

Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16 α hydroxylase from *Bupleurum falcatum*

Tessa Moses^{a,b,c,d}, Jacob Pollier^{a,b}, Lorena Almagro^e, Dieter Buyst^f, Marc Van Montagu^{a,b,1}, María A. Pedreño^e, José C. Martins^f, Johan M. Thevelein^{c,d}, and Alain Goossens^{a,b,1}

^aDepartment of Plant Systems Biology, VIB, 9052 Gent, Belgium; ^bDepartment of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Gent, Belgium; ^cDepartment of Molecular Microbiology, VIB, 3001 Leuven-Heverlee, Belgium; ^dLaboratory of Molecular Cell Biology, Institute of Botany and Microbiology, Katholieke Universiteit Leuven, 3001 Leuven-Heverlee, Belgium; ^eDepartment of Plant Biology, Faculty of Biology, Universidad de Murcia, 30100 Murcia, Spain; and ^fDepartment of Organic Chemistry, Ghent University, 9000 Gent, Belgium

Contributed by Marc Van Montagu, December 17, 2013 (sent for review April 29, 2013)

The saikosaponins comprise oleanane- and ursane-type triterpene saponins that are abundantly present in the roots of the genus *Bupleurum* widely used in Asian traditional medicine. Here we identified a gene, designated CYP716Y1, encoding a cytochrome P450 monooxygenase from *Bupleurum falcatum* that catalyzes the C-16 α hydroxylation of oleanane- and ursane-type triterpenes. Exploiting this hitherto unavailable enzymatic activity, we launched a combinatorial synthetic biology program in which we combined CYP716Y1 with oxidosqualene cyclase, P450, and glycosyltransferase genes available from other plant species and reconstituted the synthesis of monoglycosylated saponins in yeast. Additionally, we established a culturing strategy in which applying methylated β -cyclodextrin to the culture medium allows the sequestration of heterologous nonvolatile hydrophobic terpenes, such as triterpene sapogenins, from engineered yeast cells into the growth medium, thereby greatly enhancing productivity. Together, our findings provide a sound base for the development of a synthetic biology platform for the production of bioactive triterpene sapo(ge)nins.

cyclodextrin | amyirin | metabolic engineering | triterpenoid

Triterpene saponins are secondary metabolites that exhibit a large structural diversity and wide range of biological activities in many plant species (1, 2). Saponins are glycosides of sapogenins, which are composed of 30 carbon atoms arranged in 4- or 5-ring structures that are “decorated” by functional groups. Saponins are synthesized by multiple glycosylations of the sapogenin building blocks that are produced by multiple cytochrome P450-dependent monooxygenase (P450) or oxidoreductase-mediated modifications of basic backbones, such as β -amyirin (oleanane type), α -amyirin (ursane type), lupeol, and dammarenediol. These backbones are generated by oxidosqualene cyclase (OSC)-mediated cyclization of 2,3-oxidosqualene, which is also an intermediate in the synthesis of sterols in eukaryotes (3, 4). Both saponins and sapogenins include biologically active compounds or serve as starter molecules for the generation of novel, potentially bioactive structures by synthetic modification (5–7).

The genus *Bupleurum* consists of perennial herbs that are used in Asian traditional medicine, either alone or in combination with other ingredients, for the treatment of common colds, fever, and inflammatory disorders (8). Saikosaponins constitute the largest class of secondary metabolites in *Bupleurum* and can account for up to ~7% of root dry weight. Their accumulation can be further stimulated by jasmonate treatment (9). More than 120 closely related oleanane- and ursane-type saikosaponins have been identified from this genus and the oxidations at various positions suggest the presence of multiple enzymes, mainly P450s, capable of catalyzing specific modifications on the amyirin backbones (8, 10). To date, no P450 or oxidoreductase involved in triterpene saponin biosynthesis has been identified from *Bupleurum* species.

P450s that modify the β -amyirin backbone on C-11; C-12,13; C-16; C-22; C-23; C-28 or C-30 have been characterized from *Glycyrrhiza uralensis*, *Avena strigosa*, *Medicago truncatula*, *Glycine max*, *Vitis vinifera*, and *Catharanthus roseus* (11–18). Hydroxylases from *Panax ginseng* that oxidize the dammarenediol-II backbone on C-6, C-12, or C-28 (19–21), and a C-20 hydroxylase from *Lotus japonicus* (22) that modifies lupeol, have also been identified. To characterize these P450s, they have been ectopically expressed in yeast strains either producing β -amyirin or externally supplied with candidate substrates. Similarly, several OSCs have been produced and functionally analyzed in yeast. From these studies, it is clear that yeast cells cannot only be used for the characterization of novel enzymes, but possibly also as a heterologous host for the production of triterpene sapogenins (23). To date only two pilot studies have aimed at engineering of β -amyirin production in yeast (24, 25), but no efforts toward engineering of sustainable production of sapogenins or saponins in yeast have been reported.

Here, we identified and characterized CYP716Y1, a P450 from *Bupleurum falcatum* that corresponds to a C-16 α oxidase, designated according to Nelson’s nomenclature (<http://drnelson.uthsc.edu/cytochromeP450.html>). By designing triterpene-hyperproducing

Significance

Saponins are plant molecules that are produced as a chemical defense against herbivores and eukaryotic pathogens. They constitute structurally diverse, bioactive compounds composed of a 30-carbon triterpene backbone adorned with multiple functional groups and sugars. Saikosaponins are abundant saponins accumulating in the Asian medicinal plant *Bupleurum falcatum*, but none of the enzymes involved in their biosynthesis had been characterized. We identified a cytochrome P450 involved in the oxidation of saikosaponins, thereby expanding the enzyme compendium that can generate plant saponins with an extra activity. Using this enzyme compendium, we established a synthetic biology program to reconstitute saponin biosynthesis in the yeast *Saccharomyces cerevisiae* and developed a cyclodextrin-based culturing strategy to sequester triterpenes from engineered yeast cells and enhance their productivity.

