

Microbial Synthesis of the Forskolin Precursor Manoyl Oxide in an Enantiomerically Pure Form

Morten T. Nielsen,^{a,b} Johan Andersen Ranberg,^b Ulla Christensen,^a Hanne Bjerre Christensen,^a Scott J. Harrison,^a Carl Erik Olsen,^b Björn Hamberger,^b Birger Lindberg Møller,^b Morten H. H. Nørholm^a

Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark^a; Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg, Denmark^b

Forskolin is a promising medicinal compound belonging to a plethora of specialized plant metabolites that constitute a rich source of bioactive high-value compounds. A major obstacle for exploitation of plant metabolites is that they often are produced in small amounts and in plants difficult to cultivate. This may result in insufficient and unreliable supply leading to fluctuating and high sales prices. Hence, substantial efforts and resources have been invested in developing sustainable and reliable supply routes based on microbial cell factories. Here, we report microbial synthesis of (13*R*)-manoyl oxide, a proposed intermediate in the biosynthesis of forskolin and other medically important labdane-type terpenoids. Process optimization enabled synthesis of enantiomerically pure (13*R*)-manoyl oxide as the sole metabolite, providing a pure compound in just two steps with a yield of 10 mg/liter. The work presented here demonstrates the value of a standardized bioengineering pipeline and the large potential of microbial cell factories as sources for sustainable synthesis of complex biochemicals.

Terpenoids represent a highly diverse class of natural products with a multitude of applications in modern society ranging from renewable fuels to structurally complex medicinal compounds (1). For example, they constitute the active compounds in several medicinal plant species, including paclitaxel (*Taxus brevifolia*), forskolin (*Coleus forskohlii*), and artemisinin (*Artemisia annua*). Although several terpenoid-derived natural products are currently in or have passed clinical trial for treatment of a number of conditions (2), sufficient supply of the molecules remains an obstacle. Purification from naturally producing plants has proven infeasible in many cases due to low yield, while establishment of heterologous production systems has been hampered by insufficient insight into the biosynthetic routes. However, the advent and rapid development of whole-genome and transcriptome sequencing have paved the way for several recent breakthroughs. Particularly, tissue-specific metabolomics combined with RNA deep sequencing is an exceptionally promising approach for gene discovery in nonmodel organisms (3, 4).

Our interest is focused on the diterpene forskolin produced by *C. forskohlii*. Forskolin is a promising pharmaceutical that has shown potential as treatment against a wide range of diseases (5–9). Forskolin is currently approved for treatment of glaucoma and heart failure, while clinical trials against several forms of cancer are ongoing; the compound has been tested in over 800 independent assays, and in 76 cases forskolin was reported to be potently active (<2 micromolar), including activity as a D3 dopamine and relaxin receptor agonist, as an allosteric activator of Mas-related G-protein-coupled receptor X1, and as an antagonist of the androgen receptor (<http://pubchem.ncbi.nlm.nih.gov>). Finally, forskolin is extensively used in basic research as an inhibitor of protein kinases (more than 50,000 publications can be found in Web of Science [Thomson Reuters]). Recently, the combined use of metabolomics and RNA deep sequencing were successfully applied to *C. forskohlii* root cells, resulting in the identification of two diterpene synthases (diTPSs), CfTPS2 and CfTPS3, that catalyze the formation of (13*R*)-manoyl oxide (13*R*-MO) from geranylgeranyl diphosphate (GGPP) (10). 13*R*-MO is proposed to be

the first dedicated intermediate in forskolin biosynthesis (11). With this first biosynthetic knowledge in hand, establishment of a microbial system for further elucidation of the biosynthetic route as well as production of forskolin comes within reach.

From a microbial engineering perspective, terpenoids constitute an attractive class of compounds, as they are all derived from a single precursor, isopentenyl pyrophosphate (IPP), and its isomer, dimethylallyl pyrophosphate (DMAPP). Substantial efforts have therefore been made in increasing IPP supply as a driver for efficient terpenoid synthesis in microbes (12). IPP can be synthesized through two biochemical routes, the mevalonate (MEV) and the 1-deoxy-D-xylulose 5-phosphate (DXP) pathways, with the MEV pathway being prevalent in eukaryotes and archaea, while the DXP pathway dominates in eubacteria. While both routes have been subjected to numerous engineering efforts, the MEV pathway has been most successfully applied in *Escherichia coli* (12, 13). One consensus minimal set of MEV pathway enzymes required for boosting the IPP/DMAPP supply can be formulated as mevalonate kinase (e.g., *Saccharomyces cerevisiae* ERG12 [ScERG12]), phosphomevalonate kinase (e.g., ScERG8), mevalonate pyrophosphate decarboxylase (e.g., ScMVD1), and IPP isomerase (e.g., *E. coli* Idi2 [EIdi2]). Biosynthesis of diterpenes in *E. coli* faces the additional challenge that conversion of IPP to the general diterpene precursor, GGPP, is very inefficient in wild-type *E. coli* strains and therefore requires introduction of a heterologous GGPP synthase (14, 15). As previ-

Received 11 July 2014 Accepted 11 September 2014

Published ahead of print 19 September 2014

Editor: R. M. Kelly

Address correspondence to Morten H. H. Nørholm, morno@biosustain.dtu.dk.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.02301-14>.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.02301-14

ously shown for the C₁₅ sesquiterpene amorphadiene, a substantial increase in product yield can be gained by combining a heterologous farnesyl diphosphate (FPP) synthase with the minimal consensus set of MEV pathway genes (16). Likewise, a high-level supply of GGPP in *E. coli* requires the coordinated expression of a minimum of five genes: four for producing the precursor IPP/DMAPP (for example, *S. cerevisiae* *ERG12*, *ERG8*, and *MVD* and *E. coli* *idi2*) and one for a GGPP synthase. As this consensus set is independent of the target diterpene, assembling all five genes on a single plasmid would substantially simplify experimental designs such as strain screenings. The Keasling group has reported the construction and verification of a single plasmid for synthesis of farnesyl diphosphate (13). Previous attempts of adaptation of this strategy for synthesis of GGPP unfortunately resulted in a non-functional plasmid (16). We have recently described a set of design principles for standardizing molecular cloning and formulation of a cloning pipeline to simplify multigene expression in *E. coli* (17). Here, we take advantage of this bioengineering pipeline to assemble and validate a single plasmid enabling GGPP synthesis from mevalonate, providing a versatile platform for diterpene biosynthesis in *E. coli*. By utilization of the platform, we achieve *in vivo* microbial synthesis of the proposed forskolin precursor 13R-MO. Through a limited series of iterative optimization steps, stereospecific synthesis of 13R-MO is achieved and yields are increased from trace amounts to 10 mg/liter in microscale cultures. The final process yields 13R-MO in nuclear magnetic resonance (NMR) quality following a single, simple step for removal of impurities.

