Molecular insights into substrate recognition and catalysis by phthalate dioxygenase from Comamonas testosteroni

Received for publication, August 31, 2021, and in revised form, November 9, 2021 Published, Papers in Press, November 17, 2021, https://doi.org/10.1016/j.jbc.2021.101416

Jai Krishna Mahto, Neetu Neetu, Bhairavnath Waghmode, Eugene Kuatsjah, Monica Sharma, Debabrata Sircar, Ashwani Kumar Sharma, Shailly Tomar, Lindsay D. Eltis, and Pravindra Kumar.

From the 1Department of Biosciences and Bioengineering, IIT Roorkee, Roorkee, India and 2Department of Microbiology & Immunology, Life Sciences Institute, The University of British Columbia, Vancouver, Canada

Edited by Ruma Banerjee

Phthalate, a plasticizer, endocrine disruptor, and potential carcinogen, is degraded by a variety of bacteria. This degradation is initiated by phthalate dioxygenase (PDO), a Rieske oxygenase (RO) that catalyzes the dihydroxylation of phthalate to a dihydrodiol. PDO has long served as a model for understanding ROs despite a lack of structural data. Here we purified PDO$_{KF1}$ from Comamonas testosteroni KF1 and found that it had an apparent $k_{cat}/K_m$ for phthalate of 0.58 ± 0.09 μM$^{-1}$s$^{-1}$, over 25-fold greater than for terephthalate. The crystal structure of the enzyme at 2.1 Å resolution revealed that it is a hexamer comprising two stacked $α_3$ trimers, a configuration not previously observed in RO crystal structures. We show that within each trimer, the protomers adopt a head-to-tail configuration typical of ROs. The stacking of the trimers is stabilized by two extended helices, which make the catalytic domain of PDO$_{KF1}$ larger than that of other characterized ROs. Complexes of PDO$_{KF1}$ with phthalate and terephthalate revealed that Arg207 and Arg244, two residues on one face of the active site, position these substrates for regiospecific hydroxylation. Consistent with their roles as determinants of substrate specificity, substitution of either residue with alanine yielded variants that did not detectably turnover phthalate. Together, these results provide critical insights into a pollutant-degrading enzyme that has served as a paradigm for ROs and facilitate the engineering of this enzyme for bioremediation and biocatalytic applications.

Phthalate is an endocrine-disrupting pollutant and potential carcinogen that is extensively used as a plasticizer in a wide range of consumer products, including plastics, medicines, and cosmetics (1–3). As phthalate is not covalently bound in these materials, it readily leaches into the environment, potentially exposing humans and other organisms to its detrimental health effects through inhalation, ingestion, and absorption (2). Phthalate is also a major catabolic intermediate in the biodegradation of phthalate esters and some polycyclic aromatic hydrocarbons, such as fluoranthene, fluorene, and phenanthrene (4).

A wide range of bacterial strains are able to grow on phthalate, including proteobacteria such as Burkholderia cepacia DB01 (5, 6) and Comamonas testosteroni (7), and actinobacteria such as Rhodococcus jostii RHA1 (8) and Arthrobacter keyseri 12B (9). In these strains, the first step in phthalate catabolism is dihydroxylation of the aromatic diacid catalyzed by the phthalate dioxygenase (PDO). PDO$_{DB01}$ from B. cepacia DB01 and PDO$_{KF1}$ of C. testosteroni KF1 catalyze the dihydroxylation of phthalate to cis-4,5-dihydrodiol phthalate (DHP) (6) (Fig. 1A). This differs from the reaction catalyzed by the actinobacterial PDOs, which is 3,4-dihydroxylation (9). Nevertheless, in both cases, the dihydrodiol is eventually converted to protocatechuate and ultimately leads to the formation of tricarboxylic acid (TCA) cycle intermediates for energy and biomass (6). Moreover, both classes of PDOs are Rieske oxygenases (RO).

ROs are a family of enzymes best known for catalyzing the NAD(P)H-dependent dihydroxylation of aromatic compounds pollutants (10, 11). However, they catalyze a range of oxidation chemistries and have been the focus of extensive investigation, in part because of their potential use in bioremediation and biocatalysis (12, 13). RO systems comprise two or three components: an oxygenase and a reductase that transfers electrons from NAD(P)H to the oxygenase, either directly or via a ferredoxin (14). The oxygenase contains a mononuclear iron where catalysis occurs and a Rieske-type iron-sulfur cluster ([2Fe-2S]) that mediates electron transfer to the catalytic center. Structural studies on ROs such as naphthalene dioxygenase (NDO) (10, 15–19) have identified key features that are responsible for substrate specificity and have provided valuable insight into the catalytic mechanism of these enzymes (12). For example, a small N-terminal domain harbors the [2Fe-2S] cluster while the catalytic center occurs in a larger C-terminal domain (20). All ROs characterized to date are either $α_3$ trimers or have an additional small subunit to form an $α_3β_3$ hexamer.

The PDO systems of B. cepacia DB01 and C. testosteroni KF1 comprise two components: an oxygenase and a reductase, PDR, encoded by phtA and phtB, respectively. PDR contains a flavin mononucleotide and a ferredoxin-type [2Fe-2S] center (21). Work on PDO$_{DB01}$ pioneered our understanding of RO function (11, 20–34). For example, these studies were the first to establish that ROs are iron-dependent (20). Similarly, ENDOR spectroscopy of the PDO Rieske center provided the

* For correspondence: Pravindra Kumar, pravindra.kumar@bt.iitr.ac.in.
Figure 1. Phthalate catabolic pathway and phylogeny of select ROs. A, catabolism of phthalate to protocatechuate by the Pht enzymes of C. testosteroni KF1. PhtA and PhtB correspond to PDO and PDR, which catalyze the 4,5-dihydroxylation of phthalate. Protocatechuate is catabolized to TCA cycle intermediates. B, phylogenetic relationship of PDOs. Unrooted phylogenetic tree was calculated using a sequence alignment of 20 Rieske oxygenase α-subunits. The proteins are abbreviated using the protein name and strain as follows (GenPeptID): biphenyl dioxygenase from Comamonas testosteroni B356 (BPDOB356, AAC44526), Paraburkholderia xenovorans LB400 (BPDOB3400, AAB63425), and Rhodococcus jostii RHA1 (BPDOB1101, BAA06868); ethylbenzene dioxygenase from R. jostii RHA1 (EBDORHA1, BAC92718); naphthalene dioxygenase from Pseudomonas putida 9816-4 (NDODPp, P0A110), and Rhodococcus sp. NCIMB 12038 (NDODRho, AAD28100); terephthalate dioxygenase from Comamonas testosteroni E6 (tPDOE6, AU48578), and R. jostii RHA1 (tPDOE6, ABH00392); benzoate dioxygenase from Pseudomonas putida (BODOPp, WP_011600771); carbazole dioxygenase from Sphingomonas sp. GTIN11 (CarDOGTIN11, AAD45326).
first evidence that one of the cluster iron has bis-thiolate coordination while the other has bis-imidazole coordination (31). Additional spectroscopic studies of PDO revealed that substrate binding results in a decrease in the coordination number of the mononuclear iron center from six to five (26, 33).

Despite extensive efforts, the structure of PDO has been elusive. Based on gel filtration chromatography, the enzyme was initially proposed to be an α4 tetramer (20). Further analysis, including analytical centrifugation and mass spectrometry, indicated that PDO is a hexamer comprising two stacked α3 trimers, a configuration that has not been shown in crystal structures of ROs to date (24). Moreover, ROs share less than 20% amino acid sequence identity with other structurally characterized ROs. Thus, it has been difficult to identify key residues beyond the metal ligands and an acidic residue (28) proposed to mediate the interaction between the Rieske cluster and the mononuclear iron (35). An atomic-resolution structure of PDO would address a major gap in our understanding of these enzymes and provide insights into the molecular determinants of substrate specificity and regiospecificity.

Herein, we present a biochemical and structural characterization of PDO. Steady-state kinetics were used to evaluate the enzyme’s apparent specificity for phthalate. Crystal structures of PDO were solved for the resting state enzyme as well as in complex with phthalate and terephthalate. We also report the structure of a homolog, ROCH34. The structures reveal that PDO has unique structural features that rationalize the enzyme’s function. The roles of key residues in substrate specificity and catalysis were evaluated using site-directed mutagenesis. Additionally, the structures were compared to those of related oxygenases to understand the dihydroxylation reaction catalyzed by other ROs.

