

Published in final edited form as:

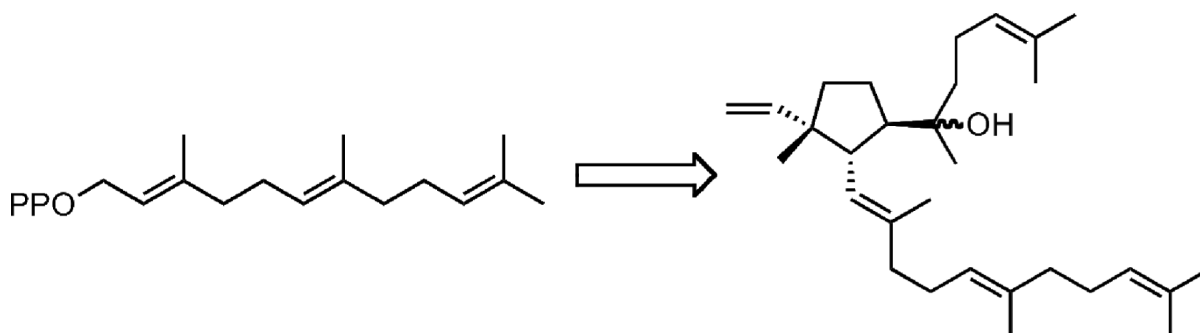
*J Org Chem.* 2009 October 2; 74(19): 7562–7565. doi:10.1021/jo9014547.

## Recombinant Squalene Synthase. Synthesis of Cyclopentyl Non-Head to Tail Triterpenes

 Jian-Jung Pan<sup>†</sup>, Tim S. Bugni<sup>‡</sup>, and C. Dale Poulter<sup>‡,\*</sup>

Department of Chemistry, University of Utah, 315 South 1400 East RM 2020, Salt Lake City, Utah 84112, and Department of Medicinal Chemistry, University of Utah, 30 South 2000 East RM 307, Salt Lake City, Utah, 84112

### Abstract



Incubation of farnesyl diphosphate with recombinant yeast squalene synthase in the absence of NADPH gives a mixture of triterpene hydrocarbons and alcohols, including botryococcene-like compounds with 1'-3 linkages between the farnesyl units. One of these molecules has an unusual cyclopentane structure similar to those recently reported in plant extracts and lakebed sediments.

Race B of the colonial green alga *Botryococcus braunii* produces large quantities of the triterpene botryococcene and related methylated derivatives with 1'-3 linkages between two farnesyl moieties.<sup>1</sup> Incorporation experiments with labeled farnesol suggest that botryococcene is derived from farnesyl diphosphate (FPP) by condensation to presqualene diphosphate (PSPP) followed by a carbocationic rearrangement of the cyclopropylcarbinyl diphosphate and reduction of the intermediate by NADPH to give the 1'-3 linked triterpene.<sup>2</sup> Thus, the reactions for biosynthesis of botryococcene appear to be closely related those of squalene.<sup>3,4</sup>

Under normal conditions, incubation of squalene synthase (SQase) with FPP and NADPH gives squalene as the only detectable product. However, if the incubation is conducted in the absence of the cofactor, the enzyme synthesizes PSPP from FPP and then catalyzes its solvolysis to give a mixture of six triterpene hydrocarbons and alcohols.<sup>3,5</sup> The major products are the 1'-1 triterpenes hydroxysqualene (HSQ), where the hydrogen in squalene derived from NADPH is replaced by a hydroxyl group derived from water, its allylic isomer isoHSQ, and cis-dehydrosqualene (DSQ), the C<sub>30</sub> analog of cis-phytoene. However, approximately 20% of the products are the 1'-3 triterpenes hydroxybotryococcene (HBO) and its allylic isomer

\*To whom correspondence should be addressed. Tel: 801-581-6685. poulter@chemistry.utah.edu.

