Evidence That *pcpA* Encodes 2,6-Dichlorohydroquinone Dioxygenase, the Ring Cleavage Enzyme Required for Pentachlorophenol Degradation in *Sphingomonas chlorophenolica* Strain ATCC 39723†

Ling Xu,‡,§ Katheryn Resing,‡ Sherry L. Lawson,‡,§ Patricia C. Babbitt,|| and Shelley D. Copley*,‡,§

Department of Chemistry and Biochemistry and Cooperative Institute for Environmental Research, University of Colorado at Boulder, Campus Box 215, Boulder, Colorado 80309-0215, and Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Received January 19, 1999; Revised Manuscript Received March 31, 1999

ABSTRACT: An enzyme that catalyzes an Fe^{2+}-dependent reaction of 2,6-dichlorohydroquinone with O_2 has been isolated from *Sphingomonas chlorophenolica* sp. strain ATCC 39723, a soil microorganism capable of complete mineralization of pentachlorophenol. The product of the reaction is too unstable to allow spectroscopic characterization, but is apparently negatively charged and retains the two chlorine atoms of the substrate. The enzyme was partially sequenced using electrospray LC−MS, and one peptide was used to search the NCBInr database. This peptide matched a part of PcpA, a protein of unknown function that is induced in *S. chlorophenolica* in response to pentachlorophenol. Several other peptides could also be mapped onto the sequence of PcpA, suggesting that the enzyme is encoded by *pcpA*. PcpA has low but significant sequence similarity to an unusual class of extradiol dioxygenases. On the basis of the sequence analysis, the Fe^{2+} and O_2 dependence of the enzyme, and the characteristics of the product, the enzyme is proposed to be a 2,6-dichlorohydroquinone dioxygenase. The position of ring cleavage has not yet been identified.

Advances in synthetic chemistry during the last century have given chemists the ability to make a vast number of novel chemical compounds, some of which are xenobiotic (foreign to life). Some xenobiotic compounds, particularly those used as pesticides, are toxic. Biodegradation of toxic xenobiotics poses a double challenge to microorganisms in soil and natural aquatic systems. Since xenobiotics are not naturally occurring compounds, metabolic pathways for achieving their degradation are not necessarily available. Therefore, biodegradation of these compounds requires the assembly of new pathways, presumably largely by recruitment of pre-existing enzymes to serve new functions. Furthermore, in order to degrade a toxic xenobiotic, the organism must have effective strategies for avoiding or mitigating the toxicity of the xenobiotic and possibly some of its intermediates as well.

Pentachlorophenol (PCP[1]), a widely used wood preservative, is an important toxic xenobiotic. In recent years, a number of microorganisms that can degrade this compound, sometimes as a sole source of carbon, have been identified. *Sphingomonas chlorophenolica*[2] is the best studied of these microorganisms (*1*, *2*). The first two enzymes in the pathway (PCP hydroxylase and tetrachlorohydroquinone dehalogenase; see Figure 1) have been purified (*3*, *4*), cloned, and sequenced (*5*, *6*), and mechanistic studies of each are underway (*7*, *8*). *S. chlorophenolica* avoids the toxic effects of the chlorinated hydroquinone intermediates in the pathway by keeping the levels of these compounds in the low micromolar range (*9*), and may avoid the interference of PCP with oxidative phosphorylation by utilizing a gradient of sodium ions rather than protons to drive oxidative phosphorylation (T. Cort and S. D. Copley, unpublished observations).

Recently, the third step in the pathway for degradation of PCP in *S. chlorophenolica* was reported to be a hydrolytic dehalogenation which converts 2,6-dichlorohydroquinone (DCHQ) to 6-chloro-1,2,4-trihydroxybenzene (CTHB) (see Scheme 1) (*10*). DCHQ is also believed to be converted to CTHB in *Streptomyces rochei* 303 (*11*) and *Azotobacter* sp. strain GP1 (*12*), two other soil microorganisms that degrade chlorinated phenols. An Fe^{2+}-dependent protein which was designated DCHQ chlorohydrolase was purified and characterized from *S. chlorophenolica* strain ATCC 39723.

† This work was supported by awards to S.D.C. from the Environmental Protection Agency (Grant R 821274-01-0) and from the Petroleum Research Fund (Grant 32782-AC4), to K.R. from the NIH (Grant AR39730), and to P.C.B. from the DOE (Grant DE-FG03-96ER6) and from the Pharmaceutical Research and Manufacturers of America Foundation (Bioinformatics Faculty Development Award).

‡ Department of Chemistry and Biochemistry, University of Colorado at Boulder.

§ Cooperative Institute for Environmental Research, University of Colorado at Boulder.

|| University of California at San Francisco.

[1] Abbreviations: AEBSF, 4-(2-aminoethyl)benzenesulfonyl fluoride; BSTFA, bis(trimethylsilyl)trifluoroacetamide; CTHB, 6-chloro-1,2,4-trihydroxybenzene; DCHQ, 2,6-dichlorohydroquinone; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GSH, glutathione; NCBInr database, National Center for Biotechnology Information nonredundant (protein) database; TCHQ, tetrachlorohydroquinone; PAGE, polyacrylamide gel electrophoresis; PCP, pentachlorophenol.

[2] This microorganism was originally classified as *Flavobacterium* sp. strain ATCC 39723, but was reclassified in 1995 as *S. chlorophenolica* sp. strain ATCC 39723 (*2*).

10.1021/bi990103y CCC: $18.00
Published on Web 05/26/1999

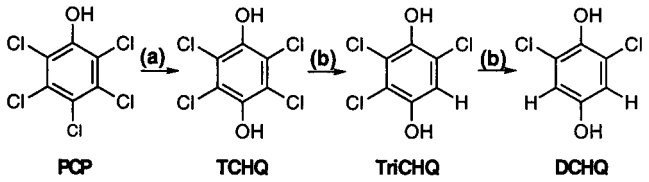
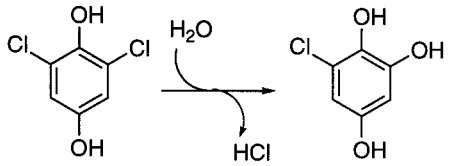


FIGURE 1: Initial part of the pathway for degradation of PCP in *S. chlorophenolica*: (a) PCP hydroxylase, O_2, and 2 NADPH (*3*) and (b) TCHQ dehalogenase and 2 GSH (*4*). PCP, pentachlorophenol; TCHQ, tetrachlorohydroquinone; TriCHQ, trichlorohydroquinone; and DCHQ, 2,6-dichlorohydroquinone.

Scheme 1



Because the only well-characterized hydrolytic aromatic dehalogenase (4-chlorobenzoyl CoA dehalogenase) does not use metal ions for catalysis (*13*), we were interested in how DCHQ dehalogenase accomplishes a comparable reaction utilizing a ferrous ion. We purified an iron-dependent enzyme with a molecular mass similar to that previously reported from *S. chlorophenolica* strain ATCC 39723. However, our results suggest that this enzyme is an O_2-dependent ring cleavage dioxygenase, rather than a hydrolytic dehalogenase. This enzyme is encoded by the *pcpA* gene and appears to belong to a new and unusual class of extradiol dioxygenases.

MATERIALS AND METHODS

Materials. CTHB was prepared as described by Latus et al. (*14*). DCHQ was prepared by reduction of 2,6-dichlorobenzoquinone (Chem Service) with ascorbic acid or dithionite. [U-^{14}C]DCHQ and [U-^{13}C]DCHQ were prepared from [U-^{14}C]PCP (Sigma) and [U-^{13}C]PCP (Cambridge Isotope Laboratories, Inc.), respectively, using the first two enzymes in the PCP degradation pathway. A crude extract of *Escherichia coli* strain CCL4 containing pCL3 (an overexpression plasmid for PCP hydroxylase) was prepared from a suspension of 3 g of cells in lysis buffer [50 mM potassium phosphate (pH 7.0) containing 5 mM DTT, 0.5 mM AEBSF, 50 μM leupeptin, 0.1 mg/mL DNase I, and 2 mg/mL lysozyme] by three passes through a French pressure cell at 1200 psi. The crude extract was dialyzed against 25 mM potassium phosphate buffer (pH 7.0) containing 2 mM DTT and 1 mM EDTA. An aliquot of the dialyzed crude extract (200 μL) was added to a 1 mL reaction mixture containing 190 μM [U-^{14}C]PCP (10.6 μCi/μmol), 1 mM NADPH, 0.13 mM FAD, 1 mM EDTA, and 40 mM potassium phosphate (pH 7.2). After 50 min, the reaction was quenched with 200 μL of 6 N HCl and the [U-^{14}C]TCHQ purified by reverse phase chromatography on a Rainin analytical C18 HPLC column (5 cm, 4.6 mm inside diameter) using a gradient of acetonitrile in 0.1% acetic acid (increasing from 0 to 85% acetonitrile over the course of 10 min). Subsequently, the [U-^{14}C]TCHQ was converted to [U-^{14}C]DCHQ in a reaction mixture containing 50 mM potassium phosphate (pH 7.2), 30 mM DTT, 1 mM glutathione, 0.2% ascorbic acid, 1 mM EDTA, and TCHQ dehalogenase purified as previously described (*15*). The [U-^{14}C]DCHQ was purified by reverse phase chromatography, evaporated to dryness, and stored at -20 °C.

