JOURNAL OF BACTERIOLOGY, Aug. 1997, p. 4850–4858 Vol. 179, No. 15
0021-9193/97/$04.00+0

Identification and Characterization of Genes Encoding Carbazole 1,9a-Dioxygenase in *Pseudomonas* sp. Strain CA10

SHIN-ICHI SATO, JEONG-WON NAM, KANO KASUGA, HIDEAKI NOJIRI, HISAKAZU YAMANE, AND TOSHIO OMORI*

Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan

Received 13 January 1997/Accepted 16 May 1997

Nucleotide sequence analysis of the flanking regions of the *carBC* genes of *Pseudomonas* sp. strain CA10 revealed that there were two open reading frames (ORFs) ORF4 and ORF5, in the upstream region of *carBC*. Similarly, three ORFs, ORF6 to ORF8, were found in the downstream region of *carBC*. The deduced amino acid sequences of ORF6 and ORF8 showed homologies with ferredoxin and ferredoxin reductase components of bacterial multicomponent dioxygenase systems, respectively. ORF4 and ORF5 had the same sequence and were tandemly linked. Their deduced amino acid sequences showed about 30% homology with large (α) subunits of other terminal oxygenase components. Functional analysis using resting cells harboring the deleted plasmids revealed that the products of ORF4 and -5, ORF6, and ORF8 were terminal dioxygenase, ferredoxin, and ferredoxin reductase, respectively, of carbazole 1,9a-dioxygenase (CARDO), which attacks the angular position adjacent to the nitrogen atom of carbazole, and that the product of ORF7 is not indispensable for CARDO activity. Based on the results, ORF4, ORF5, ORF6, and ORF8 were designated *carAa*, *carAa*, *carAc*, and *carAd*, respectively. The products of *carAa*, *carAd*, and ORF7 were shown by sodium dodecyl sulfate-polyacrylamide gel electrophoresis to be polypeptides with molecular masses of 43, 36, and 11 kDa, respectively. However, the product of *carAc* was not detected in *Escherichia coli*. CARDO has the ability to oxidize a wide variety of polyaromatic compounds, including dibenzo-*p*-dioxin, dibenzofuran, biphenyl, and polycyclic aromatic hydrocarbons such as naphthalene and phenanthrene. Since 2,2′,3-trihydroxydiphenyl ether and 2,2′,3-trihydroxybiphenyl were identified as metabolites of dibenzo-*p*-dioxin and dibenzofuran, respectively, it was considered that CARDO attacked at the angular position adjacent to the oxygen atom of dibenzo-*p*-dioxin and dibenzofuran as in the case with carbazole.

Microbial degradation of xenobiotics such as polycyclic aromatic hydrocarbons, heterocyclic polyaromatics, and halogenated aromatics and aliphatics is quite useful to detoxify and mineralize these pollutants and thus is an important goal of research related to bioremediation. Among these pollutants, polychlorinated dibenzo-*p*-dioxin (DD), dibenzofuran (DBF), and carbazole (CAR) are well known to possess strong mutagenic and toxic activities (2) and also to be recalcitrant molecules (7). A critical point in the degradative pathways of these compounds is the initial oxidation of the chemically stable aromatic ring. In the aerobic degradation of many aromatic compounds, the multicomponent aromatic ring dioxygenases are very important enzyme complexes, consisting of two or three protein components (17, 28). Multicomponent dioxygenases catalyze a redox reaction in which molecular oxygen atoms are incorporated into the aromatic ring at the expense of the oxidation of NAD(P)H (17, 28). The recalcitrance of these xenobiotic compounds is considered to be due to requirement of rare and unique aromatic ring oxygenases for degradation.

Some bacterial and fungal strains which are able to degrade DD and DBF have been reported (8, 12, 18, 42). Recently, a dioxygenase that attacks at the angular position adjacent to the hetero atom of DD and DBF was purified from *Sphingomonas* sp. strain RW1 (5). However, its nucleotide sequence was not reported. On the other hand, almost no information has been obtained on the genes involved in the degradation of CAR, although several CAR-utilizing bacteria have been isolated (13, 14, 20, 24, 31).

Previously, we isolated *Pseudomonas* sp. strain CA10 from a sample of activated sludge as a microorganism with the ability to utilize CAR as a sole source of carbon, nitrogen, and energy (31). As shown in Fig. 1, the degradation pathway of CAR in strain CA10 was proposed on the basis of identification of the metabolic intermediates of CAR and comparison with the degradative pathways of other aromatic compounds such as biphenyl and naphthalene. The dioxygenase of strain CA10 is considered to attack at the angular position adjacent to the nitrogen atom to give the dihydroxylated intermediate. This unstable intermediate is considered to be spontaneously converted to 2′-aminobiphenyl-2,3-diol. To clone the genes encoding this angular dioxygenase, carbazole 1,9a-dioxygenase (CARDO), we attempted to clone a DNA fragment containing the *meta*-cleavage enzyme gene and the adjacent regions, because the genes encoding the enzymes which catalyze the initial oxidation of aromatic compounds and the subsequent *meta*-cleavage and hydrolysis of the *meta*-cleavage compound are often clustered.

In the accompanying paper (34), we reported the cloning of the genes encoding *meta*-cleavage enzyme (*carB*) and hydrolase (*carC*), which are involved in CAR degradation of strain CA10. In this report, we describe the identification and characterization of the genes encoding CARDO, which were found to be present in the adjacent regions of *carB* and *carC*. We also report that CARDO has the ability to oxidize a wide variety of polyaromatic compounds, including DD and DBF.

* Corresponding author. Mailing address: Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan. Phone: 81(3)3812-2111, ext. 3067. Fax: 81(3)5802-3326. E-mail: aseigyo@hongo.ecc.u-tokyo.ac.jp.

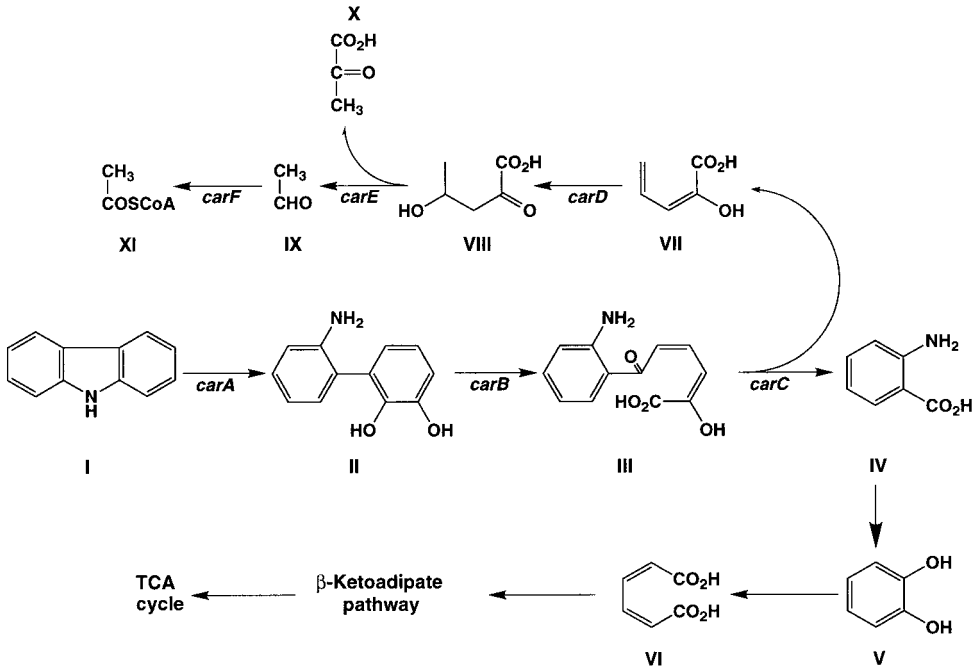

FIG. 1. Degradation pathway of carbazole in *Pseudomonas* sp. strain CA10. Gene designations: *carA*, carbazole 1,9a-dioxygenase; *carB*, 2′-aminobiphenyl-2,3-diol 1,2-dioxygenase; *carC*, 2-hydroxy-6-oxo-6-(2′-aminophenyl)-hexa-2,4-dienoic acid (*meta*-cleavage compound) hydrolase; *carD*, 2-hydroxypenta-2,4-dienoate hydratase; *carE*, 4-hydroxy-2-oxovalerate aldolase; *carF*, acetaldehyde dehydrogenase (acylating). Compounds: I, carbazole; II, 2′-aminobiphenyl-2,3-diol; III, 2-hydroxy-6-oxo-6-(2′-aminophenyl)-hexa-2,4-dienoic acid (*meta*-cleavage compound); IV, anthranilic acid; V, catechol; VI, *cis*,*cis*-muconic acid; VII, 2-hydroxypenta-2,4-dienoic acid; VIII, 4-hydroxy-2-oxovaleric acid; IX, acetaldehyde; X, pyruvic acid; XI, acetyl coenzyme A. TCA, tricarboxylic acid.

