Supporting Information

# **Cupin-domain containing protein is not essential for the alkyl salicylaldehyde formation in** *Aspergillus ustus*

Marlies Peter and Shu-Ming Li\*

Institut für Pharmazeutische Biologie und Biotechnologie, Fachbereich Pharmazie, Philipps-Universität Marburg, Robert-Koch-Straße 4, 35037 Marburg, Germany.

\* Corresponding author, e-mail: shuming.li@staff.uni-marburg.de.

# **Table of contents**





## <span id="page-3-0"></span>**Experimental section**

## <span id="page-3-1"></span>**Sequence analysis**

The genome of *Aspergillus ustus* 3.3904 was sequenced and published by Pi *et al*. in 2015. 1 The genomic sequence reported in this study is available in GenBank under the accession number JOMC01000132. To predict biosynthetic gene clusters, initial analysis was carried out with antiSMASH (http://antismash.secondarymetabolites.org/)<sup>2</sup> and 2nd find (https://biosyn.nih.go.jp/2ndfind/). Detailed investigation including annotation of genes and prediction of enzyme functions (Table S1) were conducted using BLASTp program (http://www.ncbi.nlm.nih.gov).

## <span id="page-3-2"></span>**Strains, media, and growth conditions**

All strains used and created in this study are listed in Table S2. *Escherichia coli* DH5α and *Saccharomyces cerevisiae* HOD114-2B cells were cultivated as described before. <sup>3</sup> *Aspergillus ustus* 3.3904 was purchased from China General Microbiological Culture Collection Centre (Beijing, China) and cultivated in liquid potato dextrose broth (PDB, Sigma, Germany, Darmstadt) for 3 days at 30 °C for harvesting mycelium for subsequent genomic DNA isolation. *Aspergillus nidulans* strains were grown at 37 °C on LMM agar plates for spore production with the required supplements (1 g/L uracil, 1.2 g/L uridine, 2.5 mg/L riboflavin, 0.5 mg/L pyridoxine). The LMM agar plates were prepared as given in a previous publication. <sup>3</sup> For secondary metabolite production, the *A. nidulans* strains were cultivated in liquid PDB medium at 25 °C. To quantify the metabolite production, three independent cultures per strain were set up in 50 mL PDB medium for quantification.

## <span id="page-3-3"></span>**Genomic DNA isolation**

To extract the genomic DNA, liquid PDB medium was inoculated with spores of *A. ustus* 3.3904, *P. roqueforti* FM164 and *A. nidulans* and stand for two days at 37 °C. The media were removed by centrifugation. Subsequent isolation of the genomic DNA was performed as described in a previous publication.<sup>4</sup>

## <span id="page-3-4"></span>**PCR amplification, cloning, and plasmid construction**

The generated plasmids and the used primers are listed in Tables S3 and S4, respectively. Primers were synthesised by SeqLab GmBH (Göttingen, Germany). Q5® High-Fidelity Polymerase (New England Biolabs) was used for PCR amplification as given in the manufacturer's instruction.

To construct the expression plasmid pMP024 with a segment of 13,572 bp (bp 10521 – 24091 of JOMC01000132), three fragments of 5241, 5700, and 2630 bp were amplified from the genomic DNA of *A. ustus* 3.3904. The fragments share overlapping regions of approximately 300 bp to each other and 30 bp to the vector pJN017.<sup>5</sup> The sequence of the transcription factor KIA75503 is directly cloned behind the strong *gpdA* promoter sequence. The three fragments and the SfoI-linearised pJN017 were assembled via homologous recombination in *Saccharomyces cerevisiae* HOD114-2B. 6

For cloning of the expression construct pMP046, the putative gene coding for PsaCP with a length of 725 bp and additional 500 bp at 5'-UTR and 380 bp at 3'-UTR was amplified via PCR from the genomic DNA of *A. ustus* including 30 bp overhang on both 5' and 3' to the plasmid pMP024. Insertion of *psaCP* into the NotI-linearised expression plasmid pMP024 was carried out by using homologous recombination in *Escherichia coli* DH5α.<sup>7</sup>

To generate the deletion constructs pMP030, pMP038, pMP039, and pMP040, fragments of 5'-UTR and 3'-UTR in lengths of about 620 – 730 bp were amplified from genomic DNA from *P. roqueforti* for *anuC* deletion or *A. ustus* for deletion of *psaOX1*, *psaPS*, and *psaOX2*, respectively. The vector backbone including *ampR* and URA3 was amplified from pJN017 and afpyrG from pYH-wA-pyrG, respectively.<sup>8</sup> Assembly of fragments of 5'-UTR, 3'-UTR, vector backbone and *afpyrG* was achieved in yeast as mentioned above.<sup>6</sup>

