Article

Substrate specificity of a branch of aromatic dioxygenases determined by three distinct motifs

Chengsen Cui [1,2,8], Lu-Jia Yang[1,2,8], Zi-Wei Liu [3,8], Xian Shu [4,8], Wei-Wei Zhang[5,6], Yuan Gao [5,6], Yu-Xuan Wang[3], Te Wang[3], Chun-Chi Chen [3,7], Rey-Ting Guo [3,7] ✉ & Shu-Shan Gao [1,2,4] ✉

The inversion of substrate size specificity is an evolutionary roadblock for proteins. The Duf4243 dioxygenases GedK and BTG13 are known to catalyze the aromatic cleavage of bulky tricyclic hydroquinone. In this study, we discover a Duf4243 dioxygenase PaD that favors small monocyclic hydroquinones from the penicillic-acid biosynthetic pathway. Sequence alignments between PaD and GedK and BTG13 suggest PaD has three additional motifs, namely motifs 1-3, distributed at different positions in the protein sequence. X-ray crystal structures of PaD with the substrate at high resolution show motifs 1-3 determine three loops (loops 1-3). Most intriguing, loops 1-3 stack together at the top of the pocket, creating a lid-like tertiary structure with a narrow channel and a clearly constricted opening. This drastically changes the substrate specificity by determining the entry and binding of much smaller substrates. Further genome mining suggests Duf4243 dioxygenases with motifs 1-3 belong to an evolutionary branch that is extensively involved in the biosynthesis of natural products and has the ability to degrade diverse monocyclic hydroquinone pollutants. This study showcases how natural enzymes alter the substrate specificity fundamentally by incorporating new small motifs, with a fixed overall scaffold-architecture. It will also offer a theoretical foundation for the engineering of substrate specificity in enzymes and act as a guide for the identification of aromatic dioxygenases with distinct substrate specificities.

Aromatic rings are prevalent structures in feedstocks, but their aromaticity renders them unreactive under mild reaction conditions. Only a handful of organic reactions under severe conditions have been developed to directly cleave carbon-carbon bonds of aromatic rings[1,2]. On the other hand, Nature hires aromatic dioxygenase for microbial phenol-catabolism[3,4], which includes intradiol dioxygenase, extradiol dioxygenase, hydroquinone dioxygenase, and 2,5-hydroxy-pyridine dioxygenase (Fig. 1a). Those dioxygenases degrade small monocyclic aromatic intermediates[3], which is a crucial step in the aromatic hydrocarbon biodegradation (Fig. 1a). Aromatic dioxygenases could

[1]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China. [2]National Center of Technology Innovation for Synthetic Biology, Tianjin, China. [3]State Key Laboratory of Biocatalysis and Enzyme Engineering, Hubei Hongshan Laboratory, School of Life Sciences, Hubei University, Wuhan, China. [4]State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [5]National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China. [6]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China. [7]Zhejiang Key Laboratory of Medical Epigenetics, Department of Immunology and Pathogen Biology, School of Basic Medical Sciences, Hangzhou Normal University, Hangzhou, China. [8]These authors contributed equally: Chengsen Cui, Lu-Jia Yang, Zi-Wei Liu, Xian Shu. ✉e-mail: guoreyting@hubu.edu.cn; gaoss@tib.cas.cn

overcome the stabilizing resonance energy derived from aromaticity and transform different aromatic intermediates generated by peripheral pathways to βketoadipate and intermediates into the tricarboxylic acid cycle[5]. For instance, the 4-hydroxyacetophenone catabolic pathway in *Pseudomonas fluorescens* ACB involves a nonheme-iron (II)-dependent hydroquinone dioxygenase HapCD[4,6]. Recently, fungal aromatic dioxygenases have been characterized, as fungi are significant members of lignin biodegradation in Nature[3]. Intradiol dioxygenases, protocatechuate dioxygenase, and hydroquinone dioxygenase have been characterized by *Aspergillus niger*[7,8].

In addition to their role in catabolism, aromatic cleavage is also proposed as a valuable strategy for the biosynthesis of fungal natural products[9,10]. Recently, two DUF4243 domain-containing dioxygenases BTG13 and GedK were characterized from the biosynthetic pathway of fungal natural products beticolin and geodin, respectively, which were reported to catalyze the ring-opening of anthraquinones through hydroquinone cleavage[11,12] (Fig. 1c). Notably, the reactive site of BTG13 is built by atypical iron coordination with four histidine residues, an unusual carboxylated-lysine, and water. Hydroquinone cleavage also involves the biosynthesis of fungal tetronate polyketides, such as penicillic acid (PA, **1**, Fig. 1b), pesthetoxin, nodulisporacid, lowdenic acid, etc., which all displayed diverse biological activities[9,13].

Penicillic acid (PA), a famous lactone mycotoxin produced by over 100 fungal species, was first discovered and named in 1913[14,15]. Although well-known as a food and fodder contaminant with carcinogenic toxicity, PA also showed antibacterial, antiviral, and antidiuretic pharmacological and phytotoxic properties[14,16-18]. PA has toxicity to poultry and livestock, and even probably to humans due to the transmission of the food chain, as it is one of the major mycotoxins in moldy feedingstuff[19]. The understanding of its biosynthesis would benefit the reduction of mycotoxin contaminations. Initial feeding experiments using ^{14}C-labelled substances demonstrated that orsellinic acid (OA, **2**) and 1-4-dihydroxy-6-methoxy-2-methylbenzene (**4**) were intermediates (Fig. 1b and Supplementary Fig. 1)[20,21]. However, the detailed biosynthetic pathway of **1** as well as the aromatic ring-opening mechanism remains obscure.

In this study, the biosynthetic gene cluster (BGC) and biosynthetic logics of **1** in *A. westerdijkiae* have been described through targeted gene knockout, substrate complementation, and biochemical characterization. We discover a DUF4243 domain-containing dioxygenases PaD catalyzed aromatic cleavage of the small monocyclic hydroquinones, which is distinct with reported Duf4243 dioxygenases GedK and BTG13. Through the analysis of crystallographic structures and mutagenesis of PaD and BTG13, the structural basis for the substrate specificity has been explored.

Results

Identification of the *pa* BGC

Although scientists discovered PA over 100 years, the *pa* BGC remains unknown. We isolated **1** from *Aspergillus westerdijkiae* and confirmed its structure by comparing the NMR spectra with reported data (Supplementary Figs. 17−18)[22]. Previous biogenetic analysis and labeling studies suggested orsellinic acid (OA, **2**)[23,24], 6-methyl-1,2,4-benzenetriol (**3**), and 1-4-dihydroxy-6-methoxy-2-methylbenzene (**4**) as on-pathway intermediates (Fig. 1b)[20,23]. The previously proposed biosynthetic pathway suggests that a nonreducing polyketide synthase (NR-PKS) for OA biosynthesis, a putative oxidative decarboxylase, and a methyltransferase should be encoded in the *pa* BGC[20,23] (Fig. 2f).

We used the OA synthase OpS1 from *Beauveria bassiana* as a query to search the *pa* BGC[25], which identified six NR-PKS encoding BGCs from the *A. westerdijkiae* genome. Screening six BGCs led to the identification of a candidate BGC containing seven genes. It encodes a typical NR-PKS (PaA), a flavin-dependent hydroxylase (PaB), an *O*-methyltransferase (PaC), a DUF4243 domain-containing protein (PaD), a short-chain dehydrogenases/reductases (S.D.R., PaE), a major facilitator superfamily transporter (PaF), and a transcription regulator (PaG) (Fig. 2a and Supplementary Table 3). Further genome mining towards more **1**-producing fungi revealed numerous homologous BGCs to the *pa* cluster (Supplementary Fig. 2).

