1   **Biochemical and structural characterization of an aromatic ring-hydroxylating dioxygenase for terephthalic acid catabolism**

2   William M. Kincannon,[1,†] Michael Zahn,[2,4,†] Rita Clare,[1,4] Ari Romberg,[1] James Larson,[1,] Jessica Lusty-Beech,[1,4] Brian Bothner,[1] Gregg T. Beckham,[3,4] John E.
3   McGeehan,[2,4,*] Jennifer L. DuBois,[1,4,*]

4   1. Department of Biochemistry, Montana State University, Bozeman MT, United States

5   2. Centre for Enzyme Innovation, University of Portsmouth, Portsmouth, United Kingdom

6   3. Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden CO United States

7   4. BOTTLE Consortium, Golden CO, United States

8   * Correspondence: john.mcgeehan@port.ac.uk; jennifer.dubois1@montana.edu

9   †co-first authors

10   **Classification**: Biological Sciences, Applied Biological Sciences

11   **Keywords**: plastics recycling, upcycling, Rieske oxygenase, non-heme iron, aromatic catabolism; *Comamonas* sp. E6

12   **Abstract**: Several bacteria possess components of catabolic pathways for the synthetic polyester poly(ethylene terephthalate) (PET). These proceed by hydrolyzing
13   the ester linkages of the polymer to its monomers, ethylene glycol and terephthalate (TPA), which are further converted into common metabolites. These pathways
14   are crucial for genetically engineering microbes for PET upcycling, prompting interest in their fundamental biochemical and structural elucidation. TPA dioxygenase
15   (TPADO) and its cognate reductase comprise a complex multi-metalloenzyme system that dihydroxylates TPA, activating it for enzymatic decarboxylation to yield
16   protocatechuic acid (PCA). Here, we report structural, biochemical, and bioinformatic analyses of TPADO. Together, these data illustrate the remarkable adaptation of
17   TPADO to the TPA dianion as its preferred substrate, with small, protonatable ring-2-carbon substituents being among the few permitted substrate modifications.
18   TPADO is a Rieske [2Fe2S] and mononuclear non-heme iron-dependent oxygenase (RO) that shares low sequence similarity with most structurally characterized
19   members of its family. Structural data show an α-helix associated histidine side chain that rotates into an Fe (II)-coordinating position following binding of the
20   substrate into an adjacent pocket. TPA interactions with side chains in this pocket were not conserved in homologs with different substrate preferences. The binding
21   mode of the less symmetric 2-hydroxy-TPA substrate, the observation that PCA is its oxygenation product, and the close relationship of the TPADO α subunit to that
22   of anthranilate dioxygenase allowed us to propose a structure-based model for product formation. Future efforts to identify, evolve, or engineer TPADO variants with
23   desirable properties will be enabled by the results described here.

24   **Significance Statement:** More than 400 million tons of plastic waste are produced each year, the overwhelming majority of which ends up in landfills. Bioconversion
25   strategies aimed at plastics have emerged as important components of enabling a circular economy for synthetic plastics, especially those that exhibit chemically
26   similar linkages to those found in nature, such as polyesters. The enzyme system described in this work is essential for mineralization of the xenobiotic components of
27   PET in the biosphere. Our description of its structure and substrate preferences lay the groundwork for *in vivo* or *ex vivo* engineering of this system for PET upcycling.

28   **Introduction**
29   The discovery of a bacterium that assimilates poly(ethylene terephthalate) (PET) (1-2), the synthetic polymer used for clothing, single-use plastic bottles, and carpets
30   has generated great excitement over the prospect of using biological catalysis for recycling this abundant waste product (3-7). The pathway for PET assimilation by
31   *Ideonella sakaiensis* begins with a pair of esterases that hydrolyze the polymer to its constituent monomers, ethylene glycol and terephthalate (TPA). Ethylene glycol
32   is a natural product that is metabolized by multiple bacteria (8-10). TPA resembles plant-derived aromatic compounds, but it is not widely known as a substrate for
33   bacterial growth. Because of its size and charge (-2 at pH 7), TPA must be actively transported into the cell where it is *cis*-dihydroxylated and dearomatized to yield
34   1,2-dihydroxy-3,5-cyclohexadiene-1,4-dicarboxylate (DCD) (11-17). The initial dihydroxylation is catalyzed by an O_2-dependent TPA dioxygenase (TPADO) working in
35   conjunction with an NAD(P)H, flavin, and iron-sulfur-dependent reductase. A zinc-dependent dehydrogenase finally reductively decarboxylates DCD to produce
36   protocatechuate (PCA) (**Fig. 1**) (1). Interest in using this pathway, either *in vitro* or engineered into microbes optimized for PET conversion, is motivated by the
37   societal, ecological, and economic benefits of plastics reclamation, recycling, and upcycling (7, 18, 19). Obtaining a structural and functional understanding of each of
38   the four enzymes is an essential step toward these future applications.

39   Native systems for TPA import and *cis*-dihydroxylation have thus far been identified in several bacteria, including *I. sakaiensis* and several strains that do not use PET
40   as a carbon source. The latter include *Comamonas* sp. strain E6 (12, 17), *C. testosteroni* T-2 (20), *C. testosteroni* YZW-D (13), *Delftia tsuruhatensis* T7 (19),
41   *Rhodococcus sp.* DK17 (11), *Rhodococcus jostii* RHA1(14), *Pseudomonas umsongensis* GO16 (8), and *Acinetobacter baylyi* ADP1 (TPA importer) (22). Several studies
42   have introduced TPA catabolism into microbes for metabolic engineering applications as well (23-27).

43   The TPADO enzyme is a member of a large family of Rieske oxygenases (ROs). Several members of this family permit bacteria to aerobically assimilate, and thereby
44   remediate, a wide range of environmental contaminants (28), such as naphthalene, pyrene, toluene, and chlorobenzoate. Several of these compounds resemble
45   natural metabolites, including benzoate, cinnamate, picolinate, and salicylate, all of which are also RO substrates (29). Naphthalene dioxygenase (NDO), a
46   paradigmatic RO (30-32), catalyzes stereo- and regio-selective dihydroxylations on a range of aromatic substrates. Other family members likewise catalyze an array of
47   mono- and dihydroxylations, *O*- and *N*- dealkylations (29), desaturations, and sulfoxidations, where several prior studies have suggested a structural basis for reaction
48   type or substrate preference (28, 33). ROs possess a family-defining [2Fe-2S] Rieske cluster that delivers electrons sequentially to an adjacent mononuclear non-heme
49   iron center, where O_2 is reductively activated. The ultimate electron source is NAD(P)H, which transfers a hydride to a flavin adenine dinucleotide (FAD) cofactor in a
50   separate reductase enzyme. One to two additional [Fe-S] clusters serve as 1e- shuttles to the active site (*SI Appendix*, **Fig. S1**). These clusters are found in diverse
51   protein domain architectures that have traditionally been used for subtyping ROs in types I through V (34). Among the type II family members, which include TPADO,
52   sequence conservation can be surprisingly low (35). Additionally, until recently, there were no structurally characterized examples of type II enzymes. Predicting TPA-
53   directed activity among type II ROs based solely on the primary sequence of the catalytic (α) subunit is consequently challenging, limiting efforts to prospect for
54   TPADO homologs in sequence databases.

55   Insert Figure 1 near here

56   In this work, we describe multiple crystal structures of the TPADO from *Comamonas* sp. strain E6 (17) in complex with both TPA and an *ortho*-substituted analog that
57   links TPADO to salicylate 1-hydroxylase (36) and anthranilate 1,2-dioxygenase (37) in mechanism. Further, we define catalytic parameters and describe the specificity
58   of TPADO for *para*-dicarboxylate anions that is well-explained by the binding mode observed in the structures. Finally, we show that TPADO has functionally
59   significant sequence relationships with aryl-carboxylate directed ROs which, in conjunction with the structural data, can be used to propose a TPA-recognition
60   sequence motif. The results establish this TPADO and its cognate reductase as a foundation for further protein and strain engineering efforts to achieve biological
61   upcycling of PET plastic.

62

Results

A two-enzyme system dioxygenates TPA *in vitro.* TPADO, a 196 kDa complex of TphA2 and TphA3 subunits, and its cognate reductase (TphA1, 36.4 kDa) from *Comamonas sp.* strain E6 were heterologously expressed in *Escherichia coli* (*SI Appendix*, **Fig. S2-S3**; *Supplementary Methods*), with conditions then optimized to maximize both protein and cofactor yields as guided by studies from Ballou *et al*. (38). The α and β subunits of TPADO were co-expressed from a single IPTG-inducible pET-DUET vector, where the α subunit contained a C-terminal His_6 tag that was used for affinity purification at yields >20 mg/L culture. Native mass spectrometry identified the $\alpha_3\beta_3$ oligomer as the major component with the individual monomers as minor constituents (*SI Appendix*, **Fig. S4**). The monomeric reductase was purified via a C-terminal His_6 tag, with yields >10 mg/L culture.

UV/visible, atomic absorption, and electron paramagnetic spectroscopies confirmed the presence and type of [2Fe-2S] cluster in each protein (Rieske and plant types in TPADO and the reductase, respectively) (*SI Appendix*, **Fig. S5-S7**), based on their characteristic g-values and absorbance maxima. Combined metal and protein analyses predicted approximately stoichiometric occupancy of the expected cofactors in TPADO; however, this assumes that all of the protein (by mass) was present as intact $\alpha_3\beta_3$ oligomer, and that all iron was cofactor-associated. Consequently, the concentrations of TPADO used throughout this study, reported in terms of TPADO active sites (three per $\alpha_3\beta_3$ unit), are undoubtedly overestimated since at least some catalytically inactive α and β monomers and unpopulated cofactor binding sites were likely present.

