Supplementary Information for:

Labeling studies clarify the committed step in bacterial gibberellin biosynthesis

Ryan S. Nett,[†] Jeroen S. Dickschat,[‡] Reuben J. Peters[†,*]

[†]*Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, IA 50011, USA*

[‡]*Institut für Organische Chemie und Biochemie, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. E-mail:dickschat@uni-bonn.de; Tel: +49 228 735797*

**Corresponding author. E-mail: rjpeters@iastate.edu*

Table of Contents

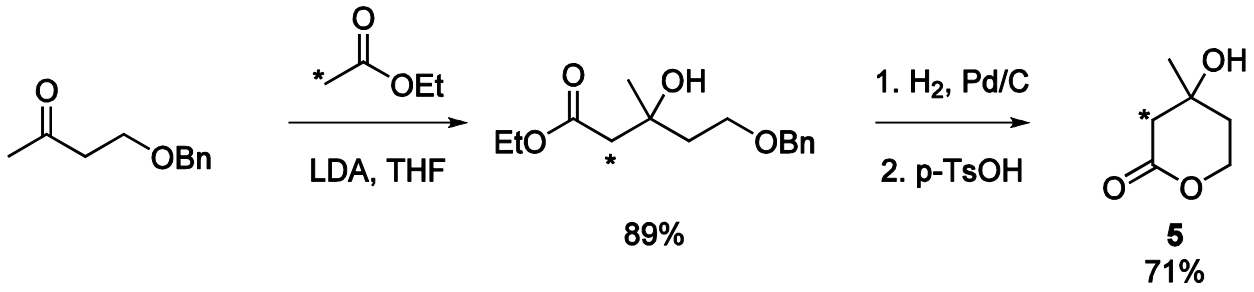
MATERIALS AND METHODS

General

All chemicals were purchased from Sigma-Aldrich and Fisher Scientific, with molecular biology tools purchased from Thermo Fisher Scientific. Unless noted otherwise, antibiotics were used at the following concentrations: tetracycline (Tc) 2 μg/mL, chloramphenicol (Cm) 20 μg/mL, carbenicillin (Cb) 25 μg/mL, spectinomycin (Sp) 25 μg/mL. Gas chromatography-mass spectrometry (GC-MS) was performed with a 3900 Saturn GC with Saturn 2100T ion trap MS (Varian) equipped with an HP-5MS column (Agilent). Relative quantification of purified compounds was determined using GC-FID (flame ionization detection) on an Agilent 6890N Network GC System equipped with an HP-1MS column (Agilent). High-performance liquid chromatography (HPLC) purification of compounds was carried out with an Agilent 1200 series HPLC equipped with a reverse-phase C-8 column (Kromasil® C8, 50 x 4.6 mm), autosampler, diode array UV detector, and fraction collector.

Synthesis of (2-^{13}C)mevalonolactone (5)

(2-^{13}C)Mevalonolactone (**5**) was synthesized as previously described[1–4]. Briefly, 4-(benzyloxy)butan-2-one was used in an aldol addition with the ester enolate of ethyl (2-^{13}C)acetate obtained by LDA deprotonation. The product ethyl 5-(benzyloxy)-3-hydroxy-3-methyl-(2-^{13}C)pentanoate was deprotected by catalytic hydrogenation, followed by lactonisation under acidic conditions (*p*-toluenesulfonic acid) to afford (2-^{13}C)-**5** in 63% overall yield. Details of the procedures and spectroscopic data of the products are given in reference 4.



Scheme S1. Synthesis of (2-^{13}C)mevalonolactone (**5**).

Preparation of ^{13}C-labeled *ent*-kaurenoic acid (1) substrate

Previous work has shown that C-2 of **5** is incorporated as C-1, C-7, C-12, and C-18 of the *ent*-kaurane backbone[4], and thus the ^{13}C-label from (2-^{13}C)-**5** would be located at those positions. ^{13}C-labeled *ent*-kaurenoic acid **1** was produced by metabolic engineering in two ways, either via sequential production of *ent*-kaurene **6** from **5** and then oxidation to **1**, or direct production of **1** from **5**. For the first method, (2-^{13}C)-**5** was fed into a previously described metabolic engineering system containing the enzymes necessary for biosynthesis of **6**.[5] Briefly, pMBI[6] (Tc^R; contains mevalonate pathway enzymes for increasing isoprenoid precursors), pGGeC (Cm^R; contains (*E,E,E*)-geranylgeranyl diphosphate synthase from *Abies grandis* and an *ent*-copalyl diphosphate synthase from *Zea maize*, for *ent*-copalyl diphosphate production) and pET-AtKS (Cb^R; contains *ent*-kaurene synthase from *Arabidopsis thaliana*, for production of **6**) were co-expressed in C41 OverExpress *E. coli* cells (Lucigen) with Tc, Cm and Cb selection. Cultures were

