Supporting information for:

Efficient heterocyclisation by (di)terpene synthases

S. Mafu, K. C. Potter, M. L. Hillwig, S. Schulte, J. Criswell and R. J. Peters

Roy J. Carver Dept. Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, Iowa 50011, United States

Table of Contents

Experimental materials and methods

All reagents were purchased from Fischer Scientific unless noted otherwise.

Enzymatic analyses were carried out using a previously described modular metabolic engineering system,[1] with inclusion of a compatible plasmid (pIRS) to increase flux through the isoprenoid precursor pathway, which also has been previously described.[2] Briefly, this system drives metabolic flux towards the general diterpenoid precursor, GGPP (**1**), and enables co-expression of di-TPSs that can be readily recombined into DESTination cassettes via the Gateway cloning system (Invitrogen), which have been inserted into a set of compatible Duet vectors (Novagen).[2] The recombinant enzymes used in this study were expressed as previously described pseudomature constructs for AgAS,[3] PpCPS/KS,[4] FfCPS/KS,[5] a class II diterpene cyclase from *Nicotiana glutinosa* (NgCLS) that produces 8α-hydroxy-CPP (**2**),[6] the AtCPS:H263A mutant that produces 8β-hydroxy-*ent*-CPP (**5**),[7] the KS from the dicot *Arabidopsis thaliana* (AtKS),[8] the KS from the monocot *Oryza sativa*, rice (OsKS),[9] the KS from the rhizobacteria *Sinorhizobium fredii* (SfKS),[10] and the miltiradiene synthase from *Salvia miltiorrhiza* (SmMS).[11]

Mutants were generated by whole-plasmid PCR amplification with overlapping mutagenic primers of pENTR/SD/ d-TOPO (Invitrogen) clones, and verified by complete gene sequencing prior to transfer by directional recombination to expression vectors, pGG-DEST for the class II diterpene cyclases (and bifunctional enzymes), and pDEST17 for the class I diterpene synthases. The resulting constructs were heterologously expressed in the C41 OverExpress strain of *Escherichia coli* (Lucigen), which is optimal for such purpose as previously described.[12] Briefly, recombinant *E. coli* was grown in liquid NZY media (10 g casein, 10 g NaCl, 5 g yeast extract, 1 g $MgSO_4$ (anhydrous) in 1 L H_2O, with pH adjusted to 7.0 using HCl) at 37 °C to OD_{600} = 0.6, then shifted to 16 °C for an hour prior to induction with 0.5 mM IPTG, followed by fermentation at 16 °C, as previously described.[13]

For product analysis via metabolic engineering the relevant di-TPSs were co-expressed using compatible pGG-DEST and (where necessary) pDEST17 based constructs. At the time of induction, cultures were supplemented with phosphate buffer (pH 7.0) to 100 mM, sodium pyruvate to 50 mM, and $MgCl_2$ to 1 mM (final concentrations), as previously described.[2] After induction and fermentation for 3 days, products were extracted by addition of an equal volume of hexanes and gentle swirling, the organic

solvent is separated out and then dried under N_2, with the residue resuspended in fresh hexanes and analyzed by GC-MS.

Analysis of products by GC-MS was carried out as previously described.[14] Briefly, using a HP-5MS column (Agilent, 0.25 µm, 0.25ID, 30 m) on a 3900 GC with Saturn 2100T ion trap MS (Varian) in electron impact (70 eV) mode, with a 1.2 mL/min He flow rate, and the following oven temperature program: 50 °C for 3 min, 15 °C/min to 300 °C, hold 3 min. Samples (1 µL) underwent splitless injection at 250 °C. MS data from 90 – 500 *m/z* was collected from 12 min. post-injection until the end of the run.

