Supporting information

Completing the puzzle of aurachin biosynthesis in *Stigmatella aurantiaca* Sg a15

Dominik Pistorius, Yanyan Li, ‡ Axel Sandmann§ and Rolf Müller

Helmholtz Institute for Pharmaceutical Research, Helholtz Center for infection Research and Department of Pharmaceutical Biotechnology, Saarland University, P.O. Box 151150, 66041 Saarbrücken, Germany
‡ Present address: Unité “Molécules de Communication et Adaptation des Microorganismes”, UMR 7245 CNRS/Muséum National d´Histoire Naturelle, 57 rue Cuvier, 75005 Paris, France
§ Present address: syskondata GmbH, Campestr. 7, 38102 Braunschweig, Germany

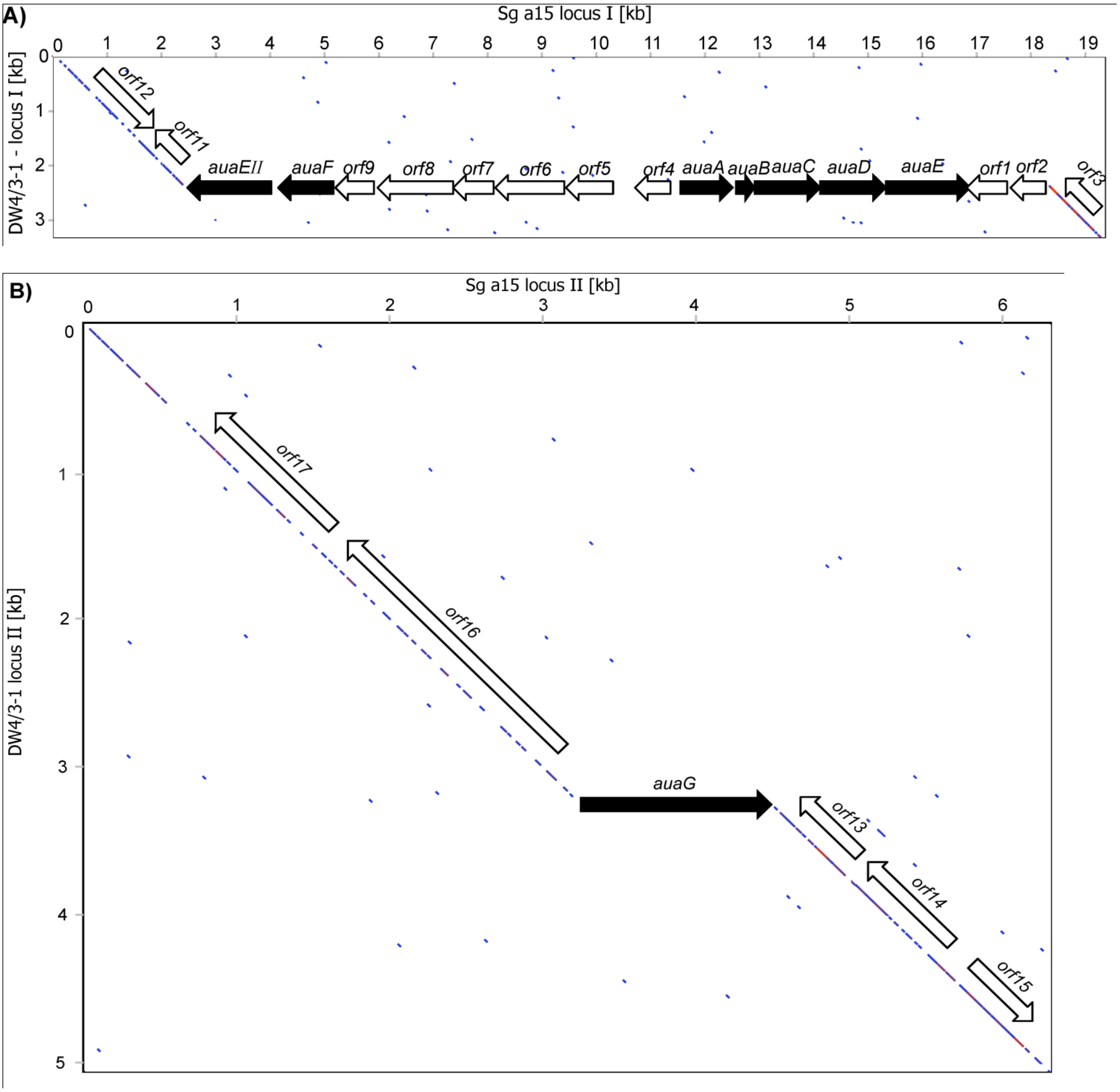
Verification of targeted gene inactivation mutants in Sg a15