Enzyme Purification. S. chlorophenolica strain ATCC 39723 was grown as previously described by Lee and Xun (*10*). *S. chlorophenolica* cells were harvested by centrifugation at 8700*g* for 25 min at 4 °C. The wet cell pellet was resuspended in 5 volumes of cold lysis buffer [50 mM potassium phosphate (pH 7.0) containing 5 mM DTT, 0.5 mM AEBSF, 50 μM leupeptin, 0.1 mg/mL DNase I, and 2 mg/mL lysozyme] and subjected to three passes through a cold French pressure cell at 12 000 psi. All subsequent operations were carried out at 4 °C. Cell debris was removed by centrifugation at 27000*g* for 15 min. The enzyme was purified using the procedure described by Lee and Xun (*10*) except that the protamine sulfate fractionation and dye chromatography steps were omitted and additional Mono Q chromatography steps were added after the gel filtration step to achieve the desired degree of purity. The enzyme was >90−95% pure as judged by SDS−PAGE. The enzyme was stored in 30% glycerol containing 5 mM DTT at -20 °C.

Enzyme Assays. Enzyme activity during protein purification was monitored in reaction mixtures containing 35 mM potassium phosphate buffer (pH 7.0), 50 μM DCHQ, 100 μM $Fe(NH_4)_2(SO_4)_2$, 0.2% ascorbate, 20 mM DTT, and 80 units each of superoxide dismutase and catalase (to convert any superoxide formed by interaction between DCHQ and O_2 to hydrogen peroxide and then to O_2 and H_2O). Reactions were carried out at 25 °C. DCHQ is stable in the absence of enzyme under these reaction conditions. After 30 min, reactions were quenched by addition of an equal volume of 1 N HCl. After centrifugation to remove particulate matter, aliquots were injected onto a Rainin analytical C18 reverse phase column (5 cm, 4.6 mm inside diameter) equilibrated with 20% acetonitrile in 0.1% acetic acid to quantitate the amount of remaining DCHQ. Assays for quantitation of enzyme activity were carried out by following the disappearance of DCHQ during the first 10% of the reaction. Reaction mixtures were comparable to those described above except that the concentration of DCHQ was 100 μM. One unit of activity is defined as the amount of enzyme required to convert 1 μmol of substrate per minute. The best results were achieved by preincubating the enzyme in the reaction mixture and initiating the reaction by adding DCHQ, presumably because this procedure allows reconstitution of the enzyme with Fe^{2+}.

Analysis of Metal Requirement. The enzyme was dialyzed against 25 mM MOPS (pH 6.9) containing 10% glycerol and 2 mM DTT for 18 h at 4 °C. Subsequently, assays were carried out in the absence of any metal ions and in the presence of 100 μM CuCl, $CuSO_4$, $MnCl_2$, $Fe_2(NH_4)_2(SO_4)_2$, or $FeCl_3$.

Assay of Chloride Release during Substrate Turnover. Chloride was measured by two procedures. The first was a modification of the method of Freier (*16*). Reaction mixtures were prepared using Tris-H_2SO_4 buffer in place of potassium phosphate buffer to reduce the background level of chloride. Aliquots (400 μL) for analysis of chloride were treated with 50 μL of concentrated nitric acid. After incubation at room temperature for 20 min, the samples were centrifuged in a microfuge for 10 min to remove precipitated protein. $AgNO_3$ (50 μL of a 0.1 N solution) was added, and the samples were incubated at room temperature for 10 min in

the dark. The turbidity due to the formation of AgCl was measured at 546 nm. Control reactions in which water was added in place of $AgNO_3$ were carried out to ensure that no signal at 546 nm was observed due to the product(s) of the enzymatic reaction. The second procedure utilized ion chromatography with indirect UV detection. Samples were injected onto a Rainin Hydropore strong anion exchange column (25 cm, 4.6 mm inside diameter) equilibrated with 2 mM phthalate buffer (pH 10). The column eluate was monitored by an absorbance detector set at 284 nm. The presence of chloride ions resulted in a negative peak at 8.4 min.

Comparison of Activity under Aerobic and Anaerobic Conditions. O_2 was removed from reaction mixtures containing 27 mM potassium phosphate buffer (pH 7.0), 0.4% ascorbic acid, 100 μM DCHQ, 20 μM $Fe(NH_4)_2(SO_4)_2$, 15 mM DTT, and 80 units each of superoxide dismutase and catalase by three cycles of freezing, evacuating, and purging with N_2 that had been passed over a column containing Dow Q-5 catalyst (Vacuum Atmospheres Co.) for removal of residual O_2. Removal of O_2 from protein solutions was more difficult because enzyme activity was lost upon freezing and thawing. Consequently, concentrated solutions of enzyme were incubated on ice under a stream of N_2 for 30 min prior to addition to duplicate O_2-free reaction mixtures at 28 °C. One of the reaction mixtures was then opened to the air. Aliquots were quenched and analyzed as described above.

Analysis of Reaction Products by UV−Vis Spectroscopy. Reaction mixtures containing 14 mM potassium phosphate (pH 7.0), 10% 2-propanol (which stabilizes the enzyme), 0.2% ascorbate, and 100 μM DCHQ were mixed in a cuvette with constant stirring at 30 °C. The absorbance was followed between 240 and 400 nm for samples with and without added enzyme.

Analysis of Reaction Products by HPLC. To analyze the product(s) formed from radiolabeled DCHQ by the enzyme, reaction mixtures were quenched and injected onto a Rainin analytical C18 reverse phase column (25 cm, 4.6 mm inside diameter) equilibrated with 0.1% acetic acid. After 5 min, the concentration of acetonitrile was increased to 20% by 30 min, and to 40% by 38 min. Under these conditions, DCHQ elutes at a retention time of 45 min. The eluate was monitored by an absorbance detector set at 210 nm, and fractions were collected for scintillation counting.

Analysis of Reaction Products by Mass Spectrometry (MS). Gas chromatography−MS (GC−MS) was carried out on a Hewlett-Packard 5988A GC−MS instrument using a fused silica capillary column (25 m) with a film of cross-linked 5% phenylmethylsilicone (0.52 μm). Samples were prepared by extraction of acidified reaction mixtures with ethyl acetate. After drying with Na_2SO_4, the solvent was removed with a stream of N_2. Samples were treated with BSTFA in acetonitrile (2:1 v/v) at room temperature for 20 min with sonication to derivatize hydroxyl and carboxyl groups (*17*). The solvent was removed with a stream of N_2 and the residue redissolved in ethyl acetate prior to injection onto the GC−MS instrument. Electrospray MS was carried out on a Perkin-Elmer Sciex API-III$^+$ triple-quadropole mass spectrometer equipped with a nebulization-assisted electrospray source. For analysis by positive mode electrospray LC−MS, the solvent was removed from the extracted product(s) and the residue was dissolved in 5% formic acid/50% methanol before being introduced onto the instrument by infusion. The orifice voltage and ion source voltage were set at 75 and 4600 V, respectively. For analysis by negative mode electrospray LC−MS, the residue was dissolved in 100 μM $(NH_4)_2CO_3$ containing 50% methanol and then introduced into the mass spectrometer by infusion, or dissolved in 0.1% acetic acid containing 5% methanol and injected onto a Vydac C18 column (15 cm, 0.32 mm inside diameter) controlled by an Applied Biosystems HPLC system coupled to the electrospray LC−MS instrument. The products were eluted with a gradient of methanol in 0.1% acetic acid (increasing by 1% methanol per minute) at a flow rate of 20 μL/min. The orifice voltage and ion source voltage were set at −45 and −4400 V, respectively.

Analysis of Reaction Products by NMR. [U-^{13}C]DCHQ was converted to product(s) in a reaction mixture containing 100 μM [U-^{13}C]DCHQ and sufficient enzyme to convert all of the substrate within 20 min at 30 °C. A control reaction mixture lacking enzyme was also prepared. The reaction mixtures were incubated for 30 min, and then acidified and extracted with ethyl acetate. The solvent was removed with a stream of N_2, and the residue dissolved in ethanol-d_6. Spectra were obtained on a Varian Unity Inova 500 MHz NMR spectrometer at a frequency of 125.832 Hz. Data were acquired for 3 h for the control sample and for 6 h for the sample which had contained enzyme.