Results

Bioinformatic analysis and phylogenetic relationship of PDO

Phylogenetic analysis of a variety of ROs and characterized ROs revealed that PDO is closely related to PDO, with which it shares 81% amino acid sequence identity (Fig. 1B). These two enzymes form a larger clade with isophthalate dioxygenases iPDO and from C. testosteroni E6 (36) and iPDO from Hydrogenophaga intermedia, a 3-chlorobenzoate dioxygenase, ChaDO, from C. testosteroni BR60 (37) as well as ROCH34, an RO of unknown function with which PDO shares ~35% amino acid sequence identity. All of these enzymes are α-type ROs (20, 36). By contrast, the actinobacterial phthalate 3,4-dioxygenases, such as PDO, are αβ-type enzymes. Accordingly, these ROs are more closely related to other αβ-type ROs such as NDO and biphenyl dioxygenase (BPDO) (38–40). The terphenyl dioxygenases tPDO from C. testosteroni E6 (41) and tPDO from R. jostii RHA1 (42) are also αβ-type ROs and cluster accordingly. Finally, the sequence alignments indicate that in PDO, the mononuclear iron is coordinated by His181, His186, and Asp343, the Rieske center is coordinated by Cys70, His72, Cys89, and His92 and that Asp178 bridges the metalloccenters (Fig. S1). This is the same residue numbering as in PDO.

Structure of phthalate dioxygenase

PDO and ROCH34 were heterologously produced as His-tagged (Ht-) proteins for structural studies. In addition, PDO was produced as an untagged protein for kinetic studies as it had higher specific activity. Ht- and PDO were purified to ~99% and ~95% apparent homogeneity, respectively (Fig. S2A), at yields of up to 20 mg protein per liter culture. Preparations of Ht-PDO and PDO contained 2.9 ± 0.2 and 2.7 ± 0.5 iron per monomer. Ht-PDO had a molecular weight of ~300 kDa according to size exclusion chromatography (SEC), consistent with a homohexameric quaternary structure, while ROCH34 had a molecular weight of ~150 kDa, consistent with a homotrimer (Fig. S2B). We also heterologously produced PDR and RO-RCH34, the cognate reductases of PDO and ROCH34 for activity assays.

The ability of the enzymes to catalyze the dihydroxylation of phthalate was evaluated using an HPLC assay. When equimolar amounts of PDO and PDR were incubated with 100 μM phthalate (retention time, tR = 2.26 min) in the presence of 1 mM NADH (0.1 M Tris, pH 8), phthalate was consumed accompanied by the formation of a product with retention time (Fig. 2A) and an m/z value of 194.95 (Fig. S3A). The m/z value of phthalate was 158.92. Overall, the tR and an m/z values are consistent with the product being cis-4,5-dihydrodiol phthalate (20). The identity of the product was further confirmed by its transformation to protocatechuate by PhcCD as described below. Product formation was dependent on the presence of NADH. Moreover, the stoichiometry of NADH oxidized to phthalate hydroxylated was 1 ± 0.4:1 (Table 1). By contrast, ROCH34 did not detectably transform phthalate (data not shown).

In addition to phthalate, PDO catalyzed the NADH-dependent transformation of terephthalate. The product’s retention time (Fig. 2B) and m/z value (Fig. S3B) are consistent with the product being cis-1,2-dihydrodiol terephthalate (41). However, the dihydroxylation of terephthalate was not well coupled to NADH oxidation: the stoichiometry of NADH oxidized to terephthalate hydroxylated was ~5:1. Finally, ROCH34 did not detectably transform terephthalate and neither

(CarDOGhin, AAL37976); 3-ketosteroid 9α-hydroxylase from Rhodococcus rhodochrous DSM43269 (Ksh, B05MS34362, B06V065); dicamba monoxygenase from S. maltophilia (DMO, AAVS3699); phthalate dioxygenase from Comamonas testosteroni KF (PDO, E670706), Rhodococcus jostii RHA1 (PDO, B0601) and Burkholderia cepacia (PDO, WP_011881604); 3-chlorobenzoate dioxygenase from Comamonas testosteroni BR6 (CbaDO, Q44256); phenoxbenzoate dioxygenase from Pseudomonas oleovorans (PooDO, Q52185); RO from Cupriavidus metallidurans CH34 (RO, AB12407); isophthalate dioxygenase from Comamonas testosteroni E6 (PDO, BAHT0269), and Hydrogenophage intermedia (PDO, CDN0362). Structurally characterized proteins are marked with red.
enzyme transformed isophthalate, 3-chlorobenzoate, 2-chlorobenzoate, 3-phenoxybenzoate, or 4-phenoxybenzoate (data not shown). The ability of PDOKF1 to transform phthalate and terephthalate was confirmed using recombinant E. coli BL21 cells. Briefly, cells harboring the phthalate catabolic genes, phtABCD, converted phthalate to protocatechuate while control E. coli BL21 did not (Fig. 2E). The identity of protocatechuate was assigned with its retention time (Fig. S3C). Cells harboring PhtABCD also transformed terephthalate to the presumed dihydrodiol, indicating that PhtC is unable to catalyze the dehydrogenation of cis-1,2-dihydrodiol terephthalate.

We next used an oxygraph assay to evaluate the apparent steady-state kinetic parameters of PDOKF1 in air-saturated buffer. PDOKF1 catalyzed oxygen consumption in the presence of NADH and phthalate (50 mM Tris, pH 8.0, at 25°C).

**Figure 2. Biochemical analyses.** A, conversion of phthalate by PDOKF1 and PDRKF1. Purified enzymes (10 μM each) were incubated with 100 μM phthalate in the presence of 100 μM NADH. The reaction product was analyzed by HPLC. B, conversion of terephthalate by purified PDOKF1 and PDRKF1. Purified enzymes (10 μM each) were incubated with 500 μM terephthalate in the presence of 100 μM NADH. C, dependence of initial velocity of oxygen consumption on the phthalate concentration in air-saturated buffer. Red lines represent fits of the Michaelis–Menten equation to the data. D, dependence of initial velocity of oxygen consumption on the terephthalate concentration in air-saturated buffer. Red lines represent fits of the Michaelis–Menten equation to the data. E, conversion of phthalate and terephthalate to protocatechuate and cis-dihydrodiol terephthalate, respectively with recombinant E. coli BL21, named R1 cells containing phthalate catabolic genes (phtABCD encoding PDO, PDR, phthalate cis-4,5-dihydrodiol dehydrogenase, and cis-4,5-diol phthalate decarboxylase, respectively).
concentration (Fig. 2, C and D). However, the apparent specificity ($k_{cat}/K_m$) of PDO$_{KF1}$ for phthalate was over 25-fold higher than for terephthalate and the apparent $K_m$ value was ~20-fold lower (Table 1). Considering these experiments measured oxygen consumption, and terephthalate turnover was poorly coupled to NADH consumption, the true difference in apparent specificity is likely greater than 100-fold.

The crystal structure of RO$_{CH34}$

Although we were unable to identify a substrate for RO$_{CH34}$, solving its structure provided a means of obtaining structures of PDO$_{KF1}$. RO$_{CH34}$ was crystallized and solved to 1.8 Å resolution with single-wavelength anomalous dispersion (SAD) phasing using Fe as anomalous scatterer. The crystals belonged to the P6$_3$ space group and contained a single monomer in the asymmetric unit. The overall statistics of data collection, refinement procedures, and protein stereochemistry are within the standard values (Table 2). The crystallographic symmetry of RO$_{CH34}$ suggests a trimeric architecture ($3_2$ arrangement) with each subunit related to another subunit by a non-crystallographic threefold axis (Fig. 3A). Each RO$_{CH34}$ protomer has 437 residues and structurally can be divided into the Rieske domain and the catalytic domain (Fig. 3B). The Rieske domain in RO$_{CH34}$ (residues 26–153) consists of eight α-helices and three β-sheet structures with a total of eight β-strands. In the Rieske center, Fe1 is coordinated with Cys69 and Cys88, at 2.2 Å and 2.3 Å distances, respectively, while Fe2 is coordinated with His71 and His91 at distances of 2.1 Å each. The two sulfide ions bridge the two iron ions and form a flat rhombic arrangement. The catalytic domain is dominated by an eight-stranded antiparallel β-sheet that extend over ten α-helices. The active site is formed by two helices (α9 and α11) and a β-sheet from the top. Three loops (LI, 246–253; LII, 291–330; LIII, 186–199) form the outer entrance of the cavity from three sides. The mononuclear iron-containing active site groove consists of residues Ser179, Glu176, His181, His186, Arg233, Asn235, and Asp349 (Fig. 3B). The mononuclear Fe atom is coordinated at 1.9 Å and 2.0 Å distances to His181 and His186, respectively, as well as bidentately to Asp349 with 2.0 Å and 2.5 Å distances (Fig. 3B). Additionally, a glutamate molecule, which originates from the mother liquor, binds the mononuclear iron in a bidentate manner (1.8 Å and 2.5 Å from each of the carboxylate oxygen atoms), completing its coordination sphere (Fig. 3B). However, the two coordination sites of the GLU are usually occupied by solvent molecules, as observed in other RO structures (16, 43).