MATERIALS AND METHODS

Bacterial strains, plasmids, media, and culture conditions. Plasmids used in this study are depicted in Fig. 2. *Escherichia coli* JM109 (40) was used as the host strain for the plasmids constructed in this study. *E. coli* strains were grown on 2xYT medium (33) at 37°C. Ampicillin and chloramphenicol at final concentrations of 50 and 30 μg/ml, respectively, were used for liquid cultures and selection on agar plates. For plate cultures, the above-described media solidified with 1.6% (wt/vol) agar were used.

DNA manipulation. Plasmid DNA was prepared from the *E. coli* host strain by the alkaline lysis method (4). Restriction endonuclease, T4 DNA ligase, and a DNA blunting kit (Takara Shuzo Co. Ltd., Kyoto, Japan) were used according to the manufacturer's instructions. DNA fragments were extracted by the glass powder method (Geneclean II kit; Bio 101 Inc., La Jolla, Calif.) as instructed by the manufacturer. Other DNA manipulation was performed according to standard methods (33).

Nucleotide sequence determination, alignment, and phylogenetic analysis. Unidirectional deletion mutants were constructed by using a kilo-sequence DNA deletion kit (Takara Shuzo) (19). Nucleotide sequence determination was carried out as described previously (34). The obtained nucleotide sequences were analyzed with DNASIS-Mac software (version 3.6; Hitachi Software Engineering Co. Ltd., Yokohama, Japan). Alignment of the sequences of CarAa with other large (α) subunits was performed with the CLUSTAL W package (39). Phylogenetic analysis was performed by using the neighbor-joining algorithm of PHYLIP (version 3.5c; obtained from J. Felsenstein, University of Washington) as described previously (34).

Resting cell reaction. Resting cells of *E. coli* strains were prepared as described previously (34). We used minimal medium (31) for washing and resuspension of *E. coli* strains. The solution of CAR (100 mg/ml) was prepared in dimethyl sulfoxide, and then a 50-μl aliquot of this solution was added to 5 ml of suspension solution. The resting cell reaction was conducted on a reciprocal shaker (300 strokes/min) at 30°C for 12 h. The extraction of metabolite was carried out as described previously (34). Each extract was analyzed by analytical thin-layer chromatography (TLC) and/or gas chromatography-mass spectrometry (GC-MS) after methylation with ethereal diazomethane.

Substrate specificity. To investigate the substrate specificity of CARDO, we conducted the resting cell reaction with *E. coli* cells carrying pUCARA, in which only *carAaAcAd* and open reading frame 7 (ORF7) were encoded (Fig. 2). In this experiment, *E. coli* carrying pUC119 was used as a control. Aromatic compounds listed in Table 1 were dissolved in ethanol or dimethyl sulfoxide, and resultant stock solutions (10 mg/ml) were added to the reaction mixture as described above. Incubation was carried out for 18 h under the conditions described in above. The reaction mixtures were similarly extracted, and the resultant extracts were analyzed by analytical TLC. GC-MS analyses were carried out to quantify the remaining substrate. To quantify the substrate remaining after the resting cell reaction, we compared the peak area for its molecular ion with that of substrate extracted from the reaction mixture prepared with *E. coli* cells harboring pUC119. The metabolites of DD and DBF were further analyzed by ^{1}H nuclear magnetic resonance (NMR) spectroscopy and direct mass spectrometry after purification by preparative TLC.

The substrate specificity of CARDO was also examined by means of resting cell reaction using strain CA10. Strain CA10 cells were grown on 2xYT medium supplemented with carbazole at a final concentration of 0.1% (wt/vol) at 30°C. The cells were washed and resuspended as described for recombinant *E. coli* cells. The metabolite was extracted and analyzed by analytical TLC as described in above.

Analytical methods. GC-MS analysis was carried out as described previously (34). Analytical TLC was performed on 0.25-mm-thick, precoated silica gel plates containing a fluorescence indicator (Merck, Darmstadt, Germany) and developed with the solvent system toluene-dioxane-acetic acid (90:25:4, by volume). For the isolation of the metabolites of DD and DBF, preparative TLC was carried out as follows. The sample was developed on precoated silica gel plate (1 mm thick; Merck) with the solvent system hexane-ethyl acetate-acetic acid (10:10:1, by volume). The desired band was scraped off and eluted with H_2O-saturated ethyl acetate. The resultant elute was loaded onto a silica gel column before evaporation and used for other analytical experiments. The ^{1}H NMR spectra were determined with a JEOL JNM-A500 spectrometer (JEOL, Tokyo, Japan) operated at 500 MHz, using tetramethylsilane as an internal standard. Mass spectra were obtained with a model M-2000 spectrometer (Hitachi, Ltd., Tokyo, Japan) equipped with a direct inlet system. The electron accelerating voltage was 70 eV.

SDS-PAGE. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed by the method of Laemmli (26), using 15% polyacrylamide gels. After the electrophoresis, protein staining of the gels was performed with Coomassie brilliant blue R-250.

Chemicals. CAR was purchased from Katayama Chemical Industries Co. Ltd., Osaka, Japan. Although this carbazole was of the highest purity commercially available, recrystallization from ethanol was done for further purification. DD was synthesized by the method described by Harms et al. (18). All other chemicals were of the highest purity commercially available.

Nucleotide sequence accession number. The nucleotide sequence data for the insert DNA fragment of pUCA1 are available from the DDBJ, EMBL, and GenBank nucleotide sequence databases under accession no. D89064.

