Supporting Information for:

A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase

Jared Criswell,[‡] Kevin Potter,[‡] Freya Shephard,[§,†] Michael H. Beale,[§] Reuben J. Peters*[,‡]

[‡]Department of Biochemistry, Biophysics, & Molecular Biology, Iowa State University, Ames, IA 50011, USA; [§]Plant Science Department, Rothamstead Research, Harpenden, Herts, AL5 2JQ, United Kingdom

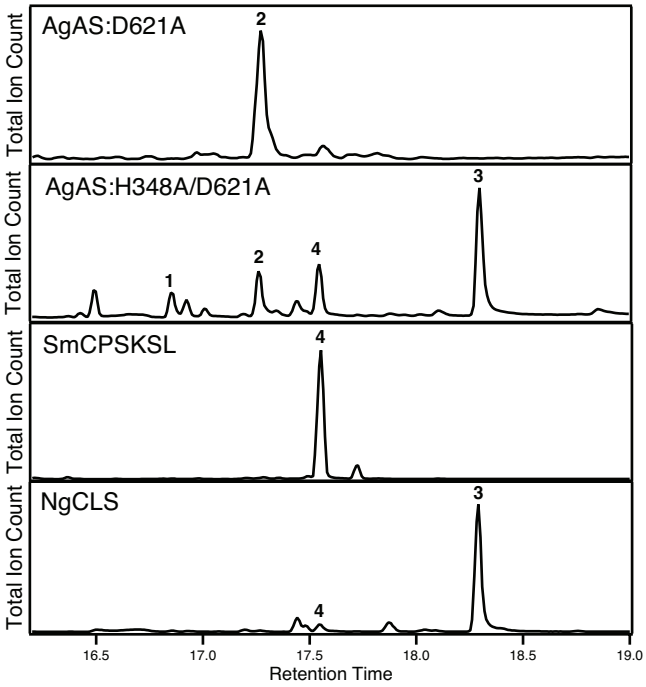

Figure S1: Effect of H348A mutation on product outcome of the class II cyclization reaction catalyzed by AgAS. GC-MS chromatograms of products from various AgAS constructs, NgCLS, or SmCPSKSL1 (as indicated). All enzymatic products are derived from expression in a previously described modular metabolic engineering system,[1] wherein the endogenous phosphatases lead to dephosphorylation, enabling extraction and detection of the parental diphosphate containing compounds as the resulting primary alcohols. Specifically, GGPP (**1**) is detected as geranylgeraniol, CPP (**2**) as copalol, 8α-hydroxy-CPP (**3**) as labda-13*E*-en-8α,15-diol, and labda-7,13*E*-dienyl diphosphate (**4**) as labda-7,13*E*-dien-15-ol.

Enzymatic analyses

The recombinant AgAS used here was the previously described pseudomature construct,[2] sub-cloned into pENTR/SD/D-TOPO (Invitrogen) as previously described.[3] Mutants were constructed by whole plasmid PCR amplification with overlapping mutagenic primers, all of which were verified by complete gene sequencing prior to transfer via directional recombination to expression vectors (pDEST17 and pGG-DEST). These were heterologously expressed in the C41 OverExpress strain of *Escherichia coli* (Lucigen), much as previously described.[2] Briefly, the recombinant *E. coli* were grown in liquid NZY media to 0.6 A_{600} at 37 °C, then shifted to 16 °C for an hour prior to induction with 0.5 mM IPTG, followed by fermentation at 16 °C. For in vitro assays, the enzymes were expressed as pDEST17 6xHis tagged constructs for ease of purification, which was accomplished much as previously described.[4] Briefly, cells from