To isolate sufficient amounts of enzymatically produced manoyl oxide (**3**) for confirmation by NMR spectral analysis, 2 x 1 L cultures of *E. coli* co-expressing NgCLS and SmMS were fermented for 3 days, and then extracted twice with an equal volume of hexanes. After extraction, the phases were separated in a separatory funnel, and the pooled hexanes dried by rotary evaporation. The resulting extract was redissolved in 10 mL of fresh hexanes and purified by passing over silica (4 g; 40-140 mesh), and **3** eluted with 5% ethyl acetate/hexanes (as determined by GC-MS analysis). This procedure was repeated four times and the resulting purified compound, with a final yield of ~5 mg, was dried under N_2 and redissolved in 0.5 mL $CDCl_3$. This sample was analyzed by NMR carried out with a Bruker Avance 700 spectrometer equipped with a 5-mm HCN cryogenic probe for ^{1}H and ^{13}C detection using standard experiments from the Bruker TopSpin version 1.4 software. One-dimensional ^{1}H spectra were acquired at 700 MHz, and one-dimensional ^{13}C spectra acquired at 175 MHz. Chemical shifts were referenced using known chloroform (^{13}C 77.23, ^{1}H 7.24 ppm) signals offset from TMS. Proton chemical shifts δ (ppm) (700 MHz, chloroform-d, 303.8 °K) were as follows: 0.813 (1H, td, 3.7, 13.1 Hz, H1a), 1.554 (1H, m, H1b), 1.393 (1H, m, H2a), 1.565 (1H, m, H2b), 1.114 (1H, td, 3.9, 13.0 Hz, H3a), 1.347 (1H, m, H3b), 0.917 (1H, dd, 2.4, 12.5 Hz, H5), 1.249 (1H, m, H6a), 1.630 (1H, m, H6b), 1.422 (1H, m, H7a), 1.803 (1H, dt, 3.2, 12.3, H7b), 1.312 (1H, dd, 4.3, 12.0, H9), 1.456 (1H, m, H11a), 1.552 (1H, m, H11b), 1.612 (1H, m, H12a), 1.741 (1H, dt, 5.5, 13.5, H12b), 5.854 (1H, dd, 10.8, 17.4 Hz, H14), 4.897 (1H, dd, 1.4, 10.8 Hz, H15*Z*), 5.123 (1H, dd, 1.4, 17.4, H16*E*), 1.252 (3H, s, H16), 1.272 (3H, s, H17), 0.836 (3H, s, H18), 0.775 (3H, s, H19), 0.765 (3H, s, H20). Carbon chemical shifts (173 MHz, chloroform-d, 303.8 °K) and assignments were as follows: δ (ppm) 39.24 (C1), 18.78 (C2), 42.34 (C3), 33.46 (C4), 56.65 (C5), 20.17 (C6), 43.49 (C7), 75.33 (C8), 55.85 (C9), 37.24 (C10), 15.52 (C11), 35.92 (C12), 73.44 (C13), 148.22 (C14), 110.49 (C15), 25.74 (C16), 28.78 (C17), 33.58 (C18), 21.57 (C19), and 15.63 (C20). These values match those previously reported for **3**.[15] After this data was obtained, it was reported that SmMS reacts with **2** to produce not only **3**, but also minor amounts of the C13 epimer **4**,[16] which could be found in our extracts as well. This enabled identification of the enantiomeric compound **6**, by comparison of the matching retention time and mass spectra (i.e., relative to **4**) in the non-chiral GC-MS analyses reported here, coupled to its derivation from **5**, the enantiomer to **2**, which is the precursor to **4**. It should be noted that all the tested *ent*-kaurene synthases (KSs) produce minor amounts of the enantiomer to **3**, *ent*-manoyl oxide, but in such small quantities that the corresponding peak is not readily detectable in the presented total ion chromatograms (e.g., this compound is produced at levels ≤ 3% of the main product, **6**). In addition, the peak labeled **9** in Fig. 5A contains small amounts of the dephosphorylated derivative of the known AgAS:H348D product labda-7,13*E*-dien-15-yl diphosphate (≤ 20% of the total peak area). Finally, although it has been reported that **3** and **4** will form spontaneously from **2′** (the dephosphorylated derivative of **2**, copal-8α,15-diol), that occurs non-specifically,[16b, 17] such that the stereoselective production of **3** by AgAS reported here is clearly enzymatically derived, with similar arguments applying to the even more selective production of **6** observed with the KSs.

References

1 A. Cyr, P. R. Wilderman, M. Determan and R. J. Peters, *J. Am. Chem. Soc.*, 2007, **129**, 6684.

2 D. Morrone, L. Lowry, M. K. Determan, D. M. Hershey, M. Xu and R. J. Peters, *Appl. Microbiol. Biotechnol.*, 2010, **85**, 1893.

3 R. J. Peters, J. E. Flory, R. Jetter, M. M. Ravn, H.-J. Lee, R. M. Coates and R. B. Croteau, *Biochemistry*, 2000, **39**, 15592.

4 K. Hayashi, H. Kawaide, M. Notomi, Y. Sakigi, A. Matsuo and H. Nozaki, *FEBS Lett.*, 2006, **580**, 6175.

5 B. Tudzynski, H. Kawaide and Y. Kamiya, *Curr. Gent.*, 1998, **34**, 234.

6 J. Criswell, K. Potter, F. Shephard, M. B. Beale and R. J. Peters, *Org. Lett.*, 2012, **14**, 5828.

7 K. Potter, J. Criswell and R. J. Peters, *Angew. Chem. Int. Ed.*, 2014, **53**, 7198.

8 S. Yamaguchi, T. Sun, H. Kawaide and Y. Kamiya, *Plant Physiol.*, 1998, **116**, 1271.

9 M. Xu, P. R. Wilderman, D. Morrone, J. Xu, A. Roy, M. Margis-Pinheiro, N. Upadhyaya, R. M. Coates and R. J. Peters, *Phytochemistry*, 2007, **68**, 312.

10 D. M. Hershey, X. Lu, J. Zi and R. J. Peters, *J. Bact.*, 2014, **196**, 100.

11 W. Gao, M. L. Hillwig, L. Huang, G. Cui, X. Wang, J. Kong, B. Yang and R. J. Peters, *Org. Lett.*, 2009, **11**, 5170.

12 R. Peters, J. Flory, R. Jetter, M. Ravn, H. Lee, R. Coates and R. Croteau, *Biochemistry*, 2000, **39**, 15592.

13 R. J. Peters, M. M. Ravn, R. M. Coates and R. B. Croteau, *J. Am. Chem. Soc.*, 2001, **123**, 8974.

14 D. Morrone, X. Chen, R. M. Coates and R. J. Peters, *Biochem. J.*, 2010, **431**, 337.

15 B. L. Buckwalter, I. R. Burfitt, A. A. Nagel, E. Wenkert and F. Näf, *Helv. Chim. Acta*, 1975, **58**, 1567.

16 (a) K. Brückner, D. Bozic, D. Manzano, D. Papaefthimiou, I. Pateraki, U. Scheler, A. Ferrer, R. C. de Vos, A. K. Kanellis and A. Tissier, *Phytochemistry*, 2014, **101**, 52; (b) I. Pateraki, J. Andersen-Ranberg, B. Hamberger, A. M. Heskes, H. J. Martens, P. Zerbe, S. S. Bach, B. L. Moller, J. Bohlmann and B. Hamberger, *Plant Physiol*, 2014, **164**, 1222.

17 P. Zerbe, B. Hamberger, M. M. Yuen, A. Chiang, H. K. Sandhu, L. L. Madilao, A. Nguyen, B. Hamberger, S. S. Bach and J. Bohlmann, *Plant Physiol*, 2013, **162**, 1073.