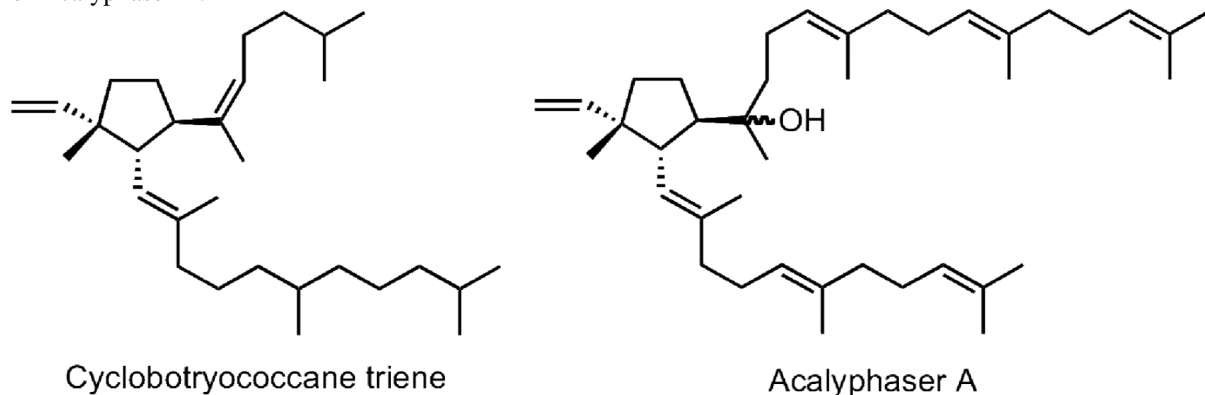
<sup>†</sup>Department of Chemistry, University of Utah.

<sup>‡</sup>Department of Medicinal Chemistry, University of Utah.

Supporting Information Available: General experimental procedures and materials; 1D <sup>1</sup>H NMR spectra; 2D <sup>1</sup>H-<sup>1</sup>H COSY, <sup>1</sup>H-<sup>13</sup>C HSQC, HMBC, and NOESY spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

isoHBO. The sixth product, formed in about 1% yield, was not identified. Thus, the catalytic machinery of squalene synthase is capable of synthesizing both 1'-1 and 1'-3 triterpenes.