grown at 37 ˚C with 225 RPM shaking to mid-log phase (OD600 = 0.5-0.9), at which point IPTG (1 mM) riboflavin (1 mM), 5-aminolevulinic acid, $FeCl_3$ (0.25 mM), and (2-^{13}C)-**5** (2.5 mM) were added and the cultures were moved to 16 ˚C with shaking at 225 RPM for 3 days. After this incubation, cultures were extracted five times with an equal volume of hexanes. These organic fractions were pooled in a round bottom flask and dried in a rotary evaporator. The round bottom flask was extracted five times with 5 mL of hexanes, and these fractions were pooled and dried under a gentle stream of N_2. This resulted in production of ~150 μg of ^{13}C-labeled **6**, with the labeling confirmed via GC-MS and comparison to an unlabeled authentic standard produced via the same metabolic engineering system described here, but with unlabeled **5** fed into the system to produce unlabeled **6** (**Figure S1**). Mass spectra of ^{13}C-labeled substrates and products contained mass ions with a characteristic "pine-tree" effect – i.e. unlabeled samples have distinct *m/z* peaks with very minor sequential +1 peaks due to natural isotope abundance, while the labeled substrate exhibited ^{13}C isotope enrichment at four carbons, leading to higher abundances of the corresponding isotopic peaks. The ^{13}C-labeled **6** was concentrated to 1 mg/mL in a 1:1 solution of MeOH and DMSO and then fed to cultures of *E. coli* transformed with pCDF-AtKO/CPR1 (Sp^R; contains the *ent*-kaurenoic acid oxidase and a cytochrome P450 reductase, CPR1, from *Arabidopsis thaliana* for production of **1**)[7] with Sp selection. These cultures were grown as described above. For extraction of **1**, cultures were acidified with HCl to pH 3 to neutralize the C-4α (C-19) carboxylate, and extracted five times with an equal volume of ethyl acetate. These organic fractions were pooled in a round bottom flask and dried in a rotary evaporator. The round bottom flask was washed five times with 5 mL of ethyl acetate, and these resuspensions were pooled and dried under a gentle stream of N_2. When dry, this extract was dissolved in 1 mL hexanes and loaded onto a 1 mL packed silica column pre-eluted with hexanes for chromatographic purification. The column was washed sequentially with 1 mL volumes of hexane-ethyl acetate solutions starting with 100% hexanes and increasing the ethyl acetate proportion by 5% in each subsequent wash solution. Aliquots from the resulting fractions were methylated with diazomethane and checked for the presence of ^{13}C-labeled **1** via GC-MS, which showed elution of **1** in the 20-50% ethyl acetate fractions. The retention time and mass spectrum of ^{13}C-labeled **1** was verified by comparison to an unlabeled, methylated standard produced via the same metabolic engineering system described here, but with unlabeled **5** fed to produce unlabeled **1** (**Figure S2**). Fractions containing ^{13}C-labeled **1** were combined, dried under N_2, dissolved in MeOH, and filtered through a 0.2 μm regenerated cellulose filter (Whatman™). **1** was then further purified for NMR analysis through HPLC. For HPLC purification, the sample was injected in 100 μL aliquots at a flow rate of 0.5 mL/min (25 ˚C) with an initial mobile phase of 50% acetonitrile in water which was held for 2 minutes. For the next 5 minutes, the mobile phase was increased to 100% acetonitrile. This was held for 16 minutes, at which point the mobile phase was switched back to 50% acetonitrile in water over a time of 1 minute, and this was held for 3 minutes. With this method, a single peak of **1** eluted at 17.1 minutes, and a fraction was collected from 16.6 to 17.5 minutes to recover purified **1**. This method resulted in ~100 μg of purified ^{13}C-labeled **1**.

The second method for production of ^{13}C-labeled **1** involved co-expressing pMBI, pGGeC, pET-AtKS, and pCDF-AtKO/CPR1 in C41 OverExpress *E. coli* cells with Tc, Cm, Cb and Sp selection (**Scheme S3**). This allowed for the direct production of **1** from **5**. These cultures were grown, induced and supplemented, with subsequent extraction and purification of **1**, as described above, with a yield of ~100 μg **1**. Ultimately, such direct production appears to be more efficient, particularly in terms of effort. Between the two methods, a total of ~200 μg purified ^{13}C-labeled **1** was obtained, with ^{13}C-enrichment at carbons 1, 7, 12 and 18, as confirmed by ^{13}C-NMR analysis (**Figure S3**).

Production of ^{13}C-labeled GA_{12}-aldehyde (3)

^{13}C-labeled GA_{12}-aldehyde (**3**) was produced by feeding ^{13}C-labeled **1** to cultures heterologously expressing CYP114 and Fd_{GA} from *Sinorhizobium fredii* NGR234, which have been shown to convert **1** into 7β-hydroxy-*ent*-kaurenoic acid (**2**) and **3**.[8] Specifically, this was done using *Sinorhizobium meliloti* 1021 (which does not contain the GA biosynthetic operon) transformed with pstb-LAFR5-CYP114-Fd_{GA} (Tc^R).[8] This recombinant strain was grown in LB-MC (Luria-Bertani media with 2.5 mM MgSO4·7H2O and 2.5 mM CaCl2·2H2O) media supplemented with Tc (15 μg/mL) and streptomycin (Sm; 500 μg/mL) at 30 ˚C with shaking at 225 rpm to mid-log phase, at which point they were supplemented with riboflavin (1 mM), 5-aminolevulinic acid, $FeCl_3$ (0.25 mM), and ^{13}C-labeled **1** (10 μM). The cultures were allowed to continue growing at 30 ˚C with 225 rpm shaking for 3 days, at which point they were acidified to pH 3 with HCl, and extracted five times with an equal volume of ethyl acetate. Labeled **3** was purified in the same fashion as **1**. For silica purification, **3** eluted in 20-80% ethyl acetate in hexanes. HPLC was the same as described above, with the exception that 0.1% formic acid was added to the water and acetonitrile mobile phases to ensure that the carboxylic acid of **3** remains protonated, and thus elutes in one distinct peak/fraction. This resulted in elution of **3** at 14.3 minutes, and a fraction was collected from 13.5 to 14.75 minutes. Purified ^{13}C-labeled **3** was verified by GC-MS with comparison to an authentic standard (purchased from OlChemIm Ltd.) (**Figure S4**). A total of 50 μg of ^{13}C-labeled **3** was produced with this process.

^{13}C-NMR

Labeled (~100 μg) and unlabeled (~5 mg) **1**, and labeled (~50 μg) and unlabeled (~2 mg) **3**, were each dissolved in 500 μL of $CDCl_3$. ^{13}C-NMR spectra were obtained at 25 ˚C on a Bruker AVIII-800 spectrometer equipped with a 5-mm HCN cryogenic probe with analysis on TopSpin 3.2 and MestReNova 10.0.2 software. For all spectra, baseline corrections were performed in MestReNova 10.0.2 using the provided Whittaker Smoother. The chemical shifts were calculated from reference to one of the peaks for the $CDCl_3$ solvent (77.23 ppm). The obtained ^{13}C-NMR spectra were verified by comparison to previously published chemical shifts for **1**[9–11] and **3**.[12] The four ^{13}C enriched positions in the labeled compounds were verified by comparison to carbon shifts of the unlabeled authentic standards.