TPA was converted by the TPADO/reductase system to an oxidized product specifically in the presence of NADH rather than NADPH. The oxygenation product could be resolved from the NAD^+ co-product by reverse phase high-performance liquid chromatography (HPLC) when subjected to a complex gradient (*SI Appendix, Supplementary Methods*). The distinctive UV/vis absorbance spectrum for the product (**Fig. 2A**) was consistent with absorbance maxima reported for DCD (15). However, this molecule is not commercially available, nor does extensive characterization exist in published literature. Here, we to chromatographically resolved the product from NAD^+ by HPLC, and obtained its UV/vis absorbance spectrum and mass spectral (MS) analyses. Under the ionization conditions used, DCD exhibited fragmentation consistent with the loss of CO_2 and/or H_2O (**Fig. 2B**). The expected exact mass of intact DCD is m/z = 199, while the predominant ion observed at the retention time of the product had an observed m/z = 137. TPA also displayed similar behavior, losing CO_2 during source ionization (*SI Appendix*, **Fig. S8**). Targeted fragmentation of the weak intensity, intact DCD and TPA peaks by LC-MS-MS yielded fragmentation profiles that were consistent with the observed source ionization fragmentation, supporting the definitive identification of the latter compound.

Steady-state kinetic parameters were measured as a function of variable [TPA] in ambient air (*SI Appendix*, **Fig. S9**). The data readily fit the Michaelis-Menten model, yielding k_{cat}(apparent) = 12 ± 0.3 min^{-1} and K_M[TPA](apparent) = 9.6 ± 1 μM. The parameters are *apparent* as saturating concentrations for NADH and O_2 have not been determined, and k_{cat} (maximal velocity x $[enzyme]^{-1}$) is likely underestimated. The k_{cat} and low-micromolar K_M are nonetheless similar to values measured for the TPADO homolog salicylate 5-monooxygenase NagGH from *Ralstonia* sp. strain U2 (k_{cat} = 7.81 min^{-1}, K_M = 22.4 μM) (39), suggesting that TPADO is well adapted to its substrate though slow in absolute terms. A total turnover number (TTN) of 704 (per apparent TPADO active site) was measured for the system via determination of unreacted TPA by HPLC when all other reactants were in excess. This TTN is low compared with several enzymes used in applied work, potentially due to over-counting of intact metallo-active sites here or to intrinsic instability of the enzymes. Either explanation suggests the need for engineered improvements to enzyme stability for future applications (40).

Insert Figure 2 near here

TPADO exhibits remarkable specificity for *para*-dicarboxylates. Many ROs are known either to accept multiple substrates (33) or to uncouple reductive O_2 activation from substrate oxygenation, expending NAD(P)H without oxygenating the organic substrate and releasing either H_2O_2 or water. The substrate or analog binds near to the mononuclear Fe(II), displacing water and opening a coordination position where O_2 can be reductively activated. The Fe/O_2 species can either productively oxygenate the substrate or, when an uncoupler is bound, break down to yield water or H_2O_2 (30, 41, 42). (See Discussion, Scheme 1.)

The specificity of TPADO was assessed by analyzing substrate consumption at a fixed time point (90 min) after incubating TPADO/reductase with stoichiometrically limiting NADH and either TPA (**Fig. 2A**) or a structural analog (**Fig. 3;** *SI Appendix*, **Fig. S12-S14**). The stoichiometry of aromatic substrate as a function of NADH usage was measured by integrating their respective HPLC peaks, using a no-substrate/no-analog control sample as a baseline. Of a series of compounds screened, only TPA and closely related derivatives with hydroxyl- and amino- substituents at the ring C2 position were oxidized (**Fig. 3**). Percent coupling was calculated as: (moles substrate consumed) (moles NADH consumed above baseline)$^{-1}$ x 100%. Both TPA and 2-hydroxy TPA (2-OH-TPA) exhibited 100% coupling, while 2-amino TPA (2-NH_2-TPA) showed 22% coupling (**Fig. 3**).

Analogs with conservative substitutions to one of the carboxylates (4-nitrobenzoic acid (4-NBA), 4-carbamoylbenzoic acid (4-CBA), and 4-formylbenzoic acid (4-FBA)) yielded no detectable oxygenated products (**Fig. 3C**). 4-FBA and 4-CBA did not stimulate NADH oxidation above baseline (*SI Appendix*, **Fig. S12-S13**), while 4-NBA acted as an efficient uncoupler, promoting full consumption of NADH without reduction in the substrate analog signal (*SI Appendix*, **Fig. S14**). Taken together, these results demonstrate an extraordinarily high level of fidelity for TPA and very closely related diacids as principal substrates of the enzyme.

Insert Figure 3 near here

TPADO converts 2-OH-TPA and 2-NH2-TPA to PCA. TPA analogs with -NH_2 or –OH substituents at the ring-2-position were converted to a product having a retention time and UV/vis absorbance spectrum identical to those obtained for a PCA analytical standard (CAS #99-50-3, **Fig. 3A-3B**, *SI Appendix* **Fig. S10B, Table S1**). 1,2-*Bis*-hydroxylation uniquely leads to an intermediate where decarboxylation and ring re-aromatization can proceed in conjunction with protonation and spontaneous loss of the ring-2 substituent (see Discussion). The observation of PCA as the product in both cases therefore indicates that hydroxylation must take place at the 1,2- (and not, for example, the 1,6-) ring carbons.

TPA dioxygenase has a canonical α3β3 structure. To understand the structural arbiters of substrate recognition, we determined the crystal structures of TPADO in the ligand-free state to 2.28 Å resolution and ligand bound structures with TPA to 2.08 Å and 2-OH-TPA to 1.95 Å resolution. During model building and refinement, clear continuous protein density was observed outside the α3β3 domains and identified as a single lysozyme molecule (*SI Appendix*, **Fig. S15**). In a rather fortuitous crystal packing, the lysozyme protein effectively acts as a crystallization chaperone that "glues" the TPADO molecules together (*SI Appendix*, **Fig. S16**), a function that, to our knowledge, has not been previously reported for lysozyme. More specifically, one lysozyme molecule forms interactions to five TPADO molecules and binds to both α- and β- subunits. TPADO forms an α3β3-heterohexamer in which three catalytic α subunits form a trimeric head-to-tail assembly atop a triad of non-catalytic β subunits (**Fig. 4A**). Like other ROs, a Rieske domain-containing a [2Fe-2S] cluster and a catalytic domain which comprises the TPA substrate and ferrous ion binding site (**Fig. 4B-C**) were located in each α-subunit (26). Residues H210, H215, and D356 coordinate the ferrous ion in the active site, which is 12.2 Å away from the [2Fe-2S] cluster of a neighboring α-subunit (**Fig. 4A**) and connected via the amino acid side chains H210, D207, and H105 (**Fig. 4B-D**). The [2Fe-2S] cluster within the same α-subunit is, by contrast, 42 Å away where it contributes to the neighboring reaction site.

The mononuclear-Fe-coordinating residues H210 and H215 are located on an α-helix spanning residues 208 to 220. In the substrate-free structure, H215 does not coordinate the ferrous ion, and in all three α-subunits, the helix conformation is the same. From residue H215 onwards, this helix is disordered in two out of the three α-subunits, while the third is stabilized by crystal contacts with a symmetry related molecule (*SI Appendix*, **Fig. S17A-B**). After crystal soaking with substrates TPA or 2-OH-TPA, the formerly disordered helix residues are ordered but partially unfolded, thereby allowing H215 to coordinate the ferrous ion (**Fig. 4E,** *SI Appendix*, **Fig.**

128   **S17C-D**). This suggests altered dynamics of the system upon substrate binding. The helix stabilized by crystal contacts in the apo structure remained in its orientation
129   upon substrate binding, and did not unfold to coordinate the iron.

130   Insert Figure 4 near here

131   TPA binding is supported by multiple ionic, hydrogen-bonding, and hydrophobic interactions. The substrate was located next to the mononuclear iron in an
132   orientation confirmed by clear unbiased F_o-F_c difference electron density (*SI Appendix*, **Fig. S18**). One of the carboxylate groups from the bound TPA forms a salt
133   bridge with R309, positioning the adjacent ring carbons 3.9 – 4.1 Å away from the reactive iron in an orientation that could promote reaction with an activated Fe/O$_2$
134   species. The side chain of I290 forms a hydrophobic-π-interaction with the aromatic ring of TPA, while the second carboxylate group forms a hydrogen bond with
135   S243. Additionally, a potential salt bridge between R390 and this carboxylate is observed, but the distance varies among the α subunits between 2.9-4.6 Å. This
136   apparent flexibility in the R390 side chain is reflected in its high B-factors and could have functional significance. For example, residue R390, as well as the α-helix
137   mentioned above and subsequent residues, form the opening of a potential site for TPA entry or product release that may have a gate function (*SI Appendix*, **Fig.**
138   **S19**). The TPA binding pocket is more open within the substrate-free structure because the α-helical residues from H215 onwards are disordered and not part of the
139   model; in total, residues 215 to 227 are missing.

140   **TPA exhibits a unique binding mode.** A search in the current structural databases, utilizing the DALI protein comparison server (43), highlighted salicylate 5-
141   monooxygenase (NagGH) (37, 42) as the closest structural homolog to TPADO with RMSD values of 1.5 Å and 1.3 Å, and sequence identities of 42% and 26%, for the α
142   and β subunits, respectively. Furthermore, the α subunits of related toluene 2,3-dioxygenases, biphenyl dioxygenases, and NDOs are significantly more divergent,
143   with RMSD values ranging between 2.4–2.8 Å. A superposition of TPADO with NagGH illustrates their similar quaternary structures (*SI Appendix*, **Fig. S20**). While no
144   substrate is present in the NagGH structure for comparison, it appears unlikely that salicylate 5-monooxygenase could bind TPA because the side chain of M257,
145   which replaces the S243-TPA hydrogen-bonding interaction observed in TPADO, would interfere with the carboxylate group of TPA (**Fig. 5A**).

146   The salt bridge involving one of the TPA carboxylate groups is conserved in both TPADO (R309) and the salicylate 5-monooxygenase (R323) from *Ralstonia* sp.(**Fig. 5A**)
147   (39), although an equivalent H-bond to R390 on the opposite end of the substrate is absent. Several additional Rieske dioxygenases possess a positively charged
148   residue (R or K) at the same position as R309 in a pocket that is otherwise largely hydrophobic, but in those cases the homologous side chain points away from the
149   substrate, as observed in the NDO structure (**Fig. 5B**) (45). Moreover, the residues involved in electron transfer from the Rieske cluster to the mononuclear iron
150   superimpose well between NagGH and NDO.

