Biodegradation **5**: 259–276, 1994.

On the origins and functions of the enzymes of the 4-chlorobenzoate to 4-hydroxybenzoate converting pathway

Debra Dunaway-Mariano[1] & Patricia C. Babbitt[2]
[1] *Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, USA*
[2] *Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, USA*

Received 2 November 1993; accepted 13 April 1994

Key words: 4-chlorobenzoate, 4-hydroxybenzoate, dehalogenase, 4-chlorobenzoyl-CoA dehalogenase, 4-chlorobenzoate CoA ligase, 4-hydroxyl benzoyl-CoA thioesterase, halogenated aromatic degradation, nucleophilic aromatic substitution

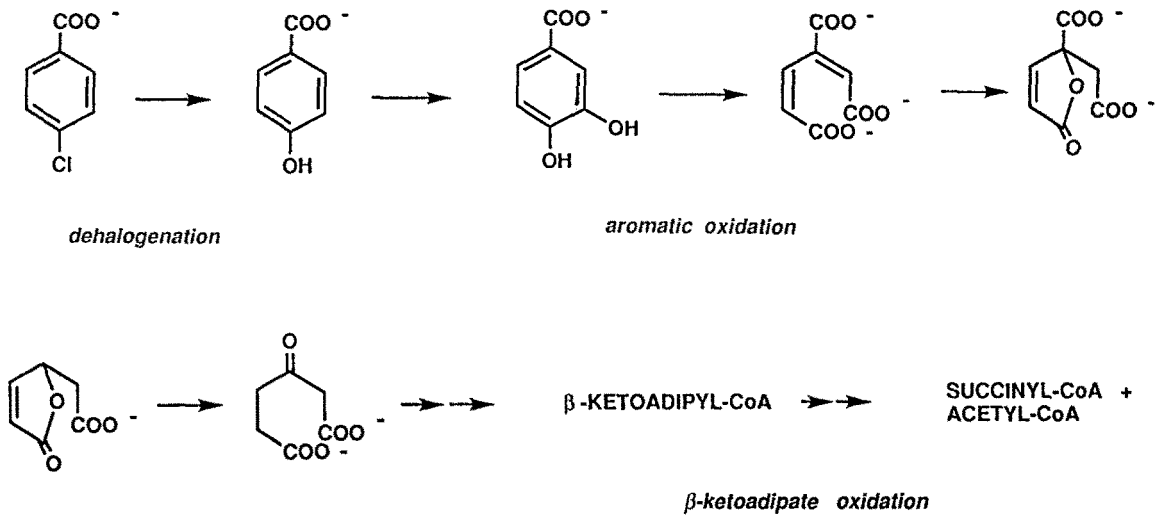
Abstract

This review examines the enzymes of 4-chlorobenzoate to 4-hydoxybenzoate converting pathway found in certain soil bacteria. This pathway consists of three enzymes: 4-chlorobenzoate: Coenzyme A ligase, 4-chlorobenzoyl-Coenzyme A dehalogenase and 4-hyroxybenzoyl-Coenzyme A thioesterase. Recent progress made in the cloning and expression of the pathway genes from assorted bacterial strains is described. Gene order and sequence found among these strains are compared to reveal independent enzyme recruitment strategies. Sequence alignments made between the *Pseudomonas sp.* strain CBS3 4-chlorobenzoate pathway enzymes and structurally related proteins contained within the protein sequence data banks suggest possible origins in preexisting β-oxidation pathways. The purification and characterization of the physical and kinetic properties of the pathway enzymes are described. Where possible a comparison of these properties between like enzymes from different bacterial sources are made.

Introduction

This review examines the enzymes of a 4-chlorobenzoate (4-CBA) degradation pathway found in certain soil bacteria. 4-Chlorobenzoate is not known to be a natural product but is introduced into the environment through its use as a precursor in the synthesis of dye stuffs, pigments and pharmaceuticals (Shimao et al. 1989). It is also produced as a by-product in the microbial breakdown of certain herbicides (Cork & Krueger 1991; Haggblom 1992) and of the ubiquitous pollutants, polychlorinated biphenyls (introduced into the environment through their former use in plastics, carbon paper, cooling systems and transformers) (Higson 1992; Abramowicz 1990; Commandeur & Parsons 1990).

In recent years several strains of 4-CBA degrading bacteria belonging to the genera *Acinetobacter* (Adriaens et al. 1989), *Alcaligenes* (Van den Tweel et al. 1986), *Pseudomonas* (Klages & Lingens 1980), *Nocardia* (Klages & Lingens 1979), *Corynebacterium* (Zaitsev et al. 1991; Groenewegen et al. 1992) and *Arthrobacter* (Marks et al. 1984; Müller et al. 1988; Shimao et al. 1989) have been isolated from 4-CBA enrichment cultures of soil or sludge samples. The potential for utilization of these 4-CBA degrading bacteria for decontamination purposes is suggested by the observation that most of the isolates grow efficiently on 4-CBA as the sole carbon source and one *Arthrobacter* strain, in particular, tolerates very high levels of 4-CBA in culture (Shimao et al. 1989). Furthermore, total degradation of 4-chlorobiphenyl has been achieved by natural (Pettigrew et al. 1990) and constructed (Adriaens et al. 1989; Furukawa & Chakrabarty 1982; Sylvestre et al. 1985) consortia as well as with pure cultures (Shields et al. 1985). *Alcaligenes* sp A2, in particular, has been shown to carry a large plasmid encoding the genes responsible for the oxidation of the biphenyl unit plus genes encoding enzymes which dehalogenate the 4-CBA by-product to the metabolite 4-hydroxybenzoate (4-HBA) (Hooper et al. 1989; Pet-



Scheme 1. Catabolism of 4-Chlorobenzoate (4-CBA).

tigrew et al. 1990; Layton et al. 1992) (see Scheme 1).

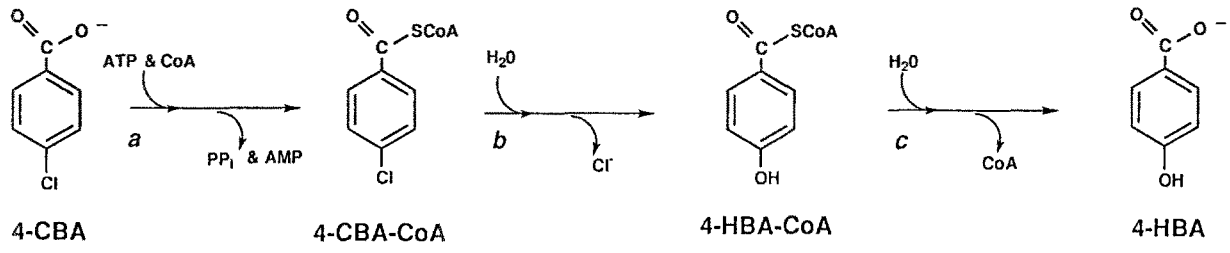
Each strain of bacteria appears to use a common pathway for 4-CBA mineralization wherein the chloroaromatic is first converted to 4-HBA which in turn is oxidized to CO_2 via aromatic metabolizing pathways. The 4-CBA degrading pathway of *Pseudomonas* sp. strain CBS3 appears to be the most thoroughly characterized. We will first focus our attention on this pathway and then examine the 4-CBA pathways found in the other known 4-CBA degrading bacteria.

4-CBA degradation in *Pseudomonas* sp. strain CBS3

Pseudomonas sp. strain CBS3 was originally isolated from soil by requiring growth on 4-CBA as the sole source of carbon (Klages & Lingens 1980). This bacterium was found to contain several dehalogenases, which include two different 2-haloalkanoic acid hydrolases (Klages et al. 1983; Schneider et al. 1991), an enzyme system that oxidizes 4-chlorophenylacetate to 3,4-dihydroxyphenylacetate (Klages et al. 1981; Markus et al. 1984) and an enzyme system that hydrolyzes 4-CBA to 4-HBA (Müller et al. 1984; Thiele et al. 1987; Scholten et al. 1991) (the subject of this review).

The chemical strategy that we have seen unfold for the biodegradation of 4-CBA in *Pseudomonas* sp. strain CBS3 involves the coupling of three catabolic pathways (Scheme 1). The first pathway converts the 4-CBA to 4-HBA in a three step process (Scholten et al. 1991) (Scheme 2). The 4-HBA thus formed is oxidized to protocatechuate and then to carboxymuconate via the ortho-cleavage pathway. The β-ketoadipate ultimately derived is further oxidized to succinyl-CoA and acetyl-CoA via the β-ketoadipate pathway. It is the 4-CBA $\rightarrow$ 4-HBA dehalogenation pathway which is unique to this strain of *Pseudomonas* and which allows 4-CBA to be mineralized in conjunction with a conventional aromatic metabolizing pathway.

Earlier attempts by other investigators to purify the 4-CBA dehalogenase activity from *Pseudomonas* sp. strain CBS3 (Thiele et al. 1987) had been unsuccessful because of the low levels of activity that were observed with cell-free extract. The cloning of the cosmid vector pPSA 843 carrying a 9.5 kb chromosomal DNA fragment encoding the dehalogenase genes from *Pseudomonas sp* strain CBS3 into *E. coli* was reported in 1986 (Savard et al. 1986). Transfer of the hybrid cosmid to the 4-CBA dehalogenase-minus strain *Pseudomonas putida* KT2440 conferred on this strain the ability to dehalogenate 4-CBA and to grow on it as the sole source of carbon. Cloning the dehalogenase genes did not, however, immediately solve the purification problem. The maximal activity observed for the cellu-



Scheme 2. Reaction Steps of the dehalogenation of 4-CBA in *Pseudomonas* sp. strain CBS3 Catalyzed by (a) 4-CBA: CoA ligase, (b) 4-CBA-CoA dehalogenase, (c) 4-HBA-CoA thioesterase.

lar extract from either clone was quite low (at least 20 times lower than measured with whole cells) and it was short lived (Savard et al. 1992). The low and transient dehalogenase activity observed with lysed cells was suspected to be the result of degradation or dilution of an unidentified cosubstrate or the disruption of a supporting membrane structure. Thus, structural homology between the dehalogenase and known proteins was examined to gain insight into the catalytic properties of the former. This was ultimately accomplished through oligonucleotide sequencing and selective expression of the dehalogenase encoding *Pseudomonas* sp. strain CBS3 DNA fragment (Scholten et al. 1991).

Analysis of dehalogenase activity in Ω insertion mutants and deletion mutants generated in the laboratory localized the dehalogenase encoding region to a 4.8 kb stretch of the original 9.5 kb cloned DNA fragment (Savard et al. 1992). Selective expression of the dehalogenase genes from the cloned 4.8 kb DNA truncation fragment in a maxicell system revealed a 30 kDa polypeptide as one of the components of the dehalogenase system (Savard et al. 1992). Selective expression of the dehalogenase genes using the more sensitive T7 polymerase promoter based expression system revealed not only the 30 kDa polypeptide but also 57 – and 15 kDa polypeptide products (Scholten et al. 1991). Deletion mutants were then used in conjunction with the selective expression systems to locate the approximate positions of the three encoding genes on the DNA fragment and the direction of transcription (Savard et al. 1992; Scholten et al. 1991).

