Genetic and chemical characterization of ibuprofen degradation by *Sphingomonas Ibu-2*

Robert W. Murdoch\(^1\) and Anthony G. Hay\(^1,2\)

\(^1\)Graduate Program in Environmental Toxicology, Institute for Comparative and Environmental Toxicology, Cornell University, Ithaca, NY 14850, USA

\(^2\)Department of Microbiology, B53A Wing Hall, Cornell University, Ithaca, NY 14850, USA

*Sphingomonas Ibu-2* has the unusual ability to cleave the acid side chain from the pharmaceutical ibuprofen and related ary lacetic acid derivatives to yield corresponding catechols under aerobic conditions via a previously uncharacterized mechanism. Screening a chromosomal library of Ibu-2 DNA in *Escherichia coli* EPI300 allowed us to identify one fosmid clone (pFOS3G7) that conferred the ability to metabolize ibuprofen to isobutylcatechol. Characterization of pFOS3G7 loss-of-function transposon mutants permitted identification of five ORFs, *ipfABDEF*, whose predicted amino acid sequences bore similarity to the large and small units of an aromatic dioxygenase (*ipfAB*), a sterol carrier protein X (SCPx) thiolase (*ipfD*), a domain of unknown function 35 (DUF35) protein (*ipfE*) and an aromatic CoA ligase (*ipfF*). Two additional ORFs, *ipfH* and *ipfI*, which encode putative ferredoxin reductase and ferredoxin components of an aromatic dioxygenase system, respectively, were also identified on pFOS3G7. Complementation of a markerless loss-of-function *ipfD* deletion mutant restored catechol production as did complementation of the *ipfF* Tn mutant. Expression of subcloned *ipfABDEF* alone in *E. coli* did not impart full metabolic activity unless coexpressed with *ipfHI*. CoA ligation followed by ring oxidation is common to phenylacetic acid pathways. However, the need for a putative SCPx thiolase (*ipfD*) and DUF35 protein (*ipfE*) in aerobic ary lacetic acid degradation is unprecedented. This work provides preliminary insights into the mechanism behind this novel ary lacetic acid-deacy lating, catechol-generating activity.

INTRODUCTION

Ibuprofen [2-(4-isobutylphenyl-propionic acid)] is a pharmaceutical with analgesic, anti-pyretic and anti-inflammatory properties. With an annual production of several kilotones (Buser *et al.*, 1999), it is the most widely used member of a diverse class of pharmaceuticals termed non-steroidal anti-inflammatory drugs (NSAIDs). Many NSAIDs such as diclofenac, naproxen, ketoprofen and flurbiprofen share a phenylacetic acid (PAA) core. Ibuprofen has been detected in bodies of water worldwide (Buser *et al.*, 1999; Farré *et al.*, 2001; Kolpin *et al.*, 2002; Stumpf *et al.*, 1999; Winkler *et al.*, 2001). Ibuprofen has also been detected in water used for irrigation (Kinney *et al.*, 2006; Pedersen *et al.*, 2003, 2005; Siemens *et al.*, 2008; Xu *et al.*, 2009) and municipal drinking water supplies (Jones *et al.*, 2005). Environmental concentrations of ibuprofen have been found to range from low part-per-trillion to low part-per-billion levels (Buser *et al.*, 1999; Farré *et al.*, 2001; Santos *et al.*, 2010).

Investigations into its environmental impacts have found that ibuprofen induced changes on the timing of spawning by medaka (Flippin *et al.*, 2007; Han *et al.*, 2010), growth/ predominance of algae and duckweed (Pomati *et al.*, 2004; Richards *et al.*, 2004), microbial diversity in aquatic mesocosms (Richards *et al.*, 2004), and riverine biofilm communities (Lawrence *et al.*, 2005) at environmentally relevant concentrations.

Little information exists regarding how ibuprofen is oxidatively metabolized by microbes. While microbially generated hydroxyibuprofen metabolites and carboxylated ibuprofen have been detected (Hanlon *et al.*, 1994; Marco-Urrea *et al.*, 2009; Quintana *et al.*, 2005; Zwiener *et al.*, 2002), these metabolites have not been linked directly with any organism’s ability to use ibuprofen as a growth or
energy source, nor did they account for the majority of the added ibuprofen.

In addition to concerns regarding environmental presence and fate of pharmaceutical compounds, there is an increasing interest in the potential for gut microbes to directly or indirectly alter the pharmacokinetics of orally administered medicines, including NSAIDs (Clayton et al., 2009; Aziz et al., 2011; Wilson, 2009).

In an effort to better understand the mechanisms involved in ibuprofen degradation, *Sphingomonas* Ibu-2 was isolated from a sewage treatment plant, and shown to use racemic ibuprofen as a sole carbon and energy source (Murdoch & Hay, 2005). Ibu-2 utilizes a classical catechol meta-cleavage-type pathway. However, the catechol that is generated from ibuprofen is isotulatorycatehol, the creation of which requires the unprecedented removal of the propionic acid side chain from the aromatic ring. Ibu-2 performs similar reactions with other arylacetic acids, including PAA, 2-phenylpropionic acid, 3- and 4-tolylacetic acids, and 2-(4-tolyl)propionic acid, converting them to the corresponding catechol (or methylcatechol). Although this is somewhat reminiscent of the removal of the carboxyl moiety from benzoate (Eaton, 1996; Fetzner et al., 1992; Jeffrey et al., 1992; Reiner, 1971), it differs from the previously characterized PAA pathways of other bacteria, which do not involve catecholic intermediates (Fernández et al., 2006; Ismail et al., 2003; Martínez-Blanco et al., 1990; El-Said Mohamed, 2000; Rost et al., 2002; Teufel et al., 2010).