For cloning of the deletion plasmid pMP043, the same strategy was used as for other deletion constructs described above. 5'-UTR and 3'-UTR of the putative gene ANIA\_7635 from *A. nidulans*, the vector backbone including *ampR* and URA3, and *afpyro* from pYWB1<sup>9</sup> were fused via homologous recombination in yeast. 6

To verify the plasmids, digestion with suitable restriction enzymes was conducted (Figure S1). The final plasmids, both expression and deletion constructs, were linearised by SwaI and transformed into *A. nidulans* protoplasts with a PEG-mediated method according to the protocols reported previously.8,10

#### <span id="page-4-0"></span>**Verification and cultivation of** *A. nidulans* **transformants**

Selection of transformants was performed directly after transformation with the subsequent cultivation using riboflavin, uracil and uridine, and pyridoxine autotrophy. For the expression strain MP13 a visual control of conidia colour change from green to white provided a first hint of correct integration in the *wA* locus. Genomic DNA was isolated from both expression and deletion strains and the correct integration was verified by PCR. Primers for verification of expression strains are designed to bind upstream or downstream of the *wA* locus and within the desired gene(s). The same approach was used for the deletion strains. Primers bind upstream or downstream of the deleted gene and within the *afpyrG* or *afpyro* gene. Additionally, amplification from upstream to downstream of a deleted gene was carried out to test its locus and to ensure its absence. The used primers are given in Table S4. The results of transformant verification are illustrated in Figure S1.

The correct transformants were cultivated in liquid PDB media for 14 days at 25 °C for detection of secondary metabolites. Samples of the cultures were extracted with equal volumes of EtOAC for three times and after evaporation dissolved in MeOH for LC-MS analysis.

#### <span id="page-5-0"></span>**Large-scale fermentation, extraction, and isolation of secondary metabolites**

To isolate compounds **1** – **4**, 6 L of PDB medium were inoculated with spores from *A. nidulans* MP13 and cultivated for 14 days at 25 °C. The mycelia were separated from the medium by filtration and washed with water for three times. Extraction was carried out using equal volume of EtOAc for three times. The crude extract was obtained after evaporation under reduced pressure. 2.6 g crude extract were subjected to silica gel column chromatography (silica gel 60, 0.04 – 0.063 mm, Roth, Karlsruhe) and separation of secondary metabolites was achieved by elution with CH<sub>2</sub>Cl<sub>2</sub>/MeOH (100:10). The obtained 9 fractions were analysed with LC-MS and further purified using semi-preparative HPLC on an Agilent HPLC 1200 equipped with an XDB-C18 column (9.4 × 250 mm, 5 µm, VDS Optilab Chromatographie Technik GmbH, Bremen, Germany).

Fraction 3 was used to purify compounds **1** and **2**. 32 mg of compound **1** were obtained after isocratic elution (H<sub>2</sub>O/CH<sub>3</sub>CN = 50/50, flow rate = 2.0 mL/min,  $\lambda$  = 280 nm,  $t_R$  = 15.2 min). Further purification of fraction 3 resulted in 9 mg of compound 2  $(H_2O/CH_3CN = 55/45$ , flow rate = 2.0 mL/min,  $\lambda$  = 310, 280 nm,  $t_R$  = 18.7 min). Compound 3 (16 mg) was isolated from fraction 2 by isocratic elution (H<sub>2</sub>O/CH<sub>3</sub>CN = 10/90, flow rate = 2.0 mL/min,  $\lambda$  = 340, 270 nm,  $t<sub>R</sub>$  = 10.5 min). Fraction 4 was purified with isocratic elution (H<sub>2</sub>O/CH<sub>3</sub>CN = 68/32, flow rate = 2.0 mL/min,  $\lambda$  = 270 nm,  $t_R$  = 7 min), leading to the isolation of compound 4 (10 mg).

## <span id="page-5-1"></span>**Cultivation of** *A. nidulans* **LO8030 with compounds 1 – 3**

Compounds **1** – **3** were dissolved in DMSO to give 100 mM stock solutions. The *A. nidulans* strain LO8030 was cultivated in 4 flasks with 10 mL PDB media at 25 °C for 2 days. 2.5 µL of **1**, **2**, or **3** were added at day 3 (25 µM). After cultivation for another 1, 3, and 5 days samples were extracted with EtOAc and analysed on LC-MS. Additionally, compounds **1** – **3** were incubated with PDB medium and treated in the same procedure as described above to test their stability.