Oligonucleotide sequencing of the cloned 4.8 kb DNA fragment identified 3 open reading frames (ORFs) corresponding to 29 847, 57 155 and 16 107 Da polypeptides arranged continuously 5' to 3' on the DNA fragment (Babbitt et al. 1992). The noncoding intervals between the 30 kDa and the 57 kDa polypeptide encoding ORFs and between the 57 kDa and 16 kDa polypeptide encoding ORFs comprise 8 and 105 bases, respectively. The ORFs were verified by comparing the predicted size and N-terminal sequences of the encoded polypeptides with the molecular weights and N-terminal amino acid sequences of the three dehalogenase polypeptide components later purified from the *E. coli* clone (Babbitt et al. 1992; Chang et al. 1992). Possible links between the dehalogenase polypeptide components and proteins whose amino acid sequences are contained in major databases were probed by carrying out primary sequence homology analyses.

Accordingly, the 30 kDa polypeptide encoded by ORF1 was first found to be related to 2-enoyl-CoA hydratases functioning in fatty acid β-oxidation (Babbitt et al. 1992). The 57 kDa polypeptide encoded by ORF2 was found to be related to a large family of ligases catalyzing acyl adenylation/thioesterification. There was no match found for the 16 kDa polypeptide encoded by ORF3.

Based upon the links made between the 57 and 30 kDa dehalogenase polypeptide components and the ligases (several of which were CoA ligases) and the 2-enoyl-CoA hydratases, respectively, the reaction sequence shown in Scheme 2 (Viz. 4-CBA: CoA ligase, 4-CBA-CoA dehalogenase and 4-HBA-CoA thioesterase) was constructed. The three enzymes represented in Scheme 2 were fractionated and identified (Scholten et al. 1991). Later, subclones were generated for overexpression of the individual genes in *E. coli* (Chang et al. 1992). The isolation of the 4-CBA-CoA ligase from the original *Pseudomonas* sp. strain CBS3 has also been reported (Löffler et al. 1992).

4-CBA to 4-HBA converting pathways in other 4-CBA degrading bacteria

Special strains SU and TM1 of *Arthrobacter* contain enzymes that catalyze hydrolytic dehalogenation of 4-CBA to 4-HBA (Marks et al. 1984; Müller et al. 1988; Ruisinger et al. 1976). The reaction (observed in cell free extracts) requires ATP and coenzyme A, indicating a dehalogenation pathway analogous to that found in *Pseudomonas* sp. strain CBS3 and shown in Scheme 2. The dehalogenase genes of *Arthrobacter* sp. strain SU were cloned and expressed in *E. coli* (Schmitz et al. 1992). In *Arthrobacter*, as was observed in *Pseudomonas* sp. strain CBS3, three ORFs, encoding the dehalogenase activity, are juxtaposed (Schmitz et al. 1992; Savard et al. 1992; Babbitt et al. 1992). ORF1 and ORF2 were suggested (based on similarities of polypeptide size and sequence to the *Pseudomonas* 4-CBA → HBA pathway enzymes) to encode the 4-CBA: CoA ligase and 4-CBA-CoA dehalogenase, respectively (Schmitz et al. 1992). The *Arthrobacter* sp. strain SU ORF3 encodes a 4-HBA-CoA thioesterase which is of similar subunit size (17 kDa) as that of the *Pseudomonas* sp. strain CBS3 4-HBA-CoA thioesterase however, the two thioesterases do not share significant sequence identity (Schmitz et al. 1992).

4-CBA to 4-HBA conversion in cellular extracts of *Acinetobacter* sp. strain 4-CB1 also occurs by initial synthesis of 4-CBA-CoA from 4-CBA, ATP and CoA (Copley & Crooks 1992). With this system the 4-CBA-CoA was shown to undergo dehalogenation to 4-HBA-CoA which in turn is hydrolyzed to 4-HBA. The amino acid sequences of the enzymes mediating these three reactions in *Acinetobacter* sp. strain 4-CB1 have not yet been reported so no comparison to the corresponding enzymes from *Arthrobacter* sp. strain SU or *Pseudomonas* sp. strain CBS3 can be made at this time. Nevertheless, the *Acinetobacter* 4-CBA-CoA dehalogenase has been purified and its physical and kinetic properties studied (Crooks & Copley 1993). A comparison of the properties of the 4-CBA-CoA dehalogenase isolated from the *Acinetobacter* sp. strain 4-CB1 and *Pseudomonas* sp. strain CBS3 is provided below.

Tsoi et al. (1991) have cloned the *Arthrobacter globiformis* genes required for 4-CBA dehalogenation in *E. coli* JM109. Although these workers were able to detect (using a minicell expression system) two protein products sized at 32 and 58 kDa from the cloned genes, they did not ascribe function. By analogy to the *Pseudomonas* sp. strain CBS3 4-CBA degrading pathway, these two proteins may function as the 4-CBA-CoA dehalogenase and 4-CBA: CoA ligase, respectively.

The 4-CBA to 4-HBA converting pathway which employs a 4-CBA: CoA ligase, 4-CBA-CoA dehalogenase and 4-HBA-CoA thioesterase thus appears to occur in both Gram positive and Gram negative 4-CBA degrading bacteria. We suspect that the chemical steps for 4-CBA → 4-HBA conversion reported to take place in the 4-CBA degrading strains *Arthrobacter* sp. strain SB8 (Shimao et al. 1989), *Coryneform* strain NTB-1 (Groenewegen et al. 1992) and *Alcaligenes* sp. strain ALP83 (Layton et al. 1992) constitute the same three-step dehalogenation pathway represented in Scheme 2.

Properties of the purified enzymes of the 4-CBA to 4-HBA converting pathway

The 4-CBA: CoA ligase

The 4-CBA: CoA ligase of *Pseudomonas* sp. strain CBS3 has been purified to homogeneity by a 6-step purification procedure (Löffler et al. 1992). The same ligase was also purified from an *E. coli* clone (Scholten et al. 1991; Chang et al. 1992). Löffler et al. (1992) and Chang et al. (1992) reported similar properties for the *Pseudomonas* and *E. coli* enzyme isolate. The enzyme displays an isoelectric point of 5.3 (Löffler et al. 1992) and has an observed molecular mass of 110–115 kDa and consists of two identical polypeptide subunits of 57 kDa (Löffler et al. 1992; Chang et al. 1992). The kinetic data taken from Löffler et al. (1992) and Chang et al. (1992) and summarized in Table 1 indicate that over the pH range of 7–9 the catalytic turnover rate of the ligase is 30–100 s^{-1}.

The 4-CBA: CoA ligase was found to require a divalent metal cofactor for catalysis. Chang et al. (1992) reported that this requirement is satisfied with Mg^{2+}, Mn^{2+} or Co^{2+} (K_m and k_{cat} values obtained using the three different metal ions are very similar) but not with Ca^{2+} or Zn^{2+}. Löffler et al. (1992), on the other hand, reported that Mg^{2+} (100%), Mn^{2+} (87%), Co^{2+} (77%), Fe^{2+} (70%), Zn^{2+} (23%) and Ni^{2+} (11%) were all effective as cofactors.

Examination of potential alternate substrates for the 4-CBA: CoA ligase showed that it is not active with aliphatic carboxylic acids including palmitate, caproate, laurate and butyrate (Chang et al. 1992). This ligase is thus distinct from the fatty acid: CoA ligases. Likewise, the substrates for 4-coumarate: CoA

Table 1. Kinetic constants of the *Pseudomonas* sp. strain CBS3 4-CBA:CoA ligase.

Substrate	K_m (μM)		k_{cat} (S^{-1})	
	pH 7.0, 25° C	pH 8.4, 30° C	pH 7.0, 25° C	pH 8.4, 30° C
4-CBA	8.50 ± 0.09^a	50×65^b	29 ± 1^a (50, 30° C)^b	107^b
CoA	70 ± 6^a	110–190^b (32 ± 40, pH 8.5, 25° C)^a	35 ± 1^a	107^b (32 ± 2, pH 8.5, 25° C)^a
MgATP	104 ± 5^a	2400–3500^b	242 ± 0.4^a	107^b

[a] Data from Chang et al. 1992 for the enzyme isolated from the *E. coli* clone.
[b] Data from Löffler et al. 1992 for the enzyme isolated from *Pseudomonas* sp. strain CBS3.

Table 2. Kinetic constants of the *Pseudomonas* sp. strain CBS3 and *Acinetobacter* sp. 4-CB1 4-CBA-CoA dehalogenase.

Substrate or inhibitor	k_{cat} (s^{-1})	K_m	K_i (μM)[c]
4-CBA-CoA	$^a0.60 \pm 0.01$ $^b1.30$	$^a3.7 \pm 0.3$	
4-bromobenzoyl-CoA	$^a1.35 \pm 0.03$ $^b2.31$	$^a4.2 \pm 0.3$	
4-iodobenzoyl-CoA	$^a1.07 \pm 0.01$	$^a6.5 \pm 0.3$	
4-fluorobenzoyl-CoA	$^a8 \times 10^{-6}$ $^b3 \times 10^{-3}$	$^b78$	$^a40 \pm 5$
2,4-dichlorobenzoyl-CoA	$^a0.511 \pm 0.005$	$^a10.4 \pm 0.4$	
3,4-dichlorobenzoyl-CoA	$^a0.052 \pm 0.001$	$^a42 \pm 3$	
4-chloro-2-nitrobenzoyl-CoA	$^a0.063 \pm 0.002$	$^a30 \pm 2$	
4-chloro-3-nitrobenzoyl-CoA	$^a0.0256 \pm 0.002$	$^a5.5 \pm 0.2$	
4-chloro-3-methylbenzoyl-CoA	$^a5 \times 10^{-4}$	$^a75 \pm 5$	$^a76 \pm 7$
benzoyl-CoA	–		$^a72 \pm 8$
CoA	–		$^a140 \pm 10$
4-CBA	–		$^a21000 \pm 2000$

[a] Data for the *Pseudomonas* sp. strain CBS3 enzyme measured at 25° C in 50 mM K^+ Hepes and 1 mM DTT (pH 7.5) (Liang et al. 1993).
[b] Data for the *Acinetobacter* sp. 4-CB1 enzyme measured at 30° C in 20 mM potassium phosphate buffer at pH 7.2 (Crooks & Copley 1993).
[c] K_i values were determined for these compounds as competitive inhibitors vs. 4-CBA-CoA.

ligase, phenylacetate: CoA ligase, 4-hydroxybenzoate: CoA ligase, 2,3-dihydroxybenzoate: CoA ligase and 2-aminobenzoate: CoA ligase are not substrates for the 4-CBA: CoA ligase (Chang et al. 1992). The following descending order of reactivity: 4-CBA (k_{cat}/K_m = 340 × 10^4 $M^{-1}s^{-1}$) > 4-bromobenzoate (k_{cat}/K_m = 170 × 10^4 $M^{-1}s^{-1}$) > 4-iodobenzoate (k_{cat}/K_m = 82 × 10^4 $M^{-1}s^{-1}$) > 4-methylbenzoate (k_{cat}/K_m = 29 × 10^4 $M^{-1}s^{-1}$) > benzoate (k_{cat}/K_m = 1 × 10^4 $M^{-1}s^{-1}$) indicates that the 4-chloro substituent plays an important role in substrate recognition perhaps by binding to and desolvating a small hydrophobic pocket on the enzyme (Chang et al. 1992). Löffler et al. (1992) have reported that the 4-CBA adduct substituted at the C(2) with a nitro substituent or at the C(3) position with a chloro or nitro substituent is not a substrate nor are the benzoates substituted at C(2) and/or C(3) with chloro, fluoro, iodo, amino or hydroxyl substituents. Overall, the ligase displays a high level of substrate specificity.