As the metabolism of ibuprofen and related PAAs by *Sphingomonas* Ibu-2 represents a new paradigm for the metabolism of PAAs, we undertook a genetic analysis to gain insight into the mechanism responsible for this unique deacylation activity. A fosmid library was constructed of Ibu-2 total DNA. A fosmid that conferred upon *Escherichia coli* the ability to generate isotulatorycatehol from ibuprofen and catechol from PAA was subjected to transposon mutagenesis. We report here the results of these efforts and propose a novel pathway for ibuprofen and PAA degradation by *Sphingomonas* Ibu-2.

**METHODS**

**Materials.** Unless otherwise noted, chemicals were purchased from Acros. Lysogeny broth (LB) was prepared as described by Sambrook et al. (1989). Mineral salts medium (MSM) was prepared as described by McCullar et al. (1994).

**Creation and screening of Ibu-2 fosmid library.** Ibu-2 chromosomal DNA was extracted from cells harvested from 100 ml of liquid MSM culture containing 2.4 mM ibuprofen by the standard alkaline lysis procedure (Sambrook et al., 1989). The fosmid library, hosted in *E. coli* EPI300, was created according to the instructions in the CopyControl® Fosmid Library Production kit (Epicentre Biotechnologies). The fosmid library was screened for the accumulation of dark brown catecholic polymers when grown in LB with 0.24 mM ibuprofen in 96-well plates. Arabinose (10 mM) was used to induce the fosmid to high copy number in all metabolic assays.

**Creation and metabolic screening of fosmid clone transposon libraries.** The EZ::TN <TET-1 >Insertion kit (Epicentre Technologies) was used to create transposon insertion mutants of chromosomal library fosmid pFOS3G7 (Table 1), which was positive for isotulatorycatehol accumulation. The reaction was packaged in phage extract (MaxPlax Lambda Packaging Extract; Epicentre Technologies), transfected into *E. coli* EPI300 and selected on LB plates containing chloramphenicol (Chl, 25 mg l$^{-1}$) and tetracycline (Tet, 12 mg l$^{-1}$). A 96-clone library of transposon mutants was then screened for loss of the ability to accumulate the brown colour associated with catechol production and polymerization. The loss-of-function mutants were characterized by sequencing DNA surrounding the site of transposon insertion using transposon-specific primers FP and RP (Epicentre Biotechnologies; Table S1 available with the online version of this paper). In addition, other non-loss-of-function mutants were also sequenced to increase the pool of sequence information. The sequence information was compiled using the SeqManager program (DNASTAR, Inc.) to build contigs. ORFs were detected using GeneQuest (DNASTAR, Inc.). The validity of the contig assemblies was confirmed by PCR mapping as detailed in Fig. S1. This sequence information can be found in the National Center for Biotechnology Institute (NCBI) database under accession number EF909268. Similarity to known and putative proteins was assessed using BLASTp (Altschul et al., 1990).

**Functional analysis of pFOS3G7 clones via HPLC.** Overnight cultures of *E. coli* EPI300 harbouring pFOS3G7 or different transposon mutants of pFOS3G7 were inoculated (10%, v/v) into 5 ml LB containing 0.24 mM ibuprofen, the appropriate antibiotics and 10 mM arabinose. The cultures were incubated at 37 ºC in a rotary shaker. Samples of 1 ml were taken at the initiation of the experiment and at 4 days and analysed for ibuprofen concentration via HPLC. The HPLC eluent, 80% methanol/20% 40 mM acetic acid, was pumped at a rate of 1 ml min$^{-1}$ using a Waters model 590 pump through a Varian Microsorb-MV C18 column (250 mm by 4.6 mm). Samples and standards were injected using a Shimadzu SIL-10A AD autoinjector and detected with a Shimadzu SPD-10A VP UV-Vis detector. The UV-Vis signal was analysed using PeakSimple (SRI Instruments). Ibuprofen was quantified by comparison with a standard curve.

GC/MS was performed on chloroform extracts of overnight cultures to characterize catecholic metabolites or other detectable metabolites that might accumulate. Samples were derivatized and analysed via GC/MS using previously described methods (Murdoch & Hay, 2005). The presence or absence of catechols was determined by HPLC analysis of washed cell suspensions that had been concentrated 20-fold and was performed as previously described (Murdoch & Hay, 2005).

**Fosmid subcloning.** Analysis of sequence from the transposon mutants implicated an approximately 5.3 kb region of fosmid pFOS3G7 (flanked by BamHI and NsiI sites at positions 1951–1956 and 17563–17560, respectively, on contig 1) as being necessary for the production of isotulatorycatehol (Fig. 1). Digestion with BamHI and NsiI yielded a 5.6 kb fragment that was gel purified away from the rest of the fosmid. The fragment was then ligated into pBBR1MC5 (Kovach et al., 1995) that had been digested with BamHI and PstI. The ligation was used to transform *E. coli* JM109 via electroporation and selected on LB Chl (25 mg l$^{-1}$) plates. The resulting plasmid, pIPFA-F (Table 1), was then harvested and used to transform *E. coli* EPI300 to yield *E. coli* EPI300 (pIPFA-F) (strain IPFA-F) to afford comparisons with the fosmid in the same genetic background.

