Characterization of a NRPS-like Protein from Pestalotiopsis fici for Aldehyde Generation

Yuanyuan Li 1,2,†, Peng-Lin Wei 1,2,†, Huomiao Ran 1, Jie Fan 1,* and Wen-Bing Yin 1,2,*

1 State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
2 Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: fanjie@im.ac.cn (J.F.); yinwb@im.ac.cn (W.-B.Y.); Tel.: +86-010-64806170 (W.-B.Y.)
† These authors contributed equally to this work.

Abstract: Nonribosomal peptide synthetase (NRPS)-like enzymes containing A-T-R domain architecture are also known as carboxylate reductases (CARs) for aldehyde generation. To identify new members of CARs, we established a virtual library containing 84 fungal CARs distributed in seven distinct clades by genome mining and phylogenetic analysis. Nine CARs, including PnlA from Pestalotiopsis fici and eight known CARs, were clustered in clade VI and proposed to catalyze the reduction of nonreducing polyketide synthase (NR-PKS)-derived aryl carboxylic acids. The recombinant protein PnlA was overproduced and purified to apparent homogeneity from Saccharomyces cerevisiae. In vitro enzyme assays of PnlA with 28 different benzoic acid derivatives (1–28) revealed the corresponding aldehyde formation in 14 cases (1–14). Comparison of conversion yields indicated the high preference of PnlA toward 3,5-dimethylorsellinic acid (DMOA, 4) and vanillic acid (10). A specificity-conferring code Q355 in PnlA was postulated by sequence alignment with the known CARs in clade VI. Our study provides an updated virtual library of fungal CAR enzymes and expands the biocatalytic selectivity of CARs.

Keywords: NRPS-like enzyme; fungal carboxylate reductase (CAR); aldehyde generation; benzoic acid derivatives; Pestalotiopsis fici

1. Introduction

Nonribosomal peptide synthetases (NRPSs) are multidomain enzymes involved in the biosynthesis of structurally and functionally diverse nonribosomal peptides (NRPs) [1]. Examples of NRPs used clinically include vancomycin, daptomycin, penicillin, cephalosporins and ergotamine [2,3]. In general, the minimal catalytic module for peptide extension comprises adenylation (A), thiolation (T), also termed peptidyl carrier protein (PCP), as well as condensation (C) domains. However, the peptide-forming C domain can be replaced by a thioesterase (TE) or reductase (R) domain, constructing two subfamilies of NRPS-like enzymes [4,5]. Recent advances in genome sequencing have indicated the widespread presence of NRPS-like genes in microorganisms, whereas only a small number were characterized for secondary metabolism. Represented by ApvA, BtyA and AtqA identified from Aspergillus terreus, NRPS-like enzymes with an A-T-TE domain architecture catalyze the dimerization of α-keto acids, leading to the formation of lactones, ketals or benzoquinones [6,7]. Differing from the TE domain, the R domain in NRPS-like enzymes functions for carboxylic acid reduction to afford aldehyde with the help of upstream substrate adenylation by the A domain at the expense of ATP, thioester linkage formation by T domain and NADPH as a hydride donor (Figure 1A) [8]. Hence, NRPS-like enzymes containing A-T-R domain architecture are also known as carboxylate reductases (CARs).

On the basis of characterization of the first CAR, NcCAR, from the fungus Neurospora crassa, CARs triggered research interest owing to the remarkable roles of...
aldehyde-containing compounds as biologically active agents and key synthetic intermediates [9–11]. However, only approximately 50 CARs were identified from microorganisms and produced recombinantly in *Escherichia coli* or *Saccharomyces cerevisiae* for biocatalytic investigations. CARs have shown considerable potential for product diversification, as A domains of CARs are differentiated to accept a broad range of substrates, including aryl carboxylic acids, aliphatic carboxylic acids, aryl-aliphatic carboxylic acids and amino acids [12]. Therefore, CARs have been connected to various types of natural products (NPs), such as the meroterpenoid LL-Z1272β [13], spiroketal epicospirocin A [14], as well as piperazine derivatives actinopolymorphol C and brasiliamide A [15,16] (Figure 1B). Exemplified by the well-studied AtCAR (ATEG_03630) from *A. terreus*, TvCAR from *Trametes versicolor* and NcCAR, CARs exhibited relaxed catalytic activities toward a variety of carboxylic acids [9,17,18]. This renders CARs ideal biocatalysts for aldehyde synthesis. However, the selective specificities of CARs have remained unclear until now.

![Diagram](attachment:carboxylic_acid_reduction_by_CARs.png)

**Figure 1.** Carboxylic acid reduction by CARs. (A) CARs containing A-T-R domain architecture catalyze two-electron reduction of carboxylates to aldehydes. (B) Representative fungal NPs involving carboxylic acid reduction by CARs in their biosynthetic pathways. They include cichorine from *Aspergillus nidulans* [19], LL-Z1272b from *Stachybotrys bisbyi* [13], epicospirocin A from *Epicoccum nigrum* [14], melanin derivative from *Ustilago maydis* [20], ascofuraneone from *Acremonium egyptiacum* [21], aspergillic acid from *Aspergillus flavus* [22], actinopolymorphol C from *A. flavus* [15] and brasiliamide A from *Penicillium brasiliameum* [16].