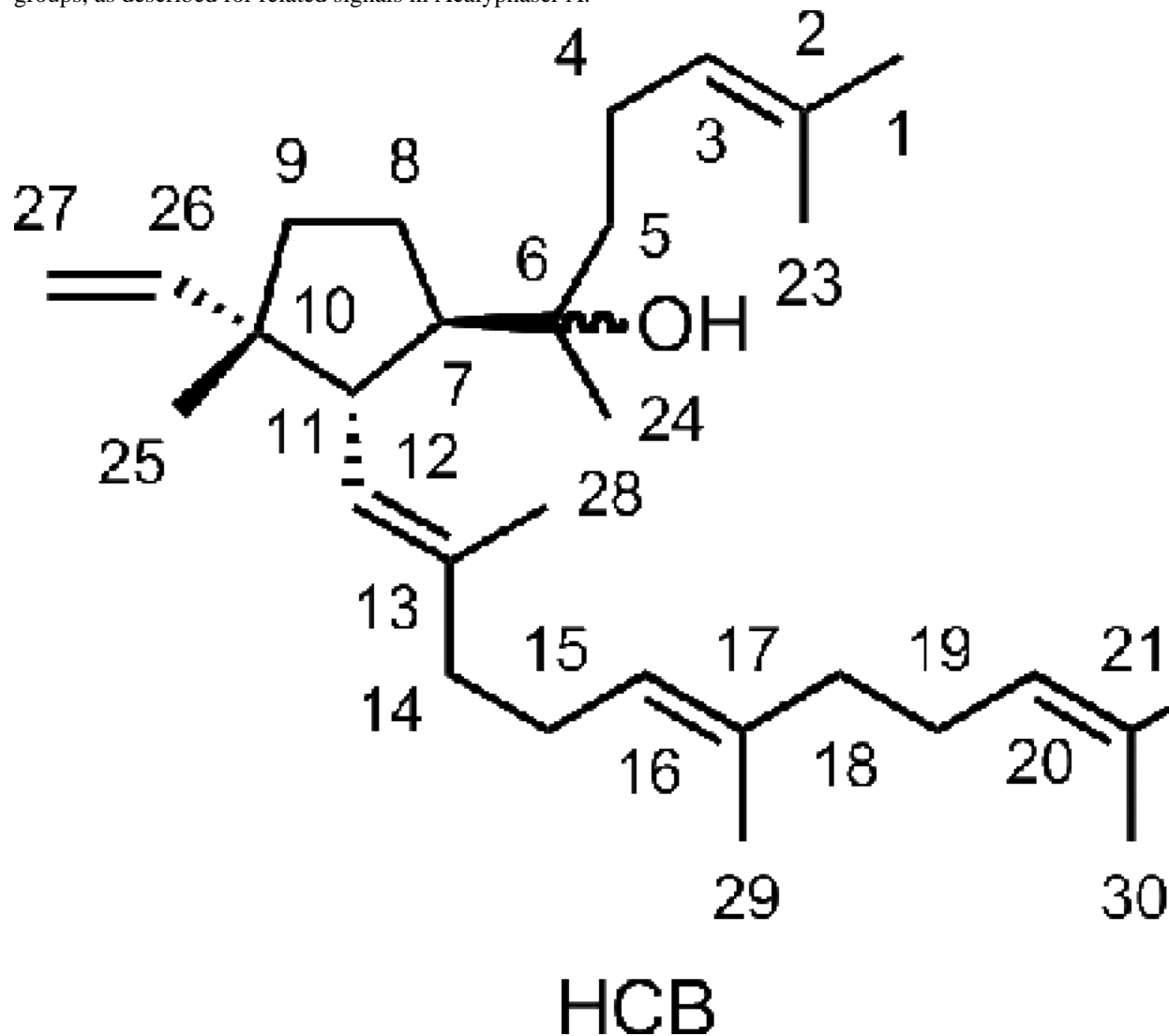
In 2000, Behrens and coworkers reported the isolation of a partially hydrogenated C<sub>30</sub> botryococcene-related triterpene, *Z,E*-7,11-cyclobotryococcan-5,12,26-triene with an unusual cyclopentane ring in the middle of the 1'-3-linked chain from organic-rich sediments obtained from Lake Cadagno in Switzerland.<sup>6</sup> They suggested that the cyclobotryococcene-like triene might be derived from PSPP. More recently, Kambara et al. reported the structure of acalyphaser A, a related cyclic tetraterpene with 1'-3 linked farnesyl and farnesylgeranyl moieties, from *Acalypha siamensis* Oliv.<sup>7</sup> *Ex Gage*, a shrub that grows in Indochina. We now report that the sixth product from the SQase-catalyzed solvolysis of PSPP is a triterpene analog of Acalyphaser A.



In conjunction with our mechanistic studies of SQase, we engineered a His<sub>6</sub>-tagged version of the yeast enzyme with truncations at the N- and C-termini, dt-ySQase, in order to increase its solubility and facilitate purification. The gene for the doubly truncated recombinant protein lacking 31 N-terminal and 24 C-terminal amino acids was ligated into pET28b+ to give plasmid pdt-ySQase, which was cloned into *E. coli* XL1-Blue. The activity of the soluble Ni-NTA purified protein (15-20 mg/L of culture) was ~17% of the less soluble enzyme with a 24 aa C-terminal truncation (mt-ySQase). HPLC analysis of the methyl t-butyl ether extracts of an incubation of dt-ySQase with FPP gave the 6 products previously seen for mt-ySQase, along with a slightly increased amount of rillingol (ROH)<sup>4</sup> and two minor products at 9.8 and 13.5 min not seen in the HPLC trace of the previous study (Figure 1). In addition, the amounts of the 1'-3 products HBO and isoHBO were substantially larger than found for mt-ySQase.<sup>3</sup>