MATERIALS AND METHODS

Bacterial strains and media. *Escherichia coli* strain NEB5 α was used for cloning and propagation of plasmids. *E. coli* strain BL21(DE3) was used for radioactive labeling studies and strain KRX (Promega, Madison, WI, USA) for production of 13R-MO. Following transformations, bacteria were propagated on Luria-Bertani (LB) agar plates supplemented with ampicillin (100 μ g/ml), spectinomycin (50 μ g/ml), kanamycin (50 μ g/ml), or combinations thereof (ampicillin, 50 μ g/ml; spectinomycin, 25 μ g/ml; kanamycin, 25 μ g/ml) when required. Liquid cultures for plasmid preparation were based on 2 $\times$ YT media, whereas precultures for assays were prepared in LB media using antibiotic concentrations as described above. The PASM-5052 defined media used for expression assays were prepared as described by Studier (18), except for the exchange of selenomethionine with methionine. For long-term incubations (more than 24 h), ampicillin was replaced by carbenicillin (50 μ g/ml) as previously described (19).

PCR and uracil excision cloning. Uracil excision-compatible PCR products were produced with 30 cycles in 50- μ l reaction mixtures using proofreading PfuX7 polymerase as previously described (20). Fragments of *ScMVD1* (NCBI Gene ID, 855779), *ScERG8* (NCBI Gene ID, 855260), and *ScERG12* (NCBI Gene ID, 855248) were amplified from *S. cerevisiae* CEN.PK113-7D genomic DNA. *E. coli* *idi2* (*Ecidi2*) (NC_000913) fragments were amplified from *E. coli* K-12 MG1655 genomic DNA. *C. forskohlii* *TPS2* (*CfTPS2*) (KF444507) and *CfTPS3* (KF444508) were amplified from previously described constructs (10). *Abies grandis* GGPP (*AgGGPP*) (AF425235.2) was codon optimized for *E. coli* and synthetically synthesized (Genscript, Piscataway, NJ, USA). USER Fusion DNA assembly was performed essentially as described by Geu-Flores et al. (21). Purified PCR products of inserts and vector were mixed in a total insert-to-vector ratio of 6:1 (volume). When several DNA fragments were cloned simultaneously, use of equimolar amounts of each fragment was approximated by agarose gel analysis. The reaction mixture was buffered with 5 $\times$ Phusion HF buffer (Invitrogen), and 1 U of USER enzyme mix (New England Biolabs, Ipswich, MA, USA) was added. The reaction mixture was incubated for 20 min at 37°C, followed by 20 min at 25°C before

transformation of chemically competent *E. coli* cells. All oligonucleotides and synthetic genes used in this study are summarized in Table S1 in the supplemental material.

Plasmid construction. The principle behind the molecular cloning of open reading frames (ORFs) is illustrated in Fig. 1 and was described previously in detail in reference 17. Briefly, ORFs were PCR amplified without start or stop codons and flanked by 9-bp linkers mediating standardized cloning into a pETDuet-1 (Merck KGaA, Darmstadt, Germany) backbone entry vector. Transcriptional units containing a T7 promoter, a consensus Shine-Dalgarno sequence, and the entry clone ORFs were reamplified using generic oligonucleotides with complementary linkers facilitating multifragment assemblies. Vector backbones used in multifragment assemblies were likewise amplified with generic oligonucleotides. All plasmids containing ORFs were verified by sequencing.

[³⁵S]Methionine labeling of proteins. Transcription and translation of ORFs used in this study were confirmed *in vivo* by the rifampin blocking technique and ³⁵S labeling (22). Briefly, precultures of BL21(DE3) cells carrying the relevant plasmids were subcultured (1:50) into 1 ml chemically defined rich medium lacking methionine (18). Following induction of T7-dependent transcription by isopropyl- β -D-1-thiogalactopyranoside (IPTG) and addition of rifampin and [³⁵S]methionine, cells were lysed and samples applied on precast 4 to 20% SDS-PAGE (Expedeon, San Diego, CA, USA). Upon completion of the electrophoresis, the gels were dried and the labeled proteins were visualized using a CyclonePl scanner (PerkinElmer, Waltham, MA, USA).

Assay conditions. For chromatographic assays, bacteria from single-plate colonies were inoculated into precultures of 500 μ l LB and grown 16 h at 37°C with 300 rpm shaking. For assaying manoyl oxide production, 40 μ l LB preculture was inoculated into 2 ml freshly prepared PASM-5052 defined medium (1:50 dilution) and grown in 10-ml microtiter plates at 37°C and 300 rpm to an optical density (OD) of 0.3. At this point, synthesis of recombinant proteins was induced by addition of rhamnose (1 mM) and IPTG (0.4 mM). Rhamnose stimulates synthesis of the T7 RNA polymerase in strain KRX, whereas IPTG alleviates the *lac* repressor from the *lac* operator present in all gene expression plasmids. Mevalonolactone (9 mM) was added at this point, when required. All metabolite concentrations given represent the final concentration in the assay. After induction, cultures were incubated at either 25°C and 150 rpm for 16 h or 16°C and 150 rpm for 112 h. Assays were performed with two independent colonies each analyzed in at least three replicates.