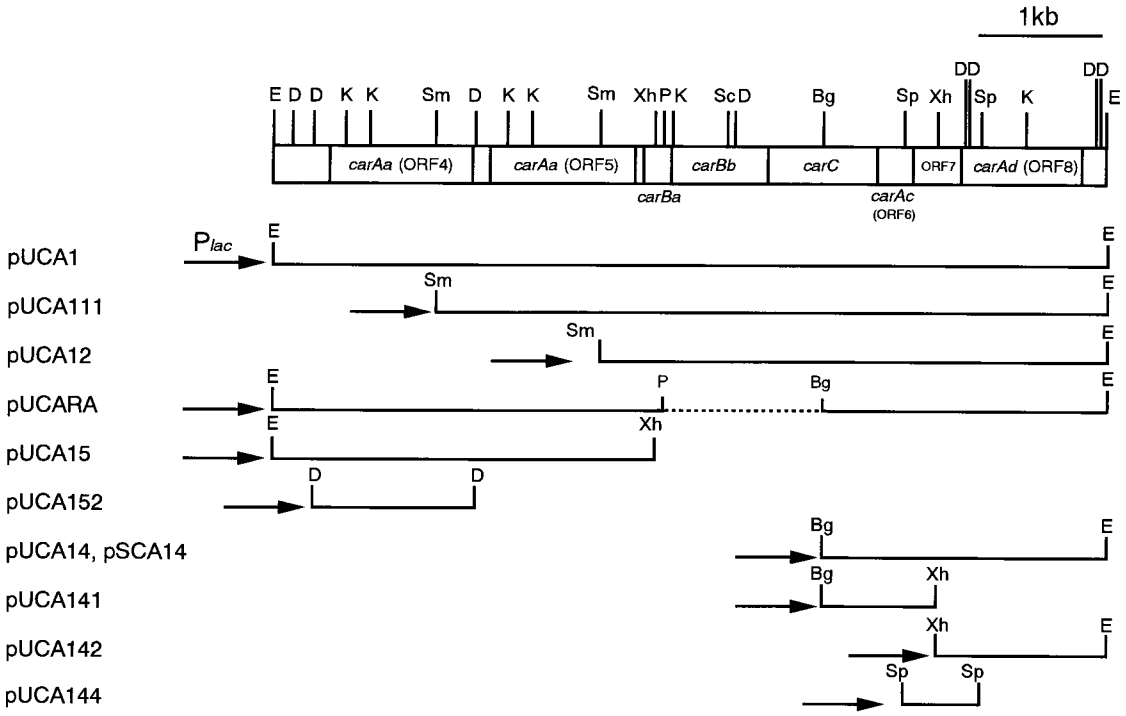

FIG. 2. Physical map of the 6.9-kb *Eco*RI insert of pUCA1 and of the different cloned DNA fragments described in this study. The boxes in the physical map indicate the sizes and locations of the ORFs derived from the DNA sequence analysis. Only restriction sites of relevant enzymes are shown. The location and orientation of the *lac* promoter of the pUC119 and pSTV29 vectors are indicated by arrows. The dotted line for the pUCARA insert indicates the deleted DNA fragment. Restriction site designations: E, *Eco*RI; D, *Dde*I; Sm, *Sma*I; Xh, *Xho*I; Bg, *Bgl*II; Sp, *Sph*I; K, *Kpn*I; H, *Hin*dIII; P, *Pst*I; Sc, *Sac*II.

RESULTS

Detection of CARDO activity. We observed that *E. coli* harboring pUCA1 (Fig. 2) (34) which contained *carBC* and the adjacent regions had the ability to form indigo on 2xYT plates. This phenomenon was reported for naphthalene dioxygenase (10) and cumene dioxygenase (1, 9), suggesting that pUCA1 contains genes encoding CARDO. Therefore, we assayed for conversion of CAR in the resting cell reaction using *E. coli* JM109 harboring pUCA1. The ethyl acetate extract of the reaction mixture was subjected to GC-MS analysis after methylation with ethereal diazomethane, with the result that the retention time and full-scan mass spectrum of the methylated derivative of the product were consistent with those of authentic anthranilic acid methyl ester (data not shown). In this analysis, we could not detect other intermediates such as 2′-aminobiphenyl-2,3-diol and 2-hydroxy-6-oxo-6-(2′-aminophenyl)-hexa-2,4-dienoic acid.

Nucleotide sequence analysis. Since in the companion paper (34) we reported the nucleotide sequence of the 2.6-kb *Sma*I-*Sph*I fragment (Fig. 2) in which *carB* and *carC* gene were located, we determined the nucleotide sequences of the 2.6-kb *Eco*RI-*Sma*I fragment and the 2.7-kb *Sph*I-*Eco*RI fragment, which were the upstream and downstream regions, respectively, of the 2.6-kb *Sma*I-*Sph*I fragment. We identified an additional five ORFs (ORF4 to ORF8) by computer analysis as shown in Fig. 2 and 3. The Shine-Dalgarno (SD) sequence (35) was found in the upstream region of the putative initiation codon of each ORF. Surprisingly, the nucleotide sequences of ORF4 and ORF5 were identical, and the 89-bp-long 5′ upstream sequence of ORF4 was the same as that of ORF5 except for one base. Similarly, the 22-bp-long 3′ downstream sequence of ORF4 was the same as that of ORF5. The 3′ and 5′ termini of other putative ORFs were also found in the upstream region of ORF4 and downstream region of ORF8, respectively.

Functional analysis of ORFs. To determine the functions of the ORFs, we performed homology searches and the following experiments.

(i) ORF6. The nucleotide sequence of ORF6 showed 58.0, 57.1, and 57.1% homologies with the ferredoxin components of chlorobenzene dioxygenase (41), biphenyl dioxygenase (27), and toluene dioxygenase (43), respectively. The consensus sequence of Rieske-type iron-sulfur proteins for the binding of a [2Fe-2S] cluster (32) was also found in ORF6. Since these results suggest that ORF6 encodes the ferredoxin component of CARDO, ORF6 was designated *carAc.*

(ii) ORF8. ORF8 showed 53.0 and 49.3% homologies with ferredoxin reductase components of toluene 3-monooxygenase (6) and naphthalene dioxygenase (25), respectively. The consensus sequence of plant-type iron-sulfur proteins for the binding of a [2Fe-2S] cluster (**C**X_4**CXXC**X_{29}**C**) (30) was located at a position close to the N-terminus of the deduced amino acid sequence of ORF8. The consensus sequence previously proposed to be involved in the binding of flavin adenine dinucleotide (FAD) and NAD (29) was also found in the C terminus of the deduced amino acid sequence of ORF8. These observations suggest that ORF8 encodes the ferredoxin reductase component of CARDO, and thus ORF8 was designated *carAd.*

(iii) ORF4, -5, and -7. ORF4 and ORF5 had identical sequences and were tandemly linked (Fig. 3A), and the deduced amino acid sequence showed about 30% homologies with large subunits of other terminal oxygenase components. The consensus sequence of Rieske-type iron-sulfur proteins for the binding of a [2Fe-2S] cluster (**CXH**X_{16-17}**CXXH**) (32) was