The structure of the compound that elutes at ~6.5 min was established from its high-resolution electrospray mass spectrum (C<sub>30</sub>H<sub>50</sub>ONa, M + Na<sup>+</sup> at 449.3760; calc. 449.3759); from its 1D <sup>1</sup>H and 2D <sup>1</sup>H-<sup>1</sup>H COSY and <sup>1</sup>H-<sup>13</sup>C HSQC and HMBC NMR spectra (Table 1); and by comparisons with the related spectra for acalyphaser A.<sup>7</sup> Obvious connections between the <sup>1</sup>H NMR spectrum of the unknown, which we assign as hydroxycyclopentylbotryococcene (HCB), and that of acalyphaser A are the resonances for the C26-C27 vinyl group and a doublet of doublets (apparent triplet, *J* = 10.6 Hz) at 2.47 ppm for the C11 proton (see Table 1). A COSY spectrum showed correlations between H11/H7, H11/H12, H7/H8, and H8/H9. In addition, HMBC cross peaks were seen for H11/C26, H11/C13, H11/C6, H11/C10, H11/C8, H25/C26 and H25/C9. These observations clearly indicated a five-membered ring in HCB, and a quaternary C10 center with vinyl and methyl substituents. Furthermore, the HMBC correlations between H24/C6, H24/C7 and H24/C5, and the chemical shift of C6 confirmed that C6 is a quaternary carbon with an attached OH group. Once the ring system was established, the remaining assignments of repeated isoprenyl units were patterned after those reported for acalyphaser A. The *E*-stereochemistry of the C12-C13 and C16-C17 internal

trisubstituted double bonds was assigned from the  $^{13}\text{C}$  chemical shifts of the attached methyl groups, as described for related signals in Acalyphaser A.<sup>7</sup>



The relative configuration of C7, C10, and C11 in HCB was established by a combination of coupling constants and NOE correlations as described for acalyphaser A.<sup>7</sup> The starting point was the observation of a large coupling constant (10.6 Hz) between H7 and H11, indicating a *trans* arrangement. The NOE observed between H11 and H24 provided additional support for this *trans* orientation. Molecular modeling of two possible configurations at C10 indicated the possibility of observing an NOE between H11 and H25 for both configurations. However, an NOE between H26 and H7 would only be expected for the configuration shown above. Thus, the relative configuration at C7, C10 and C11 was *R*\*, *S*\*, *S*\*, which is the same as that of cyclobotryococcane triene and Acalyphaser A. Since yeast SQase synthesizes (1*R*,2*R*,3*R*)-PSPP from FPP<sup>8</sup> and the quaternary cyclopropane center remains intact during the solvolysis reaction, it follows that HCB is the 7*R*,10*S*,11*S* enantiomer. The stereochemistry at C6 remains undefined.

Scheme 1 shows a mechanism for formation of the products from the dt-ySQase-catalyzed solvolysis of PSPP to give compounds with 1'-1 (squalene family) and 1'-3 (botryococcane family) carbon skeletons. The hydroxylation reactions that give HSQ and HBO are highly

stereoselective and their stereochemistries are consistent with the known stereoelectronic behavior of cyclopropylcarbinyl cations.<sup>3</sup> Likewise, the *7R,10R,11S* orientation of substituents on the cyclopentyl ring in HCB is consistent with a direct cyclization of the *c1'-2-3* cyclopropylcarbinyl cation to its tertiary cyclopentylcarbinyl isomer. However, the observation of stereoselectivity in an enzyme-catalyzed reaction cannot be used to exclude mechanisms based on non-enzymatic model reactions that are not stereoselective.<sup>9</sup> For example, model reactions for the hydroxylations that give HBO and HSQ are substantially less stereoselective than their enzymatic counter parts, indicative of competition between direct hydroxylation of the *c1'-2-3* and *c1'-1-2* cyclopropylcarbinyl cations to give HBO and HSQ, respectively, and rearrangement of the carbocations to their allylic isomers, followed by hydroxylation.<sup>3</sup> Detection of small amounts of the other enantiomers of HSQ and HBO among the products of the enzyme-catalyzed solvolysis is consistent with conformational rather than stereoelectronic control of the stereochemistry of the hydroxylation steps. Inspection of models suggests that the C7-C11 bond in HCB is more likely formed by intramolecular alkylation of the C6-C7 double bond in the allylic *1'-3* cation than in its *c1'-2-3* cyclopropylcarbinyl isomer.

