Sunlight-driven Environmental Benign Production of Bioactive Natural Products with Focus on Diterpenoids and the Pathways Involved in their Formation

Dan Luo, Birger Lindberg Møller, and Irini Pateraki*

Abstract: Diterpenoids are high value compounds characterized by high structural complexity. They constitute the largest class of specialized metabolites produced by plants. Diterpenoids are flexible molecules able to engage in specific binding to drug targets like receptors and transporters. In this review we provide an account on how the complex pathways for diterpenoids may be elucidated. Following plant pathway discovery, the compounds may be produced in heterologous hosts like yeasts and *E. coli*. Environmentally contained production in photosynthetic cells like cyanobacteria, green algae or mosses are envisioned as the ultimate future production system.

Keywords: Alcohol dehydrogenases · Cytochrome P450s · Ingenol mebutate · Light-driven production · Terpene synthases

1. Introduction

1.1 *Plant Diterpenoids and their Application as Pharmaceuticals*

Plants produce a wide variety of natural products that play important roles in both general and specialized metabolism.[1,2] Terpenoids constitute the largest and most diverse class of bio-active natural products in the plant kingdom.[3] Based on the number of five-carbon building blocks embedded in their structures, terpenoids are categorized as monoterpenoids (C10), sesquiterpenoids (C15), diterpenoids (C20), sesterterpenoids (C25), triterpenoids (C30), tetraterpenoids (C40) and polyterpenoids (>C40) (Fig. 1). Plants produce more than 11,000 different diterpenoids[4] and they display a wide range of biological activities. Some are used directly as therapeutics for treatment of human diseases whereas others have served as lead compounds (Fig. 2).[5] Paclitaxel, first isolated from the bark of the Pacific yew, *Taxus brevifolia*, is a well-known anticancer agent widely used today as a therapeutic in combinatorial treatments of ovarian cancer, breast cancer, lung cancer, Kaposi sarcoma, cervical cancer, and pancreatic cancer.[6,7] Forskolin, a labdane-type diterpene produced from the roots of *Coleus forskohlii*, is used in the treatment of glaucoma and heart failure based on its activity as a cyclic AMP booster.[8,9] Ginkgolides from the leaves of *Ginkgo biloba* L. are a series of diterpene lactones with anti-platelet-activating antagonist ac-
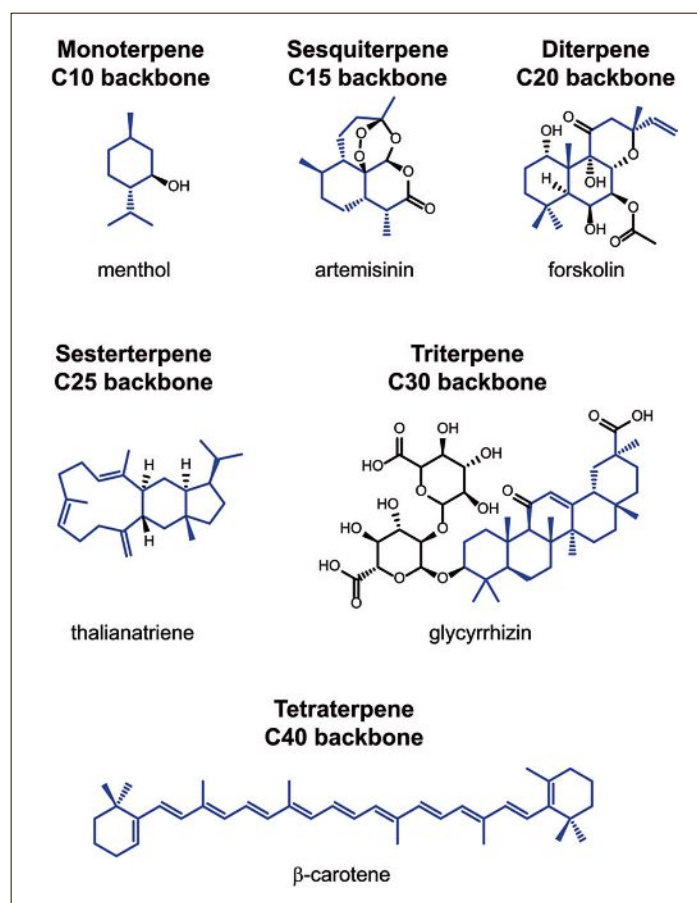

Fig 1. Examples of common plant terpenoid subclasses.