The biocatalytic potential of CARs encouraged us to identify additional new members for aldehyde generation. We conducted genome mining of CARs in fungi and established a fungal virtual CAR library containing 84 fungal CARs with 55 new CARs. These CARs were clustered into seven distinct clades by phylogenetic analysis. A new CAR, termed PnlA, was identified from *Pestalotiopsis fici* by phylogenetic analysis. PnlA showed relaxed catalytic activities toward 14 aryl carboxylic acid derivatives but with varying levels of preferences (Scheme 1). Combined with multiple sequence alignments, we proposed a key specificity-conferring code of PnlA.
2. Materials and Methods

2.1. Strain, Media and Culture Conditions

The strains used in this study are summarized in Table S1. *P. fici* CGMCC3.15140 was grown on potato dextrose agar (PDA, BD Difco™, 24 g/L) at 25 °C [23,24]. *Escherichia coli* strain DH5α was cultured in an LB medium (w/v, 1% NaCl, 1% tryptone and 0.5% yeast extract) with appropriate antibiotics for standard DNA manipulation [25]. *Saccharomyces cerevisiae* strain BJ5464-NpgA (MATα ura3-52 his3-D200 leu2-D1 trp1 pep4::HIS3 prb1 D1.6R can1 GAL) was incubated on YPD medium (w/v, 2% peptone, 1% yeast extract and 2% dextrose) for routine growth and on SD-drop medium (w/v, 0.5% Bacto™ Casamino Acids, 0.17% yeast nitrogen base, 0.5% ammonium sulfate and 2% dextrose) with appropriate nutrition as required for gene heterologous expression at 30 °C [26,27]. The substrates used in this study were purchased from Beijing Hua Wei Rui Ke Chemical Co., Ltd. (Beijing, China), Beijing Wokai Biotechnology Co., Ltd. (Beijing, China), Beijing Bailingwei Technology Co., Ltd. (Beijing, China) and Shanghai Yuanye Biotechnology Co., Ltd. (Shanghai, China).

2.2. Computer-Assisted Sequence Analysis

Using the known NcCAR and AtCAR as templates, we searched the CAR protein sequences from the NCBI database (http://www.ncbi.nlm.nih.gov/protein, accessed on 15 August 2022) and presented 84 CARs from 28 fungal species. All amino acid sequences of CARs were aligned by ClustalW using MEGA 7.0 for cluster analysis. The phylogenetic

---

**Scheme 1.** Mechanism of the PnIA-catalyzed reduction of aryl carboxylic acid derivatives. Fourteen aryl carboxylic acid derivatives (1–14) were converted to the corresponding aldehydes (1′–14′).
tree was constructed based on maximum likelihood (ML) analysis using RAxML 8.2.10 and visualized in iTOL v6 (https://itol.embl.de/, accessed on 20 August 2022). Multiple sequence alignments were performed with the program ClustalW and visualized with ESPript 3.2 (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi, accessed on 20 August 2022) to identify strictly conserved amino acid residues.

2.3. RNA Isolation and cDNA Synthesis

For isolation of RNA from *P. fici*, the fungus was cultivated on PDA at 25 °C for 7 days. Total RNA from the mycelia of *P. fici* was extracted by the TranZol™ kit (Transgen Biotech, Beijing, China) [28]. RNA integrity was confirmed by electrophoresis on the Tris-boric acid-EDTA (TBE) agarose gel and a Quawell Q3000 nucleotide analyzer (Quawell). The single-strand cDNA was synthesized by a FastKing RT kit (Tiangen Biotech, Beijing, China) according to the manufacturer’s protocol.

2.4. Plasmid Construction

The primers and plasmids are given in Table S1. For PnlA protein production under ADH2 promoter (ADH2p) in *S. cerevisiae*, TransStart® FastPfu DNA polymerase (Transgene Biotech, Telangana, India) was used for amplification. The PCR mixture comprised 10 µL 5 × TransStart® FastPfu buffer, 0.2 mM dNTPs, 0.2 µM forward/reverse primers PFICI364-XW55 F/R, 1 µL *P. fici* cDNA, 2.5 units TransStart® FastPfu DNA polymerase and ddH2O in a final volume of 50 µL. PCR was performed on a T100™ thermal cycler (Bio-Rad, Hercules, CA, USA) as follows: 95 °C for 2 min; 35 cycles of 95 °C for 20 s, 57 °C for 20 s and 72 °C for 2 min; followed by 72 °C for 5 min. The amplified fragment was purified with a DNA clean and concentrator kit (Zymo Research, Irvine, CA, USA) and cloned into the vector pXW55 with Q5 ultra-fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA) using a quick-change method [29,30], generating pYLYY7 (Figure S1). The plasmid pYLYY7 was confirmed by enzyme validation and sequencing (Sangon Biotech, Shanghai, China).

2.5. Protein Overproduction and Purification, as Well as Enzyme Assays

The plasmid pYLYY7 was transformed into *S. cerevisiae* BJ5464-NpgA for PnlA protein overproduction. Yeast-competent cell preparation and transformation were performed with an S.c. Easy comp™ transformation kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s protocol. *S. cerevisiae* transformants were grown in 10 mL of SC-Ura dropout media for 2 days and inoculated to 3.2 L of YPD medium. Cells were grown at 30 °C and 220 rpm for 3 days. The cells were harvested by centrifugation (4000 rpm, 8 min, 4 °C), resuspended in 20 mL lysis buffer (50 mM NaH2PO4, 150 mM NaCl and 10 mM imidazole; pH 8.0) and lysed by SPEX freezer mill (SPEX SamplePrep, Metuchen, USA). The recombinant His6-tagged PnlA was purified by Ni-NTA affinity chromatography (Qiagen, Hilden, Germany) according to the published procedures [31]. The purified recombinant protein was concentrated into a storage buffer (50 mM Tris–HCl, 100 mM NaCl, pH 7.9) containing 10% glycerol and analyzed on SDS-PAGE.