Incubation of 6β,7β-dihydroxy-*ent*-kaurenoic acid (4) and kaurenolides in cultures expressing CYP114-Fd_{GA}

To determine whether or not 6β,7β-dihydroxy-*ent*-kaurenoic acid (**4**) is a possible intermediate in the ring contraction reaction, we first analyzed samples in which *S. meliloti* 1021 cultures co-expressing CYP114 and Fd_{GA} were incubated with unlabeled **1**, much as previously described for the incubations in the "**Production of ^{13}C-labeled GA_{12}-aldehyde (3)**" section. After extraction, purification, and methylation, these samples were silylated using BSA+TMCS+TMSI (3:2:3; Sigma-Aldrich) with incubation at 90 ºC for 30 min, then dried down, resuspended in n-hexane, and analyzed with GC-MS. Close analysis revealed the presence of a small amount of **4**, as determined by comparison to an authentic standard. In an attempt to determine if **4** serves as an intermediate or as a by-product, cultures co-expressing CYP114 and Fd_{GA} were incubated with **4**, much as described above. However, no detectable turnover to **3** or any downstream GAs was detected within this system. Incubations with kaurenolide and 7β-hydroxykaurenolide were performed in the same fashion, and also had no detectable turnover. Purified 6β,7β-dihydroxy-*ent*-kaurenoic acid (**4**), kaurenolide, and 7β-hydroxykaurenolide were kindly provided by Peter Hedden (Rothamsted Research, U.K.).

REFERENCES

(1) Lawson, J. A.; Colwell, W. T.; DeGraw, J. I.; Peters, R. H.; Dehn, R. L.; Tanabe, M. *Synthesis (Stuttg)*. **1975**, *11*, 729–730.

(2) Tanabe, M.; Peters, R. H. *Org. Synth.* **1981**, *60*, 92.

(3) Dickschat, J. S.; Citron, C. A.; Brock, N. L.; Riclea, R.; Kuhz, H. *European J. Org. Chem.* **2011**, No. 18, 3339–3346.

(4) Citron, C. A.; Brock, N. L.; Tudzynski, B.; Dickschat, J. S. *Chem. Commun. (Camb).* **2014**, *50*, 5224–5226.

(5) Morrone, D.; Lowry, L.; Determan, M. K.; Hershey, D. M.; Xu, M.; Peters, R. J. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1893–1906.

(6) Martin, V. J. J.; Pitera, D. J.; Withers, S. T.; Newman, J. D.; Keasling, J. D. *Nat. Biotechnol.* **2003**, *21*, 796–802.

(7) Mafu, S.; Jia, M.; Zi, J.; Morrone, D.; Wu, Y.; Xu, M.; Hillwig, M. L.; Peters, R. J. *Proc. Natl. Acad. Sci.* **2016**, *113*, 5–10.

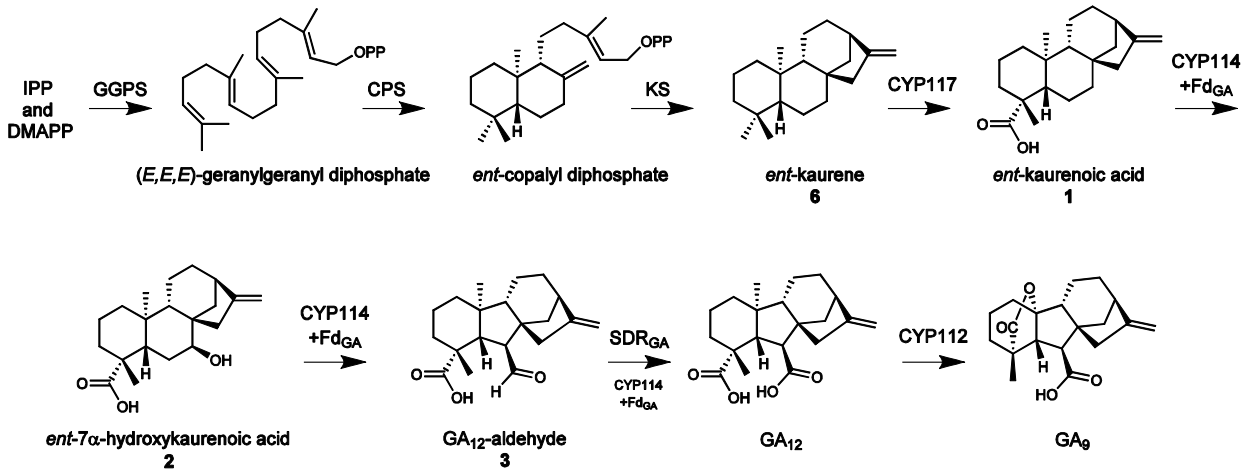
(8) Nett, R. S.; Montanares, M.; Marcassa, A.; Lu, X.; Nagel, R.; Charles, T. C.; Hedden, P.; Rojas, M. C.; Peters, R. J. *Nat. Chem. Biol.* [Online early access]. 10.1038/nchembio.2232. Published Online: November 14, 2016. http://dx.doi.org/10.1038/nchembio.2232 (accessed November 14, 2016).

(9) Yamasaki, K.; Kohda, H.; Kobayashi, T.; Kasai, R.; Tanaka, O. *Tetrahedron Lett.* **1976**, *17*, 1005–1008.

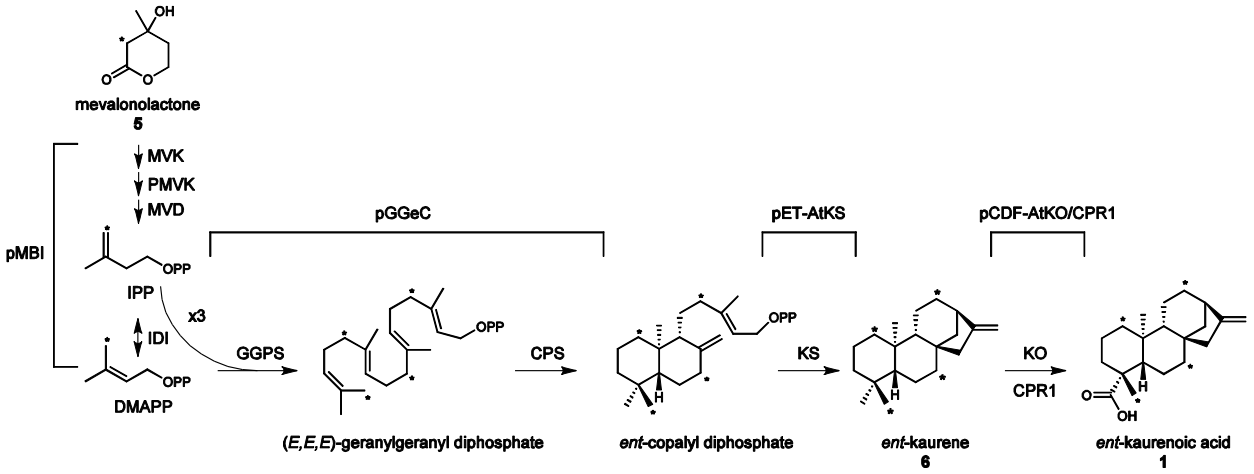
(10) Hutchison, M.; Lewer, P.; Macmillan, J. *J. Chem. Soc.* **1984**, No. 14, 2363–2366.