## <span id="page-5-2"></span>**LC-MS analysis**

For LC-MS monitoring of secondary metabolites, an Agilent 1260 HPLC equipped with VDSpher PUR100 C18-M-SE column (150 × 2.0 mm, 3 μm, VDS optilab Chromatographie Technik GmbH) and a Bruker microTOF QIII mass spectrometer was used as reported in a previous publication.<sup>3</sup> Elution profile A corresponds to a method with a linear gradient of acetonitrile in water (5 – 100%) within 30 minutes. Elution profile B uses a linear gradient of acetonitrile in water  $(5 - 100\%)$  within 10 minutes. A flow rate of 0.3 mL min<sup>-1</sup> was used. For determination of the exact masses, positive ion mode electrospray ionization (ESI) was selected with a capillary voltage set to 4.5 kV and a collision energy of 8.0 eV. 5 mM sodium formate was used in each run for mass calibration. The masses were scanned in the range of *m/z* 100.000 – 1500.000. Evaluation of the data was conducted by using Compass DataAnalysis 4.2 software (Bruker Daltonik, Bremen, Germany).

#### <span id="page-6-0"></span>**NMR analysis**

The NMR spectra of the purified compounds **1** – **4** were recorded on a JOEL ECA-500 MHz spectrometer (JOEL, Tokyo, Japan) with CDCl<sub>3</sub> as solvent. Analysis of all spectra was performed by using MestReNova 14.0.0 (Mestrelab Research).

#### <span id="page-6-1"></span>**Structural elucidation**

The isolated compounds **1** – **4** were determined by comprehensive interpretation of their MS and NMR data (Figures S3–19). Compound **1** was assigned to 6-propyl salicyl alcohol by comparison of its NMR data with the reported data. <sup>11</sup> A singlet for two protons of the methylene group (C-7) was observed at  $\delta_H$  4.91 ppm in the <sup>1</sup>H NMR spectra and  $\delta_C$  at 60.2 ppm in the <sup>13</sup>C{<sup>1</sup>H} NMR spectra.<sup>12</sup> The structures of compounds **2** and **3** were deduced from **1** as alkylated salicylic acid and alkylated salicylaldehyde, respectively. A singlet for one proton of the aldehyde group of 3 was observed at δ<sub>H</sub> 11.94 ppm and δ<sub>c</sub> at 195.4 ppm.<sup>12</sup> Compound 4 showed signals for a dihydro-γ-pyrone moiety comparable to those described previously.<sup>13</sup> Signals for three methylene groups and one methyl group in the high field of the <sup>1</sup>H NMR ( $\delta_H$ )  $0.91 - 2.21$  ppm) and another methylene group at δ<sub>H</sub> 3.83 ppm can be assigned to δ<sub>c</sub> at 13.7  $-$  37.2 ppm and δ<sub>c</sub> at 58.9 ppm, respectively, with HSQC correlations. Interpretation of the <sup>1</sup>H-<sup>1</sup>H COSY spectra revealed the presence of a propyl side chain and a hydroxyethyl group at C-3. The HMBC data confirmed the connection of the hydroxyethyl group at C-3 and the propyl group at C-7 to the dihydro-γ-pyrone moiety.

#### <span id="page-6-2"></span>**Physicochemical properties**

6-Propyl salicyl alcohol (**1**): white powder. UV (deduced from HPLC analysis) λmax 230, 280 nm. <sup>1</sup>H NMR (500 MHz; CDCl3): δ 7.09 (m, 1H), 6.73 (d, *J* = 8.2, 1H), 6.70 (d, *J* = 7.6, 1H), 4.91 (s, 2H), 2.55 (m, 2H), 1.52 (m, 2H), 0.94 (t, *J* = 7.3, 3H). <sup>13</sup>C NMR (125 MHz; CDCl3): δ 156.7, 141.1, 128.9, 122.9, 121.8, 114.6, 60.2, 35.4, 24.9, 14.1. HRMS (ESI) m/z: [M+Na] + Calcd for  $C_{10}H_{14}O_2$ Na, 189.0886; found 189.0893.

6-Propyl salicylic acid (**2**): white powder. UV (deduced from HPLC analysis) λmax 248, 284, 314 nm. <sup>1</sup>H NMR (CDCl3, 500 MHz): δ 7.36 (m, 1H), 6.87 (d, *J* = 8.3, 1H), 6.77 (d, *J* = 7.6, 1H), 2.96 (m, 2H), 1.64 (m, 2H), 0.98 (t, *J* = 7.3, 3H). <sup>13</sup>C NMR (CDCl3, 125 MHz): δ 175.5, 163.8, 147.5, 135.5, 123.0, 116.0, 110.6, 38.6, 25.3, 14.4. HRMS (ESI) m/z: [M+H] <sup>+</sup> Calcd for  $C_{10}H_{13}O_3$ , 181.0859; found 181.0856.