HCB is a member of the *1'-3*-linked botryococcene family of triterpenes. Other members include the parental acyclic hydrocarbon, botryococcene, acyclic C<sub>31</sub>-C<sub>37</sub> methylated derivatives of botryococcene, and cyclohexyl derivatives, probably formed by cyclization during the methylation reactions.<sup>1,10-12</sup> In contrast to the cyclohexyl botryococcenes, the cyclopentane ring in HCB is an extension of the rearrangement reactions of PSPP that generate the botryococcene skeleton. Thus, in addition to catalyzing formation of *1'-1*, *1'-3*, and *c1'-1-3* structures, SQase also catalyzes the *1'-3* to cyclopentylcarbinyl cyclization. It is likely that a close relative of SQase synthesizes the presumptive *1'-3*-linked fully unsaturated precursor of cyclobotryococan-5,12,26-triene found in lakebed sediments.

Acalyphaser A is a diterpene constructed from C<sub>15</sub> farnesyl and C<sub>25</sub> farnesylgeranyl units rather than the more common pattern of two C<sub>20</sub> geranylgeranyl units. Our results suggest that the unusual diterpene is synthesized by reactions similar to those proposed in Scheme 1 for HCB. However, the C<sub>40</sub> (*1R,2R,3R*)-cyclopropylcarbinyl diphosphate, which is isomeric with prephytoene diphosphate, is formed from FPP and farnesylgeranyl diphosphate (FGPP) by addition of the farnesyl unit in FPP to the C2-C3 double bond in FGPP.

## Experimental Section

### Construction of *pd*t-*y*SQase

The open reading frame (orf) for *dt-y*SQase with a 5' NdeI and a 3' HindIII restriction site was obtained from a pET28+ plasmid containing a *mt-y*SQase insert by PCR using the forward primer 5'-AAAACCGCATATGGATCAGTCCACGCTCCATATCTCTTG-3' with a NdeI restriction site (underlined) and the reverse primer 5'-AAAAGCCAAGCTTATTACTTGTACTCTTCTTCTTGTTGGTTGG-3' with a HindIII restriction site (underlined). The start and stop codons are indicated in boldface. The resulting orf, which encoded *dt-y*SQase, containing the amino acids from D32 to K420 in *y*SQase, was cloned into the NdeI/HindIII site of pET28b+. The target sequence was amplified by PCR using Easy-A PCR under the following conditions: initial denaturation at 95 °C for 120 s, followed by 30 cycles of denaturation (95 °C, 35 s), annealing (53.3 °C, 40 s) and elongation (72 °C, 80 s), and a final extension at 72 °C for 10 min. The amplified PCR product (~1.2 kb) was gel-purified and ligated onto the NdeI/HindIII site of pGEM-T Easy. The resulting plasmid, pGEM*dt-y*SQase, was transformed into electro-competent XL1-Blue cells, which were then plated on the Luria-Bertani (LB) plates containing ampicillin (50 µg/mL) and tetracycline (12.5 µg/mL). Several colonies were picked from the plates and their plasmid DNA was extracted and sequenced.

The 1.2 kb NdeI/HindIII insert in pGEMdt-ySQase was removed and ligated into the NdeI/HindIII restriction sites of pET28b+ to give pdt-ySQase. The plasmid was transformed into electro-competent XL1-Blue cells, which were then plated on LB containing kanamycin (30  $\mu\text{g}/\text{mL}$ ). Several colonies were picked, and their plasmid DNA was isolated and sequenced to verify the correct insert. The encoded protein contained a N-terminal His<sub>6</sub> tag, a thrombin cleavage site, and dt-ySQase. The plasmid was transformed into expression strain *E. coli* BL21 (DE3).