PCR strand overlap extension was employed to generate a construct that could coexpress ipFLH, a putative ferredoxin reductase gene, and ipFL, a ferredoxin gene, both of which were also located on pFOS3G7. Primers ipFLFeDoxFsew and ipFLFeDoxRsew2 were used to amplify ipFL.
and primers ipfFeDoxRedFsew and ipfFeDoxRedRsew to amplify ipfH (Table S1). The ipf reverse primer and the ipfH forward primer were designed with artificial 19 bp 5' homologous regions. Each primer set was used to generate amplicons using pfu polymerase with pFOS3G7 serving as template. The PCR conditions were as follows: denaturation at 98 °C, annealing at 50 °C and 1.5 min extension at 72 °C, for 30 cycles. Following agarose gel purification, the 1.4 kb ipfH and 0.6 kb ipf products were combined together, their artificial homology regions were allowed to anneal, and then they were subjected to three cycles of no-primer PCR extension using the same reaction conditions. Following self-primering and extension, primers ipfFeDoxFsew and ipfFeDoxRedRsew were added and 30 further PCR cycles were performed, yielding a combined 2.0 kb ipfH and ipfH product. This blunt-ended strand extension product was adenylated by adding Taq polymerase and ATP following purification and then cloned into pGEM-easy to create pIPFHI. The resultant plasmid was sequence verified and used to transform both E. coli EP1300, Chl' when hosted in E. coli EP1300, Chl' sequence verified and used to transform both E. coli EP1300, Chl' when hosted in E. coli EP1300, and E. coli IPFHI to yield strain IPFHI and into IPFA-FHI.

**Table 1. Strains and plasmids used in this study**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sphingomonas</em> lba-2</td>
<td>Isolated from Ithaca, NY, USA, sewage treatment plant via enrichment for growth on ibuprofen</td>
<td>Murdoch &amp; Hay (2005)</td>
</tr>
<tr>
<td><em>Pseudomonas putida</em> U</td>
<td>Phenylacetic acid degrader with phenylacetyl-CoA ligase activity</td>
<td>Martínez-Blanco et al. (1990)</td>
</tr>
<tr>
<td><em>E. coli</em> JM109</td>
<td>recA1 subE44 endA1 hsdR17 gyrA96 relA1 thi Δ(lac-proAB) F' (traD36 proAB lacZ M15)</td>
<td>Sambrook et al. (1989)</td>
</tr>
<tr>
<td><em>E. coli</em> EP1300</td>
<td>F' Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK Δ⁻ rpsL (Str⁺) supG trfA tonA</td>
<td>Epicentre</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCC1FOS</td>
<td>Copy control fosmid vector, inducible to high copy number with arabinose</td>
<td>Epicentre</td>
</tr>
<tr>
<td>pGEM-T Easy</td>
<td>T-cloning vector, Amp'</td>
<td>Promega</td>
</tr>
<tr>
<td>pBBR1MCS</td>
<td>Broad-host-range cloning vector, Chl'</td>
<td>Kovach et al. (1995)</td>
</tr>
<tr>
<td>pIPFA-F</td>
<td>pBBR1MCS with 5.6 kb BamHI/NsiI fragment from pCC1FOS 3G7 bearing ipfABDEF</td>
<td>This study</td>
</tr>
<tr>
<td>pGEM: ipfDrbs</td>
<td>pGEM-T Easy with ipfDrbsF/ipfDrR PCR product including a three-frame stop codon and a ribosome-binding site</td>
<td>This study</td>
</tr>
<tr>
<td>pGEM: ipfF</td>
<td>pGEM-T Easy with ipfFF/ipfFR PCR ampiclon of ipfF</td>
<td>This study</td>
</tr>
<tr>
<td>pIPFHI</td>
<td>pGEM-T Easy with 2 kb ipfH PCR ampiclon created by strand overlap extension of ipfH and ipfF amplicons</td>
<td>This study</td>
</tr>
<tr>
<td>pKD4</td>
<td>Contains the template for generating FLP recombinase target lambda red kanamycin resistance PCR fragment; Kan'</td>
<td>Datsenko &amp; Wanner (2000)</td>
</tr>
<tr>
<td>pKD46</td>
<td>Heat-inducible FLP recombinase expression plasmid; Amp'</td>
<td>Datsenko &amp; Wanner (2000)</td>
</tr>
<tr>
<td>pFOS3G7Tn: ipfA</td>
<td>pFOS3G7 with Tn5 insertion in ipfA, transposon library clone F1, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn: ipfB</td>
<td>pFOS3G7 with Tn5 insertion in ipfA, transposon library clone H6, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn: ipfD</td>
<td>pFOS3G7 with Tn5 insertion in ipfD, transposon library clone A2, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn: ipfE</td>
<td>pFOS3G7 with Tn5 insertion in ipfE, transposon library clone E6, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn: ipfF</td>
<td>pFOS3G7 with Tn5 insertion in ipfF, transposon library clone F10, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn B8</td>
<td>pFOS3G7 with Tn5 insertion in ipfA, transposon library clone B8, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn H8</td>
<td>pFOS3G7 with Tn5 insertion in ipfA, transposon library clone H8, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn C1</td>
<td>pFOS3G7 with Tn5 insertion in ipfD, transposon library clone C1, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn F9</td>
<td>pFOS3G7 with Tn5 insertion in ipfD, transposon library clone F9, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn D4</td>
<td>pFOS3G7 with Tn5 insertion in ipfD, transposon library clone D4, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn E7</td>
<td>pFOS3G7 with Tn5 insertion in ipfE, transposon library clone E7, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn E8</td>
<td>pFOS3G7 with Tn5 insertion in ipfF, transposon library clone E8, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn C9</td>
<td>pFOS3G7 with Tn5 insertion in ipfF, transposon library clone C9, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7Tn B1</td>
<td>pFOS3G7 with Tn5 insertion in ipfF, transposon library clone B1, Tet' Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7ΔipfA</td>
<td>pFOS3G7 with markerless deletion of ipfA, Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7ΔipfB</td>
<td>pFOS3G7 with markerless deletion of ipfB, Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7ΔipfD</td>
<td>pFOS3G7 with markerless deletion of ipfD, Chl'</td>
<td>This study</td>
</tr>
<tr>
<td>pFOS3G7ΔipfE</td>
<td>pFOS3G7 with markerless deletion of ipfE, Chl'</td>
<td>This study</td>
</tr>
</tbody>
</table>
promote higher copy numbers of the fosmid prior to the introduction of substrate and avoid negative selection effects exerted by potentially toxic catecholic metabolites, substrate addition (1 mM final concentration) was delayed until the cultures had entered stationary phase. Test substrates were ibuprofen and phenylacetate, both of which have been previously shown to be deacylated to the corresponding catechols by Sphingomonas Ibu-2 (Murdoch & Hay, 2005). The media consisted of LB and 100 mg Chl ml\(^{-1}\) and/or 150 mg ampicillin ( Amp\(^{\beta}\) ) 1\(^{-1}\). Following the addition of test substrate, cultures were incubated at 37 °C for 18 h.