PaA is a canonical NR-PKS consisting of starter-unit acyl transferase (SAT), ketosynthase (KS), acyltransferase (AT), product template (PT), twin acyl carrier proteins (ACP), and *N*-terminal thioesterase (TE) domains (Supplementary Fig. 3)[26]. The PaA domain organization is
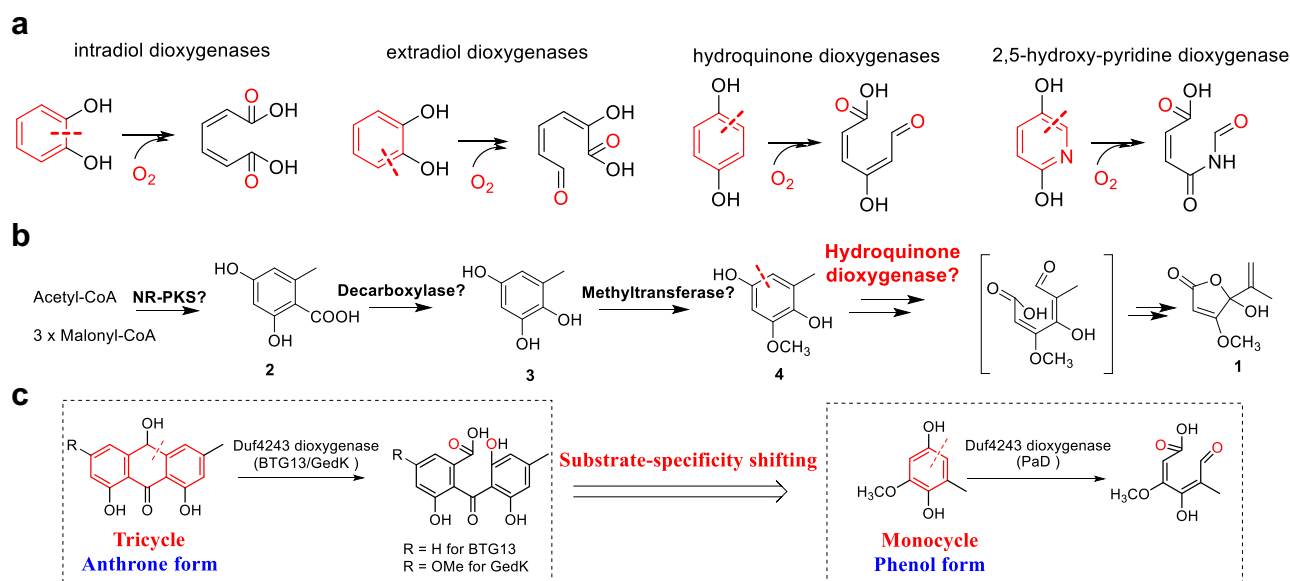

Fig. 1 | Microorganism-derived aromatic dioxygenases. a Aromatic cleavage by intradiol dioxygenase, extradiol dioxygenase, hydroquinone dioxygenase, and 2,5-hydroxy-pyridine dioxygenase. **b** Biosynthetic pathway of penicillic acid proposed by previous studies. A nonreducing polyketide synthase (NR-PKS) generates orsellinic acid (OA, **2**), followed by a decarboxylase and methyltransferase catalyze the formation of hydroquinone **4** (6-methylbenzene-1,2,4-triol). **c** Substrate-specificity shifting of different Duf4243 dioxygenases with their specific substrates. BTG and GedK catalyze the oxidative ring opening of the tricycle hydroquinone, while PaD accepts the monocyclic hydroquinone.
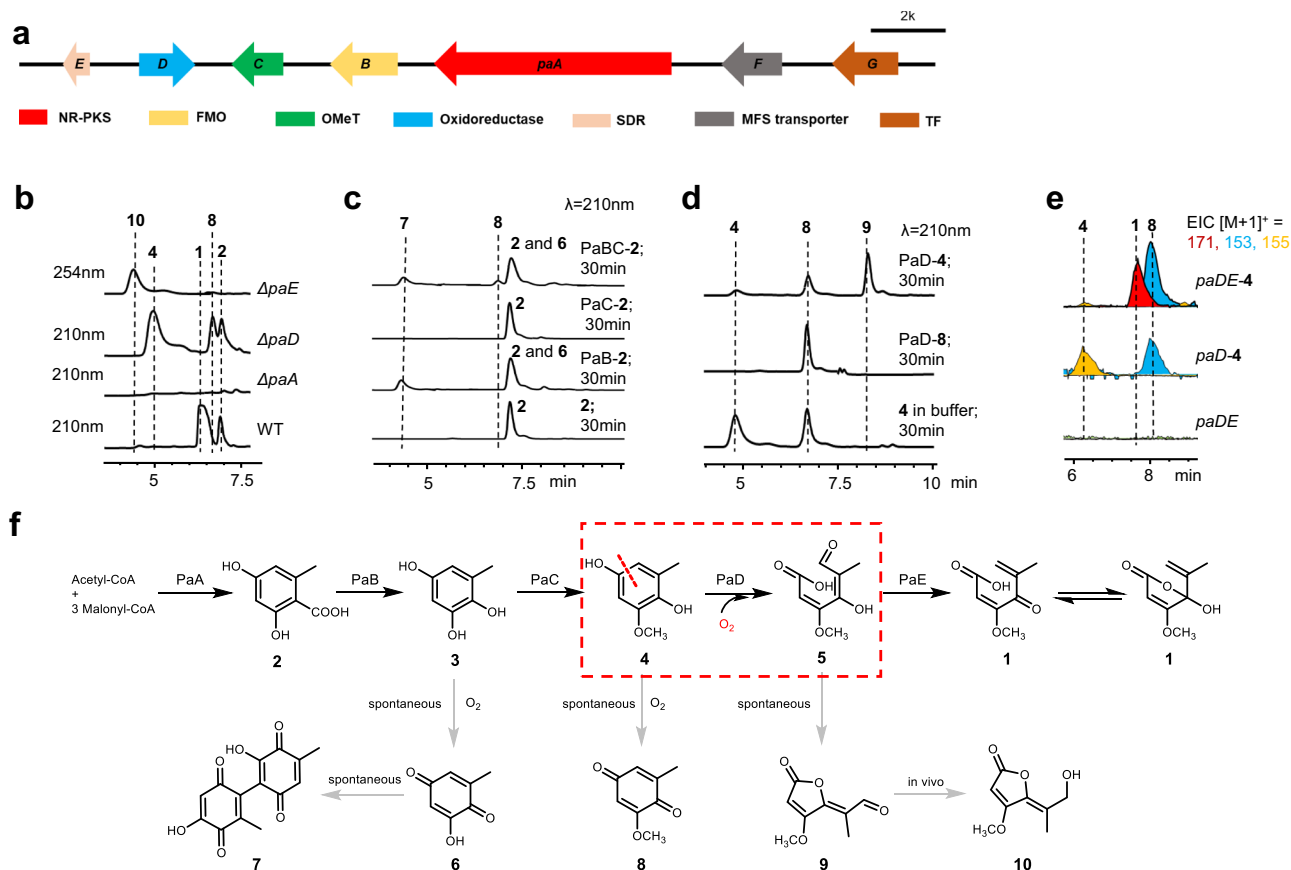
Fig. 2 | Proposed biosynthetic pathway of penicillic acid (1). a Penicillic acid (**1**) biosynthetic gene cluster (BGC) in *A. westerdijkiae*. NR-PKS, nonreducing polyketide synthase; FMO, flavin-containing monooxygenase; OMeT, *O*-methyltransferase; SDR, short chain reductase; MFS, major facilitator superfamily; TF, transcription factor. **b** Product profiles of wild-type and mutants (*ΔpaA*, *ΔpaD*, and *ΔpaE*) of *A. westerdijkiae*. **c** In vitro assays of PaB and PaC with **2**. **d** In vitro assays of PaD with **4** or **8**. **e** LC-MS analysis of product profiles of compound **4** fed to *S. cerevisiae* expressing *paD* and *paE*. *S. cerevisiae* containing empty pXW55 and pXW06 vectors was used as a control. **f** Proposed biosynthetic pathway of **1**.

similar to the reported OA synthases (Supplementary Fig. 3). When we knocked out *paA* through double crossover recombination by using the hygromycin-resistance gene *hyg* as a marker (Supplementary Fig. 4), the resultant mutant lost the production of **1** (Fig. 2b). Furthermore, with the chemical complementation of *ΔpaA* with **2**, the production of **1** was restored (Supplementary Fig. 5). All those results confirmed i) the identified *pa* cluster is responsible for the PA biosynthesis and ii) PaA is responsible for the biosynthesis of the first biogenetic intermediate OA (**2**), consistent with previous labeling studies[23].