The 4-CBA-CoA dehalogenase

The 4-CBA-CoA dehalogenases from *Acinetobacter* sp. strain 4-CB1 and from the *E. coli* clone expressing the *Pseudomonas* sp. strain CBS3 gene have been purified and characterized (Chang et al. 1992; Crooks & Copley, unpublished data). Both dehalogenases appear to be 120 kDa homotetramers of 30 kDa subunits. Neither enzyme seems to require a divalent metal ion for activity. The K_m value for 4-CBA-CoA is 34 μM in the case of the *Acinetobacter* enzyme and 4 μM in the case of the *Pseudomonas* enzyme. The turnover rates for the two enzymes, however, are the same, 1 s^{-1}

```
110.Ec          MKQQGTTLPANNHTLKQYAFFAGMLSSLKKQKWRKGMSESLHLTRNGSILEITLDRPKA-   59
enoyl.Rn.mit    MAALRALLPRACNSLLSPVRCPEFRRFASGANFQYIITE--KKGKNSSVGLIQLNRPKAL   58
orf257.Rc       ---------------------------MSYHTIRYEISE--------GLAVITLDRPEVM   25
4CBA.deha.Asp   ----------------------MSSNSDHHISVEHTD--------GVATIRFTRPSKH   28
4CBA.deha.Psp   --------------------------MYEAIGHRVED--------GVAEITIKLPRHR   24
dhna.Ec         ---------------------------------MKIFDE-----TYNGIAKITINRPEVH   22
enoyl.Pf        --------------------MIYEGKAITVTALESGIVE---LKFDLKGESVNKFNRLTL   37
enoyl.Rn.per    -----------------------------------MAE---YLRLPHSLAMIRLCNPPV   21

                                                    .  .
110.Ec          NAIDAKTSFEMGEVFLNFRDDPQLRVAIITGAGEKFFSAGWDLKA-A-----AEGEAPDA  113
enoyl.Rn.mit    NALCNGLIEELNQALETFEEDPAV-GAIVLTGGEKAFAAGADIK--E-----MQNRTFQD  110
orf257.Rc       NALNAAMRHELTAALHRARGEA---RAIVLTGSGRAFCSGQDLGDGA-----AEGLNLET   77
4CBA.deha.Asp   NAASGQLLLETLEALYRLESDD-SVGAIVLTGEGAVFSAGFDLEEVPMGPA-SEIQSHFR   86
4CBA.deha.Psp   NALSVKAMQEVTDALNRAEEDD-SVGAVMITGAEDAFCAGFYLREIPLDKGVAGVRDHFR   83
dhna.Ec         NAFTPKTVAEMIDAFADARDDQNVGVIVLAGAGDKAFCSGGDQKVRG----HGGYVGDDQ   78
enoyl.Pf        NELRQAVDAIKADASVKGVIVSSGK---------DVFIVGADIT--EFVENFKLPDAELI   86
enoyl.Rn.per    NAVSPTVIREVRNGLQKAGSDHTVKAIVICGAN-GNFCAGA-----------DIHGFSAF   69

                                  .  ..  .  .  ..           .          .
110.Ec          DFGPGGFAGLTEIFNLDKPVIAAVNGYAFGGAFELALAADFIVCADNASFALPEAKLGIV  173
enoyl.Rn.mit    CYSGKFLSHWDHITRIKKPVIAAVNGYALGGGCELAMMCDIIYAGEKAQFGQPEILLGTI  170
orf257.Rc       VLREEYEPLLQAIYSCPLPVLAAVNGAAAGAGANLALAADVVIAAQSAAFMQAFTRIGLM  137
4CBA.deha.Asp   LKALYYHAVIHMLARIEKPTLAAINGPAVGGGLGMSLACDLAVCTDRATFLPAWMSIGIA  146
4CBA.deha.Psp   IGALWWHQMIHKIIRVKRPVLAAINGVAAGGGLGISLASDMAICADSAKFVCAWHTIGIG  143
dhna.Ec         IPRLNVLDLQRLIRVIPKPVVAMVSGYAIGGGHVLHIVCDLTIAADNAIFGQTGPKVGSF  138
enoyl.Pf        AGNLEANKIFSDFEDLNVPTVAAINGIALGGGLEMCLAADFRVMADSAKIGLPEVKLGIY  146
enoyl.Rn.per    TPGLALGSLVDEIQRYQKPVLAAIQGVALGGGLELALGCHYRIANAKARVGLPEVTLGIL  129

                  .   .                    .
110.Ec          PDSGGVLRLPKILPPAIVNEMVMTGRRMGAEEALRWGIVNRVVSQAELMDNARELAQQLV  233
enoyl.Rn.mit    PGAGGTQRLTRAVGKSLAMEMVLTGDRISAQDAKQAGLVSKIFPVETLVEEAIQCAEKIA  230
orf257.Rc       PDAGGTWWLPRQVGMARAMGMALFAEKIGAEEAARMGLIWEAVPDVDFEHHWRARAAHLA  197
4CBA.deha.Asp   NDASSSFYLPRIVGYRRAMEWLLTNRTLGADEAYEWGVVNRVFSEADFQSRVGEIARQLA  206
4CBA.deha.Psp   NDTATSYSLARIVGMRRAMELMLTNRTLYPEEAKDWGLVSRVYPKDEFREVAWKVARELA  203
dhna.Ec         DAGYGSGYLARIVGHKKAREIWYLCRQYNAQEALDMGLVNTVVPLEQLEEETIKWCEEML  198
enoyl.Pf        PGFGGTVRLPRLIGVDNAVEWIASGKENRAEDALKVSAVDAVVTADKLG-AAALDLIKRA  205
enoyl.Rn.per    PGARGTQLLPRVVGVPVALDLITSGKYLSADEALRLGILDAVVKSDPV--EEAIKFAQKI  187

110.Ec          NSAPLAIAALKEIYRTTSEMPVEEAYRYIRSGVLKHYPSVLHSEDAIEGPLAFAEK--RD  291
enoyl.Rn.mit    NNSKIIVAMAKESVNAAFEMTLTEGNKLEK----KLFYSTFATDDRREGMSAFVEK--RK  284
orf257.Rc       RGPSAAFAAVKKAFHAGLSNPLPAQLALEA----RLQGELGQSADFREGVQAFLEK--RP  251
4CBA.deha.Asp   AAPTHLQGLVKNRIQEGSSETLESCTEHEV----QNVIASVGHPHFAERLAMFRSKEMRS  262
4CBA.deha.Psp   AAPTHLNVMAKERFHAGWMNPVEECTEFEI----QNVIASVTHPHFMPCLTRFLDGH-RA  258
dhna.Ec         EKSPTALRFLKAAFNADTDGLAGIQQFAG-----DATLLYYTTDEAKEGRDSFKEK--RK  251
enoyl.Pf        ISGELDYKAKRQPKLEKLKLNAIEQMMAFE--TAKGFVAGQAGPNYPAPVEAIKTI--QK  261
enoyl.Rn.per    I--DKPIEPRRIFNKPVPSLPNMDSVFAEA--IAKVRKQYPG---VLAPETCVRSI--QA  238

110.Ec          PVWKGR----------------------------------  297
enoyl.Rn.mit    ANFKDH----------------------------------  290
orf257.Rc       PHFTGR----------------------------------  257
4CBA.deha.Asp   SALAVDLDAVCGGR----------------------------  276
4CBA.deha.Psp   DRPQVELPAGV----------------------------  269
dhna.Ec         PDFGQFPRFP----------------------------  261
enoyl.Pf        AANFGRDKALEVEAAGFAKLAKTSASNCLIGLFLNDQEL  300
enoyl.Rn.per    SVKHPYEVGIKEEEKLFMYLRASG---------------  262
```

Fig. 1.

←

Fig. 1. Alignment of 4-Chlorobenzoate-CoA dehalogenase sequences from *Arthrobacter* species strain SU (4CBA.deha.Asp) *Pseudomonas* species CBS-3 (4CBA.deha.Psp) with homologous proteins. Each sequence shown in the alignment is identified as described below. Residues that are conserved in all 8 sequences shown in the alignment are designated by dots (•). Three stretches of sequence that exhibit particularly high sequence similarity among all of the aligned proteins are boxed. Sequences are listed by their abbreviation in the Figure, followed by the designations used to identify each protein in the Figures and text. Figure designation, protein names, species, Genbank Locus #, and the original reference for each sequence are also included in that order where known. The "GB Locus #" designations can be used to obtain the sequences using database search tools available from NCBI. All database searches were performed by email using the NCBI blast and fasta network services, email address: blast@ncbi.nih.gov. 110.Ec, ORF 56, *Escherichia coli*, ECO110K, Yura, T. et al. 1992; enoyl.Rn.mit, enoyl-CoA hydratase, *Rattus norvvegicus* (mitochondria), RNECH, Minami-Ishii, N. Et al. 1989; orf257.Rc, enoyl-CoA hydratase homolog, *Rhodobacter capsulatus*, RC 257, Beckman, D.L. and Kranz, R.G. 1991; enoyl.Pf, fatty acid β-oxidation multienzyme complex, α-chain, *Pseudomas fragi*, S38338, Sato, S. et al. 1992; dhna.Ec, DHNA synthase, *Escherichia coli*, ECODHNASYN, Sharma, V. et al. 1992; enoyl.Rn.per, enoyl-CoA hyratase/3-hydroxyl-CoA dehydrogenase, *Rattus norvegicus* (peroxisomes), RATPECOA, Osumi, T. et al. 1985; 4CBA.deha.Asp, 4-chlorobenzoate-CoA dehalogenase, *Arthrobacter sp.* st. SU, ARGFCBABC, Schmitz, A. et al. 1992; 4CBA.deha.Psp, 4-chlorobenzozte-CoA dehalogenase, *Pseudomonas sp.* CBS-3, Babbitt, P.C. et al. 1992.

(Table 2). The leaving group mobility has been studied for both enzymes. While both enzymes convert the 4-bromobenzoyl-CoA twice as fast as the 4-CBA-CoA, the *Pseudomonas* enzyme has much more difficulty dehalogenating the 4-fluorobenzoyl-CoA substrate ($k_{cat} = 8 \times 10^{-6} s^{-1}$) (Liang et al. 1993) than does the *Acinetobacter* enzyme ($k_{cat} = 3 \times 10^{-3} s^{-1}$) (Crooks & Copley 1993), suggesting a difference in mechanism and/or reaction energetics. Structure/activity measurements made with 4-CBA-CoA analogs bearing electron donating or withdrawing substituents at the benzoate ring C(2) or C(3) (Table 2) have suggested the importance of steric/solvation effects on the *Pseudomonas* dehalogenase reaction but have failed to provide insight into the mechanism of the dehalogenation reaction (Liang et al. 1993). The inhibition constants measured for benzoyl-CoA (72 μM), CoA (140 μM) and 4-CBA (21 mM) compared to the K_m value measured for 4-CBA-CoA (4 μM) suggests the dominant role played by the CoA moiety in substrate anchoring to the *Pseudomonas* dehalogenase (Liang et al. 1993).

Known substrates for 2-enoyl-CoA hydratase were tested as potential substrates (hydration) for the *Pseudomonas* 4-CBA-CoA dehalogenase and conversely, 4-CBA-CoA was tested as a potential substrate (dehalogenation) for a 2-enoyl-CoA hydratase. 4-CBA-CoA dehalogenase was not active with crotonyl-CoA, α-methylcrotonyl-CoA and β-methylcrotonyl-CoA (Chang et al. 1992). Likewise, the 2-enoyl-CoA hydratase, crotonase was not an active dehalogenase towards 4-CBA-CoA (Liang & Dunaway-Mariano, unpublished data). Thus, the 4-CBA-CoA dehalogenase is distinct from its closest structural and catalytic analogue (as discussed below), the 2-enoyl-CoA hydratase.