---

Table 1

| Substrate       | $k_{cat}$ (s$^{-1}$) | $K_m$ (μM) | $k_{cat}/K_m$ (mM$^{-1}$s$^{-1}$) | Coupling $^b$ |
|-----------------|---------------------|------------|---------------------------------|----------------|
| Phthalate       | 2.1 ± 0.1           | 3.6 ± 0.4  | 583 ± 87                        | 1.0 ± 0.4      |
| Terephthalate   | 2.0 ± 0.1           | 90 ± 12    | 22 ± 4                          | 5.0 ± 0.2      |

$^a$ Experiments were performed using air-saturated 50 mM Tris, pH 8.0, at 25 °C. The reported parameters are based on oxygen consumption.

$^b$ Relative amount of NADH oxidized per mol of aromatic substrate hydroxylated.

---

Table 2

| Data processing and refinement statistics | RO$_{CH34}$ | PDO$_{KF1}$-native | PDO$_{KF1}$-phthalate | PDO$_{KF1}$-terephthalate |
|------------------------------------------|------------|---------------------|-----------------------|---------------------------|
| Resolution range                         | 38.8–1.84  | 88.87–2.11          | 53.29–2.74            | 53.06–3.07                |
| Wavelength (Å)                           | 1.74       | 0.96                | 0.68                  | 0.69                      |
| Space group                              | P 63       | C 1 2 1             | C 1 2 1               | C 1 2 1                   |
| Unit cell dimensions                     | 103.4 103.4| 178.2 121.4         | 179.0 122.2           | 179.2 122.7               |
| α β γ (°)                                | 90 90 120  | 90 100.9 90        | 90 101.1 90           | 90 101.1 90               |
| Completeness (%)                         | 99.9       | 98.84               | 97.88                 | 97.35                     |
| $R_{merge}$ (%)                          | 3.7        | 10                  | 6                     | 8.8                       |
| I/σ(I)                                   | 9.4        | 3.9                 | 12.6                  | 8.6                       |
| Refinement                               | 40.967     | 191.509             | 88.624                | 63.340                    |
| Reflections used in refinement           | 2026       | 9784                | 4362                  | 3147                      |
| R-work $^c$                              | 0.16       | 0.17                | 0.2                   | 0.26                      |
| R-free $^c$                              | 0.21       | 0.22                | 0.28                  | 0.3                       |
| Number of non-hydrogen atoms             | 3875       | 22.171              | 30.545                | 19.601                    |
| Macromolecules                           | 3508       | 19.389              | 19.274                | 19.311                    |
| Ligands                                  | 27         | 154                 | 42                    | 42                        |
| Solvent                                  | 340        | 2628                | 1229                  | 248                       |
| Protein residues                         | 437        | 2459                | 2446                  | 2454                      |
| RMS (bonds (Å))                          | 0.011      | 0.013               | 0.008                 | 0.008                     |
| RMS (angles (°)) $^c$                    | 1.78       | 2.11                | 1.58                  | 1.6                       |
| Favored (%)                              | 97.46      | 94.99               | 91.54                 | 88.39                     |
| Allowed (%)                              | 2.51       | 4.52                | 7.5                   | 9.64                      |
| Outliers (%)                             | 0.23       | 0.49                | 0.95                  | 1.98                      |
| Average B-factor (Å)                     | 43.99      | 37.7                | 50.73                 | 44.09                     |
| Macromolecules (Å)                       | 43.31      | 36.29               | 51.03                 | 44.36                     |
| Phthalate/terephthalate (Å)              | 77.25      | 77.25               | 78.5                  | 43.8                      |
| Ligands (Å)                              | 60.09      | 44.1                | 47.51                 | 46.16                     |
| Solvent (Å)                              | 49.72      | 47.69               | 46.16                 | 22.59                     |

$^a$ $R_{merge} = \sum |<F_{obs}> - |F_{calc}>|/|F_{calc}>.$

$^b$ $R = \sum |F_{calc}> - |F_{obs}>|/|F_{calc}>.$ The $R_{free}$ is the $R$ calculated on the 5% reflections excluded for refinement.

$^c$ RMS is root mean square.
The structure of PDOKF1

PDOKF1 was crystallized and its structure was determined at 2.1 Å resolution using molecular replacement and the ROCH34 structure as a template. The overall data collection and refinement statistics are within the standard values (Table 2). The PDOKF1 crystal belongs to the C2 space group and contained six monomers in the asymmetric unit. The overall crystallographic symmetry operation shows that PDOKF1 is hexameric, consistent with the SEC measurement. The hexamer has a star-shaped arrangement in which one α3 is stacked on top of, and 60° off-set from, the other (Fig. 4B). The modeled PDOKF1 protomer can be divided into the Rieske domain and the catalytic domain (Fig. 4A). Typical head-to-tail interaction interface between the subunits consists of many charged residues that form up to 28 hydrogen bonds and 16 salt bridges as calculated by PISA-CCP4i (44). The overall structure of PDOKF1 is very similar to that of ROCH34, with an RMSD of 0.98 Å over 274 Ca atoms. The Rieske domain of PDOKF1 is very similar to that of other α3 type ROs despite sharing less than 20% amino acid sequence identity with them over this domain. For example, the Rieske domains of PDOKF1 and DMO (45) have an RMSD of 1.0 Å over 90 Ca atoms (Fig. 4D). More specifically, the Rieske domain of PDOKF1 consists of five α-helices and three beta-sheet structures with a total of ten β-strands. The Fe1 of the Rieske cluster is coordinated with Cys70 and Cys89, at distances of 2.3 Å each. For its part, Fe2 is coordinated with His72 and His92 at distances of 2.2 Å each. The two sulfide ions bridge the two iron ions and form a flat rhombic arrangement typical of ROs (Fig. 4C). Within a single protomer, the mononuclear iron is located at 44 Å from the Rieske center (nearer iron). However, as typical of ROs, it is only at 12 Å (nearer iron) from the mononuclear iron of the adjacent α subunit within the trimer (39, 45, 46). The mononuclear iron is also seen to be very distant (44 Å) from the Rieske center of the stacked protomer (Fig. S4). Finally, an N-terminal helix extension present in CarDO, which mediates interactions with a ferredoxin, is absent in PDOKF1, consistent with the absence of this component in the PDOKF1 system.

In contrast to the Rieske domain, the catalytic domain of PDOKF1 shows important differences with respect to that of other ROs (Fig. 4E). Most strikingly, this domain contains 286 residues (154–439), significantly more than any other...
structurally characterized RO. For example, the catalytic domains of DMO and CarDO comprise 220 and 240 residues, respectively. Much of the additional size is due to an extended region that contains protruded helices and loops (Zone I and II in Fig. 4E). Together, these protrusions form a bipod-like structure and enable the stacking of each α protomer with two protomers from the neighboring trimer (Fig. S5A). More specifically, this extended region comprising residues 285 to 338 forms two helices, each of which interacts with a separate protomer (α4 and α5). These interactions include a hydrogen bond and a salt bridge formation (Fig. S5B). The residues that mediate trimer stacking in PDOKFI are largely conserved among proteobacterial PDOS, including PDODBO1 (Fig. S1), consistent with the latter being a homohexamer (24). ROCH34 harbors residues (292–300) that align with the extended region (Fig. 4E); however, they adopt a different structure than that of PDOKFI (Fig. S5C). Additionally, the residues constituting this helix are not conserved. Especially, ROCH34 lacks Glu293 and Thr294, which form critical intertrimer interactions in PDOKFI (Fig. S1). Overall, these differences are consistent with ROCH34 being a trimer rather a homohexamer.