Extraction procedure. Metabolite extraction was performed by the following procedure: 500 μ l of bacterial culture was transferred to a 2-ml high-performance liquid chromatography (HPLC) vial (Mikrolab, Aarhus, Denmark), prior to application of an overlay of 500 μ l hexane containing internal standard (1-eicosene, 1 mg/liter; Sigma-Aldrich). Extractions were carried out at room temperature with shaking (240 rpm on a table shaker [IKA-KS 130 basic]). After 1 or 16 h of extraction, 300 μ l of the hexane phase was transferred to clean HPLC vials and subjected to gas chromatography analysis. Cultures for purification of 13R-MO were performed as described above, except that the total volumes (2 ml) of 240 cultures were pooled in a 2-liter Erlenmeyer flask and extracted with 500 ml of hexane.

Analytical method. Gas chromatography/mass spectrometry (GC-MS) analysis was performed as previously described (23). Extracts were analyzed on a Shimadzu GCMS-QP2010 Ultra instrument with an Agilent HP-5MS column (30 m by 0.25 mm [inner diameter], 0.25- μ m film thickness) installed. One microliter of extract was injected in the injection port set at 250°C in splitless mode. The gas flow was set to 2.2 ml/min and the purge flow to 1 ml/min, using helium as carrier gas. The GC program was set to the following: hold for 1 min at 100°C, ramp 10°C/min to 250°C, followed by a second ramp with 20°C/min to a final temperature of 310°C, which was held for 2 min. Ion source and interface temperature were set to 230°C and 150°C, respectively, and the solvent cutoff was set for 8 min. Mass analyzer was set in scan mode from 50 Hz to 350 Hz and a scan rate of 625 u/s. Data analysis was carried out with the GC-MS solution version

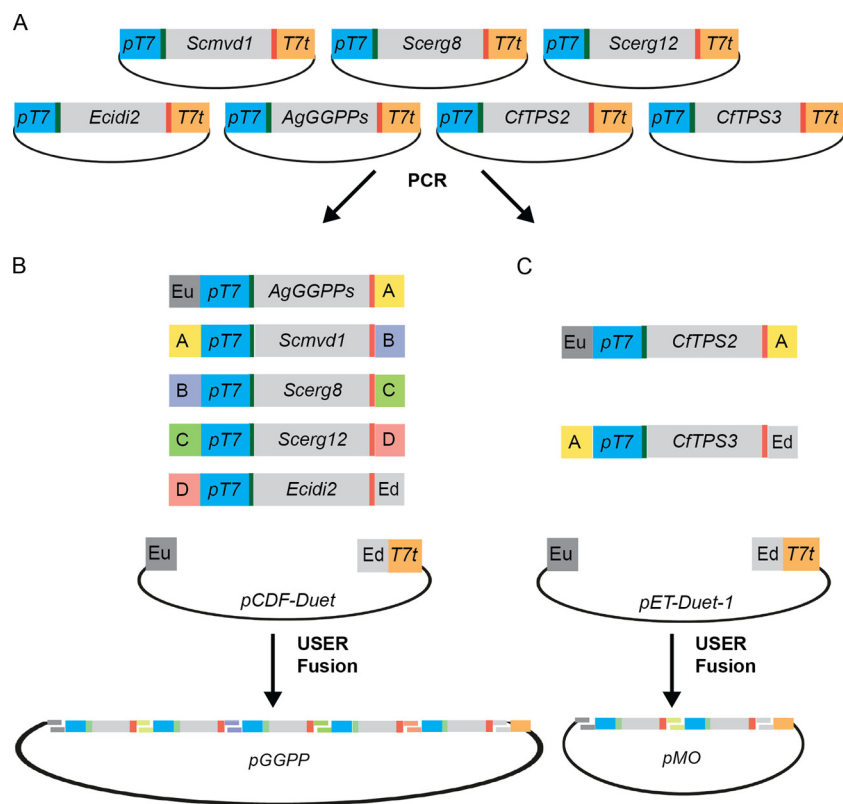


FIG 1 Schematic overview of the gene assembly procedure. (A) Open reading frames are cloned individually into standardized entry expression vectors. (B and C) Assembly of pGGPP (B) and pMO (C). Transcriptional units consisting of a T7 promoter, a consensus Shine-Dalgarno sequence, and the open reading frame are PCR amplified with generic oligonucleotides that anneal outside the different ORFs in the flanking regions common to all entry plasmids. The generic oligonucleotides facilitate specific assembly through defined linker regions (regions A to D, Eu, and Ed). Transcriptional units are assembled in a single cloning reaction by USER Fusion. Start and stop codons of open reading frames are indicated by green and red bars, respectively. The T7 promoter and terminator are denoted pT7 and T7t, respectively.

2.50 SU3 (Shimadzu Corporation). Compound identification was done by comparison to reference spectra databases (Wiley Registry of Mass Spectral Data, 8th ed., July 2006 [John Wiley & Sons]) and comparison of retention indices found in reference 23. 13R-MO was quantified relatively by comparison of the TIC peak area normalized to the peak area of internal standard (1 mg/liter 1-eicosene). Absolute quantification was performed by comparison of normalized peak areas as above with the standard curve containing purified 13R-MO. The standard curve was fitted with linear fit after log-log transformation.

Single-step purification of 13R-MO. The hexane extract was concentrated by rotor evaporation (Buchi, Switzerland) with settings at 35°C and 220 mbar. The concentrated sample was subsequently purified by solid-phase extraction (Dual Layer Florisil/Na₂SO₄ 6 ml PP SPE Tube; Supelco Analytical, Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's protocol. In brief, the sample was applied to the column and washed with hexane before removal of residual solvent in a gentle nitrogen gas stream. The sample was eluted in deuterated chloroform (CDCl₃) and analyzed by NMR for structural verification.

NMR analysis and quantification. NMR analysis was performed using a 400 MHz Bruker Avance 400 spectrometer, and the structure was confirmed by comparison of the observed ¹H and ¹³C NMR spectra with literature data (Tables 1 and 2) (24, 25). CDCl₃ was used as a solvent. The amount of 13R-MO present in the NMR sample was determined by integration after adding 1.00 mg of dimethyl formamide (DMF). The integrals of one of the vinyl protons and one of the DMF methyl groups were used for the calculations. The NMR sample was subsequently used for creating a GC-MS standard curve.