Characterizing 4 as the substrate for the hydroquinone cleavage

A decarboxylase and methyltransferase were proposed to convert the PKS product **2** to **4**, the substrate proposed for the hydroquinone cleavage (Fig. 1b and Supplementary Fig. 1)[26,27]. Thus, the proposed decarboxylase PaB was purified from *E. coli* BL21 (Supplementary Fig. 6) and incubated with **2** and FAD. The reaction yielded two new products, **6** (MW = 138) and its putative oxidative dimer **7** (MW = 274) (Fig. 2c, and Supplementary Fig. 38). **6** was not isolable due to its instability. But **7** was successfully purified from large-scale assays and further characterized by NMR data (Supplementary Table 4 and Figs. 23–26). Based on the structure of **7** and the MS of **6** (two units less than that of **3**), we proposed **6** as the spontaneous oxidative product of **3** (Fig. 2f), and a similar oxidative sequence from **3** to **7** is present in the biochemical studies for oosporein enzymes[25]. Thus, PaB is an oxidative decarboxylase for converting **2** to **3**.

To demonstrate PaC catalyzing the following methylation of **3** (Fig. 2f), we purified PaC from *E. coli* BL21 (Supplementary Fig. 6). As **3** and **6** were not isolable, we used **2** incubating with PaB and PaC. However, in addition to **6** and **7**, a new compound **8** was observed instead of **4** (Fig. 2c). NMR and MS characterization confirmed **8** is indeed *O*-methylated and an oxidative product of **4** (Supplementary Figs. 27, 28). Thus, we assumed that PaC catalyzes **3** to generate **4**, which spontaneously oxidizes to **8** due to its highly active phenol structure (Fig. 2f).

To confirm the hypothesis that **4** is the actual PaC product as well as the substrate for the following hydroquinone cleavage, we performed gene knock-outs of the rest genes in the *pa* BGC. The *paD* knock-out accumulated three major compounds **2** and **8**, as well as the missing product **4** in the reaction of PaC, which are all verified by NMR characterization (Fig. 2b and Supplementary Figs. 21, 22). As expected, the spontaneous transformation from **4** to **8** can be observed in PBS buffer (Fig. 2d), confirming that **4** oxidizes to **8** in the in vitro reactions. Both complementations of equal amounts of **4** and **8** to *ΔpaA* restored the production of **1**. However, the titer of **1** is significantly higher in the feeding of **4** (Supplementary Fig. 5), hinting that the slow restoration of **1** in the feeding of **8** is due to the endogenous reduction of **8** to **4**. Finally, when we incubated **2** and adenosylmethionine (SAM) with PaC, no new product was detected (Fig. 2c), excluding the promiscuity of PaC and confirming the reaction sequence from PaB to PaC to generate **4** for the following hydroquinone cleavage.
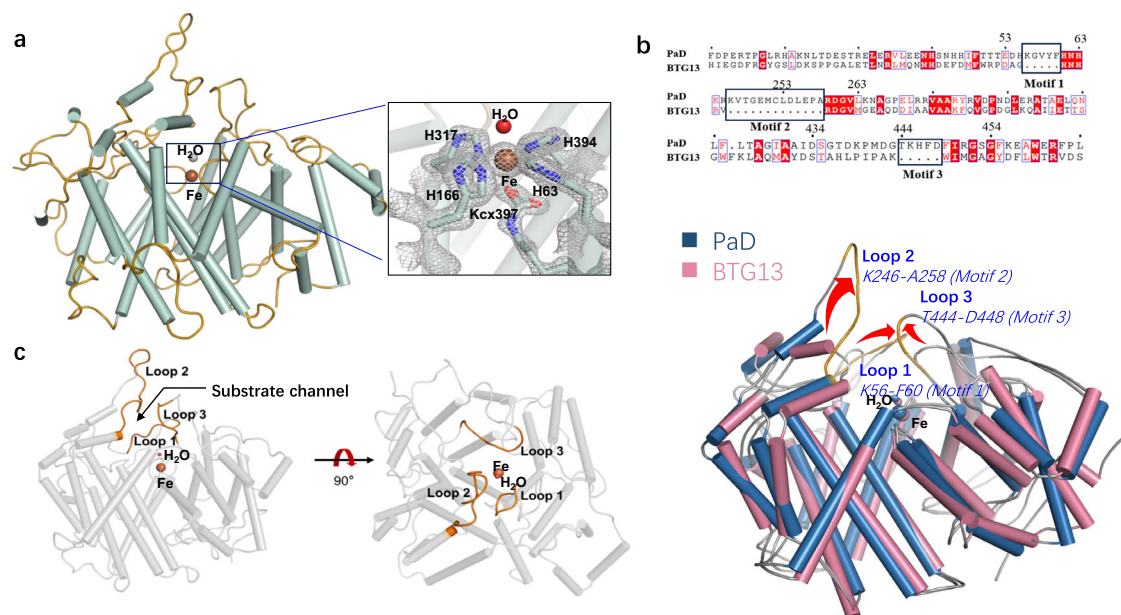
Fig. 3 | Crystal structure of PaD with a lid-like tertiary structure. a Crystal structure of PaD (PDB: 8Z4Q). H63, H166, H317, H394, K397, and one water molecule coordinate the iron. The electron density 2Fo-Fc maps of the ligands coordinated with Fe are shown with a gray-colored mesh, and contoured at 2.5 σ. **b** Sequence and structure comparison of PaD with BTG13. Structural comparison was depicted by the overlay of the overall structure of BTG13 and PaD. The three distinct motifs 1-3 are marked with black boxes. The three loops 1-3 determined by motifs 1-3 are marked in gold. **c** The locations of loops 1-3 in PaD structure viewed from different positions.

Identifying PaD as a hydroquinone dioxygenase

The gene knock-out of *paD* (Fig. 2b) suggests PaD catalyzing the hydroquinone cleavage of **4**. To confirm PaD's role, we incubated **4** with PaD purified from *E. coli* without adding any cofactors. In addition to **8**, the reaction showed the emergence of a new compound **9** (Fig. 2d), characterized as a lactone without aromaticity by NMR data (Supplementary Table 5 and Figs. 29−32). **9** is likely to be a shut product formed by spontaneous lactonization of **5**, the linear dioxygenation-product generated by PaD (Fig. 2f). Indeed, feeding **9** to *ΔpaA* mutant cannot restore production of **1** but developed a product **10** with the terminal aldehyde reduced (Supplementary Figs. 5,33,34,42), potentially mediated by an endogenous reductase. Thus, PaD catalyzes the aromatic cleavage of **4** via a dioxygenase pathway.

Finally, the *paE* knock-out also accumulated a significant amount of **10**, suggesting that PaE is the immediate enzyme after PaD on the pathway. Since **5** is not isolable, we fed **4** to *Saccharomyces cerevisiae* BJ5464 containing both *paE* and *paD* to demonstrate the function of SDR-like protein PaE. The whole-cell transformation clearly shows the yield of **1**, suggesting that PaE is a dual-function protein responsible for the reduction and dehydration of **5** to generate the linear **1** (Fig. 2e). Finally, spontaneous cyclization of linear **1** would result in cyclic **1** in the intracellular environment. In conclusion, we totally clarified the biosynthetic pathway of PA, and identified PaD, a Duf4243 dioxygenase, catalyzing the critical hydroquinone cleavage of **4**.

A lid-like tertiary structure composed of loops 1-3 (motifs 1-3) in *apo*-PaD structure

Identifying PaD as a hydroquinone dioxygenase raises a question about its difference from the recently characterized GedK and BTG13, which are also Duf4243 dioxygenases and catalyzes C_{4a}–C_{10} bond cleavage of anthraquinone under the cooperation with a reductase[11,12]. However, unlike PaD which accepts the monocyclic hydroquinone, BTG13 and GedK demonstrated activity towards the bulky tricyclic hydroquinone (Fig. 1c). Sequence alignment revealed low similarity between PaD and BTG13/GedK, with an identity of less than 30%. Notably, the alignment indicated the absence of three distinct motifs in GedK and BTG13: K56~F60 (motif 1), K246~A258 (motif 2), and T444~D448 (motif 3) in PaD (Fig. 3b and Supplementary Fig. 10). To confirm that motifs 1-3 are essential for the PaD specificity, seven modified proteins by truncating motifs 1-3 were purified and all displayed no activity towards **4**. We also synthesized 7 different genes of modified BTG13 by introducing motifs 1-3 (Supplementary Table 7 and Fig. 45). However, soluble protein could be obtained from only one gene that contains motif 1, which shows no activity towards **4** (Supplementary Table 7 and Fig. 45).