151   The structure of TPADO bound to 2-OH-TPA is close to identical to the complex with TPA but surprisingly, the electron density for this ligand strongly suggests that the
152   2-hydroxy group is oriented towards the hydrophobic part of the active site and not towards the polar part where hydrogen bonding interactions with the carbonyl
153   oxygen of V205 would appear possible (*SI Appendix*, **Fig. S21**). It is possible that this binding orientation of 2OH-TPA promotes catalysis by substrate destabilization,
154   that it favors product release, or that it represents a non-productive binding mode. To investigate the latter possibility, we aligned the 2-OH-TPA bound structure with
155   a product-bound NDO structure (PDB ID 1O7P) (31). From this alignment (**Fig. 5B-C**), the *cis*-dioxygenated carbons of the (*1R,2S*)-*cis*-1,2-dihydroxy-1,2-
156   dihydronaphthalene product appear to align with the 1- and 2-carbons of 2-OH-TPA. These carbons are consequently implicated as the sites of hydroxylation based
157   on the observed formation of PCA as the ultimate product, suggesting that 2-OH-TPA occupies a productive binding mode. This also suggests, by analogy, that the
158   stereochemistry of product of the TPA reaction is predicted to be (*1S,2R*)-*cis*-1,2-dihydroxy-3,5-cyclohexadiene-1,4-dicarboxylate (DCD).

159   Insert Figure 5 near here

160   **Sequence similarity network (SSN) analyses relate the TPADO α-subunit to a subgroup of ROs that hydroxylate hydrophilic monoaryls.** Classic RO subtyping by
161   Kweon *et al*. focused on the different domain organizations that mediate electron flow between the reductase and RO active site (34). A subsequent structure-based
162   sequence alignment of 121 then-available catalytic α subunits was carried out by Eltis *et al*. (35). This analysis resulted in a first-ever phylogenomic map and
163   suggested how gene fusion events gave rise to the Rieske/mononuclear Fe α subunit.

164   Just over 45,000 α-subunit homologs are currently known, permitting subtyping of a much broader scope. Their sequences, members of Pfam PF00848, were
165   submitted to an "all-by-all" comparison via the Enzyme Function Initiative's (EFI) network analysis algorithm (**Fig. 6**, *SI Appendix*, **Fig. S22**) (46). Consistent with earlier
166   findings from Eltis *et al*., we observed that a single subgroup of the massive and diverse family has been experimentally oversampled in prior work (cluster #2, *SI*
167   *Appendix*, **Fig. S22**). This node contains several aromatic/polyaromatic hydrocarbon hydroxylases of interest for bioremediation purposes, including NDO and
168   biphenyl dioxygenase (BPO) (47).

169   Insert Figure 6 near hereBeyond this large subgroup, ROs with similar substrate and/or reaction types, where known, were grouped together in this analysis,
170   suggestive of sequence-based family subdivision along functional lines (*SI Appendix,* **Fig. S22**). The cluster of sequences to which the α-subunit of TPADO (TphA2)
171   belongs (**Fig. 6**) also contains NagG, the α-subunit of salicylate 5-hydroxylase, consistent with identification of the latter via the DALI search for TPADO structural
172   relatives. Additionally, other functionally annotated members of the TPADO subfamily were all associated with aryl carboxylate substrates (anthranilate (37),
173   picolinate (48), and salicylate) like TPA, although TPA is the only dicarboxylate of the group.

174   A sequence alignment of the α-subunits of these annotated members of the TphA2 sequence cluster showed that, of all the residues making direct contact with the
175   substrate in the TPADO structure, only R309 is conserved. Key portions of this alignment are shown in **Fig. 6B** and *SI Appendix*, **Fig. S23** for TPA consuming organisms
176   including *I. sakaiensis*, where the degree of conservation is especially high. As highlighted in **Fig. 5A**, the analogous arginine in NagG (R323) is proposed to form a salt
177   bridge with the lone carboxylate of the substrate, positioning it for monooxygenation at the ring-5-carbon (yielding gentisate (49)). This active site arginine is also
178   conserved in the α-subunits of hydroxypicolinate 3-monooxygenase (Pic3B) (48), anthranilate 1,2-dioxygenase (AndAc) (37), and salicylate 1-hydroxylase (AhdA1c;
179   catechol product) (50, 51). A similar role in positioning the substrate via the carboxylate could be proposed for the conserved arginine in the other enzymes, though
180   the sites of ring mono- or di-oxygenation vary (**Fig. 6A**). In TPADO, N224, S243, and R390 interact with the carboxylate at the non-reactive end of TPA but are not
181   conserved across the subfamily. This observation suggests that they may be important for productively positioning dicarboxylate substrates.

182   **Discussion**
183   Plastic bioconversion using microbial enzymes depends on understanding and ultimately improving the properties of the responsible enzymes. Hydrolyzable ester
184   linkages are ubiquitous in biology, and aromatic compounds bearing polar substituents are abundant in the metabolic pathways of diverse organisms. The presence
185   of both ester linkages and an aromatic building block with a polar substituent in PET is consistent with this plastic constituting a readily available carbon source for
186   bacterial consumption in the biosphere.

187   TPADO is an α$_3$β$_3$ non-heme Rieske oxygenase which catalyzes the NADH-dependent dioxygenation of TPA, the aromatic subunit of PET. Like the better-studied
188   cytochrome P450s, ROs catalyze a wide range of oxidations and oxygenations of diverse substrates, increasing their water solubility and activating them for further
189   metabolism (52). However, unlike cytochromes P450, certain members of the RO family are capable, perhaps uniquely, of catalyzing the *cis*-dihydroxylation of an aryl
190   ring. The reaction depends on the proper positioning of the substrate relative to an activated O$_2$ species that forms at a mononuclear iron center, where both O-
191   atoms are poised to react on the same side of the plane defined by the substrate's aromatic ring (33). Dihydroxylation results in both the loss of aromaticity and the
192   formation of two new chiral centers in the product, DCD. This product, in turn, is well-situated for further catabolism via an exergonic step in which CO$_2$ is produced,
193   NADH regenerated, and the ring rearomatized to yield a valuable and versatile metabolite, PCA (53). This elegant metabolic arrangement, in which reductant is

194 intrinsically recycled within the pathway, suggests that, in spite of the complexity of the TPADO/reductase system and ROs in general, this
195 system offers a compelling starting point for engineering an efficient route for TPA bioconversion to PCA.

196 PET is an abundant, xenobiotic source of environmental TPA, suggesting that TPA might not be the primary but perhaps a secondary substrate of TPADO. We noted
197 that the bacterial TPADO studied here was nonetheless highly efficient and extraordinarily specific for TPA (K_M = 9.6 µM, k_{cat}/K_M(TPA) = 2.1 M^{-1}s^{-1}), coupling NADH
198 oxidation to TPA dihydroxylation with 100% fidelity and excluding a variety of structurally related compounds as potential substrates (**Fig. 3**). This observation is
199 consistent with an early report on the *Comamonas* E6 TPADO indicating that neither the *iso-* nor the *ortho-*benzene dicarboxylate regioisomers of TPA (*para*-benzene
200 dicarboxylate) were substrates for the enzyme (17). Two major exceptions identified here are 2-OH-TPA and, to a lesser extent, 2-NH$_2$-TPA, where TPADO catalyzed
201 the conversion of each to PCA.

202 Each of the benzene dicarboxylate regioisomers is used in the production of plastics, whether as a repeating subunit in PET, as a synthetic precursor for monomers
203 used in plastics, or as a noncovalently bound plasticizer (52). The *iso-* and *ortho-* benzene dicarboxylates (phthalates) serve one or both of the latter functions in
204 several plastics. *Ortho*-phthalate has been of special concern as an endocrine disruptor with the potential to leach out of consumer products and into water (54).
205 Bacteria that can degrade each of these compounds have been identified (55), indicating that enzymatic adaptation to these plastic-relevant compounds is not
206 unique to one or a few strains. Recent work even suggests possible biogenic sources for and derivatives of phthalates (56), which could have helped drive RO
207 diversification.

208 The *Comamonas* E6 strain has served as a paradigmatic plastic monomer degrader, capable of using each of the benzene dicarboxylate regioisomers as a sole source
209 of carbon and energy (55), via separate operons encoding three distinct RO enzymes. While the TPADO has an α3β3 subunit structure, the *ortho-* phthalate
210 dioxygenase from *Comamonas E6* displays α3α3 hexameric structure. It is possible that *iso*-PDO and *ortho*-PDOs from related strains have similar architecture. The
211 strain E6 *iso-/ortho-* phthalate dioxygenase α subunits share 32% identity with one another, forming a separate lineage from the α subunits of TPADO and other α3β3
212 ROs with which they share ≤19% identity (57). Accordingly, the α subunits of TPADO and a recently reported *ortho*-phthalate dioxygenase structure superimpose
213 poorly (*Comamonas testosteroni* KF1 phthalate dioxygenase (PDO), *SI Appendix*, **Fig. S24 (55)**). The active site of this PDO exhibited ionic/hydrogen-bonding
214 interactions between the bound *ortho*-phthalate carboxylates and an Arg-Arg-Ser triad that is unconserved in the TPADO sequence or tertiary structure. Hydrophobic
215 interactions were observed surrounding the benzylic ring. TPA was able to bind with some observed hydroxylation in the same PDO pocket, though with 80%
216 uncoupling and a much lower k_{cat}/K_M than for *ortho*-phthalate. These results suggest that TPADO and PDO are each well-adapted to their preferred benzene
217 dicarboxylate regioisomer.