6-Propyl salicylaldehyde (**3**): pale yellow oil. UV (deduced from HPLC analysis) λmax 228, 270, 340 nm. <sup>1</sup>H NMR (CDCl3, 500 MHz): δ 11.94 (1H, s), 10.30 (s, 1H), 7.40 (m, 1H), 6.82 (d, *J* = 8.5, 1H), 6.73 (d, *J* = 7.5, 1H), 2.89 (m, 2H), 1.68 (m, 2H), 0.99 (t, *J* = 7.3, 3H). <sup>13</sup>C NMR (CDCl3, 125 MHz): δ 195.4, 163.5, 147.3, 137.5, 121.4, 118.0, 116.1, 33.9, 26.4, 14.0. HRMS (ESI) m/z: [M+H]<sup>+</sup> Calcd for C10H13O2, 165.0910; found 165.0909.

(4): white powder. UV (deduced from HPLC analysis)  $\lambda_{\text{max}}$  224, 270 nm. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500) MHz): δ 5.33 (s, 1H), 4.57 (ddt, *J* = 12.8, 8.5, 4.3, 1H), 3.83 (m, 2H), 2.46 (dd, *J* = 16.8, 12.9, 1H), 2.40 (dd, *J* = 16.8, 4.3, 1H), 2.21 (m, 2H), 2.06 (m, 1H), 1.90 (m, 1H), 1.57 (m, 2H), 0.94 (t, *J* = 7.3, 3H). <sup>13</sup>C NMR (CDCl3, 125 MHz): δ 193.1, 177.6, 104.5, 77.0, 58.9, 41.2, 37.2, 36.8, 19.9, 13.7. HRMS (ESI) m/z: [M+H] <sup>+</sup> Calcd for C10H17O3, 185.1172; found 185.1164.

# **Supplementary Tables**

**Table S1.** Sequence identities of proteins encoded by *psa* cluster from *A. ustus* and ANIA\_7635 from *A. nidulans* with those of known clusters

<span id="page-8-0"></span>

<span id="page-8-1"></span>Accession numbers are given in parenthesis.



# <span id="page-9-0"></span>Table S2. Strains used in this study

CGMCCC: China General Microbiological Culture Collection Center (Beijing, China)



<span id="page-10-0"></span>**Table S3.** Plasmids used in this study

<span id="page-11-0"></span>**Table S4.** Oligonucleotide primers used in this study



**Table S4.** Oligonucleotide primers used in this study (continued)



# <span id="page-13-0"></span>**Supplementary Figures**



<span id="page-13-1"></span>**Figure S1.** PCR verification of plasmids for heterologous expression and deletion constructs as well as of the obtained mutants.

To prove the presence of the target genes, PCR verifications were performed with genomic DNA of the heterologous expression (HE) *A. nidulans* transformants. The primers were designed to bind outside of the *wA* locus and inside the desired gene(s). To verify the deletion strains, the correct site-specific integration (corresponding 5'- and 3'-regions) and amplification of the locus were tested, to prove the absence of the gene and the presence of the inserted selection marker. The GeneRuler DNA Ladder mix from Thermo Fisher served as size standard for DNA fragments. The primers are given in Table S4.



<span id="page-14-0"></span>**Figure S2.** LC-MS analysis of compounds **1** – **3** after incubation in PDB medium or cultivation with *A. nidulans* LO8030.

Incubation of PDB medium with compounds **1** – **3** (A) and cultivation of *A. nidulans* LO8030 with compounds **1** – **3** for 1 day. The analysis was carried out with elution profile B and UV absorptions are illustrated at 270 – 340 nm.



<span id="page-15-0"></span>Figure S3. <sup>1</sup>H NMR spectrum of 1 in CDCl<sub>3</sub> (500 MHz).



<span id="page-16-0"></span>Figure S4. <sup>13</sup>C NMR spectrum of 1 in CDCl<sub>3</sub> (125 MHz).



<span id="page-17-0"></span>Figure S5. HSQC spectrum of 1 in CDCl<sub>3</sub>.



<span id="page-18-0"></span>Figure S6. <sup>1</sup>H-<sup>1</sup>H COSY spectrum of 1 in CDCl<sub>3</sub>.



<span id="page-19-0"></span>Figure S7. HMBC spectrum of 1 in CDCl<sub>3</sub>.



<span id="page-20-0"></span>Figure S8. <sup>1</sup>H NMR spectrum of 2 in CDCl<sub>3</sub> (500 MHz).