### Expression and purification of recombinant dt-ySQase

A 1 L culture of *E. coli* pdt-ySQase/ BL21(DE3) in LB-kanamycin was shaken at 37 °C until OD<sub>600</sub> reached 0.7-0.8 at which point protein expression was induced with 1 mM IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside). Shaking was continued for 5 h at 30 °C.

The cells were harvested by centrifugation at 4,000 rpm at 4 °C for 15 min. The harvested cell paste was re-suspended in lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, containing 300 mM NaCl, 10 mM imidazole, 10 mM BME and protease inhibitors aprotinin, leupeptin and Pefabloc SC). The cells were disrupted by sonication on ice and spun down at 14,000  $\times$  g at 4 °C for 15 min. The cell debris was discarded and the supernatant was applied to a Ni-NTA column at 4 °C. The column was washed with 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, containing 300 mM NaCl, 20 mM imidazole, and 10 mM BME, and the protein was eluted with 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, containing 300 mM NaCl, 250 mM imidazole, and 10 mM BME. The eluted protein was dialyzed against 50 mM HEPES, pH 7.5, 10 mM BME and concentrated to ~3 mg/mL by centrifugation using a Amicon Ultra, Millipore (10,000 MWCO) membrane. The freshly prepared protein was used immediately.

### Incubations with dt-ySQase

A typical incubation was carried out as follows: 6 mg (13.8  $\mu\text{mol}$ ) of FPP and 30 mg of dt-ySQase (0.63  $\mu\text{mol}$ ) were incubated in 30 mL of 50 mM MOPS buffer, pH 7.2, containing 10 mM MgCl<sub>2</sub> at 30 °C for 3 h. The final concentrations of FPP and dt-ySQase were 0.46 mM and 1 mg/mL, respectively. The progress of the reactions was followed by measuring the increase in the UV absorption of dehydrosqualene (DSQ) at 300 nm.

### HPLC purification of products

After 3 h, the reaction mixture was saturated with solid NaCl and extracted with MTBE (3  $\times$  30 mL). The layers were cleanly separated by centrifugation at 3,000  $\times$  g. The MTBE extracts were combined and concentrated under a stream of nitrogen. The residues from 6 individual incubations were dissolved in a hexane and combined. The combined residues were purified by normal-phase HPLC on a Rainin Microsorb-MV silica gel column with isocratic elution (MTBE/hexanes, 1:19 v/v). Products were detected by UV absorbance at 214 nm. Fractions for individual peaks were collected, solvent was removed with a stream of nitrogen, and the residue was stored at -20 °C until analyzed.

### NMR Analysis of HCB

NMR data were obtained for a ~ 150  $\mu\text{g}$  sample at 600 MHz using a probe with a cryogenically cooled <sup>1</sup>H channel. Standard vendor-supplied pulse sequences were used for gCOSY, CIGAR-HMBC, gChsqc, and double-pulsed-field-gradient spin echo (DPFGSE) NOE experiments. For the gChsqc, 2.0 ms, 200 ppm STUD pulses were used for the <sup>13</sup>C 180° pulses. DPGFSE NOE experiments were acquired using a 500 ms mixing time. About 300  $\mu\text{g}$  of HCB was dissolved in 0.6 mL of CDCl<sub>3</sub> or CD<sub>2</sub>Cl<sub>2</sub> for the NMR measurements.