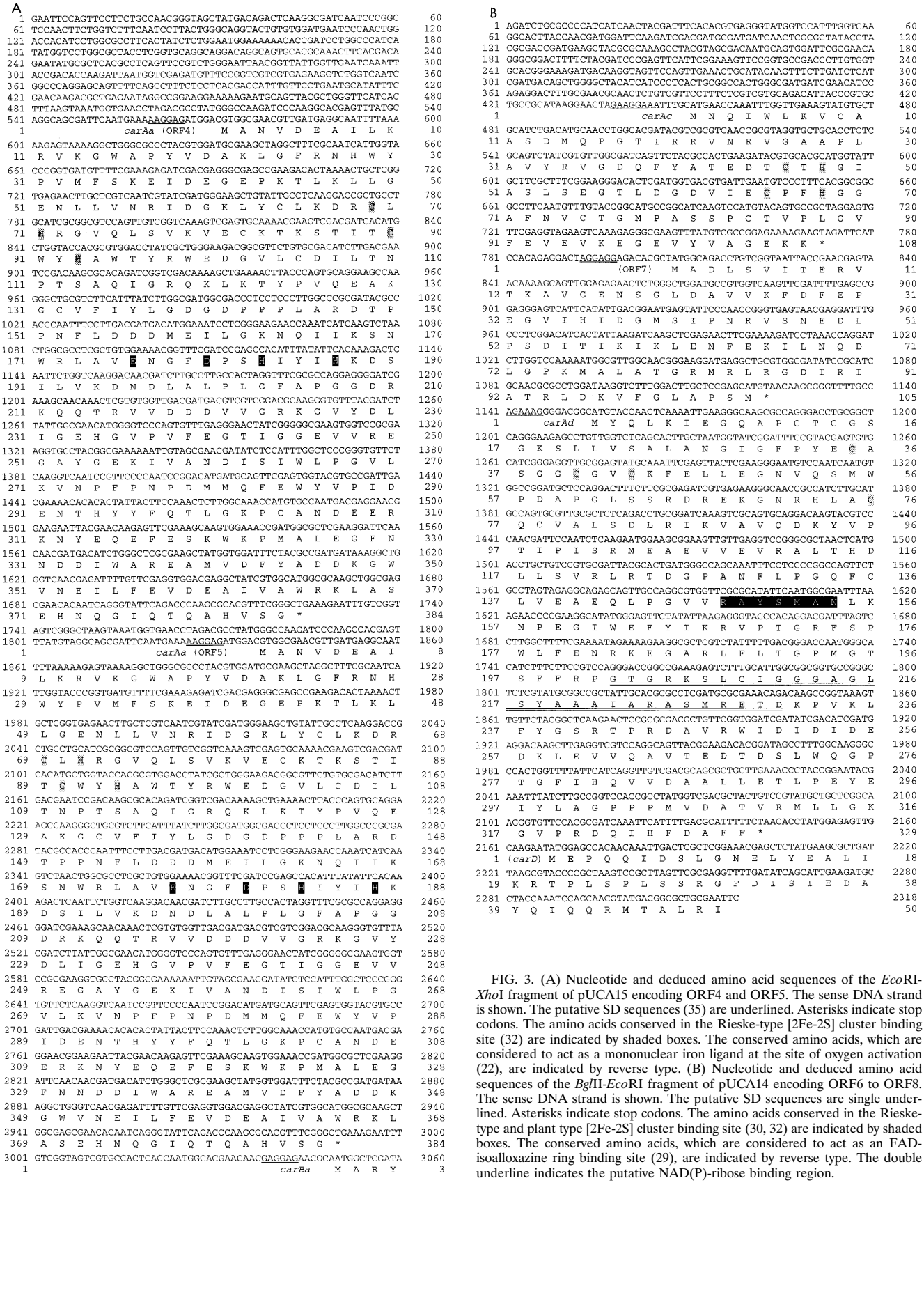

FIG. 3. (A) Nucleotide and deduced amino acid sequences of the *Eco*RI-*Xho*I fragment of pUCA15 encoding ORF4 and ORF5. The sense DNA strand is shown. The putative SD sequences (35) are underlined. Asterisks indicate stop codons. The amino acids conserved in the Rieske-type [2Fe-2S] cluster binding site (32) are indicated by shaded boxes. The conserved amino acids, which are considered to act as a mononuclear iron ligand at the site of oxygen activation (22), are indicated by reverse type. (B) Nucleotide and deduced amino acid sequences of the *Bgl*II-*Eco*RI fragment of pUCA14 encoding ORF6 to ORF8. The sense DNA strand is shown. The putative SD sequences are single underlined. Asterisks indicate stop codons. The amino acids conserved in the Rieske-type and plant type [2Fe-2S] cluster binding site (30, 32) are indicated by shaded boxes. The conserved amino acids, which are considered to act as an FAD-isoalloxazine ring binding site (29), are indicated by reverse type. The double underline indicates the putative NAD(P)-ribose binding region.

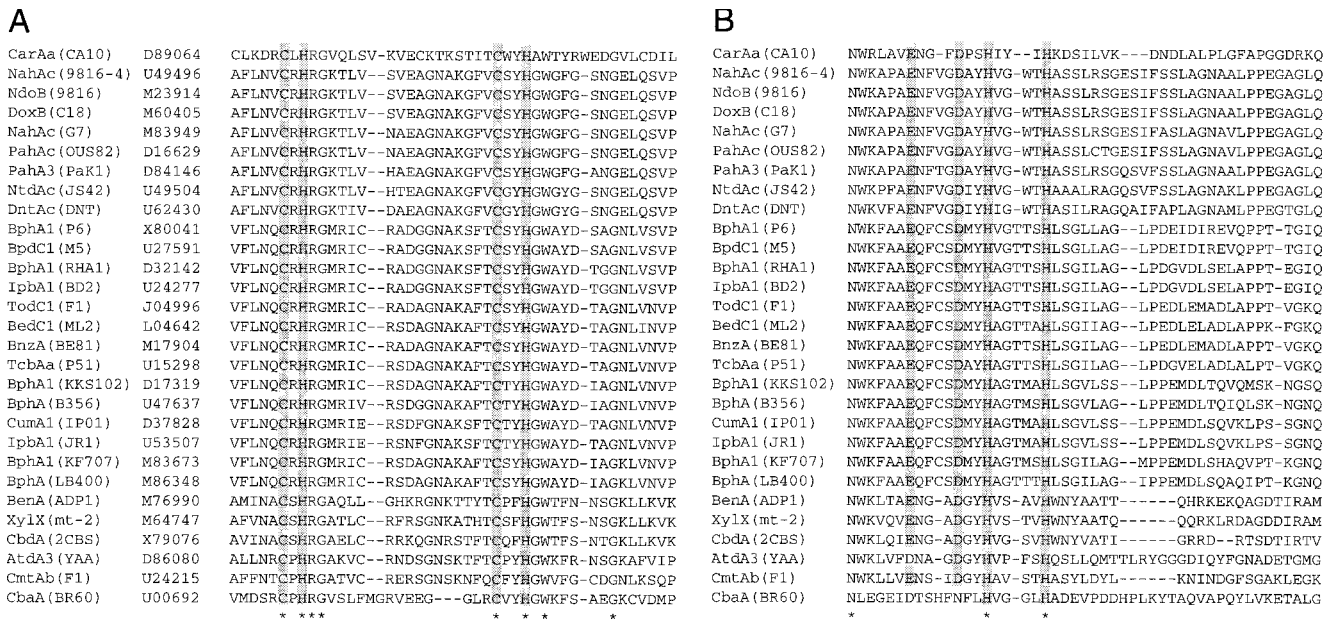

FIG. 4. Alignments of the large subunits of various terminal dioxygenase components with the postulated *carAa* gene products. Sequences of proteins that are homologous to the consensus sequence for the Rieske-type [2Fe-2S] cluster binding site (A) and that are homologous to the potential mononuclear iron binding region (B) are shown. Amino acid residues identical in all proteins are indicated by asterisks. Gaps introduced into the alignment are indicated by hyphens. The amino acids conserved for the Rieske-type [2Fe-2S] cluster binding site (32) and mononuclear iron coordination site (22) are indicated by shaded boxes in panels A and B, respectively. The DDBJ, EMBL, and GenBank nucleotide sequence accession numbers are indicated after the names of enzymes and bacterial strains (in parentheses) in panel A.

found in ORF4 and ORF5, as shown in Fig. 4A. The critical motif $\mathbf{G}X_{3-4}\mathbf{D}X_2\mathbf{H}X_{4-5}\mathbf{H}$, which was considered to act as a mononuclear iron ligand at the site of oxygen activation (22), was also conserved in ORF4 and ORF5 (Fig. 4B). These results suggested the possibility that ORF4 and ORF5 encode the large subunit of the terminal oxygenase component of CARDO. To confirm this possibility, two deletion mutants were constructed. In resting cell reactions using the deletion mutants, the Δ1.3-kb *Eco*RI-*Sma*I mutant (pUCA111 [Fig. 2]), in which only ORF4 was deleted, could metabolize CAR to anthranilic acid. On the other hand, in the reaction mixture using resting cells of the Δ2.6-kb *Eco*RI-*Sma*I mutant (pUCA12 [Fig. 2]), in which both ORF4 and ORF5 were deleted, CAR was not metabolized at all (data not shown). These results indicated that the protein encoded by ORF4 and ORF5 is indispensable for the expression of CARDO activity and supported the possibility that ORF4 and ORF5 encode the subunit of the terminal oxygenase component of CARDO.