To determine the activity of PnlA, the enzyme assays (100 µL) containing 50 mM Tris-HCl (pH 8.5), 10 mM MgCl2, 1 mM substrate, 10 mM ATP, 2 mM NADPH, 0.7–6% (v/v) glycerol, up to 5% (v/v) DMSO and 35 µg of purified PnlA were carried out at 30 °C for 12 h. The reaction mixtures were terminated by extraction with 200 µL ethyl acetate twice and evaporated under reduced pressure to afford crude extracts. Before injecting to liquid chromatography–mass spectrometry (LC-MS), the crude extracts were dissolved in methanol and centrifuged at 13,000 rpm for 15 min. Conversion yields of the enzyme reactions were calculated from peak areas of aldehyde products and aryl carboxylic acid substrates as analyzed on LC-MS.
2.6. LC-MS Analysis

LC-MS analysis was performed on an Agilent HPLC 1200 series system equipped with a single-quadrupole mass-selective detector and an Agilent 1100LC MSD model G1946D mass spectrometer using a Venusil XBP C18 column (3.0 by 50 mm, 3 μm, Bonna-Agela Technologies, Tianjin, China). Water (A) with 0.1% (v/v) formic acid and acetonitrile (B) were used as the solvents at a flow rate of 0.5 mL/min. The substances were eluted with a linear gradient from 2% to 30% (v/v) solvent B in 30 min, then washed with 100% (v/v) solvent B for 5 min and equilibrated with 2% (v/v) solvent B for 5 min. The mass spectrometer was set in electrospray positive-ion mode for ionization.

3. Results

3.1. Construction of a Fungal Virtual CAR Library

To identify new members, we searched for fungal CARs using NcCAR and At-CAR as templates through the BLASTP program [9,17]. A total of 84 proteins from 28 fungal species were revealed as CARs with the typical A-T-R domain architecture (Figure 2). 29 of them have been previously reported for carboxylic acid reduction by genome mining and in vitro enzyme assays [4,8,32]. Ten CARs containing a well-known AtCAR and nine other CARs were mined in A. terreus, which is consistent with a previous study [4]. The genome of A. flavus harbors eight CAR-encoding genes with three known CARs, LnaA, LnbA and AsaC, accepting L-Tyr, Leu or Ile as substrates [15,22]. Four to six CARs were first found in the genomes of four Penicillium species, including P. camemberti, P. chrysogenum, P. expansum and P. roqueforti, whereas none of them was connected to the secondary metabolite (SM) biosynthesis. With respect to an endophytic fungus, P. fici, an A-T-R domain structure was observed in thirteen enzymes, including PFICI_01589, PFICI_00364, PFICI_05312, PFICI_05398, PFICI_06834, PFICI_08544, PFICI_08738, PFICI_10568, PFICI_11160, PFICI_11332, PFICI_11756, PFICI_12508 and PFICI_12763. However, neither biosynthetic nor biochemical investigation was reported for these thirteen CARs [23,28].

Hence, we carried out a sequence comparison among these 84 CARs by phylogenetic analysis, leading to the construction of a CAR similarity network (Figure 2). Seven distinct clades were observed, with three relatively large clades containing 17 (clade I), 22 (clade III) and 20 sequences (clade VII), respectively. On the basis of the strict specificities of 03009 and 06235 for Ala/Val/Ser [33], AsaC for Leu/Ile [22], LnaA, LnbA and HqlA for L-Tyr [15,34] and BrsA for L-Phe [16], we proposed that the CARs clustered in clade I were involved in amino acid reduction. Represented by Nps3 and Lys2, CARs in clade II are L-aminoadipate reductases [35]. In clade III, only NcCAR and TtCAR were reported with relaxed substrate tolerances for reducing ary1-aliphatic carboxylic acids, aliphatic carboxylic acids and a few heterocyclic aromatic carboxylic acids [8]. Pub8 from Fusarium fujikuroi catalyzes the reduction of trans-2-hexenoic acid in the biosynthesis of fusaric acid [36]. Differing from other clades, CARs in clade V are mainly identified from basidiomycetes [18,33]. In addition, no CAR was known in clades IV and VII until now.

A newly identified PFICI_00364, hereafter termed PnlA, was found in clade VI, together with eight other known CARs, including StbB, AscB, CicB, Esp4, AtCAR, PMAA_062890, ATEG_07380 and Pks5 [13,14,17,19–21,37]. Interestingly, the eight CARs shared similarities on their catalytic preferences for the ary1 carboxylic acid derivatives, such as orsellinic acid (OA, 1), 5-methylorsellinic acid (5-MOA, 2) and 3-methylorsellinic acid (3). Pks5 from Ustilago maydis reduced 1 to produce orsellinaldehyde (1′), undergoing melanin biosynthesis [20]. StbB and AscB showed activity toward grifolic acid, a meroterpenoid derived from 1 [13,21]. AtCAR, CicB and Esp4 catalyzed the carboxylic acid reduction of 2, but only AtCAR was characterized for its broad substrate scope [14,17,19,37].
Figure 2. Phylogenetic analysis of 84 CAR protein sequences from fungi. Seven distinct clades are shown in different colors. CARs in clade I, II and VI are mainly responsible for the reduction of amino acids, L-aminoadipate and orsellinic acid derivatives, respectively. CmlP characterized in *Streptomyces venezuelae* was used as the outgroup to root the tree [38]. Multiple alignments were performed by ClustalW with MEGA 7.0. The phylogenetic tree was constructed based on maximum likelihood (ML) analysis using RAxML 8.2.10 and visualized in iTOL v6. Amino acid sequences were obtained from the GenBank database.