## Supplementary Material

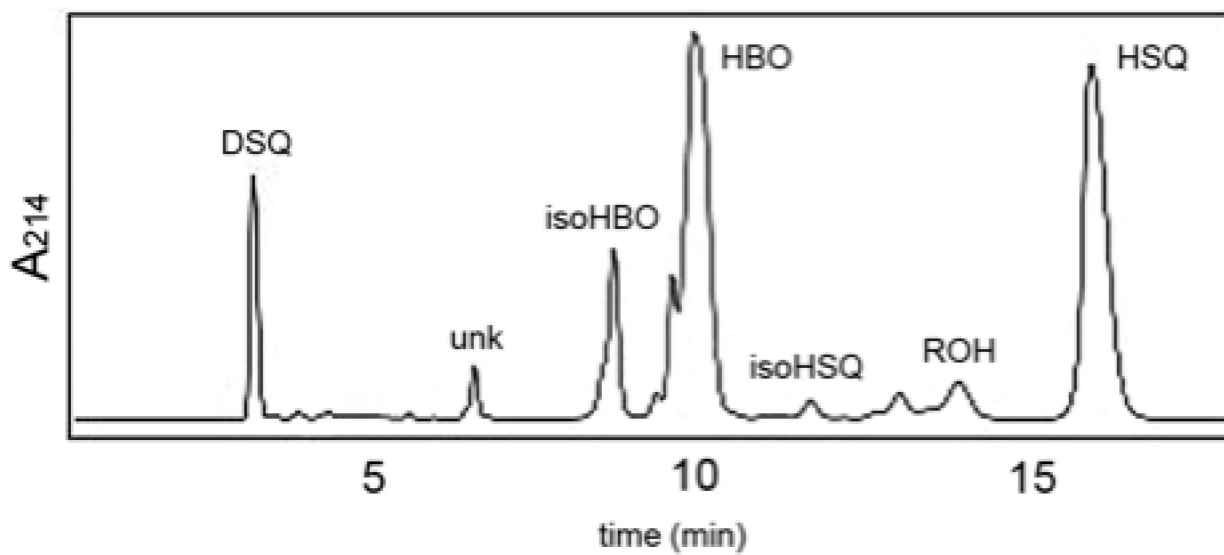
Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

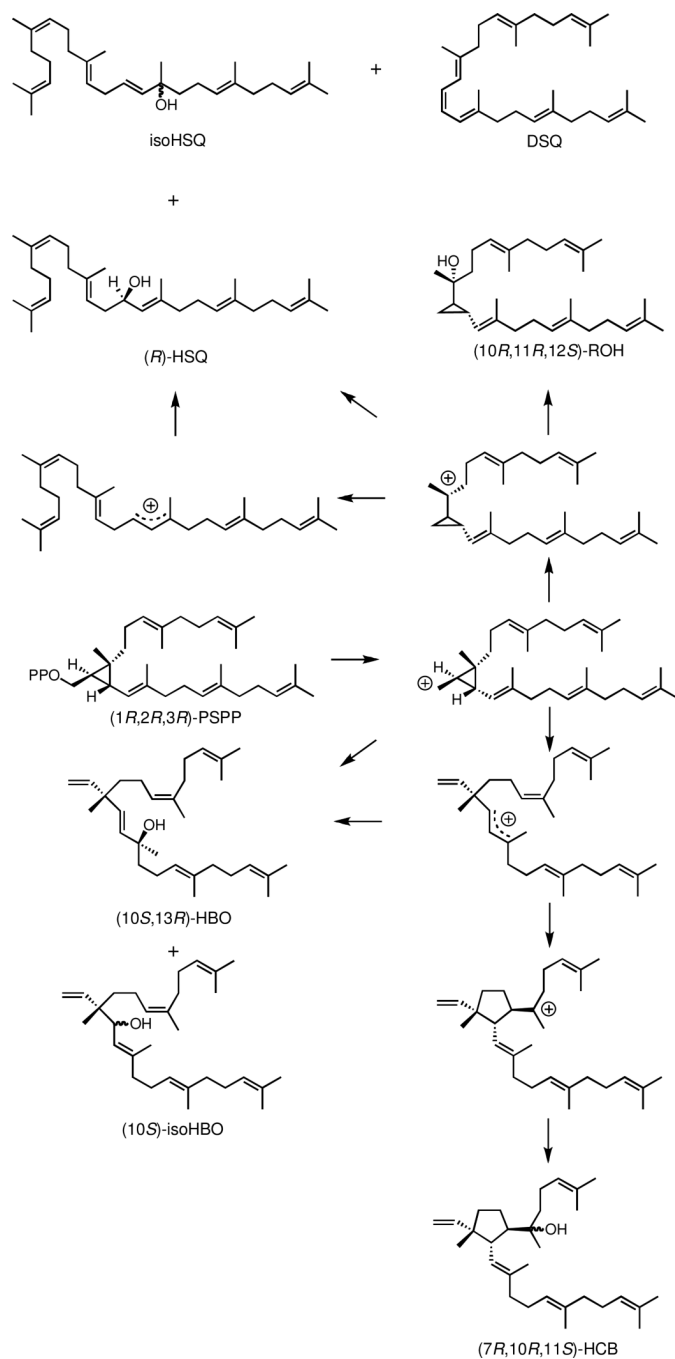
This work was supported by NIH grant GM 21328 (CDP) and NSF instrument grants RR-13030 and DBI-0002806.