overnight fermentation were harvested by centrifugation, lysed by gentle sonification in lysis buffer (50 mM Bis-Tris, pH 6.8, 150 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 10% glycerol), with the resulting lysate clarified by centrifugation (15,000*g* × 20 min. at 4 °C). The tagged enzymes were purified over Ni-NTA His-bind resin (Novagen), in batch mode, washing with 20 mM imidazole and elution by 250 mM imidazole in column buffer (50 mM Bis-Tris, pH 6.8, 1 mM DTT). Enzymatic assays for class II activity were carried out much as previously described.[5] Enzymatic products were investigated by expression in our previously described modular metabolic engineering system,[1] which induces production of GGPP in *E. coli* (using pGG-DEST based constructs), with further engineering increasing yields and enabling straightforward extraction of the products resulting from co-expressed diterpene cyclases,[6] as depicted in Figures 2 and S1. Briefly, the products resulting from 3 day fermentations were extracted with an equal volume of hexanes, dried under N_2, resuspended in fresh hexanes, and then filtered prior to analysis by gas chromatography with mass spectra detection (GC-MS), using a 3900 GC with Saturn 2100T ion trap MS (Varian) equipped with HP5-ms column (Agilent), as previously described.[7] To isolate sufficient amounts of labd-13*E*-en-8α,15-diol for conformation by NMR spectral analysis, 3 x 1 L cultures were fermented, extracted twice with an equal volume of hexanes, with the phases separated in a separatory funnel, and the pool hexanes dried by rotary evaporation. The resulting extract was redissolved in 10 mL fresh hexanes and purified using a Reveleris automated flash chromatography system, much as previously described.[8] The resulting fractions were analyzed by GC-MS, and that containing the targeted labd-13*E*-en-8α,15-diol dried under N_2, yielding ~ 3 mg, which was redissolved in 0.5 mL $CDCl_3$. This sample was analyzed by NMR, carried out much as previously described.[8] Briefly, using a Bruker Avance 700 spectrometer equipped with a 5-mm HCN cryogenic probe for ^1H and ^{13}C, structural analysis was performed using one-dimensional ^1H, and two-dimensional DQF-COSY, HSQC, HMQC, HMBC, and NOESY experiment spectra acquired at 700 MHz, and one-dimensional ^{13}C and DEPT135 spectra (174 MHz) using standard experiments from the Bruker TopSpin version 1.4 software. Correlations from the HMBC spectra were used to propose a partial structure, while resonance signals between protonated carbons were obtained from DQF-COSY data to complete the partial structure and assign proton chemical shifts. The structure was further verified using HSQC, and DEPT135 spectra to confirm assignments. Chemical shifts were referenced using known chloroform (^{13}C 77.23, ^1H 7.24 ppm) signals offset from TMS (Table S1), and compared to those previously reported.[9]

References

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

(2) Peters, R. J.; Flory, J. E.; Jetter, R.; Ravn, M. M.; Lee, H.-J.; Coates, R. M.; Croteau, R. B. *Biochemistry* **2000**, *39*, 15592.

(3) Prisic, S.; Peters, R. J. *Plant Physiol.* **2007**, *144*, 445.

(4) Zhou, K.; Peters, R. J. *Phytochemistry* **2009**, *70*, 366.

(5) Mann, F. M.; Prisic, S.; Davenport, E. K.; Determan, M. K.; Coates, R. M.; Peters, R. J. *J. Biol. Chem.* **2010**, *285*, 20558.

(6) Morrone, D.; Lowry, L.; Determan, M. K.; Hershey, D. M.; Xu, M.; Peters, R. J. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1893.

(7) Morrone, D.; Hillwig, M. L.; Mead, M. E.; Lowry, L.; Fulton, D. B.; Peters, R. J. *Biochem. J.* **2011**, *435*, 589.
(8) Wu, Y.; Hillwig, M. L.; Wang, Q.; Peters, R. J. *FEBS Lett.* **2011**, *585*, 3446.
(9) Koukoulitsa, C.; Zervou, M.; Demetzos, C.; Mavromoustakos, T. *Bioorg Med Chem* **2008**, *16*, 8237.

Table S1: ^{1}H and ^{13}C NMR data for labd-13*E*-en-8α,15-diol.