Protein Sequencing by Mass Spectrometry. Protein samples were treated with 0.1 M DTT in 0.1 M Tris-HCl (pH 8.1) containing 8 M urea at 50 °C under N_2 to reduce disulfide bonds, and then treated with 0.1 M iodoacetamide for 30 min in the dark to alkylate free thiols. The sample was then subjected to PAGE. After staining with Coomassie blue, the band containing the enzyme was cut out, destained, and digested with trypsin in the gel as described by Rosenfeld et al. (*18*). Tryptic digests were eluted from the gel by two treatments with 200 μL of 60% acetonitrile/0.1% trifluoroacetic acid. For MALDI analyses, 1 μL of the extracted tryptic digest was coprecipitated with 1 μL of MALDI Quality matrix solution (α-cyano-4-hydroxy-*trans*-cinnamic acid in methanol, obtained from HP) on a MALDI plate and analyzed on a Perseptive Biosystems Voyager-DE STR mass spectrometer. Peptides were sequenced on a SciEx API-III$^+$ triple-quadropole mass spectrometer (LC−MS/MS) by on-line collision-induced dissociation. The extracted tryptic digest was dried on a Speed-Vac, redissolved in 0.1% formic acid, and then injected onto a POROS R120 C4 capillary column on an Applied Biosystems HPLC instrument coupled to the mass spectrometer, which is equipped with a nebulization-assisted electrospray source and a high-pressure collision cell. Peptides were selected in the first quadrupole and bombarded with Ar gas to induce fragmentation. In some cases, peptides were methylated as described by Hunt et al. (*19*). This procedure results in methylation of aspartates and glutamates, as well as the C-terminal carboxylate, and assists in the interpretation of spectra by indicating the number and positions of carboxylate-bearing residues in the peptide.

Analysis of Protein Molecular Masses by Electrospray LC−MS. Formic acid was added to a solution of pure enzyme to give a final concentration of 0.1%. The sample was then injected onto a POROS R120 C4 capillary column. The column was washed with 0.1% formic acid for 10 min at a

flow rate of 20 μL/min to remove buffer salts, and then connected to the electrospray LC−MS instrument described above. The protein was eluted with a gradient of acetonitrile in 0.1% formic acid (increasing by 1.7% acetonitrile per minute) at a flow rate of 20 μL/min.

Analysis of Metal Ion Content. The metal ion content of the enzyme was determined by inductively coupled plasma emission spectroscopy at the Colorado State University Soil Testing Laboratory. The enzyme was purified to homogeneity and then dialyzed against 20 mM potassium phosphate buffer (pH 7.8) containing 2 mM DTT which had been prepared in Chelex-treated water. The enzyme and a sample of the dialysate were submitted for metal analysis.

Sequence Analysis. The revised PcpA sequence (SP M55159) was used in a BLAST 2.0 search (*20*) of the NCBInr database to identify homologues of PcpA. The five most closely related proteins (PID g1881356, PID g2529479, PID g2415394, PID g2632008, and SP L43135; *E* values between 3×10^{-23} and 7×10^{-9}) were defined as the PcpA subgroup and used as query sequences for further analysis.[3] The Shotgun algorithm (Pegg and P. C. Babbitt, manuscript submitted) was used to identify proteins that are likely to be distant relatives of this cluster of proteins. Relationships among the pcpA subgroup proteins and these candidate sequences were evaluated using multiple sequence alignments and motif analysis. Multiple alignments were generated using either ClustalW (*21*) or Pileup (*22*). Motif analysis was performed using the MEME algorithm (*23*) on the PcpA subgroup of proteins, a set of extradiol dioxygenases, including those described by Eltis and Bolin (*24*), and others identified by BLAST and Shotgun analyses, a set of fosfomycin resistance proteins identified by BLAST analysis, and the set of glyoxalases described by Cameron et al. (*25*).

RESULTS

Purification and Characterization of the Enzyme. An enzyme which catalyzes the disappearance of DCHQ was purified from *S. chlorophenolica* strain ATCC 39723 with an overall yield of 4% and a total purification factor of 18-fold. From 10 g of wet cells obtained from 6 L of cell culture, 1.8 mg of protein was obtained. Analysis by electrospray LC−MS indicates that the enzyme has a molecular mass of 36 522 Da (data not shown). The DCHQ chlorohydrolase described by Lee and Xun was reported to have a molecular mass of about 40 000 Da as judged by gel filtration and 42 500 Da as judged by SDS−PAGE (*10*). The specific activity of our enzyme (using purification and reaction conditions which have not yet been optimized) was 0.3 unit/mg, where 1 unit is the amount of enzyme required to convert 1 μmol of substrate per minute. Lee and Xun report a specific activity for the DCHQ chlorohydrolase of 0.016 unit/mg using slightly different assay conditions (*10*).[4] In agreement with the previous report, enzyme activity was found to be stimulated by the addition of Fe^{2+}. The correspondence between the substrates, molecular masses, and iron dependence of the enzymes isolated by Lee and Xun and by us suggests that these two proteins are identical.

[3] The sequences of PID g2529479 and PID g2415394 are 99% identical, so only the former was included for further analysis.

[4] The published specific activity is 15.96 units/mg, and 1 unit is defined as the amount of enzyme required to convert 1 nmol of substrate per minute.

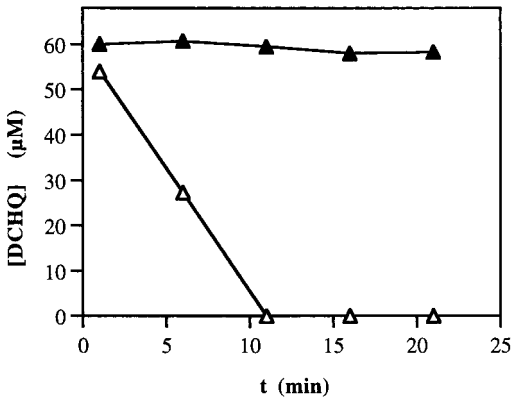


FIGURE 2: Effect of O_2 on the rate of conversion of DCHQ by the purified enzyme. Reaction mixtures contained 100 μM DCHQ, 20 μM $Fe(NH_4)_2(SO_4)_2$, 15 mM DTT, 50 units each of superoxide dismutase and catalase, and 20 μg of purified enzyme in a total volume of 300 μL. The concentration of remaining DCHQ was monitored by reverse phase HPLC. Reactions were carried out under N_2 (▲) or in air (△).

Only Fe^{2+} was found to support enzyme activity. Other transition metals, such as Fe^{3+}, Mn^{2+}, Cu^{+}, and Cu^{2+}, were ineffective. The degree of stimulation by Fe^{2+} varied depending upon the length of time required to complete the purification, suggesting that the endogenous metal ion is lost during purification. Indeed, analysis by inductively coupled plasma emission spectroscopy of the metal content of enzyme that was purified to an especially high degree to remove any contaminating metalloproteins showed only low levels of both iron (0.15 mol/mol of enzyme) and copper (0.13 mol/mol of enzyme). In particular, the level of iron was barely above the background. Organic cofactors such as NADPH, pterins, and α-ketoglutarate did not stimulate enzyme activity (data not shown).

The UV−vis spectrum of the purified enzyme exhibited a maximum at 280 nm. No long-wavelength absorbance characteristic of iron−sulfur clusters or organic cofactors, such as flavins or hemes, was observed.

Effect of O_2 on Enzyme Activity. The effect of O_2 on enzyme activity was evaluated by comparing the rates of DCHQ disappearance under ambient air and under N_2. As shown in Figure 2, the enzyme shows an absolute dependence on O_2 for substrate turnover.