Comparison among the amino acid sequences of small (β) subunits of catalytic terminal oxygenases of various bacterial dioxygenase systems revealed that their phylogenetic relationships are similar to those found for large subunits, but with lower degrees of homology (3). No significant homology of ORF7 with other degradative genes was observed, but there was a possibility that ORF7 encodes a small subunit of the terminal oxygenase component. To examine this possibility, the following experiment was conducted. A 2.3-kb *Bgl*II-*Eco*RI fragment carrying ORF6 to ORF8 was inserted into pSTV29, the resultant plasmid being designated pSCA14. Frameshift mutation in ORF7 was introduced in pSCA14 at the *Xho*I site by *Xho*I digestion, blunting the ends with T4 DNA polymerase, and ligation at the blunt-ended *Xho*I restriction site, the resultant plasmid being designated pSCA14X. We also constructed a plasmid in which the 3.2-kb *Eco*RI-*Xho*I fragment carrying ORF4 and ORF5 was inserted into pUC119, and it was named pUCA15 (Fig. 2). The mixture of pUCA15 and pSCA14 or that of pUCA15 and pSCA14X was introduced into *E. coli* JM109, and the CARDO activities of the recombinant *E. coli* cells were assayed by the resting cell reactions. As the result of GC-MS analysis, the same metabolite, 2′-aminobiphenyl-2,3-diol, was identified in both reaction mixtures (data not shown). No difference was observed in the yield of 2′-aminobiphenyl-2,3-diol between the clones having an in-

TABLE 1. Substrate specificity of CARDO

Substrate	R_f value[a]		Substrate remaining (%)
	Substrates	Products	
Carbazole	0.74	0.56	1.27
Dibenzofuran	0.92	0.38	48.5
Dibenzothiophene	0.86	0.52	NT[c]
Fluorene	0.93	0.81, 0.68, 0.63	80.8
N-Methylcarbazole	0.92	0.37	NT
N-Ethylcarbazole	0.92	0.38	NT
Dibenzo-*p*-dioxin	0.94	0.42	3.95
Phenazine	0.70	ND[b]	NT
Phenoxazine	0.75	0.56	1.09
Phenothiazine	0.80	0.51	7.62
Phenoxathiin	0.92	0.43	25.0
Xanthene	0.92	0.41	52.8
Biphenyl	0.95	0.43	0.00
Naphthalene	0.93	0.37	23.6
Phenanthrene	0.90	0.68, 0.42	5.43
Anthracene	0.92	0.70, 0.44	8.75
Fluoranthene	0.89	0.73, 0.44	NT

[a] Analytical TLC was developed with the solvent system toluene-dioxane-acetic acid (90:25:4, by volume).
[b] ND, not detected.
[c] NT, not tested.

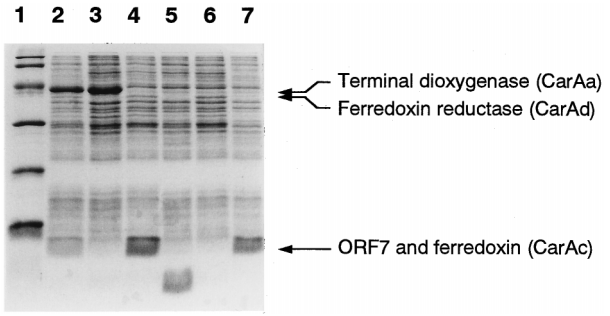

FIG. 5. Detection of the products of *carAa*, *carAc*, *carAd*, and ORF7. Total cellular proteins of *E. coli* strains were analyzed by SDS-PAGE. Lanes: 1, molecular mass standards of 94, 67, 43, 30, 20.1, and 14.4 kDa (top to bottom); 2, JM109(pUCARA); 3, JM109(pUCA152); 4, JM109(pUCA14); 5, JM109 (pUCA141); 6, JM109(pUCA142); 7, JM109(pUCA144).

tact ORF7 and frameshift mutation-introduced ORF7. These results suggest that ORF7 is not indispensable to the CARDO activity and that the presence of ORF4 (or ORF5) with ferredoxin and ferredoxin reductase components is essential for the expression of CARDO activity. Thus, we concluded that ORF7 does not encode the small subunit of the terminal dioxygenase component of CARDO and that CarAa acts as a subunit of terminal oxygenase of CARDO; that is, this terminal oxygenase component consists of a single subunit. Thus, both ORF4 and ORF5 were designated *carAa*.

Substrate specificity of CARDO. To analyze the substrate specificity of CARDO, we constructed a recombinant plasmid which encoded *carAaAcAd* but did not encode the *carBaBb* and *carC*. The plasmid vector pUC119 was digested with restriction endonuclease *Pst*I and blunted with T4 DNA polymerase, and the resultant linear DNA fragment was self-ligated. The 6.9-kb *Eco*RI insert of pUCA1 was reinserted into this pUC119 derivative. The resultant plasmid was digested with *Pst*I and *Bgl*II, and the ends of the resultant DNA fragment were blunted. The blunt-ended restriction sites were ligated, and the resultant plasmid was designated pUCARA.

By means of a resting cell reaction using *E. coli* JM109 harboring pUCARA, the conversion of other aromatic compounds by CARDO was investigated. The resultant products were analyzed by TLC. *E. coli* JM109 harboring pUC119 was used as a control. The R_f values of the aromatic compounds tested in this study were determined (Table 1). *E. coli* JM109 harboring pUCARA could oxidize many aromatic compounds but not phenazine. The R_f values of all of the products were lower than those of the corresponding substrates, suggesting that all substrates were oxidized. Thus, CARDO was found to have a broad substrate range. In the case of DD, DBF, phenoxazine, phenothiazine, and xanthene, the detected metabolites clearly turned brown on TLC plates in a few days, suggesting the possibility that the metabolites were easily oxidized with air. The mass spectrum of the product from DD revealed a molecular ion peak at *m/z* 218 (M^+, 76) and prominent fragment ion peaks at *m/z* 199 ($M^+ - 19$, 5), 171 ($M^+ - 47$, 11), 125 ($M^+ - 93$, 13), 110 ($M^+ - 108$, 74), and 94 ($M^+ - 124$, 100). ^{1}H NMR data (500 MHz, CD_3OD/C_6D_6 = 100:15, by volume; tetramethylsilane as an internal standard) were as follows: δ = 6.35 (H-4, $J_{4,5}$ = 8.0 Hz, $J_{4,6}$ = 2.0 Hz), 6.60 (H-5, $J_{5,6}$ = 8.0 Hz), 6.63 (H-6), 6.74 (H-5′, $J_{5',6'}$ = 7.5 Hz), 6.82 (H-6′), 6.93 (H-4′, $J_{4',5'}$ = 8.0 Hz, $J_{4',6'}$ = 1.5 Hz), and 6.96 (H-3′, $J_{3',4'}$ = 1.5 Hz) ppm. Since these data were identical with those reported by Wittich et al. (42), it was determined that DD was converted to 2,2′,3-trihydroxydiphenyl ether by CARDO. This result indicates that DD was dihydroxylated by CARDO at the angular and adjacent carbon atoms and that the dihydroxylated intermediate was spontaneously converted to 2,2′,3-trihydroxydiphenyl ether. The mass spectrum of the product from DBF revealed a molecular ion peak at *m/z* 202 (M^+, 100), and prominent fragment ion peaks at *m/z* 174 ($M^+ - 28$, 15), 149 ($M^+ - 53$, 22), 131 ($M^+ - 71$, 17), 115 ($M^+ - 87$, 20), and 91 ($M^+ - 111$, 18). Since this mass fragmentation pattern was identical to that reported by Fortnagel et al. (12), this product from DBF was determined to be 2,2′,3-trihydroxybiphenyl. This result indicates that CARDO is also able to attack at the angular and its adjacent carbon atoms of DBF as is the case for DD.