Author contributions: T.M., J.P., M.V.M., J.M.T., and A.G. designed research; T.M., J.P., L.A., and D.B. performed research; M.A.P. and J.C.M. contributed new reagents/analytic tools; T.M., J.P., D.B., J.C.M., and A.G. analyzed data; and T.M., J.P., D.B., J.C.M., and A.G. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database [accession no. [NC963423](https://www.ncbi.nlm.nih.gov/nuclot/NC_0963423) (CYP716Y1)].

¹To whom correspondence may be addressed. E-mail: alain.goossens@psb.vib-ugent.be or marc.vanmontagu@ugent.be.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323369111/-DCSupplemental.

starter strains, optimizing culturing conditions for triterpene synthesis, and using the *CYP716Y1* gene in a combinatorial synthetic biology program, we established a platform that allows us to produce and sequester triterpene saponin in culture medium and to reconstitute a full saponin synthetic pathway in yeast cells.

Results

Transcript Profiling of Jasmonate-Treated *B. falcatum* Roots Reveals *CYP716Y1*. To identify saponin biosynthetic genes, expression analysis on roots of methyl jasmonate (MeJA)-treated *B. falcatum* plants was performed with cDNA-amplified fragment length polymorphism–based transcript profiling (26). The expression of 18,695 transcript tags was monitored over time and 1,771 MeJA-responsive tags (hereafter referred to as “BF tags”) were isolated. Direct sequencing of the reamplified BF tags gave good-quality sequences for 1,217 fragments. BF tags corresponding to genes encoding enzymes known to catalyze early steps in triterpene saponin biosynthesis, such as β -amyrin synthase (bAS), displayed similar expression patterns, suggesting tight coregulation, and reached maximum expression levels 4–24 h post-MeJA elicitation (Fig. 1A). The gene corresponding to tag BF567 was coregulated with these genes and matched a P450 with 48% and 47% sequence similarity with the *P. ginseng* CYP716A53v2 and *M. truncatula* CYP716A12, respectively. The full-length open reading frame corresponding to the BF567 tag (hereafter called “*CYP716Y1*”) was cloned from a *B. falcatum* cDNA library (27). Phylogenetic analysis confirmed that *CYP716Y1* belongs to the CYP716 family of P450s involved in triterpene saponin biosynthesis (Fig. 1B and Fig. S1).

Generation of Triterpene-Producing Yeast Strains. To determine the *in vivo* enzymatic activity of *CYP716Y1*, we engineered *Saccharomyces cerevisiae* strains for hyperproduction of triterpenes, analogous to a previously described strain (24). All yeast strains generated in this study are listed in Table S1. First, we replaced the native promoter of lanosterol synthase with a methionine-repressible promoter in BY4742 to generate strain TM1, which in the presence of 1 mM methionine accumulated 60% less ergosterol than the parent strain. Next, the gene encoding a truncated feedback-insensitive copy of isoform 1 of the *S. cerevisiae* 3-hydroxy-3-methylglutaryl-CoA reductase (tHMG1) was cloned into the high-copy-number plasmid pESC-URA to generate pESC-URA[*GAL10/tHMG1*], which was transformed into TM1 to create strain TM5. We generated strains TM2 and TM3 from TM1 by expressing *bAS* of *M. truncatula* (GenBank accession no. AJ430607) (28) or *Glycyrrhiza glabra* (GenBank accession no. AB037203) (29), respectively using plasmid pESC-URA[*GAL10/tHMG1;GAL1/bAS*] (Fig. 2). Gas chromatography–mass spectrometry (GC-MS) analysis of cell extracts derived from 72-h-old TM2 and TM3 cultures revealed a single peak at 27.2 min with an electron ionization (EI) pattern corresponding to a β -amyrin standard, which was not observed in the GC chromatograms of the control strain TM5 (Fig. S2A and B). β -Amyrin accumulation was higher in TM3 (36.2 ± 3.9 mg/L) than in TM2 (19 ± 1.0 mg/L).

In parallel, we generated a lupeol-producing yeast strain TM6 from TM1 by expressing the *Arabidopsis thaliana* isoform 1 of lupeol synthase (*AtLUS1*) (GenBank accession no. U49919) (30) by means of plasmid pESC-URA[*GAL10/tHMG1;GAL1/AtLUS1*] (Fig. 2 and Fig. S2C). We quantified 46.3 ± 2.4 mg/L lupeol from cells of TM6 that additionally cyclized 2,3-oxidosqualene to 3β , 20-dihydroxylupane (31). Strain TM33 was obtained by expressing the *Centella asiatica* dammarenediol synthase gene (*CaDDS*) (GenBank accession no. AY520818) (32) from the pESC-URA [GAL10/tHMG1;GAL1/CaDDS] plasmid in strain TM1. In our hands, this strain produced α -amyrin, β -amyrin, and dammarenediol-II in a ratio of 88:11:1 (Fig. 2 and Fig. S2D). Hence, we used TM33 as an α -amyrin–producing strain.