The 4-HBA-CoA thioesterase

The *Pseudomonas* sp. strain CBS3 4-HBA-CoA thioesterase expressed in *E. coli* was isolated as an apparent 66 kDa homotetramer of 16 kDa subunits (Chang et al. 1992). The 4-HBA-CoA thioesterase was shown to be inactive with the aliphatic acyl-CoA thioesters including palmitoyl-CoA, hexanoyl-CoA and acetyl-CoA, and thus, to be catalytically distinct from the thioesterases involved in fatty acid metabolism. The relative reactivity of benzoyl-CoA derivatives toward thioesterase catalysis was found to be 4-HBA-CoA ($k_{cat}/K_m = 1.5 \times 10^6\ M^{-1}s^{-1}$)>> benzoyl-CoA ($k_{cat}/K_m = 13 \times 10^2\ M^{-1}s^{-1}$) $\approx$ 4-CBA-CoA ($k_{cat}/K_m = 6.4 \times 10^2\ M^{-1}s^{-1}$). These results suggest that the thioesterase is highly discriminating and recognizes 4-HBA-CoA, in part, through interaction with the 4-OH substituent. Furthermore, the 4-HBA moiety of the 4-HBA-CoA appears to play a dominant role (K_i values for 4-HBA and 4-HBA-N-acetyl-L-cysteine are 34 ± 4 and 52 ± 7, respectively; Taylor & Dunaway-Mariano, unpublished data) in substrate anchoring.

Thus, as observed with the ligase and dehalogenase of the 4-CBA dehalogenation pathway, the thioesterase displays a high level of substrate specificity. The kinetic properties of the three pathway enzymes give no indication that these enzymes serve alternate physiological functions in the host cell. Therefore, we have pursued the idea that the genes encoding these three enzymes were recruited from a preexisting metabolic pathway(s) and retooled to recognize and transform the new metabolites.

Origin of the 4-CBA degrading pathway

Given the existence of several examples of bacteria which mineralize 4-CBA by conversion to the 4-HBA using the same three step pathway (represented in Scheme 2) we wonder if the 4-CBA → 4-HBA pathway enzymes found in each of the different 4-CBA degrading bacterial strains share the same origin or are examples of different recruiting strategies. We are continuing to investigate and contemplate the origins of these pathways and specifically, to determine if the 4-CBA → 4-HBA pathway enzymes are 'retooled' enzymes and, if this is so, to determine from which pathway or pathways these enzymes were recruited. Our first investigation of the lineage of the 4-CBA → 4-HBA pathway enzymes involved the identification of proteins with which they share significant sequence identity thus allowing connections to other proteins to be made based upon structural homology (Babbitt et al. 1992). Our recent findings are described below.

Alignment of the 4-CBA-CoA dehalogenase with homologous proteins

The sequences of the *Pseudomonas* sp. strain CBS3 and *Arthrobacter* sp. strain SU 4-CBA-CoA dehalogenases are aligned with the sequences of a subset of structurally related proteins as shown in Fig. 1. The two 4-CBA-CoA dehalogenase sequences are 48% identical and 65% similar (calculated using the same definitions for similarity as those used by the alignment algorithm (Smith 1990)). The similarity extends throughout these two sequences but concentrates in several long stretches of very high identity. The percent identities (see Table 3) found to exist between pairs of all of the sequences shown in the alignment is relatively low, ranging between 16% and 32% for all pairs except for the dehalogenase pair. In the overall alignment, the observed sequence similarities are much less global than the similarities existing between the two dehalogenase sequences. Evidence that a primary structural relationship exists for the entire set of proteins is observed in three main clusters of conserved residues (boxed in Fig. 1). Another group of proteins that exhibit sequence similarities with the dehalogenase, the Δ^3-cis-Δ^2-trans-enoyl-CoA isomerases are not included in the alignment shown in Fig. 1 because their overall sequence similarity with the dehalogenases is low in comparison to those of the other proteins shown in the alignment. An earlier sequence alignment which included a representative of the isomerase sequences is given in Babbitt et al. (1992).

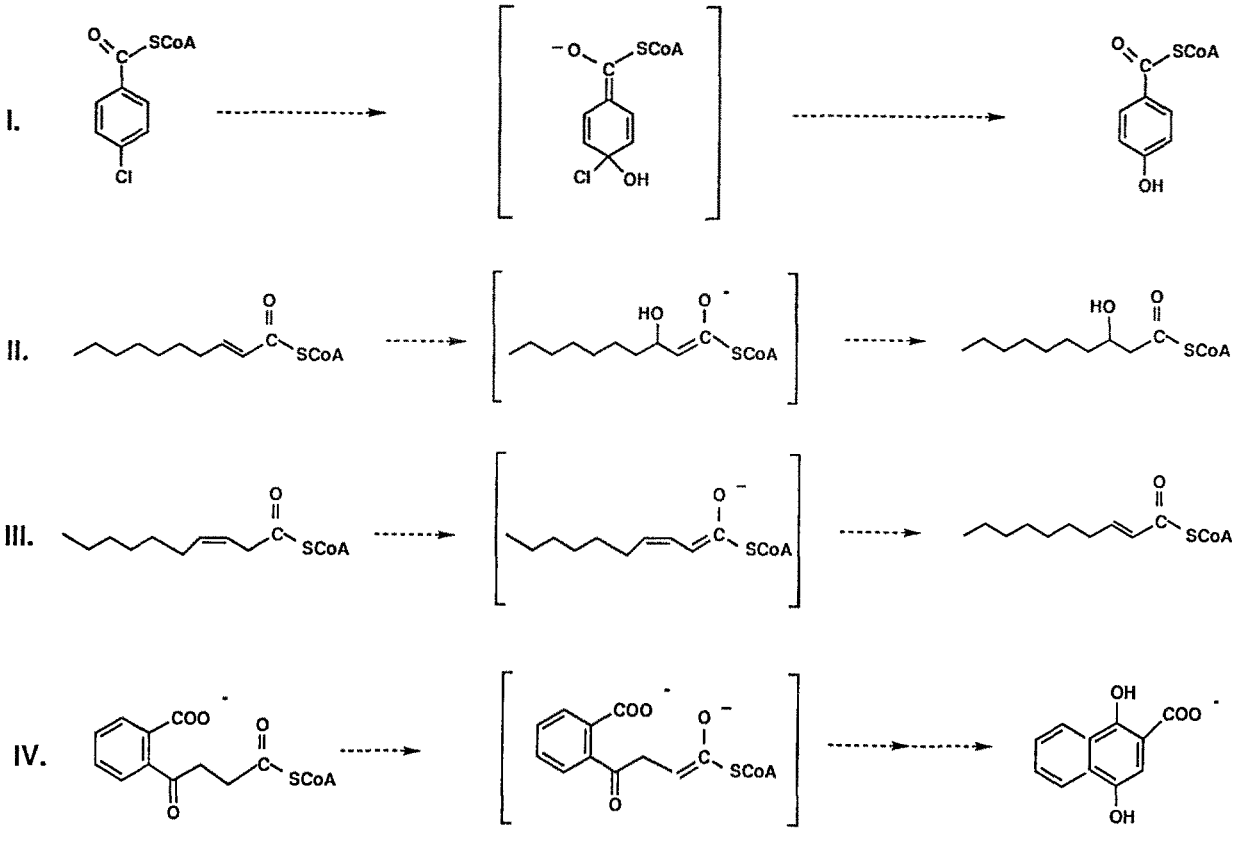
A possible link between the mechanisms of action of the dehalogenase, 2-enoyl-CoA hydratase, Δ^3-cis, Δ^2-trans-enoyl-CoA isomerase and dihydroxynaphthoate synthase is apparent from the nature of the reactions that they catalyze (see Scheme 3). For each reaction, one would expect excess electron density to accumulate at the carbon adjacent to the CoA-thioester during the course of catalytic turnover. For the extreme case of a stepwise reaction proceeding in the absence of acid catalysis, a thioester enolate intermediate would be formed. Independent of whether the reaction proceeds by a concerted or stepwise pathway, with or without acid catalysis, we might expect that a specialized active site surface, functioning to bind the thioester moiety of the substrate and to polarize (through interaction with positively charged amino acid side chains) the C = O bond, to be conserved among this group of enzymes.

It is of interest to note that we were unable to find significant similarities between the protein sequences shown in Fig. 1 and the sequences of proteins which catalyze related reactions in β-keto-CoA thioesters. Attempted alignments of the HMG-CoA ligase, acetyl-CoA acetyl transferase, 3-ketoacyl-CoA thiolase, and citrate synthetase sequences failed. Proteins catalyzing reactions of carboxylate substrates proceeding through aci-acid intermediates were also found to be structurally unrelated. This latter group includes mandelate racemase, fumerase, aconitase and enolase. Nor were aspartate ammonia lyase, argininosuccinate lyase or 3-hydroxyl'-3-methyl-glutaryl lyase found to be structurally related. Finally, neither of the 4-CBA-CoA dehalogenases appear to share significant sequence identity with known 2-haloalkanoic dehalogenase sequences including the two 2-alkanoic dehalogenases identified in the *Pseudomonas* sp. strain CBS3 (Schneider et al. 1991).

In summary, the 4-CBA-CoA dehalogenases from *Pseudomonas* sp. strain CBS3 and *Arthrobacter* sp. strain SU appear to share common ancestry with the hydratase and isomerase of the fatty acid β-oxidation pathway and with the dihydroxynapthoate synthase of the menaquinone (vitamin K_2) biosynthetic pathway. It is noteworthy that the 4-CBA → 4-HBA dehalogenation, fatty acid β-oxidation and menaquinone pathways also contain structurally related CoA ligases (see below).

Table 3. The percent identities calculated for each pair of sequences aligned in Fig. 1. The abbreviations used for the sequences are defined in Fig. 1.

Sequence	enoyl.Rn.mit	orf257.Rc	4CBA.deha.Asp	4CBA.deha.Psp.	dhna.Ec	enoyl.Pf	enoyl.Rn.per
110.Ec	31	32	25	23	28	20	21
enoyl.Rn.mit		32	24	25	29	18	22
orf257.Rc			30	32	28	18	19
4CBA.deha.Asp				48	20	16	19
4CBA.Psp.					25	17	19
dhna.Ec						16	20
enoyl.Pf							26



Scheme 3. Parallel reaction pathways catalyzed by 4-CBA-CoA dehalogenase (I), 2-enoyl-CoA hydratase (II), Δ^3-cis-Δ^2-trans-enoyl-CoA isomerase (III) and dihydroxynapthoate synthase (IV).

Alignment of the 4-CBA: CoA ligase with homologous proteins

An alignment of the two 4-CBA: CoA ligases from *Pseudomonas* sp. strain CBS3 and *Arthrobacter* sp. strain SU is shown in Fig. 2a. Overall, the sequence similarities are global with only a few long stretches of identity or dissimilarity. The two sequences are 38% identical and 58% similar.

Based on our most recent (August 1993) database searches, 49 protein sequences that exhibit sequence similarities with the two 4-CBA: CoA ligases have been identified. It is noteworthy that the recently determined sequence (Altenschmidt & Fuchs 1992) of the


enn.Fs
gramI.Bb
angR.Va
ala.Lc
HC.tox
ad.red.Sc
4CBA.lig.A
4CBA.lig.P
phosph.Sv
acet.Nc
luci.PHp
rfbL.Vc
aas.Ec
acyl.Po
110.Ec
EntE.Ec
baiB.Esp

Fig. 2A.