To identify the amino-acid sequences of motifs 1-3 in the PaD structure, we acquired the X-ray crystallographic structure of ligand-free -PaD. The structure of PaD consisted of 22 α-helices (α1 to α22) and four 3_{10} helices (η1-η4) (Fig. 3a). Like BTG13, the active site of PaD was formed by an unconventional iron ion that interacts with four conserved histidine residues (H63, H166, H317, and H394), an uncommon carboxylated-lysine (K397), and one water molecule (Fig. 3a, Supplementary Figs. 7 and 12). Previous density functional theory (DFT) calculations of BTG13 suggest the carboxylated-lysine could increase the electron-donating ability of Fe(II), thereby facilitating the formation of Fe(III)−$O_2^{\cdot-}$ species, which is essential for the following C-C bond cleavage[28].

The primary and higher-order structures of the two proteins exhibit significant similarity, except for three additional loops, referred to as loop1, loop2, and loop3, respectively, which are structurally determined by motifs 1-3 (Fig. 3b). Loops 1-3 converge at the top of the active pocket (Fig. 3b): Loop 2 noticeably extends upwards, while loops 1 and 3 stretch towards the center above, forming a narrow substrate channel and dramatically increasing the depth of the active pocket (Fig. 3b, c). As shown in Fig. 3c, the Fe center is located directly below the catalytic active pocket. Loops 1 and 2 are spatially close to each other and the opening of the pocket is significantly constricted, which could impede the entry of large substrates (Fig. 3c). In conclusion, loops 1-3 form a lid, resulting in a significant alteration of the tertiary structure of the upper sub-structure.
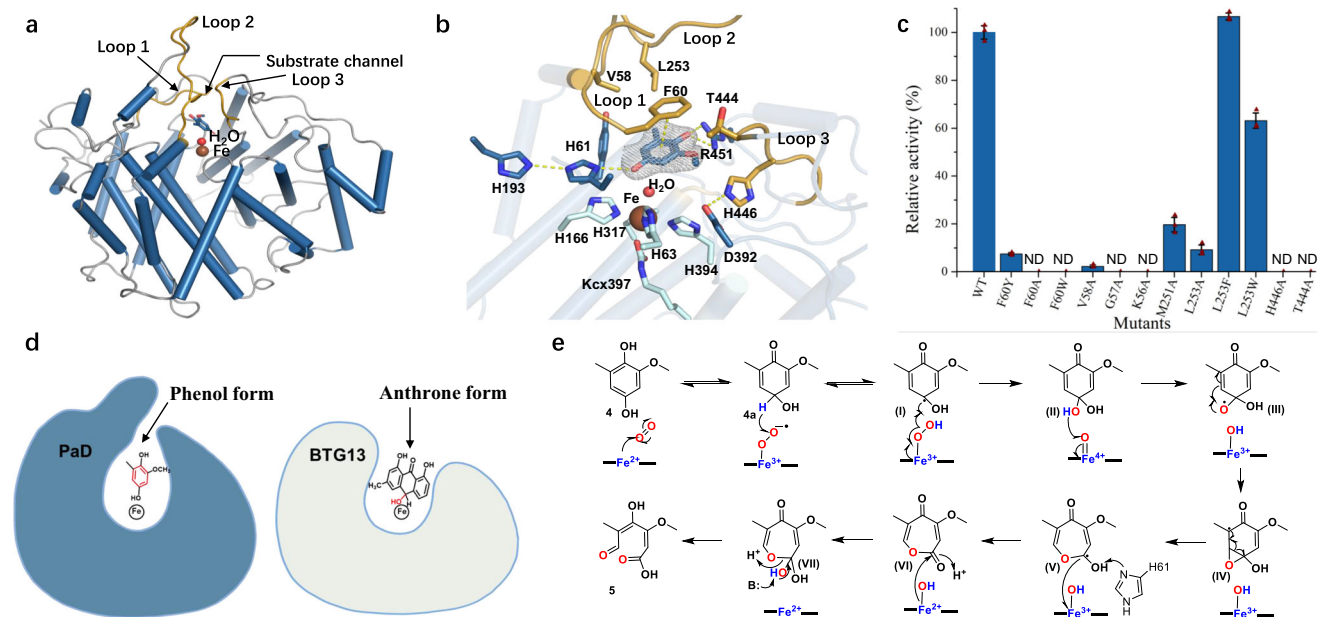
Fig. 4 | Crystal structure analysis of PaD-4 complex. a Overall structure of PaD-**4** complex (PDB: 8Z4R). Loops 1-3 are marked in gold. **b** The active site view of PaD-**4** complex. The electron density *2Fo-Fc* maps of substrate **4** in PaD are shown with a gray-colored mesh, and contoured at 1 σ. **c** Relative activities of PaD mutants located on the loops 1-3. The yield of **9** is quantified, defining the activity of wild-type PaD at 100%. Three biological parallel (*n* = 3) replicates are performed and presented as triangular points. The error bars are presented as standard deviation (SD). ND: not detected. Source data are provided as a Source Data file. **d** Comparison of the active-site shape and the substrate-binding mode in PaD and BTG13. The BTG13 complex structure was stimulated by Discovery Studio Client v19.1.0. **e** Proposed catalytic mechanism for PaD.

Essential roles of key amino acids of loops 1-3

The PaD-**4** complex was acquired at a high resolution (Fig. 4a and Supplementary Fig. 13), indicating that the 1-OH group and the neighboring C-C bond are positioned in close proximity to the Fe center, allowing for direct cleavage targeting. When the substrate binds to PaD, the distance from the water ligand to the Fe center shifts from 3.2 Å to 2.3 Å, and the side chain of R451 rotates, causing the guanidinium group to approach the hydroxyl and methoxy groups of **4** and facilitating the formation of hydrogen bonds (Supplementary Fig. 14a). An in-depth examination of the intricate arrangement indicates that loop 1 primarily encompasses the substrate (Fig. 4a, b). F60 on loop 1 engages in direct substrate interaction via π-π stacking (Fig. 4b). Substituting F60 with A and W results in a significant loss of activity, but substituting it with the structurally similar aromatic amino acid Y retains approximately 10% of the activity (Fig. 4c and Supplementary Fig. 8). When hydrophobic amino acids such as V58 in loop 1 and L253 in loop 2 are mutated to alanine, the activity is completely lost (Fig. 4b, c and Supplementary Fig. 8). Nevertheless, substituting L253 with either F or W, both of which possess substantial hydrophobic groups, preserves most of the activity (Fig. 4b, c and Supplementary Fig. 8). In addition, the mutants of G57A and K56A on loop1 completely lost the catalytic activity, and the activity of M251A on loop2 was significantly reduced to 20% (Supplementary Fig. 8). The above results imply that these amino acids at the interface of loops 1 and 2 probably uphold the stability of the substrate channel and the upper sub-structure through hydrophobic or van der Waals interactions (Supplementary Fig. 14b). Loop 3 comprises various polar amino acid residues, such as T444 and H446. These residues form hydrogen bonds with adjacent amino acids, thereby contributing to the hydrogen bond network within the pocket. This interaction plays a crucial function in stabilizing either the substrate or Fe core, either directly or indirectly (Fig. 4b). Mutating T444 and H446 also results in a complete loss of activity (Fig. 4c and Supplementary Fig. 8).

Comparison of PaD-4 complex with the stimulated BTG13 complex structure

The simulated BTG13 complex structure (Supplementary Fig. 15) and the PaD complex structure (Fig. 4a, and Supplementary Figs. 13 and 14a) are very similar in terms of the substrate binding mode. Three hydrogen interactions of hydroquinone **4** with H61 and R451, and H61 with H193 are expected to be present in the stimulated BTG complex (Fig. 4b)[11]. Those hydrogen interactions are essential in orienting the substrate at a correct position for the targeting cleavage of C-C bond[22]. Additionally, the terminal hydroxyl groups of hydroquinone substrates are all directed towards the Fe center (Fig. 4d and Supplementary Fig. 15).

On the other hand, PaD and BTG13 have fundamentally different substrate specificity (Fig. 1c). Unlike BTG13 which has a wide-opening pocket without any barrier between the substrate and the outside environment (Fig. 4d and Supplementary Fig. 15), PaD's loops 1-3 operate as a lid, securely enclosing the substrate within the active pocket (Fig. 4d). While both enzymes facilitate dioxygenation processes on hydroquinones, it is worth noting that PaD's substrate **4** is considerably smaller in the phenol form (Figs. 1c and 4d). Previous studies demonstrated that the substrate of BTG13 is more stable in an anthrone form, with an energy difference of 3.4 kcal/mol compared to the naphthol form[28] (Fig. 4d). Accordingly, in the PaD complex, the phenolic hydroxyl group is oriented towards the Fe, whereas in the simulated BTG13-substrate complex, it is a secondary alcohol (Fig. 4d). In conclusion, the inclusion of motifs 1-3 in PaD adapts to the variations in substrate size and equilibrium form in the natural evolution.