Detection of the products of *carAa*, *carAc*, *carAd*, and ORF7. To confirm if the ORFs could be translated to proteins with the predicted sizes in *E. coli*, the cellular proteins from the recombinant *E. coli* cells harboring plasmids encoding the respective genes were loaded onto an SDS–15% polyacrylamide gel (Fig. 5). Expression of ORF4 encoded in a 1.3-kb *Dde*I insert of pUCA152 yielded a peptide with a molecular mass of 43 kDa, which corresponds to the predicted molecular mass of the *carAa* gene product (lane 3). Expression of ORF5 encoded in a 2.0-kb *Dde*I insert of pUCA153 also yielded the 43-kDa protein (data not shown). In a similar analysis of the *Bgl*II-*Eco*RI insert of pUCA14 encoding *carAc*, *carAd*, and ORF7, two proteins of 11 and 36 kDa were detected (lane 4). The 36-kDa protein was also expressed in the clone containing the *Xho*I-*Eco*RI insert of pUCA142 encoding *carAd* (lane 6). Since the calculated molecular masses of the product of ORF7 and CarAc are similar (11,366 and 11,551 Da, respectively), it remains to be clarified whether the 11-kDa protein is the product from ORF7, *carAc*, or a mixture of both. However, an 11-kDa protein was also expressed in the clone containing the 0.7-kb *Sph*I insert of pUCA144 encoding ORF7 (lane 7), and the 11-kDa protein corresponding to CarAc was not detected in the clone containing the *Bgl*II-*Xho*I insert of pUCA141 (lane 5). These results suggested that the band corresponding to the 11-kDa protein in lanes 2 and 4 was mostly due to the gene product of ORF7. The ca. 6-kDa protein in lane 5 is considered to be the N-terminal region of the product of ORF7.

DISCUSSION

In this study, we cloned the genes encoding CARDO, a member of a growing number of known bacterial enzymes that oxidize aromatic compounds. CARDO is a multicomponent enzyme system similar to naphthalene dioxygenases (36) and toluene dioxygenases (43). However, CARDO is a unique dioxygenase that attacks at the angular position adjacent to the nitrogen atom in CAR. Though the purification of the dibenzofuran 4,4a-dioxygenase which degrade DD and DBF has been reported (5), the nucleotide sequence of the genes encoding this dioxygenase system was not reported. This is the first paper reporting the cloning of the genes encoding dioxygenase that attacks at the angular position adjacent to the hetero atoms in heterocyclic aromatic compounds. It should also be noted that the nucleotide and deduced amino acid sequences of CarAa are unique and that the homologies with large subunits of terminal oxygenase components of other previously known dioxygenases were only 27 to 30%.

The gene organization of *car* gene cluster has been revealed to be *carAaAaBaBbCAc*(ORF7)*Ad*, as shown in Fig. 2. In other bacteria, the components of dioxygenase such as *bph* (27, 37), *cum* (1), and *nah* (36) are tandemly linked, although the gene orders of the respective components are varied. The gene encoding the ferredoxin reductase component in the biphenyl

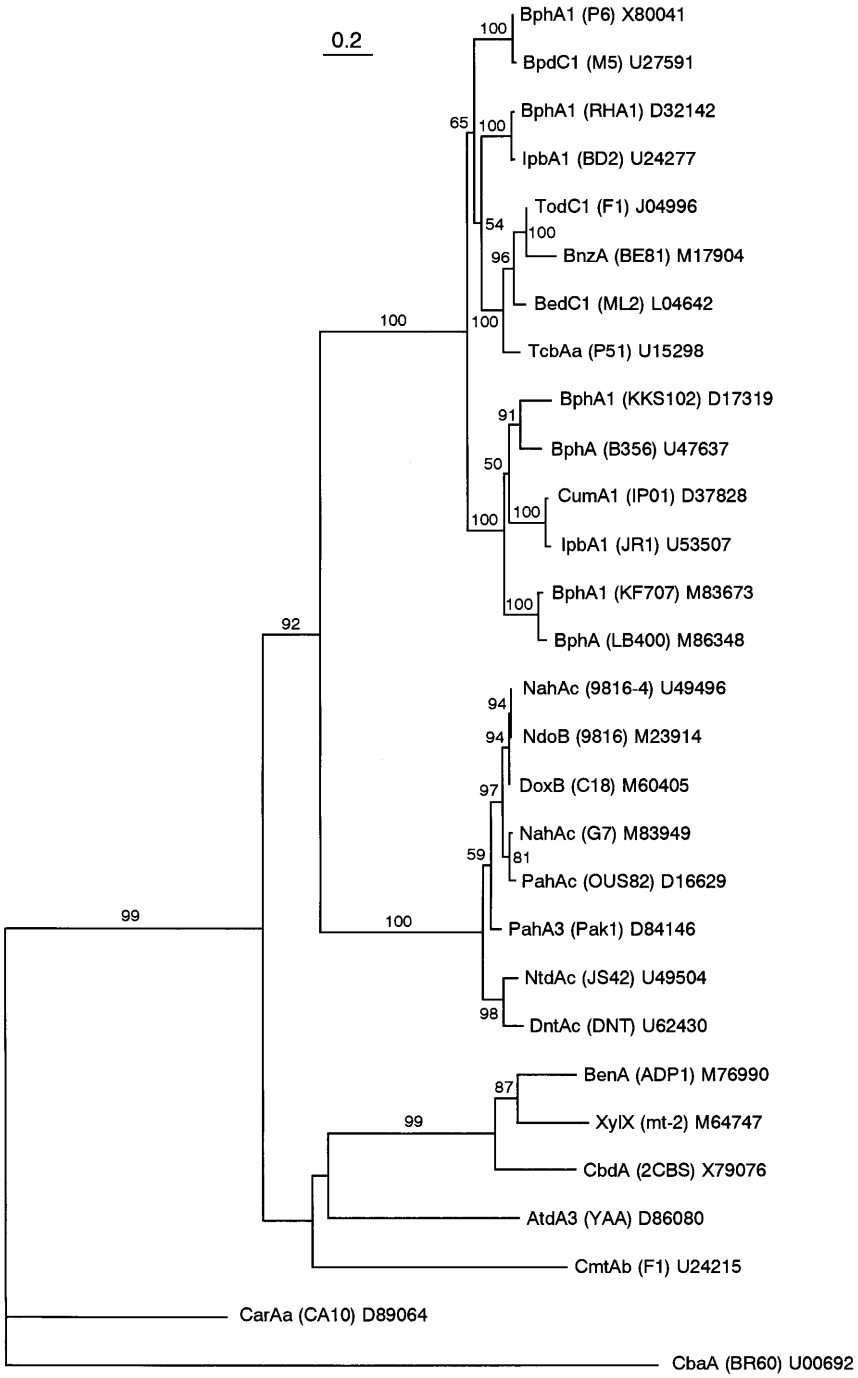

FIG. 6. Phylogenetic tree of large (α) subunits of terminal oxygenase components of multicomponent dioxygenase systems. The names of bacterial strains (in parentheses) and the DDBJ, EMBL, and GenBank accession numbers are indicated after the enzyme names. CbaA was used as an outgroup. The scale bar denotes 0.2 substitution per site. The tree was constructed by the neighbor-joining method, and bootstrap values above 50 from 100 resamplings are shown for each node.

dioxygenase in *Pseudomonas* sp. strain KKS102 is separated by a 3.5-kb DNA fragment from genes encoding the other components (23). The genes encoding the terminal oxygenase component (CarAa) of CARDO are also separated from the genes encoding the ferredoxin component (CarAc) and ferredoxin reductase component (CarAd) by a 2-kb DNA fragment encoding a *meta*-cleavage enzyme (CarB) and *meta*-cleavage compound hydrolase (CarC). The gene organization of *car* gene cluster suggests the possibility that gene rearrangements also occurred in the *car* gene cluster through the evolutionary divergence in the natural environment as suggested by Werlen et al. (41).

Since a homology search of DNA sequence databases showed that the 5′ terminus of the putative ORF located in the downstream region of the *carAd* gene is similar to the DNA sequence that encodes 2-hydroxypenta-2,4-dienoate hydratase of *Pseudomonas* sp. strain LB400 (21), this putative ORF is tentatively designated *carD*. On the basis of the gene organi-

zation in other degradative gene clusters (15, 21), it is quite possible that the genes involved in the *meta*-cleavage pathway (*carD*, *carE*, and *carF*) are encoded in the downstream region of *carAd*. On the other hand, the 3′ terminus of the putative ORF was found in the upstream region of *carAa*, but no significant homology was observed. To clarify the whole structure of *car* gene cluster, further gene walking is now under way.