←

Fig. 2A. Alignment of the 4-Chlorobenzoate:CoA ligase sequences from *Arthrobacter* species strain SU (4CBA.deha.Asp) and *Pseudomonas* species CBS3 (4CBA.deha.Psp) with 15 other homologous sequences. Each sequence shown in the alignment is identified as described below. Five stretches of sequence that exhibit significant sequence similarity in all 17 of the proteins shown are boxed in bold lines. Part of the boxed region shown in the uppermost left of the Figure spans the sequence motif highlighted in Figure 2b and discussed further in the text. The other regions shown by boxes around only a few of the 17 sequences in the alignment represent regions of sequence similarity that help to distinguish the two sequence sub-families discussed in the text. The alignment shown is that determined by the algorithm PIMA (Smith, 1990) except that some additional gaps have been added (designated by '=') to align motifs of high similarity among all of the sequences but that were left offset by the algorithm. In all cases, these motifs can be aligned by PIMA as shown in the Figure when alternate sets of proteins are used as the dataset. While the alignment shown exhibits several short regions of high sequence similarity among all of the sequences, the overall low degree of sequence similarity leads to an enormous number of possible alignments. For this reason, no further changes in the alignment (beyond the introduction of the gaps mentioned above) were made over that determined as optimal by the PIMA algorithm. Thus, careful examination of the alignment shows that addition or deletion of some gaps would result in better local alignments for some pairs of sequences. These changes were not made because we could not determine a single 'best' alignment given the low degree of overall sequence similarity in the dataset.

2 aminobenzoate: CoA ligase from denitrifying *Pseudomonas* sp. is not related to this group.

The 51 sequences represent 37 unique protein functions. A subset of the 51 sequences was used to generate a sequence alignment (shown in Fig. 2a) with the two 4-CBA: CoA ligases. This alignment shows a selected region of these sequences (spanning 250 amino acid residues in length and roughly corresponding to the middle section of the 4-CBA: CoA ligase sequence) which displays the highest level of sequence identities. Included in the alignment shown in Fig. 2a are only those sequences which share less than 30% sequence identity with any other protein in the alignment (excluding the two 4-CBA: CoA ligase sequences). The gramicidin synthetase I sequence (designated 'gram I.Bb' in Fig. 2a) for example, was used in the alignment to represent a family of protein sequences which share comparatively high sequence identity among themselves. This family of proteins includes gramicidin synthetase II, tyrocidine synthetase I, the entire family of α-aminoadipyl-cysteinyl-valine (ACV) synthetases (representing at least 4 species and 12 homologous domains), surfactin synthetase and the *Ent*F gene product.

The most pronounced region of sequence similarity among the proteins aligned in Fig. 2a spans amino acids # 161–170 in the *Pseudomonas* sp. strain CBS3 4-CBA: CoA ligase sequence (4CBA.lig.Psp). This motif (whose function is described below) can be recognized in all 51 sequences. An alignment of this conserved region, which includes all of the variations (17 total) exhibited by these 51 protein sequences, is shown in Fig. 2b. The high degree of sequence conservation observed for this motif is exceptional, particularly in the light of the broad range of substrates, functions and chemistries these proteins represent.

From the sequence alignment presented in Fig. 2a it is evident that these proteins share sequence similarity beyond that exhibited in the highly conserved motif of Fig. 2b. Regions of high similarity in the seventeen proteins shown in the alignment are boxed. Qualitative estimates of sequence similarity can be determined from the percent identities existing between all pairs of sequences included in the alignment. This analysis suggests that the six sequences shown at the top of the alignment (enn.Fs, gramII.Bb., angR.Va, ala.Lc, HC.tox.Cc, ad.red.Sc), which are 25–35% identical to each other over the truncated sequence lengths shown, might be grouped together in one sub-family. In addition, the seven sequences shown at the bottom of the alignment (luci,PHp, rfbL.Vc, aas.Ec, acyl.Po, 110.Ec, EntE.Ec, baiB.Esp) which exhibit a similar level of sequence identity among themselves (25–38% identical) might be grouped together as a second sub-family. In contrast, the 4-CBA: CoA ligase sequences (4CBA.lig.A, 4CBA.lig.P), and the sequences designated 'phosph.Sv' and 'acet.Nc' exhibit only 9–20% sequence identity with the other sequences for the regions shown in the alignment.

Unfortunately, the low degree of global sequence similarity among these sequences makes it impossible to construct a single 'correct' alignment that would allow the generation of a phylogeny from these data. The somewhat speculative grouping into sub-families is supported, however, by the observation that short regions of some of these sequences exhibit locally high degrees of similarity. An understanding of the importance of these sub-grouping awaits availability of more sequences that can be used to generate a single 'correct' overall alignment. Finally, we note that the 4-CBA: CoA ligase from the *Arthrobacter* sp. strain SU appears to share more sequence similarity with the proteins in the alignment in some of these highly con-

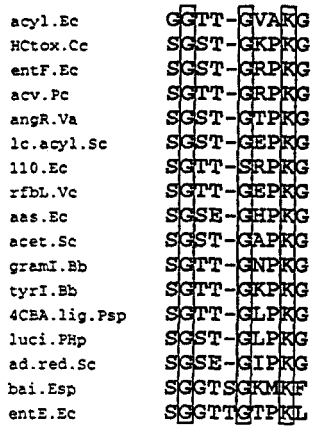
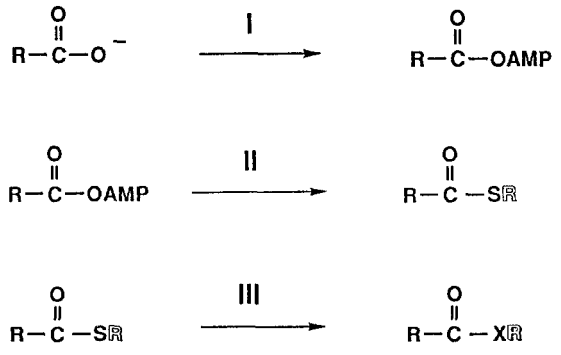


Fig. 2B. Alignment of the highly similar stretch of sequence spanning residues 161–170 in the 4CBA:CoA ligase from *Pseudomonas* sp. strain CBS with all other unique versions of this motif represented by the 51 sequences we found to share significant sequence similarity. The three residues conserved in all of these versions are boxed. Most of these sequence motifs represents more than one protein, although only one protein name is listed for each motif in Figure 2a. Sequences are listed below for both Figures 2a and 2b in alphabetical order of their abbreviations, followed by the designations used to identify each protein in the Figures and text as described in the legend for Figure 1. 110.Ec=ORF 56, *Escherichia coli*, ECO110K, Yura, T. et al. 1992; 4CBA.lig.Asp, 4-chlorobenzoate: CoA ligase, *Arthrobacter sp.*st.SU, ARGFCBABC, Schmitz, A. et al. 1992; 4CBA.lig.Psp, 4-chlorobenzoate:CoA ligase, *Pseudomonas sp.* CBS-3, Babbitt, P.C. et al. 1992; aas,Ec,2-acylglycerophosphoethanolamine acyltransferase/acyl-acyl carrier protein synthetase, *Escherichia coli*, ECOAAS, Jackowski, S. et al. 1993; acet.Bs, acetyl-CoA synthetase, *Bacillus subtilis*, BACACUCBA, Grundy, F.J. Et al. 1993; acyl.Po, acyl CoA synthetase, *Pseudomonas oleovorans*, POOCT, van Beilen, J. B. et al. 1192; ad.red.Sc, alpha-aminoadipate reductase, *Saccharomyces cerevisiae*, YSCLS2A, Morris, M.E. and Jinks-Robertson, S. 1991; ala.Lc, D-alanine-activating enzyme, *Lactobacillus casei*, LBADAAE, Heaton, M.P. and Neuhaus, F.C. 1992; angR. Va, transacting transcriptional activator, *Vibrio anguillarum*, VIBANGRA, Farrell, D.H. et al. 1190; bai.Esp, baiB protein, *Eubacterium sp.*st. VPI 12708, EUBBAIOA, Mallonee, D.H. et al. 1990; enn.Fs, enniatin synthetase, *Fusarium scirpi*, FSESYN1A, Haese, A. et al. 1993; entE.Ec, enterobactin synthetase component E, *Escherichia coli*, ECENTB, Staab, J.F. et al. 1989; gramI.Bb, gramicidin S synthetase 1, *Bacillus brevis* (Nagano), BACGS1, Hori, K. et al. 1989; HC.tox.Cc, HC-toxin synthetase, *Cochliobolus carbonum*, CCLHTS1X, Scott-Craig, J.S. et al. 1992; luci.PHp, photinus-luciferin 4-monooxygenase, *Photinus pyralis*, PPYLUC, de Wit, J.R. et al. 1987; phosph.Sv, phosphinothricin synthetase A, *Streptomyces viridochromogenes*, SVPTT, Wohllben, W. et al. 1992; rfbL.Vc, rfbL protein, *Vibrio cholerae*, VCRFBAT, Manning, P.A. 1991.

served regions than does the 4-CBA: CoA ligase from the *Pseudomonas* sp. strain CBS3.

The question which emerges from the recognition of structural relatedness between this large group of proteins is: what common functional feature(s), if any,



Scheme 4. A summary of the types of reactions catalyzed by the proteins represented in Fig. 2. (I) Adenylation, (II) Thioesterification and (III) Formation of an ester or amide.

do these proteins share? Recently, based upon a smaller group of some of the proteins aligned in Fig. 2a we (Scholten et al. 1991; Babbitt et al. 1992) and others (Toh 1990; Turgay et al. 1992; van Beilen et al. 1992; Masuda et al. 1989; Staab et al. 1989; Rusnak et al. 1991; Hori et al. 1991) have suggested the possibility that they may share a common function. Most of these proteins are enzymes which catalyze the adenylation of carboxylate substrates with ATP (the possible exceptions, lacking defined catalytic function, are angR.Va, 110.Ec, phosph.Sv, 110.Ec, rfbL.Vc, baiB.Esp and orf1.Pa). A reoccurring theme among these adenylate forming enzymes is the intermediate or final acyl transfer to a thiol, that being enzyme linked phosphopantetheine or bound CoA, respectively (see Scheme 4; Babbitt et al. 1992). Thus, the chemistry conserved among these proteins is likely to be related to acyl adenylate formation and possibly to thioester formation.

In recent studies, we (Chang 1994) have carried out site directed mutagenesis of conserved residues of the sequence motif highlighted in Fig. 2b and of the conserved Glu residue at amino acid position 306 of the *Pseudomonas* sp. strain CBS3 4-CBA: CoA ligase (see Fig. 2a) to test the possible role(s) of these residues in acyl adenylate formation and in thioester formation. Kinetic analysis of the wild type enzyme and the G163I, G166I, P168A, K169M and E306Q mutants of the *Pseudomonas* sp. strain CBS3 4-CBA: CoA ligase revealed that these amino acid residues function in acyl adenylate formation. The mutations increased the 4-CBA and ATP K_m values from 14 μM and 308 μM for wild type to 105 μM and 1314 μM for G163I, 73 μM and 981 for G166I, 20 μM and 830 μM for K169M and

253 μM and 321 μM for E306 (the P168A mutant was inactive) while not effecting the K_m for CoA so much. Single turnover experiments showed that the rate of the adenylation half reaction had been effected (k = 135 s^{-1} for wild type *vs* 0.04–25 s^{-1} for the mutants) while the rate of the thioesterification step was unaltered (100 s^{-1} for wild type and mutants). Thus, Glu # 306 and the motif of Fig. 2b appear to be primarily involved in an acyl adenylate formation. It is of interest to note that this motif is absent from the sequences of the family of the acyl adenylate forming enzymes, the aminoacyl tRNA synthetases. This distinction coupled with the global sequence identity found among the proteins of Fig. 2a suggests that the origin of these (largely) thioester ester forming enzymes may have involved divergent evolution from a common ancestor distinct from the ancestor of the aminoacyl tRNA synthetases.