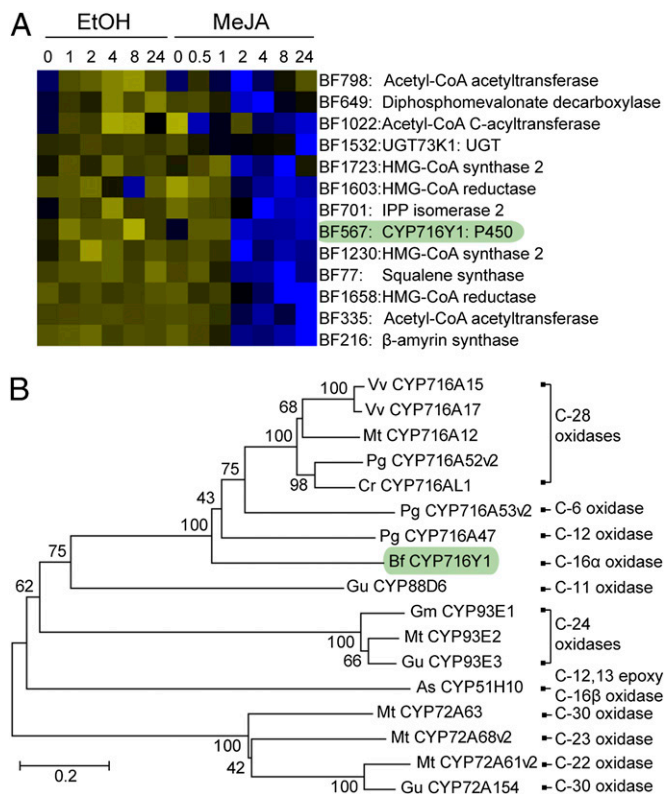


Fig. 1. Identification of *CYP716Y1* by transcript profiling of MeJA-treated *B. falcatum* roots. (A) Coregulation of *CYP716Y1* (in green) with other triterpene saponin biosynthetic enzymes. Cluster of the *B. falcatum* transcriptome comprising tags corresponding to genes reported to be involved in triterpene biosynthesis. Treatments and time points (in hours) are indicated at the top. Transcriptional activation and repression relative to the average expression level are represented by blue and yellow boxes, respectively. (B) Phylogenetic analysis of *CYP716Y1* (in green) and other P450s involved in triterpene saponin biosynthesis. The percentage of replicate trees that clustered together in the bootstrap test is indicated to the left of the branches. The scale bar gives the number of amino acid substitutions per site. The enzymatic activity of the P450s is indicated on the right. The amino acid sequences of the P450s were retrieved from GenBank (www.ncbi.nlm.nih.gov/genbank). As, *A. strigosa*; Bf, *B. falcatum*; Cr, *C. roseus*; Gm, *G. max*; Gu, *G. uralensis*; Mt, *M. truncatula*; Pg, *P. ginseng*; Vv, *V. vinifera*.

In Vivo Enzymatic Activity of *CYP716Y1*. Strains TM7 and TM26 expressing the *Arabidopsis* cytochrome P450 reductase (CPR) *AtATR1* (At4g24520) with or without *CYP716Y1*, respectively, were generated from the β -amyrin–producing strain TM3. In the GC chromatograms, a unique peak eluting at 31.8 min was observed in cell extracts of TM7 but not of TM26 (Fig. S3A). The EI pattern of this peak corresponded to a D- or E-ring–hydroxylated derivative of β -amyrin (Fig. S3B). As *CYP716Y1* was tentatively annotated as a homolog of *M. truncatula* *CYP716A12* (GenBank accession no. FN995113) that encodes a P450 catalyzing the three-step oxidation of β -amyrin at C-28 (16), we compared the EI pattern of this peak with an erythrodiol (28-hydroxy- β -amyrin) standard (Fig. S3D). Additionally, strain TM10 expressing *CYP716A12* and *AtATR1* was generated and the GC-MS profile of its cell extract was compared with that of TM7. A peak corresponding to the elution time and EI pattern of erythrodiol occurred at 32.5 min in TM10 extracts (Fig. S3C) but not in extracts of TM7 and TM26, suggesting that *CYP716Y1* and *CYP716A12* hydroxylate β -amyrin at different positions. From the literature we identified the oleanane-type triterpene saponin reported in the *Bupleurum* species (8) and postulated that *CYP716Y1* hydroxylates the D- or E-ring of β -amyrin at C-16 or C-21, respectively (Fig. S3B).

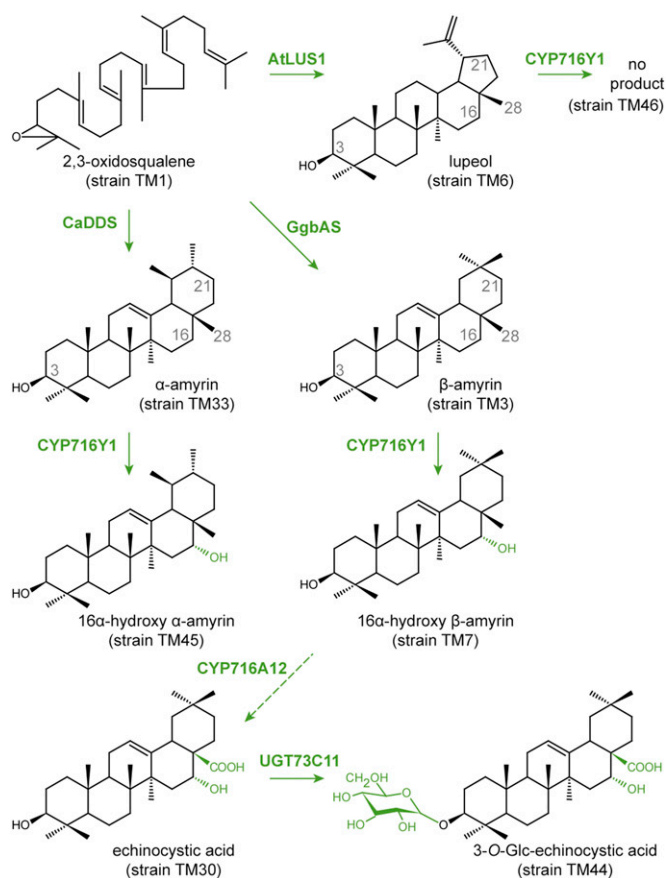


Fig. 2. Triterpene saponin biosynthesis pathway engineered in yeast. The triterpene saponin precursor 2,3-oxidosqualene is cyclized to different saponin backbones, such as lupeol, β -amyrin, or α -amyrin. These backbones can be modified by multiple P450s before ultimately being glycosylated by UGTs. The structures of the triterpene saponin intermediates accumulating in the engineered yeasts are shown, and the yeast strains producing these compounds are indicated between parentheses and listed in Table S1. The enzymes catalyzing these reactions and the modifications resulting from their activity are highlighted in green. Dashed arrows mark multiple enzymatic steps. AtLUS1, *A. thaliana* lupeol synthase 1; CaDDS, *C. asiatica* dammarenediol synthase; CYP716A12, *M. truncatula* P450 that catalyzes a multistep oxidation at C-28; CYP716Y1, *B. falcatum* P450 that hydroxylates C-16; GgbAS, *G. glabra* β -amyrin synthase; Glc, glucose; UGT73C11, *B. vulgaris* UGT that catalyzes glucosylation at C-3.