The Product Formed from DCHQ Is Not CTHB. DCHQ chlorohydrolase was reported to convert DCHQ to CTHB. The identification of the product was based upon the correspondence of the HPLC retention times and UV−vis spectra of the product and authentic CTHB. Since these observations are not adequate for confirming the identity of the product, we revisited this issue with the goal of obtaining mass spectral data for the product. However, we found that the product formed by our enzyme is not CTHB. CTHB is not very stable in the presence of O_2, so we carried out the reaction using conditions under which CTHB could certainly be detected if it were formed. A reaction mixture containing 2.7 μM radiolabeled DCHQ was incubated on ice to slow the decomposition of CTHB. The amount of enzyme added was sufficient to complete the reaction in less than 8 min. Aliquots were withdrawn at 0, 2, 5, and 8 min, the reactions quenched with acid, and the samples injected onto a reverse phase column. Fractions were collected and subjected to scintillation counting to quantify the levels of substrate and

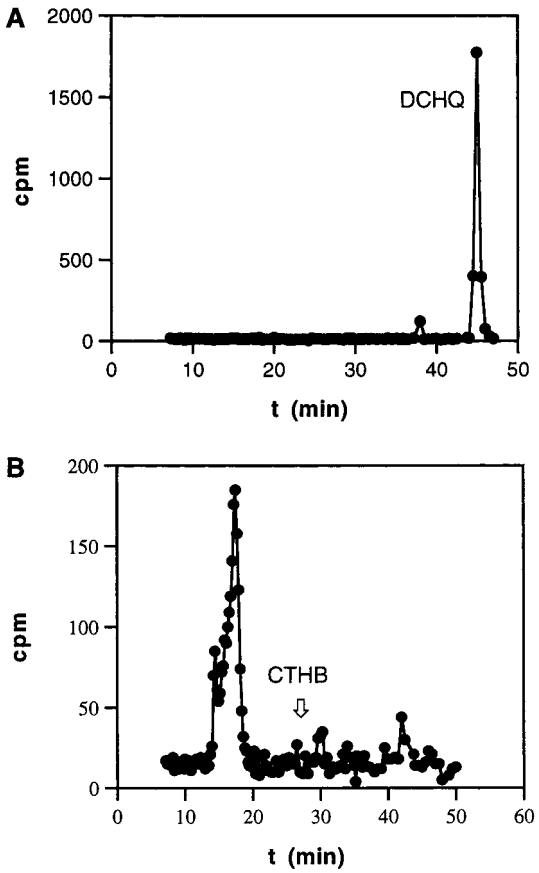


FIGURE 3: Chromatograms of reaction mixtures at (A) $t = 0$ and (B) $t = 8$ min. Reaction mixtures containing [U-^{14}C]DCHQ were incubated on ice, quenched with acid, and injected onto a reverse phase column. Fractions were collected and subjected to scintillation counting.

product(s). Data for the 0 and 8 min samples are shown in Figure 3. No radioactivity was detected at the retention time for CTHB at any time. A simple kinetic analysis ensured that CTHB would have accumulated to a significant level under these conditions if it had been formed. The proposed formation and decomposition of CTHB can be treated as a series of two consecutive pseudo-first-order irreversible reactions as shown in eq 1.

$$\text{DCHQ} \xrightarrow{k_1} \text{CTHB} \xrightarrow{k_2} \text{decomposition products} \quad (1)$$

The conversion of DCHQ to CTHB behaves as a pseudo-first-order reaction because the concentration of DCHQ (2.7 μM) is below the apparent $K_{M,DCHQ}$ for the enzyme [which is estimated to be 35 μM on the basis of the report of Lee and Xun (*10*) and our preliminary measurements], and therefore, the observed rate constant for the reaction will be given by eq 2

$$k_1 = (\text{apparent } k_{cat}/K_{M,DCHQ})[\text{enzyme}] \quad (2)$$

in which the apparent $k_{cat}/K_{M,DCHQ}$ refers to the $k_{cat}/K_{M,DCHQ}$ at ambient O_2 concentrations. Since the concentration of the enzyme is constant, the first step behaves according to simple pseudo-first-order kinetics. The pseudo-first-order rate constant for this step was found to be 0.73 min^{-1}. The decomposition of CTHB was found to obey first-order kinetics, although this reaction, which undoubtedly involves reaction of CTHB with O_2, is likely also to be pseudo-first-order in reality. The observed rate constant for CTHB decomposition in the reaction mixture on ice was 0.13 min^{-1}. These rate constants can be substituted into eq 3

$$[\text{CTHB}] = \frac{[\text{DCHQ}]_0 k_1}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t}) \quad (3)$$

which describes the accumulation of the intermediate in a kinetic scheme such as that shown in eq 1. Thus, the concentration of CTHB at 2, 5, and 8 min would be expected to be 66, 61, and 44% of the original DCHQ concentration, respectively. These levels would have been easily detectable. Therefore, we conclude that the product of the reaction is not CTHB.

Release of Chloride during Substrate Turnover. We measured the concentration of chloride in samples removed from reaction mixtures containing DCHQ and enzyme at various times. Chloride was measured using both the $AgNO_3$ procedure and the ion chromatography procedure. Both procedures indicated that 0.5−0.6 equiv of chloride was released during turnover of DCHQ, either during incubation with the enzyme or during the subsequent analytical procedures (data not shown).

Attempts To Characterize the Reaction Product. The product formed from DCHQ by the enzyme appears to be quite unstable. When reaction mixtures containing [U-^{14}C]-DCHQ were monitored by HPLC, we observed the appearance of a broad peak containing ^{14}C with a retention time of about 18 min (see Figure 3). The shape and width of this peak suggested that multiple species were present. Only 64% of the expected amount of radioactivity was recovered in this peak after 8 min, suggesting that additional decomposition products were also formed but did not elute from the HPLC column during the time that fractions were collected. These data indicate that the initial reaction product is unstable. Even after 2 min (data not shown), it has decomposed to multiple products. Most of these products are apparently negatively charged, since they can be extracted into ethyl acetate from acidic solutions, but not from neutral solutions (data not shown).

The conversion of DCHQ to product(s) was monitored using UV−vis spectroscopy. As the reaction proceeded, the absorbance from 300 to >400 nm increased significantly, but a discrete absorbance band was not discernible (data not shown).

Numerous efforts to obtain more definitive spectral characterization of the reaction product were thwarted by its instability. The product was extracted into ethyl acetate from the acidified reaction mixture, treated with BSTFA at room temperature, and subjected to GC−MS, but a product could not be identified. Electrospray LC−MS using both positive ion and negative ion detection was similarly unsuccessful. We also attempted to characterize the products using ^{13}C NMR, but no peaks could be observed in the spectrum of the extracted product, although the peaks due to [U-^{13}C]-DCHQ in the spectrum of the extracted control sample were clearly resolved.

Partial Sequencing of the Enzyme Using Mass Spectrometry and Identification of the Gene Encoding the Enzyme. The enzyme was subjected to PAGE, and the gel fragment

