooxygenation catalyzed by toluene dioxygenase from Pseudo-

monas putida. Biochemistry 27, 1360–1367.


modulated reactions of putidamonooxin: the nature of the active

oxygen species formed, and its reaction mechanism. Eur J Biochem

181, 189–197.


monadaceae, a new family encompassing the Acidovorax rRNA

complex, including Varivorax paradoxus gen. nov., comb. nov., for


and interaction between Alcaligenes and Pseudomonas species from


instability, plasmid gene deletion and recombination in Alcaligenes


oxygenase component in benzoate-1,2-dioxygenase system from


phage cloning vectors and host strains: nucleotide sequences of the


Pseudomonas putida F1: nucleotide sequence of the toaC1C2B:ADE