Finally, we generated strain TM45 expressing *CYP716Y1* and a control strain TM47 from the α -amyrin-producing strain TM33. GC chromatograms showed the accumulation of hydroxylated α -amyrin in TM45, but not in TM47 (Fig. S3 E and F). In contrast, no hydroxylated derivative of lupeol or $3\beta,20$ -dihydroxylupane could be detected in strain TM46, expressing *CYP716Y1* in the lupeol-producing strain TM6, indicating that lupeol is not a substrate for *CYP716Y1*.

Engineering Hydroxylated Amyrin Synthesis. Effect of CPR:P450 ratios on the *CYP716Y1* activity. Hydroxylated amyryns accumulated at low levels in strains TM7 and TM45 and the available amyryn pools seemed hardly consumed, suggesting inefficient *in vivo* conversion. The endoplasmic reticulum (ER)-localized CPRs are flavoproteins that serve as electron donor proteins for several ER oxygenases, including P450s (33). Optimal interaction between CPR and P450 is essential to allow the reducing equivalents from NADPH to pass from the CPR to the P450 and guarantee P450 functionality (34). To assess whether different CPR:P450 ratios would affect *CYP716Y1* performance, we generated strains

expressing *CYP716Y1* from a high-copy number plasmid together with *AtATR1* from either an integrated (1 copy per cell), low-copy (3–5 copies per cell), or high-copy (10–40 copies per cell) number plasmid. Strain TM9 with *AtATR1* on the low-copy plasmid accumulated the highest hydroxylated β -amyryn levels, whereas the lowest accumulation occurred in strain TM8 expressing the integrated copy of *AtATR1* (Fig. S4). Although the actual CPR:P450 ratios were not biochemically measured, the product accumulation observed after attempting to alter the ratios through the vector copy numbers is in line with the P450 being in excess of the CPR, as reported for mammalian and yeast systems (34). Alternatively, the reaction rate of P450s can be (further) enhanced by interaction with cytochrome *b*₅, as recently demonstrated in a yeast strain engineered for artemisinin production (35), but this was not assessed here.

Cyclodextrin facilitates sequestration and stimulates production of triterpene saponogenins. Intracellular accumulation of saponogenins might be detrimental to cell viability or alternatively, exert feedback inhibition on the P450 activity as reported for human and *Streptomyces* P450s (36, 37), possibly accounting for the observed low hydroxy-amyryn accumulation in *CYP716Y1*-expressing yeasts. Cyclodextrins (CDs) are cyclic oligosaccharides composed of α -D-glucopyranoside units that can be represented by a toroidal topology (Fig. S5). CDs are extensively used for solubilizing and stabilizing pharmaceuticals because their toroidal interior is more hydrophobic, enabling them to host apolar molecules, such as cholesterol, whereas the hydrophilic exterior allows their solubilization in aqueous environments (38, 39). We hypothesized that CDs could also sequester saponogenins, facilitating their export to the growth medium and eventually stimulating their production.

First, we determined whether methyl- β -cyclodextrin (M β CD), the most commonly used CD, could sequester triterpenes. GC-MS analysis of extracts from cells and spent medium of strain TM3 treated or not with 5 or 25 mM M β CD revealed that both ergosterol and β -amyryn could be detected in all of the cell extracts, but only in the spent medium of M β CD-treated samples (Fig. 3A), indicating that M β CD was capable of sequestering and promoting export of both sterol- and saponenin-type triterpenes. Although M β CD concentration and sequestering capacity correlated positively (Fig. 3A), the lowest concentration was used in further experiments for practical reasons. With 5 mM M β CD, we quantified 20.4 ± 10.5 mg/L and 37.3 ± 3.6 mg/L β -amyryn from cells and spent medium of TM3, respectively (Fig. 3B). The total concentration of β -amyryn in M β CD-treated cultures was 1.6-fold higher than in untreated cultures.

Next, we investigated whether timing and frequency of M β CD application during the yeast cultivation procedure affected β -amyryn production (Fig. 3C). We applied M β CD at a concentration of 5 mM at different times during culturing, thereby obtaining seven treatment conditions, including the mock-treated control. On day 1, to samples I1, I2, and I3, M β CD was added concomitantly with culture inoculation in galactose medium to induce heterologous gene expression. On day 2, to samples I2 and I3, M β CD was added again simultaneously with methionine; and on day 3, 24 h after the methionine repression, it was supplemented to sample I3 only. To samples R1, R2, and AR1, M β CD was added on day 2 only, days 2 and 3, and day 3 only, respectively. The spent medium of all samples was extracted on day 4 and β -amyryn quantified (Fig. 3C). The frequency of M β CD application and triterpene amounts correlated directly. For practical reasons, condition R2 was judged the most effective and used in subsequent assays to assess M β CD's capacity to sequester hydroxylated β -amyryn. When M β CD was applied to strain TM9, hydroxylated β -amyryn occurred only in the spent medium (Fig. 3D), suggesting that the saponogenins were sequestered completely from the yeast cells. In addition, 5.5-fold more saponenin could be extracted from the M β CD-treated than from untreated TM9 cells.