For direct visualization of catechols, ferric chloride was added to 150 μl of culture to a final concentration of 1.5 mM in 96-well plate format. Additionally, HPLC was used to directly quantify substrate and catecholic metabolites. A 40:60 methanol/40 mM acetic acid eluent was used for separation of phenylacetate (10.6 min) and catechol (4.3 min). A detection wavelength of 220 nm was used for the aromatic acids while a detection wavelength of 280 nm was used for the catechol. An eluent of 70:30 methanol/40 mM acetic acid was used to separate ibuprofen (13.2 min) and a peak at 3 min presumed to be isobutylcatechol due to its high absorbance at 280 nm. Standard curves were used to quantify the analytes. Because no standard exists for isobutylcatechol, the catechol standard curve was used to approximate isobutylcatechol concentration.

**Complementation of pFOS3G7Tn mutants.** pFOS3G7Tn : ipfF was complemented by cloning ipfF into pGEM-easy to give pGEM : ipfF (Table 1) using primers ipfFF/ipfFR (Table S1) that had been designed so as to include the native ribosome-binding site (Table 1). Attempts to complement the other four genes (ipfABDE) using the same approach described above were unsuccessful (results not shown). To reduce the influence of possible polar effects introduced by the Tn5 cassette, markerless mutants were created using the lambda red protocol described by Datsenko & Wanner (2000), except that 10 mM arabinose was added during both the recovery and the initial plating. Ip6A lambda, Ip6B lambda and Ip6D lambda primer sets, each with 36 bp 5’ regions homologous to DNA immediately flanking start and stop codons of the target gene (Table S1), were used to generate insertion cassettes with pKD4 as template (Datsenko & Wanner, 2000). Insertions and deletions were confirmed by size analysis of PCR amplicons using the ipfA, ipfB or ipfD primer sets, each of which flanks the targeted gene region, but are located outside of the deleted portions (Table S1).

To address the possibility that the lack of complementation was due to lack of efficient translation, complements of the deletion mutants were created using primers with artificial stop codons and E. coli-optimized ribosome-binding sites added to the 5’ end of the forward primers (Table S1). PCR was performed using pufS polymerase followed by mnoadenylation of the resulting blunt-ended product by adding Taq polymerase and ATP following purification. The resulting plasmids were sequenced and used to transform pFOS3G7 deletion mutants or transposon mutants and screened for catechol accumulation.

**In silico analyses.** Fosmid transposon library clone sequences were assembled into contiguous units using SeqMan and searched for ORFs using GeneQuest. Translated ORFs were subjected to BLASTX analysis of these ORFs, including accession numbers and identity/similarity percentages, are as follows: (1) catechol, 2,3-dioxygenase, P47228, 90/94%; (2) 4-oxaloacrylate tautomerase, Q9RHM8, 84/95%; (3) peniplasmic binding protein of ABC transport system, P21175, 25/46%; (4) catechol, 2,3-dioxygenase, P11122, 69/74%; (5) 4-hydroxy-2-oxovalerate aldolase, O85977, 90/95%; (6) 4-oxaloacrylate decarboxylase, Q9KWS3, 53/73%; (7) 4-oxaloacrylate isomerase, Q9RHM8, 34/59%; (8) plant-like ferredoxin, P23103, 49/65%; (8) dehydrogenase, P23102, 54/69%. Large black arrows represent ORFs with high similarity to conserved transposase genes. The locations of the BamHI and NsiI restriction sites used in the generation of pIPFA–F on contig 1 at positions 1951–1956 and 7565–7560, respectively, are indicated.
followed by two washes with 10 mM phosphate buffer (pH 7.4) and resuspension in 1–2 ml of sonication buffer [40 mM potassium phosphate, 20% (v/v) glycerol, 1 mM PMSF, 1 mM DTT, pH 7.4]. For Ibu-2 and *P. putida* U, the concentrated cell mass was then sonicated (Branson Sonifier 450; Branson Ultrasonics) using three 1-min cycles at maximum output with 1-min rest time on ice in between each cycle. For *E. coli* crude extract preparation, approximately 0.1 g of 0.1 mm glass beads was added and the pellets were bead-beaten for 3 min (MiniBeadbeater-8; Biospec Products). In all cases, the cell lysate was then spun for 20 min at 15 000 g at 4°C. The protein content of the supernatant containing the crude extract was quantified using the Bio-Rad Protein Assay Kit with BSA as a standard (Bio-Rad Laboratories). Phenylacetate CoA ligase assays were performed as described by Martinez-Blanco *et al.* (1990) and quantified with a phenylacetate ferric chloride extinction coefficient of 0.9 mM⁻¹ cm⁻¹.

