Aromatic compounds which carry sulfonic acid substituents are generally considered xenobiotic compounds which are rather recalcitrant against microbial degradation (1, 57). Nevertheless, several bacterial strains which utilize sulfonated benzenes or naphthalenes as sole sources of carbon and energy have been isolated (5, 7, 9, 31, 34, 35, 36, 55, 58, 59). Thus, it was shown that a coculture of *Hydrogenophaga* sp. strain S1 and *Agrobacterium radiobacter* strain S2 mineralized 4-aminobenzenesulfonate (sulfanilate) in a syntrophic association via 4-sulfocatechol (4SC), which was oxidized in both strains by an ortho-cleavage mechanism to 3-sulfumonocatechol (10, 17, 18) (Fig. 1). The 4SC-converting dioxygenase was purified to homogeneity from *A. radiobacter* strain S2, and it was shown that the purified enzyme also converted protocatechuate. The corresponding enzyme activity was therefore tentatively termed the protocatechuate 3,4-dioxygenase type II (P34O-II) (18, 23). After growth with 4SC, strain S2 also induced a classical P34O (P34O-I) which oxidized protocatechuate but not 4-sulfocatechol. These genes were part of a protocatechuate-degradative operon which strongly resembled the isofunctional operon from the protocatechuate-degrading strain *Agrobacterium tumefaciens* A348 described previously by D. Parke (FEMS Microbiol. Lett. 146:3–12, 1997). The second P34O (P34O-II), encoded by the pcaH2G2 genes, was functionally expressed and shown to convert protocatechuate and 4-sulfocatechol. A comparison of the deduced amino acid sequences of PcaH-I and PcaH-II, and of PcaG-I and PcaG-II, with each other and with the corresponding sequences from the P34Os, from other bacterial genera suggested that the genes for the P34O-II were obtained by strain S2 by lateral gene transfer. The genes encoding the P34O-II were found in a putative operon together with two genes which, according to sequence alignments, encoded transport proteins. Further downstream from this putative operon, two open reading frames which code for a putative regulator protein of the IcR family and a putative 3-carboxymuconate cycloisomerase were identified.

**Characterization of the Genes for Two Protocatechuate 3,4-Dioxygenases from the 4-Sulfocatechol-Degrading Bacterium *Agrobacterium radiobacter* Strain S2**

MATTHIAS CONTZEN AND ANDREAS STOLZ*

Institut für Mikrobiologie, Universität Stuttgart, 70569 Stuttgart, Germany

Received 24 April 2000/Accepted 4 July 2000

The genes for two different protocatechuate 3,4-dioxygenases (P34Os) were cloned from the 4-sulfocatechol-degrading bacterium *Agrobacterium radiobacter* strain S2 (DSMZ 5681). The *pcaHI* genes encoded a P34O (P34O-I) which oxidized protocatechuate but not 4-sulfocatechol. These genes were part of a protocatechuate-degradative operon which strongly resembled the isofunctional operon from the protocatechuate-degrading strain *Agrobacterium tumefaciens* A348 described previously by D. Parke (FEMS Microbiol. Lett. 146:3–12, 1997). The second P34O (P34O-II), encoded by the *pcaH2G2* genes, was functionally expressed and shown to convert protocatechuate and 4-sulfocatechol. A comparison of the deduced amino acid sequences of PcaH-I and PcaH-II, and of PcaG-I and PcaG-II, with each other and with the corresponding sequences from the P34Os, from other bacterial genera suggested that the genes for the P34O-II were obtained by strain S2 by lateral gene transfer. The genes encoding the P34O-II were found in a putative operon together with two genes which, according to sequence alignments, encoded transport proteins. Further downstream from this putative operon, two open reading frames which code for a putative regulator protein of the IcR family and a putative 3-carboxymuconate cycloisomerase were identified.