Alignment of the 4-HBA-CoA thioesterase with homologous proteins

Recent database searches have failed to identify sequence data that would alter our earlier report that no known protein sequences are homologous to the 4-hydroxybenzoate-CoA (4-HBA-CoA) thioesterase from *Pseudomonas* sp. CBS3 (Babbitt et al. 1992) including the putative 4-HBA-CoA thioesterase component of the dehalogenase system of *Arthrobacter* sp. strain SU (Schmitz et al. 1992). The lack of homology between the two 4-HBA-CoA thioesterases is particularly remarkable in light of the high degree of similarity observed between the 4-CBA-CoA dehalogenase and 4-CBA: CoA ligase components of the two 4-CBA → 4-HBA pathways.

While there is no obvious sequence similarity that exists between the *Pseudomonas* thioesterase and the putative thioesterase from *Arthrobacter*, the latter protein does share significant sequence similarity with at least two other proteins (Schmitz et al. 1992). These are the hypothetical protein (ent.orf4.Ec) of MW 15 kD (Liu et al. 1989) encoded by a gene residing downstream of the *Ent*E, *Ent*B, and *Ent*A of the *E. coli* enterobactin biosynthetic pathway operon and the protein encoded by ORF2 of the comA gene (comp2.Bs) (Weinrauch et al. 1989), associated with the development of genetic competence in *Bacillus subtilis*. An alignment of these proteins is shown in Fig. 3 (attempts to align the *Pseudomonas* sp. strain CBS3 thioesterase sequence with this multiple alignment failed, indicating that there is no significant relationship between it and this group of proteins). The *Arthrobacter* sp. strain SU thioesterase sequence (4CBA.thio.Asp in Fig. 3) exhibits 32% sequence identity with the protein encoded by ORF2 of the *com*A gene (comp 2.Bs) and 26% identity with protein encoded by the ORF associated with the *Ent*A/B genes (ent.orf4.Ec). The comp2.Bs sequence is 40% identical to the ent.orf4.Ec sequence.

The alignment shown in Fig. 3 reveals a definite relationship between the structures of the three proteins as evidenced by the high sequence identity located in a stretch of 13 residues in the central region of the sequences (boxed region in Fig. 3). A number of individual residues are also conserved in all three sequences. Both the relatively high degree of sequence similarity among these proteins and the fact that all are of unusually small size suggests that they derive from a common ancestor. Unfortunately, the function of neither the comp2.Bs or ent.orf4.Ec protein has been clearly defined (although it has been established that they are transcribed during induction of their respective pathways). Thus, a more definitive understanding of their relationship to the putative thioesterase from *Arthrobacter* awaits characterization of their catalytic functions.

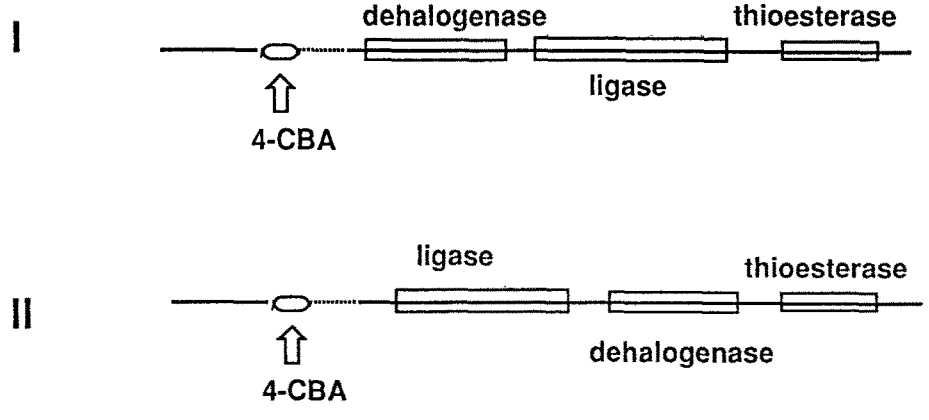
Recruitment of the 4-CBA → 4-HBA pathway enzymes in Pseudomonas sp. strain CBS3 and Arthrobacter sp. strain SU

The 4-CBA → 4-HBA converting pathways observed in the *Pseudomonas* sp. strain CBS3 and *Arthrobacter* sp. strain SU each consists of a 4-CBA: CoA ligase, a 4-CBA-CoA dehalogenase and a 4-HBA-CoA thioesterase. The chemical strategy used by these two organisms to replace the halogen substituent of 4-CBA with a hydroxyl group (to allow catabolism of the aromatic ring) is, therefore, the same. Here we return to our earlier question of whether these two sets of pathway enzymes were recruited from the same preexisting metabolic pathway(s).

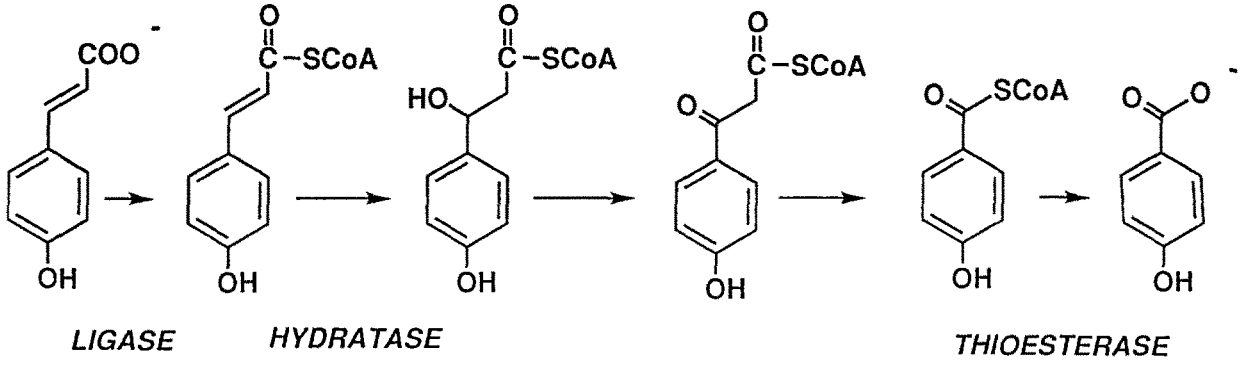
First, it might be argued on the basis of the high level of sequence identity shared between the 4-CBA: CoA ligases and between the 4-CBA-CoA dehalogenases of the *Pseudomonas* and *Arthrobacter* 4-CBA → 4-HBA pathways that these two pathways have a common origin. On the other hand, the lack of sequence homology that one might otherwise expect to exist between the *Pseudomonas* and *Arthrobacter* 4-HBA-CoA thioesterases would suggest that at least the thioesterase components of the two pathways have

```
comp2.Bs          ----------------------MDMKHTLLEALGIEIVENTAERCVAVMPVDHRTVQPF   37
ent.orf4.Ec       ----------MIWKRHLTLDELNATSDNTMVAHLGIVYTRLGDDVLEAEMPVDTRTHQPF   50
4CBA.thio.Asp     MHRTSNGSHATGGNLPDVASHYPVAYEQTLDGTVGFVIDEMTPERATASVEVTDTLRQRW   60

comp2.Bs          GYLHGGASVALAETAARPGAQNLIDHTTQACVGLEINANHLKSVKEGTVKAIAEPVHIGR   97
ent.orf4.Ec       GLLHGGASAALAETLG-SMAGFMMTRDGQCVVGTELNATHHRPVSEGKVRGVCQPLHLGR  109
4CBA.thio.Asp     GLVHGGAYCALAEMLATEATVAVVHEKGMMAVGQSNHTSFFRPVKEGHVRAEAVRIHAGS  120

comp2.Bs          TTIVYHIHIYDEQERLICISRCTLAVIKK--  126
ent.orf4.Ec       QNQSWEIVVFDEQGRRCCTCRLGTAVLG---  137
4CBA.thio.Asp     TTWFWDVSLRDDAGRLCAVSSMSIAVRPRRD  151
```

Fig. 3. Alignment of the 4-chlorobenzoate thioesterase sequence (4CBA.thio.Asp) from *Arthrobacter* sp. strain SU (Schmitz et al. 1992) with a hypothetical protein associated with genes involved in enterobactin synthesis, ent.orf4.Ec (Liu et al. 1989) and the comA gene product, comp2.Bs (Weinrauch et al. 1989). Positions at which residues are identical in all three sequences are designated by dots (•). The boxed region exhibits particularly high sequence similarity among all three proteins.



Scheme 5. Representation of the order observed for the 4-CBA → 4-HBA pathway genes found in *Pseudomonas* sp. strain CBS3 (I) and in *Arthrobacter* sp. strain SU (II). Both sets of genes are induced by 4-CBA.



Scheme 6. The chemical steps of 4-coumarate β-oxidation in plants.

separate origins. Furthermore, the difference in gene ordering on the 4-CBA → 4-HBA operons found in the *Pseudomonas* sp. strain CBS3 (Savard et al. 1992) and the *Arthrobacter* sp. strain SU (Schmitz et al. 1992) (see Scheme 5) suggest different modes of gene recruitment for construction of the two 4-CBA → 4-HBA pathways.

Elucidation of the gene sources involved in the development of the 4-CBA → 4-HBA pathways of the *Pseudomonas* and *Arthrobacter* systems may be aided by the recent finding of several examples of structural homology between two of the three 4-CBA → 4-HBA pathway enzymes and two other enzymes catalyzing parallel reactions in a known metabolic pathway. First, as pointed out earlier, the fatty acid β-oxidation pathway contains: (i) fatty acid: CoA ligase which shares homology with the 4-CBA-CoA ligases and (ii) 2-enoyl-CoA hydratase which shares homology with the 4-CBA-CoA hydratases (Babbitt et al. 1992). Second, the enterobactin biosynthetic pathway contains enterobactin synthetase component E (i.e., 2,3-dihydroxybenzoate: CoA ligase) which shares homology with the 4-CBA: CoA ligases (Staab et al. 1989) and a putative thioesterase (Liu et al. 1989) which shares homology with the *Arthrobacter* 4-HBA-CoA thioesterase (Schmitz et al. 1992). Third, we find structural homology between the 4-CBA: CoA ligases and 4-CBA-CoA dehalogenases and the menaquinone pathway enzymes O-succinyl benzoate: CoA ligase and the dihydroxynaphthoic acid synthetase (Sharma et al. 1992; Driscoll & Taber 1992), respectively. Finally, we note that the 4-CBA: CoA ligases and 4-CBA-CoA dehalogenases are structurally related to two proteins of an unidentified pathway encoded by the *E. coli* K12 genome (0–2.4 min; Yura et al. 1992). For these latter two operons we find examples of both patterns of gene ordering i.e., in the menaquinone operon the two genes of interest are ordered in the same way as the dehalogenase and ligase encoding genes of the *Pseudomonas* 4-CBA → 4-HBA operon, while in the *E. coli* K12 genome the gene ordering is found to match that of the *Arthrobacter* 4-CBA → 4-HBA operon (Scheme 5).

Presently, we can not ascribe the origin of the 4-CBA → 4-HBA pathway enzymes to one particular metabolic pathway. Our current findings do, however, provide ample evidence that multiple proteins have been coconscribed and retooled to catalyze the reactions required for the dehalogenation of 4-CBA. Furthermore, we have seen clear structural and/or functional links between the enzymes of the 4-CBA → 4-HBA pathway and several biosynthetic and biodegradative pathways. Perhaps the closest chemical analogy to the 4-CBA → 4-HBA pathway is the 4-coumarate → 4-hydroxybenzoate oxidation pathway found in plants and shown in Scheme 6. Unfortunately, the proteins (with the exception of the 4-coumarate: CoA ligase) of this pathway and their encoding genes have not been characterized for a detailed comparison to be made at this time.