**RESULTS**

**Ibu-2 fosmid library**

*E. coli* EPI300(pFOS3G7) was the only clone of the 900-clone Ibu-2 fosmid library that accumulated visible brown coloration when exposed to ibuprofen in liquid media. The presence of an acetylated derivative of isobutylcatechol in ethyl acetate extracts of acetylated culture supernatant from *E. coli* EPI300(pFOS3G7) grown in LB and exposed to 2.4 mM ibuprofen was confirmed via GC/MS. With a retention time of 15.7 min and major peaks (relative abundance) of 123(99), 166(100), 208(17) and 250(4), it was identical to that which had been previously observed in Sphingomonas Ibu-2 culture extract (Murdoch & Hay, 2005).

**Screening and characterization of *E. coli* EPI300(pFOS3G7) transposon library and deletion mutants**

When the 96-clone transposon library of *E. coli* EPI300(pFOS3G7) was screened for catechol accumulation in the presence of ibuprofen, 17 clones no longer accumulated the characteristic dark catecholic polymerization product, indicating a loss of function. In addition, a single transposon mutant that exhibited less dark coloration was selected for further analysis. Sequencing of the DNA surrounding the transposon insertion and analysis of the sequence information revealed that four of the loss-of-function *E. coli* EPI300(pFOS3G7) transposon insertion clones had the transposon in the vector. Loss of function in these mutants was possibly due to reduction of copy number, loss of the insert or other unidentified mutations, but was not further investigated. The remaining 13 loss-of-function mutants harboured a transposon in a 5.3 kb region of Ibu-2 DNA. Sequence analysis (DNASTAR, Inc.) revealed that this 5.3 kb region contained five ORFs (Fig. 1) with sequence similarities to genes encoding enzymes for aromatic acid degradation and other catabolic activities described in Table 2. Following additional sequencing to increase sequence quality and confidence (data not shown), the full 5.3 kb region was redundantly covered in both forward and reverse directions. All five ORFs had representative knockouts amongst the loss-of-function mutants. Sequencing of the transposon mutant that generated less colour in the presence of ibuprofen allowed us to identify the ORF *ipfH*, which had similarity to ferredoxin reductase components of aromatic dioxygenase systems. Additional sequencing of randomly selected Tn5 mutants allowed the fortuitous identification of *ipfI*, which bears similarity to ferredoxins of aromatic dioxygenase systems. Altogether, sequencing of the fosmid transposon clones allowed for the assembly of two contigs of 11741 and 8229 bp (Fig. 1). These assemblies were confirmed by PCR mapping (Fig. S1). Sequence information can be found under NCBI submission EF090268.

Deletion mutants of *ipfABDE* were successfully constructed (as indicated by PCR analysis, data not shown) to eliminate the possibility of polar effects. The *ipfF* gene was not targeted for deletion because sequence analysis showed it to be downstream of the other genes and so it was unlikely that the *ipfTn5* phenotype caused any polar effects.

**Metabolic analyses of clones and constructs**

Ibuprofen disappearance assays revealed several trends. First, they clearly demonstrated that expression of pFOS3G7 in *E. coli* EPI300 caused the disappearance of ibuprofen from the culture supernatant (Fig. 2a). Tn : *ipfF* and markerless deletion mutants of the individual *ipf* genes (*ΔipfA, ΔipfB, ΔipfD* and ΔipfE) failed to produce catecholic metabolites and eliminated significantly less (*P*<0.025) ibuprofen than intact pFOS3G7 (Fig. 2). Complementation of pFOS3G7*ΔipfD* with pGEM : *ipfD* and pFOS3G7Tn : *ipfF* with pGEM : *ipfF* restored catechol generation and ibuprofen disappearance (Fig. 2) although we were unable to complement the other mutants.

When *ipfABDEF* were subcloned into pBBR1MCS to create pIPFA-F, *E. coli* harbouring this plasmid produced trace amounts of isobutylcatechol when grown in the presence of ibuprofen. This was detectable only by GC/MS (data not shown), but no polymerization products were visible nor was isobutylcatechol detectable via HPLC analysis. In contrast, IPFA-FHI cultures incubated with ibuprofen produced a dark brown pigment (Fig. 3). Neither vector on its own nor the vectorless controls produced any detectable dark pigment. HPLC analysis revealed that IPFA-FHI caused the disappearance of significantly more ibuprofen and phenylacetate than the single vector or vectorless controls (*P*<0.025). Additionally, HPLC analysis revealed the production of catechol in the IPFA-FHI dual-vector cultures supplied with phenylacetate, and isobutylcatechol in the cultures supplied with ibuprofen. The disappearance of phenylacetate in the dual vector culture (0.58 mM) corresponded closely to the accumulation of catechol (0.37 mM) (Fig. 3). During the same time frame, 0.42 mM ibuprofen disappeared from the IPFA-FHI culture and approximately 0.25 mM isobutylcatechol appeared (Fig. 3).
Table 2. Two most similar functionally characterized proteins from the Swiss-Prot database as determined by BLASTP analysis of ipfABDEFHI

The contig and location where each ORF is found are indicated. The E-value represents the strength of the similarity with a lower value representing stronger similarity. Conserved domains present were detected using the CDD tool from NCBI. NA, Not applicable.