TABLE 1 ¹³ C NMR spectra showing chemical shifts of annotated carbons of 13R-MO			
No. of carbons	¹³ C NMR chemical shift (ppm) from:		M
	Literature ^a	This work	
1	39.12	39.018	t
2	18.62	18.564	t
3	42.23	42.117	t
4	33.32	33.238	s
5	56.55	56.432	d
6	20.03	19.956	t
7	43.4	43.267	t
8	75.09	75.091	s
9	55.6	55.64	d
10	37.11	37.011	s
11	15.41	15.304	t
12	35.73	35.714	t
13	73.25	73.208	s
14	148.07	147.976	d
15	110.21	110.255	t
16	28.71	28.556	q
17	25.58	25.524	q
18	33.4	33.362	q
19	21.39	21.348	q
20	15.41	15.421	q

^a Reference spectra obtained from González et al. (24).

TABLE 2 ^1H -NMR spectra showing chemical shifts of ^1H in 13R-MO and coupling constants^a

Carbon environment type	^1H chemical shift (ppm) from:		Coupling constant
	Literature ^a	This work	
—CH ₃	0.83	0.806	s
—CH ₃	0.83	0.816	s
—CH ₃	0.9	0.877	s
—CH ₃	1.31	1.294	s
—CH ₃	1.33	1.313	s
=CH	4.94	4.938	q, $J = 11$ & 2 Hz
=CH	5.16	5.164	q, $J = 17$ & 2 Hz
=CH ₂	5.92	5.894	q, $J = 17$ & 11 Hz

^a Reference spectra obtained from González et al. (24).

RESULTS

Assembly of a two-plasmid system for producing 13R-MO from mevalonolactone. Open reading frames of MEV pathway genes from *S. cerevisiae* (*ScERG12*, *ScERG8*, and *ScMVD1*), IPP isomerase from *E. coli* (*Ecidi2*), GGPP synthase from *Abies grandis* (*AgGGPPs*) (26), and diterpene synthases from *C. forskohlii* (*CfTPS2* and *CfTPS3*) (10) were initially cloned into an entry expression vector allowing assessment of proper expression and a standardized multigene assembly (Fig. 1A). All plant-derived ORFs were codon optimized for *E. coli* and truncated to remove plastid-targeting signals according to signal peptide predictions by the ChloroP algorithm (27). Single-ORF expression

plasmids were transformed into BL21(DE3), and transcription and translation were verified by [^{35}S]methionine labeling in the presence of rifampin as previously described (28) (see Fig. S1 in the supplemental material). From these entry clones, *ScMVD1*, *ScERG8*, *ScERG12*, *Ecidi2*, and *AgGGPPs* were subsequently assembled into the plasmid pGGPP for GGPP synthesis, while *CfTPS2* and *CfTPS3* were assembled into pMO for conversion of GGPP into 13R-MO (Fig. 1B). In the setup employed here, the T7 promoter and a consensus Shine-Dalgarno sequence were maintained as transcriptional units in front of every ORF encoding the entry fragment in the assemblies (Fig. 1B).

Cotransformation of pGGPP and pMO enables *in vivo* synthesis of 13R-MO. In order to establish the system and test the functionality of the plasmids, pMO was cotransformed with either a plasmid harboring the *AgGGPP* synthase alone (pSyn), pGGPP, or an empty vector control (pCDFDuet-1) into the *E. coli* strain KRX. Subsequently, assays were performed in a defined rich medium based on lactose and glycerol (18). In order to test for pGGPP-mediated conversion of mevalonolactone to GGPP, cultures were split in two and mevalonolactone (9 mM) was added as a supplement to one culture containing pGGPP at the time of induction. After 16 h of induction at 25°C, nonpolar metabolites were extracted with hexane and analyzed with gas chromatography coupled to mass spectrometry (GC-MS). Consistent with previous findings (15), introduction of the *AgGGPP* synthase, either individually in pSyn or as part of pGGPP, enabled synthesis of trace amounts of two diterpene-related compounds, compounds 1 and 2 (Fig. 2A and B). Supplementation of mevalonolactone to