Position	δ_H	δ_C
1a **1b**	1.57 (1H, m) 0.89 (1H, dd, *J* = 11.8, 3.6 Hz)	40.3
2a **2b**	1.51 (1H, dt, *J* = 13.6, 3.6 Hz) 1.35 (1H, m)	19.0
3a **3b**	1.30 (1H, m) 1.07 (1H, dd, *J* = 13.6, 3.6 Hz)	42.6
4		33.8
5	0.85 (1H,dd, *J* = 12.2, 2.0 Hz)	56.7
6a **6b**	1.57 (1H, m) 1.19 (1H, m)	21.2
7a **7b**	1.78 (1H, dt, *J* = 12.3, 2.9 Hz) 1.30 (1H, m)	45.2
8		74.7
9	0.98 (1H, t, *J* = 3.6 Hz)	61.9
10		
11a **11b**	1.45 (1H, ddd, *J* = 9.9, 7.4, 3.6 Hz) 1.31 (1H,)	24.2
12	2.01 (2H, m)	43.5
13		141.7
14	5.36 (1H, t, *J* = 6.8 Hz)	123.7
15	4.07 (2H, m)	60.0
16	1.61 (3H, s)	17.1
17	1.05 (3H, s)	24.5
18	0.79 (3H, s)	34.0[a]
19	0.70 (3H, s)	22.1
20	0.71 (3H, s)	16.1

[a]This value differs slightly (34.0 versus 36.6) from that reported by ref. 9.

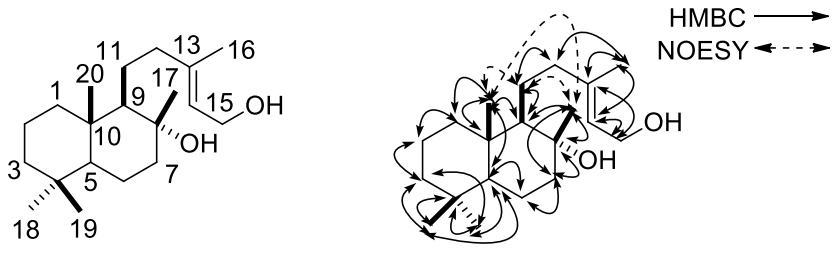

Figure S2: Numbering and selected HMBC and NOESY correlations for labd-13*E*-en-8α,15-diol.