Aromatic compounds which carry sulfonic acid substituents are generally considered xenobiotic compounds which are rather recalcitrant against microbial degradation (1, 57). Nevertheless, several bacterial strains which utilize sulfonated benzenes or naphthalenes as sole sources of carbon and energy have been isolated (5, 7, 9, 31, 34, 35, 36, 55, 58, 59). Thus, it was shown that a coculture of *Hydrogenophaga* sp. strain S1 and *Agrobacterium radiobacter* strain S2 mineralized 4-aminobenzenesulfonate (sulfanilate) in a syntrophic association via 4-sulfocatechol (4SC), which was oxidized in both strains by an ortho-cleavage mechanism to 3-sulfumonocatechol (10, 17, 18) (Fig. 1). The 4SC-converting dioxygenase was purified to homogeneity from *A. radiobacter* strain S2, and it was shown that the purified enzyme also converted protocatechuate. The corresponding enzyme activity was therefore tentatively termed the protocatechuate 3,4-dioxygenase type II (P34O-II) (18, 23). After growth with 4SC, strain S2 also induced a classical P34O which did not oxidize 4SC.

P34Os, (protocatechuate:oxygen 3,4-oxidoreductase; EC 1.13.11.3) are of central importance in the biodegradation of many aromatic compounds by bacteria (25). The enzymes catalyze the intradiol addition of molecular oxygen, cleaving the aromatic ring and forming 3-carboxy-

© Copyright 2000, American Society for Microbiology. All Rights Reserved.
ments obtained with T4 DNA ligase. Thus, intramolecular ligation of these DNA
pcaG1, a partial inverse PCR was performed (41). The template was prepared by
pBluescript II SK(+) (33).

Partial inverse PCR. For the determination of the complete sequence of
pcaG1, a partial inverse PCR was performed (41). The template was prepared by
digesting chromosomal DNA of strain S2 with HindIII and religating the frag-
ments obtained with T4 DNA ligase. Thus, intramolecular ligation of these DNA
fragments should result in circular DNA, which was then used as a template for
the following PCR. Primers for this PCR were deduced from the partial se-
quence of pcaG1 present on the 11-kb PstI fragment previously obtained and
were facing in both directions outwards from the known DNA sequence. This
resulted in the amplification of an approximately 2-kb fragment, which was used
to finally obtain an approximately 2.5-kb EcoRI fragment of chromosomal DNA
after Southern hybridization. This fragment was cloned and finally sequenced.

Hybridization procedures. A digoxigenin DNA labeling and detection kit was
used according to the instructions of the supplier (Boehringer). The hybridiza-
tion temperature was set to 63°C.

Nucleotide sequence analysis. The DNA sequence was determined by dideoxy-
chain termination with double-stranded DNAs of overlapping subclones in an
automated DNA-sequencing system (ALF-Sequencer; Amersham-Pharmacia)
with fluorescently labeled primers.

Sequence analysis. database searches, and comparisons were done with the
PCGene software package (release 6.85) and the BLAST Search at the National
Center for Biotechnology Information (BLASTX) (3). The alignments of the
P34Os were obtained with the program CLUSTAL using the default parameters.

Expression of P34O-I and P34O-II in E. coli. For expression in E. coli,
pcaHGI and pcaHGI2 were inserted into pET11a (54) under the control of the
phase T7 promoter. The DNA segments encompassing pcaHGI or pcaHGI2
were amplified by PCR with simultaneous introduction of an Ndel site upstream
and a BamHI site downstream of pcaHGI or pcaHGI2. The oligonucleotide
primers were 5′-ACGC-CATATG-AGCAACCAGCCACCGA-3′ and 5′-ATAA-G
GATCC-CTCATGCGGTCGTATAC-3′ were used for the amplification of
pcaHGI. The primers used for the amplification of pcaHGI2 were 5′-ATTT-
-CATATG-GCCTTGTTCCTCCCCGGG-3′ and 5′-ATAAT-GATCC-CTCAC
AATTACCTTGCGC-3′. The amplified products were cleaved with Ndel and
BamHI and ligated into pET11a. E. coli DH5α was transformed with the resulting
plasmids. The plasmids were subsequently isolated and introduced into E. coli
BL21(DE3) pLysS by transformation.