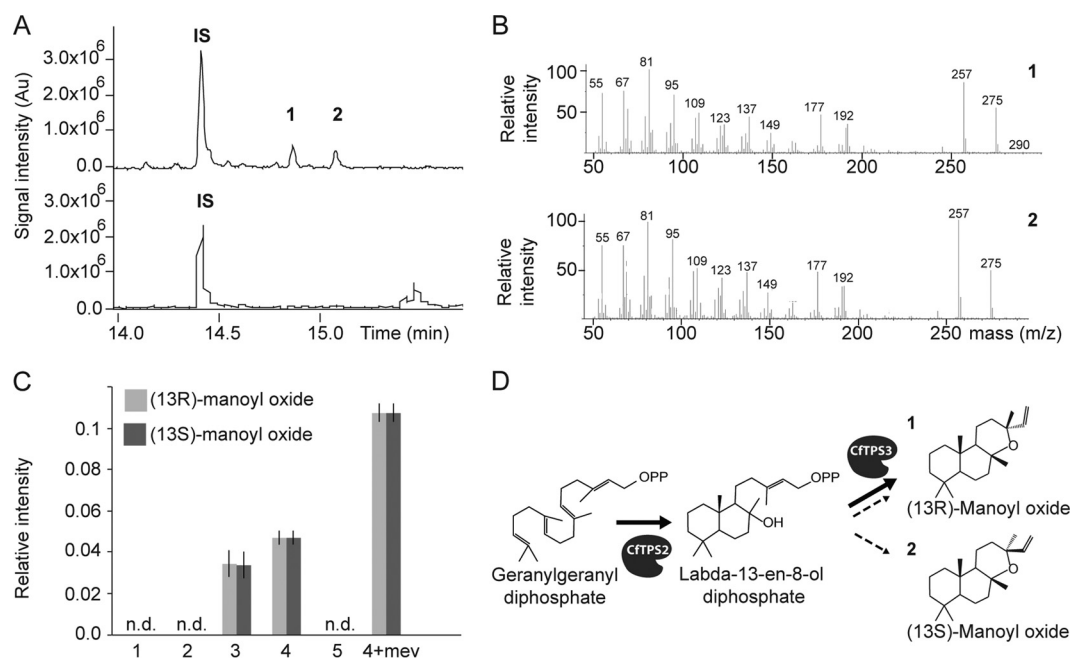


FIG 2 *In vivo* synthesis of two manoyl oxide isomers in *E. coli*. (A) Total ion GC-MS chromatograms (TIC) of hexane extract from *E. coli* transformed with either pGGPP and pMO (top) or empty vectors (bottom). IS, internal standard (1 mg/liter 1-eicosene). (B) Mass spectra of metabolites 1 and 2 identified as 13R- and 13S-MO, respectively, based on comparison to the Wiley Registry of Mass Spectral Data (see Fig. S2 in the supplemental material). (C) Relative quantification of 13R- and 13S-MO synthesis in *E. coli*. Yields are normalized to a 1-mg/liter 1-eicosene internal standard. n.d., neither 13R- nor 13S-MO was detectable. Numbers on the x axis refer to the combination of plasmids present in the analyzed strain, as follows: 1, pET-Duet/pCDF-Duet; 2, pMO/pCDF-Duet; 3, pMO/pSyn; 4, pMO/pGGPP; 5, pET-Duet/pGGPP; 4+mev, pMO/pGGPP supplemented with 9 mM mevalonolactone. Data represent the averages for four independent replicates with standard deviations. (D) Biosynthetic route to 13R-MO. Solid arrows indicate enzymatic reactions. The enzyme catalyzing each step is depicted next to the reaction arrow. The dashed arrows indicate the intrinsic degradation of labda-13-en-8-ol diphosphate into equimolar amounts of 13R- and 13S-MO occurring in the absence of *CfTPS3*. Phosphate groups in the chemical structures are abbreviated as "P".

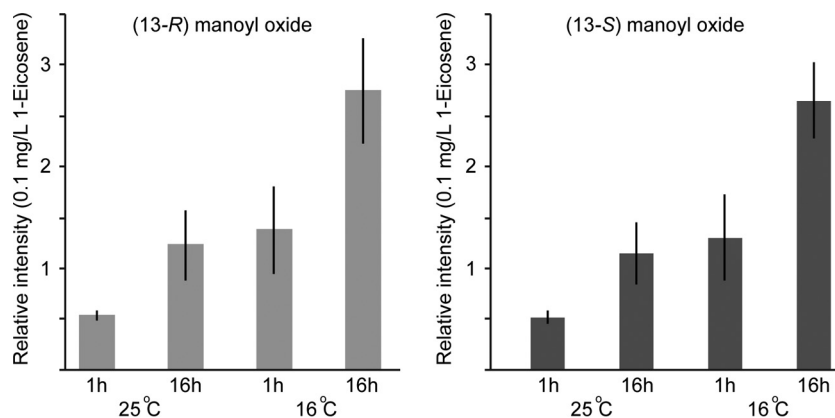


FIG 3 Impact of iterative improvements of culture conditions and extraction time on manoyl oxide yields. *E. coli* transformed with pGGPP and pMO was cultivated at either 25°C for 16 h postinduction or at 16°C for 112 h postinduction. Cultures were extracted with hexane for either 1 or 16 h. Graphs depict the average yields of the two manoyl oxide isomers from two independent colonies each analyzed in four independent replicates. Error bars denote standard deviations.

pGGPP-containing cultures resulted in a modest, yet significant, increase in the yields of both compounds (Fig. 2C). The compounds were tentatively identified as 13R-MO and 13S-MO, respectively, by comparison of mass spectra with the Wiley Registry of Mass Spectral Data, 8th edition, database (see Fig. S2 in the supplemental material), as well as previously published data (23). Absolute identification of 13R-MO by NMR was achieved using the final process described below (Table 1; see also Fig. S3 in the supplemental material). To our knowledge, this is the first report of *in vivo* biosynthesis of 13R-MO, the proposed precursor for forskolin, in microbial systems and therefore an important milestone in efforts toward establishment of a microbial cell factory for forskolin synthesis. Moreover, the mevalonate-mediated increase in diterpene synthesis strongly indicates that all enzymes required for conversion of mevalonate to GGPP are functionally expressed from pGGPP and this plasmid might therefore serve as a future platform for diterpene synthesis in *E. coli*. The observed cooccurrence of 13S-MO with 13R-MO has previously been associated with insufficient activity of the class I diTPS *CfTPS3* (10) (Fig. 2) and therefore indicated that the current process is suboptimal.

Process optimization facilitates increased yield. Several studies have demonstrated that substantial improvements of microbial production systems can be achieved using a simple set of process optimization steps, including alteration of culture parameters and recovery procedures (13, 16, 29). Although no clear consensus protocol has emerged, several groups reported success with extended cultivation at low temperatures. Moreover, recovery by hexane extraction has previously been demonstrated to yield highly pure diterpene preparations from *E. coli* (16). A comparison between short incubation at medium temperature (16 h, 25°C) and longer incubation at low temperature (112 h, 16°C) as well as between short (1-h) and extended (16-h) hexane extraction was undertaken. For this experiment, *E. coli* KRX was transformed with pGGPP and pMO and all cultures were supplemented with mevalonolactone (9 mM) upon induction. Incubation as well as extraction time had a significant impact on diterpene yield compared to the initial protocol, contributing a 2-fold yield increase each (Fig. 3). These yield improvements proved to be cumulative when combined, resulting in an approximately 5-fold total in-

crease over the initial protocol. Under these experimental conditions, the process, however, still resulted in formation of both 13S- and 13R-MO in approximately equal amounts. At this stage, the expression level of *CfTPS3* from pMO was examined in detail. Remarkably, when pMO was transformed into BL21(DE3) and subjected to [³⁵S]methionine labeling, only *CfTPS2* expression was detected (see Fig. S1 in the supplemental material). Since it was already established that both *CfTPS2* and *CfTPS3* could be transcribed and translated in *E. coli* when expressed alone (see Fig. S1 in the supplemental material), these findings indicated that transcription from the second T7 promoter is impeded in pMO. To test this hypothesis, we established a three-plasmid system consisting of pGGPP, p*CfTPS2*, and p*CfTPS3*. Indeed, when the three plasmids were cotransformed into *E. coli* KRX cells and cultivated in consistence with the optimal parameters identified above, biosynthesis of enantiomerically pure 13R-MO was obtained (Fig. 4A). Moreover, the flux through the pathway was dramatically increased, resulting in a >2,000-fold yield increase over the initial protocol.