<table>
<thead>
<tr>
<th>ORF</th>
<th>Contig (position)</th>
<th>Accession no.</th>
<th>Description</th>
<th>Identity (%) (similarity, %)</th>
<th>Query coverage (%)</th>
<th>E-value</th>
<th>Conserved domain(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipfA</td>
<td>1 (2074–3378)</td>
<td>ABY62759.1</td>
<td>Indole 3-acetic acid dioxygenase IaaC</td>
<td>34 (51)</td>
<td>94</td>
<td>1E-62</td>
<td>cd03469: Rieske non-haem iron oxygenase (RO)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P37333.3</td>
<td>Biphenyl 2,3-dioxygenase BphA</td>
<td>29 (45)</td>
<td>93</td>
<td>6E-45</td>
<td></td>
</tr>
<tr>
<td>ipfB</td>
<td>1 (3485–4021)</td>
<td>AAB62286.1</td>
<td>p-Cumate 2,3 dioxygenase CmtAc</td>
<td>33 (47)</td>
<td>84</td>
<td>1E-16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>P37334.3</td>
<td>Biphenyl 2,3 dioxygenase BphE</td>
<td>31 (50)</td>
<td>85</td>
<td>4E-15</td>
<td>cd00667: ring-hydroxylating dioxygenases β subunit</td>
</tr>
<tr>
<td>ipfD</td>
<td>1 (4030–5229)</td>
<td>YP_158078.1</td>
<td>Benzoylsuccinyl-CoA thiolase subunit BbsB</td>
<td>34 (50)</td>
<td>53</td>
<td>5E-26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q07598.1</td>
<td>Propanoyl-CoA C-acyltransferase SCPx</td>
<td>30 (40)</td>
<td>64</td>
<td>8E-23</td>
<td></td>
</tr>
<tr>
<td>ipfE</td>
<td>1 (5242–5640)</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
<td>pfam12172: domain of unknown function 35 rubredoxin-like zinc ribbon. pfam01796: domain of unknown function 35 oligonucleotide binding fold</td>
</tr>
<tr>
<td>ipfF</td>
<td>1 (5665–7245)</td>
<td>O07610.2</td>
<td>Long-chain fatty acid CoA ligase LcfL</td>
<td>27 (47)</td>
<td>96</td>
<td>2E-33</td>
<td>cl10450: LuxE superfamily</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q9S725.2</td>
<td>4-Coumarate CoA ligase 4cl2</td>
<td>29 (47)</td>
<td>98</td>
<td>9E-51</td>
<td></td>
</tr>
<tr>
<td>ipfH</td>
<td>2 (3798–5052)</td>
<td>P43494.2</td>
<td>Rhodocoxin reductase</td>
<td>38 (52)</td>
<td>89</td>
<td>4E-43</td>
<td>pfam07992: pyridine nucleotide–disulphide oxidoreductase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P37337.2</td>
<td>Biphenyl dioxygenase system ferredoxin reductase</td>
<td>37 (52)</td>
<td>87</td>
<td>2E-37</td>
<td></td>
</tr>
<tr>
<td>ipfI</td>
<td>1 (10818–11150)</td>
<td>P0A185.2</td>
<td>Naphthalene 1,2-dioxygenase ferredoxin NdoA</td>
<td>39 (60)</td>
<td>90</td>
<td>2E-20</td>
<td>cd03528: non-haem iron oxygenase (RO)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Q0T1X9.1</td>
<td>Cinnamic acid dioxygenase ferredoxin HcaC</td>
<td>42 (59)</td>
<td>93</td>
<td>5E-20</td>
<td>ferredoxin</td>
</tr>
</tbody>
</table>
Coenzyme A ligase activity

CoA ligase activity in Ibu-2 cell-free extracts was determined using both PAA and ibuprofen as substrates. Ibu-2 crude extract catalysed phenylacetyl-CoA ligation at a mean (±SD) rate of 9.5 ± 2.0 nmol (mg protein)^{-1} min^{-1}. Prior exposure of Ibu-2 to ibuprofen had no detectable effect on CoA ligase activity (P=0.36). The positive control P. putida U produced similar levels of phenylacetyl-CoA ligase activity as Ibu-2 and did not yield any detectable ibuprofen CoA-ligation product (data not shown). Crude extract from E. coli EPI300(pFOS3G7) or E. coli JM109(pGEM:ipfF) contained phenylacetyl-CoA activities of 19.7 ± 1.1 and 34.0 ± 6.5 nmol (mg protein)^{-1} min^{-1}, respectively, which were significantly (P<0.05) different from the 0.5 ± 0.1 nmol (mg protein)^{-1} min^{-1} in vectorless EPI300 extracts. The increase in OD during CoA ligase assays containing ibuprofen and extracts from Ibu-2 and from E. coli JM109(pGEM:ipfF) was only 5% of that observed with

Ibuprofen remaining (%)

(a)

(b)

Fig. 2. (a) Percentage of 0.24 mM ibuprofen remaining after 2 days of incubation in E. coli EPI300(pFOS3G7), loss-of-function mutants and the two successful complementation constructs (ipfF and ipfD) as determined by HPLC analysis; n=3, standard deviations were too small to be visualized effectively (typically <1% of the means). Control: E. coli EPI300 with no vector. (b) Catecholic polymer accumulation in E. coli EPI300(pFOS3G7), loss-of-function mutants and the two successful complement constructs grown in LB with 0.24 mM ibuprofen and 1.5 mM ferric chloride.