3.2. Overproduction and In Vitro Characterization of PnlA

Encouraged by the reported biocatalytic activities of CARs in clade VI, we proceeded to identify PnlA by in vitro biochemical investigation. PnlA comprising three exons of 3261 bp was amplified from cDNA of *P. fici*, cloned into pXW55 under ADH2p and overexpressed in *S. cerevisiae* BJ5464-NpgA cells (Figure S1) [26]. Purification with the aid of Ni-NTA agarose resin resulted in 2.7 mg of recombinant PnlA per liter of standard culture. Consistent with its expected mass of 122.1 kDa, the C-terminal His6-tagged protein was confirmed on SDS-PAGE (Figure S1). To determine the activity of PnlA, we first incubated it with one of the most well-accepted substrates by clade VI members, 1, in the presence of ATP,
NADPH and MgCl\(_2\). LC-MS analysis of the incubation mixture showed a new peak at the retention time of 12.5 min in comparison with the negative control containing denatured enzyme (Figure 3). The product peak was identified as the corresponding aldehyde (1') due to its [M + H]\(^+\) ions at m/z 153.1, which was 16 Da less than the carboxylic acid-containing 1 (Scheme 1 and Table 1). This result confirmed the newly characterized PnlA as a CAR for aryl carboxylic acid reduction.

**Figure 3.** LC-MS analyses of the incubation mixtures of PnlA with different aryl carboxylic acid derivatives (1–14). 1'–14' are the corresponding aldehyde products. UV absorptions at 280 nm are illustrated for the enzyme assays of PnlA with 1–13, and those at 340 nm are illustrated for the enzyme assay of PnlA with 14. CK is the reaction mixture with denatured PnlA. Additional data on the enzyme assays of PnlA with 15–28 are presented in Figure S2.
Table 1. Specific activities of PnlA toward different benzoic acid derivatives.

| Substrate                                      | [M + H]^+ of PnlA Product | Conversion |
|------------------------------------------------|---------------------------|------------|
| Calculated                                     | Measured                  |            |
| orsellinic acid (1)                            | 153.1                     | 153.1      | 22.6%     |
| 5-methylorsellinic acid (5-MOA, 2)             | 167.1                     | 167.0      | 19.2%     |
| 3-methylorsellinic acid (3-MOA, 3)             | 167.1                     | 166.9      | 58.8%     |
| 3,5-dimethylorsellinic acid (DMOA, 4)          | 181.1                     | 181.0      | 100.0%    |
| 2,4-dihydroxybenzoic acid (5)                  | 139.0                     | 139.3      | 16.3%     |
| 4-hydroxybenzoic acid (6)                      | 123.0                     | 122.9      | 45.0%     |
| 4-methoxybenzoic acid (7)                      | 137.1                     | 137.0      | 40.0%     |
| 4-butoxybenzoic acid (8)                       | 179.1                     | 179.0      | 4.1%      |
| 3,4-dihydroxybenzoic acid (9)                  | 139.0                     | 139.0      | 53.9%     |
| vanillic acid (10)                             | 153.1                     | 152.9      | 95.8%     |
| isovanillic acid (11)                          | 153.1                     | 152.9      | 7.7%      |
| 3,4-diaminobenzoic acid (12)                   | 137.1                     | 136.8      | t.a.      |
| 4-amino-3-nitrobenzoic acid (13)               | 167.0                     | 166.8      | 6.3%      |
| 4-amino-2-chlorobenzoic acid (14)              | 156.0                     | 156.0      | 22.6%     |
| benzoic acid (15)                              | 107.0                     | n.d.       | 0         |
| salicylic acid (16)                            | 123.0                     | n.d.       | 0         |
| 3-hydroxybenzoic acid (17)                     | 123.0                     | n.d.       | 0         |
| 2-nitrobenzoic acid (18)                       | 152.0                     | n.d.       | 0         |
| 3-nitrobenzoic acid (19)                       | 152.0                     | n.d.       | 0         |
| 4-nitrobenzoic acid (20)                       | 152.0                     | n.d.       | 0         |
| phthalic acid (21)                             | 151.0 or 135.0            | n.d.       | 0         |
| cinnamic acid (22)                             | 133.1                     | n.d.       | 0         |
| coumaric acid (23)                             | 149.1                     | n.d.       | 0         |
| caffeic acid (24)                              | 165.1                     | n.d.       | 0         |
| ferulic acid (25)                              | 179.1                     | n.d.       | 0         |
| 4-hydroxyphenylacetic acid (26)                | 137.1                     | n.d.       | 0         |
| 2-hydroxyphenylacetic acid (27)                | 137.1                     | n.d.       | 0         |
| 4-aminophenylacetic acid (28)                  | 136.1                     | n.d.       | 0         |
| t.a., trace amount; the conversion yield was less than 1%; n.d., not detected by LC-MS.

The CAR encoding gene pnlA was located together with a transcription factor (TF), PFICI_00365, and a non-reducing polyketide synthase (NR-PKS), PFICI_00366, in the biosynthetic gene cluster (BGC) (Figure S3). The assembly combination of NR-PKS and CAR was also found for other CAR members in clade VI [13,14,17,19–21,37]. In contrast, PFICI_00366, containing a domain structure of a starter-unit acyltransferase (SAT), a ketosynthase (KS), an acyltransferase (AT), a product template (PT), an acyl carrier protein (ACP), a methyltransferase (MT) and a TE [39], shares relatively high identities of 34–37% with AusA, PrhL, AdrD, Trt4 and AndM on the amino acid level, all of which catalyze the formation of 3,5-dimethylorsellinic acid (4) [40–44]. Thus, we proposed that PFICI_00366 also catalyzes the formation of 4 as an NR-PKS assembly line, which was subsequently reduced by PnlA for aldehyde (4') formation (Figure S3). To prove the activity of PnlA toward 4, we carried out an in vitro enzyme assay in analogy to that toward 1. Surprisingly, 4 was completely converted to the corresponding aldehyde product, 4', with [M + H]^+ ions at m/z 181.0 (Figure 3 and Table 1). The remarkable conversion yield indicated that the NR-PKS product of PFICI_00366 4 could be the natural substrate of PnlA. To the best of our knowledge, no CAR was previously discovered to accept aryl carboxylic acid containing three methyl substituents, implying the catalytic potential of PnlA to study the effect of trimethylation on substrate specificity.
3.3. Flexible Biocatalytic Activities of PnlA