Two of the PCR verification reactions involved one primer annealing in the genome either up- or downstream of the targeted gene and the second primer annealing in the vector backbone (TOPO in or TOPO out). Products for these reactions were specific for mutants whereas the wild type control must not give a product. The third PCR reaction was performed with the two primers up- and downstream of the targeted gene. In this case the wild type control gave a product whereas the mutants did not since the inserted vector backbone is too large to be amplified under the chosen conditions. Thereby the targeted inactivation of single genes could be verified from the pattern of PCR products. To this purpose the following primers were used: TOPO in (5´-CCT CTA GAT GCA TGC TCG AGC-3´) and TOPO out (5´-TTG GTA CCG AGC TCG GAT CC-3´), *auaF* ko ver up (5´-TCA CGA TGC CGA AGA ACG AC-3´) and *auaF* ko ver down (5´-TCA ACC TCA TCC CGG TCC TC-3´), *auaG* ko ver up (5´-ACC TCA GGA AAC CTC AAA GGG-3´) and *auaG* ko ver down (5´-CAT CCA CGC GAT CGC TGG ACA-3´), *auaH* ko ver up (5´-TCG CTC CAT GGC TCG CTG GG-3´) and *auaH* ko ver down (5´-GAT CCG CGAT GTG TGC ATG A-3´), *auaJ* ko ver up (5´-GAC TTT CGG ATC CAC AAG ATG-3´) and *auaJ* ko ver down (5´-GGA AAG GCG TTC CGA GCA CCG-3´), *auaI/J* ko ver up (5´-GGA AAA GGA CAT CGC AAC CTG-3´) and *auaI/J* ko ver down (5´-CGA CTT GAG GAG CGT CTC GTA-3´), *orf6* ko ver up (5´-GGA GGG TCA TCA AGC TGA GGA-3´) and *orf6* ko ver down (5´-CGT CAT TCC CTT GAC CTA CCT C-3´), *orf11* ko ver up (5´-CCT ACG GTT TCG AGC ACG ATG-3´) and *orf11* ko ver down (5´-GGA AGT CGA GAG CAC CCT CAT-3´), *orf16* ko ver up (5´-GGT GAA GGT TCA GCC TTC CC-3´) and *orf16* ko ver down (5´-GTA GTC GGA GAG ATC GGA GAG G-3´), *orf20* ko ver up (5´-TCT CAC CCA AGT GAT CAA AAT C-3´) and *orf20* ko ver down (5´-GAC GTG TGA CCC ACC AGC AAG-3´).

Table S1. Proteins encoded in the loci I-III of *Stigmatella aurantiaca* Sg a15 discussed in the main text with their closest homologs.

Protein (*gene*)	Size (aa / bp)	Closest homologue, origin (identity [%], similarity [%]; Access no.)
AuaA (*auaA*)	327 / 981	UbiA prenyltransferase, *Chloroherpeton thalassium* ATCC 35110 (28, 50; YP_001997422)
AuaB (*auaB*)	85 / 255	acyl carrier protein, *Aeromonas salmonicida subsp. salmonicida* A449 (31, 62; YP_001140793)
AuaC (*auaC*)	407 / 1221	β-ketoacyl-ACP synthase II, *Candidatus Kuenenia stuttgartiensis* (43, 63; CAJ74368)
AuaD (*auaD*)	408 / 1224	β-ketoacyl-ACP synthase II, *Candidatus Kuenenia stuttgartiensies* (33, 50; CAJ74369)
AuaE (*auaE*)	498 / 1494	benzoate-CoA ligase family, *Rhodopseudomonas palustris* BisB5 (38, 53; YP_568017)
AuaEII (*auaEII*)	517 / 1551	benzoate-CoA ligase family, *Geobacter metallireducens* GS-15 (48, 65; YP_385097)