Other than the insertions that mediate trimer stacking, the catalytic domain is similar to that of other ROs. Briefly, the domain comprises V-shaped β-sheets formed from eight
antiparallel β-strands. These sheets are surrounded by nine α-helices (Fig. 4A). The active site itself is formed by two α-helices and two β-sheets facing each other. A broad channel extends from the surface of the protein, providing substrates with access to the catalytic mononuclear iron. The mononuclear iron ion is coordinated to His181 and His186 at 2.1 Å each and is bidentate to Asp343 with 2.2 Å and 2.3 Å distances (Fig. 4D). Unlike ROCH34, the last two coordination sites of mononuclear iron are occupied by two water ligands, W1 and W2, at distances of 2.1 Å and 2.0 Å, respectively (Fig. 4D). As described previously, the coordination sphere of the mononuclear iron is conserved within ROs (38, 46). Asp178 bridges the mononuclear iron and the Rieske center of the neighboring subunit (Fig. 4C) as reported in PDO_DB01 (28). More specifically, Asp178 forms a hydrogen bond with His181, a mononuclear iron ligand, and His91 of the neighboring subunit where it coordinates the FeS cluster (Fig. 4C). The substrate-binding pocket is defined by Arg207, Arg231, Ser179, and Arg244 on one face and Phe278, Phe280, and Phe339 on the other face. These residues interact with phthalate as described below and are conserved in proteobacterial PDOs, including PDO_DB01 (Fig. S1).

The structure of PDO_kf1 complexes with phthalate and terephthalate

To identify residues responsible for binding phthalate, we soaked crystals of PDO_kf1 with the substrate and determined a structure of the complex to 2.7 Å resolution. The main chain conformation in the PDO_kf1:phthalate complex is nearly identical to that of substrate-free PDO_kf1 (RMSD 0.24 Å over 366 Ca atoms). The phthalate molecule could be modeled in one chain of PDO_kf1 structure that matched the observed electron density with a real space correlation coefficient (RSCC) of 0.84 (Fig. 5A). In contrast to the distorted octahedral geometry of mononuclear iron in the substrate-free

Figure 5. Structure of phthalate and terephthalate bound PDO_kf1 complexes to reveal the protein interactions that are important for correctly positioning substrate. A, electron density omit maps (contoured at 3σ) at 2.7 Å, 3.1 Å resolution for phthalate and terephthalate, respectively in the active site. B, the coordination state of the mononuclear iron in presence of phthalate and terephthalate. C, the phthalate molecule is anchored in the active site of PDO_kf1 via interactions that are highlighted here. D, similarly, the terephthalate molecule is anchored in the active site of PDO_kf1 via interactions that are highlighted here. The interactions are represented with black dotted lines. Residues and bond distances (in Å) are labeled.
enzyme, the mononuclear iron is pentacoordinate in the PDOKF1:phthalate complex (Fig. 5B), as observed in case of PDOPDB01 with magnetic circular dichroism (26, 33). Phthalate is positioned in the active site such that the carbon atoms that are dihydroxylated, C4 and C5, are 4.3 Å and 4.5 Å, respectively, away from mononuclear iron, consistent with earlier reported (32) distance (4.3 Å–6.5 Å) between phthalate and mononuclear iron. A solvent ligand is also present, between the substrate and the mononuclear iron, 2.9 Å and 3.0 Å from C4 and C5 carbon atoms, respectively (Fig. 5B). There is sufficient space between the metal ion and the substrate for O₂ to displace the solvent ligand as suggested for NDO and other ROs (15, 47). Interestingly, Arg244 interacts with both carboxylate groups of the phthalate molecule, being positioned at 2.7 Å and 3.2 Å from C1 and C2 carboxylate groups, respectively. The C2 carboxylate also interacts with the side chain of Arg207, at a distance of 3.0 Å (Fig. 5C). These carboxylate groups are also hydrogen-bonded to the side chain hydroxyls of Ser179 and Ser182. These salt bridges and hydrogen-bond networks indicate that the carboxylate groups are critical binding determinants. Finally, Phe280 and Phe339, located on the other face of the active site, form π–π stacking interaction with the substrate’s aromatic ring (Fig. 5C).

The key phthalate-binding residues are conserved among PDOs from proteobacterial strains (Fig. S1). However, these residues are not all conserved in ROCH34. Thus, Ser182, Arg207, Phe280, and Phe339 correspond to Ala182, Trp201, Met271, and Leu346, respectively, in ROCH34 (Fig. S1). These substitutions are consistent with ROCH34’s inability to turnover phthalate. Finally, the sequence analysis reveals that Arg207 and Arg244 of PDOKF1 correspond to Arg218 and Met283, respectively, in PDOHRA1 as discussed below. This highlights the different architectures of the PDOs that catalyze 4,5-dihydroxylation and 3,4-dihydroxylation, respectively.

To investigate the structural basis of why phthalate is a better substrate for PDOKF1 than terephthalate, we soaked PDOKF1 crystals with terephthalate and solved the structure of the resulting complex to 3.1 Å. In the PDOKF1:terephthalate structure, the ligand terephthalate also could be modeled in one chain of the PDOKF1 structure and has an RSCC value of 0.87. Like the PDOKF1:phthalate complex, the structure of the PDOKF1:terephthalate complex is very similar to that of the substrate-free enzyme (RMSD 0.26 Å over 367 Ca atoms). Interestingly, no density corresponding to a metal-bound solvent or O₂ species was observed, perhaps due to the low resolution. The terephthalate molecule is positioned in the active site such that the carbon atoms that are dihydroxylated, C1 and C2, are at 3.4 Å and 3.6 Å, respectively from the mononuclear iron (Fig. 5B). Notably, the orientation of the terephthalate in the active site is consistent with that required for productive catalysis. Moreover, several of the interactions between the enzyme and the substrate observed in the PDOKF1:phthalate structure are conserved in the PDOKF1:terephthalate structure. The substrate’s C4 carboxylate interacts with the side chains of Arg244 and Arg207 at distances of 2.8 Å and 3.0 Å, respectively, analogous to the interactions between the C2 carboxylate of phthalate and these residues (Fig. 5D). Similarly, the aromatic ring of terephthalate forms a nonpolar interaction with Ile256 and π–π stacking interactions with Phe339 and Phe280 (Fig. 5D), as observed for phthalate. The striking difference between the bound substrates concerns the second carboxylate: in the case of terephthalate, the C1 carboxylate does not form hydrogen bonds or salt bridges with any residues. The markedly fewer interactions are consistent with the lower specificity of the enzyme for terephthalate.

**Active site variants**

The roles of active site residues in determining the substrate specificity of PDOKF1 were probed using site-directed mutagenesis. More specifically, we generated the R207A and R244A variants of the enzyme and purified them as described for the wild-type enzyme. Preparations of both variants contained 2.9 ± 0.3 Fe per monomer. Using the HPLC-based assay, neither variant detectably transformed phthalate, indicating that these two residues are critical determinants of substrate specificity (Fig. 6).

![Figure 6. Site directed mutagenesis and additional activity of PDOKF1.](image-url)

**Figure 6. Site directed mutagenesis and additional activity of PDOKF1.** Reactions of PDOKF1 wild-type and variants with phthalate. Purified enzymes (10 μM each) were incubated with 10 μM PDOKF1, 100 μM phthalate in the presence of 1 mM NADH. The reaction product was analyzed by HPLC.
Structure of phthalate dioxygenase

Modeling of PDO_{RHA1} and iPDO_{E6}

To better understand the structural features that determine the regiospecificity of PDOs, we modeled PDO_{RHA1}, an enzyme that catalyzes 3,4-dihydroxylation, and compared it with the structure of PDO_{KF1}. A homology model of PDO_{RHA1} generated using the structure of NDO_{Rho} (19) has an RMSD of 0.15 Å over 325 aligned Cα atoms with the structure of NDO_{Rho}. Despite significant architectural differences between actinobacterial PDOs and proteobacterial PDOs, the active site features are similar in terms of their constituent residues. As in PDO_{KF1}, the PDO_{RHA1} active site features basic residues on one face and nonpolar residues on the other, positioned to interact with the phthalate's carboxyl moieties and aromatic ring, respectively (Fig. 7A). Consistent with enzyme's regiospecificity, docking phthalate into the active site using AutoDock Vina yielded a pose with the atoms to be hydroxylated, C3 and C4, positioned 3.4 Å and 3.1 Å, respectively, from a water molecule, which is at 2.2 Å distance from mononuclear iron and occupies the presumed binding site for dioxygen (Fig. 7A). More specifically, Arg218, which corresponds to Arg207 in PDO_{KF1}, interacts with the C2 carboxylate group of phthalate, while Lys224 and Lys242 residues interact with the C1 carboxylate. These interactions orient the phthalate molecule for its regiospecific 3,4-dihydroxylation (Fig. 7A). Comparison with the PDO_{KF1}:phthalate complex revealed that Arg244 in PDO_{KF1} is replaced with Met283 in PDO_{RHA1}, and that this prevents the orientation of the substrate to enable 4,5-dihydroxylation (Fig. 7B).