Proposed mechanism for PaD

As shown in Fig. 4e, **4** can be tautomerized to **4a** in the reaction buffer. From **4a**, the hydrogen abstraction from $sp3$-hydridized carbon by Fe(III)−O$_2$·⁻ generates (**I**), and the following O-O bond cleavage leads to the formation of intermediate (**II**). The resultant Fe(IV) = O species abstracts a hydrogen from (**II**) to generate (**III**), followed by the formation of the epoxide intermediate (**IV**) through the oxygen radical

attack. From (**IV**), the cleavage of the C-C bond of the epoxide and a step of proton-coupled electron transfer, with the aid of the H61, generates the lactone (**VI**). (**VI**) is converted to (**VII**) by the nucleophilic attack of the hydroxyl group of Fe(II)-OH on the lactone group. Finally, the product **5** is generated by the C−O bond cleavage of (**VII**). In conclusion, the proposed mechanism for PaD here is similar to the most favorable pathway of BTG13, which features a hydrogen abstraction from the $sp3$-hybridized substrate carbon and hires a lactone intermediate.

PaD like proteins (PaDLs) featuring motifs 1-3 are promiscuous towards small hydroquinones

To probe PaD's substrate preference, we incubated PaD with diverse hydroquinones **S1**−**S13** (Supplementary Fig. 9a), which are industrial chemical materials and can cause environmental pollution. Indeed, we found the aromatic-cleavage activity of PaD with methyl-, methoxy-, and trimethyl-hydroquinones (Fig. 5a and Supplementary Fig. 9b). Another remarkable property of PaD is its ability to cleave chlorinated substrates (Fig. 5a). Interestingly, we observed a similar dehydration product as **5** in the reactions with **S6**, which shows a MW of 166 with 18 units less than the proposed dioxygenation product (Supplementary Fig. 9c). We also incubated PaD with *ortho*- and *meta*-hydroxyl substituted phenols, which all show no transformation (Supplementary Fig. 44), confirming the hydroquinone structure is essential to the PaD activity.

We built a fungal protein database containing proteins from 10,239 fungal genomes (see methods for details). Genome mining in this database identified 8701 PaD Like proteins (PaDLs) overspread 7303 fungal species, with at least 20% sequence identity to PaD. To explore the hydroquinone-cleavage potential, we constructed a sequence similarity network (SSN) of selected 2730 PaDLs (Fig. 5c). Based on the SSN, we synthesized two genes of PaDL1 and PaDL2 located in the same cluster of PaD (Fig. 5b), which show 62% and 66% sequence identities to PaD respectively (Supplementary Fig. 11). Sequence analysis indicates that PaD and PaDL1 and PaDL2 share the same motifs 1-3 with PaD (Supplementary Fig. 10), suggesting that they may also be dioxygenases towards small hydroquinones. We transformed both genes of PaDL1 and PaDL2 into *E. coli* BL21(DE3) cells and obtained their proteins under standard expression conditions. We evaluate their activity towards substrates **4** and **S1**−**S13**. Gratifyingly, both enzymes showed activity towards at least one of those substrates (Fig. 5a). Notably, PaDL2 displayed considerable activity towards **S2** and **S11** by transforming both substrates completely (Supplementary Fig. 9b). Finally, in addition to PaDL1 and PaDL2, we have obtained more PaDLs (PaDLs 3-5) from the SSN network, which have been supplemented to Supplementary Fig. 10.

The phylogenetic analysis revealed that PaD and its related proteins PaDL1 and PaDL2 form a distinct clade apart from the known Duf4243 oxygenases (Supplementary Fig. 11). After confirming that PaD had the ability to degrade **S6** (Supplementary Fig. 9), we proceeded to acquire the complex structure of PaD in conjunction with **S6**. Figure 6 demonstrates that the position where **S6** binds in the pocket is nearly identical to that of PaD's natural substrate **4** (Supplementary Fig. 16). Similar to PaD-**4**, the terminal 1-OH faces the Fe center, indicating that the C-C cleavage occurs between the C1 and C2 positions (Fig. 6). The spatial distances between **S6** and key amino acids F60, R451, H61, etc. are 4.0 Å, 3.0 Å, and 2.8 Å, respectively (Fig. 6b). These distances are equivalent to those between these amino acids and the natural substrate **4** (Fig. 6b and Supplementary Fig. 16). The spatial coordinates of other amino acids, namely L253, V58, T444, and H446, are also fundamentally identical (Fig. 6). Overall, the PaD-**S6** complex structure effectively elucidates the substrate selectivity of PaD's catalytic activity towards monocyclic hydroquinones.

PaD homolog MtaD featuring motifs 1-3 is involved in the biosynthesis of multicolosic acid

Like penicillic acid **1**, multicolosic acid (Fig. 5b) is also produced in nature via oxidative cleavage of polyketide derived aromatic intermediates[29]. By utilizing the protein sequence of PaD as a reference, we successfully discovered a putative BGC for multicolosic acid (**11**) in *P. sclerotiorum* (Fig. 5b). It encodes 7 proteins, including a cytochrome P450 oxidoreductase (MtaA), a high reducing PKS (MtaB), a thioesterase (MtaC), a flavin-dependent monooxygenase (MtaE), an NR-PKS (MtaF), a GMC-type oxidoreductase (MtaG), and a Duf4243-containing protein MtaD. MatD shows 48.7% sequence identity to PaD and shares the same motifs 1-3. The BGC and biosynthesis of multicolosic acid were previously unclear, although labeling studies already suggested a hydroquinone cleavage involved in its biosynthesis[29,30]. We hypothesized that the PaD homolog MtaD catalyzes this cleavage (Fig. 5b). To confirm our hypothesis, we purified MtaD from *E. coli*
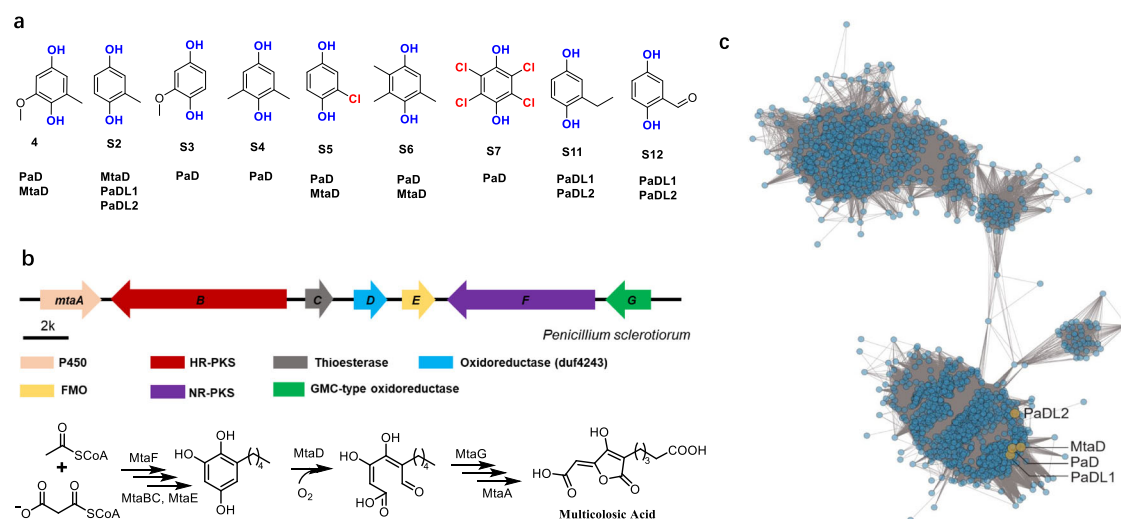