AuaF (auaF)	340 / 1020	Rieske oxygenase AmbP, *Sorangium cellulosum* So ce10 (38, 54; ABK32265)
AuaG (auaG)	383/1152	FAD-dependent oxidoreductase, *Myxococcus xanthus* DK 1622 (53, 66; YP_631595)
AuaH (auaH)	334/1005	NAD dependent epimerase/dehydratase family protein, *Myxococcus xanthus* DK 1622 (57, 68; YP_633987)
AuaI (auaI)	129/390	putative steroid delta-isomerase, *Candidatus Nitrospira defluvii* (42, 55; YP_003797490)
AuaJ (auaJ)	454/1365	FAD dependent oxidoreductase, *Ktedonobacter racemifer* DSM 44963 (43, 59; ZP_06971993)
Orf1 (*orf1*)	225 / 675	carboxymethylenebutenolidase, *Rhodopseudomonas palustris* CGA009 (46, 57; NP_946545)
Orf2 (*orf2*)	221 / 663	hypothetical protein, *Flavobacterium johnsoniae* UW101 (42, 59; YP_001195226)
Orf3 (*orf3*)	267 / 801	putative secreted hydrolase, *Stigmatella aurantiaca* DW4/3-1 (97, 97; ZP_01466699)
Orf4 (*orf4*)	208 / 624	transcriptional regulator, *Bradyrhizobium japonicum* USDA 110 (42, 63; NP_768894)
Orf5 (*orf5*)	297 / 891	ABC-type transport system, *Halothermothrix orenii* H 168 (39, 57; YP_002510063)
Orf6 (*ofr6*)	427 / 1281	ABC-type lipoprotein release transporter, permease component, *Desulfobacterium autotrophicum* HRM2 (29, 51; YP002605141)
Orf7 (*orf7*)	242 / 726	outer membrane lipoprotein-sorting protein LolF, *Desulfobacterium autotrophicum* HRM2 (28, 47; YP002603286)
Orf8 (*orf8*)	478 / 1434	hypothetical protein, *Desulfovibrio sp.* ND132 (23, 39; ZP_08110832)
Orf9 (*orf9*)	242 / 726	ABC transporter ATP-binding protein, *Myxococcus xanthus* DK 1622 (54, 70; YP_630650)
Orf10 (*orf10*)	340 / 1020	Rieske [2Fe-2S] domain protein, *Stigmatella aurantiaca* DW4/3-1 (34, 53; ZP_01461542)
Orf11 (*orf11*)	268 / 804	alpha/beta superfamily hydrolase, *Stigmatella aurantiaca* DW4/3-1 (82, 90; YP_003956397)
Orf12 (*orf12*)	498 / 1494	hypothetical protein STAUR_6814, *Stigmatella aurantiaca* DW4/3-1 (78, 87; YP003956398)
Orf13 (orf13)	173/522	ABC transporter ATP-binding protein, *Stigmatella aurantiaca* DW4/3-1 (96, 98; ZP_01464426)
Orf14 (orf14)	175/528	kinase associated protein phosphatase, *Stigmatella aurantiaca* DW4/3-1 (95, 98; ZP_01464423)
Orf15 (orf15)	169/510	transcription elongation factor GreA, *Stigmatella aurantiaca* DW4/3-1 (95, 96; ZP_01464405)
Orf16 (orf16)	616/1851	hypothetical protein, *Stigmatella aurantiaca* DW4/3-1 (78, 86; YP_003953554)
Orf17 (orf17)	345/1038	fatty acid desaturase domain protein, *Stigmatella aurantiaca* DW4/3-1 (85, 91; ZP_01464414)
Orf18 (orf18)	1291/3876	SNF2/helicase domain protein, *Stigmatella aurantiaca* DW4/3-1 (87, 92; YP_003954938)
Orf19 (orf19)	201/606	TetR family transcriptional regulator, *Rhodopseudomonas palustris* DX-1 (42, 56; YP004108268)
Orf20 (orf20)	292/879	L-lactate dehydrogenase, *Stigmatella aurantiaca* DW4/3-1 (93, 97; ZP_01465399)
Orf21 (orf21)	395/1188	beta-lactamase, *Stigmatella aurantiaca* DW4/3-1 (57, 74; ZP_01463978)


A)
Sg a15 locus I [kb]
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
DW4/3-1 - locus I [kb]
0
1
2
3
orf12
orf11
auaEII
auaF
orf9
orf8
orf7
orf6
orf5
orf4
auaA
auaB
auaC
auaD
auaE
orf1
orf2
orf3
B)
Sg a15 locus II [kb]
0
1
2
3
4
5
6
DW4/3-1 locus II [kb]
0
1
2
3
4
5
orf17
orf16
auaG
orf13
orf14
orf15