To gain insight into the active site features of iPDO_{E6}, we modeled the enzyme's structure using the ROCH_{34} coordinates. The homology model of iPDO_{E6} has an RMSD of 0.19 Å over 309 aligned Cα atoms. Comparison with the structures of PDO_{KF1} and ROCH_{34} revealed significant similarities and differences in the active sites. First, Arg244, which interacts with a phthalate carboxylate in PDO_{KF1}, is conserved in iPDO_{E6} (Arg234), and ROCH_{34} (Arg233) (Fig. 7, D and E). The iPDO_{E6} active site also harbors His257 that is positioned to interact with the meta-carboxylate group of the isophthalate. In the modeled structure, this residue corresponds to Asn266 of PDO_{KF1}. Importantly, iPDO_{E6} does not have a second Arg in the active site to bind an ortho-carboxylate. Finally, Val178, Phe249, and Phe333 are positioned to stabilize the aromatic ring of isophthalate. These residues correspond to Ser182, Arg207, and Phe339 of PDO_{KF1}, respectively (Fig. 7, D and E). This analysis indicates that Arg234 and His257 are critical determinants for the specificity of iPDO_{E6}. Interestingly, ROCH_{34} also harbors His257. However, the residues predicted to interact with the substrate's aromatic ring in iPDO_{E6}, Val178, Val245, Phe249, and Phe333, correspond to Ala182, Ile246, Gln248, and Leu345 in ROCH_{34}, suggesting this is a major reason why ROCH_{34} is unable to turnover isophthalate (Fig. 7E).

Discussion

The structural characterization of PDO_{KF1} validates and extends pioneering work done on PDO_{DB01}. Specific features of PDO_{DB01} that are validated by the structure of PDO_{KF1} include the hexameric configuration of the former (24) and the bridging of the metallocenters by Asp178 (35). More particularly, because all functionally important residues are conserved between these two enzymes, the structure of PDO_{KF1} provides a framework for characterizing the molecular determinants of PDO_{DB01}'s properties. For example, PDO_{DB01} contains the regions and specific residues that stabilize hexamer formation in PDO_{KF1}, and these residues are unique to this clade of PDOs. Accordingly, we predict that the PDO_{DB01} hexamer adopts the same star-shaped structure as the PDO_{KF1} hexamer. Similarly, all the substrate-binding residues of PDO_{KF1} are conserved in PDO_{DB01}, including Arg207 and Arg244, providing a basis for understanding the substrate specificity and regiospecificity of PDO_{DB01}.

The structural data together with the functional characterization of the variants establish that Arg244 and Arg207 are critical determinants of substrate specificity in the proteobacterial PDOs. The structure of the PDO_{KF1}:phthalate complex revealed that these two residues on one face of the active site form salt bridges with carboxylate groups of phthalate. These residues are conserved among proteobacterial PDOs, highlighting their importance in correctly positioning phthalate for its regiospecific hydroxylation. Further, the positioning of these residues is such that they are unable to properly orientate either terephthalate or isophthalate in the active site, as indicated by PDO_{KF1}'s poor turnover of terephthalate and inability to turn over isophthalate.

The data for the PDO_{KF1} and PDO_{KF1}:phthalate structures corroborate the spectroscopic data showing the substrate-dependent change in the coordination geometry of PDO_{DB01}. Magnetic circular dichroism spectroscopic analyses indicated that in substrate-free PDO_{DB01}, the mononuclear iron is hexacoordinate and becomes pentacoordinate upon substrate binding (26, 33). Moreover, NMR evidence showed that the dissociating ligand is a water molecule (34). Consistent with these, the data presented here establish that the mononuclear iron exists in a distorted octahedral six-coordimate geometry in the resting state PDO_{KF1} (Fig. 3B) and is five-coordinate in the PDO_{KF1}:phthalate complex (Fig. 5B). The mononuclear iron is six-coordinate in the resting state of NDO as well (48). The change in coordination state induced by substrate binding is thought to be mechanistically important as it enables the mononuclear iron to bind O2 in a side-on manner (13). Additionally, in PDO_{DB01}, mononuclear iron was reported (11) to be in the ferrous state at the end of the reaction during single turnover experiments. Based on this, it was proposed that two electrons for the reaction are provided by two Rieske centers and that the resulting stoichiometry is one product molecule produced per two Rieske centers oxidized (24). This phenomenon was proposed to be a consequence of a stacked trimer architecture placing two Rieske centers close to each mononuclear iron, facilitating electron transfer. The PDO_{KF1} structure demonstrates that the stacked trimer architecture does not position a second Rieske center sufficiently close to the mononuclear iron (Fig. S4) to enable transfer of second electron. However, with the current lack of biochemical and
The terephthalate-transforming activity of PDOKF1 is unlikely to be physiologically relevant. First, the dihydroxylation of terephthalate was not well coupled with NADH oxidation, indicating that the active site of PDOKF1 does not accommodate terephthalate well. This is confirmed by the structure of the PDOKF1:terephthalate complex, which reveals that one of the two Rieske centers for one mononuclear iron cannot be proven.

Figure 7. Structural comparison and active site analysis of PDORHA1 and iPDOE6. A, active site of PDORHA1 with docked phthalate molecule, shown in salmon-colored stick. Residues and bond distances (in Å) are labeled. B, superposition of the structures of PDORHA1 and PDOKF1 (green). C, stereoview of superposition of the structures of PDORHA1 and PDOKF1. D, superposition of the structures of iPDOE6 and PDOKF1 (green). The isophthalate docked model of iPDOE6 is shown in violet-colored stick. E, superposition of the structures of iPDOE6 and ROCH34 (Cyan). The residues are labeled. F, stereoview of superposition of the structures of iPDOE6 and PDOKF1. G, stereoview of superposition of the structures of iPDOE6 and ROCH34.
the terephthalate carboxylates does not form any salt bridges with active site residues. Second, recombinant cells containing the phthalate catabolic pathway transformed terephthalate to the cis-dihydrodiol. While this confirms that PDO_{K1} can dihydroxylate terephthalate, it also established that the subsequent phthalate catabolic enzymes are unable to transform the cis-diol product to protocatechuate. Finally, as *C. testosteroni* K1 possesses a bona fide terephthalate catabolic pathway (41), terephthalate hydroxylation activity of PDO_{K1} is expected to not be physiologically relevant.

The structural comparison of PDO_{K1} with a homology model of PDO_{RHA1} highlights possible determinants of the regiospecificity of PDO_{RHA1}. Specifically, while the overall structures of PDO_{RHA1} and PDO_{K1} are different, most of the active site residues are identical. Notably, Arg207, a key phthalate-binding residue in PDO_{K1}, corresponds to Arg218 in PDO_{RHA1}. An important difference in the active sites of these enzymes is Met283 in PDO_{RHA1}, which corresponds to Arg244 of PDO_{K1}. In PDO_{RHA1}, the phthalate binding pose equivalent of PDO_{K1} would have a steric conflict between the phthalate and Met283 and thus prevent the phthalate molecule to get 4,5-dihydroxylated. Additionally, the interactions of Arg218, Lys224, and Lys242 in PDO_{RHA1} with the C1 and C2 carboxylates of phthalate are critical to orientating the substrate for 3,4-dihydroxylation.

While the physiological role of RO_{CH34} is unclear, the functional and structural data are consistent with it not being either a PDO or an iPDO. The crystal structure of RO_{CH34} depicted common features of ROS related to mononuclear iron-binding residues and electron transfer residues. However, the structural analysis indicated that active site residues in RO_{CH34} differ from those of PDO_{K1}. More specifically, the lack of Arg207, Phe280, and Phe339 in RO_{CH34} likely explains this enzyme’s inability to transform phthalate. Comparison of the crystal structure of the PDO_{K1} complex to the modeled iPDO_{E6} structure revealed structural features that explain the different reactivities of these two enzymes. Most strikingly, the occurrence of polar residue His257 in iPDO_{E6} at the position corresponding to Asn266 of PDO_{K1} is consistent with the specificity of iPDO_{E6} for isophthalate. The RO_{CH34} structure
lacks the equivalent of aromatic ring stabilizing residues Phe333, Phe249, Val178, Ile246 of iPDOE6, likely explaining its inability to bind isophthalate. Unfortunately, the genomic context of the gene encoding ROCH34 (protein_id, ABF12407) provides no insight into the function of the enzyme: the adjacent genes are predicted to encode an alcohol dehydrogenase (ABF12404) and an aldehyde dehydrogenase (ABF12403). Finally, although ROCH34 shares sequence similarity with phenoxybenzoate dioxygenase, ROCH34 did not transform either 3-phenoxybenzoate or 4-phenoxybenzoate.