218 A large-scale network analysis of >45,000 sequences of RO catalytic α-subunits was generated, aimed at understanding the origins of TPA bioconversion. Consistent
219 with the unique homohexameric structures identified by the Eltis group, PDO α−subunits from *Comamonas* are sufficiently sequence-distant that they are not
220 grouped with the same PFAM family (PF00848) and are not found in this SSN. Analysis of the network identified a subcluster containing the closest sequence relatives
221 to TPADO (**Fig. 6, SI Appendix, S24**). While only a few members of the subcluster are characterized in the literature, they are all known to catalyze reactions with aryl-
222 carboxylic acids which structurally resemble TPA. Many of the nodes in the cluster shown in **Fig. 6A** are annotated as similar to AhdA1c, the catalytic subunit of
223 salicylate-1-monooxyenase (IPR043264). A multiple sequence alignment revealed a cluster-wide conserved arginine which forms a salt bridge to one of the
224 carboxylate groups of TPA (R309, **Fig. 4**), and which is proposed to engage in a similar interaction with salicylate based on structural characterization of salicylate 5-
225 monooxygenase (NagGH) (39). An additional set of residues (N224, S243, and R390) have side chains within hydrogen bonding distance of the second carboxylate in
226 the TPA and 2-OH-TPA co-structures with TPADO (**Fig. 4-5**) described here. These are not conserved in the other functionally annotated sequences highlighted in **Fig.**
227 **6A-6B**. This conserved motif may therefore offer a means of identifying diverse TPADOs from sequence databases. Additional interesting findings resulting from the
228 sequence similarity network (SSN) analysis cannot be discussed here at length. We are supplying the network file for readers to explore the vast sequence space.
229 Many of the clusters in the SSN are comprised entirely of sequences with no known function, suggesting a plethora of catalytic diversity remains to be discovered.

230 A structure-based mechanism can be proposed to explain the observed products of the TPADO-catalyzed reactions with TPA, 2-OH-TPA, and 2-NH$_2$-TPA considering
231 the data presented here and two additional observations. First, the network analysis revealed a close sequence relationship between the α subunits of TPADO and
232 anthranilate dioxygenase from *Burkholderia cepacia* DBO1 (37). This organism can grow with anthranilic acid as a sole carbon source, where it is proposed that its RO
233 functions to dihydroxylate anthranilate, which then spontaneously deaminates and decarboxylates to yield PCA (37). Second, extensive prior work has shown that
234 NDO, perhaps the best-studied RO, can catalyze a variety of oxidations depending on the position of the substrate in the WT and variant enzymes relative to the site
235 of O$_2$ activation. Aryl carbons that are closest to the mononuclear iron (31, 58) are generally prioritized for hydroxylation.

236 A mechanism (**Scheme 1A**), taking these observations and prior work with ROs into account, would proceed as follows. Substrate binding near to the mononuclear Fe
237 in fully-reduced TPADO is expected to displace an iron-bound water molecule from TPADO (30, 41, 42). Here, we observed an unexpectedly large change in the
238 conformation of a helix containing two Fe-ligating histidine residues in response to substrate binding (**Fig. 4C**), consistent with the role of the substrate in permitting
239 binding and reductive activation of O$_2$. The initially formed ferric-η^1-superoxy intermediate (49) or its ferryl product could in principle be the oxygenating species. This
240 species would attack the nearest available substrate carbon (Fe-ring-C2 = 3.6 Å). Addition of a second active site electron from the Rieske cluster yields the ring-1,2-
241 epoxide, adjacent to the mononuclear ferric-hydroxyl. Nucleophilic attack of the hydroxyl at ring-C1 forms the hydrogenated, dearomatized, *cis*-diol product.

242 Dioxygenation of the ring-1,2 carbons of 2-OH-TPA by this route would yield an unstable gem-diol at the ring-C2 position. This is expected to readily dehydrate and
243 decarboxylate to yield PCA and CO$_2$ in the presence of an aqueous proton source to react with the hydroxyl leaving group. An analogous mechanism has been
244 postulated for the close sequence relative of TPADO, anthranilate dioxygenase, in which the substrate, a 2-amino benzoate, is initially dihydroxylated and
245 hydrogenated at the ring-1,2 carbons. Protonation of the C2-NH$_2$ group would catalyze the breakdown of the product to yield catechol, CO$_2$, and ammonia, which
246 would rapidly acquire a proton to form ammonium cation under neutral, aqueous conditions. The dicarboxylate analog of anthranilate (2-NH$_2$-TPA) serves as a
247 substrate of TPADO, with the corresponding deaminated product PCA, though the poorer amino leaving group leads to a lower level of productive turnover compared
248 with 2-OH-TPA. The close relationship between TPADO and anthranilate dioxygenase, and the proximity of the reactive Fe(II) to the 1,2 carbons of TPA or 2-OH-TPA in
249 the TPADO structures presented here (**Fig. 5C**) suggests an analogous route to DCD or PCA production in TPADO (**Scheme 1B**).

250 Together, these observations connect the experimentally determined binding interactions between TPADO and its substrates to a RO sequence subtype and a
251 potential sequence motif specific for *para*-aryl-dicarboxylate substrates like TPA. These, in turn provide strong support for a proposed pathway for the TPADO
252 catalyzed reaction that can now be optimized for future applied work.

253 **Materials and Methods**
254 TPADO and the reductase were heterologously expressed in *E. coli* and isolated in high yields via nickel affinity chromatography. Enzyme activity was monitored
255 continuously via UV/vis monitoring of NADH disappearance and discontinuously via separation of reaction components via HPLC. Identities of starting material and
256 products were detected with both mass-spectrometric and diode array UV-visible detectors. Quantification of reaction materials was carried out by integrating peak
257 areas from chromatograms recorded at distinct absorption maxima. The crystal structures were solved by molecular replacement using structural homologs for the α-
258 and β-subunits. TPADO was crystallized in the substrate-free state and soaked with substrates TPA and 2-OH-TPA overnight. Coordinates for the resulting apo and
259 substrate-bound structures have been deposited in the Protein Data Bank (PDB IDs: 7Q04, 7Q05, 7Q06). Structural superpositions of TPADO with NagG (7C8Z) and
260 NDO (1O7P) were generated and visualized in PyMol. An SSN of the family of proteins that the catalytic domain of TPADO belongs to (PF00848) was generated with
261 EFI-EST web tools and visualized in Cytoscape. Detailed methods are provided in the *SI Appendix*.

References

1.   S. Yoshida, *et al.*, A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351**, 1196-1199 (2016).
2.   U. T. Bornscheuer, Feeding on plastic. *Science* **351**, 1154-1155 (2016).
3.   V. Tournier et al., An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* **580**, 216-219 (2020).
4.   R. Wei, W. Zimmermann, Biocatalysis as a green route for recycling the recalcitrant plastic polyethylene terephthalate. *Microb. Biotech.* **10**, 1302-1307 (2017).5.   L. D. Ellis, *et al.*, Chemical and biological catalysis for plastics recycling and upcycling. *Nature Cat.* **4**, 539-556 (2021).
6.   A. J. Martin, C. Mondelli, S. D. Jaydev, J. Perez-Ramirez, Catalytic processing of plastic waste on the rise. *Chem* **7**, 1487-1533 (2021).
7.   S. R. Nicholson, N. A. Rorrer, A. C. Carpenter, G. T. Beckham, Manufacturing energy and greenhouse gas emissions associated with plastics consumption. *Joule* **5**, 673-686 (2021).
8.   T. Narancic, *et al.*, Genome analysis of the metabolically versatile *Pseudomonas umsongensis* GO16: The genetic basis for pet monomer upcycling into polyhydroxyalkanoates. *Microb. Biotech.* **14**, 2463-2480 (2021).
9.   W. J. Li, *et al.*, Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440. *Environ. Microbiol.* **21**, 3669-3682 (2019).
10.   M. A. Franden, *et al.*, Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. *Metab. Eng.* **48**, 197-207 (2018).
11.   K. Y. Choi, *et al.*, Molecular and biochemical analysis of phthalate and terephthalate degradation by *Rhodococcus* sp strain DK17. *FEMS Microbiol. Lett.* **252**, 207-213 (2005).
12.   M. Sasoh, *et al.*, Characterization of the terephthalate degradation genes of *Comamonas* sp, strain E6. *Appl. Environ. Microbiol.* **72**, 1825-1832 (2006).
13.   Y. Z. Wang, Y. M. Zhou, G. J. Zystra, Molecular analysis of isophthalate and terephthalate degradation by *Comamonas testosteroni* YZW-D. *Environ. Health Persp.* **103**, 9-12 (1995).
14.   H. Hara, L. D. Eltis, J. E. Davies, W. W. Mohn, Transcriptomic analysis reveals a bifurcated terephthalate degradation pathway in *Rhodococcus* sp. strain RHA1. *Journal of Bacteriology* **189**, 1641-1647 (2007).
15.   H. R. Schläfli, M. A. Weiss, T. Leisinger, A. M. Cook, Terephthalate 1,2-dioxygenase system from *Comamonas testosteroni* T-2: Purification and some properties of the oxygenase component. *J. Bacteriol.* **176**, 6644-6652 (1994).
16.   D. Kasai, M. Kitajima, M. Fukuda, E. Masai, Transcriptional regulation of the terephthalate catabolism operon in *Comamonas* sp. strain E6. *Appl. Environ. Microbiol.* **76**, 6047-6055 (2010).
17.   Y. Fukuhara, D. Kasai, Y. Katayama, M. Fukuda, E. Masai, Enzymatic properties of terephthalate 1,2-dioxygenase of *Comamonas* sp. strain E6. *Biosci., Biotech., Biochem.* **72**, 2335-2341 (2008).
18.   J. Zheng, S. Suh, Strategies to reduce the global carbon footprint of plastics. *Nature Clim. Ch.* **9**, 374-378 (2019).
19.   A. Singh, *et al.*, Techno-economic, life-cycle, and socioeconomic impact analysis of enzymatic recycling of poly(ethylene terephthalate). *Joule* **5**, 2479-2503 (2021).
20.   H. R. Schlafli, M. A. Weiss, T. Leisinger, A. M. Cook, Terephthalate 1,2-dioxygenase system from *Comamonas-testosteroni*-T-2 - purification and some properties of the oxygenase component. *J. Bact.* **176**, 6644-6652 (1994).
21.   T. Shigematsu, K. Yumihara, Y. Ueda, S. Morimura, K. Kida, Purification and gene cloning of the oxygenase component of the terephthalate 1,2-dioxygenase system from *Delftia tsuruhatensis* strain T7. *FEMS Microbiol. Lett.* **220**, 255-260 (2003).
22.   I. Pardo, *et al.*, Gene amplification, laboratory evolution, and biosensor screening reveal MucK as a terephthalic acid transporter in *Acinetobacter baylyi* ADP1. *Metab. Eng.* **62**, 260-274 (2020).
23.   T. Tiso, *et al.*, Towards bio-upcycling of polyethylene terephthalate. *Met. Eng.* **66**, 167-178 (2021).
24.   H. T. Kim, *et al.*, Biological valorization of poly(ethylene terephthalate) monomers for upcycling waste PET. *ACS Sustainable Chem. & Eng.* **7**, 19396-19406 (2019).
25.   H. T. Kim, *et al.*, Chemo-biological upcycling of poly(ethylene terephthalate) to multifunctional coating materials. *ChemSusChem* **14**, 4251-4259 (2021).
26.   D. H. Kim, *et al.*, One-pot chemo-bioprocess of PET depolymerization and recycling enabled by a biocompatible catalyst, betaine. *ACS Catal* **11**, 3996-4008 (2021).
27.   A. Z. Werner, *et al.*, Tandem chemical deconstruction and biological upcycling of poly(ethylene terephthalate) to β-ketoadipic acid by *Pseudomonas putida* KT2440. *Met. Eng.* **67**, 250-261 (2021).
28.   D. J. Ferraro, L. Gakhar, S. Ramaswamy, Rieske business: Structure-function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Comm.* **338**, 175-190 (2005).
29.   V. Venturi, F. Zennaro, G. Degrassi, B. C. Okeke, C. V. Bruschi, Genetics of ferulic acid bioconversion to protocatechuic acid in plant-growth-promoting *Pseudomonas putida* WCS358. *Microbiology-(UK)* **144**, 965-973 (1998).
30.   M. D. Wolfe, J. V. Parales, D. T. Gibson, J. D. Lipscomb, Single turnover chemistry and regulation of O-2 activation by the oxygenase component of naphthalene 1,2-dioxygenase. *J. Biol. Chem.* **276**, 1945-1953 (2001).
31.   A. Karlsson, *et al.*, Crystal structure of naphthalene dioxygenase: Side-on binding of dioxygen to iron. *Science* **299**, 1039-1042 (2003).
32.   S. M. Resnick, K. Lee, D. T. Gibson, Diverse reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. *J. Indust. Microbiol.* **17**, 438-457 (1996).
33.   D. J. Ferraro, A. Okerlund, E. Brown, S. Ramaswamy, One enzyme, many reactions: Structural basis for the various reactions catalyzed by naphthalene 1,2-dioxygenase. *IUCRJ* **4**, 648-656 (2017).
34.   O. Kweon, *et al.*, A new classification system for bacterial rieske non-heme iron aromatic ring-hydroxylating oxygenases. *BMC Biochem.* **9**, 11 (2008).
35.   J. K. Capyk, L. D. Eltis, Phylogenetic analysis reveals the surprising diversity of an oxygenase class. *J. Biol. Inorg. Chem.* **17**, 425-436 (2012).