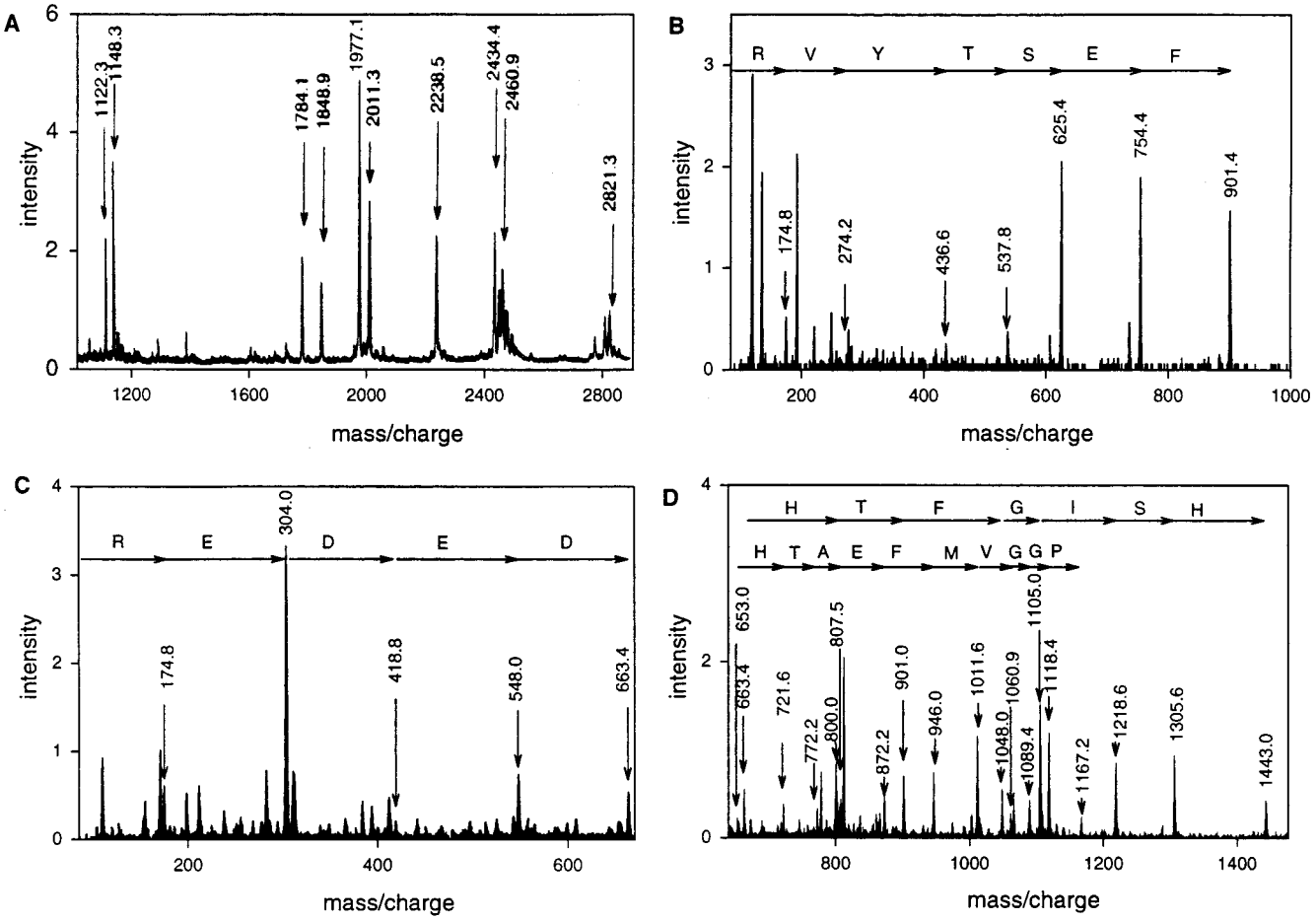


FIGURE 4: (A) MALDI mass spectrum of the in-gel digested peptides. Labeled ions indicate those with masses corresponding to the predicted masses of the peptides shown in Table 1. (B) MS/MS spectrum of the doubly charged ion of peptide 5 ($MH_2^{2+} = 560.6$). For clarity, only the y^+ ion series (corresponding to the C-terminal fragments generated by cleavage of amide bonds) is shown for the sequence [G]FESTYVR. (C and D) MS/MS spectrum of the triply charged ion of peptide 7 ($MH_3^{3+} = 812.0$). The spectrum is divided into two panels, covering *m*/*z* 100−670 (C) and 650−1450 (D). In panel C, the y^+ ions are shown for the partial sequence DEDER. In panel D, the y^+ ions are shown for the partial sequence HSiGFTH and y^{2+} ions are shown for the partial sequence PGGVMFEATH. (i represents either I or L.)

Table 1: Sequences of Peptides Derived by Tryptic Digestion of Enzyme[a]

peptide	calcd mass for MH^+ (Da)	exptl mass for MH^+ (Da)	sequence
MS1		3781.8[b]	(1628.4^{+2})iPiYHiYFADEi(762.0^{+2})
MS2		2460.7[b]	(1588.2^{+1})GHiNA(1095.4^{+1})
MS3		1815.8[b]	(949^{+1})VPiEHAQR
MS4		1148.0[b,c]	(944.2^{+1})HSWT(432.3^{+1})
MS5	774.8	773.9[c]	iGEEGNR
MS6		2080.9[d]	(961.8^{+2})TESGTiiD(518.2^{+2})[c]
MS7	2824.1	2821.8[b]	QGSWTiAEGiiHHGAFAVPDMDiQAR
MS8		1976.8[b,c]	(1460.0^{+1})TEGVGFTDF(505.2^{+1})
MS9	1122.2	1121.6[b,c]	GYFESTYVR
MS10	2434.6	2434.6[b,c]	TPGGVMFEATHSiGFTHDEDER
MS11	563.7	563.5[b]	GMDiK

[a] Calculated masses are given only for peptides which were sequenced completely. The masses of I and L are the same, so these residues cannot be distinguished and are consequently listed as i. Masses listed before and after the individual sequences correspond to the masses of fragments at which sequencing was ended and begun, respectively. [b] From LC−MS data. [c] From MALDI data. [d] From LC−MS analysis of methylated peptide.

containing the enzyme was treated with trypsin. The tryptic peptides were eluted from the gel and analyzed using a MALDI mass spectrometer (see Figure 4A). Five peptides were selected and sequenced completely by electrospray LC−MS/MS, and six more were sequenced in part (see Table 1). The spectra of selected peptides are shown in Figure 4.

Peptide 9 (GYFESTYVR) was used as a query sequence to search the NCBInr database using the MS-TAG 2.0 software (http://prospector.ucsf.edu/mshome.html). The query sequence showed a match with the protein encoded by *pcpA*, a gene in *S. chlorophenolica* ATCC 39723 previously identified by Xun and Orser (*26*) as a gene of unknown function that is induced by PCP. Several other peptides could be

mapped onto the *pcpA* sequence. However, one peptide (10 in Table 1) could not be mapped onto the predicted PcpA sequence in the reported reading frame. The C-terminal part of the peptide could be mapped onto the nucleotide sequence downstream of the reported coding region for *pcpA*, but in a different reading frame. Another peptide (11 in Table 1) could not be mapped onto the reported sequence. These data suggest that the enzyme is encoded by *pcpA*, but that there are errors in the published sequence. Recently, a revised sequence was submitted to GenBank under accession number M55159.[5] This sequence is in better agreement with our mass spectral data.

Search for Homologues of PcpA. The revised PcpA sequence was used in a BLAST 2.0 (*27*) search of the NCBInr protein database to identify homologues of PcpA. The five most closely related sequences (SP s43691, PID g2632008, SP L43135, PID g2529479, and PID g1881356) show statistically significant similarities to PcpA (*E* values between 3×10^{-23} and 7×10^{-9}) and represent proteins of unknown function. No clues as to the function of these proteins could be obtained from the surrounding genes. The Shotgun algorithm (S. C.-H. Pegg and P. C. Babbitt, manuscript submitted) was used to identify proteins that are likely to be distant relatives of this cluster of proteins. Shotgun identified a number of high-scoring proteins that are members of the vicinal oxygen chelate fold superfamily (*28*−*30*). This superfamily comprises extradiol dioxygenases, glyoxalases (also known as lactoylglutathione lyases), and fosfomycin resistance proteins. The unifying characteristic of this superfamily is the use of an active site transition metal to activate a substrate that is chelated to the metal center.

Since the PcpA subgroup is only distantly related to the known members of the vicinal oxygen chelate fold superfamily,[6] as judged by the BLAST and Shotgun results, we analyzed the metal-binding regions of the dioxygenases, glyoxalases, and fosfomycin resistance proteins (see Table 2) to determine which class of proteins was most similar to the PcpA subgroup. The metal-binding regions characteristic of each class were identified by a combination of structural comparisons, multiple sequence alignments, and identification of motifs in appropriate subsets of the superfamily using MEME (*31*). Bergdoll et al. (*30*) have shown that three of the four ligands to the active site zinc in human glyoxalase can be superimposed with the three ligands to the active site iron in 2,3-dihydroxybiphenyl 1,2-dioxygenase, despite the low sequence similarity between these proteins and the occurrence of domain swapping in the glyoxalase but not in the dioxygenase. The ligands in Table 2 are arranged accordingly. (The dioxygenases have only three protein ligands because three sites on the iron are required to bind O_2 and two groups on the substrate.) The multiple sequence alignments that were used were the structure-validated alignment of extradiol dioxygenases by Eltis and Bolin (*24*), a ClustalW (*21*) alignment of the extradiol dioxygenases most closely related to PcpA (chosen from the Shotgun output, data not shown), the structure-validated alignment of glyoxalases by Cameron et al. (*25*), a ClustalW alignment of fosfomycin resistance proteins and glyoxalases (data not shown), and a ClustalW alignment of the PcpA cluster of proteins with the most closely related extradiol dioxygenase (cf. Figure 5). Motifs found by MEME in subsets of the superfamily were compared with the structure-validated multiple sequence alignments to identify those containing ligands to the metal. These motifs contained between 12 and 25 bits of information. The information content of motifs found in the set of dioxygenases was similar to that of motifs found in the larger set containing the dioxygenases and the PcpA subgroup. Overall, the information content of the motifs for all the sets that were analyzed was lower than might be expected because of the domain duplication typical of the VOC superfamily (*24*, *30*). Thus, MEME found not only the true metal-binding regions but also the corresponding regions in the C-terminal domain of the protein that are more divergent and no longer bind metal ions. The information from the multiple sequence alignments and MEME analyses was combined to identify the most highly conserved region surrounding each metal ligand (see Table 2).

The motif analysis shown in Table 2 suggests that the PcpA subgroup is more similar to the extradiol dioxygenases than to the glyoxalases or fosfomycin resistance proteins. The region surrounding ligand 2 in the extradiol dioxygenases contains a highly conserved HH motif, in which the second H is the ligand. In glyoxalases, the corresponding region contains a highly conserved GHI motif. All of the members of the PcpA subgroup of proteins contain the HH motif characteristic of extradiol dioxygenases. Furthermore, the PcpA subgroup proteins lack the motif corresponding to ligand 4 in the glyoxalases. However, an important difference between the PcpA subgroup proteins and the dioxygenases is the lack of the motif corresponding to ligand 1 in the PcpA subgroup proteins. This finding suggests that there will be differences between the binding of iron at the active sites of these two classes of proteins.