## References

1. Wolf FR, Nemethy EK, Blanding JH, Bassham JA. *Phytochem* 1985;24:733–737.
2. Huang Z, Poulter CD. *J. Am. Chem. Soc* 1989;111:2713–2715.
3. Jarstfer MB, Zhang DL, Poulter CD. *J. Am. Chem. Soc* 2002;124:8834–8845. [PubMed: 12137536]
4. Blagg BSJ, Jarstfer MB, Rogers DH, Poulter CD. *J. Am. Chem. Soc* 2002;124:8846–8853. [PubMed: 12137537]
5. Zhang, D.-l.; Poulter, CD. *J. Am. Chem. Soc* 1995;117:1641–1642.
6. Behrens A, Schaeffer P, Bernasconi S, Albrecht P. *Org. Lett* 2000;2:1271–1274. [PubMed: 10810725]
7. Kambara H, Yamada T, Tsujioka M, Matsunaga S, Tanaka R, Ali HI, Wiart C, Yusof M, Hassan H, Hanifah A, Fauzi ZM, Mazlan NH, Jay M, Kunishima M, Akaho E. *Chem. & Biodivers* 2006;3:1301–1306. [PubMed: 17193244]
8. Popjak G, Edmond J, Wong S-M. *J. Am. Chem. Soc* 1973;95:2713–2714.
9. Poulter CD. *Acc. Chem. Res* 1990;23:70–77.
10. Casadevall E, Metzger P, Puech M-P. *Tetrahedron Lett* 1984;25:4123–4126.
11. Metzger P, David M, Casadevall E. *Phytochem* 1987;26:129–134.
12. Huang Z, Poulter CD, Wolf FR, Somers TC, White JD. *J. Am. Chem. Soc* 1988;110:3959–3964.



**Figure 1.**  
HPLC trace of products from dt-ySQase catalyzed solvolysis of PSPP.



**Scheme 1.**  
Mechanism for synthesis of non-head-to-tail triterpenes by squalene synthase.



Table 1

<sup>1</sup>H and <sup>13</sup>C NMR data for HCB

#	$\delta^1\text{H}$	$\delta^{13}\text{C}$
1	1.68 (s)	25.4
2		131.2
3	5.07-5.11 (m)	124.7
4	1.97-2.13 (m)	21.9
5	1.36-1.48 (m)	41.2
6		74.7
7	2.04-2.13 (m)	53.0
8	1.36-1.48 (m)/ 1.76-1.85 (m)	25.2
9	1.36-1.48 (m)/ 1.76-1.85 (m)	37.2
10		50.5
11	2.47 (t, $J=10.6$ )	50.3
12	5.07-5.11 (m)	127.2
13		137.1
14	1.97-2.13 (m)	40.0
15	1.97-2.13 (m)	26.6
16	5.07-5.11 (m)	124.3
17		135.3
18	1.97-2.13 (m)	39.8
19	1.97-2.13 (m)	26.6
20	5.07-5.11 (m)	124.7
21		131.2
22	1.67 (s)	25.4
23	1.61 (s)	17.4
24	1.11 (s)	23.6
25	1.02 (s)	24.6
26	5.88 (dd, $J=17.5, 10.9$ )	143.6
27	4.98 (dd, $J=10.9, 1.5$ )/ 4.92 (dd, $J=17.5, 1.5$ )	111.1
28	1.70 (s)	16.2
29	1.61 (s)	17.2
30	1.61 (s)	17.4