Given the activities of PnlA toward its probably natural substrate 4 and another well-accepted substrate by CARs, 1, we speculated that PnlA also showed flexible biocatalytic activities toward a variety of substrates. Therefore, we performed enzyme assays of PnlA with 26 different carboxylic acid derivatives (2, 3 and 5–28) (Figure S4). LC-MS analysis of the incubation mixtures showed that PnlA accepted 12 substrates (2, 3 and 5–14) for the corresponding aldehyde generation (2′, 3′ and 5′–14′) (Figure 3). In contrast to AtCAR with the best accepted substrate, 2, the activity of PnlA toward 2 dramatically decreased by ~80% in comparison with that toward 4. However, the removal of the 5-methyl group in 4 (3) led to a reduced effect of the substrate specificity, with a decreasing conversion from 100% to 58.8%. In addition, only 22.6% of 1 and 16.3% of 2,4-dihydroxybenzoic acid (5) were consumed by PnlA. The biocatalytic activities of PnlA toward 1–5 indicated the importance of dimethylation at positions C3 and C5 in contributing the substrate specificity of the A domain.

When the 2-hydroxyl moiety was removed from 5, the specific activity of PnlA toward 4-hydroxybenzoic acid (6) obviously increased, with a conversion yield of 45.0%. The methoxylation at the C4 position (7) led to a slightly lower conversion yield comparing with that with 6. However, the activity of PnlA significantly decreased by 10-fold in the presence of larger substitution at the C4 position, like the butoxy moiety (8). This might be caused by the steric hindrance. In comparison with 6, dihydroxylation at the C3 and C4 positions slightly increased the activity of PnlA (53.9%). Interestingly, PnlA showed much better acceptance for vanillic acid (10), with a conversion yield of 95.8%, which was only lower than the acceptance for its natural substrate, 4, whereas only 7.7% of isovanillic acid (11) was consumed by PnlA. In contrast to the 4-hydroxy-containing benzoic acids as mentioned above, no substrate consumption was observed for benzoic acid (15), salicylic acid (16) and 3-hydroxybenzoic acid (17) (Figure S2). These results indicate the crucial role of 4-hydroxyl substituent in benzoic acid for the substrate specificity because hydroxyl moiety interacted with the key residues of the A domain to form hydrogen bonds.

PnlA can accept 3,4-diaminobenzoic acid (12) and 4-amino-3-nitrobenzoic acid (13) to form the corresponding aldehydes according to LC-MS analysis but afford products only with trace amounts (Figure 3 and Table 1). In comparison, better acceptance was observed for 4-amino-2-chlorobenzoic acid (14), with a conversion yield of 22.6%, proving the importance of 4-hydroxyl substitution in benzoic acid for PnlA specificity. In addition, no aldehyde product was detected for the substrates including nitrobenzoic acids (18–20), phthalic acid (21), cinnamic acid derivatives (22–25) and other aryl-aliphatic carboxylic acids (26–28) (Figure S2).

Considering the differing substrate specificities of the CARs in clade VI, we further carried out the multiple sequence alignments (Figures 4 and S5). The conserved residues of CARs can be found in PnlA, such as G260, V264, T304, L324 and L352, which were also described for AtCAR. However, residue differences at the 355 position were revealed, e.g., Q355 in PnlA, F355 in StbB and AscB, as well as H355 in six other proteins. Previous studies indicated H355 as a key residue of AtCAR for recognizing 2 as its predominant substrate. This explains its high conservation in Esp4 and CicB, with the substrate preference of 2 [14,19]. In contrast, StbB and AscB contained F355 at this site, which might be relevant to their preference for the prenylated substrate [13,21]. Thus, we elucidated Q355 as a crucial residue for the A domain of PnlA in the catalytic process.
The conversion of 4 to 4′ by a CAR was identified for the first time in this study.

4. Discussion

Aldehyde-containing products, such as vanillin, benzaldehyde and cinnamaldehyde, have attracted the attention of researchers owing to their importance in the fragrance and flavoring industries and their high reactivities via nucleophilic addition [32,45]. Chemical reduction strategies of carboxylic acids normally result in low yields of aldehyde formation and require the introduction of protecting groups to limit over-reduction or off-target reduction. Alternatively, biocatalytic strategies, including carboxylic acid reduction and alcohol oxidation, can efficiently provide the targeted aldehyde products [10]. Enzymatic reduction of carboxylic acids via CARs was reported as earlier as 1959 for NcCAR characterization [9]. Genome mining and the techniques of genetic manipulation have led to approximately 50 identified CARs with a broad substrate scope for accepting aromatic, polycyclic and heteroaromatic carboxylic acids; aliphatic carboxylic acids; and amino acids [8]. Inspired by the catalytic potential of CARs, we constructed a fungal virtual CAR library containing 84 enzymes with 55 new CARs from 28 species for the first time in this study.