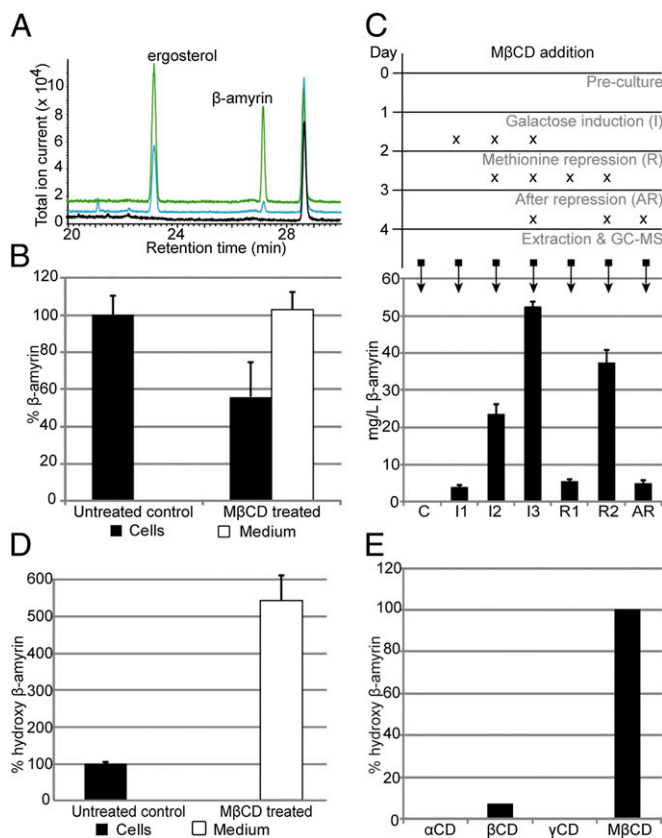


Fig. 3. Enhanced accumulation of CYP716Y1 products by application of CDs. (A) Sequestration of triterpenes from yeast cells by M β CD. Overlay of GC chromatograms showing ergosterol and β -amyryn in the spent medium of strain TM3 treated with 5 mM (blue) or 25 mM (green) M β CD and untreated control (black). (B) Enhancement of triterpene productivity by M β CD. Quantification of β -amyryn sequestered from cells of strain TM3 upon 5 mM M β CD treatment showed a 1.6-fold higher β -amyryn amount in the M β CD-treated culture with 100% corresponding to 36.2 ± 3.9 mg/L (mean \pm SD, $n = 3$) of β -amyryn in the untreated control. (C) M β CD dose-dependent sequestration of triterpenes. Quantification of β -amyryn from spent medium of strain TM3 treated with M β CD on day 1 (I1); days 1 and 2 (I2); days 1, 2, and 3 (I3); day 2 (R1); days 2 and 3 (R2); day 3 (AR); and an untreated control (C). (D) Effects of M β CD on hydroxy- β -amyryn sequestration and productivity. Relative amounts of hydroxy- β -amyryn quantified from the cells and spent medium of strain TM9 expressing CYP716Y1 treated or not with M β CD. (E) Specific sequestration of triterpenes by M β CD. Relative comparison of the hydroxy- β -amyryn quantified from spent medium of strain TM9 treated with different CD variants (Fig. S5). Quantitation and error bars correspond to mean and SD values, respectively ($n = 3$).

Finally, we compared the efficiency of sequestration with different types of CDs on strain TM9. The most common CD variants are α , β , and γ with 6, 7, and 8 D-glucopyranoside units, respectively (Fig. S5). We applied α CD, β CD, γ CD, or M β CD at a concentration of 5 mM under condition R2. Sequestering was observed only with β CD and its methylated derivative with the highest amount extracted from the M β CD-treated culture (Fig. 3E). The specificity of β CD could be a consequence of the variable internal toroidal diameters of the CD variants, resulting from the number of glucopyranoside units they contain (Fig. S5). Likewise, the high affinity of saponin derivatives for the methylated β CD compared with the nonmethylated β CD could be related to their higher interior hydrophobicity.

Combinatorial Biochemistry Toward in Vivo Production of Saponins in Yeast. In combinatorial biochemistry, genes from different organisms are combined in a heterologous host to produce novel bioactive compounds (40). As CYP716Y1 and CYP716A12 oxidize

β -amyryn at different positions, we reasoned that combining the enzymes in strain TM3 could result in a combinatorial compound corresponding to an intermediate of the saponin biosynthetic pathways from *M. truncatula* or *B. falcatum* that normally does not accumulate in these species. We generated strain TM30 in which CYP716Y1 and CYP716A12 were expressed from a high-copy number plasmid to produce a self-processing polyprotein in which the two enzymes are linked via a 2A oligopeptide (41). GC chromatograms of the extracts from the spent medium of TM30, TM9, and TM17 cultured with M β CD showed a single unique peak at 40.5 min in strain TM30 but not in TM9 and TM17 (Fig. 4A). This additional peak corresponds to β -amyryn oxidized at two positions, presumably a hydroxy oleanolic aldehyde based on the EI-MS spectra (Fig. S3G), the differences with an available 3 β ,16 α -dihydroxyolean-12-en-28-oic acid (also called “echinocystic acid”) standard (Fig. S3H and I), and the predicted enzymatic activities of CYP716A12 (aldehyde on C-28) and CYP716Y1 (hydroxyl on C-16 or C-21).