(11) Taveepanich, S.; Muangsin, N.; Saithong, S.; Pakawatchai, C.; Chaichit, N.; Roengsumran, S.; Petsom, A. *Nat. Prod. Res.* **2010**, *24*, 1050–1058.

(12) Lewer, P.; Macmillan, J. *Phytochemistry* **1984**, *23*, 2803–2811.



Scheme S2. GA biosynthesis in bacteria**.** IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GGPS, geranylgeranyl diphosphate synthase; CPS, *ent*-copalyl diphosphate synthase; KS, *ent*-kaurene synthase.



Scheme S3. Production of ^{13}C-labeled *ent*-kaurenoic acid (**1**) in an *E. coli* metabolic engineering system. Co-expression of several plasmids (containing the corresponding genes within each bracket) allows for a full biosynthetic pathway to **1** from mevalonolactone (**5**) in *E. coli*. Feeding labeled **5** to such recombinant cultures allows for the specified incorporation of ^{13}C. MVK, mevalonate kinase; PMVK, phosphomevalonate kinase; MVD, mevalonate diphosphate decarboxylase; IDI, isopentenyl diphosphate δ-isomerase; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GGPS, geranylgeranyl diphosphate synthase (from *Abies grandis*); CPS, *ent*-copalyl diphosphate synthase (from *Zea mays*); KS, *ent*-kaurene synthase (from *Arabidopsis thaliana*); KO, *ent*-kaurene oxidase (from *A. thaliana*); CPR1, cytochrome P450 reductase (from *A. thaliana*).

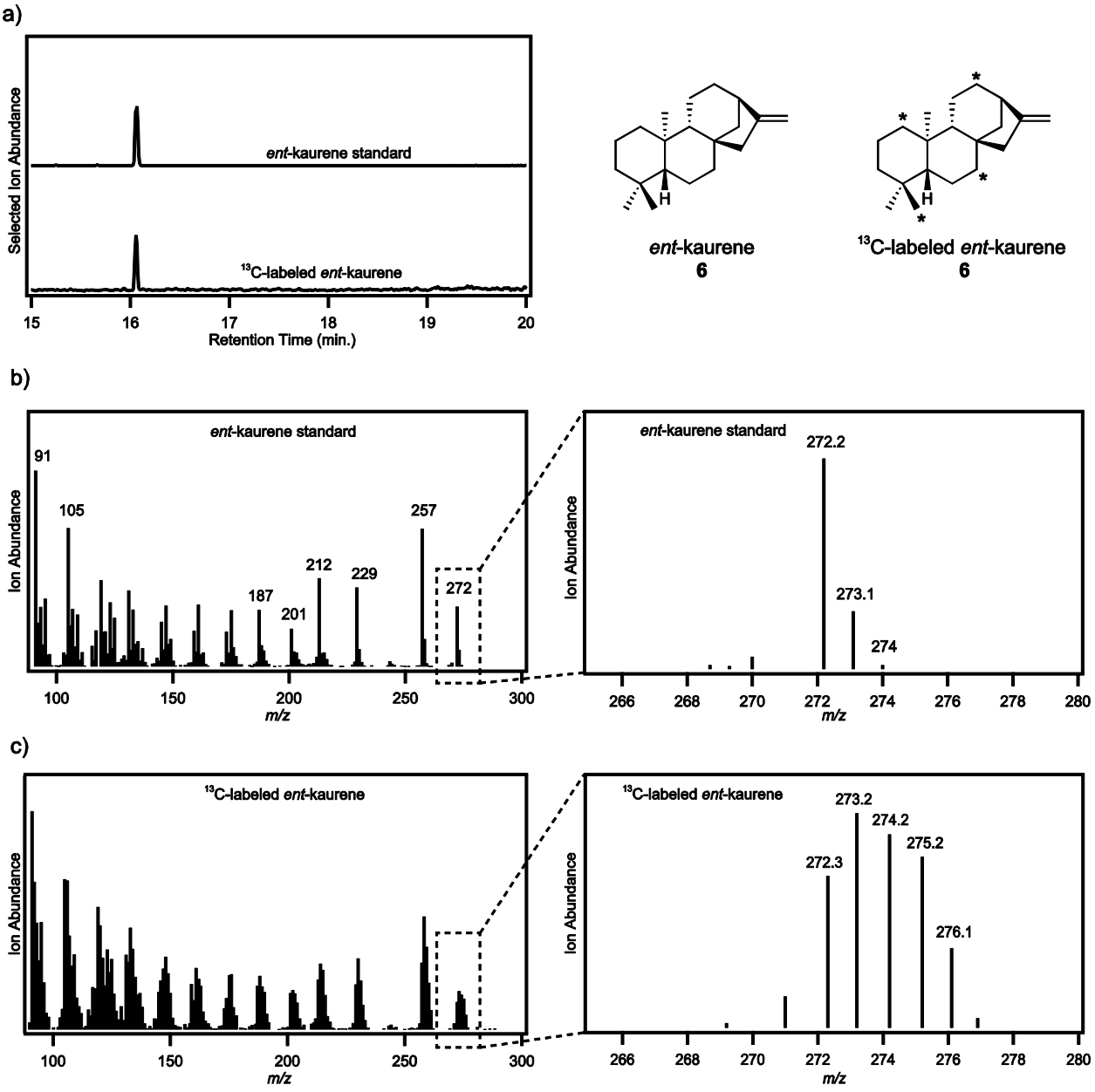


Figure S1. **Verification of ^{13}C-labeled *ent*-kaurene (6) via GC-MS**. a) Gas chromatography shows that purified ^{13}C-labeled **6** has the same retention time as an authentic standard of **6** (selected ion mass *m/z* = 257). The mass spectrum of b) the authentic standard is similar to that of c) ^{13}C-labeled **6**, with the notable exception of isotope-enriched peaks in ^{13}C-labeled **6**. This is evident at the molecular ion (M^+) peak of **6** (*m/z* = 272).