333 36. K. V. Ambrose, *et al.*, Functional characterization of salicylate hydroxylase from the fungal endophyte *Epichloë festucae*. *Sci. Reports* **5**, 10939 (2015).

334 37. H. K. Chang, P. Mohseni, G. J. Zylstra, Characterization and regulation of the genes for a novel anthranilate 1,2-dioxygenase from *Burkholderia cepacia*
335   DBO1. *J. Bacteriol.* **185**, 5871-5881 (2003).

336 38. S. Jaganaman, A. Pinto, M. Tarasev, D. P. Ballou, High levels of expression of the iron-sulfur proteins phthalate dioxygenase and phthalate dioxygenase
337   reductase in *Escherichia coli. Protein Express. Purif.* **52**, 273-279 (2007).

338 39. Y.-J. Hou, Y. Guo, D.-F. Li, N.-Y. Zhou, Structural and biochemical analysis reveals a distinct catalytic site of salicylate 5-monooxygenase NagGH from Rieske
339   dioxygenases. *Appl. Environ. Microbiol.* **87**, e01629-20 (2021).

340 40. T. A. Rogers,A. S. Bommarius, Utilizing simple biochemical measurements to predict lifetime output of biocatalysts in continuous isothermal processes.
341   *Chem. Eng. Sci.* **65**, 2118-2124 (2010).

342 41. M. D. Wolfe, *et al.*, Benzoate 1,2-dioxygenase from *Pseudomonas putida:* Single turnover kinetics and regulation of a two-component Rieske dioxygenase.
343   *Biochemistry* **41**, 9611-9626 (2002).

344 42. T. Ohta, S. Chakrabarty, J. D. Lipscomb, E. I. Solomon, Near-IR MCD of the nonheme ferrous active site in naphthalene 1,2-dioxygenase: Correlation to
345   crystallography and structural insight into the mechanism of Rieske dioxygenases. *J. Am. Chem. Soc.* **130**, 1601-1610 (2008).

346 43. L. Holm, Using DALI for protein structure comparison. *Methods in Molec. Bio. (Clifton, N.J.)* **2112**, 29-42 (2020).

347 44. T. Fang,N. Y. Zhou, Purification and characterization of salicylate 5-hydroxylase, a three-component monooxygenase from *Ralstonia* sp. strain U2. *Applied*
348   *Microbiol. Biotech.* **98**, 671-679 (2014).

349 45. D. J. Ferraro, A. L. Okerlund, J. C. Mowers, S. Ramaswamy, Structural basis for regioselectivity and stereoselectivity of product formation by naphthalene
350   1,2-dioxygenase. *J. Bact.* **188**, 6986-6994 (2006).

351 46. J. A. Gerlt, *et al.*, Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim.*
352   *Biophys. Acta* **1854**, 1019-1037 (2015).

353 47. C. L. Colbert, *et al.*, Structural characterization of *Pandoraea pnomenusa* B-356 biphenyl dioxygenase reveals features of potent polychlorinated biphenyl-
354   degrading enzymes. *PloS One* **8**, e52550 (2013).

355 48. J. Qiu, *et al.*, Identification and characterization of a novel pic gene cluster responsible for picolinic acid degradation in *Alcaligenes faecalis* JQ135. *J. Bact.*
356   **201**, e00077-00019 (2019).

357 49. M. S. Rogers, J. D. Lipscomb, Salicylate 5-hydroxylase: Intermediates in aromatic hydroxylation by a Rieske monooxygenase. *Biochemistry* **58**, 5305-5319
358   (2019).

359 50. O. Pinyakong, H. Habe, T. Yoshida, H. Nojiri, T. Omori, Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of
360   *Sphingobium* sp. strain P2. *Biochem. Biophys. Res. Commun.* **301**, 350-357 (2003).

361 51. Y. Jouanneau, J. Micoud, C. Meyer, Purification and characterization of a three-component salicylate 1-hydroxylase from *Sphingomonas* sp. strain CHY-1.
362   *Appl. Environ. Microbiol.* **73**, 7515-7521 (2007).

363 52. F. P. Guengerich, Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* **14**, 611-650
364   (2001).

365 53. C. S. Harwood, R. E. Parales, The beta-ketoadipate pathway and the biology of self-identity. *Ann. Rev. Microbiol.* **50**, 553-590 (1996).

366 54. R. Jamarani, H. C. Erythropel, J. A. Nicell, R. L. Leask, M. Marić, How green is your plasticizer? **10**, 834 (2018).

367 55. Y. Fukuhara, *et al.*, Characterization of the isophthalate degradation genes of *Comamonas* sp. strain E6. *Appl. Environ. Microbiol.* **76**, 519-527 (2010).

368 56. R. N. Roy, Bioactive natural derivatives of phthalate ester. *Crit. Rev. Biotech.* **40**, 913-929 (2020).

369 57. J. K. Mahto, *et al.*, Molecular insights into substrate recognition and catalysis by phthalate dioxygenase from C*omamonas testosteroni. J. Biol. Chem.* **297**,
370   101416 (2021).

371 58. R. E. Parales, *et al.*, Substrate specificity of naphthalene dioxygenase: Effect of specific amino acids at the active site of the enzyme. *J. Bacteriol.* **182**, 1641-
372   1649 (2000).

373

374   Figure Captions

375   **Figure 1. Pathway for enzymatic conversion of TPA to PCA. (A)** Organization of the operon of genes encoding a TPA sensing transcription factor, TPA transporter
376   domain, dioxygenase, reductase, and dehydrogenase. **(B)** TPADO (orange) uses molecular oxygen to dihydroxylate TPA, and reducing equivalents are supplied by the
377   reductase (yellow). The product, DCD, is converted to protocatechuate via a Zn dependent dehydrogenase (blue).

378   **Figure 2. TPADO converts TPA to DCD.** (**A**) HPLC chromatograms of a reaction mixture containing TPADO, its reductase, NADH, and TPA in air after 0 (top) and 90 min.
379   The data illustrate disappearance of NADH with concomitant formation of new peaks at 12 min (assigned as DCD) and 13 min (NAD$^+$). Peak assignments were made
380   based on identical retention times and UV/vis spectra (insets) relative to known standards (**Fig S10-S11**), except for DCD, for which no standards were available. In
381   that case, the product was assigned based on (**B**) its mass spectrum, where the parent ion (m/z = 199.02) and fragments predicted to result from loss of water (m/z =
382   181.01), CO_2 (m/z = 155.01), or both (m/z = 137.03) were readily identified. (See also *SI Appendix*, **Fig. S8**.)