A multiple sequence alignment of the PcpA subgroup proteins and the most closely related extradiol dioxygenase identified by Shotgun (catechol 2,3-dioxygenase from *Rhodococcus rhodochrous*, S43691) was generated using ClustalW. The percent identities between PcpA and the other unidentified proteins range between 26 and 31%. Percent identities between PcpA and the closest sequence neighbors among the extradiol dioxygenases cannot be credibly determined since they fall below 20%, obviating generation of a credible overall alignment except in regions where there is a high degree of sequence conservation. Despite the great divergence between these proteins, the N-terminal regions and the regions corresponding to ligands 2 and 3 [assigned on the basis of the alignment of 35 extradiol dioxygenases produced by Eltis and Bolin (*24*) and the results shown in Table 2] align quite well and show complete conservation of the putative ligands, as well as several other nearby residues. The residues in the alignment corresponding to predicted ligands to the iron have been boxed, and conserved residues that are believed to be involved in hydrogen bonding interactions in the active site based upon the work of Eltis and Bolin (*24*) are denoted with an asterisk.

[5] The annotation to entry M55159 assigns the function of the gene product as DCHQ chlorohydrolase.

[6] Both Blast searches and Shotgun analysis show that the sequence closest to all members of the PcpA subgroup is s43691, a characterized extradiol dioxygenase [catechol 2,3-dioxygenase (EC 1.13.11.2) from *Rhodococcus rhodochrous* strain CTM]. When the sequences of the PcpA subgroup were used as queries, Blast 2.0 *E* values for s43691 range from 2×10^{-5} to 0.9995.

Table 2: Comparison of Metal-Binding Regions in Members of the Vicinal Oxygen Chelate Fold Superfamily[a]

family	metal	Ligand 1	Seqs.	Ligand 2	Seqs.	Ligand 3	Seqs.	Ligand 4	Seqs.
Dioxygenase	Fe+2 or Mn+2	His165 in S43691	LGHFVR IGHFVR AGHYFY LGHYVL FGHYVL FGHIVV MGHYVL IGHCIL LGHCIV IGHVLI LGHLVV LGHVVL FDHALM FDHCLM FDHCAL LDHIAL LDHLLL LDHALL LDHLNI LDHINL LDHLNI LSHVLL LDHFNQ FAHVVL LAHFVL IAHIVL FDHALM FDHCLL	His227 in S43691	RIHHFML RVHHFLQ RLHHFML KTHHFML KMHHLMI RINHLMF RLNHLML VLGHMMV WFNHLML RLHHVSF KFHHVSF KIHHTSF GVHHISF GLHHIAF RLHHIAY KLHHLCY GWHHAAW RMHHVAF GMHHTAY GLAHTAF EFEHVGL RMHHVAF RLNHFML GIVHLMI	Glu278 in S43691	PSGVEVEYGW PSGFELEFGW PSGFAVEYGW PSGFAVEFGW PSGWLWEFGW PSGWLIEPGW PSGFSIEYGW PAGFAVELGF PSGNRNEVFC PSGNRNETFS PSGNRNETFA PGGNRVELFT PGGNRIELFG PWGSFCEYSA PDGHRIEIYT PDGNFVELQI PDGNMVELQI PDGHMLEVFC PSGNRNEVFC PSGNRSEVFC PGGNRVELFT PSGFMIEYGW PGGFDIEFGC	none	
Glyoxylase	Zn+2	Gln33A in human	QQTMLRVKD QQTMFRIKD NHTMLRVKD NHTCLRVKD NHTMIRVKD NHTMVRVKD NHTMIRIKN LHTMLRVGD LHTMIRVGD LHVVYRVGD CQVMLRVGD	His126B in human	GHIGIAVPD GHIGVTVDD GHICVSVPD GHICFSVSD GHICFTVDN GHVCISVDN GHICISCDD GHIGLSVDN GHIALSVDN GHIAIGVDD GHIALGVED GHFAISTQD AQIAIGTDD	Glu172B in human	PDGYWIEILN PDGYWIEIFD PDGYWVEVIQ PDGYWIELIT PDNYWIELVS PDNYWVEVIE PDGYSIEVVP PDGYKIELIE PDGYKIELIE PDGYKIEFIE PDGYMIELIQ PDGYKIELIQ PDGYTFELIQ PDGWKTVLVD PDGWKQVLVD	Glu99A in human	ATLELTHNWG STLELTHNWG GVLELTHNHG GVLELTHNEG GILELTYNFG GLLELTHNWG SVLELTHNWG AVIELTYNWG AEIELTYNWG AVIELTHNWG FVVELTYNYG IVLELTYNYG
Fosfomycin resistance	Mn+2		NHITYSVSN NHLLFSVSH NHLTLAVSD ?		?		PDGHKFEFHT LDGHKLELHT PDGHKLELHV		Motif not present
PcpA cluster	Fe+2		Motif not present	His227 in PcpA	*IIHHGAF* SVHHVAF TVHHVAF TVHHIAW	Glu276 in PcpA	*PGGVMFEATH* HGEILFEIAT KGGILFEIAT PNGILFELST		Motif not present

[a] Ligands to active site metal ions are underlined. PcpA sequences are shown in italics. Dioxygenase sequences represent the set of 23 analyzed by Eltis and Bolin (*26*) and the dioxygenases closest to PcpA (see the text).

DISCUSSION

We have isolated a protein from *S. chlorophenolica* ATCC 39723 that catalyzes the Fe^{2+}- and O_2-dependent conversion of DCHQ to an unstable product that could not be characterized by mass spectrometry or ^{13}C NMR spectroscopy. HPLC analysis demonstrated that the product was not CTHB, as previously reported (*10*). Other clues that this protein does not perform the previously reported reaction were provided by the requirement for O_2 and the detection of only substoichiometric amounts of chloride in reaction mixtures. (Since it is quite difficult to remove all O_2 from aqueous reaction mixtures, we suspect that the presence of some residual O_2 in the previously reported experiments may have supported substrate turnover.) The substoichiometric release of chloride suggests that the enzyme is not, in fact, a dehalogenase, since a dehalogenation reaction would result in the release of stoichiometric amounts of chloride. We attribute the partial release of chloride from the reaction product to nonenzymatic decomposition reactions, only one or some of which result in the loss of chloride.

Given the O_2 dependence of the enzyme, the enzyme is likely to be a monooxygenase or dioxygenase. The cofactor requirements of the enzyme do not coincide with those of known monooxygenases or dioxygenases that hydroxylate aromatic substrates. Phenol monooxygenases utilize either a flavin (*32*) or a heme cofactor (*33*) to catalyze the NADPH-dependent hydroxylation of phenols. Iron-dependent phenylalanine and tyrosine hydroxylases require pterin cofactors (*34*). Toluene 4-monooxygenase is a multicomponent enzyme system that requires NADH (*35*). The iron-dependent dioxygenases that introduce two hydroxyl groups onto aromatic substrates are oligomeric enzymes that contain iron−sulfur clusters in addition to mononuclear non-heme iron and require NADPH (*36*, *37*). The cofactor requirements of iron-