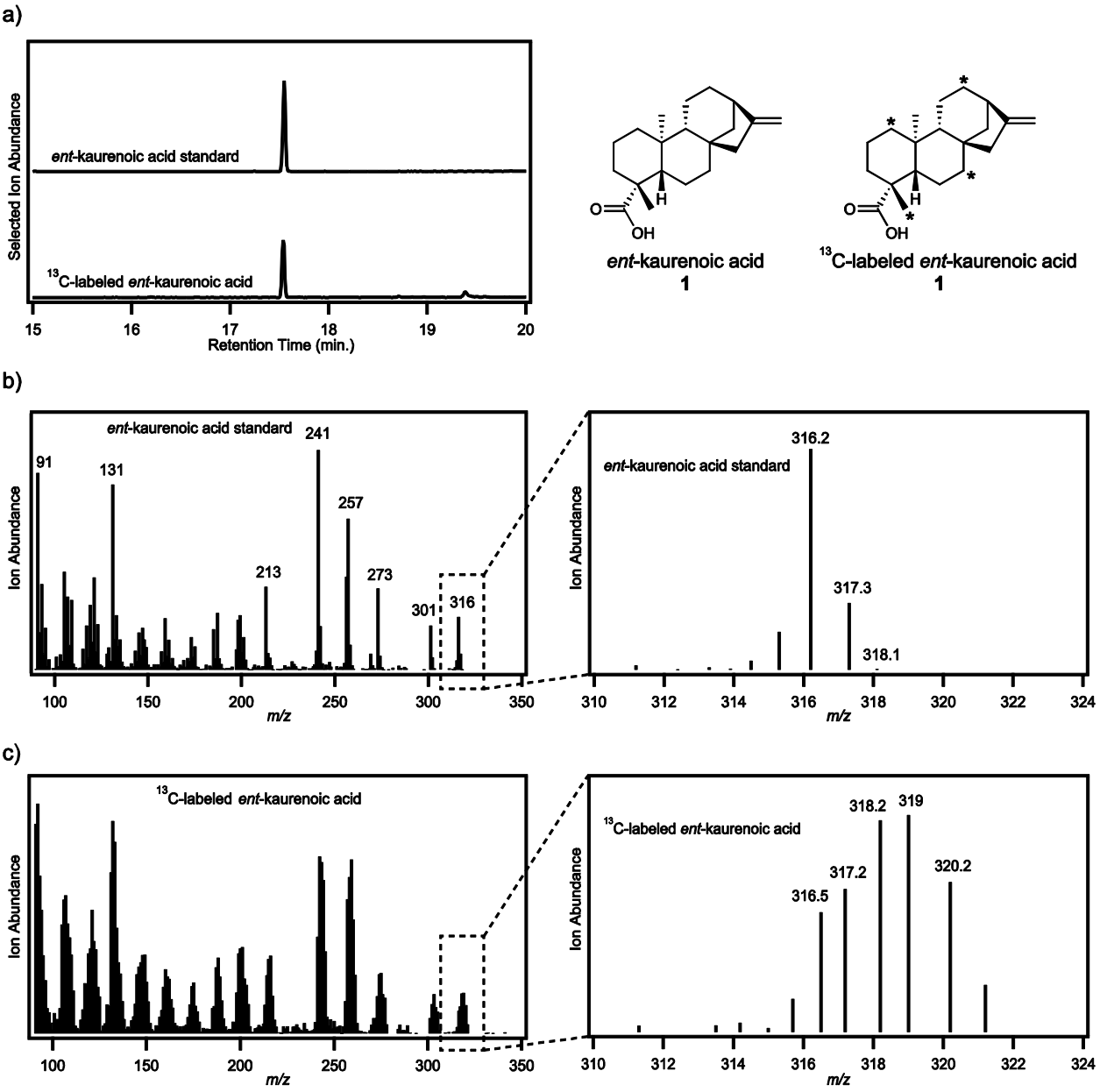


Figure S2. Verification of ^{13}C-labeled *ent*-kaurenoic acid (1) via GC-MS. a) Gas chromatography of methyl ester derivatives shows that purified ^{13}C-labeled **1** has the same retention time as an authentic standard of **1** (selected ion mass = *m/z* 241 for unlabeled **1** and *m/z* 244 for labeled **1**). The mass spectrum of b) the authentic standard is similar to that of c) ^{13}C-labeled **1**, with the notable exception of isotope-enriched peaks in ^{13}C-labeled **1**. This is evident at the molecular ion (M^+) peak of **1** (*m/z* = 316).