Recovery, structure validation, and absolute quantification.

An advantage of the *E. coli*-based biosynthesis of diterpenes is high purity and therefore simple downstream processing. Indeed, with the hexane-based extraction utilized in this study, merely three compounds were observed with GC-MS in addition to 13R-MO. Two of these (compounds 4 and 5) could be attributed to impurities arising from the instrument (data not shown) and are thus experimental artifacts rather than genuine compounds present in the extract. The one remaining metabolite, compound 3, was identified as indole by comparison of retention time and mass spectra with an authentic standard (Fig. 4B). Indole is known to be a signaling molecule in *E. coli*, with substantial influence on metabolism and membrane potential (30, 31), and has been reported to inhibit terpene synthesis in *E. coli* (29). Thus, elimination of indole formation might not only simplify recovery but also increase yield. Recently, it was discovered that *E. coli* very efficiently converts extracellular tryptophan to indole (32). The defined medium used in this study is supplemented with high concentrations (10 mM) of free amino acids, including tryptophan, to support high-level protein synthesis. Adjustment of the medium composition by omission of tryptophan eliminated indole accumulation,

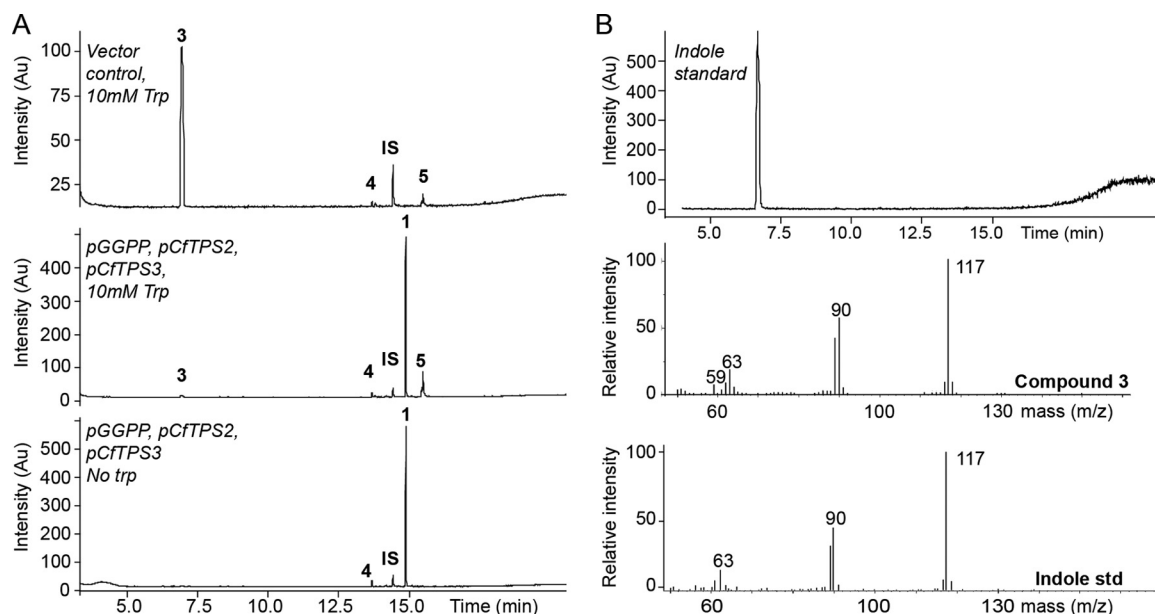


FIG 4 Identification and elimination of indole in *E. coli* extracts. (A) Total ion chromatograms from GC-MS analysis of extracts from *E. coli*. The plasmid combination in question is listed in the top left corner of each chromatogram along with the initial concentration of tryptophan (Trp or trp) in the media. The metabolites observed were identified as 13R-MO (1), indole (3), and impurities from the instrument (4 and 5). IS, internal standard (1 mg/liter 1-eicosene). (B) Total ion chromatogram of authentic indole standard (top) as well as mass spectra of compound 3 (middle) and indole standard (bottom). Based on a comparison of retention time and mass spectra with authentic standard data, compound 3 was identified as indole.

resulting in an extract with 13R-MO as the sole metabolite observed by GC-MS analysis (Fig. 4A). Importantly, the overall yield of 13R-MO was unaffected. Following a simple removal of impurities by solid-phase extraction, the recovered 13R-MO was sufficiently pure for absolute identification by NMR analysis (Table 1). The yield of the established process was quantified absolutely to 10 mg/liter by GC-MS using a standard curve of purified 13R-MO as a reference. Typical dephosphorylation products (and results of endogenous phosphatase activity) of the pathway intermediates GGPP and 8-hydroxy-copalol diphosphate are geranyl-geraniol and dihydroxy-copalol. Under the conditions used, both compounds are detectable with our analytical setup, yet no traces were found, indicating that those pathway intermediates do not accumulate.