Next, to synthesize saponins, we generated strain TM44 that expressed two self-processing polyproteins, one with the two P450s and one consisting of AtATR1 and the recently reported *Barbarea vulgaris* UDP-dependent glucosyltransferase (UGT) UGT73C11 (GenBank accession no. JQ291614) (42). In parallel, strains TM48 and TM49 were designed expressing the self-processing polyprotein of AtATR1 with UGT73C11 together with either CYP716Y1 or CYP716A12 alone. Liquid chromatography (LC)–MS chromatograms of cell extracts revealed accumulation of two saponins in strain TM44 (Fig. 4B), demonstrating that yeast can be engineered to produce monoglucosylated saponins. As one of these compounds also occurred in strain TM49 expressing CYP716A12 alone, it probably corresponds to 3-O-Glc-oleanolic

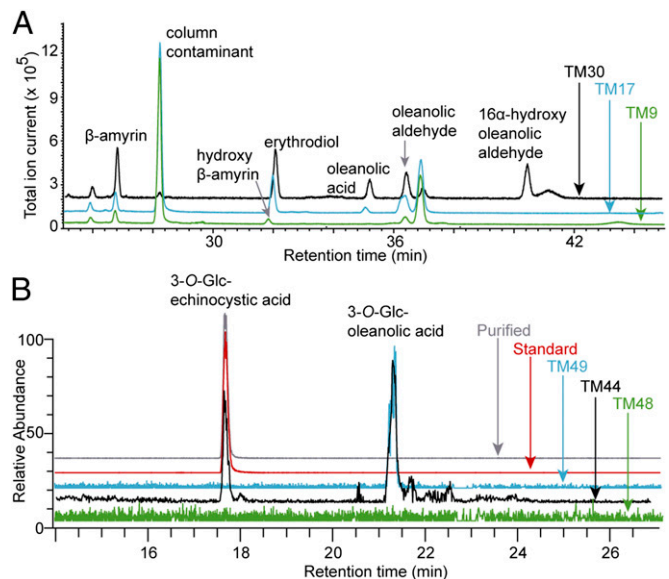


Fig. 4. In vivo combinatorial production of 3-O-Glc-echinocystic acid in yeast. (A) Overlay of GC chromatograms showing the accumulation of hydroxyl oleanolic aldehyde in strain TM30 expressing CYP716Y1 and CYP716A12, but not in strains TM9 and TM17 expressing either of the P450s alone. (B) Overlay of LC-MS-extracted ion chromatograms of formate adducts showing the accumulation of 3-O-Glc-echinocystic acid and 3-O-Glc-oleanolic acid in strain TM44 expressing both CYP716Y1 and CYP716A12. Strain TM49 expressing only CYP716A12 accumulates 3-O-Glc-oleanolic acid, whereas strain TM48 expressing only CYP716Y1 does not produce any glucosylated product. For 3-O-Glc-echinocystic acid, 3-O-Glc-oleanolic acid, and 3-O-Glc-3 β ,16 α -dihydroxyolean-12-ene the mass ranges 678.0–679.0 Da, 662.9–663.4 Da, and 648.0–649.0 Da, respectively were screened in all chromatograms.

acid based on the reported enzymatic activities (Fig. 4B). No M β CD was applied in these experiments, indicating that it was not necessary to allow accumulation of saponins. As no glucosylation of the hydroxy- β -amyirin was observed in strain TM48, β -amyirin might need to undergo a certain number of oxidations before it can be accepted by UGT73C11 as an in vivo substrate.

CYP716Y1 Catalyzes Hydroxylation of Amyrins on C-16 in an α -Configuration. To unambiguously establish the enzymatic activity of CYP716Y1, we purified the second monoglucosylated saponin from strain TM44 and investigated its structure by 1D and 2D NMR. Complete assignment was not achieved, but the NMR analysis conclusively established that the product of the combined action of bAS, CYP716A12, CYP716Y1, and UGT73C11 was α -hydroxylated at C-16 of the triterpene base structure (Fig. 2). Further validation was obtained by (i) the near identity of the NMR data with that measured for a commercial 3-O-Glc-echinocystic acid standard (Figs. S6 and S7) and (ii) the comigration of the TM44-specific compound and the 3-O-Glc-echinocystic acid standard by LC-MS analysis (Fig. 4B), definitively establishing that CYP716Y1 hydroxylates β -amyirin on C-16 in an α -configuration.

Discussion

The pharmacological activity of the genus *Bupleurum*, used in Asian traditional medicine, is generally attributed to the presence of saikosaponins, the most abundant secondary metabolites produced by this genus. The saikosapogenins can contain oxidative modifications at the C-11, C-16, C-21, C-23, C-28, C-29, and C-30 positions (8). Through transcript profiling of MeJA-treated *B. falcatum* roots, we identified CYP716Y1, encoding a P450 that is transcriptionally coregulated with known triterpene saponin biosynthetic genes and that catalyzes the C-16 α hydroxylation of amyirin, a catalytic activity previously not linked to any gene product in the plant kingdom. Hence, the P450 compendium that can oxidize the amyirin backbone is expanded, now covering seven positions on the triterpene backbone (11–18) and including two enzymes that catalyze hydroxylation at C-16, in β - (CYP51H10) (12) and α -configuration (CYP716Y1). Like many reported P450s (15, 17), CYP716Y1 can hydroxylate α - and β -amyirin, both pentacyclic triterpenes.

We generated a high-titer β -amyirin-producing *S. cerevisiae* strain (TM3) by expressing bAS from *G. glabra*. By means of a previously published engineering strategy (24) but with a more efficient bAS, we improved the β -amyirin production by sixfold and reached levels of 36 mg/L. Even our less efficient strain, expressing a *M. truncatula* bAS, produced threefold more β -amyirin (19 mg/L) than the previously published strain (24). This hyperproducing strain was used for the ectopic expression of CYP716Y1, causing accumulation of a hydroxylated product of β -amyirin. However, the production levels were low and only a fraction of the available amyirin pool seemed to be converted. Several reasons may account for this observation, such as the suboptimal folding, recycling, or localization of the P450 protein or, alternatively, the inappropriate localization of the catabolic product, thereby inhibiting P450 activity or affecting cell viability. The latter reflects the concerns that triterpene production might not be as amenable to engineering efforts as production of volatile sesquiterpenes and monoterpenes, molecules that can readily diffuse out of cells (24). Here, we have tackled this issue and designed a culturing strategy in which triterpenes are efficiently transferred from yeast cells into the culture medium by the use of M β CD. We show that these molecules are capable of sequestering hydrophobic triterpene sapogenins into the growth medium, thereby relieving the yeast cells from the load of toxic heterologous compounds, which, in turn, dramatically increased production levels. As the application of M β CD allows compound sequestration from living cells without damaging the integrity of the production host (43), this principle could be

applied to continuous culture systems to produce large amounts of any valuable hydrophobic compound, including triterpene sapogenins.