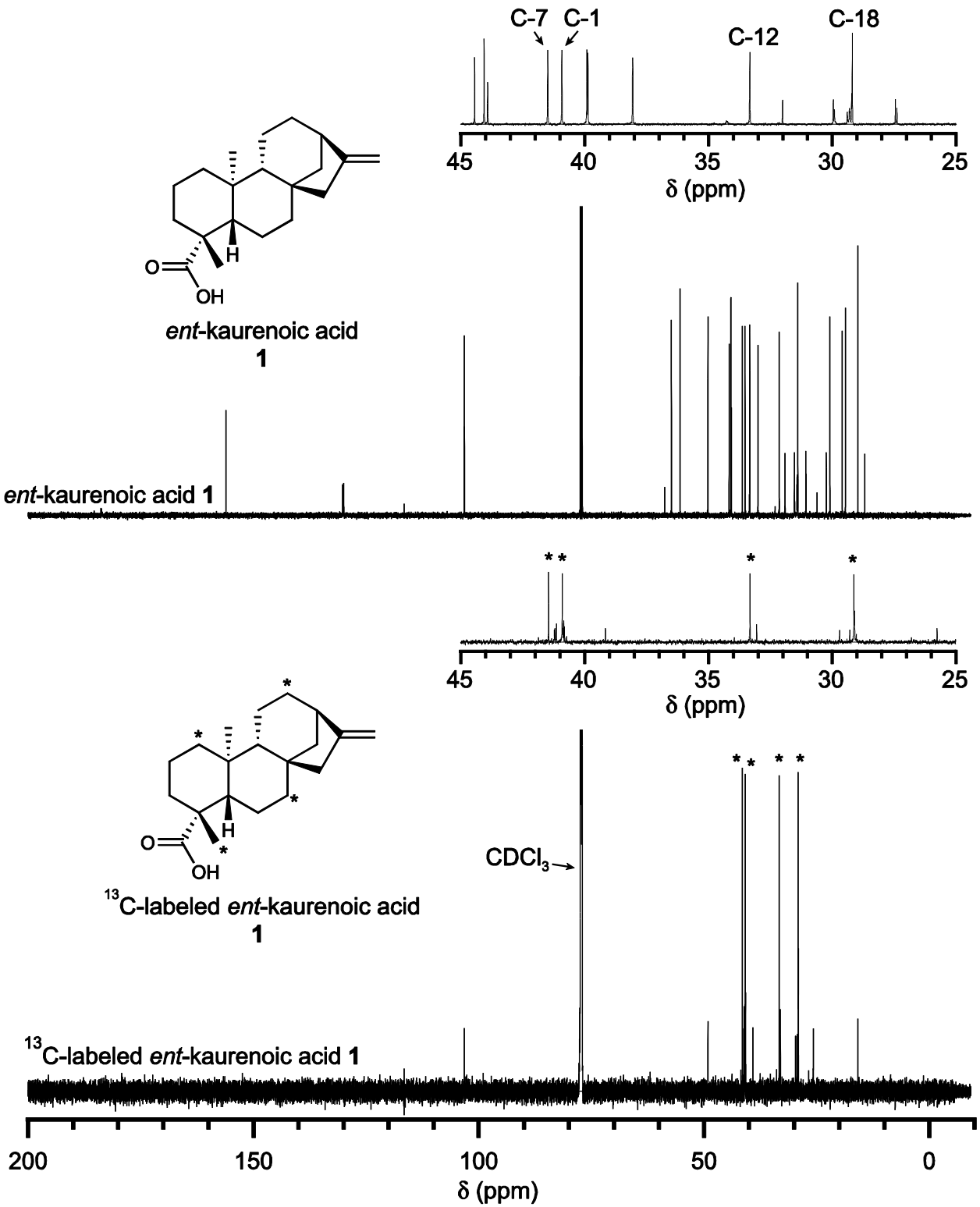


Figure S3. ^{13}C-NMR of labeled *ent*-kaurenoic acid (1) substrate. ^{13}C-labeled **1** was confirmed to have labels incorporated as carbons C-1, C-7, C-12, and C-18 through ^{13}C-NMR comparison to an authentic standard (800 MHz, $CDCl_3$ for both). ^{13}C-labeled carbons are indicated with an asterisk (*). Chemical shifts for both unlabeled (standard) and ^{13}C-labeled **1** are presented in Tables S1 and S2, respectively.

Table S1. ^{13}C-NMR chemical shifts for unlabeled *ent*-kaurenoic acid (1). Chemical shifts are referenced to $CDCl_3$ (77.23 ppm).

carbon number	δ (ppm)
19	183.8
16	156.1
17	103.2
5	57.2
9	55.3
15	49.2
8	44.4
13	44.1
4	43.9
7	41.5
1	40.9
10	39.9
14	39.9
3	38.1
12	33.3
18	29.2
6	22.0
2	19.3
11	18.6
20	15.8

Table S2. Enriched/Observed ^{13}C-NMR chemical shifts for ^{13}C-labeled *ent*-kaurenoic acid (1). Chemical shifts are referenced to $CDCl_3$ (77.23 ppm).