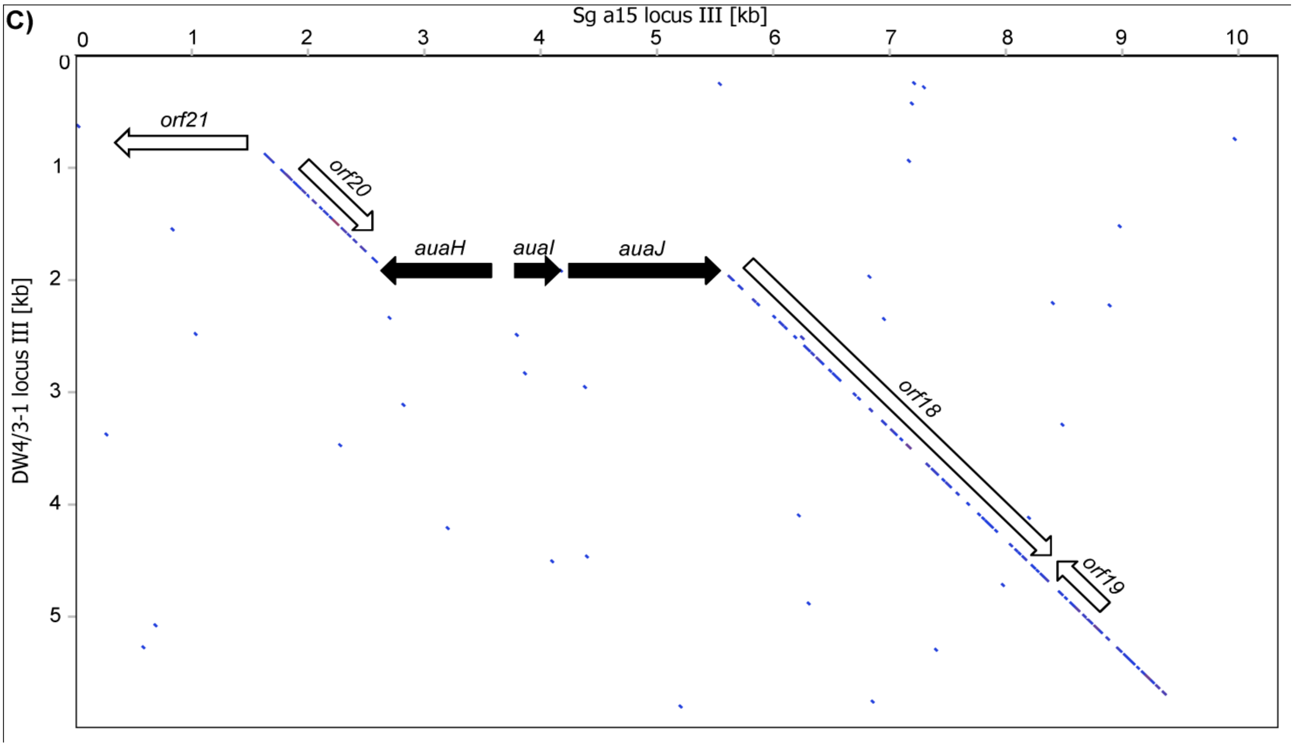
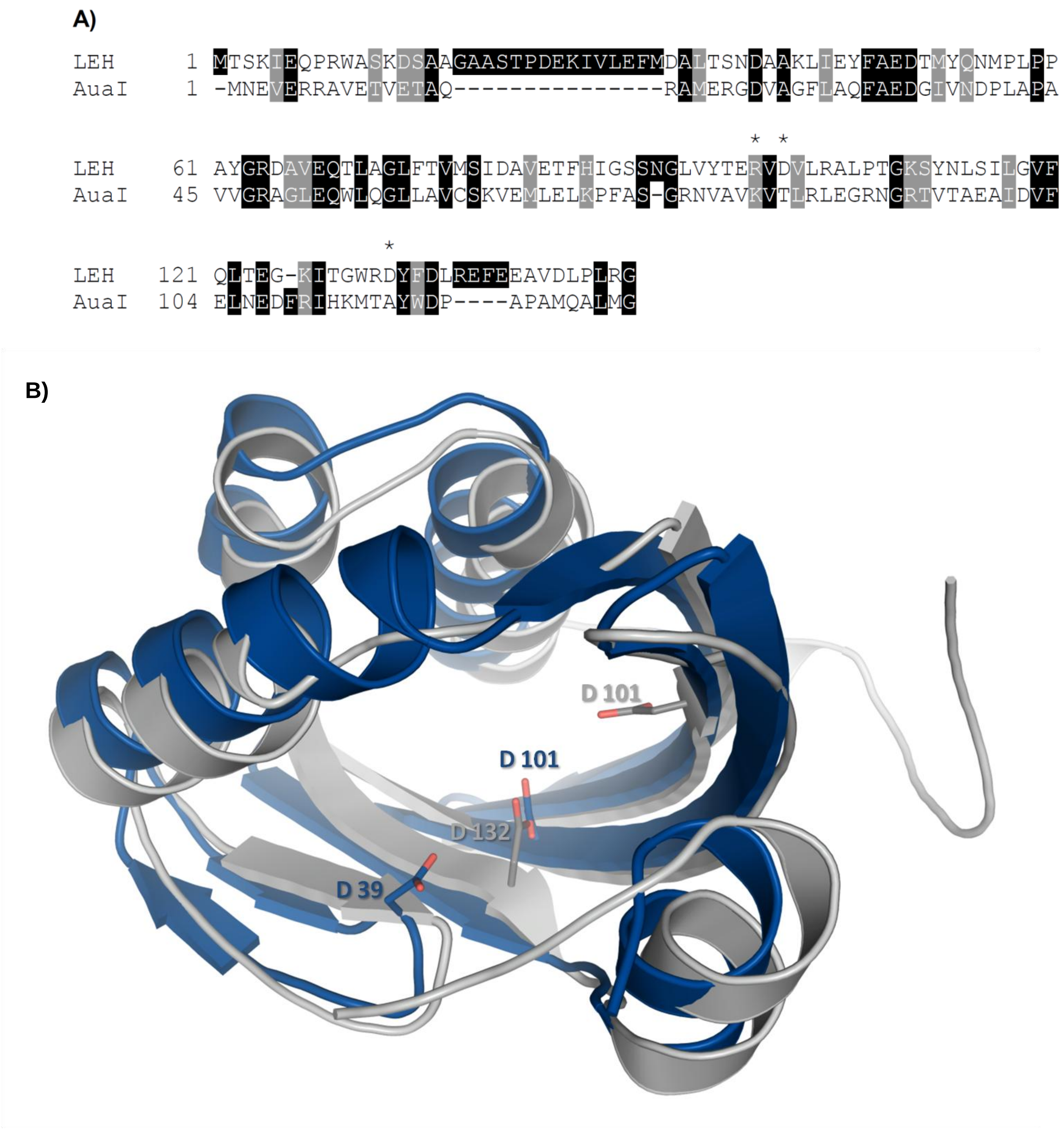