Figure S3: ^{1}H 1D spectrum

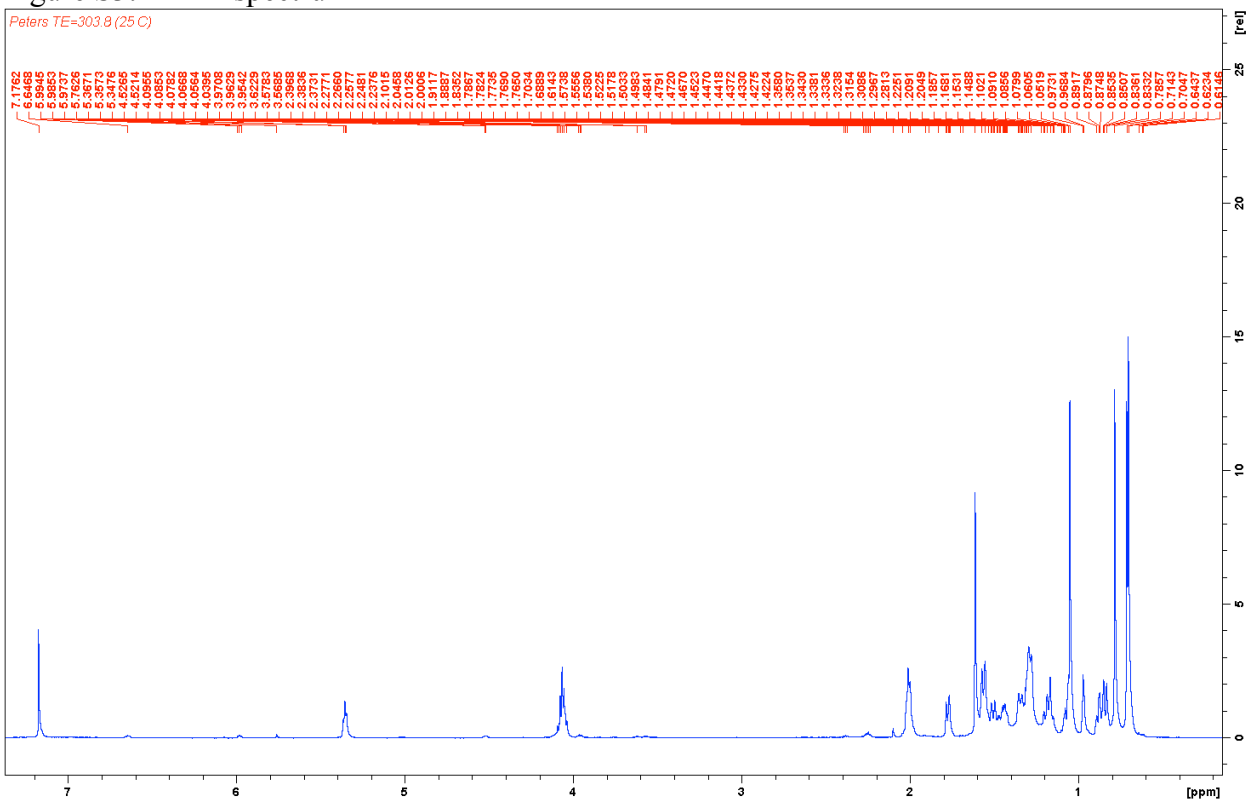

Figure S4: ^{13}C 1D spectrum