DISCUSSION

Biosynthesis of high-value terpenoids in microbial systems is emerging as an adequate and sustainable production platform, with several studies reporting yields of diterpene olefins in the scale of grams per liter (29, 33–35; reviewed in reference 36). A uniting feature of these studies is that yields in unoptimized microscale cultures are meager, typical in the low milligram per liter range: 0.15 mg/liter (33), 2.6 mg/liter (16), 0.005 mg/liter (15), or 4 mg/liter (35). A highly successful route for yield increase in microscale has been improvement of the precursor pool by either introduction of the mevalonate pathway from yeast (13, 16, 35) or modulation of the endogenous methylerythritol 4-phosphate (MEP) pathway by promoter engineering (29) or stacking of gene copies (33). Moreover, culture scale-up into controlled bioreactors has proven essential to obtaining gram scale yields (34, 35). Our aim is to use microbial systems for elucidation of biosynthetic pathways to high-value diterpenoids and simultaneously establish

a microbial cell factory. Since gene discovery is likely to require a substantial screening effort, it was crucial for us to develop a process compatible with microscale cultures. We describe here the assembly of a single plasmid harboring five ORFs encoding enzymes sufficient for efficient GGPP synthesis from mevalonate in *E. coli*. Maintaining all these enzymatic functions on one plasmid substantially increases the flexibility in experimental design by serving as a platform to facilitate high-throughput screening of novel diTPS candidates in different microbial strains. For our proof of principle, the process was optimized from affording trace amounts of two isomers, 13R-MO and 13S-MO, to obtaining 10 mg/liter of enantiopure 13R-MO, corresponding to a >2,000-fold increase. As such, our findings are comparable to previous findings regarding yield increase in microscale cultures (16, 33, 35), which hints that further engineering and culture scale-up can substantially increase the yield of 13R-MO, although the previously proposed cytotoxic effects (37–40) may present a challenge. However, our interest in 13R-MO is exclusively as a precursor for forskolin, and thus maintaining the microscale format is crucial at this stage. Moreover, Ajikumar and coworkers observed that the strain developed for efficient production of the olefin taxadiene (1 g/liter) (29) merely produced 50 mg/liter of the next biosynthetic intermediate taxadien-5 α -ol, thus indicating that careful balancing of enzyme activities might be most relevant once a complete pathway is assembled. We expect that the microbial synthesis process described here will serve as a foundation for full elucidation of the forskolin biosynthetic pathway by providing complex precursor molecules for *in vitro* assays and facilitate high-throughput *in vivo* screening. In due course, the microbial process may also be developed into a cell factory for sustainable production of forskolin and other labdane-type terpenes.

ACKNOWLEDGMENTS

This work was supported by the Novo Nordisk Foundation, UNIK Synthetic Biology, and an ERC Advanced grant: ERC-2012-AdG 323034 LightDriven P450s.

We thank Lone Riisberg for excellent technical assistance and Susanna Seppälä for critical reading of the manuscript.

It is noted that parts of the discoveries presented are the subject of a patent application.