Novel mechanism for metabolism of ibuprofen

PAA, but was significantly greater than the negative controls (P<0.05). The lack of an extinction coefficient for ibuprofen-CoA prevented a more precise calculation of ibuprofen CoA-ligation rate. Both PAA and ibuprofen CoA-ligase activities were dependent upon the presence of ATP and Mg^{2+} (data not shown).
DISCUSSION

Seven genes, ipfABDEFHI, identified on fosmid pFOS3G7 from a Sphingomonas Ibu-2 chromosomal library, were found to be sufficient to encode ibuprofen and PAA deacetylation activity when expressed in E. coli. Briefly, ipfA and ipfB appear to encode the large and small subunits, respectively, of an aromatic ring dioxygenase while ipfH and ipfI display similarity to the ferredoxin reductase and ferredoxin subunits of the electron transport chain of an aromatic ring dioxygenase, respectively. Analyses of ipfD and ipfE suggested similarities to genes encoding sterol carrier protein X (SCPx) and domain of unknown function 35 (DUF35), respectively, two poorly characterized classes of proteins. The ipfF gene has limited sequence similarity to genes encoding well-characterized CoA ligases used in fatty acid metabolism and in the anaerobic metabolism of benzoic acid (Table 2).

The phenylacetyl- and ibuprofen-CoA ligase activities of Ibu-2, E. coli EPI300(pFOS3G7) and E. coli (pGEM : ipfF) suggest that CoA ligation by Ipff is the first step performed in this pathway (Fig. 4). Despite this activity, ipfF showed no similarity to human xenobiotic/medium-chain fatty acid : CoA ligase, which is the only other enzyme that has been shown to be capable of performing ibuprofen CoA ligation (Vessey et al., 1996). Our finding that ibuprofen degradation was dependent on CoA ligation activity is consistent with well-characterized PAA catabolic pathways, which are also dependent on CoA ligation (Garcia et al., 2000; Teufel et al., 2010).

Following CoA ligation, PAA has recently been shown to undergo ring epoxidation by a mixed function oxidase encoded by paaABCDE in E. coli K-12 and P. putida (Teufel et al., 2010, 2012). Ring oxidation is also evident in the metabolism of ibuprofen by Ibu-2 based on the production of catechol, which we found to be dependent upon functioning copies of ipfAB in E. coli EPI300(pFOS3G7), although they do not bear detectable sequence similarity to the paa genes. Other indirect evidence for the role of a multicomponent oxygenase in the ring hydroxylation of PAA and ibuprofen is our observation that E. coli harbouring pIPFA-F (strain IPFA-F) only produced trace quantities of catechol and isobutyric acid, respectively, unless ipfH, which encode a putative ferredoxin and ferredoxin reductase, respectively, were present (Fig. 3). This dependency is consistent with the fact that many ring-hydroxylic aromatic dioxygenases require a reductase component for activity (Butler & Mason, 1997; Mason & Cammack, 1992) and also offers indirect support for the involvement of ipfAB, the putative aromatic ring dioxygenase subunits, in the pathway (Fig. 4).

At this point analogy to the paa pathways ceases because Ibu-2 produces catechols from alkyl aromatic acids and cleaves them oxidatively, whereas the paa pathway produces a CoA-oxepin, which is cleaved hydrolytically (Teufel et al., 2010, 2012). Despite numerous attempts, we were unable to detect any other intermediates that might unequivocally establish the exact nature of the steps between CoA ligation and catechol/isobutyric acid formation. However, given that production of these catechols by E. coli EPI300(pFOS3G7) was dependent on the presence of ipfD and E (Fig. 2) we hypothesize that IpfD and E are responsible for the acid side chain removal based on the role that related proteins play in acyl group transfer (Fig. 4) (Bangera & Thomashow, 1999; Kube et al., 2004; Kühner et al., 2005; Leuthner & Heider, 2000; Verhoeven & Jakobs, 2001; Wanders et al., 1997; Westin et al., 2007).

The predicted amino acid sequence of ipfD contains a highly conserved domain similar to that found in the poorly characterized SCPx family (E-value = 1.02 × 10⁻51) (Stolowich et al., 2002). Detailed molecular analyses of mammalian SCPx family members have demonstrated their involvement in two specific metabolic reactions in eukaryotes, namely bile-acid synthesis (Takeuchi et al., 2004) and the β-oxidation of pristanoyl-CoA, a branched-chain fatty acid (Fig. 5) (Verhoeven & Jakobs, 2001; Wanders et al., 1997; Westin et al., 2007). In both cases, SCPx is specifically involved in the β-oxidation of an α-methyl β-keto fatty acid. This is particularly relevant to ibuprofen as it is also an α-methyl acid.