Fig. 5 | Characterizing PaD-like dioxygenases (PaDLs) as a family of hydroquinone dioxygenases. a Promiscuity profile substrate scope of PaD, MtaD, PaDL1, and PaDL2. PaDL1 and PaDL2 were selected from the same cluster of PaD. The active proteins are marked below the tested hydroquinones. **b** The BGC and proposed biosynthetic pathway of multicolosic acid. **c** Sequence similarity network (SSN) analysis of PaDLs. The SSN uses 2730 PaDLs obtained from uniport database by searching for Duf4243-domain containing proteins.
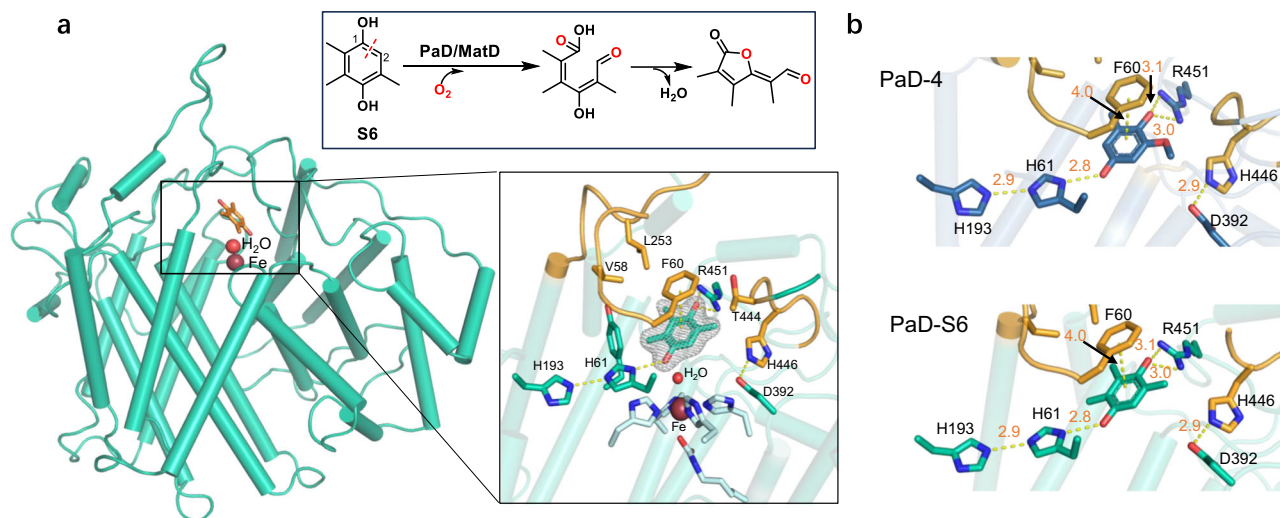
Fig. 6 | Crystal structure analysis of PaD-S6. a The overall structure and substrate-binding site view of PaD-**S6** complex (PDB: 8Z4S). Loops 1-3 are marked in gold. The cleavage-site of PaD toward **S6** and the generated product is shown at the top of the PaD-**S6** structure complex, which was proposed based on the structure and mass spectroscopy analysis. The electron density *2Fo-Fc* maps of the bound hydroquinone **S6** in PaD are shown with a gray-colored mesh, and contoured at 1 σ. **b** The comparison of the spatial distances from the bound hydroquinone to key amino acids in the PaD-**4** and PaD-**S6** structure complexes.

(Supplementary Fig. 6), which shows considerable substrate promiscuity towards different monocyclic hydroquinones (Fig. 5a and Supplementary Fig. 9b). Finally, we proposed a concise pathway for **11** based on the MtaD function (Fig. 5b and Supplementary Fig. 43).

Discussion

To the best of our knowledge, there are rare instances of protein engineering that have successfully altered substrate selectivity from a bulky tricyclic structure to a smaller monocyclic one. The substrate specificity of enzymes has a substantial impact on the composition and proportion of substrates, directly determining the resulting value of the products[31]. Although random mutagenesis and structure guided evolution have demonstrated their utility in the recent redesign of enzymes[32], attempts to make even minor modifications in enzyme substrate-specificity have proven to be challenging[33]. We now lack the ability to predict in advance the effects of structural alterations on substrate specificity. To address this issue, it is necessary to investigate the molecular and structural basis of substrate-specific processes in various enzymes.

In the current study, we reported the BGC as well the biosynthetic logic for PA in *A. westerdijkiae*. During this course, we identified a new type of Duf4243 dioxygenase PaD, catalyzing the critical aromatic cleavage towards the monocyclic hydroquinone. We systematically analyzed its differences in sequence and structure from known Duf4243 dioxygenases, characterizing a lid-like tertiary structure composed of loops 1-3 (motifs 1-3), which leads to a distinctive active pocket. Key amino acids on loops 1-3 support substrate entering, binding, and activation through various interactions. Guided by motifs 1-3, we identified more PaDLs, which either participate in the biosynthesis of natural products or possess a broad capability to degrade small hydroquinone compounds. Overall, our study demonstrates how nature has incorporated new motifs to support the drastic shifting of substrate specificity, which may represent a theoretical foundation for the direct evolution of different natural enzymes, and the promiscuous PaDLs uncovered here represent a large group of new Fe (II)-dependent HQDOs widely distributed in Nature.

Methods

General materials, Strains, Cultivation Conditions

A. westerdijkiae NRRL 3174 was obtained from Agricultural Research Service Culture Collection (NRRL). Restriction enzymes and Q5 DNA polymerase were purchased from New England Biolabs (USA). DNA sequencing and synthesis of primers were completed at Tsingke Biological Technology (Beijing, China). NMR spectra were recorded on Bruker AVANCE III 500 instrument (500 MHz for ^1H, 125 MHz for ^{13}C NMR). Cellulase and snailase were purchased from Yakult (Japan) and Solarbio (Beijing, China), respectively. HPLC grade methanol and acetonitrile were purchased from Thermo Fisher Scientific (Waltham, Massachusetts, USA). Other analytical reagents were purchased from Sinopharm Chemical Reagent Beijing Co., Ltd.

The PDA medium (20 g glucose, 4 g potato extract, 15 g agar, 1 L H_2O) was used for the sporulation and fermentation of *A. westerdijkiae*. YES medium (20 g yeast extract, 150 g sucrose, 0.5 g $MgSO_4 \cdot 7H_2O$, 20 g agar, 1 L H_2O) was used for protoplast regeneration of *A. westerdijkiae*. The strain of *E. coli* was cultured in LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl, 1 L H_2O). Competent cell of *E. coli* BL21 (DE3) was used for protein expression and purification.

Preparation of gDNA, RNA, and cDNA from *A. westerdijkiae*

To extract the genomic DNA and RNA of *A. westerdijkiae*, the strain was grown on PDA medium at 28 °C for 3 days and the mycelia were collected. Then the mycelia were frozen in liquid nitrogen and the genomic DNA was extracted following the instructions provided in the Omega Fungi DNA Extraction Kit. The RNA extraction was performed using the Omega Fungi RNA Extraction Kit following manufacturer's instructions. The cDNA was obtained from total RNA using random primers of EasyScript® One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen).

Construction of plasmids

The target genes (*paB−E*) were amplified from gDNA of *A. westerdijkiae* using the primers in Supplementary Table 1. Restriction enzymes XhoI and NdeI were used to obtain the linearized vector pET28a, and then linked to the target DNA fragments by Gibson assembly method[34]. The linked products were transformed into the competent cell of *E. coli* BL21 (DE3). The site-directed mutagenesis of PaD was constructed by Fast Mutagenesis System (TransGen Biotech) using the plasmid of pET28a-*paD* as the template. The single clones of correct transformants were used to further express and purify protein. The proteins PaD and PaE was obtained using the *S. cerevisiae* protein expression system. The plasmids pXW55 (Ura marker) and pXW06 (Trp marker)

were used for construction of the heterologous expression plasmids. Restriction enzymes SpeI/PmlI and NdeI/PmeI were performed to digest pXW55 and pXW06, respectively. The linearized vectors were assembled with the target DNA fragments by Gibson assembly method. All primers for amplifying target sequences and constructing plasmids in this study are listed in Supplementary Tables 1 and 2.

Construction of knock-out mutations in *A. westerdijkiae*

The 2k bp 5′- and 3′-flanking regions of target gene were amplified from *A. westerdijkiae* genomic DNA. The donor DNA fragments were obtained by fusing the flanking regions of target gene and hygromycin resistance gene and integrated into the plasmid of pUC57. 10 mL of enzyme-containing solution (20 mg/mL cellulase and 20 mg/mL snailase) was used to prepare protoplasts[35]. Then the plasmids were transformed into wild-type *A. westerdijkiae* by polyethylene glycol (PEG) mediated protoplast transformation, and protoplasts regeneration was done on YES medium containing 200 mg/mL hygromycin. The identification of transformants was carried out by PCR using the testing primers in Supplementary Table 1 and illustrated in Supplementary Fig. 3.