Acknowledgements

The editorial assistance provided by Pauli G. Kosclskey is gratefully acknowledged. The authors also wish to thank Shelly D. Copley and Gwen P. Crooks for providing data on the *Acinetobacter* sp. strain 4-CB1 4-CBA-CoA dehalogenase prior to publication. This work was supported by NIH Grant GM 28688 to D.D.M. and by a grant from the Cohn Medical Research Fund of the University of California at San Francisco to P.C.B.

References

Abe T, Fujino T, Fukuyama R, Minoshima S, Shimizu N, Toh N, Hiroyuki S & Yamamoto T (1992) Human long-chain acyl-CoA synthetase: Structure and chromosomal location. J. Biochem. 111: 123–128

Abramowicz DA (1990). Crit. Rev. Biotechnol. 10: 241–

Adriaens P, Kohler H-PE, Kohler-Staub D & Focht DD (1989) Bacterial dehalogenation of chlorobenzoates and coculture biodegradation of 4, 4'-dichlorobiphenyl. Appl. Environ. Microbiol. 55: 887–892

Altenschmidt U & Fuchs G (1992) Novel aerobic 2-aminobenzoate metabolism. Eur. J. Biochem. 205: 721–727

Babbitt PC, Kenyon GL, Martin BM, Charest H, Sylvestre M, Scholten JD, Chang K-H, Liang PH & Dunaway-Mariano D (1992) Ancestry of the 4-chlorobenzoate dehalogenase: analysis of amino acid sequence identities among families of acyl: adenyl ligases, enoyl-CoA hydratases/isomerases and acyl-CoA thioesterases. Biochemistry 31: 5594–5604

Becker-André M, Schulze-Lefert P & Hahlbrock K (1991) Structural comparison, modes of expression, and putative *cis*-acting elements of the two 4-coumarate: CoA ligase genes in potato. J. Biol. Chem. 266: 8551–8559

Beckman DL & Kranz RG (1991) A bacterial homology to the mitochondrial enoyl-CoA hydratase. Gene 107: 171–172

Birch PR, Sims PF & Broda P (1992) Nucleotide sequences of a gene from *Phanerochaete chrysosporium* that shows homology to the xxi gene of *Aspergillus nidulans*. DNA Sequence 2: 319–323

Black PN, DiRusso CC, Metzger AK & Heimert TL (1992) Cloning, sequencing, and expression for the *fadD* gene of *Escherichia coli* encoding acyl CoA synthetase. J. Biol. 267: 25513–25520

Chang K-H (1994) Purification of a novel 4-chlorobenzoate dehalogenase system and characterization of the 4-chlorobenzoate: CoA ligase. (1994) Ph.D. Thesis, University of Maryland

Chang K-H, Liang P-H, Beck W, Scholten JD & Dunaway-Mariano D (1992) Isolation and characterization of the three polypeptide components of 4-chlorobenzoate dehalogenase from *Pseudomonas* sp. strain CBS-3. Biochemistry 31: 5605–5610

Commandeur LCM & Parsons JR (1990) Degradation of halogenated aromatics. Biodegradation 1: 207–

Connerton IF, Fincham JRS, Sandeman RA & Hynes MJ (1990) Comparison and cross-species expression of the acetyl-CoA synthetase genes of the Ascomycete fungi, *Aspergillus nidulans* and *Neurospora crassa*. Mol. Microbiol. 4: 451–460

Copley SD & Crooks GP (1992) Enzymic dehalogenation of 4-chlorobenzoyl Coenzyme A in *Acinetobacter* sp. strain 4-CB1. Appl. Environ. Microbiol. 58: 1385–1387

Cork DJ & Krueger JP (1991) Microbial transformations of herbicides and pesticides. Adv. App. Microbiol. 36: 1–67

Coque JJ, Martin JF, Calzada JG & Liras P (1991) The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase and pcbC of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremoium chrysogenum* and *Penicillin chrysogenum*. Mol. Microbiol. 5: 1125–1133

Cosmina P, Rodriguez F, de Ferra F, Grandi G, Perego M, Venema G & van Sinderen D (1993) Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. Mol. Microbiol. 8: 821–831

Crooks GP & Copley SD (1993) A surprising effect of leaving group on the nucleophilic aromatic substitution reaction catalyzed by 4-chlorobenzoyl-CoA dehalogenase. J. Amer. Chem. Soc. 115: 6422–6423

De Virgilio C, Burckert N, Barth G, Neuhaus JM, Boller & Wiemken A (1992) Cloning and disruption of a gene required for growth on acetate but not on ethanol: the acetyl-CoA synthetase gene of *Saccharomyces cerevisae*. Yeast 8: 1043–1051

DeWet JR, Wood KV, DeLuca M, Helinski DR & Subramani S (1987) Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell Biol. 7: 725–737

Diez B, Gutierrez S, Barredo JL, van Solingen P, van der Voort LH & Martin JF (1990) The cluster of penicillin biosynthetic genes: Identification and characterization of the pcbAB gene encoding the alpha aminoadipyl-cysteinyl-valine synthetase and linkage to the pcbC and penDE genes. J. Biol. Chem. 265: 16358–16365

Dirusso CC (1990) Primary sequence of the *Escherichia coli* fadBA operon, encoding the fatty acid-oxidizing multienzyme complex, indicates a high degree of homology to eucaryotic enzymes. J. Bacteriol. 172: 6459–6468

Driscoll JR & Taber HW (1992) Sequence organization and regulation of the *Bacillus subtilis men*BE operon. J. Bacteriol. 174: 5063–5071

Duronio RJ, Knoll LJ & Gordon JI (1992) Isolation of a *Saccharomyces cerevisiae* long chain fatty acyl: CoA synthetase gene (*FAA1*) and assessment of its role in protein N-myristoylation. J. Cell Biol. 117: 515–529

Eggen RI, Geerling ACM, Boshoven ABP & de Vos WM (1991) Cloning, sequence analysis, and functional expression of the acetyl CoA synthetase gene from *Methanothrix soehngenii in Escherichia coli*. J. Bacteriol. 173: 6383–6389

Farrell DH, Mikesell P, Actis LA & Crosa JH (1990) A regulatory gene, angR, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 *cro* and regulation by iron. Gene 86: 45–51

Fujino T & Yamamoto T (1992) Cloning and functional expression of a novel long-chain acyl-CoA synthetase expressed in brain. J. Biochem. 111: 197–203

Fuma S, Fujishima Y, Corbell N, D'Souza C, Nakano MM, Zuber P & Yamane K (1993) Nucleotide sequence of 5' portion of xi that contains the region required for competence establishment in *Bacillus subtilus*. Nucleic Acids Research 21: 93–97

Furukawa K & Chakrabarty AM (1982) Involvement of plasmids in total degradations of chlorinated biphenyls. Appl. Environ. Microbiol. 44: 619–626

Garre V, Murrillo FJ & Torres-Martinez S (1992) unpublished submission to Genbank

Groenewegen PEJ, van den Tweel WJJ & de Bont JAM (1992) Anaerobic bioformation of 4-hydroxybenzoate from 4-chlorobenzoate by the *coryneform* bacterium NJB-1. Appl. Microbiol. Biotechnol. 36: 541–547

Grundy FJ, Waters DA, Takova TY & Henkin TM (1993) Identification of genes involved in utilization of acetate and acetoin in *Bacillus subtilus*. Unpublished submission to GenBank

Gutierrez S, Diez B, Montenegro E & Martin JF (1991) Characterization of the *Cephalosporium acrenonium* pcbAB gene encoding alpha-aminoadipyl-cysteinyl-valine synthetase, a large multidomain peptide synthetase: linkage to the pcbC gene as a cluster of early cephalosporin biosynthetic genes and evidence of multiple functional domains. J. Bacteriol. 173: 2354–2365

Haese A, Schubert M, Hermann M & Zocher R (1993) Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalyzing N-methyldepsipeptide formation in *Fusarium scirpi*. Mol. Microbiol. 7: 905–914

Haggblom MM (1992) Microbial breakdown halogenated aromatic pesticides and related compounds. FEMS Microbiol. Reviews 103: 29–72

Heaton MP & Neuhaus FC (1992) Biosynthesis of D-alanyl-lipoteichoic acid: cloning, nucleotide sequence, and expression of the *Lactobacillus casei* gene for the D-alanine-activating enzyme. J. Bacteriol. 17: 4707–4717

Higson FK (1992) Microbial degradation of biphenyl and its derivatives. Adv. App. Microbiol. 37: 135–165

Hooper SW, Dockendorf JC & Sayler GS (1989) Characteristics and restriction analysis of the 4-chlorobiphenyl catabolic plasmid pSS50. Appl. Environ. Microbiol. 55: 1286–1288

Hori K, Yamamoto Y, Minetoki T, Kurotsu T, Kanda M, Miura S, Okamura K, Furuyama J & Saito Y (1989) Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. J. Biochem. 106: 639–645

Hori K, Yamamoto Y, Tokita K, Saito F, Kurotsu T, Kanda M, Okamura K, Furuyama J & Saito Y (1991) The nucleotide sequence for a proline-activating domain of gramicidin S synthetase gene from Bacillus brevis. J. Biochem. 110: 111–119

Jackowski S, Jackson PD & Rock CO (1993) Sequence and expression of the 2-acylglycerophosphoelanolamine acyltransferase/acyl-acyl carrier protein synthetase (aas) gene and its role in lysophospholipid acylation in *Escherichia coli*. Unpubished submission to Genbank

Klages U, Krauss S & Lingens F (1983) 2-Haloalkonoic acid dehalogenase from 4-chlorobenzoate degrading *Pseudomonas* sp. CBS-3. Hoppe-Seyler Z. Physiol. Chem. 364: 529–535

Klages U & Lingens F (1980) Degradation of 4-chlorobenzoic acid by a *Pseudomonas* sp. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. C 1: 215–223

— (1979) Degradation of 4-chlorobenzoic acid by a *Norcardia* species. FEMS Microbiol. Lett. 6: 201–203

Klages U, Markus A & Lingens F (1981) Degradation of 4-chlorophenylacetic acid by a *Pseudomonas* species. J. Bacteriol. 146: 64–68

Kraetzchumar J, Krause M & Marahiel MA (1989) Gramicidin S biosynthesis operon containing the structural genes grsA and

graB has an open reading frame encoding a protein homologous to fatty acid thioesterases. J. Bacteriol. 171: 5422–5429

Layton AC, Sanseverino J, Wallace W, Corcoran C & Sayler GS (1992) Evidence for 4-chlorobenzoic acid dehalogenation mediated by plasmids related to pSS50. Appl. Environ. Microbiol. 58: 399–402

Liang P-H, Yang G & Dunaway-Mariano D (1993) Specificity of 4-chlorobenzoate-Coenzyme A dehalogenase catalyzed dehalogenation of halogenated aromatics. Biochemistry 32: in press

Liu J, Duncan K & Walsh CT (1989) Nucleotide sequence of a cluster of *Escherichia coli* enterobactin biosynthesis genes: identification of *ent*A and purification of its product 2, 3-dihydro-2, 3-dihydroxybenzoate dehydrogenase. J. Bacteriol. 171: 791–798

Löffler F, Müller R & Lingens F (1992) Purification and properties of 4-halobenzoate-Coenzyme A Ligase from *Pseudomonas* sp. CBS3. Biol. Chem. Hoppe-Seyler 373: 1001–1007