Fig. S1 Synteny comparison of **A)** locus I, **B)** locus II, and **C)** locus III in *S. aurantiaca* Sg a15 and *S. aurantiaca* DW4/3-1. Genes on the diagonal show synteny between both strains, whereas horizontally arranged genes are only found in Sg a15.



```
                                                 *   *
LEH      1 MTSKIEQPRWASKDSAAGAASTPDEKIVLEFMDALTSNDAAKLIEYFAEDTMYQNMPLPP
AuaI     1 -MNEVERRAVETVETAQ---------------RAMERGDVAGFLAQFAEDGIVNDPLAPA

                                                 * *
LEH     61 AYGRDAVEQTLAGLFTVMSIDAVETFHIGSSNGLVYTERVDVLRALPTGKSYNLSILGVF
AuaI    45 VVGRAGLEQWLQGLLAVCSKVEMLELKPFAS-GRNVAVKVTLRLEGRNGRTVTAEAIDVF

                        *
LEH    121 QLTEG-KITGWRDYFDLREFEEAVDLPLRG
AuaI   104 ELNEDFRIHKMTAYWDP----APAMQALMG
```

Fig. S2 Comparison between limonene-1,2-epoxide hydrolase (LEH) from *Rhodococcus erythropolis* DCL14 and AuaI. **A)** Sequence alignment of LEH and AuaI. Active site of LEH are marked by asterisks. **B)** Superimposition of the LEH crystal structure (depicted in gray, active site residues D101 and D132 identified to be involved in the one-step push-pull epoxide hydrolase mechanism established are shown as sticks) and the model computed for AuaI by PHYRE (depicted in blue, residues D39 and D101, that exhibit comparable distances to each other as D101 and D132 in the LEH structure, are shown as sticks). The crystal structure of LEH can be accessed at the Protein Data Bank with the identity code 1NWW.

Fig. S3 Structures of aurachins E-L and P