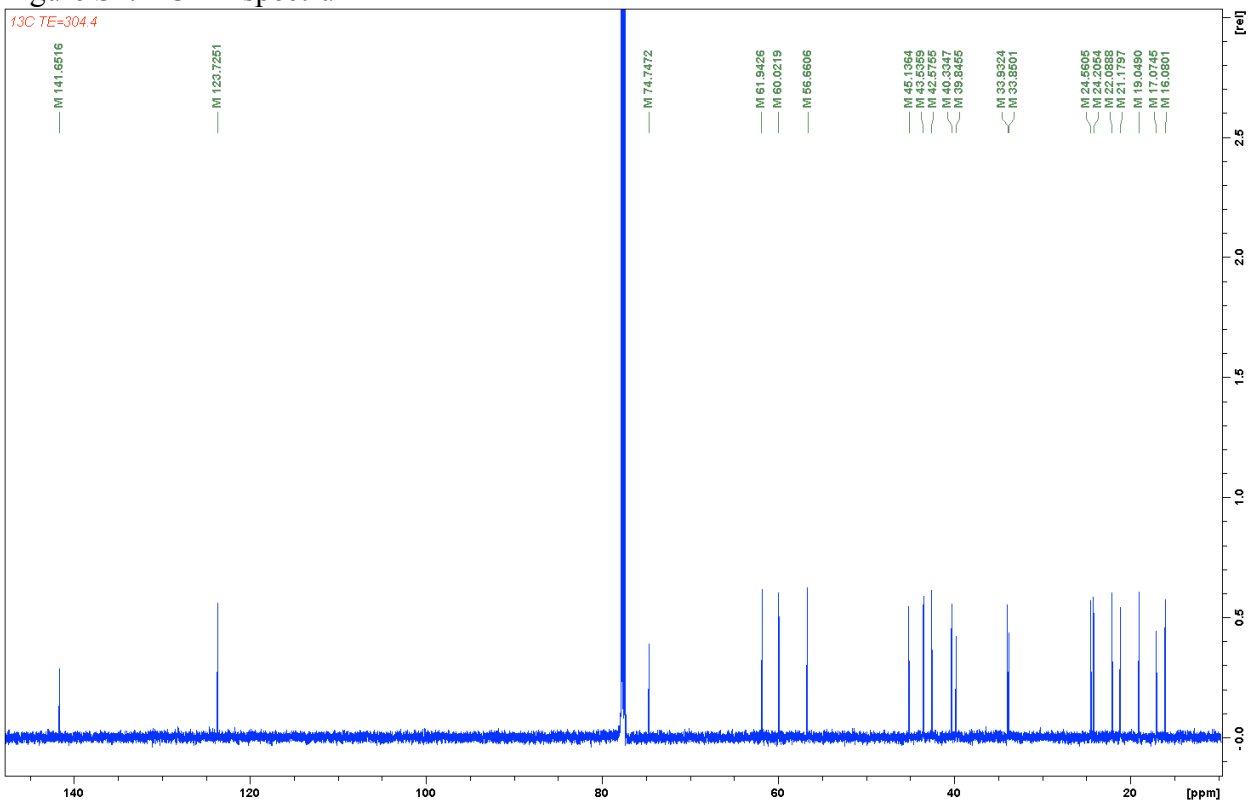

Figure S5: HMBC 2D spectrum

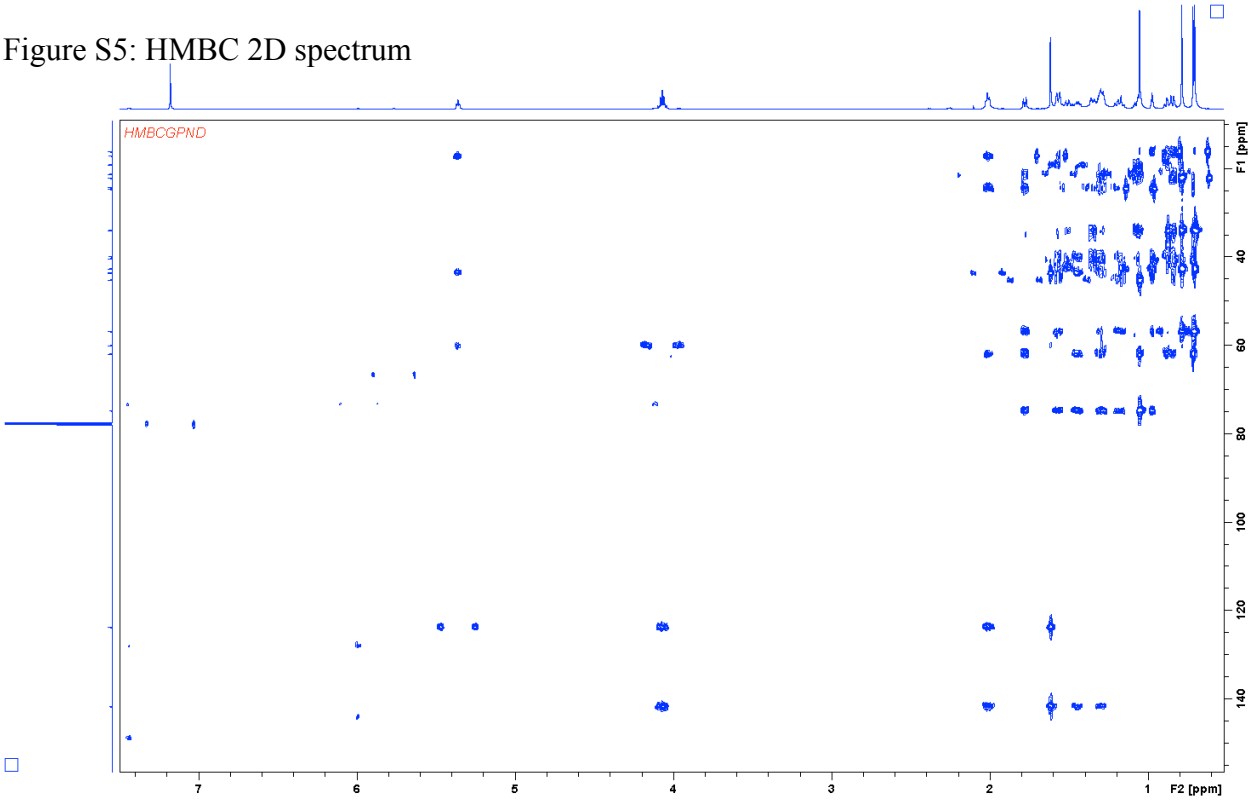

Figure S6: NOESY 2D spectrum