Lozoya E, Hoffmann H, Douglas C, Schulz W, Scheel D & Hahlbrock K (1988) Primary structures and catalytic properties of isoenzymes encoded by the two 4-coumarate: CoA ligase genes in parsley. Eur. J. Biochem. 176: 661–667

MacCabe AP, van Liempt H, Palissa H, Unkles SE, Riach MDR, Pfeifer E, von Döhren H & Kinghorn JR (1991) Delta (L-alpha-aminoadipyl)-L-cysteine-D-valine synthetase from *Aspergillus nidulans*. J. Biol. Chem. 266: 12646–12654

Mallonee DH, White WB & Hylemon PB (1990) Cloning and sequencing of a bile acid-inducible operon from *Eubacterium* sp. strain VPI 12708. J. Bacteriol. 172: 7011–7019

Manning PA (1991) Unpublished submission to Genbank

Marks TS, Smith ARW & Quirk AV (1984) Degradation of 4-chlorobenzoic acid by *Arthrobacter* sp. Appl. Environ. Microbiol. 48: 1020–1025

Markus A, Klages U, Krauss S & Lingens F (1984) Oxidation and dehalogenation of 4-chlorophenylacetate by two-component enzyme systems from *Pseudomonas* sp. strain CBS-3. J. Bacteriol. 160: 618–621

Masuda T, Tatsumi H & Nakano E (1989) Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, *Luciola cruciata*. Gene 77: 265–270

Minami-Ishii N, Taketani S, Osumi T & Hashimoto T (1989) Molecular cloning and sequence analysis of the cDNA for rat mitochondrial enoyl-CoA hydratase. Eur. J. Biochem. 185: 73–78

Morris ME & Jinks-Robertson S (1991) Nucleotide sequence of the LYS2 gene of *Saccharomyces cerevisiae*: homology to *Bacillus brevis* tyrocidine synthetase I. Gene 98: 141–145

Müller R, Thiele J, Klages U & Lingens F (1984) Incorporation of [^{18}O] water into 4-hydroxybenzoic acid in the reaction of 4-chlorobenzoate dehalogenase from *Pseudomonas* sp. CBS3. Biochem. Biophys. Res. Commun. 124: 178–182

Müller R, Oltmanns RH & Lingens F (1988) Enzymatic dehalogenation of 4-chlorobenzoate by extracts from *Arthrobacter* sp. Su DSM 20407. Biol. Chem. Hoppe-Seyler 369: 567–571

Osumi T, Ishii N, Hijakata M, Kamijos K, Ozasa H, Furuta S, Miyazawa S, Kondo K, Inoue K, Kagamiyama H & Hashimoto T (1985) Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-CoA: hydratase-3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. J. Biol. Chem. 260: 8905–8910

Palossari PM, Vihinen M, Mäntsälia PI, Alexson SEH, Pihlajanismi T & Hiltunen JK (1991) Amino acid sequence similarities of the mitochondrial short chain Δ^3, Δ^2-enoyl-CoA isomerase and peroxisomal multifunctional Δ^3, Δ^2-enoyl-CoA isomerase, 2-enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase enzyme in rat liver. J. Biol. Chem. 266: 10750–10753

Pettigrew CA, Breen A, Corcoran C & Sayler GS (1990) Chlorinated biphenyl mineralization by individual populations and consortia of fresh water bacteria. Appl. Environ. Microbiol. 56: 2036–3045

Priefert H & Steinbüchel A (1992) Identification and molecular characterization of the aceryl-CoA synthetase gene (*aco*E) of *Alcaligenes eutrophus*. J. Bacteriol. 174: 6590–6599

Ruisinger S, Klager L & Lingens F (1976) Abbau der 4-chlorobenzoesäune durch eine *Arthrobacter*-species. Arch. Microbiol. 110: 253–256

Rusnak F, Sakaitani M, Drueckhammer D, Reichert J & Walsh CT (1991) Biosynthesis of the *Escherichia coli* siderophore enterobactin: sequence of the *ent*F gene, expression and purification of entF, and analysis of covalent phosphophantetheine. Biochemistry 30: 2916–2927

Sato S, Hayashi M, Imamura S, Ozeki Y & Kawaguchi A (1992) Primary structures of the genes, *faoA* and *faoB*, from *Pseudomonas fragi* B-0771 which encode the two subunits of the HDT multienzyme complex involved in fatty acid β-oxidation. J. Biochem. 111: 8–15

Savard P, Charest H, Sylvestre M, Shareck F, Scholten JD & Dunaway-Mariano D (1992) Expression of the 4-chlorobenzoate dehalogenase genes of *Pseudomonas* sp. CBS3 in *Escherichia coli* and identification of the gene translation products. Can. J. Microbiol. 38: 1074–1083

Savard P, Péloquin L & Sylvestre M (1986) Cloning of *Pseudomonas* sp. strain CBS3 genes specifying dehalogenation of 4-chlorobenzoate. J. Bacteriol. 168: 81–85

Schmitz A, Gartemann K-H, Fiedler J, Grund E & Eichenlaub (1992) Cloning and sequence analysis of genes for dehalogenation of 4-chlorobenzoate from *Arthrobacter* sp. strain SU. Appl. Environ. Microbiol. 58: 4068–4071

Schneider B, Müller R, Franck R & Lingens F (1991) Complete nucleotide sequences and comparisons of the structural genes of two 2-haloalkanoic acid dehalogenase from *Pseudomonas* sp. strain CBS-3. J. Bacteriol. 173: 1530–1535

Scholten JD, Chang K-H, Babbitt PC, Charest H, Sylvestre M & Dunaway-Mariano D (1991) Novel enzymic hydrolytic dehalogenation of a chlorinated aromatic. Science 253: 182–185

Scott-Craig JS, Panaccione DG, Pocard J-A & Walton JD (1992) The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7-kilobase open reading frame. J. Biol. Chem. 267: 26044–26049

Sharma V, Suvarna K, Meganthan R & Hudspeth MES (1992) Menaquinone (Vitamin K2) biosynthesis: Nucleotide sequence and expression of the *menB* gene from *E. coli*. J. Bacteriol. 174: 5057–5062

Shields MS, Hooper SW & Sayler GS (1985) Plasmid-mediated mineralization of 4-chlorobiphenyl. J. Bacteriol. 163: 882–889

Shimao M, Onishi S, Mizumori S, Kato N & Sakazawa C (1989) Degradation of 4-chlorobenzoate by facultatively alkalophilic *Arthrobacter* sp. strain SB8. Appl. Environ. Microbiol. 55: 478–482

Smith DJ, Earl AJ & Turner G (1990) The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillin chrysogenum* is a 421,073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthesis. EMBO Journal 9: 2743–2750

Smith RF & Smith TF (1990) Automatic generation of primary sequence patterns from sets of related protein sequence. Proc. Natl. Acad. Sci. USA 87: 118–

Staab JF, Elkins MF & Earhart CF (1989) Nucleotide sequence of the *Escherichia coli* entE gene. FEMS Microbiol. Lett. 59: 15–20

Steele MI, Lorenz D, Hatter K, Park A & Sokatch JR (1992) Characterization of the mmsAB operon of *Pseudomonas aeruginosa* PAO encoding methylmalonate-semaldehyde dehydrogenase and 3-hydroxyisobutyrate dehydrogenase. J. Biol. Chem. 267: 13585–13592

Suuki H, Kawarabayasi Y, Kondo J & Abe T (1990) Structure and regulation of rat long-chain acyl-CoA synthetase. J. Biol. Chem. 265: 8681–8685

Sylvestre M, Masse R, Ayotte C, Messier T & Fauteux J (1985) Total biodegradation of 4-chlorobiphenyl (4CB) by a two membered bacterial culture. Appl. Microbiol. Biotechnol. 21: 192–195

Tatsumi H, Kajiyama N & Nakano E (1992) Molecular cloning and expression in *Escherichia coli* of a cDNA clone encoding luciferase of a firefly, *Luciola lateralis*. Biochim. Biophys. Acta 1131: 161–165

Thiele J, Müller R & Lingens F (1987) Initial characterization of 4-chlorobenzoate dehalogenase from *Pseudomonas* sp. CBS3. FEMS Microbiol. Lett. 41: 115–119

Toh H (1990) N-Terminal halves of gramicidin 4 synthetase 1 and tyrocidine synthetase 1 as novel members of the firefly luciferase family. Protein Sequences Data Analysis 3: 512–521

Tomioka Y, Hirose A, Moritani H, Hishinuma T, Hashimoto T & Mizugaki M (1992) cDNA cloning of mitochondrial delta-3, delta-2-enoyl-CoA isomerase of rat liver. Biochim. Biophys. Acta 1130: 109–112

Tsoi T, Zaïtsev GM, Platnikova EG, Kosheleva IA & Boronia AM (1991) Cloning and expression of *Athrobacter globiformis* 4-chlorobenzoate-4-hydroxylase in *Escherichia coli*. FEMS Microbiol. Lett. 81: 165–170

Turgay K, Krause M & Marahiel MA (1992) Four homologous domains in the primary structure of GrsB are related to domains in a super-family of adenylate-forming enzymes. Mol. Microbiol. 6: 529–546

Van Beilen JB, Eggink G, Enequist H, Box R & Witholt B (1992) DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alk*JKL gene of *Pseudomonas oleovorans*. Mol. Microbiol. 6: 3121–3136

Van den Tweel WJJ, Terburg N, Kok JB & deBont JAM (1986) Bioformation of 4-hydroxybenzoate from 4-chlorobenzoate by *Alcaligenes denitrificans* NTB-1. Appl. Microbiol. Biotechnol. 25: 289–294

Weckermann R, Fuerbass R & Marahiel MA (1988) Complete nucleotide sequence of the *tycA* gene coding the tyrocidine synthetase 1 from *Bacillus brevis*. Nucleic Acids Research 16: 11841

Weinrauch Y, Guillen N & Dubnau DA (1989) Sequence and transcription mapping of *Bacillus subtilis* competence genes *com*B and *com*A, one of which is related to a family of bacterial regulatory determinants. J. Bacteriol. 171: 5362–5375

Wohlleben W, Alijah R, Dorendorf J, Hillemann D, Nussbaumer B & Pelzer S (1992) Identification and characterization of phosphinorhricin-tripeptide biosynthetic genes in *Streptomyces viridochromogenes*. Gene 115: 127–132

Wood KV, Lam YA, Seliger HH & McElroy WD (1989) Complementary DNA coding clickbeetle luciferases can elicit bioluminescence of different colors. Science (Washington, D.C.) 244: 700–702

Yamakawa A, Nishizawa M, Fujiware KT, Kawai S, Kawasaki H, Suzuki K & Takenawa T (1991) Molecular cloning a sequencing of cDNA encoding the phosphatidylinositol kinase from rat brain. J. Biol. Chem. 266: 17850–17583

Yura T, Mori H, Nagata T, Ishihama A, Fujita N, Iisono K, Mizobuchi K & Nakata A (1992) Systematic sequencing of the *Escherichia coli* genome: analysis of the 0–2.4 min region. Nucleic Acids Research 20: 3305–3308

Zaitsev GM, Tsoi TV, Grishenkov VG, Plotnikova EG & Boronia AM (1991) Genetic control of degradation of chlorinated benzoic acids in *Arthrobacter globiformis, corynebacterium sepedoniuum* and *Pseudomonas cepacia* strains. FEMS Microbiol. Lett. 81: 171–176

Zhao Y, Kung SD & Dube SK (1990) Nucleotide sequence of rice 4-coumarate: CoA ligase gene, 4-CL. 1. Nucleic Acids Research 18: 6144