Preparation of cell extracts. Cell suspensions in 50 mM Tris-HCl buffer (pH
8.0) were disrupted by using a French press (Amicon; SLM Instruments Inc.,
Urbana, Ill.) at 1.1 × 108 Pa. Cells and cell debris were removed by centrifuga-
tion at 100,000 × g for 30 min at 4°C.

Protein estimation and enzyme assays. The protein content of cell extracts
was determined by the method of Bradford (6). Bovine serum albumin was used as
a standard. One unit of enzyme activity is defined as the amount of enzyme that
converts 1 μmol of substrate per min. P34O-I enzyme was measured by the
procedure of Stanier and Ingraham (52) but using Tris-HCl buffer (50 mM, pH
8.0). For the assay of P34O-II activity, protocatechuate was replaced by 4SC as
substrate (18).

Sequence analysis, database searches, and comparisons were done with the
protein or peptide Amino acid sequence Deduced primer sequence

<table>
<thead>
<tr>
<th>Protein or peptide</th>
<th>Amino acid sequence</th>
<th>Deduced primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>P34O type I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N terminus of PcaG-I</td>
<td>TPSQTGYGP</td>
<td>3′-TGG GGS AGM GTY TGG CAS CGG GGS(5)′−5′</td>
</tr>
<tr>
<td>PcaH-I</td>
<td>TIKPPFYYPW</td>
<td>5′-ACS ATY AAR CCS GGY CCR TAY CCS TGG-3′</td>
</tr>
<tr>
<td>P34O type II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N terminus of PcaG-II</td>
<td>TVGGFYAY</td>
<td>3′-TGN CAN CCN GGN AAR ATR CGN TA(R)-5′</td>
</tr>
<tr>
<td>N terminus of PcaH-II</td>
<td>EVPAEYP</td>
<td>5′-GGN GTN CCN GGN GAR TAY CCN-3′</td>
</tr>
</tbody>
</table>

a The oligonucleotides were deduced from the consensus sequences of different P34Os. The amino acid sequence for PcaH-I corresponded to amino acids 9 to 16, and that of PcaG-I corresponded to amino acids 141 to 149, of the respective subunits of the P34O from P. putida (21).
b The oligonucleotides were deduced from the amino-terminal amino acid sequences of PcaG-II (amino acids 14 to 21) and PcaH-II (amino acids 9 to 15) from A. radiobacter strain S2 (23).

RESULTS

Cloning of the genes for the P34O-I. Previous biochemical studies suggested that A. radiobacter strain S2 synthesized a classical P34O (P34O-I), with the ability to oxidize protocate-
chuate but not 4SC, and also a second, modified type of P34O (P34O-II) which converted protocatechuate and 4SC. For both enzymes the amino-terminal amino acid sequences of the sub-
units (PcaH and PcaG) had been determined (23). The se-
quencies for P34O-I served for the design of oligo-
nucleotide primers, but no specific amplification was
obtained. Therefore, the PCR experiments were repeated using oligonucleotide primers from amino-terminal amino acid
sequences and conserved regions within the iron-binding sites of other known P34Os (Table 1). This resulted in the amplification of a fragment (409 bp) which showed significant sequence homology to other P34Os. The PCR fragment was labeled and used to clone an 11-kb PstI fragment from the genomic DNA of strain S2 in pBluescript II SK(+) after Southern hybridization, resulting in plasmid pMC171 (Fig. 2).

The genes for a P34O (pcaHG) had previously been cloned from *Agrobacterium tumefaciens* A348 (43). A 405-bp EcoRI-EcoRV fragment was obtained from plasmid pARO162, encoding parts of *pcaHG* from *A. tumefaciens* A348. This fragment was also labeled and shown to hybridize with the 11-kb PstI fragment from strain S2 cloned in pMC171.