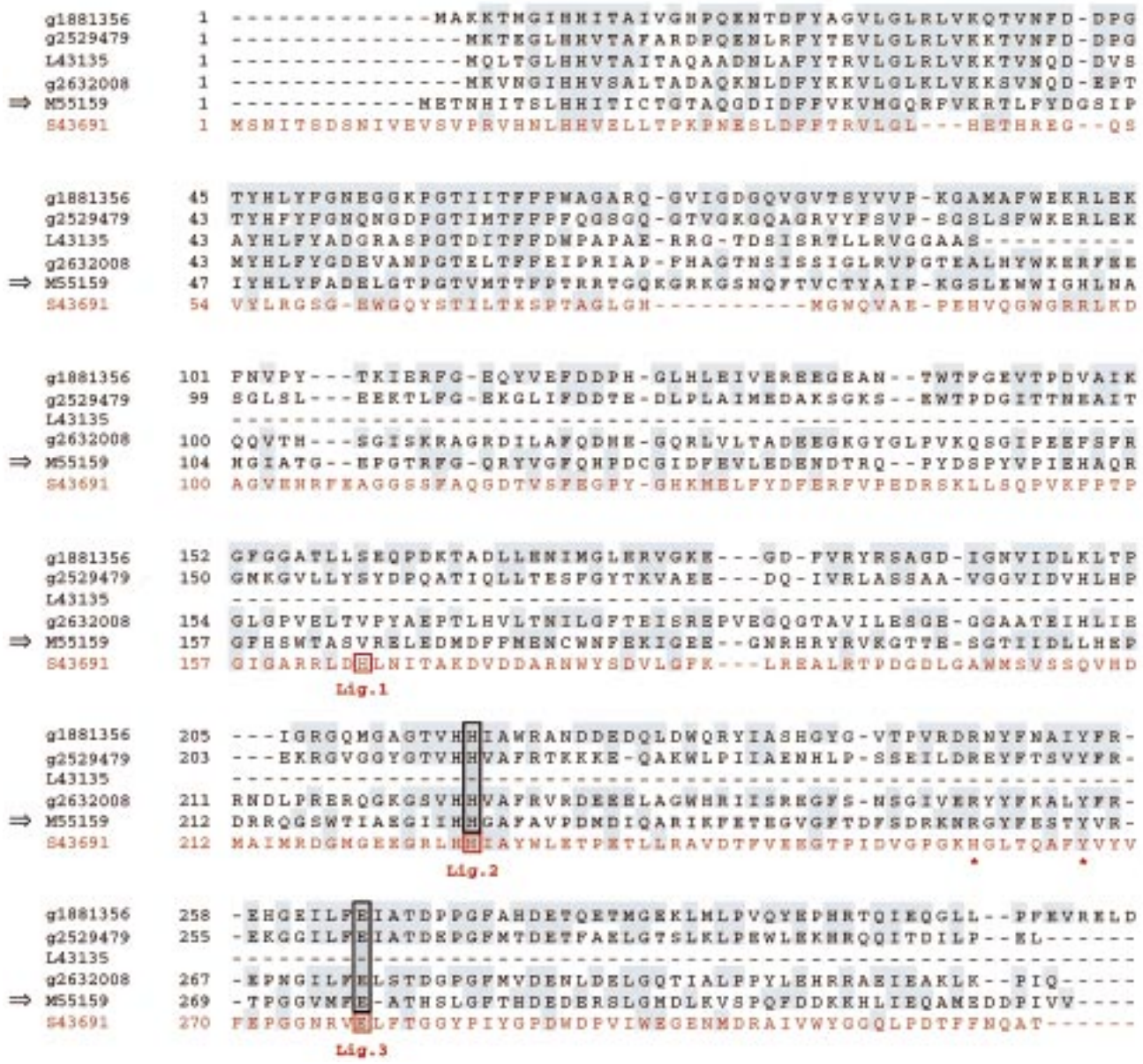


FIGURE 5: Alignment of PcpA (SP M55159) with the following related proteins: SP s43691, catechol 2,3-dioxygenase from *Rhodococcus rhodochrous* (strain CTM) plasmid TC1; PID g2632008, putative dioxygenase from *Bacillus subtilis* (YckA); SP L43135, putative dioxygenase fragment from *Methylobacterium extorquens*; PID g2529479, unknown protein from *B. subtilis* (YolE); and PID g1881356, unknown protein from *B. subtilis* (YdfO). Positions at which >65% of the residues are homologous [defined using the Dayhoff PAM250 matrix in the SeqVu alignment editing package, J. Gardner, The Garvan Institute of Medical Research, Sydney, Australia (1995)] are shaded gray. Conserved residues predicted to be ligands to the iron (*26*) are boxed; conserved residues predicted to be involved in hydrogen bonding interactions in the active site of S43691 are denoted with an asterisk.

dependent extradiol dioxygenases that cleave catechols or hydroquinones adjacent to one of the hydroxyl groups are simpler (*37*). These enzymes, like our enzyme, require only Fe^{2+} and O_2. Thus, the cofactor requirements of the enzyme are consistent with those of the extradiol ring cleavage dioxygenases.

The characteristics of the product are also most consistent with its formation via a ring cleavage reaction. The product of a ring cleavage dioxygenase reaction would be a carboxylic acid, and the products of a monooxygenation reaction or a dioxygenation reaction that adds two hydroxyl groups to the ring would be uncharged. Thus, the observation that the reaction products are likely to be negatively charged, since they can only be extracted into organic solvents after acidification of the reaction mixture, suggests that a ring cleavage reaction is most likely. The UV−vis spectrum of the product(s), which exhibits a long-wavelength absorbance above 300 nm, is also consistent with the existence of a ring cleavage product. Ring cleavage products typically absorb more strongly and at longer wavelengths than the catechol or hydroquinone reactants (*38*−*40*). Finally, as described below, the ring cleavage products formed from DCHQ would be expected to be very unstable.

To provide further insight into the function of the enzyme, we sequenced several peptides and used one as a query sequence to search the NCBInr database. A match to the *pcpA* gene product was found. The mapping of several peptides onto the translated sequence of the *pcpA* gene leaves little doubt that the enzyme is encoded by this gene. This assignment is further supported by the recent report that a

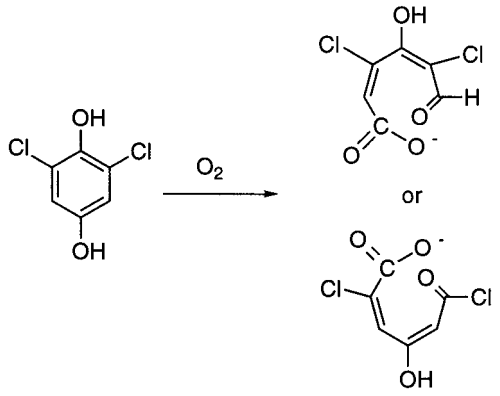


FIGURE 6: Proposed function of the enzyme encoded by *pcpA*.

strain of *S. chlorophenolica* in which *pcpA* had been knocked out accumulates DCHQ when grown in the presence of PCP (*41*), suggesting that *pcpA* encodes the enzyme which converts DCHQ to the next compound in the pathway.

The original published sequence of PcpA (*10*) is incomplete. Since the sequence of one peptide indicates a shift in reading frame, it appears that a sequencing error led to a shift to a reading frame with a premature stop codon. The actual gene sequence apparently extends for a considerable distance beyond the reported stop codon, since the exact mass of the protein as determined by electrospray LC−MS is 36 522 Da, but the mass predicted from the published PcpA sequence is only 30 746 Da. Recently, a revised sequence has been submitted to GenBank which is more consistent with our mass spectral data in that the error leading to the premature stop codon has been corrected and that the differences between the peptide sequences and the protein sequence have been resolved. However, the mass predicted from the revised sequence (36 515 Da) differs slightly from the experimental mass, and the reason for this discrepancy is not known.

PcpA is not closely related to any characterized proteins. The Shotgun algorithm, which allows the identification of distant relatives, showed that the PcpA subgroup (the cluster of proteins of unknown function defined earlier that shows statistically significant relationships to PcpA) is likely to be related to proteins in the vicinal oxygen chelate fold superfamily. The sequences of the putative metal-binding regions of the PcpA cluster of proteins and the identity of the transition metal required for PcpA suggest that these proteins are most similar to the extradiol dioxygenases. Most of the known extradiol dioxygenases cleave catechol substrates, but enzymes that cleave hydroquinones, such as gentisate (*38*), have also been identified. The products of cleavage of aromatic substrates by extradiol dioxygenases are often unstable (*38*), and spectroscopic characterization of the product has rarely been possible (*39*). In many cases, the identification of the product has rested upon HPLC retention times and the characteristic UV−vis changes due to the cleavage of the aromatic ring (*40*, *42*, *43*), or on the identities of further metabolites in crude extracts (*44*).

The similarities between the PcpA sequence and those of extradiol dioxygenases, the requirements for O_2 and Fe^{2+}, and the anionic nature and instability of the product strongly suggest that the enzyme is a ring cleavage dioxygenase. Thus, we propose that the enzyme is a DCHQ dioxygenase that catalyzes the oxidative cleavage of DCHQ (see Figure 6). We cannot determine whether cleavage occurs between the 1- and 2-positions, the 3- and 4-positions, or both. The observed substoichiometric release of chloride could in theory be due to formation of products resulting from cleavage at both positions, followed by the rapid hydrolysis of the acyl chloride produced by cleavage between the 1- and 2-positions. Although this possibility cannot be ruled out, it seems rather unlikely that the active site would accommodate the substrate in two binding orientations sufficiently well to produce substantial amounts of two products. The substoichiometric chloride release and extreme instability of the product are most consistent with cleavage between the 3- and 4-positions. Cleavage between the 1- and 2-positions would produce an acyl chloride, which would rapidly hydrolyze to form 2-chloromaleylacetate, releasing a stoichiometric amount of chloride. Furthermore, 2-chloromaleylacetate is moderately stable. Kaschabek and Reineke have synthesized 22 substituted maleylacetates as substrates for maleylacetate reductase (*45*). After hydrolysis of the dienelactone precursor at pH ~12, 2-chloromaleylacetate was stable for several hours. The product formed by cleavage between the 3- and 4-positions would be 2,4-dichloro-3-hydroxy-*cis*,*cis*-muconic semialdehyde, the tautomer of which is 3,5-dichloromaleylacetaldehyde. Substituted maleylacetates (which differ from maleylacetaldehydes only in the oxidation state of carbon 6) with halogen substituents in the 3-position are extremely unstable. Kaschabek and Reineke report that, after hydrolysis of the corresponding dienelactone precursors, these compounds were stable for ≤1 min. Thus, cleavage between the 3- and 4-positions seems most likely.