REFERENCES

- Schwab W, Fuchs C, Huang F-C. 2012. Transformation of terpenes into fine chemicals. *Eur. J. Lipid Sci. Technol.* 115:3–8. <http://dx.doi.org/10.1002/ejlt.201200157>.
- Newman DJ, Cragg GM. 2012. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* 75:311–335. <http://dx.doi.org/10.1021/np200906s>.
- Zerbe P, Hamberger B, Yuen MMS, Chiang A, Sandhu HK, Madilao LL, Nguyen A, Hamberger B, Bach SS, Bohlmann J. 2013. Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant Physiol.* 162: 1073–1091. <http://dx.doi.org/10.1104/pp.113.218347>.
- Xiao M, Zhang Y, Chen X, Lee E-J, Barber CJS, Chakrabarty R, Desgagné-Penix I, Haslam TM, Kim Y-B, Liu E, MacNevin G, Masada-Atsumi S, Reed DW, Stout JM, Zerbe P, Zhang Y, Bohlmann J, Covello PS, De Luca V, Page JE, Ro D-K, Martin VJJ, Facchini PJ, Sensen CW. 2013. Transcriptome analysis on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.* 166:122–134. <http://dx.doi.org/10.1016/j.biotech.2013.04.004>.
- Wagh K, Patil P, Surana S, Wagh V. 2011. Forskolin: upcoming anti-glaucoma molecule. *J. Postgrad. Med.* 58:199–202. <http://dx.doi.org/10.4103/0022-3859.101396>.
- Yoneyama M, Sugiyama A, Satoh Y, Takahara A, Nakamura Y, Hashimoto K. 2002. Cardiovascular and adenylate cyclase stimulating effects of colforsin daropate, a water-soluble forskolin derivative, compared with those of isoproterenol, dopamine and dobutamine. *Circ. J.* 66:1150–1154. <http://dx.doi.org/10.1253/circj.66.1150>.
- Li Z, Wang J. 2006. A forskolin derivative, FSK88, induces apoptosis in human gastric cancer BGC823 cells through caspase activation involving regulation of Bcl-2 family gene expression, dissipation of mitochondrial membrane potential and cytochrome c release. *Cell Biol. Int.* 30:940–946. <http://dx.doi.org/10.1016/j.cellbi.2006.06.015>.
- Kikura M, Morita K, Sato S. 2004. Pharmacokinetics and a simulation model of colforsin daropate, new forskolin derivative inotropic vasodilator, in patients undergoing coronary artery bypass grafting. *Pharmacol. Res.* 49:275–281. <http://dx.doi.org/10.1016/j.phrs.2003.09.012>.
- Alasbahi R, Melzig M. 2010. Plectranthus barbatus: a review of phytochemistry, ethnobotanical uses and pharmacology—part 1. *Planta Med.* 76:653–661. <http://dx.doi.org/10.1055/s-0029-1240898>.
- Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Moller BL, Bohlmann J, Hamberger B. 2014. Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. *Plant Physiol.* 164:1222–1236. <http://dx.doi.org/10.1104/pp.113.228429>.
- Asada Y, Li W, Terada T, Kuang X, Li Q, Yoshikawa T, Hamaguchi S, Namekata I, Tanaka H, Koike K. 2012. Labdane-type diterpenoids from hairy root cultures of *Coleus forskohlii*, possible intermediates in the biosynthesis of forskolin. *Phytochemistry* 79:141–146. <http://dx.doi.org/10.1016/j.phytochem.2012.03.007>.
- Kirby J, Keasling JD. 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol.* 60:335–355. <http://dx.doi.org/10.1146/annurev.arplant.043008.091955>.
- Martin V, Pitera DJ, Withers ST, Newman JD, Keasling JD. 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* 21:796–802. <http://dx.doi.org/10.1038/nbt833>.
- Vallon T, Ghanegaonkar S, Vielhauer O, Müller A, Albermann C, Sprenger G, Reuss M, Lemuth K. 2008. Quantitative analysis of isoprenoid diphosphate intermediates in recombinant and wild-type *Escherichia coli* strains. *Appl. Microbiol. Biotechnol.* 81:175–182. <http://dx.doi.org/10.1007/s00253-008-1707-8>.
- Cyr A, Wilderman PR, Determan M, Peters RJ. 2007. A modular approach for facile biosynthesis of labdane-related diterpenes. *J. Am. Chem. Soc.* 129:6684–6685. <http://dx.doi.org/10.1021/ja071158n>.
- Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ. 2010. Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl. Microbiol. Biotechnol.* 85:1893–1906. <http://dx.doi.org/10.1007/s00253-009-2219-x>.
- Nielsen MT, Madsen KM, Seppälä S, Christensen U, Riisberg L, Harrison SJ, Møller BL, Nørholm MHH. 2014. Assembly of highly standardized gene fragments for high-level production of porphyrins in *Escherichia coli*. *ACS Synth. Biol.* <http://dx.doi.org/10.1021/sb500055u>.
- Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41:207–234. <http://dx.doi.org/10.1016/j.pep.2005.01.016>.
- Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nørholm MHH. 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* 10:21. <http://dx.doi.org/10.1186/1472-6750-10-21>.
- Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* 35:e55. <http://dx.doi.org/10.1093/nar/gkm106>.
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185:60–89. [http://dx.doi.org/10.1016/0076-6879\(90\)85008-C](http://dx.doi.org/10.1016/0076-6879(90)85008-C).
- Adams RP. 2007. Identification of essential oil components by gas chromatography/mass spectrometry, 4th ed. Allured Publishing Corporation, Carol Stream, IL.
- González AG, Arteaga JM, Bretón JL, Fraga BM. 1977. Five new labdane diterpene oxides from *Eupatorium jhanii*. *Phytochemistry* 16:107–110. [http://dx.doi.org/10.1016/0031-9422\(77\)83023-9](http://dx.doi.org/10.1016/0031-9422(77)83023-9).
- Almqvist S, Enzell CR, Wehrli FW. 1974. Carbon-13 NMR studies of labdane diterpenoids. *Acta Chem. Scand.* 29:695–702.
- Burke C, Croteau R. 2002. Geranyl diphosphate synthase from *Abies grandis*: cDNA isolation, functional expression, and characterization. *Arch. Biochem. Biophys.* 405:130–136. [http://dx.doi.org/10.1016/S0003-9861\(02\)00335-1](http://dx.doi.org/10.1016/S0003-9861(02)00335-1).
- Emanuelsson O, Nielsen H, von Heijne G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8:978–984. <http://dx.doi.org/10.1110/ps.8.5.978>.
- Nevin DE, Pratt JM. 1991. A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides from the T7-promoter. *FEBS Lett.* 291:259–263. [http://dx.doi.org/10.1016/0014-5793\(91\)81297-L](http://dx.doi.org/10.1016/0014-5793(91)81297-L).
- Ajlikumar PK, Xiao WH, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. 2010. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330:70–74. <http://dx.doi.org/10.1126/science.1191652>.
- Chimerel C, Field CM, Piñero-Fernandez S, Keyser UF, Summers DK. 2012. Indole prevents *Escherichia coli* cell division by modulating membrane potential. *Biochim. Biophys. Acta* 1818:1590–1594. <http://dx.doi.org/10.1016/j.bbame.2012.02.022>.
- Wang D, Ding X, Rather PN. 2001. Indole can act as an extracellular signal in *Escherichia coli*. *J. Bacteriol.* 183:4210–4216. <http://dx.doi.org/10.1128/JB.183.14.4210-4216.2001>.
- Li G, Young KD. 2013. Indole production by the tryptophanase TnaA in *Escherichia coli* is determined by the amount of exogenous tryptophan. *Microbiology* 159:402–410. <http://dx.doi.org/10.1099/mic.0.064139-0>.
- Leonard E, Ajlikumar PK, Thayer K, Xiao W-H, Mo JD, Tidor B, Stephanopoulos G, Prather KLJ. 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U. S. A.* 107:13654–13659. <http://dx.doi.org/10.1073/pnas.1006138107>.
- Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD. 2009. High-level production of amorpho-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* 4:e4489. <http://dx.doi.org/10.1371/journal.pone.0004489>.
- Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerre F, Pineda V, Rocci L, Daviet L. 2012. Toward a biosynthetic route to sclareol and amber odorants. *J. Am. Chem. Soc.* 134:18900–18903. <http://dx.doi.org/10.1021/ja307404u>.

36. Brück T, Kourist R, Loll B. 2014. Production of macrocyclic sesqui- and diterpenes in heterologous microbial hosts: a systems approach to harness nature's molecular diversity. *ChemCatChem* 6:1142–1165. <http://dx.doi.org/10.1002/cctc.201300733>.
37. Angelopoulou D, Demetzos C, Dimas C, Perdetzoglou D, Loukis A. 2001. Essential oils and hexane extracts from leaves and fruits of *Cistus monspeliensis*. Cytotoxic activity of ent-13-epi-manoyl oxide and its isomers. *Planta Med.* 67:168–171. <http://dx.doi.org/10.1055/s-2001-11497>.
38. de las Heras B, Hoult J. 1994. Noncytotoxic inhibition of macrophage eicosanoid biosynthesis and effects on leukocyte functions and reactive oxygen species of 2 novel antiinflammatory plant diterpenoids. *Planta Med.* 60:501–506. <http://dx.doi.org/10.1055/s-2006-959559>.
39. Demetzos C, Mitaku S, Couladis M, Harvala C, Kokkinopoulos D. 1994. Natural metabolites of ent-13-epi-manoyl oxide and other cytotoxic diterpenes from the resin “LADANO” of *Cistus creticus*. *Planta Med.* 60: 590–591. <http://dx.doi.org/10.1055/s-2006-959584>.
40. Dimas K, Demetzos C, Mitaku S, Vaos B, Marselos M, Tzavaras T, Kokkinopoulos D. 1999. Cytotoxic activity and antiproliferative effects of a new semisynthetic derivative of ent-3 beta-hydroxy-13-epi-manoyl oxide on human leukemic cell lines. *Anticancer Res.* 19:4065–4072.