**Determination of the nucleotide sequence of pcaH1G1 from strain S2.** DNA sequencing showed that approximately 200 bp of *pcaG1* was missing at the 3′ end of the 11-kb PstI fragment cloned in pMC171. The missing part of *pcaG1* was obtained after partial inverse PCR (41), resulting in plasmid pS2-I-39B. After sequencing of both inserts, the complete sequence of *pcaH1G1* from *A. radiobacter* strain S2 was obtained (Fig. 2). A comparison of the deduced amino acid sequences of PcaH-I and PcaG-I from strain S2 with the sequences submitted recently by Parke to the National Center for Biotechnology Information data bank (accession number U32867) for the homologous proteins from *A. tumefaciens* A348 demonstrated 96.3 and 96.6% sequence identity, respectively.

The amino-terminal sequences of PcaH-I and PcaG-I from strain S2 which were deduced from the nucleotide sequences showed considerable differences from the sequences published previously by Hammer et al. (23) after amino acid sequencing. To test the identity of *pcaH1G1* with the genes encoding the P34O-I, *pcaH1G1* were functionally expressed in *E. coli* using a phage T7-promoter system. Cultures of *E. coli* BL21(DE3)/pLysS with the expression plasmid were grown at 37°C in 200 ml of Luria-Bertani medium plus 50 μg of ampicillin per ml and 10 mg of FeCl3 per ml to an optical density (510 nm) of about 1. The formation of the P34O was then induced by the addition of 0.4 mM IPTG (isopropyl-β-D-thiogalactopyranoside). The cultures were further incubated at 30°C on a rotary shaker. Cells were harvested by centrifugation 90 min after the addition of IPTG, and the cells were resuspended in 1.8 ml of Tris-HCl buffer (pH 8, 50 mM) and disrupted with a French press. Enzyme activities were measured spectrophotometrically with protocatechuate and 4SC as described in Materials and Methods. The extract oxidized protocatechuate (specific activity, 0.09 U/mg of protein) but not 4SC (<0.001 U/mg of protein). It was therefore concluded that indeed the genes encoding the P34O-I had been cloned and sequenced. The recombinant *E. coli* strain expressed only low specific activities of P34O-I. This has been repeatedly observed for recombinant P34Os from various sources (14, 21, 46).

**Analysis of the gene cluster with pcaH1G1.** A gene cluster including *pcaHG* had been previously described for *A. tumefaciens*, which differed in several aspects from the pathways for protocatechuate degradation in the well-studied gram-negative bacteria *Acinetobacter calcoaceticus* and *Pseudomonas putida* (45). Therefore, the organization of the *pca* genes in *A. radiobacter* strain S2 was determined by sequencing the 11-kb DNA fragment (partially only by single-strand sequencing), and it was found that the gene order within strain S2 appears to be identical to the one proposed by Parke (45) for *A. tumefaciens*. Furthermore, a comparison of the sequence of PcaQ determined by Parke (42, 44) with that of the corresponding gene product from strain S2 showed 91% sequence identity.