If the product is indeed formed by cleavage between the 3- and 4-positions, our inability to obtain spectroscopic confirmation of the product structure is not unexpected. The instability of this product can be attributed to several factors. First, the initially formed ring cleavage product contains two *cis* double bonds, and is thus the least stable isomer of the muconic semialdehyde. Isomerization to the two possible *cis*,*trans* isomers and the *trans*,*trans* isomer is likely. Second, the position of the keto−enol equilibrium is uncertain, but both tautomers may be present in solution. Third, closely related compounds such as maleylacetone are known to form lactones (*40*, *46*). Fourth, polymerization of these compounds is possible. Fifth, the aldehyde group may be susceptible to oxidation. Sixth, vinylic substitution of the chlorine at the 3-position with a hydroxyl group derived from H_2O would result in the formation of a β-keto acid, which would be expected to undergo facile decarboxylation and/or subsequent substitution of the remaining chlorine. Vinylic substitution might be catalyzed by the intramolecular carboxylate group. Thus, there are multiple avenues for decomposition of the ring cleavage product which result in the formation of mixtures, no component of which is abundant enough to be detected.

The reaction catalyzed by DCHQ dioxygenase differs in two obvious respects from those which are typical of the extradiol dioxygenases. First, this is the only known extradiol dioxygenase whose primary role is to cleave a dichlorinated aromatic compound. Since chlorine substituents generally impede the reaction of aromatic compounds with O_2, ring cleavage reactions usually occur when all or all but one of the chlorine atoms have been removed by prior dehaloge-

nation reactions. Second, in all cases studied so far, extradiol dioxygenases have been shown to bind their substrates in a bidentate fashion, either through adjacent hydroxyl groups in the cases of catechol 2,3-dioxygenase and protocatechuate 4,5-dioxygenase (*47*) or through a hydroxyl group and an adjacent carboxylate in the case of gentisate dioxygenase (*39*). However, in DCHQ, only one hydroxyl group will be able to bind to the iron atom. Thus, substrate binding in DCHQ dioxygenase must depart from the norm typical of the extradiol dioxygenases. This variation in substrate binding may be related to our finding that the N-terminal ligand typical of extradiol dioxygenases is not found in the PcpA subgroup of proteins and to other sequence characteristics that distinguish the PcpA subgroup from previously characterized members of the superfamily. The binding characteristics and catalytic properties of the enzyme will be explored further in future studies.

ACKNOWLEDGMENT

We are grateful to Dr. Bob Barkley for assistance with the GC−MS analysis, Mr. Mark Chappel for assistance with the NMR analysis, Professor Natalie Ahn for the use of her mass spectrometers, and Dr. Richard Johnson, Dr. Darla McCarthy, and Dr. Kandiah Anandarajah for helpful discussions.

REFERENCES

1. Ederer, M. M., Crawford, R. L., Herwig, R. P., and Orser, C. S. (1997) *Mol. Ecol. 6*, 39−49.
2. Nohynek, L. J., Suhonen, E. L., Nurmiaho-Lassila, E.-L., Hantula, J., and Salkinoja-Salonen, M. (1995) *Syst. Appl. Microbiol. 18*, 527−538.
3. Xun, L., and Orser, C. S. (1991) *J. Bacteriol. 173*, 4447−4453.
4. Xun, L., Topp, E., and Orser, C. S. (1992) *J. Bacteriol. 174*, 8003−8007.
5. Orser, C., Lange, C. C., Xun, L., Zahrt, T. C., and Schneider, B. J. (1993) *J. Bacteriol. 175*, 411−416.
6. Orser, C. S., Dutton, J., Lange, C., Jablonski, P., Xun, L., and Hargis, M. (1993) *J. Bacteriol. 175*, 2640−2644.
7. McCarthy, D. L., Navarrete, S., Willett, W. S., Babbitt, P. C., and Copley, S. D. (1996) *Biochemistry 35*, 14634−14642.
8. McCarthy, D. L., Louie, D. F., and Copley, S. D. (1997) *J. Am. Chem. Soc. 119*, 11337−11338.
9. McCarthy, D. L., Claude, A., and Copley, S. D. (1997) *Appl. Environ. Microbiol. 63*, 1883−1888.
10. Lee, J.-Y., and Xun, L. (1997) *J. Bacteriol. 179*, 1521.
11. Zaborina, O., Latus, M., Eberspächer, J., Golovleva, L. A., and Lingens, F. (1995) *J. Bacteriol. 177*, 229−234.
12. Latus, M., Seitz, H.-J., Eberspächer, J., and Lingens, F. (1995) *Appl. Environ. Microbiol. 61*, 2453−2460.
13. Crooks, G. P., Xu, L., Barkley, R. M., and Copley, S. D. (1995) *J. Am. Chem. Soc. 117*, 10791−10798.
14. Latus, M., Seitz, H.-J., Eberspächer, J., and Lingens, F. (1995) *Appl. Environ. Microbiol. 61*, 2453−2460.
15. McCarthy, D. L., Navarrete, S., Willett, W. S., Babbitt, P. C., and Copley, S. D. (1996) *Biochemistry 35*, 14634−14642.
16. Freier, R. K. (1974) in *Wasseranalyse*, 2nd ed., pp 45−46, Walter de Gruyter, Berlin.
17. Gehrke, C. W., and Patel, A. B. (1997) *J. Chromatogr. 130*, 103−114.
18. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) *Anal. Biochem. 203*, 173−179.
19. Hunt, D. F., Yates, J. R., III, Shabanowitz, J., Winston, S., and Hauer, C. R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6233−6237.
20. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
21. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
22. Feng, D. F., and Dolittle, R. F. (1987) *J. Mol. Evol. 25*, 351−360.
23. Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers, in *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp 28−36, AAAI Press, Menlo Park, CA.
24. Eltis, L. D., and Bolin, J. T. (1996) *J. Bacteriol. 178*, 5930−5937.
25. Cameron, A. D., Olin, B., Ridderström, M., Mannervik, B., and Jones, T. A. (1997) *EMBO J. 16*, 3386−3395.
26. Xun, L., and Orser, C. S. (1991) *J. Bacteriol. 173*, 2920−2926.
27. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
28. Babbitt, P. C., and Gerlt, J. A. (1997) *J. Biol. Chem. 272*, 30591−30594.
29. Bernat, B. A., Laughlin, T., and Armstrong, R. (1997) *Biochemistry 36*, 3050−3055.
30. Bergdoll, M., Eltis, L. D., Cameron, A. D., Dumas, P., and Bolin, J. T. (1998) *Protein Sci. 7*, 1661−1670.
31. Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers, in *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp 28−36, AAAI Press, Menlo Park, CA.
32. van Berkel, W. J. H., and Müller, F. (1990) in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., Ed.) pp 1−30, CRC Press, Boca Raton, FL.
33. Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) *Chem. Rev. 96*, 2841−2887.
34. Kappock, T. J., and Caradonna, J. P. (1996) *Chem. Rev. 96*, 2659−2756.
35. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry 35*, 9106−9119.
36. Mason, J. R., and Cammack, R. (1992) in *Annual Review of Microbiology* (Ornston, L. N., Balows, A., and Greenberg, E. P., Eds.) pp 277−305, Annual Reviews, Inc., Palo Alto, CA.
37. Que, J., L., and Ho, R. Y. N. (1996) *Chem. Rev. 96*, 2607−2624.
38. Harpel, M. R., and Lipscomb, J. D. (1990) *J. Biol. Chem. 265*, 6301−6311.
39. Harpel, M. R., and Lipscomb, J. D. (1990) *J. Biol. Chem. 265*, 22187−22196.
40. Kaschabek, S. R., Kasberg, T., Müller, D., Mars, A. E., Janssen, D. B., and Reineke, W. (1998) *J. Bacteriol. 180*, 296−302.
41. Chanama, S., and Crawford, R. L. (1997) *Appl. Environ. Microbiol. 63*, 4833−4838.
42. Crawford, R. L., Hutton, S. W., and Chapman, P. J. (1975) *J. Bacteriol. 121*, 794−799.
43. Boldt, Y. R., Sadowsky, M. J., Ellis, L. B., Que, L., Jr., and Wackett, L. P. (1995) *J. Bacteriol. 177*, 1225−1232.
44. Wittich, R.-M., Wilkes, H., Sinnwell, V., Francke, W., and Fortnagel, P. (1992) *Appl. Environ. Microbiol. 58*, 1005−1010.
45. Kaschabek, S. R., and Reineke, W. (1995) *J. Bacteriol. 177*, 320−325.
46. Fowler, J., and Seltzer, S. (1970) *J. Org. Chem. 35*, 3529−3532.
47. Arciero, D. M., and Lipscomb, J. D. (1986) *J. Biol. Chem. 261*, 2170−2178.

BI990103Y