383   **Figure 3. TPADO permits only structurally conservative substitutions on TPA analogs.** Consumption of NADH and substrate analog following incubation in air with
384   the TPADO/reductase system was measured via HPLC. **(A-B)** HPLC chromatograms measured at the outset (top) and 90-minute endpoint (bottom) of the reaction
385   with 2-OH-TPA (**A**) and 2-NH$_2$-TPA (**B**). A star indicates the peak correlating to the adjacent UV/vis callout. Retention time and UV/vis spectra of the products for both
386   the 2-OH-TPA and 2-NH$_2$-TPA reactions match those of PCA, though the latter achieved substantially lower coupling. The observed PCA product for both 2-substituted
387   TPA analogs definitively indicates 1,2-dioxygenation. **(C)** Analogs 4-FBA (**1**) and 4-CBA (**2**) showed no NADH consumption, while 4-NBA (**3**) showed 100% consumption
388   of NADH with no aromatic product formation (0% coupled). 2-NH$_2$-TPA (**4**) and 2-OH-TPA (**5**) showed 22% and 100% coupling, respectively.

389   **Figure 4. TPADO structure shows TPA binding interactions. (A, C)** Top view of the $\alpha_3\beta_3$-heterohexamer. The β subunits are colored in shades of green and the α
390   subunits are colored in pink, yellow, and blue. Iron-sulfur cluster and mononuclear iron cofactors shown as spheres, exhibiting a head-to-tail inter-subunit pathway
391   for electron transfer. **(B, D)** Bottom view shows the trimeric architecture of the α subunits. The bound TPA substrate is shown in yellow for the blue colored α subunit.
392   Salt bridges with R309 and R390 and a hydrogen bonding interaction with S243 are shown. The ferrous ion, which is coordinated by the side chains of residues H210,
393   H215, and D356, is connected to the Rieske-type [2Fe-2S] cluster of a neighboring α subunit (pink). **(E)** Superposition of two α subunits of the TPA-bound structure
394   shows two different conformations of an α-helix within the active site which contains the ferrous ion coordinating histidines H210 and H215. The pink α-helix is
395   stabilized by crystal contacts, whereas the blue α-helix is kinked to allow the coordination of the ferrous ion by H215 when bound to TPA.

396   **Figure 5. Active site structure of TPADO with 2-OH-TPA bound is shown superimposed on the structure of (A) NagGH (A) and (B, C) product-bound NDO.** (A)
397   Superposition of the blue colored TPADO α subunit with the α subunit of its closest structural homolog salicylate 5-monooxygenase NagGH (grey) which has no
398   substrate bound (PDB ID 7C8Z) highlights the unique roles played by S243 and R390 in accommodating the substrate (39). (**B**) Superposition with the α subunit of
399   NDO (orange) with its dioxygenated product bound (PDB ID 1O7P) (31). The lower label numbers belong to NDO. (**C**) Closer views of the overlay of NDO product
400   (orange) with 2-OH-TPA substrate and their relative positions to the ferrous ion. Overlap of the 1,2-diol of portion of the NDO product with the ring-1,2 carbons of 2-
401   OH-TPA suggests these are the sites of hydroxylation.

402   **Figure 6. Sequence relationships of aryl carboxylate Rieske oxygenases. (A)** A single cluster from a full SSN generated for Pfam00848 (*SI Appendix*, **Fig. S22**)
403   containing 1,812 RO α subunit sequences, including TphA2 (α-TPADO) and others with characterized functions relating to aryl carboxylate oxygenation (indicated by
404   circles with color), is shown. To the right of the cluster are structures illustrating where the enzymes are proposed to hydroxylate their aryl carboxylate substrates
405   (O$_2$-derived hydroxyl groups highlighted in yellow). Intermediates are shown in brackets to clarify where hydroxyls are installed prior to their spontaneous breakdown
406   to yield CO_2, H_2O, and catechol (AhdA1c) or CO_2, NH_3 (protonating to form NH$_4^+$ at neutral pH) and catechol (AndAc). **(B)** A multiple sequence alignment of the five
407   functionally characterized homologs from (A). Orange boxes and arrows indicate residues proposed to interact with a substrate. R309 is highly conserved in these 5
408   sequences and across the cluster, suggesting it has a functionally significant role. N224, S243, and R390 are only found in TPADO. I290 is conserved in TphA2, AhdA1c,
409   and AndAc.

410   **Scheme 1. A canonical mechanism for ROs adapted for TPADO. (A)** The fully reduced, resting active site (1) has a ferrous hexacoordinate mononuclear iron and
411   reduced Rieske cluster (Rieske^{red}). TPA binding (2) displaces a water molecule and primes the enzyme for reaction with O$_2$. The positions of residues involved in
412   anchoring the substrate (R390, R309) are indicated. O$_2$ binds in an end-on fashion to form (3) the ferric-superoxy intermediate. One electron donation from the Rieske
413   cluster forms a peroxy adduct, which can either heterolytically cleave to yield (4) a high valent ferryl or (5) release H_2O_2 following di-protonation in the unproductive,
414   uncoupling pathway. A species resembling (3) or (4) (shown together in dashed box) may react with TPA to form (6) an epoxide and ferric-hydroxy intermediate. The
415   ferric-OH attacks the epoxide, which opens and protonates to form the product. Re-reduction and hydration of the active site occurs as the product departs,
416   returning the active site to its starting state. The *cis-1S,2R*-DCD isomer is drawn as the presumptive product, based on the observed products of the reactions
417   between NADH/O$_2$ and 2NH$_2$-TPA or 2OH-TPA. **(B)** Hydroxylation of the ring-1,2 carbons of 2-NH$_2$-TPA yields an intermediate analogous to DCD. Arrows show the
418   pathway for re-aromatization and loss of CO_2 + NH_3.

A *Comamonas* sp. strain E6

1 kbp

☐ *tphR* transcription factor
◼ *tphC* TPA transport domain
◼ *tphA2* dioxygenase α
◼ *tphA3* dioxygenase β
◼ *tphB* dehydrogenase
◼ *tphA1* reductase