The combination of phylogenetic analysis of 84 fungal CARs and enzyme/substrate correlations of 29 known CARs prompted us to postulate the substrate scopes of fungal CARs (Figure 2). For example, 17 CARs clustered in clade I represent a toolbox for catalysis of amino acid reduction, probably undergoing piperazine core formation as the products of LnaA, LnaB, HqlA and BrsA [15,16,34]. Nine CARs clustered in clade VI, including a newly identified PnlA from P. fici and eight other known CARs, demonstrating substrate preferences for orsellinic acid derivatives, which were assembled by NR-PKS. An NR-PKS-encoding gene, PFICI_00366, was also found in the same cluster with pnlA. Therefore, we proposed PFICI_00366 for assembly of 4 based on its high identities with the reported biosynthetic enzymes and the subsequent reduction to generate its aldehyde product (4′) (Figure S3) [40–44]. Notably, the conversion of 4 to 4′ by a CAR was identified for the first time in this study, to the best of our knowledge. These results imply a strategy to generate aldehyde products by coexpression of an NR-PKS gene with a CAR gene.
In analogy to AtCAR, in vitro biochemical investigation of PnlA also revealed a broad substrate scope for accepting 14 of 28 benzoic acid derivatives (Figure 3 and S4). 4 and 10 are the best accepted substrates of PnlA, with conversion yields of 100% and 95.8%, respectively (Table 1). The comparison of the substrate consumption indicated the importance of 4-hydroxyl substitution in benzoic acid for the substrate specificity of PnlA, as this moiety functioned in the interaction with the A domain to construct hydrogen bonds. Furthermore, we postulated a key residue, Q355 of PnlA, contributing to its biocatalytic specificity by multiple sequence alignments. In contrast, eight other CARs in clade VI carry F355 and H355, which might determine their substrate preferences for 1 and 2 or the prenylated 1. Although cinnamic acid (22) is one of the most popular CAR substrates in previous studies, no consumption of 22 or its derivatives (23–25) can be observed in the incubation mixtures with PnlA.

In summary, we established a fungal virtual CAR library containing 84 fungal CARs with 55 new CARs. Their substrate scopes can be postulated on the basis of phylogenetic analysis and the known CAR substrate scopes. The recombinant protein production in S. cerevisiae and in vitro biochemical investigation promoted us to identify PnlA from P. fici as a new CAR for aldehyde generation. We proved the flexible biocatalytic activities of PnlA toward 14 benzoic acid derivatives (1–14) and uncovered its specificity for 4-hydroxylated benzoic acids. A key specificity-conferring code, Q355, was postulated for the A domain of PnlA. Our study provides a promising biocatalyst for aldehyde generation.

**Supplementary Materials:** The following supporting information can be downloaded at: https://https://www.mdpi.com/article/10.3390/jof8101001/s1, Table S1: Plasmids, strains and primers; Figure S1: Overproduction of PnlA in S. cerevisiae; Figure S2: LC-MS analyses of the incubation mixtures of PnlA with different benzoic acid derivatives (15–28), but no corresponding aldehyde product was observed; Figure S3: The biosynthetic gene cluster (BGC) containing pnlA and its proposed biosynthetic pathway; Figure S4: All substrates screened for PnlA biocatalytic activity in this study; Figure S5: Alignment of fungal CARs located in clade VI, including PnlA, StbB, AscB, CicB, Esp4, AtCAR, PMAA_062890, ATEG_07380 and Pks5.

**Author Contributions:** Conceptualization, W.-B.Y.; methodology, W.-B.Y. and J.F.; software, P.-L.W.; validation, J.F. and H.R.; formal analysis, Y.L. and P.-L.W.; investigation, Y.L. and P.-L.W.; data curation, Y.L., P.-L.W. and J.F.; writing—original draft preparation, Y.L., P.-L.W. and J.F.; writing—review and editing, W.-B.Y. and J.F.; visualization, J.F. and P.-L.W.; supervision, W.-B.Y.; project administration, W.-B.Y.; funding acquisition, W.-B.Y., J.F. and H.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by National Key Research and Development Program of China (2020YFA0907800), the National Natural Science Foundation of China (31861133004) and the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (ZDBS-LY-SM016), as well as the Biological Resources Program, Chinese Academy of Sciences (KFJ-BRP-009-005). Jie Fan (YJ20200201, 2020M680720 and 2022T150689) and Huomiao Ran (YJ20200309 and 2021M693362) were supported by the China Postdoctoral Science Foundation.

**Institutional Review Board Statement:** This study does not contain any studies with human or animal subjects.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in this manuscript and can be requested from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.
References

1. Adrover-Castellano, M.L.; Schmidt, J.J.; Sherman, D.H. Biosynthetic cyclization catalysts for the assembly of peptide and polyketide natural products. *ChemCatChem* 2021, 13, 2095–2116. [CrossRef] [PubMed]

2. Neethu, K.; Karthik, L.; Li, Z. Nonribosomally and ribosomally synthesized bioactive peptides (NRPSs and RPPs) from Actinobacteria. In *Actinobacteria: Microbiology to Synthetic Biology*; Karthik, L., Ed.; Springer Nature: Singapore, 2022; pp. 87–99.

3. Stüßmuth, R.D.; Mainz, A. Nonribosomal peptide synthesis-principles and prospects. *Angew. Chem. Int. Ed.* 2017, 56, 3770–3821. [CrossRef] [PubMed]

4. Guo, C.J.; Wang, C.C. Recent advances in genome mining of secondary metabolites in *Aspergillus terreus*. *Front. Microbiol.* 2014, 5, 717. [CrossRef] [PubMed]

5. Zhu, J.; Chen, W.; Li, Y.Y.; Duan, J.; Liu, Y.; Shi, G.Y.; Xie, C.; Wang, H.X.; et al. Identification and catalytic characterization of a nonribosomal peptide synthetase-like (NRPS-like) enzyme involved in the biosynthesis of ecosides from *Streptomyces* sp. *LZ35*. *Gene* 2014, 546, 352–358. [CrossRef] [PubMed]