carbon number	δ (ppm)
7	41.5
1	40.9
12	33.3
18	29.1

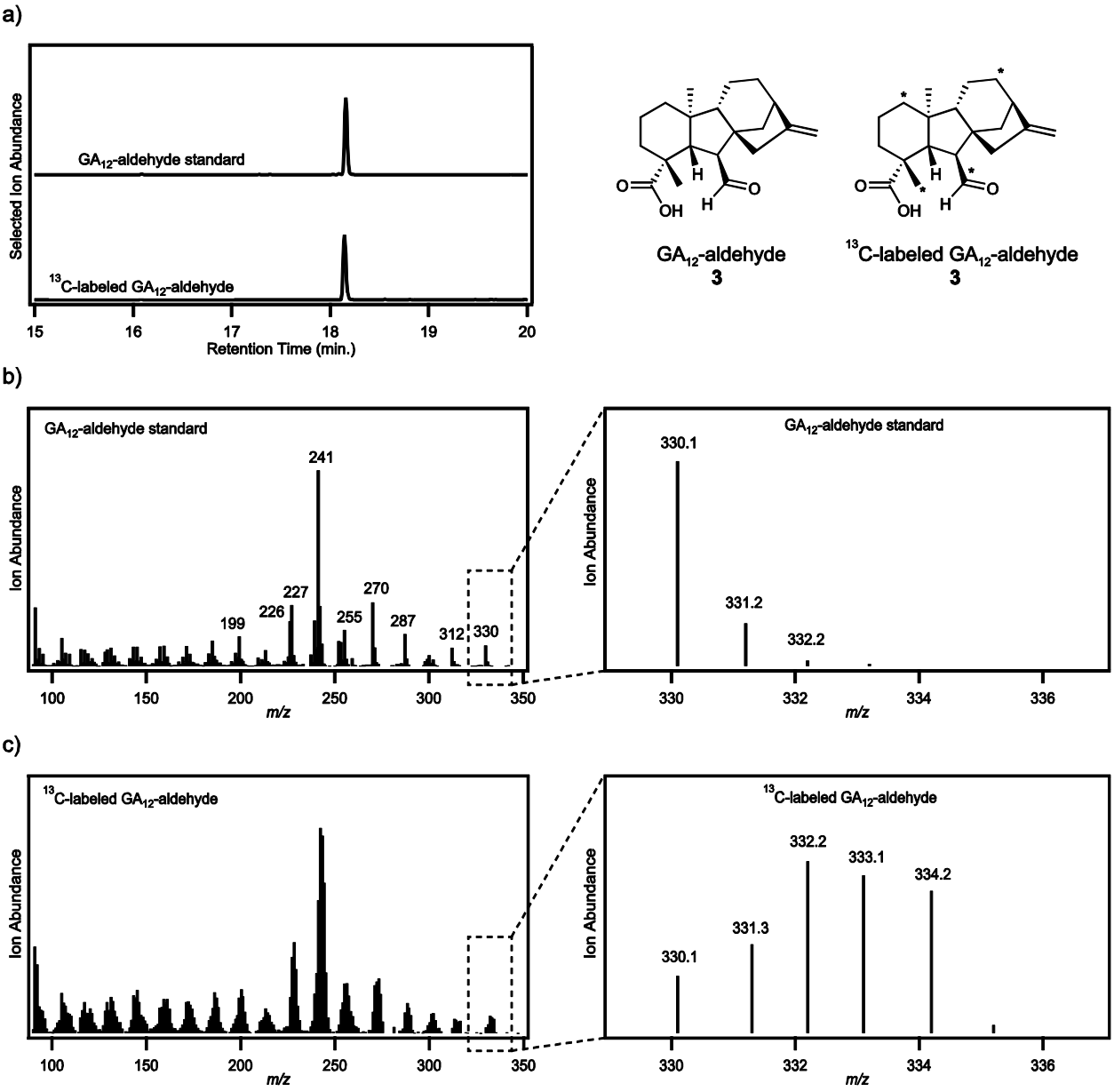


Figure S4. Verification of ^{13}C-labeled GA$_{12}$-aldehyde (3) via GC-MS. a) Gas chromatography of the methyl ester derivatives shows that purified ^{13}C-labeled **3** has the same retention time as an authentic standard of **3** (selected ion mass = *m/z* 241 for unlabeled **3** and *m/z* 242 for labeled **3**). The mass spectrum of b) the authentic standard is similar to that of c) ^{13}C-labeled **3**, with the notable exception of isotope-enriched peaks in ^{13}C-labeled **3**. This is evident at the molecular ion (M^+) peak of **3** (*m/z* = 330). Chemical shifts for both unlabeled (standard) and ^{13}C-labeled **3** are presented in Tables S3 and S4, respectively.

Table S3. ^{13}C-NMR chemical shifts for unlabeled GA_{12}-aldehyde (3). Chemical shifts are referenced to $CDCl_3$ (77.23 ppm). *Note that the identity of C-6 and C-7 are presented here based upon the results of this study.

carbon number	δ (ppm)
7*	205.9
19	181.3
16	157.2
17	106.7
5	58.2
6*	57.9
9	57.2
8	50.3
4	45.2
10	44.2
15	43.6
1	40.3
13	38.3
3	37.7
14	37.2
12	32.1
18	30.3
2	19.9
11	17.4
20	15.6

Table S4. Enriched/Observed ^{13}C-NMR chemical shifts for ^{13}C-labeled GA_{12}-aldehyde (3). Chemical shifts are referenced to $CDCl_3$ (77.23 ppm).