Finally, we have advanced triterpene engineering in two additional aspects. First, we have demonstrated combinatorial triterpene sapogenin biosynthesis by combining CYP716Y1 from *B. falcatum* and CYP716A12 from *M. truncatula* in our yeast strain expressing *G. glabra* bAS and *A. thaliana* CPR, leading to the production of oxidized amyrins, sequestered in the medium after CD application. Second, we have been able to reconstitute a full saponin synthetic pathway in yeast. By supertransforming yeast strains that produce oxidized amyrins with the *B. vulgaris* UGT73C11, encoding a 3-O-glucosyltransferase (42), the corresponding monoglucosylated saponins accumulated in the yeast cells, independently of a need for CDs.

In conclusion, we have shown the versatile ability of yeast cells to produce natural or nonnatural sapogenins and saponins, through the utilization of M β CD as a sequestering agent for the hydrophobic sapogenins and by reconstitution of a full pathway to synthesize glycosylated saponins by the combined expression of OSC, P450, and UGT genes. Considering that the current strains have undergone little systems engineering, for instance with regard to the fluxes in the endogenous precursor or competing pathways, we trust that there is still much room for improvement of product yields. Hence, we believe that we have set an excellent base for a synthetic biology program toward the establishment of an efficient platform for the commercial production of valuable bioactive sapo(ge)nins.

Materials and Methods

Chemicals. β -Amyirin, erythrodiol, and 3-O-Glc-echinocystic acid were purchased from Extrasynthese; random M β CD from CAVASOL; and α CD, β CD, and γ CD from Cyclolab.

Cultivation and Elicitation of *B. falcatum*. *B. falcatum* seeds (purchased from www.sandmountainherbs.com) were germinated in soil. Two-week-old seedlings were transferred to aerated hydroponic medium containing 1 g/L 10-30-20 salts (Scotts). The pH of the medium was monitored daily and maintained at 6.5 with potassium hydroxide (KOH). Plants were grown in a 16-h:8-h light:dark regime at 21 °C. After 3 wk, the plants were treated with 50 μ M MeJA or an equivalent amount of ethanol as control by adding the solution directly to the hydroponic medium. For transcript profiling (SI Materials and Methods), roots were harvested 0, 0.5, 1, 2, 4, 8, and 24 h after treatment, frozen in liquid nitrogen, and stored at –70 °C. For each sample, three individual plants were pooled.

Cloning of CYP716Y1. Primers P11+P12 were used to screen for the full-length coding sequence of CYP716Y1 in a *B. falcatum* Uncut Nanoquantity cDNA library (custom made by Invitrogen) as reported (27). The full-length CYP716Y1 was amplified by PCR for Gateway cloning into pDONR221 with primers P21+P22. All primers used are listed in Table S2. Phylogenies were analyzed as described (SI Materials and Methods).

Generation and Culturing of Yeast Strains. Yeast strain TM1 and all derivatives were generated from the BY4742 strain (SI Materials and Methods). Precultures were grown at 30 °C with shaking at 250 rpm for 18–20 h in synthetic-defined (SD) medium with appropriate dropout (DO) supplements (Clontech). To induce gene expression, the precultures were washed and inoculated into SD Gal/Raf medium with appropriate DO supplements (Clontech) to a starting optical density of 0.25 on day 1. The cultures were further incubated for 24 h before addition of 1 mM methionine on day 2 and incubated for 48 h. For CD treatment, M β CD was added at a final concentration of 5 or 25 mM; and α CD, β CD, or γ CD were added at a final concentration of 5 mM.

Metabolite Extraction for GC-MS. Organic extracts of yeast cells or spent medium were prepared for identification and quantification of sapogenins. Cells were separated from the culture and lysed with equal volumes of 40% (wt/vol) KOH and 50% (vol/vol) ethanol by boiling for 10 min. The supernatant corresponded to the spent medium. The lysed cells and spent medium were extracted thrice with hexane. The organic phases were pooled,

vaporized to dryness, and trimethylsilylated for GC-MS analysis (*SI Materials and Methods*).

Metabolite Extraction for LC-MS. Yeast cells were lysed with the YeastBuster protein extraction reagent (Novagen) and extracted thrice with ethyl acetate. The organic phases were pooled, dried, and dissolved in 100 μ L water for LC-MS analysis (*SI Materials and Methods*).

NMR Analysis. The 3-O-Glc-echinocystic acid was purified from strain TM44 and its structure determined by NMR analysis (*SI Materials and Methods*).