<span id="page-21-0"></span>Figure S9. <sup>13</sup>C NMR spectrum of 2 in CDCl<sub>3</sub> (125 MHz).



<span id="page-22-0"></span>Figure S10. HSQC NMR spectrum of 2 in CDCl<sub>3</sub>.



<span id="page-23-0"></span>Figure S11. <sup>1</sup>H-<sup>1</sup>H COSY NMR spectrum of 2 in CDCl<sub>3</sub>.



<span id="page-24-0"></span>Figure S12. HMBC NMR spectrum of 2 in CDCl<sub>3</sub>.



<span id="page-25-0"></span>Figure S13. <sup>1</sup>H NMR spectrum of 3 in CDCl<sub>3</sub> (500 MHz).



<span id="page-26-0"></span>Figure S14. <sup>13</sup>C NMR spectrum of 3 in CDCl<sub>3</sub> (125 MHz).



<span id="page-27-0"></span>Figure S15. <sup>1</sup>H NMR spectrum of 4 in CDCl<sub>3</sub> (500 MHz).



<span id="page-28-0"></span>Figure S16. <sup>13</sup>C NMR spectrum of 4 in CDCl<sub>3</sub> (125 MHz).



<span id="page-29-0"></span>Figure S17. HSQC spectrum of 4 in CDCl<sub>3</sub>.



<span id="page-30-0"></span>Figure S18. <sup>1</sup>H-<sup>1</sup>H COSY NMR spectrum of 4 in CDCl<sub>3</sub>



<span id="page-31-0"></span>Figure S19. HMBC spectrum of 4 in CDCl<sub>3</sub>

## <span id="page-32-0"></span>**References**

- B. Pi, D. Yu, F. Dai, X. Song, C. Zhu, H. Li and Y. Yu, *PLoS. One.*, 2015, **10**, e0116089.
- K. Blin, S. Shaw, A. M. Kloosterman, Z. Charlop-Powers, G. P. van Wezel, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2021, **49**, 29.
- M. Peter, Y. Yang and S.-M. Li, *RSC Adv.*, 2022, **12**, 28171–28177.
- M. Peter, Z. X. Zhang, Y. Yang and S. M. Li, *Org. Lett.*, 2024, **26**, 6670–6674.
- F. Kindinger, J. Nies, A. Becker, T. Zhu and S.-M. Li, *ACS Chem. Biol.*, 2019, **14**, 1227– 1234.
- L. Mojardín, M. Vega, F. Moreno, H. P. Schmitz, J. J. Heinisch and R. Rodicio, *Fungal Genet. Biol.*, 2018, **111**, 16–29.
- A. P. Jacobus and J. Gross, *PLoS. One.*, 2015, **10**, e0119221.
- W. B. Yin, Y. H. Chooi, A. R. Smith, R. A. Cacho, Y. Hu, T. C. White and Y. Tang, *ACS Synth. Biol.*, 2013, **2**, 629–634.
- D. J. Janzen, H. Wang and S.-M. Li, *J. Nat. Prod.*, 2023, **86**, 1779–1785.
- J. Zhou and S.-M. Li, *Appl. Microbiol. Biotechnol.*, 2021, **105**, 9181–9189.
- D. Weber, S. Gorzalczany, V. Martino, C. Acevedo, O. Sterner and T. Anke, *Z. Naturforsch.*, 2005, **60c**, 467–477.
- J. Nies, H. Ran, V. Wohlgemuth, W. B. Yin and S.-M. Li, *Org. Lett.*, 2020, **22**, 2256–2260.
- R. Yang, J. Feng, H. Xiang, B. Cheng, L.-D. Shao, Y.-P. Li, H. Wang, Q.-F. Hu, W.-L. Xiao, Y. Matsuda and W.-G. Wang, *J. Am. Chem. Soc.*, 2023, **145**, 11293–11300.
- M. R. Green and J. Sambrook, *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2012, 4th.
- G. Gillot, J. L. Jany, R. Dominguez-Santos, E. Poirier, S. Debaets, P. I. Hidalgo, R. V. Ullan, E. Coton and M. Coton, *Food Microbiol.*, 2017, **62**, 239–250.
- Y. M. Chiang, M. Ahuja, C. E. Oakley, R. Entwistle, A. Asokan, C. Zutz, C. C. Wang and B. R. Oakley, *Angew. Chem., Int. Ed.*, 2016, **55**, 1662–1665.
- P. Xiang, B. Kemmerich, L. Yang and S.-M. Li, *Org. Lett.*, 2022, **24**, 6072–6077.