6. Hühner, E.; Backhaus, K.; Kraut, R.; Li, S.-M. Production of α-keto carboxylic acid dimers in yeast by overexpression of NRPS-like genes from *Aspergillus terreus*. *Appl. Microbiol. Biotechnol.* 2018, 102, 1663–1672. [CrossRef]

7. Guo, C.-J.; Knox, B.P.; Sanchez, J.E.; Chiang, Y.-M.; Bruno, K.S.; Wang, C.C.C. Application of an efficient gene targeting system linking secondary metabolites to their biosynthetic genes in *Aspergillus terreus*. *Org. Lett.* 2013, 15, 3562–3565. [CrossRef]

8. Winkler, M.; Ling, J.G. Biocatalytic carboxylate reduction—Recent advances and new enzymes. *ChemCatChem* 2022, e202200441. [CrossRef]

9. Farmer, V.C.; Henderson, M.E.K.; Russell, J.D. Reduction of certain aromatic acids to aldehydes and alcohols by *Polytrichum versicolor*. *Biochim. Biophys. Acta* 1959, 35, 202–211. [CrossRef]

10. Ran, H.; Li, S.-M. Fungal benzene carbaldehydes: Occurrence, structural diversity, activities and biosynthesis. *Nat. Prod. Rep.* 2021, 38, 240–263. [CrossRef]

11. Winkler, M. Carboxylic acid reductase enzymes (CARs). *Curr. Opin. Chem. Biol.* 2018, 43, 23–29. [CrossRef]

12. Kuhnert, E.; Navarro-Muñoz, J.C.; Becker, K.; Stadler, M.; Collemare, J.; Cox, R.J. Secondary metabolite biosynthetic diversity in the fungal family *Hypoxylaceae* and *Xylaria hypoxylon*. *Stud. Mycol.* 2021, 99, 100118. [CrossRef] [PubMed]

13. Li, C.; Matsuda, Y.; Gao, H.; Hu, D.; Yao, X.S.; Abe, I. Biosynthesis of LL-Z1272β: Discovery of a new member of NRPS-like enzymes for aryl-aldehyde formation. *ChemBioChem* 2016, 17, 904–907. [CrossRef] [PubMed]

14. Zhu, G.; Hou, C.; Yuan, W.; Wang, Z.; Zhang, J.; Jiang, L.; Karthik, L.; Li, B.; Ren, B.; Lv, K.; et al. Molecular networking assisted discovery and biosynthesis elucidation of the antimicrobial spiroketals epicospirocins. *Chem. Commun.* 2020, 56, 10171–10174. [CrossRef]

15. Forseth, R.R.; Amaike, S.; Schwenk, D.; Affeldt, K.J.; Hoffmeister, D.; Schroeder, F.C.; Keller, N.P. Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. *Angew. Chem. Int. Ed.* 2013, 52, 1590–1594. [CrossRef]

16. Yuan, B.; Liu, D.; Guan, X.; Yan, Y.; Zhang, J.; Zhang, Y.; Yang, D.; Ma, M.; Lin, W. Piperazine ring formation by a single-module NRPS and cleavage by an α-KG-dependent nonheme iron dioxygenase in brasilimide biosynthesis. *Appl. Microbiol. Biotechnol.* 2020, 104, 6149–6159. [CrossRef] [PubMed]

17. Wang, M.; Zhao, H. Characterization and engineering of the adenylation domain of a NRPS-like protein: A potential biocatalyst for aldehyde generation. *ACS Catal.* 2014, 4, 1219–1225. [CrossRef]

18. Winkler, M.; Winkler, C.K. *Trametes versicolor* carboxylate reductase enzyme uncovered. *Monatsh. Chem.* 2016, 147, 575–578. [CrossRef]

19. Sanchez, J.F.; Entwistle, R.; Corcoran, D.; Oakley, B.R.; Wang, C.C.C. Identification and molecular genetic analysis of the chiorine gene cluster in *Aspergillus nidulans*. *MedChemComm* 2012, 3, 997–1002. [CrossRef]

20. Reyes-Fernández, E.Z.; Shi, Y.-M.; Grün, P.; Bode, H.B.; Bölker, M. An unconventional melanin biosynthesis pathway in *Ustilago maydis*. *Appl. Environ. Microbiol.* 2021, 87, e01510–e01520. [CrossRef]

21. Araki, Y.; Awakawa, T.; Matsuzaki, M.; Cho, R.; Matsuda, Y.; Hoshino, S.; Shihohara, Y.; Yamamoto, M.; Kido, Y.; Inaoka, D.K.; et al. Complete biosynthetic pathways of ascofuracnone and ascochlorin in *Acremonium egyptiacum*. *Proc. Natl. Acad. Sci. USA* 2019, 116, 8269–8274. [CrossRef]

22. Lebar, M.D.; Cary, J.W.; Majumdar, R.; Carter-Wientjes, C.H.; Mack, B.M.; Wei, Q.; Uka, V.; De Saeger, S.; Di Mavungu, J.D. Identification and functional analysis of the aspergillic acid gene cluster in *Aspergillus flavus*. *Fungal Genet. Biol.* 2018, 116, 14–23. [CrossRef] [PubMed]

23. Zhang, P.; Wang, X.; Fan, A.; Zheng, Y.; Liu, X.; Wang, S.; Zou, H.; Oakley, B.R.; Keller, N.P.; Yin, W.B. A cryptic pigment biosynthetic pathway uncovered by heterologous expression is essential for conidial development in *Pestalotiopsis fici*. *Mol. Microbiol.* 2017, 105, 469–483. [CrossRef] [PubMed]

24. Wang, X.; Zhang, X.; Liu, X.; Xiang, M.; Wang, W.; Sun, X.; Che, Y.; Guo, L.; Liu, G.; Guo, L.; et al. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genom.* 2015, 16, 28. [CrossRef] [PubMed]