carbon number	δ (ppm)
7	205.9
1	40.3
12	32.1
18	30.3

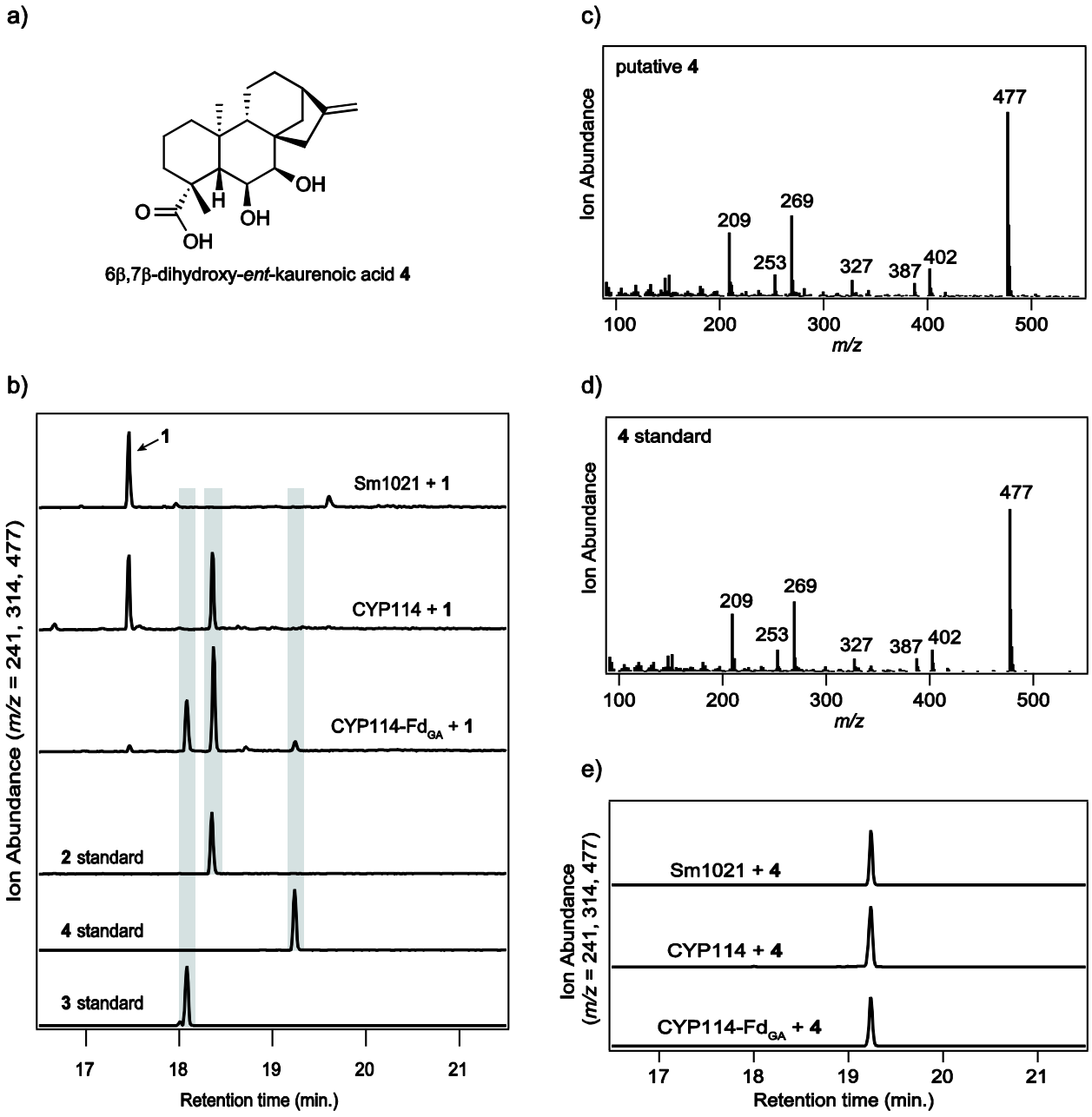


Figure S5. 6β,7β-dihydroxy-*ent*-kaurenoic acid (4) is produced in incubations of *ent*-kaurenoic acid (1) with cells co-expressing CYP114 and Fd$_{GA}$. a) Chemical structure of **4**. b) When incubated with cells co-expressing CYP114 and Fd$_{GA}$, **1** is converted into a small amount of **4**, along with the major products **2** and **3**, as shown here in gas chromatograms with comparison to authentic standards of **2**, **3**, and **4**. c) The mass spectrum of the putative **4** peak corresponds with that of d) an authentic standard. e) However, incubation of **4** in cells expressing either CYP114 or CYP14 + Fd$_{GA}$ does not result in any measured turnover to **3** or any downstream GAs. In this experiment, samples were derivatized to form methyl esters and trimethylsilyl-ethers for GC-MS analysis, and thus the chromatograms and mass spectra shown in this figure are representative of such derivatives.

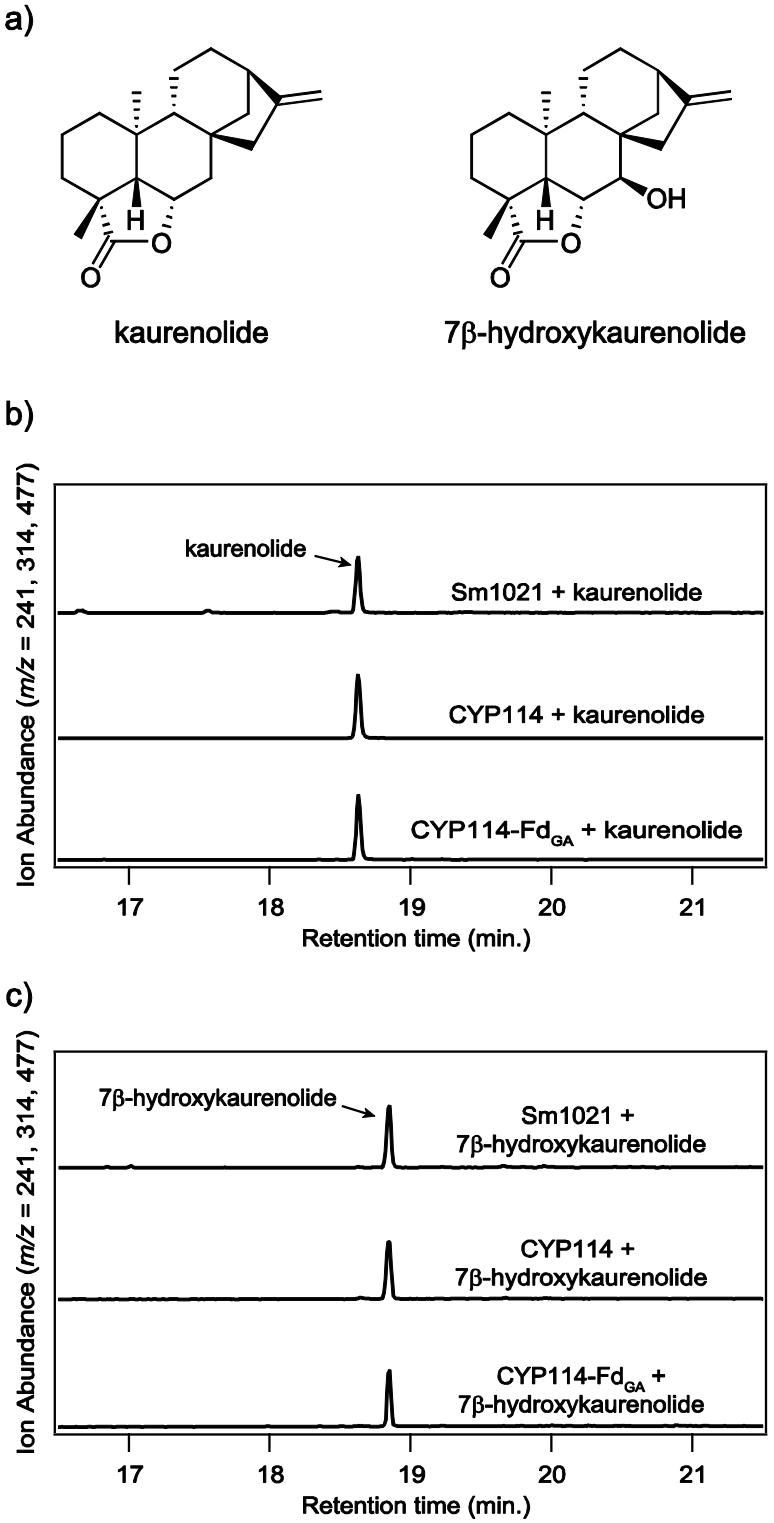


Figure S6. Incubation of kaurenolide and 7β-hydroxykaurenolide in cells co-expressing CYP114 and Fd$_{GA}$. a) Chemical structure of kaurenolide and 7β-hydroxykaurenolide. Neither incubations with b) kaurenolide or c) 7β-hydroxykaurenolide result in conversion to GA intermediates, as demonstrated here with gas chromatograms of the extracts, which were derivatized to form methyl esters and/or trimethylsilyl-ethers (e.g., 7β-hydroxykaurenolide is analyzed as the trimethylsilyl-ether derivative).