- Augustin JM, Kuzina V, Andersen SB, Bak S (2011) Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry* 72(6):435–457.
- Osborn A, Goss RJM, Field RA (2011) The saponins—Polar isoprenoids with important and diverse biological activities. *Nat Prod Rep* 28(7):1261–1268.
- Phillips DR, Rasbery JM, Bartel B, Matsuda SP (2006) Biosynthetic diversity in plant triterpene cyclization. *Curr Opin Plant Biol* 9(3):305–314.
- Abe I (2007) Enzymatic synthesis of cyclic triterpenes. *Nat Prod Rep* 24(6):1311–1331.
- Pollier J, Goossens A (2012) Oleanolic acid. *Phytochemistry* 77(5):10–15.
- Salvador JAR, Moreira VM, Gonçalves BMF, Leal AS, Jing Y (2012) Ursane-type pentacyclic triterpenoids as useful platforms to discover anticancer drugs. *Nat Prod Rep* 29(12):1463–1479.
- Sporn MB, et al. (2011) New synthetic triterpenoids: Potent agents for prevention and treatment of tissue injury caused by inflammatory and oxidative stress. *J Nat Prod* 74(3):537–545.
- Ashour ML, Wink M (2011) Genus *Bupleurum*: A review of its phytochemistry, pharmacology and modes of action. *J Pharm Pharmacol* 63(3):305–321.
- Kim YS, et al. (2011) Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. *Planta* 233(2):343–355.
- Sui C, et al. (2011) Transcriptome analysis of *Bupleurum chinense* focusing on genes involved in the biosynthesis of saikosaponins. *BMC Genomics* 12:539.
- Seki H, et al. (2008) Licorice β -amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc Natl Acad Sci USA* 105(37):14204–14209.
- Geisler K, et al. (2013) Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc Natl Acad Sci USA* 110(35):E3360–E3367.
- Fukushima EO, et al. (2013) Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant Cell Physiol* 54(5):740–749.
- Shibuya M, et al. (2006) Identification of β -amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J* 273(5):948–959.
- Fukushima EO, et al. (2011) CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol* 52(12):2050–2061.
- Carelli M, et al. (2011) *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell* 23(8):3070–3081.
- Huang L, et al. (2012) Molecular characterization of the pentacyclic triterpenoid biosynthetic pathway in *Catharanthus roseus*. *Planta* 236(5):1571–1581 [Note added: *Planta* 236(5):1583].
- Seki H, et al. (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* 23(11):4112–4123.
- Han J-Y, Hwang H-S, Choi S-W, Kim H-J, Choi Y-E (2012) Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol* 53(9):1535–1545.
- Han J-Y, Kim H-J, Kwon Y-S, Choi Y-E (2011) The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol* 52(12):2062–2073.
- Han J-Y, Kim M-J, Ban Y-W, Hwang H-S, Choi Y-E (2013) The involvement of β -amyrin 28-oxidase (CYP716A52v2) in oleanane-type ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol* 54(12):2034–2046.
- Krokida A, et al. (2013) A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol* 200(3):675–690.
- Moses T, Pollier J, Thevelein JM, Goossens A (2013) Bioengineering of plant (tri)terpenoids: From metabolic engineering of plants to synthetic biology *in vivo* and *in vitro*. *New Phytol* 200(1):27–43.
- Kirby J, Romanini DW, Paradise EM, Keasling JD (2008) Engineering triterpene production in *Saccharomyces cerevisiae- β -amyrin synthase from *Artemisia annua*. *FEBS J* 275(8):1852–1859.*
- Madsen KM, et al. (2011) Linking genotype and phenotype of *Saccharomyces cerevisiae* strains reveals metabolic engineering targets and leads to triterpene hyper-producers. *PLoS ONE* 6(3):e14763.
- Breyne P, et al. (2003) Quantitative cDNA-AFLP analysis for genome-wide expression studies. *Mol Genet Genomics* 269(2):173–179.
- Pollier J, González-Guzmán M, Ardiles-Díaz W, Geelen D, Goossens A (2011) An integrated PCR colony hybridization approach to screen cDNA libraries for full-length coding sequences. *PLoS ONE* 6(9):e24978.
- Suzuki H, Achnine L, Xu R, Matsuda SPT, Dixon RA (2002) A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. *Plant J* 32(6):1033–1048.
- Hayashi H, et al. (2001) Cloning and characterization of a cDNA encoding β -amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biol Pharm Bull* 24(8):912–916.
- Herrera JBR, Bartel B, Wilson WK, Matsuda SPT (1998) Cloning and characterization of the *Arabidopsis thaliana* lupeol synthase gene. *Phytochemistry* 49(7):1905–1911.
- Kushiro T, et al. (2006) Stereochemical course in water addition during LUP1-catalyzed triterpene cyclization. *Org Lett* 8(24):5589–5592.
- Kim OT, et al. (2009) Characterization of a dammarenediol synthase in *Centella asiatica* (L.) Urban. *Plant Physiol Biochem* 47(11–12):998–1002.
- Laursen T, Jensen K, Møller BL (2011) Conformational changes of the NADPH-dependent cytochrome P450 reductase in the course of electron transfer to cytochromes P450. *Biochim Biophys Acta* 1814(1):132–138.
- Reed JR, Backes WL (2012) Formation of P450-P450 complexes and their effect on P450 function. *Pharmacol Ther* 133(3):299–310.
- Paddon CJ, et al. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496(7446):528–532.
- Gavira C, et al. (2013) Challenges and pitfalls of P450-dependent (+)-valencene bio-conversion by *Saccharomyces cerevisiae*. *Metab Eng* 18(7):25–35.
- Walczak RJ, Dickens ML, Priestley ND, Strohl WR (1999) Purification, properties, and characterization of recombinant *Streptomyces* sp. strain C5 DoxA, a cytochrome P-450 catalyzing multiple steps in doxorubicin biosynthesis. *J Bacteriol* 181(1):298–304.
- Stella VJ, He Q (2008) Cyclodextrins. *Toxicol Pathol* 36(1):30–42.
- Christian AE, Haynes MP, Phillips MC, Rothblat GH (1997) Use of cyclodextrins for manipulating cellular cholesterol content. *J Lipid Res* 38(11):2264–2272.
- Pollier J, Moses T, Goossens A (2011) Combinatorial biosynthesis in plants: A (p)review on its potential and future exploitation. *Nat Prod Rep* 28(12):1897–1916.
- de Felipe P, et al. (2006) *E unum pluribus*: Multiple proteins from a self-processing polyprotein. *Trends Biotechnol* 24(2):68–75.
- Augustin JM, et al. (2012) UDP-glycosyltransferases from the UGT73C subfamily in *Barbarea vulgaris* catalyze saponin 3-O-glucosylation in saponin-mediated insect resistance. *Plant Physiol* 160(4):1881–1895.
- Sabater-Jara AB, Pedreño MA (2013) Use of β -cyclodextrins to enhance phytosterol production in cell suspension cultures of carrot (*Daucus carota* L.). *Plant Cell Tissue Organ Cult* 114(2):249–258.