It was revealed that the 1,263-bp DNA region containing the *carAa* gene was tandemly duplicated except for one base and that there were only three bases (GGC) between these tandemly linked DNA regions, as shown in Fig. 3. Such a gene structure has not been observed in other degradative gene clusters. Because CarAa was highly expressed in the recombinant *E. coli* cells harboring pUCA1 (duplicated *carAa* genes) or pUCA111 (single *carAa* gene), no significant difference in CARDO activity was observed between these clones. The significance of this duplication of *carAa* genes in strain CA10 cells remains to be clarified.

Expression of genes encoded in the *Eco*RI insert of pUCARA yielded 43-, 36-, and 11-kDa proteins, which are considered to correspond to *carAa*, *carAd*, and ORF7 gene products, respectively (Fig. 5, lane 2). As for CarAc, we could not clearly detect the gene product in *E. coli* harboring pUCARA, pUCA14, and pUCA141 (Fig. 5, lanes 2, 4, and 5, respectively). This might be due to weak expression as in the case of the *bphA3* gene product in *Pseudomonas pseudoalcaligenes* KF707 (37). The component of benzene dioxygenase of *P. putida* ML2 has intracellular α subunit of iron-sulfur protein/ferredoxin/reductase molar ratio of 1:0.45:0.8, suggesting that the multicomponent dioxygenase is an enzyme complex with a rather loose association between the oxygenase component, ferredoxin, and ferredoxin reductase (38). Although the expression level of CarAc in recombinant *E. coli* was low, the amount of CarAc expressed in *E. coli* might be sufficient to transport electrons to the terminal component, or CarAc might be more highly expressed in strain CA10 and able to transport electrons efficiently.

The terminal oxygenase components of bacterial aromatic ring dioxygenase systems reported previously consist of one or two subunits. The terminal oxygenase components of biphenyl dioxygenase (27, 37), naphthalene dioxygenase (36), and toluene dioxygenase (43) were reported to consist of large and small subunits. The recombinant *E. coli* cells harboring the plasmid encoding *carAa*, *carAc*, ORF7, and *carAd* had CARDO activity. The frameshift mutation in ORF7 did not affect the CARDO activity; therefore, ORF7 is not considered to be involved in the expression of CARDO activity. CarAc and CarAd were found to be ferredoxin and ferredoxin reductase components by homology search and alignment analyses. Based on the foregoing data, the terminal oxygenase component of CARDO was determined to consist of a single subunit (CarAa). The phylogenetic tree of the large (α) subunits of terminal oxygenases showed that CARDO was an evolutionarily novel dioxygenase (Fig. 6). The aromatic ring dioxygenases were classified into three classes in terms of the number of constituent components and the nature of the redox centers (28). Class I consists of two-component enzymes in which the first protein is a reductase containing both a flavin and a [2Fe-2S] redox center, and the second component is the terminal oxygenase; class II consists of three-component enzymes in which the flavin and [2Fe-2S] redox centers of the reductase are on the separate flavoprotein and ferredoxin, respectively; and class III consists of three-component enzymes in which the reductase contains both a flavin and a [2Fe-2S] redox center but also requires a second [2Fe-2S] center on a ferredoxin for electron transfer to the terminal oxygenase. Further subdivision is based on the type of flavin (flavin mononucleotide or FAD) in the reductase, the coordination of the [2Fe-2S] center in the ferredoxin, and the number of terminal oxygenase subunits. The consensus sequence of Rieske-type iron-sulfur proteins for the binding of a [2Fe-2S] cluster was found in the deduced amino acid sequences of the terminal oxygenase (CarAa) and ferredoxin (CarAc) of CARDO (Fig. 3 and 4A). The ferredoxin reductase (CarAd) of CARDO possesses the consensus sequence of the plant-type iron-sulfur proteins for the binding of a [2Fe-2S] cluster and the putative FAD- and NAD-binding site (Fig. 3B). From the conserved deduced amino acid sequence, CARDO seems to have a class III electron transport chain like that of naphthalene dioxygenase (36). It has been reported that naphthalene dioxygenase consists of three components, ferredoxin reductase, ferredoxin, and terminal oxygenase, consisting of large and small subunits (36). Because the terminal dioxygenase of CARDO consists of a single subunit, it is considered that CARDO constitutes a new subfamily in this class III dioxygenase family.

In this study, it was shown that CARDO attacked at the angular position adjacent to the hetero atom of heterocyclic aromatic compounds such as CAR, DD, and DBF. This type of dioxygenation was followed by the spontaneous cleavage of the bond between the hetero atom and the adjacent angular position. It was also shown that CARDO had the ability to metabolize biphenyl and polycyclic aromatic hydrocarbons, such as naphthalene, phenanthrene. Identification of the products from these various types of aromatic compounds by CARDO will provide information on the manner of dioxygenation catalyzed by this novel dioxygenase.

Although strain CA10 did not grow on DD, DBF, fluorene, or naphthalene (31), resting cells of strain CA10 were able to convert these compounds (data not shown). However, the ability of the recombinant *E. coli* harboring pUCARA to convert these compounds was much higher than that of strain CA10 (data not shown). Because DD and DBF were converted by CARDO to biphenyl and diphenyl ether derivatives which are easily degraded by many bacteria, the transformants harboring pUCARA were useful tools for remediation of DD and DBF in a closed system. In a preliminary experiment, CARDO could degrade polychlorinated dibenzo-*p*-dioxin (data not shown). Therefore, it is considered possible to remediate sites contaminated with this compound and/or polychlorinated dibenzofuran by using this recombinant *E. coli* strain along with derivative metabolite-degrading microorganisms, such as polychlorinated biphenyl degraders. Considering that CARDO has broad substrate specificity, this recombinant *E. coli* could become a very useful tool for bioremediation of various aromatic compound-contaminated sites.

While the small subunits of terminal oxygenase have been implicated in substrate recognition (16), Erickson and Mondello (11) reported the large subunit of biphenyl dioxygenase to be responsible for the substrate specificity. The terminal oxygenase component of CARDO consists of a single subunit of CarAa protein, suggesting that the substrate recognition site is located in CarAa, which corresponds to the large subunits of other aromatic ring dioxygenases. Crystallization and determination of the three-dimensional structure of CarAa will provide valuable information on the substrate recognition site and the mechanism of dioxygenation at the angular positions adjacent to the hetero atoms in CAR, DD, DBF, and so on.

REFERENCES

1. **Aoki, H., T. Kimura, H. Habe, H. Yamane, T. Kodama, and T. Omori.** 1996. Cloning, nucleotide sequence, and characterization of the genes encoding enzymes involved in the degradation of cumene to 2-hydroxy-6-oxo-7-meth-