B

NADH → NAD⁺

PCA

Ox — Flavin 2Fe—2S — Red

2Fe—2S Fe²ᐟ³⁺ — Ox / Red

TPA — O₂ → DCD

NADH CO₂ — Zn²⁺ — NAD⁺

A

2-OH-TPA

PCA

B

2-NH2-TPA

PCA

C

4-FBA

NADH consumed 0%
Coupling N/A

4-CBA

NADH consumed 0%
Coupling N/A

4-NBA
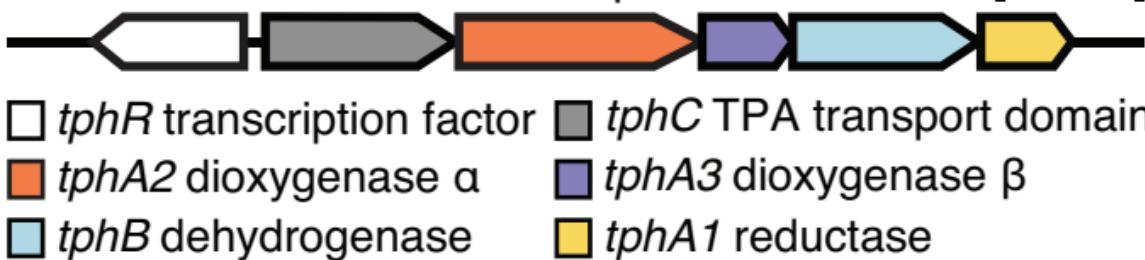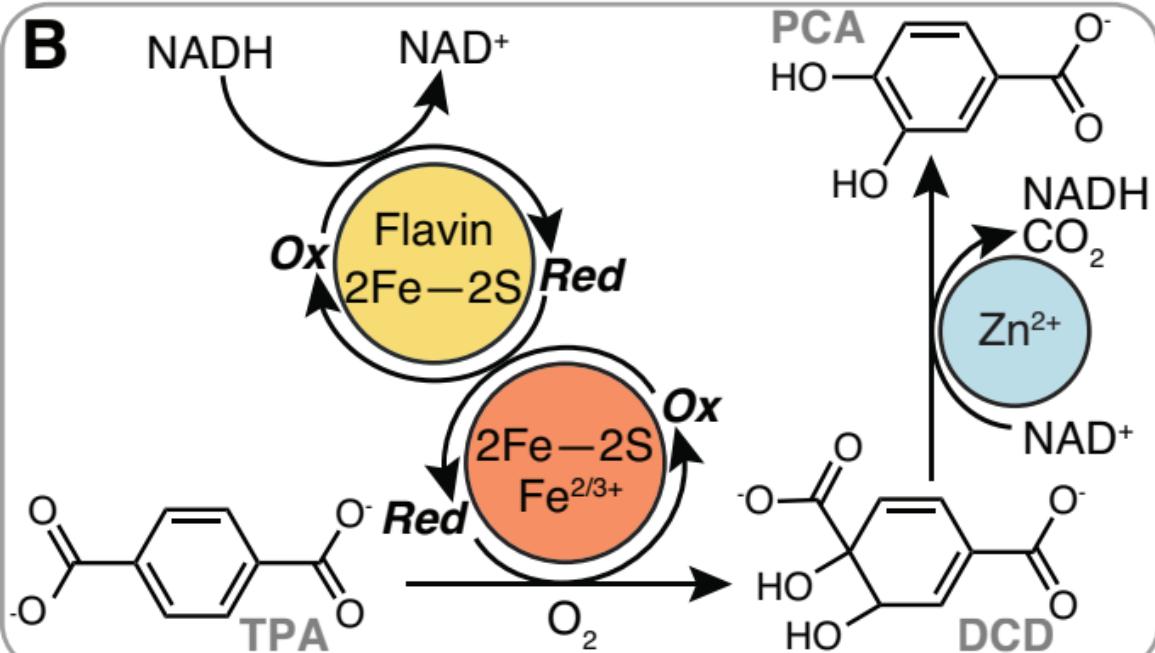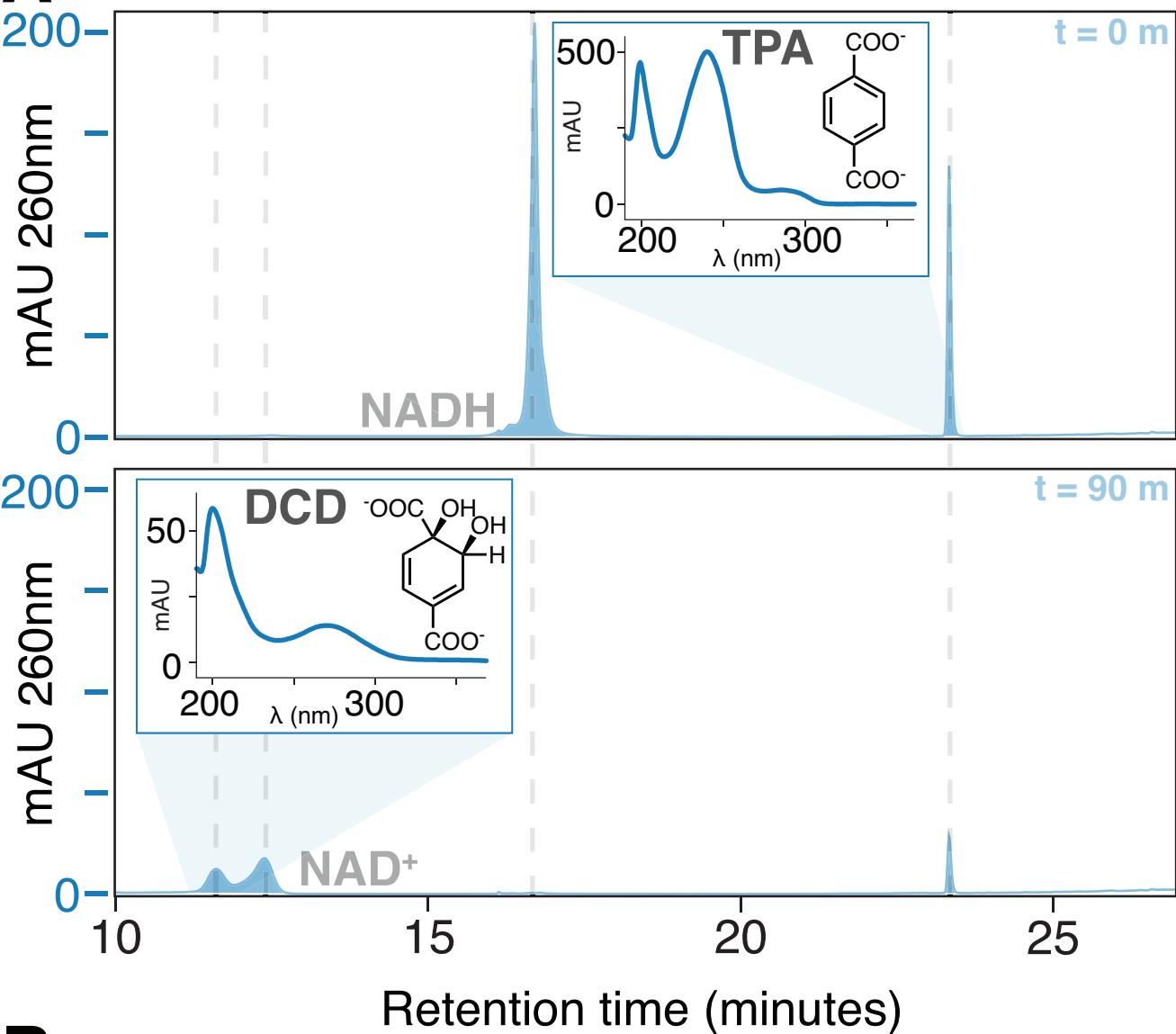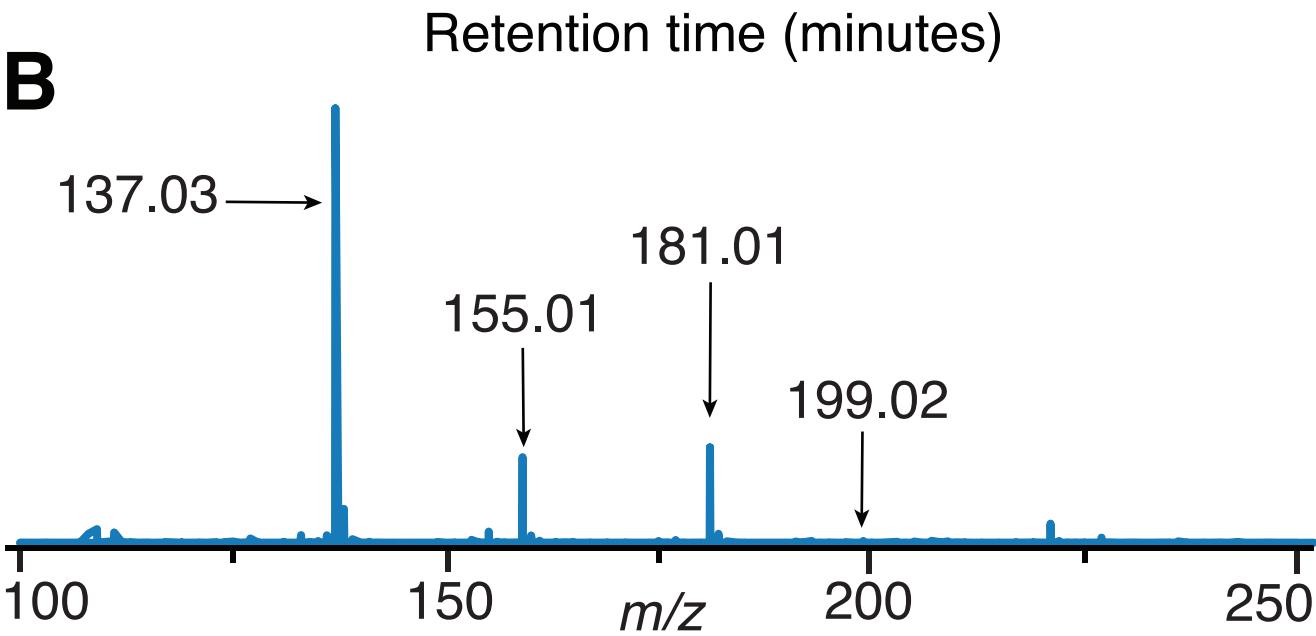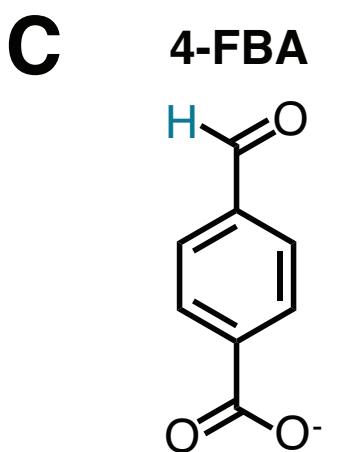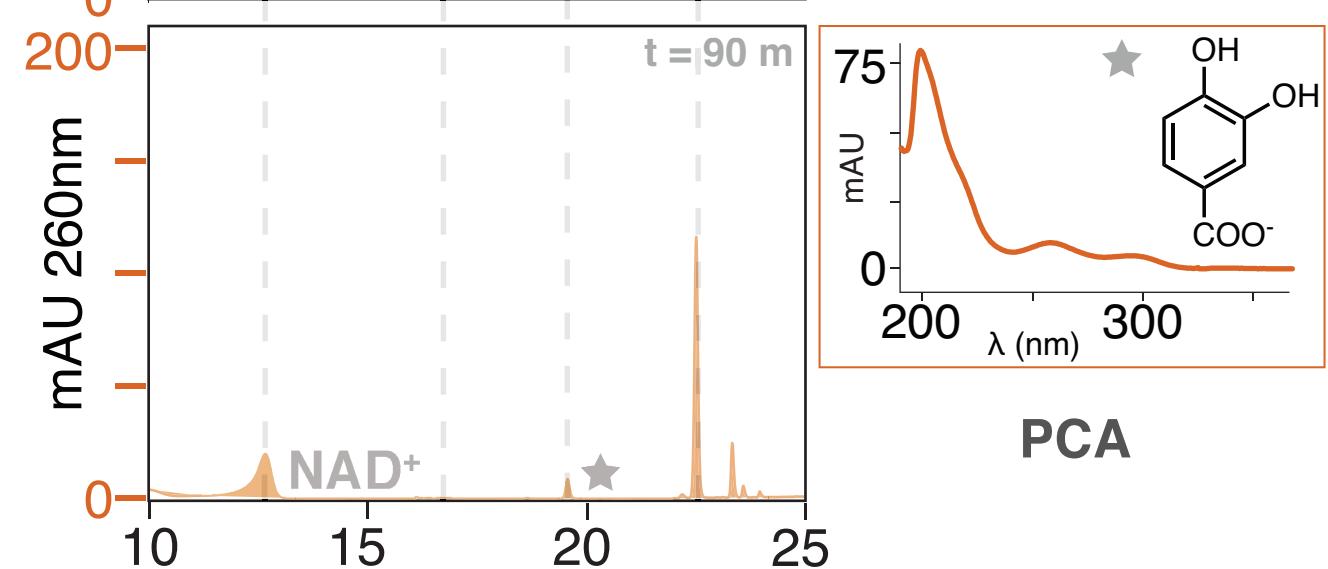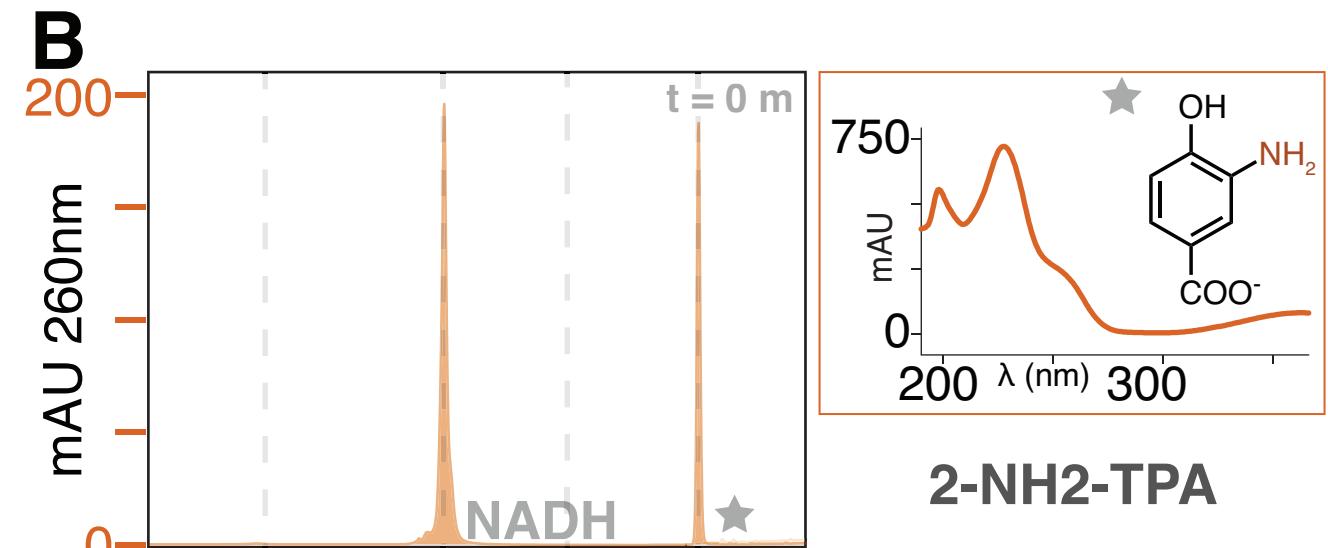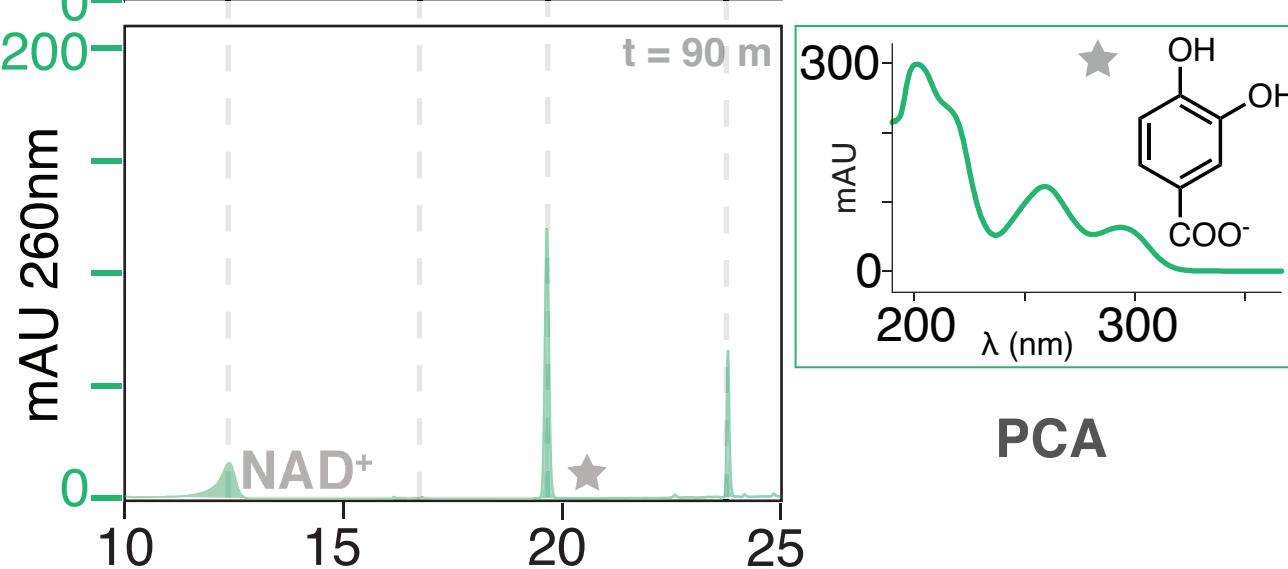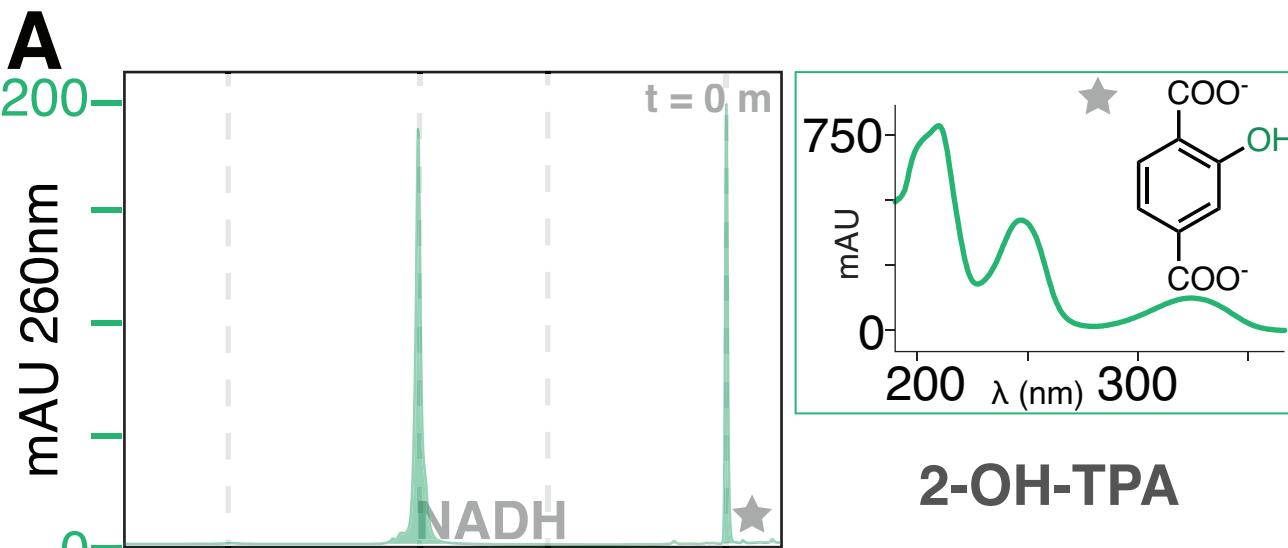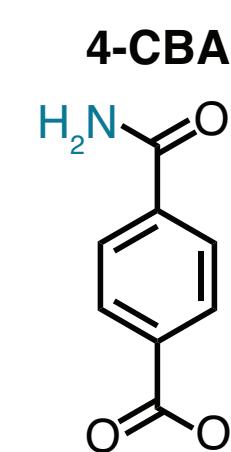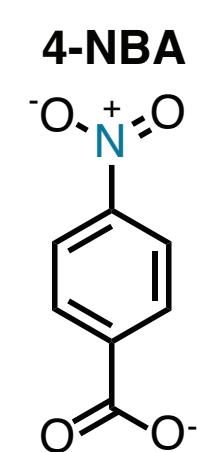
NADH consumed 100%
0% coupled