In conclusion, this study provides important new insights into the dihydroxylation of phthalate by proteobacterial PDOs. The structures and mutagenesis data not only provide insights into the molecular details of PDOKF1, but also provide structural basis for better understanding many years of work on PDODB01. The findings also provide a structural basis for understanding the regiospecific 4,5-dihydroxylation by proteobacterial PDOs versus 3,4-dihydroxylation of actinobacterial PDOs.

### Experimental procedures

#### Chemicals, reagents, and bacterial strains

All reagents were of analytical grade. Restriction enzymes, T4 DNA ligase used for cloning were from New England Biolabs. Phusion polymerase used was from Thermo scientific. Water for buffers was purified using a Merck Synergy Water Purification System to resistance of at least 18.2 MΩ. C. testosteroni KF1 was purchased from the German Collection of Microorganisms and Cell Cultures DSMZ. Cupriavidus metallidurans CH34 was purchased from the Microbial Type Culture Collection (MTCC).

#### Cloning and expression

DNA was purified, propagated, cloned, and amplified using standard protocols (49). Genes encoding RO components were amplified from genomic DNA of C. testosteroni KF1 and megaplasmid DNA of C. metallidurans CH34 using primers listed in Table S1. N- and C- terminal primers contained Ndel and XhoI sites, respectively. The PCR amplicons and vectors (either pET-28c or pET-41b) were digested with the appropriate restriction enzymes, gel purified, and ligated with T4 DNA ligase. The phthalate catabolic gene fragment (phtA encoding PDO, phtB encoding PDR, phtC encoding phthalate cis-4,5-dihydrodiol dehydrogenase, and phtD encoding cis-4,5-diol phthalate decarboxylase) was amplified from genomic DNA of C. testosteroni KF1 using primers listed in Table S1. The amplified fragment was digested with Ncol and HindIII enzymes and cloned into the pET-28c vector. The recombinant vector was transformed into competent E. coli DH5α cells and subsequently to E. coli BL21 λ(DE3) cells. The ligation products were transformed into chemically competent E. coli DH5α cells and spread onto lysogeny broth (LB) agar plates containing kanamycin (50 µg/ml). Multiple single colonies were picked for each construct and grown in 10 ml LB supplemented with kanamycin at 37 °C and 220 rpm overnight. Plasmid DNA was extracted and a positive clone was confirmed by restriction digestion and DNA sequencing. Further, the confirmed plasmid was transformed into chemically either competent E. coli BL21 λ(DE3) cells and spread onto LB agar plates containing kanamycin (50 µg/ml). A single transformed colony was inoculated into 10 ml of LB supplemented with kanamycin (50 µg/ml) and incubated at 37 °C and 220 rpm for 80 h. Cells were grown in LB supplemented with kanamycin (50 µg/ml) at 37 °C and 200 rpm until OD_{600} = 0.6, upon which culture was transferred to 16 °C and induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) for 16 h. Cells were pelleted (6000g, 10 min, 4 °C) and stored at −80 °C.

### Purification of recombinant PDOKF1 and PDRKF1

For kinetic characterization, PDOKF1 and ROCH34 were produced in E. coli BL21 λ(DE3) with pET41PDOKF1 and pET41ROCH34 plasmid constructs, respectively, as the His-tagged form of these enzyme did not show any detectable activity. Cell culture collected from 1 l was resuspended in 10 ml purification buffer (20 mM Tris, pH 7.4) containing 5 mM PMSF, and cells were lysed using Constant Cell Disruptor at 20 psi. Cell debris was removed by centrifugation at 12,000 rpm for 1 h at 4 °C. The PDOKF1 was purified using a MonoQ 10/100 Gl column (GE Healthcare). The protein was eluted with a linear gradient from 0.2 to 0.6 M NaCl in 120 ml of 20 mM Tris, pH 8.0. Fractions containing PDOKF1 and ROCH34 were pooled, dialyzed into 20 mM Tris, pH 7.4, concentrated to ~20 mg/ml, flash-frozen as beads in liquid N2, and stored at −80 °C until needed.

The ROCH34-Ht, RO-RCH34-Ht, PDOKF1-Ht, and PDRKF1-Ht were produced in E. coli BL21 λ(DE3) with pET28ROCH34, pET28RO-RCH34, pET28PDOKF1, and pET28PDRKF1 plasmid constructs, respectively. Cells were grown, cell extracts were produced as described above and resuspended in 10 ml purification buffer (20 mM Tris, pH 7.4, NaCl 300 mM) containing 5 mM PMSF, and cells were lysed using constant Cell Disruptor operated at 20 psi. The supernatant solutions were incubated with the pre-equilibrated Ni-NTA affinity column. Bound proteins were eluted by an imidazole gradient from 100 to 500 mM with purification buffer and analyzed on an SDS-PAGE gel. Fractions containing pure protein bands were concentrated using an Amicon 30 kDa Millipore filter. N-terminal His tag was removed by incubation with TEV protease (1:10 M ratio) in dialysis buffer (100 mM sodium phosphate, pH 7.4; 50 mM NaCl). The cleaved protein mixture was purified by using a Reverse Ni-NTA affinity column. The purity of the protein was checked on SDS-PAGE. Tag-free proteins were concentrated using an Amicon 30 kDa Millipore filter and dialyzed with 50 mM HEPES, pH 7.4, 50 mM NaCl.

### Analytical methods

Protein concentration was determined using the Micro BCA protein assay kit (Pierce) using bovine serum albumin as a standard. UV–Visible spectra were recorded using a Cary 60 spectrophotometer. Iron concentrations were determined
**Structure of phthalate dioxygenase**

spectrophotometrically using the Ferene-S assay and FAS solution as a standard (50).

**Phylogenetic analysis of RO sequences**

Twenty sequences of ROs were selected for phylogenetic analysis based on biochemical characterization and preferred substrate. Sequences of α subunits were aligned using MUSCLE. Phylogenetic reconstruction and validation were performed using the neighbor-joining (NJ) method based on bootstrap analysis with 1000 replications using the Jukes–Cantor distance model in molecular evolutionary genetics analysis (MEGA) 7 software (51). Both maximum-likelihood and minimum-evolution methods were also employed to test the robustness of the tree.

**Biochemical activity**

The PDOF1 activity was determined by measuring the consumption of phthalate by high-pressure liquid chromatography (HPLC). HPLC experiments were carried out on an X-Bridge (Waters) C18 column (5 μm, 150 × 4.6 mm) using a Waters HPLC system equipped with a binary pump (Waters; Binary pump 1525) and photodiode array (Waters; PDA 2998) detector. Chromatograms were analyzed on a Windows 10 platform with an EMPower software ver. 3 feature release 4 (Waters). The 1 ml assay mixture contained 100 mM Tris, pH 8.0 containing 100 μM Fe(NH₄)₂(SO₄)₂.6H₂O, 200 μM substrate, 100 μM NADH, 10 μM each of PDOF1 and PDRF1. Portions (100 μl) of the reaction mixture were removed at various sampling times and quenched with 100 μl acetonitrile (100%) prior to analysis. The mobile phase used for HPLC analysis was a mixture of acidified water (68%) and acetonitrile (32%). The HPLC system was operated at a flow rate of 1 ml/min and the metabolites were detected at 240 nm. For substrate preference activity, PDOF1 was incubated with 100 μM of phthalate, isophthalate, terephthalate, 3-chlorobenzoate, 2-chlorobenzoate, 3-phenoxybenzoate, or 4-phenoxybenzoate under the assay conditions described above, and the reaction mixtures were analyzed by HPLC.

Steady-state kinetic assays were performed by monitoring the consumption of O₂ using a Clark-type polarographic O₂ electrode OXYG1 (Hansatech) connected to a circulating water bath. Assays were performed in 1 ml of air-saturated 50 mM Tris, pH 8.0 at 25 °C, with 0.25 μM each of PDOF1 and PDRF1 and initiated by adding the substrate. The electrode was calibrated according to the manufacturer's instructions using air-saturated water and O₂-depleted water via the addition of sodium dithionite. Reaction velocities were corrected for the background rate of O₂ consumption, recorded prior to aromatic substrate addition. Steady-state kinetic parameters were evaluated by fitting the Michaelis–Menten equation to the initial velocities.