Protein expression and purification

The protein expression plasmids pET28a-*paB*, pET28a-*paC*, pET28a-*paD*, and pET28a-*paE* were respectively transformed into competent cells *E. coli* BL21 (DE3). The corresponding transformant was cultured in a 10 mL LB liquid medium containing 35 μg/mL kanamycin (kana) at 37 °C, shaking at 220 rpm overnight. The seed culture was then transferred into 1 L LB medium (35 μg/mL kana) and cultured at 37 °C and 220 rpm to an OD_{600} of 0.4. Isopropyl-β-D-thiogalactoside (IPTG, 0.1 mM) was added to induce protein expression at 16 °C for 16 h. Protein purification was performed using Nickel-NTA affinity chromatography (Solarbio, Beijing, China) at 4 °C with standard protocols[36]. The cells were harvested by centrifugation at $2600 \times g$ for 10 min at 4 °C. The precipitate was resuspended in 30 mL cold binding buffer (20 mM Tris-HCl, pH 8.0, containing 0.3 M NaCl and 20 mM imidazole), and disrupted using an ultrasonic processor (SX-605D, Henglong instrument Co. Ltd., Changzhou, China). The cell lysate was centrifuged at $15,000 \times g$ for 1 h at 4 °C to remove cell debris. The supernatant was mixed with Ni-NTA agarose resin, followed by eluting with washing buffer (20 mM Tris−HCl, pH 8.0, containing 0.3 M NaCl, 40 mM imidazole, and 10%glycerin) using a gravity flow column. The bonded protein was eluted with eluting buffer containing 0.25 M imidazole and concentrated using an Amicon Ultra (Merck Millipore) at $2600 \times g$ for 30 min at 4 °C. The protein concentration was determined by Bradford analysis using bovine serum albumin (BSA) as a standard. The purified protein was stored at −20 °C.

Determination of Fe (II) in PaD

The iron element in PaD was determined using inductively coupled plasma-optical emission spectroscopy (ICP-OES, Optima 5300DV). The purified PaD protein was replaced with ultrapure water for 3 times and used to determine the Fe content, and the samples without PaD were treated similarly as a control.

In vitro assays of PaB, PaC, PaD, and PaDLs1-2

The PaB-catalyzed reaction was carried out in a 50 μL reaction mixture containing 50 μM PaB, 2 mM **2**, 50 μM FAD, 4 mM NADPH, and 50 mM PBS (pH 7.0). The PaBC-catalyzed reaction was carried out in a 50 μL reaction mixture containing 50 μM PaB, 50 μM PaC, 2 mM **2**, 50 μM FAD, 2 mM SAM, 4 mM NADPH, and 50 mM PBS (pH 7.0). And the PaD and PaDLs1-2-catalyzed reactions were carried out in a 50 μL reaction mixture containing 50 μM PaD, 2 mM substrate, and 50 mM PBS (pH 7.0). All reactions were carried out at 25 °C and quenched with methanol. protein precipitate was removed by centrifugation at $15000 \times g$ for 15 min and then the supernatant was analyzed by LC-MS.

Negative controls were performed with boiled proteins (incubated in boiling water for 15 minutes) as other components remained as above.

LC-MS analysis was performed on an AGILENT-1200HPLC/6520QTOFMS (USA) system using a C18 analytical column (Gemini 150 × 2.0 mm, particle size 3 μm; Phenomenex). The gradient was as follow: 0 min, 5% A; 14 min, 100% A; 16 min, 100% A; 16.5 min, 10% A; 20 min, 10% A. Flow rate: 0.3 mL/min.

Chemical complementation and feeding studies

For chemical complementation of *ΔpaA* strain of *A. westerdijkiae* with compounds (solubilized in DMSO), spores of *ΔpaA* were inoculated in PDA plate together with 0.2 mg compounds **2**, **4**, **8**, **9** and **10**, and further cultured for 3 days at 30 °C. The mycelia and medium were extracted for LC-MS analysis.

To test the function of *paE* gene in vivo, *paD* and *paE* were expressed in yeast, and compound **4** was fed to yeast for in vivo transformation. Yeast harboring plasmids pXW55-*paE* and pXW06-*paD* were initially cultured in 5 mL liquid drop-out medium overnight, and 1 mL of yeast cells was transferred to 10 mL of liquid YPD medium for an additional 16 h culture. Then the cells were harvested and resuspended by 1 mL fresh YPD medium with 0.2 mg compound **3** added. After 24 h, the culture was extracted for LC-MS analysis.

Purification of compounds 1, 2, 4, 7, 8, 9, and 10

Compounds **1**, **4**, **8**, and **10** were isolated from wide-type strains and mutants. The fungus was cultured in PDB medium, shaking at 250 rpm and 37 °C for 3 days. The culture was extracted three times with ethyl acetate (EtOAc) to obtain the crude extract. The crude extract was separated on Sephadex-LH20 eluting with MeOH: CH_2Cl_2 (1:1) to get a sub-fraction. Then semi-prepared HPLC was performed to yield **2** (white powder, 37 mg from 2 L culture), **4** (white powder, 53 mg from 2 L culture), **1** (white powder, 7 mg from 0.5 L culture), **8** (red powder, 44 mg from 2 L culture), and **10** (white powder, 15 mg from 2 L culture). To isolate compound **7**, a 40 ml reaction mixture of PaB and **2** was carried out in PBS buffer solution (pH 7.0) for 1 h, and the reaction was quenched with methanol. After the solution was concentrated by rotary evaporation, it was further purified by semi-preparative HPLC to yield product **7** (red powder, 5.4 mg). Similarly, compound **9** (white powder, 2 mg) was isolated from a 4 ml reaction solution of PaD and **4**. The structures of those obtained compounds were elucidated by comprehensive NMR spectra and MS assignments, which are presented in the supplementary information as shown in Supplementary Figs. S17−42).

High-performance liquid chromatography (HPLC) analysis was performed on a Waters 2695 (USA) system coupled with a photodiode array detector and a C_{18} analytical column (Sepax GP-C18, 250 × 4.6 mm, particle size 5 μm; Sepax Technologies). The mobile phase consisted of acetonitrile (CH_3CN) and water (H_2O) containing 0.1% formic acid respectively. The gradient was as follow: 0 min, 10% A; 25 min, 100% A; 30 min, 100% A; 35 min, 10% A; 40 min, 10% A. Flow rate: 1 mL/min.

Network analysis of PaDL-related genes

A total of 2730 PaDLs were identified by blasting the querying PaD protein sequence against NR (Non-Redundant Protein Sequence Database) database. The sequence similarity network was constructed using an online enzyme similarity tool (https://efi.igb.illinois.edu/efi-est/). The parameters with Alignment Score Threshold = 60; Minimum = 0; Maximum = 1100 were used for SSN construction. The SSN network was exhibited by Cytoscape.

Crystallization and structure determination of PaD

Crystals of PaD were grown by using the hanging-drop vapor-diffusion method. Drops composed of a 1:1 ratio of protein (15 mg/mL in 20 mM Tris pH 7.0, 150 mM NaCl) to reservoir solution were equilibrated against reservoir solution (0.1 M Tris pH 8.0, 25% PEG 1000, 0.1 M

Sodium Malonate pH 8.0). Large rhombus crystals (~100 μm) grew within one week. To cryoprotect the crystals, the mother liquor was replaced by a solution containing 90% v/v reservoir solution (0.1 M Tris pH 8.0, 25% PEG 1000, 0.1 M Sodium Malonate pH 8.0) and 10% glycerol. Crystals treated in this way belong to space group P62 with unit cell dimensions a = b = 120.209 Å, c = 93.305 Å, and have a single molecule per crystallographic asymmetric unit. The X-ray diffraction datasets were collected at beamline TPS 05 A of the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan) and processed by using HKL2000[37]. The crystal structure of PaD was solved with a method of molecular replacement (MR) by using the Alpha Fold predicted PaD structure as the search model. Subsequent model adjustment and refinement were conducted by using Refmac5[38,39]. Five percent of randomly selected reflections were set aside for calculating R_{free} as a monitor of model quality. The complex structure with substrate and **S6** were determined by MR method with Phaser using the refined PaD structure as a search model. All figures of the protein structures were prepared by using the PyMOL program (http://pymol.sourceforge.net/). Data collection and refinement statistics are summarized in Supplementary Table 6.