2-NH₂-TPA
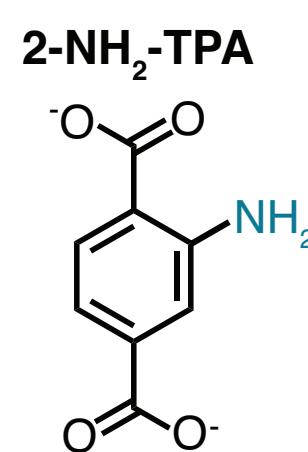
NADH consumed 100%
22% coupled

2-OH-TPA
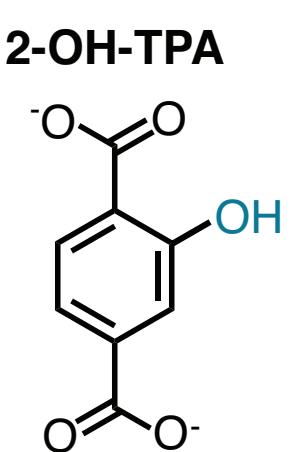
NADH consumed 100%
90% coupled

A

12.2Å

B

I290   S243
R309
R390
D356   H215
        H210   D207
                H105
        C102
        C82
                H84

C   D

E

D356
H210   H215
H215
H210

A

B

	215												**245**
TphA2	H M F F T T F E L N R L S Q K G G V I V D E S G G H V S Y S M I D R
AhdA1c	H L F H N T F G L Y R S T Q T G K S L M D S E K . R H S L L Y S I A A
AndAc	H L F H T T F N I F R V G M K A R S I P D A N H G L H S I I T V T K T
NagG	H T W F S T F G L W R A D N K S E L K M D A K F . R H A A M I S T R G
HybB	H T W F V T F G L W R A D N K S A L R M D D R H . A A M I S T R G
PicB3	H S F L V V F G L L V A G N K S Q M V V D P V H G R H G Y M A S A K S

	285													**315**
TphA2	G V T L Q I L S V F P G F V L Q Q I Q N S I A V R Q L L P K S I S S S
AhdA1c	G I T L V I L A V F P N L V L Q Q I Q N T L A V R Q I V P K G Q E A F
AndAc	D C T N H I Q P I F P Q L V I Q Q I H N T L V A R Q I L P K G P D N F
NagG	P T A V . M T T I F P S V I I Q Q V N S V S T R H I Q P N G H G S F
HybB	P T A V . M M T I F P S V I I Q Q V N S V S T R H I Q P N D E G S F
PicB3	P W S V T M Q T I W P N L I V Q R E M N T L G V R Q I V P N G P N S L

	375														**405**
TphA2	A V I E M G G D H E G S S E G R A T E T S V R G F W K A Y R K H M G Q
AhdA1c	S F I A M G G G R A E D A D H L V T E G A I V G F W D S Y E K L V G F
AndAc	S V I E M S R G N P E Q Q D T V I T E S L I R K F W V G Y Q K L M G Y
NagG	T V I E M G G H E I G D T D H M V T E T L I R G M Y D Y W R K V M G E
HybB	A V A E L G G Y D V E N T E H M V T E T L I R G M Y E Y W E Y K R Q M E I
PicB3	N I I K L D E S K L G T S N T L I S E A A I R S M Y Q Y Y R Q V M D L