**Site-directed mutagenesis**

PCR amplification using mutagenic primers was used to introduce the desired substitution. In all, 25 μl reactions composed of 1x Phusion buffer (HF), 100 μM dNTPs, 2 ng/μl plasmid template, 1 μM primer (Table S1), 1 U Phusion DNA polymerase, and 1% DMSO. The parent template was digested using DpnI upon the visual confirmation of a successful amplification using agarose gel electrophoresis. The DpnI digest reaction was incubated at 37 °C for 1 h and 10 μl reaction volume was transformed into DH5α chemically competent cells. Mutations were verified by sequencing.

**Whole cell biotransformation of phthalate and terephthalate**

Cell-mediated substrate hydroxylation was detected and analyzed with HPLC. One mL of overnight recombinant BL21 λ(DE3) culture was added to two flasks containing 1 mM phthalate and terephthalate in 50 ml of MINERAL MEDIUM (BRUNNER), and IPTG (1 mM final concentration) was added. Cultures were grown for 4 h at 30 °C and centrifuged, and the supernatant was analyzed by HPLC and the cells were disrupted using Cell Disruptor (Constant Systems) and extracted using acetonitrile (GC-grade). The cell debris was removed with centrifugation. The extracted samples were concentrated to 500 μl and then filtered through 0.22 μm membrane filters. Filtered samples were analyzed with HPLC as described above. The HPLC experiment was carried out on a SunFire (Waters) C18 column (5 μm, 250 × 4.6 mm).

**Crystallization of ROCH₃₄ and PDOF1**

ROCH₃₄ was concentrated up to 18 to 20 mg/ml and subjected to crystallization trials using the sitting drop vapor diffusion method at 20 °C. The initial crystals appeared in the Morpheus screen. The best diffracting crystals were obtained in an optimized condition, which contained 50 mM glycine, 25 mM l-lysine, 20 mM glutamate, and 14% polyethylene glycol 4000 reservoir solution, and protein-to-reservoir ratio of 1:1 was used. ROCH₃₄ crystals were obtained at 20 °C. For data collection, ROCH₃₄ crystals were cryoprotected in a reservoir solution containing 15% ethylene glycol and were flash-frozen in a nitrogen stream at 100 K. PDOF1 was concentrated up to 12 to 15 mg/ml and subjected to crystallization trials using the sitting drop vapor diffusion method at 20 °C. The best diffracting crystals were obtained in an optimized condition that contained 100 mM HEPES buffer pH 7.5, 25 mM CaCl₂, 25 mM MgCl₂, 100 mM sodium potassium tartrate, and 12% polyethylene glycol 4000 as a reservoir solution with protein-to-reservoir ratio of 1:1. Within 2 days, deep red-colored PDOF1 crystals were obtained at 20 °C. For data collection, crystals were cryoprotected in a reservoir solution containing 40% ethylene glycol and were flash-frozen in a nitrogen stream at 100 K. PDOF1 crystals in complex with phthalate and terephthalate were obtained by soaking the crystals in cryoprotectant solutions containing 5 mM phthalate or 5 mM terephthalate for 5 min at 25 °C before freezing and data collection.

**X-ray data collection and structure determination**

Diffraction data were collected at the Home Source, Macromolecular crystallography Unit, IIC, IIT Roorkee, European Synchrotron Radiation Facility (ESRF), and Elettra.
Sincrotrone Trieste. The data quality was evaluated using PHENIX.XTRIAGE (52). The structure of RO\textsubscript{CH\textsubscript{3}} was solved by a SAD phasing using Auto-Rickshaw (53) and was used as a search model for molecular replacement (MR) for PDOK\textsubscript{K1} data, which was performed with Phaser-MR (54) of CCP4i. Iterative rounds of model-building in COOT (55) and refinement of atomic coordinates and B-factors in refmac5 (56) allowed for the correct placement of sidechains and loops. The NCS restraints were used throughout the refinement of PDOK\textsubscript{K1} data. The data collection and refinement statistics are summarized in Table 2. The phthalate and terephthalate ligands could be modeled in only one chain of PDOK\textsubscript{K1} structure. All figures of protein and ligand structures were prepared using PyMol (57) and UCSF Chimera (58).

**Homology modeling and molecular docking**

Homology modeling was performed with Modeller 9.21 (59). The RO\textsubscript{CH\textsubscript{3}} structure served as a template for the homology modeling of iPDO\textsubscript{E6}, while the NDO\textsubscript{Rho} (PDB id: 2B1X) structure served for PDO\textsubscript{RHA1}. The alignment (Fig. S1) was produced with the STAMP algorithm (60) implemented in Multiseq (61). Two passes were performed with the following parameters: similarity set to 3, comparison residues set to 10, and slow scan option as performed in Capyk & Eltis, 2012 (62). The program ESPript (63) was used for the visualization of multiple sequence alignments. Molecular docking of isophthalate and phthalate was carried out by docking using AutoDock Vina (64). The ligand molecules of iPDO\textsubscript{E6} were prepared using ChemSketch (65). Autogrid4 was utilized for the grid selection, and the dimensions of the box were set to 6 Å × 7 Å × 7 Å for iPDO\textsubscript{E6} and 6 Å × 6 Å × 6 Å for PDO\textsubscript{RHA1}. The coordinates’ center points were at x = 23.91, y = 7.65, and z = 45.48 for iPDO\textsubscript{E6} and x = 25.12, y = 26.72, and z = 48 for PDO\textsubscript{RHA1}. A total of ten docking poses were evaluated based on the proper distances of the substrates to the receptors, and the best pose with the binding energy of −5.9 kcal mol\(^{-1}\) for iPDO\textsubscript{E6} and −7.2 kcal mol\(^{-1}\) for PDO\textsubscript{RHA1} was selected.

**Data availability**

The RO\textsubscript{CH\textsubscript{3}} and PDOK\textsubscript{K1} coordinates and structure factors have been deposited in the Protein Data Bank (http://www.rcsb.org/pdb) with accession numbers 7FHR, 7FJL, 7V25, 7V28. Other data are available from the corresponding author upon reasonable request.

**Supporting information**—This article contains supporting information.

**Acknowledgments**—The Authors thank MCU, IIT Roorkee; DBT, India; ESRF (Grenoble), France; and Elettra Synchrotron Trieste, Italy for facilitating with X-ray data collection. P. K. thank IIT Roorkee for providing funds for the anaerobic chamber through SMILE grant.

**Author contributions**—P. K. conceptualization; E. K., L. D. E., A. K. S., S. T., D. S., N. N., B. W., M. S., J. K. M., and P. K. investigation; P. K., L. D. E., J. K. M., and N. N. methodology; P. K. writing—original draft; J. K. M., L. D. E., N. N., E. K., and P. K. writing—review and editing.

**Funding and additional information**—This work was funded by National Bioscience Award by the Department of Biotechnology (DBT), India (Project No. BT/HRD/NBA/37/01/2015 (VIII)) and WTI 2019 grant by the Department of Science and Technology (DST), India (Project No. DST/TMD/EWO/WTI/2K19/EWFH/2019/8 (G)) to P. K. J. K. M. wants to thank DBT, Govt. of India, N. N. to CSIR, India and B. W and M. S. to MHRD, India for providing fellowship.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: BPDO, biphenyl dioxygenase; DHP, cis-4,5-dihydrodiol phthalate; HPLC, high-pressure liquid chromatography; NDO, naphthalene dioxygenase; PDO, phthalate dioxygenase; RO, Rieske oxidase; TCA, tricarboxylic acid.