ylocta-2,4-dienoic acid in *Pseudomonas fluorescens* IP01. J. Ferment. Bioeng. **81:**187–196.
2. **Arcos, J. S., and M. F. Argus.** 1968. Molecular geometry and carcinogenic activity of aromatic compounds. New perspectives. Adv. Cancer Res. **11:**305–471.
3. **Asturias, J. A., E. Diaz, and K. N. Timmis.** 1995. The evolutionary relationship of biphenyl dioxygenase from Gram-positive *Rhodococcus globerulus* P6 to multicomponent dioxygenases from Gram-negative bacteria. Gene **156:** 11–18.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
5. **Bünz, P. V., and A. M. Cook.** 1993. Dibenzofuran 4,4a-dioxygenase from *Sphingomonas* sp. strain RW1: angular dioxygenation by a three-component enzyme system. J. Bacteriol. **175:**6467–6475.
6. **Byrne, A. M., J. J. Kukor, and R. H. Olsen.** 1995. Sequence analysis of the gene cluster encoding toluene-3-monooxygenase from *Pseudomonas pickettii* PKO1. Gene **154:**65–70.
7. **Chemicals Inspection and Testing Institute Japan.** 1992. Biodegradation and bioaccumulation: data of existing chemicals based on the CSCL, Japan. Japan Chemical Industry Ecology-Toxicology and Information Center, Tokyo, Japan.
8. **Cerniglia, C. E., J. C. Morgan, and D. T. Gibson.** 1979. Bacterial and fungal oxidation of dibenzofuran. Biochem. J. **180:**175–185.
9. **Eaton, R. W., and K. N. Timmis.** 1986. Characterization of a plasmid-specific pathway for catabolism of isopropylbenzene in *Pseudomonas putida* RE204. J. Bacteriol. **168:**123–131.
10. **Ensley, B. D., B. J. Ratzkin, T. D. Osslund, M. J. Simon, L. P. Wackett, and D. T. Gibson.** 1983. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. Science **222:**167–169.
11. **Erickson, B. D., and F. J. Mondello.** 1993. Enhanced biodegradation of polychlorinated biphenyls after site-directed mutagenesis of a biphenyl dioxygenase gene. Appl. Environ. Microbiol. **59:**3858–3862.
12. **Fortnagel, P., H. Harms, R.-M. Wittich, S. Krohn, H. Meyer, V. Sinnwell, H. Wilkes, and W. Francke.** 1990. Metabolism of dibenzofuran by *Pseudomonas* sp. strain HH69 and the mixed culture HH27. Appl. Environ. Microbiol. **56:**1148–1156.
13. **Gieg, L. M., A. Otter, and P. M. Fedorak.** 1996. Carbazole degradation by *Pseudomonas* sp. LD2: metabolic characteristics and the identification of some metabolites. Environ. Sci. Technol. **30:**575–585.
14. **Grosser, R. J., D. Warshawsky, and J. R. Vestal.** 1991. Indigenous and enhanced mineralization of pyrene, benzo[*a*]pyrene, and carbazole in soils. Appl. Environ. Microbiol. **57:**3462–3469.
15. **Habe, H., T. Kimura, H. Nojiri, H. Yamane, and T. Omori.** 1996. Cloning and nucleotide sequences of the genes involved in the *meta*-cleavage pathway of cumene degradation in *Pseudomonas fluorescens* IP01. J. Ferment. Bioeng. **81:**247–254.
16. **Harayama, S., M. Rekik, and K. N. Timmis.** 1986. Genetic analysis of a relaxed substrate specificity aromatic ring dioxygenase, toluene 1,2-dioxygenase encoded by TOL plasmid pWW0 of *Pseudomonas putida*. Mol. Gen. Genet. **202:**226–234.
17. **Harayama, S., and M. Kok.** 1992. Functional and evolutionary relationships among diverse oxygenases. Annu. Rev. Microbiol. **46:**565–601.
18. **Harms, H., R.-H. Wittich, V. Sinnwell, H. Meyer, P. Fortnagel, and W. Francke.** 1990. Transformation of dibenzo-*p*-dioxin by *Pseudomonas* sp. strain HH69. Appl. Environ. Microbiol. **56:**1157–1159.
19. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
20. **Hisatsuka, K., and M. Sato.** 1994. Microbiol transformation of carbazole to anthranilic acid by *Pseudomonas stutzeri*. Biosci. Biotechnol. Biochem. **58:** 213–214.
21. **Hofer, B., S. Backhaus, and K. N. Timmis.** 1994. The biphenyl/polychlorinated biphenyl-degradation locus (*bph*) of *Pseudomonas* sp. LB400 encodes four additional metabolic enzymes. Gene **144:**9–16.
22. **Jiang, H., R. E. Parales, N. A. Lynch, and D. T. Gibson.** 1996. Site-directed mutagenesis of conserved amino acids in the alpha subunit of toluene dioxygenase: potential mononuclear non-heme iron coordination sites. J. Bacteriol. **178:**3133–3139.
23. **Kikuchi, Y., Y. Nagata, M. Hinata, K. Kimbara, M. Fukuda, K. Yano, and M. Takagi.** 1994. Identification of the *bphA4* gene encoding ferredoxin reductase involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. strain KKS102. J. Bacteriol. **176:**1689–1694.
24. **Kobayashi, T., R. Kurane, K. Nakajima, Y. Nakamura, K. Kirimura, and S. Usami.** 1995. Isolation of bacteria degrading carbazole under microaerobic condition, i.e., nitrogen gas substituted conditions. Biosci. Biotechnol. Biochem. **59:**932–933.
25. **Kurkela, S., H. Lehväslaiho, E. T. Palva, and T. H. Teeri.** 1988. Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB9816. Gene **73:**355–362.
26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
27. **Masai, E., A. Yamada, J. M. Healy, T. Hatta, K. Kimbara, M. Fukuda, and K. Yano.** 1995. Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **61:**2079–2085.
28. **Mason, J. R., and R. Commack.** 1992. The electron-transport proteins of hydroxylating bacterial dioxygenases. Annu. Rev. Microbiol. **46:**277–305.
29. **Neidle, E. L., C. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama.** 1991. Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. J. Bacteriol. **173:**5385–5395.
30. **Otaka, E., and T. Ooi.** 1989. Examination of protein sequence homologies. 5. New perspectives on evolution between bacterial and chloroplast-type ferredoxins inferred from sequence evidence. J. Mol. Evol. **29:**246–254.
31. **Ouchiyama, N., Y. Zhang, T. Omori, and T. Kodama.** 1993. Biodegradation of carbazole by *Pseudomonas* spp. CA06 and CA10. Biosci. Biotechnol. Biochem. **57:**455–460.
32. **Rieske, J. S., D. H. Maclennan, and R. Coleman.** 1964. Isolation and properties of an iron-protein from the (reduced coenzyme Q)-cytochrome C reductase complex of the respiratory chain. Biochem. Biophys. Res. Commun. **15:**338–344.
33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Sato, S. I., N. Ouchiyama, T. Kimura, H. Nojiri, H. Yamane, and T. Omori.** 1997. Cloning of genes involved in carbazole degradation of *Pseudomonas* sp. strain CA10: nucleotide sequences of genes and characterization of *meta*-cleavage enzymes and hydrolase. J. Bacteriol. **179:**4841–4849.
35. **Shine, J., and L. Dalgarno.** 1975. Determination of cistron specificity in bacterial ribosomes. Nature **254:**34–38.
36. **Simon, M. J., T. D. Osslund, R. Saunders, B. D. Ensley, S. Suggs, A. Harcourt, W.-C. Suen, D. L. Cruden, D. T. Gibson, and G. J. Zylstra.** 1993. Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strain G7 and NCIB 9616-4. Gene **127:**31–37.
37. **Taira, K., J. Hirose, S. Hayashida, and K. Furukawa.** 1992. Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. J. Biol. Chem. **267:**4844–4853.
38. **Tan, H.-M., C. L. Joannou, C. E. Cooper, C. S. Butler, R. Cammack, and J. R. Mason.** 1994. The effect of ferredoxin$_{BED}$ overexpression on benzene dioxygenase activity in *Pseudomonas putida* ML2. J. Bacteriol. **176:**2507–2512.
39. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
40. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
41. **Werlen, C., H. E. Kohler, and J. R. van der Meer.** 1996. The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. J. Biol. Chem. **271:**4009–4016.
42. **Wittich, R.-M., H. Wilkes, V. Sinnwell, W. Francke, and P. Fortnagel.** 1992. Metabolism of dibenzo-*p*-dioxin by *Sphingomonas* sp. strain RW1. Appl. Environ. Microbiol. **58:**1005–1010.
43. **Zylstra, G. J., and D. T. Gibson.** 1989. Toluene degradation by *Pseudomonas putida* F1: nucleotide sequence of the *todC1C2BADE* genes and their expression in *Escherichia coli*. J. Biol. Chem. **264:**14940–14946.