25. Sommers, C.H.; Raikowski, K.T. Inactivation of *Escherichia coli* Biosynthesis of LL-Z1272109, DH5α, and O157:H7 suspended in Butterfield’s phosphate buffer by gamma irradiation. *J. Food Sci.* 2008, 73, M87–M90. [CrossRef] [PubMed]

26. Bond, C.; Tang, Y.; Li, L. *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases. *Fungal Genet. Biol.* 2016, 89, 52–61. [CrossRef] [PubMed]
27. Ma, S.M.; Li, J.W.-H.; Choi, J.W.; Zhou, H.; Lee, K.K.M.; Moorthie, V.A.; Xie, X.; Kealey, J.T.; Da Silva, N.A.; Vederas, J.C.; et al. Complete reconstitution of a highly reducing iterative polyketide synthase. Science 2009, 326, 589–592. [CrossRef]

28. Zhou, S.; Zhang, P.; Zhou, H.; Liu, X.; Li, S.-M.; Guo, L.; Li, K.; Yin, W.-B. A new regulator RsdA mediating fungal secondary metabolism has a detrimental impact on asexual development in Pestalotiopsis fici. Environ. Microbiol. 2019, 21, 416–426. [CrossRef]

29. Bok, J.W.; Keller, N.P. Fast and easy method for construction of plasmid vectors using modified quick-change mutagenesis. Methods Mol. Biol. 2012, 944, 163–174. [CrossRef]

30. Zou, Y.; Garcia-Borràs, M.; Tang, M.; Hirayama, Y.; Li, D.; Li, L.; Watanabe, K.; Houk, K.N.; Tang, Y. Enzyme-catalyzed cationic epoxide rearrangements in quinolone alkaloid biosynthesis. Nat. Chem. Biol. 2017, 13, 325–332. [CrossRef]

31. Yu, X.; Li, S.-M. Prenyltransferases of the dimethylallyltryptophan synthase superfamily. Methods Enzymol. 2012, 516, 259–278. [CrossRef]

32. Butler, N.; Kunjapur, A.M. Carboxylic acid reductases in metabolic engineering. J. Biotechnol. 2020, 307, 1–14. [CrossRef] [PubMed]

33. Brandenburger, E.; Braga, D.; Kombrink, A.; Lackner, G.; Gressler, J.; Künzler, M.; Hoffmeister, D. Multi-genome analysis identifies functional and phylogenetic diversity of basidiomycete adenylate-forming reductases. Fungal Genet. Biol. 2018, 112, 55–63. [CrossRef] [PubMed]

34. Yang, C.; Xu, Y.; Xu, K.; Tan, G.; Yu, X. Preparation of new halogenated diphenyl pyrazine analogs in Escherichia coli by a mono-module fungal nonribosomal peptide synthetase from Penicillium herquei. Tetrahedron Lett. 2018, 59, 3084–3087. [CrossRef]

35. Hai, Y.; Chen, M.; Huang, A.; Tang, Y. Biosynthesis of mycotoxin fusaric acid and application of a PLP-dependent enzyme for chemoenzymatic synthesis of substituted L-pipeolic acids. J. Am. Chem. Soc. 2020, 142, 19668–19677. [CrossRef]

36. Wang, M.; Beissner, M.; Zhao, H. Aryl-aldehyde formation in fungal polyketides: Discovery and characterization of a distinct biosynthetic mechanism. Chem. Biol. 2014, 21, 257–263. [CrossRef]

37. Fernández-Martínez, L.T.; Borsetto, C.; Gomez-Escribano, J.P.; Bibb, M.J.; Al-Bassam, M.M.; Chandra, G.; Bibb, M.J. New insights into chloramphenicol biosynthesis in Streptomyces venezuelae ATCC 10712. Antimicrob. Agents Chemother. 2014, 58, 7441–7450. [CrossRef]

38. Herbst, D.A.; Townsend, C.A.; Maier, T. The architectures of iterative type I PKS and FAS. Nat. Prod. Rep. 2018, 35, 1046–1069. [CrossRef]

39. Lo, H.C.; Entwistle, R.; Guo, C.J.; Ahuja, M.; Szewczyk, E.; Hung, J.H.; Chiang, Y.M.; Oakley, B.R.; Wang, C.C. Two separate gene clusters encode the biosynthetic pathway for the meroterpenoids austinol and dehydroaustinol in Aspergillus nidulans. J. Am. Chem. Soc. 2012, 134, 4709–4720. [CrossRef]

40. Matsuda, Y.; Iwabuchi, T.; Fujimoto, T.; Awakawa, T.; Nakashima, Y.; Mori, T.; Zhang, H.; Hayashi, F.; Abe, I. Discovery of key dioxygenases that diverged the paraherquonin and acetoxydehydroaustin pathways in Penicillium brasiliense. J. Am. Chem. Soc. 2016, 138, 12671–12677. [CrossRef]

41. Matsuda, Y.; Awakawa, T.; Abe, I. Reconstituted biosynthesis of fungal meroterpenoid andrastin A. Tetrahedron Lett. 2013, 69, 8199–8204. [CrossRef]

42. Matsuda, Y.; Awakawa, T.; Abe, I. Complete biosynthetic pathway of anditomin: Nature’s sophisticated synthetic route to a complex fungal meroterpenoid. J. Am. Chem. Soc. 2014, 136, 15326–15336. [CrossRef] [PubMed]

43. Kohlpaintner, C.; Schulte, M.; Falbe, J.; Lappe, P.; Weber, J. Aldehydes, aliphatic and araliphatic. In Ullmann’s Encyclopedia of Industrial Chemistry; John Wiley & Sons: Hoboken, NJ, USA, 2000.