**References**

1. Lin, H., Ge, R.-S., Chen, G.-R., Hu, G.-X., Dong, L., Lian, Q.-Q., Hardy, D. O., Sottas, C. M., Li, X.-K., and Hardy, M. P. (2008) Involvement of testicular growth factors in fetal Leydig cell aggregation after exposure to phthalate in utero. Proc. Natl. Acad. Sci. U. S. A. 105, 7218–7222
2. Xie, M., Wu, Y., Little, J. C., and Marr, L. C. (2016) Phthalates and alternative plasticizers and potential for contact exposure from children’s backpacks and toys. J. Expo. Sci. Environ. Epidemiol. 26, 119–124
3. Melnick, R. L. (2001) Is peroxisome proliferation an obligatory precursor step in the carcinogenicity of di (2-ethylhexyl) phthalate (DEHP)? Environ. Health Perspect. 109, 437–442
4. Kasai, D., Iwasaki, T., Nagai, K., Araki, N., Nishi, T., and Fukuda, M. (2019) 2, 3-Dihydroxybenzoate meta-cleavage pathway is involved in o-phthalate utilization in Pseudomonas sp. strain PTH10. J. Bacteriol. 191, 109–111
5. Keyser, P., Pujar, B. G., Eaton, R. W., and Ribbons, D. W. (1976) Biodegradation of the phthalates and their esters by bacteria. Environ. Health Perspect. 18, 159
6. Chang, H.-K., and Zylstra, G. J. (1998) Novel organization of the genes for phthalate degradation from Burkholderia cepacia DBO1. J. Bacteriol. 180, 6529–6537
7. Nakazawa, T., and Hayashi, E. (1977) Phthalate metabolism in Pseudomonas testosteroni: Accumulation of 4, 5-dihydroxyphtalate by a mutant strain. J. Bacteriol. 131, 42–48
8. Patrauchan, M. A., Florizone, C., Dosanjh, M., Mohn, W. W., Davies, J., and Eltis, L. D. (2005) Catabolism of benzoate and phthalate in Rhodococcus sp. strain RHA1: Redundancies and convergence. J. Bacteriol. 187, 4050–4063
9. Eaton, R. W. (2001) Plasmid-encoded phthalate catabolic pathway in Arthrobacter keyseri sp. strain RHA1: Redundancies and convergence. J. Bacteriol. 183, 3689–3703
10. Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2003) Crystal structure of naphthalene dioxygenase: Side-on binding of dioxygen to iron. Science 299, 1039–1042
11. Tarasev, M., and Ballou, D. P. (2005) Chemistry of the catalytic conversion of phthalate into its cis-dihydrodiol during the reaction of oxygen with the reduced form of phthalate dioxygenase. Biochemistry 44, 6197–6207
12. Lukowski, A. L., Liu, J., Bridwell-Rabb, J., and Narayan, A. R. H. (2020) Structural basis for divergent C–H hydroxylation selectivity in two Rieske oxygenases. Nat. Commun. 11, 2991
13. Kovaleva, E. G., and Lipscomb, J. D. (2008) Versatility of biological non-heme Fe (II) centers in oxygen activation reactions. Nat. Chem. Biol. 4, 186–193
Structure of phthalate dioxygenase

14. Rogers, M. S., and Lipscomb, J. D. (2019) Salicylate 5-hydroxylase: Intermediate in aromatic hydroxylation by a Rieske monoxygenase. Biochemistry 58, 5305–5319

15. Kauppi, B., Lee, K., Carredano, E., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (1998) Structure of an aromatic-ring-hydroxylating dioxygenase–naphthalene 1, 2-dioxygenase. Structure 6, 571–586

16. Carredano, E., Karlsson, A., Kauppi, B., Choudhury, D., Parales, R. E., Parales, J. V., Lee, K., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2000) Substrate binding site of naphthalene 1, 2-dioxygenase: Functional implications of indole binding. J. Mol. Biol. 296, 701–712

17. Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2005) NO binding to naphthalene dioxygenase. J. Biol. Inorg. Chem. 10, 483–489

18. Ferraro, D. J., Okerlund, A. L., Mowers, J. C., and Ramaswamy, S. (2006) Structural basis for regioselectivity and stereoselectivity of product formation by naphthalene 1,2-dioxygenase. J. Bacteriol. 188, 6986–6994

19. Galáhr, L., Malik, Z. A., Allen, C. C., Lipscomb, D. A., Larkin, M. J., and Ramaswamy, S. (2005) Structure and increased thermostability of Rhodococcus sp. naphthalene 1, 2-dioxygenase. J. Biol. Chem. 187, 7222–7231

20. Batie, C. J., LaHaié, E., and Ballou, D. P. (1987) Purification and characterization of phthalate oxygenase and phthalate oxygenase reductase from Pseudomonas cepacia. J. Biol. Chem. 262, 1510–1518

21. Correll, C. B., Batie, C. J., Ballou, D. P., and Ludwig, M. L. (1992) Phthalate dioxygenase reductase: A modular structure for electron transfer from pyridine nucleotides to [2Fe–2S]. Science 258, 1604–1614

22. Tarasev, M., Pullela, S., and Ballou, D. P. (2009) Distal end of 105–125 loop – a putative reductase binding domain of phthalate dioxygenase. Arch. Biochem. Biophys. 487, 10–18

23. Coulter, E. D., Moon, N., Batie, C. J., Dunham, W. R., and Ballou, D. P. (1999) Electron paramagnetic resonance measurements of the ferrous mononuclear site of phthalate dioxygenase substituted with alternate divalent metal ions: Direct evidence for ligation of two histidines in the copper(II)-reconstituted protein. Biochemistry 38, 11062–11072

24. Tarasev, M., Kaddis, C. S., Yin, S., Loo, J. A., Burgner, J., and Ballou, D. P. (2007) Similar enzymes, different structures: Phthalate dioxygenase is an a3a3 stacked hexamer, not an n383 trimer like “normal” Rieske oxygogenases. Arch. Biochem. Biophys. 466, 31–39

25. Gassner, G. T., and Ballou, D. P. (1995) Preparation and characterization of a truncated form of phthalate dioxygenase reductase that lacks an iron-sulfur domain. Biochemistry 34, 13460–13471

26. Gassner, G. T., Ballou, D. P., Landrum, G. A., and Whittaker, J. W. (1993) Magnetic circular dichroism studies on the mononuclear ferrous active site of phthalate dioxygenase from Pseudomonas cepacia show a change of ligation state on substrate binding. Biochemistry 32, 4820–4825

27. Gassner, G., Wang, L., Batie, C., and Ballou, D. P. (1994) Reaction of phthalate dioxygenase reductase with NADH and NAD: Kinetic and spectral characterization of intermediates. Biochemistry 33, 12184–12193

28. Pinto, A., Tarasev, M., and Ballou, D. P. (2006) Substitutions of the “bridging” aspartate 178 result in profound changes in the reactivity of the Rieske center of phthalate dioxygenase. Biochemistry 45, 9032–9041

29. Batie, C. J., and Ballou, D. P. (1990) Phthalate dioxygenase. Methods Enzymol. 188, 61

30. Tsang, H. T., Batie, C. J., Ballou, D. P., and Penner-Hahn, J. E. (1989) X-ray absorption spectroscopy of the [2-iron-2-sulfur] Rieske cluster in Pseudomonas cepacia phthalate dioxygenase. Determination of core dimensions and iron ligation. Biochemistry 28, 7233–7240

31. Gurbriel, R. I., Batie, C. J., Sivara, M., True, A. E., Fee, J. A., Hoffman, B. M., and Ballou, D. P. (1989) Electron-nuclear double resonance spectroscopy of nitrogen-15-enriched phthalate dioxygenase from Pseudomonas cepacia proves that two histidines are coordinated to the [2Fe–2S] Rieske-type clusters. Biochemistry 28, 4861–4871

32. Tierney, D. L., Gassner, G. T., Luchinat, C., Bertini, I., Ballou, D. P., and Penner-Hahn, J. E. (1999) NMR characterization of substrate binding in the phthalate dioxygenase system. Biochemistry 38, 11051–11061

33. Pavel, E. G., Martins, L. J., Ellis, W. R., Jr., and Solomon, E. I. (1994) Magnetic circular dichroism studies of exogenous ligand and substrate binding to the non-heme ferrous active site in phthalate dioxygenase. Chem. Biol. 1, 173–183
macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221
53. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. Acta Crystallogr. D Biol. Crystallogr. 65, 1089–1097
54. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674
55. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
56. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367
57. DeLano, W. L. (2002) Pymol: An open-source molecular graphics tool. CCP4 News. Protein Crystallogr. 40, 82–92
58. Potterson, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612
59. Webb, B., and Sali, A. (2016) Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinformatics 54, 5.6.1–5.6.37
60. Russell, R. B., and Barton, G. J. (1992) Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. Proteins 14, 309–323
61. Roberts, E., Eargle, J., Wright, D., and Luthey-Schulten, Z. (2006) MultiSeq: Unifying sequence and structure data for evolutionary analysis. BMC Bioinformatics 7, 1–11
62. Capyk, J. K., and Eltis, L. D. (2012) Phylogenetic analysis reveals the surprising diversity of an oxygenase class. J. Biol. Inorg. Chem. 17, 425–436
63. Gouet, P., Courcelle, E., Stuart, D., and Metoz, F. (1999) ESPript: Analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305–308
64. Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461
65. Hunter, A. D. (1997) ACD/ChemSketch 1.0 (Freeware); ACD/ChemSketch 2.0 and its Tautomers Dictionary, and 3D Plug-Ins; ACD/HNMR 2.0; ACD/CNMR 2.0, ACS Publications, Washington, DC