**Cloning of the genes for P34O-II.** PCR using primers (Table 1) derived from the amino-terminal parts of PcaH-II and PcaG-II determined previously (23) resulted in the amplification of an approximately 700-bp DNA fragment. This fragment was used as a probe after digoxigenin labeling and gave a strong hybridization signal with an approximately 5-kb fragment of a SacII digest and a 4- to 5-kb fragment of a PstI digest of the total DNA from strain S2. Furthermore, a second, weaker signal with an approximately 11-kb fragment of the PstI digest was observed. This suggested that the probe also hybridized with the same 11-kb PstI fragment which was present in plasmid pMC171. The 5-kb SacII fragment was cloned in *pca* plasmid pMC171 (Fig. 2). DNA sequencing of the insert in pMC172-2 showed that part of the 5′ end of *pcaH2* was missing on that clone. Therefore the insert DNA was used to amplify a probe and to identify another plasmid from a gene library constructed with chromosomal DNA partially digested with PstI. The resulting plasmid (pS2GB-II-b3) contained the complete *pcaH2* gene. The *pcaH2G2* genes were expressed using a T7 expression system as described above for the *pcaH1G1* genes, and it was shown that the encoded dioxygenase was able to convert 4SC and protocatechuate (specific activities, 0.05 and 0.04 U/mg of protein, respectively). The cultures of *pcaH2G2* demonstrated that both genes overlapped for 11 bp. This has not been previously observed for *pcaHG* genes from other sources. The deduced amino acid sequences of PcaH-I and PcaH-II shared only 40% sequence identity, and PcaG-I and PcaG-II shared only 38% sequence identity.

**Analysis of the gene cluster with pcaH2G2.** The inserts in plasmids pMCS2-2 and pS2GB-II-b3 were completely sequenced. In this way the putative gene for a 3-carboxymuconate cycloisomerase (tentatively designated *pcaB2*) was identified. We are currently trying to express *pcaB2* in order to...
demonstrate its function in the 4SC degradative pathway. Surprisingly, pcaH2G2 and pcaB2 were separated by approximately 4.3 kb which did not show any homology to known genes from the β-ketoadipate pathway. A BLASTX search with the sequences downstream of pcaG2 revealed two putative open reading frames (ORFs) (ORF1 and ORF2 in Fig. 2) with the highest degree of identity to two proteins (DctP and DctM) from Rhodobacter capsulatus. These proteins belong to a high-affinity transport system for C4-dicarboxylates (TRAP transporters) (20). Approximately 3,100 bp downstream of pcaG2 could be identified ORF3, which was transcribed in the opposite direction to pcaH2G2, ORF1, ORF2, and pcaB2 (Fig. 2). ORF3 coded for a putative protein which showed the highest degree of homology to transcriptional regulators of the IclR family. This group of regulators has already been described repeatedly as regulators in the protocatechuate metabolism.

FIG. 3. Sequence alignment of the subunits of P34Os. Positions that are identical in all sequences of PcaH and PcaG are highlighted by black boxes. Those identical in all P34Os but different in the sequences of P34O-I are marked by open boxes. The accession numbers (references) for the published sequences of the enzymes from A. calcoaceticus, A. tumefaciens, B. cepacia, P. putida, R. opacus, and Streptomyces coelicolor are L05770 (24), U32867, M30791 (60), L14836 (21), AF003947 (16), and AL079355 (47). (The function of the genes from S. coelicolor has not been proven experimentally.)
olism of organisms such as *A. calcoaceticus*, *A. tumefaciens*, *P. putida*, and *Rhodococcus opacus* (13, 16, 22, 45, 48).

**DISCUSSION**

The results shown in the present study and the biochemical data published previously (18, 23) clearly demonstrate that *A. radiobacter* strain S2 synthesizes two different types of P34Os. The alignment of the amino acid sequences and the deduced dendrograms showed that both enzymes clearly clustered together with the previously determined sequences of the P34Os from organisms such as *A. calcoaceticus*, *P. putida*, *Burkholderia cepacia*, and *R. opacus* (Fig. 3 and 4) (16, 21, 24, 60) and were much more distantly related to other groups of intradiol dioxygenases such as catechol and chlorocatechol 1,2-dioxygenases. Thus, it appears that strain S2 synthesizes a distinct type of P34O (the type II enzyme) which is especially adapted to the degradation of the xenobiotic 4SC but which also shows significant activity with the natural substrate protocatechuate.