![Fig. 4. Parent compounds (I) and corresponding catechols (IV) were produced by Sphingomonas Ibu-2 and detected by GC/MS and/or HPLC (Murdoch & Hay, 2005). Expression of ipfABDEFHI in E. coli was demonstrated in this study to be sufficient for the deacylation of phenylacetic acid and ibuprofen (R1 = methyl, R2 = isobutyl, R3 = H). The identity of metabolite II was determined for phenylacetic acid and ibuprofen via ipfF CoA assays, while the identity of metabolite III is hypothesized based on the putative identities of ipfABHI.](Image)
Although there are over 2000 SCPx-type bacterial genes in the NCBI database, only four have been characterized to any degree: ditF (Pseudomonas abietaniphila), phIC (P. putida) (Bangera & Thomashow, 1999), ORF-2 from the camphor catabolic cluster (Rhodococcus sp. NCIMB 9784) (Groban et al., 2001; Roberts et al., 2004), and bbsB from both Aromatoleum aromaticum EbN1 (Kube et al., 2004; Kühner et al., 2005) and Thauera aromatica (Leuthner & Heider, 2000). As part of a pathway for the metabolism of toluene to benzoyl-CoA under anaerobic conditions, BbsB in conjunction with BbsA performs the decondensation of an α-aliphatic β-keto CoA adduct in a manner that mirrors the role of eukaryotic SCPx in the metabolism of bile acids and branched fatty acids (Fig. 5).

Although the branched nature of several of these substrates is consistent with the similar reaction expected to be catalysed by IpfD, it is unclear how the ring of ibuprofen will affect the susceptibility of the β-carbon to nucleophilic attack by the sulfur atom of CoA as the putative hydroxylated β-carbon (position 1 on the ring) could not be further oxidized to a keto. Further clouding the mechanistic picture is our finding that despite significant sequence similarity, IpfD does not share the catalytic residues of thiolases and eukaryotic SCPx-type proteins: the latter share two primary catalytic residues that have been shown to be crucial for thiolase activity, a cysteine in the 80–100 aa region and a histidine in the 350–400 region (Bangera & Thomashow, 1999). Alignment of IpfD and other bacterial SCPx thiolases against well-characterized SCPx and FadA proteins reveals that IpfD lacks the N-terminal catalytic cysteine and the C-terminal catalytic histidine, both of which are present in BbsB (Liu et al., 1999; Pagani et al., 1994) (data not shown). Thus, despite the overall similarity of IpfD to existing thiolases and the similarity of the branched side chain of ibuprofen to that of SCPx substrates, more work needs to be done before a mechanistic explanation for the role of IpfD in ibuprofen metabolism can be established.

Conserved domain analysis of the predicted amino acid sequence of ipfE revealed that it contains a DUF35 motif (E-value = 3.71 × 10⁻¹¹). A functionally uncharacterized representative of the DUF35 family from Sulfolobus solfataricus was recently crystallized (Krishna et al., 2010). Based on structural features and contextual analysis, the authors speculated that DUF35 proteins probably play an acyl-CoA carrier role. Such an acyl-carrier function is consistent with the hypothesized role of IpfE as a partner interacting with IpfD in the ipf pathway (Fig. 4) and with the role of the few DUF35 proteins whose function has been characterized.

In the benzylsuccinate pathway, BbsA, which contains a DUF35 motif, is required for the thiololytic decondensation of benzylsuccinate-CoA catalysed by the SCPx thiolase BbsB (Fig. 5) (Kube et al., 2004; Kühner et al., 2005; Leuthner & Heider, 2000). Our genetic and sequence data suggest that IpfD, a putative thiolase, may partner with IpfE, a protein which contains a DUF35 domain, just as the thiolase BbsB partners with the DUF35 protein BbsA. Cooperation between an SCPx thiolase and a DUF35 protein has also been reported for PhIB and PhIC, which are involved in acetylation of an aromatic polyketide in P. putida (Bangera & Thomashow, 1999).
Given our genetic evidence and the similarity of IpfDE to known thiolase/DUF35 pairs we predict that they act upon dearamatized 1,2-diol-ibuprofen-CoA, an \(\alpha\)-methyl-\(\beta\)-hydroxyl-CoA fatty acid, removing a propionyl-CoA group (Fig. 4).

Analysis of additional sequence data from the pFOS3G7 transposon library suggested that the ipf genes were situated amongst a number of genes encoding ring fission and further metabolism of the ring cleavage product (Fig. 1). This is consistent with the previous identification of isobutylcyclic and putative catecholic ring-cleavage products in ibuprofen-grown Sphingomonas Ibu-2 cultures (Murdoch & Hay, 2005).

In summary, based on our preliminary biochemical evidence and the putative function implied from sequence similarity, we suggest that ibuprofen is first CoA-ligated by the CoA ligase IpfF, then dihydroxylated by the multi-component oxygenase IpfABHI, and that IpfD and E catalyse removal of the acyl-CoA group to yield catechol (Fig. 4). Removal of the acidic side chain via putative oxidation of the 1-position of the aromatic rings of an alkyl aromatic acid is unprecedented, but is reminiscent of the removal of an acyl group during \(\beta\)-oxidation of fatty acids (Trotter, 2001) such as that performed by the well-characterized Fad operon of E. coli (Black et al., 1992; Campbell & Cronan, 2002; Kunau et al., 1995). It differs from \(\beta\)-oxidation, however, in that the formation of a \(\beta\)-keto group is not possible at the 1-carbon of the ring even if it is de aromatized. Thus, this step represents a new paradigm in aromatic acid metabolism and will require additional work before the exact nature of the mechanism is understood.

ACKNOWLEDGEMENTS

R.W.M was supported in part by NIH-NIEHS Environmental and Molecular Toxicology training grant ES 07052-27 and the SUNY Graduate Diversity Fellowship Program. Anastasia Kazimirova provided invaluable technical support in the creation and screening of fosmid and fosmid transposon libraries. Amy Risen designed the strand overlap extension primers and provided additional sequence data. Hanh Nguyen performed additional enzyme assays and PCR mapping.

REFERENCES


Edited by: R. E. Parales