Structure-complex simulation of BTG-substrate

To get the binding model of the BTG13-substrate complex, we performed molecular docking on the basis of the reported crystal structure of BTG13 (PDB ID: 7Y3W)[11]. Since the crystal structure of BTG13 was a homodimer, the molecular docking the structure of chain A was performed by using the CDOCKER procedure of Discovery Studio Client v19.1.0.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The authors declare that all the data supporting the findings of this study are available within the main text, and Supplementary Information documents. The X-ray structural data of proteins PaD, PaD-**4**, and PaD-**S6** in this study have been deposited in the Protein Data Bank (www.rcsb.org) with accession codes 8Z4Q, 8Z4R, and 8Z4S, respectively. Source data are provided with this paper, and additional data supporting the findings of this study are available from the corresponding author on request. Source data are provided with this paper.

References

1. Jakoobi, M. & Sergeev, A. G. Transition-metal-mediated cleavage of C-C bonds in aromatic rings. *Chem.-Asian J.* **14**, 2181–2192 (2019).
2. Qiu, X. et al. Cleaving arene rings for acyclic alkenylnitrile synthesis. *Nature* **597**, 64–69 (2021).
3. Harayama, S., Kok, M. & Neidle, E. L. Functional and evolutionary relationships among diverse oxygenases. *Annu. Rev. Microbiol.* **46**, 565–601 (1992).
4. Tang, H. et al. Systematic unraveling of the unsolved pathway of nicotine degradation in *Pseudomonas. PLoS Genet* **9**, e1003923 (2013).
5. Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds—from one strategy to four. *Nat. Rev. Microbiol.* **9**, 803–816, (2011).
6. Mariëlle J. H. et al. Hydroquinone dioxygenase from *Pseudomonas Fluorescens* Acb: A novel member of the family of Nonheme-Iron(Ii)-Dependent Dioxygenases. *J. Bacteriol.* **190**, (2008).
7. Ronnie, J. M. et al. Discovery of novel P-Hydroxybenzoate-M-Hydroxylase, Protocatechuate 3,4 Ring-Cleavage Dioxygenase, and Hydroxyquinol 1,2 Ring-Cleavage Dioxygenase from the filamentous fungus *Aspergillus niger. ACS Sustain. Chem. Eng.* **7**, 19081–19089 (2019).
8. Semana, P. & Powlowski, J. Four aromatic intradiol ring cleavage dioxygenases from *Aspergillus niger. Appl. Environ. Microbiol.* **85**, e01786–01719 (2019).
9. Bugg, T. D. H. & Winfield, C. J. Enzymatic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Nat. Prod. Rep.* **15**, 513–530 (1988).
10. Tang, M. C. et al. Oxidative cyclization in natural product biosynthesis. *Chem. Rev.* **117**, 5226–5333 (2017).
11. Hou, X. et al. Discovery of the biosynthetic pathway of Beticolin 1 reveals a novel non-heme iron-dependent oxygenase for anthraquinone ring cleavage. *Angew. Chem. Int. Ed.* **61**, e202208772 (2022).
12. Qi, F. et al. Bienzyme-catalytic and dioxygenation-mediated anthraquinone ring opening. *J. Am. Chem. Soc.* **143**, 16326–16331 (2021).
13. Vieweg, L. et al. Recent advances in the field of bioactive tetronates. *Nat. Prod. Rep.* **31**, 1554–1584 (2014).
14. Frisvad, J. C. A critical review of producers of small lactone Mycotoxins: Patulin, Penicillic Acid and Moniliformin. *World Mycotoxin J.* **11**, 73–100 (2018).
15. Alsberg, C. & Black, O. F. Contributions to the study of maize deterioration: biochemical and toxicological investigations of *Penicillium Puberulum* and *Penicillium Stoloniferum*. (US Government Printing Office, 1913).
16. Chan, P. K. & Hayes, A. W. The hepatobiliary toxicity of penicillic acid, a carcinogenic mycotoxin. *Dev. Toxicol. Environ. Sci.* **8**, 507–516, (1980).
17. Nguyen, H. T. et al. Antibacterial activities of penicillic acid isolated from *Aspergillus Persii* against various plant pathogenic bacteria. *Lett. Appl. Microbiol.* **62**, 488–493 (2016).
18. Vieira, G. et al. Isolation and agricultural potential of penicillic acid against citrus Canker. *J. Appl. Microbiol.* **132**, 3081–3088 (2022).
19. Ciegler, A. & Kurtzman, C. P. Penicillic acid production by blue-eye fungi on various agricultural commodities. *Appl. microbiol.* **20**, 761–764 (1970).
20. Sekiguchi, J., Katayama, S. & Yamada, Y. 6-Methyl-1,2,4-Benzenetriol, a new intermediate in penicillic acid biosynthesis in *Penicillium Cyclopium. Appl. l Environ. Microbiol.* **53**, 1531–1535 (1987).
21. Seto, H., Cary, L. W. & Tanabe, M. Letter: Utilization of ^{13}C-^{13}C coupling in structural and biosynthetic studies. Iv. Penicillic Acid. *J. Antibiot. (Tokyo)* **27**, 558–559 (1974).
22. He, J. et al. Cytotoxic and other metabolites of inhabiting the rhizosphere of Sonoran desert plants. *J. Nat. Prod.* **67**, 1985–1991 (2004).
23. Axberg, K. & Catenbeck, S. Intermediates in the penicillic acid biosynthesis in *Penicillium Cyclopium. Acta Chem. Scand. B.* **29**, 749–751, (1975).
24. Axberg, K. & Gatenbeck, S. The enzymic formation of penicillic acid. *FEBS Lett.* **54**, 18–20 (1975).
25. Feng, P., Shang, Y., Cen, K. & Wang, C. Fungal biosynthesis of the bibenzoquinone oosporein to evade insect immunity. *Proc. Natl Acad. Sci. USA* **112**, 11365–11370 (2015).
26. Chooi, Y. H. & Tang, Y. Navigating the fungal polyketide chemical space: from genes to molecules. *J. Org. Chem.* **77**, 9933–9953 (2012).
27. Bentley, R. & Keil, J. G. Tetronic acid biosynthesis in Molds. Ii. Formation of Penicillic acid in *Penicillium Cyclopium. J. Biol. Chem.* **237**, 867–873, (1962).
28. Deng, Z. et al. Mechanistic investigation on C–C bond cleavage of anthraquinone catalyzed by an atypical nonheme iron-dependent dioxygenase BTG13. *ACS Catal.* 797–811, (2024).
29. Gudgeon, J. A., Holker, J. S. & Simpson, T. J. Use of singly and doubly labelled ^{13}C-Acetate in the Elucidation of the structures and biosynthesis of multicolic and multicolosic acids, new Tetronic Acids from *Penicillium Multicolor. J. Am. Chem. Soc. Chem. Commun.* **16**, 636–638 (1974).

30. Holker, J. S., Kaneda, M., Ramer, S. E. & Vederas, J. C. Biosynthesis of multicolosic acid, a polyketide metabolite from *Penicillium Multicolor*: occurrence of large ^{18}O-induced B-isotope shifts in ^{13}C NMR spectra. *J. Am. Chem. Soc. Chem. Commun.* 1099–1100, (1987).

31. Robinson, P. K. Enzymes: Principles and biotechnological applications. *Essays Biochem* **59**, 1–41 (2015).

32. Arnold, F. H. Directed evolution: bringing new chemistry to life. *Angew. Chem. Int. Ed. Engl.* **57**, 4143–4148 (2018).

33. Oue, S., Okamoto, A., Yano, T. & Kagamiyama, H. Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. *J. Biol. Chem.* **274**, 2344–2349, (1999).

34. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

35. Wang, L., Wang, Y., Wang, Q. & Liu, Y. Construction of genetic transformation system of *Aspergillus Ochraceus* Using Polyethylene glycol mediated method. *J. Agric. Biotechnol.* **24**, 928–936 (2016).

36. Cheng, J. et al. High-level extracellular production of A-Cyclodextrin Glycosyltransferase with recombinant *Escherichia Coli* BL21 (De3). *J. Agric. Food Chem.* **59**, 3797–3802 (2011).

37. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method. Enzymol.* **276**, 307–326 (1997).

38. Murshudov, G. N. et al. Refmac5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367 (2011).

39. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).

Acknowledgements

Author contributions

C.C., L.Y., Z.L., X.S., and S.G. conceived the idea, carried out investigation and data analysis, and co-wrote the manuscript. W.Z. and Y.G. performed the bioinformation analysis. Y.W., T.W., C.C.C., and R. G. conducted the crystallization and structure determination. All authors analyzed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52101-2.

Correspondence and requests for materials should be addressed to Rey-Ting Guo or Shu-Shan Gao.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.