The presence of two different P34Os in one bacterial strain has been previously shown for a *Moraxella* sp., which synthesized two different P34Os after growth on different substrates (53). From dendrograms comparing both subunits of the P34Os from strain S2 with the respective subunits of the P34Os from other organisms, it appears that the two isoenzymes from strain S2 are not more closely related to each other then to the P34Os from other organisms (Fig. 4). This suggests that the evolution of the P34O-II has not taken place in *Agrobacterium* but that the relevant genes have been acquired from another organism. This is also suggested by the different GC contents observed for the type I genes (58.6 and 60.5% for *pcaH1* and *pcaG1*, respectively) and the type II genes (51.1 and 53.3% for *pcaH2* and *pcaG2*, respectively). For members of the genus *Agrobacterium*, an average GC content of 57 to 63% has been found (28). Different origins of the two types of P34O were also suggested by the different organizations of the genes which accompany the two types of P34Os. While the protocatechuate operon which encodes the type II enzyme is obviously organized identically to the operon from *A. tumefaciens* A348 studied by Parke (43–45), the type II genes are found in a totally different type of genetic organization.

Although they are based on the analysis of just one example, a few first conclusions about the evolution of the new 4SC pathway may be drawn at the present state of the investigations. First, it appears that only small modifications are necessary for the adaptation of a classical P34O (and presumably also of a carboxymuconate cycloisomerase) to allow the transformation of the respective sulfonated structural analogues. This was surprising, as it has been suggested previously that sulfo substituents (like chloro or nitro groups) confer a xenobiotic character to synthetic compounds because the electron-withdraw-
ing character of these substituents generates an electron deficiency and thus make these compounds less susceptible to oxidative catalolysis (29). The clear homology of the P340-I and P340-II suggests that the two enzymes convert their substrates by basically identical catalytical mechanisms. This suggests that the oxidation of the sulfonated catechol is not limited by the general mechanism of enzymatic ring fission but that more probably simple steric problems limit the oxidation of 4SC. Whether this is also true for other sulfonated compounds or polysulfonated substrates remains to be investigated.

From the alignment of the P340-II with the classical P340s, only nine amino acids could be identified in both subunits which are conserved among all P340-Is but which are different in the P340-II (Fig. 3). For one of these residues (K134), which is a lysine in PcaG-II but an arginine in all other P340s, it has been recently shown for the enzyme from A. calcoaceticus that a mutation of this residue (R133H) allowed the P340 to convert catechol, which is not oxidized by the wild-type enzyme (11). Furthermore, structural data from the crystal structure of the P340 from P. putida suggested that the exchange at position 154 of PcaH-II, which is a valine in PcaH-II but a tryptophan in all other P340s, may also be relevant, because this residue is close to the carboxylic group of protocatechuate in the active center of the P340 from P. putida (38). We are currently trying to identify the amino acid exchanges which are responsible for the ability of the P340-II to convert 4SC.

While the structural genes for pcaH2G2 and pcaB2 clearly resemble their classical counterparts from the degradative pathway of protocatechuate, the putative operon structure of the type II genes seems to be fundamentally different from the currently known (supra-)operonic structures of the type II genes. Thus, it may be possible that the gene products from ORF1 and ORF2 are required for the uptake of 4SC or other intermediates of the 4SC pathway by A. radiobacter strain S2 (10).

The existence of a second type of P340 which is especially adapted to the degradation of the (xenobiotic) 4SC clearly resembles the situation observed for the degradation of chlorocatechols. It has been shown that some bacterial strains are able to synthesize a classical β-ketoacid pathway for the degradation of catechol and a second set of evolutionarily related chlorocatechol 1,2-dioxygenases and chloromuconate cycloisomerases which are specifically adapted to the degradation of chlorinated catechols and muconates (50). Thus, it appears that the degradation of substituted benzenesulfonates is accomplished in nature by another novel variation of the β-ketoacid pathway.

REFERENCES


23. Hammer, A., A. Stolz, and H.-J. Knackmuss. 1996. Purification and characterization of a novel type of protocatechuate 3,4-dioxygenase with the ability