Correspondence: Assist. Prof. Dr. I. Pateraki
Plant Biochemistry Laboratory
Center for Synthetic Biology
Department of Plant and Environmental Sciences
University of Copenhagen, Thorvaldsensvej 40
DK-1871 Frederiksberg C, Copenhagen, Denmark
E-mail: eipa@plen.ku.dk

tivity.[10] For many years, they have been used in the treatment of cerebrovascular and cardiovascular diseases.[11] Ingenol mebutate is a macrocyclic diterpene from *Euphorbia peplus* exhibiting remarkable antitumor and antileukemic activity. A gel formulation of ingenol mebutate was approved for the topical treatment of actinic keratosis, a precondition of squamous-cell carcinoma.[12] With all the complex structures discovered in nature and especially from plants, diterpenoids are recognized as an especially attractive source of promising drug leads.

2. Biosynthesis of Diterpenoids

Despite their highly diverse structures, all terpenoids are biosynthesized from two common C5 isoprenoid units, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP and DMAP can be coupled and elongated through sequential head-to-tail condensations catalyzed by prenyltransferases, resulting in formation of linear prenyl diphosphate intermediates differing in the number of incorporated five-carbon building blocks. Two of these C5 units are condensed to form geranyl di-

phosphate (GPP, 10 carbons). Subsequent 1'-4 additions of isopentenyl diphosphate result in formation of farnesyl pyrophosphate (FPP, 15 carbons) and geranylgeranyl pyrophosphate (GGPP, 20 carbons). GPP, FPP and GGPP are the precursors of monoterpenoids, sesquiterpenoids and diterpenoids, respectively. FPP is also involved in the biosynthesis of triterpenoids (30 carbons). Squalene, the general precursor of triterpenoids contains six isoprene units and is formed from two molecules of FPP (Fig. 3).[13–17]

The universal C5 precursor isopentenyl diphosphate (IPP) can be synthesized *via* two different pathways: the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. In plants, the MVA pathway is located in the cytosol and mainly provides the C5 isoprenoid units for the biosynthesis of sesquiterpenoids and triterpenoids, while the MEP pathway locates in the plastids, providing the C5 isoprenoid units for mono- and diterpenoid biosynthesis (Fig. 4).[15–19]

This review is focused on the biosynthesis of diterpenoids in plants from the general precursor GGPP generated in the chloroplast by the MEP pathway. Metabolic crosstalk and interaction between the MEP and MVA pathways have been revealed in several plants species including *Nicotiana* and *Arabidopsis*.[20,21] In the biosynthesis of diterpenoids, the C20 precursor derived from the MEP pathway is transformed by diterpene synthases (diTPSs) into different types of backbones. This initial cyclization reaction is followed by oxidative decorations catalyzed by cytochrome P450 monooxygenases (P450s), and further modifications catalyzed *e.g.* by short-chain alcohol dehydrogenases (ADHs), acyltransferases and glucosyltransferases.[22]

3. Diterpene Synthases (diTPSs)

The initial step in diterpenoid biosynthesis is the formation of the basic backbone, which is mostly carried out by a family of enzymes known as diterpene synthases or cyclases (diTPSs). DiTPSs facilitate cyclization and/or rearrangement of the linear C20 prenyl diphosphate substrate into an enormous diversity of diterpene backbones through different carbocation reactions. These carbocation-driven reactions encompass a series of similar electrophilic steps: firstly, the generation, transformation, and stabilization of highly reactive carbocations; then the ultimate quenching of reactive carbocations by deprotonation or nucleophile capture. The structural features of the diterpene backbones formed arise from the diverse conformations of substrate and intermediates enforced by each individual
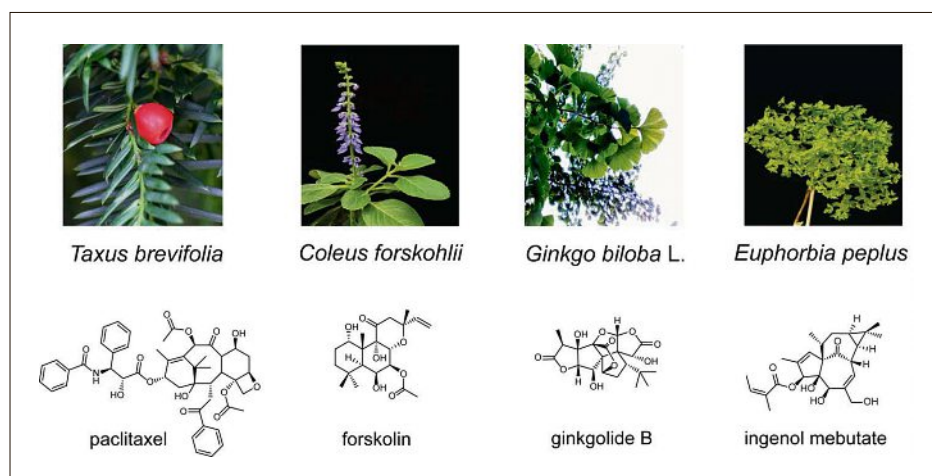

Fig 2. Examples of plant diterpenoids with important biological activities and the plants producing them.
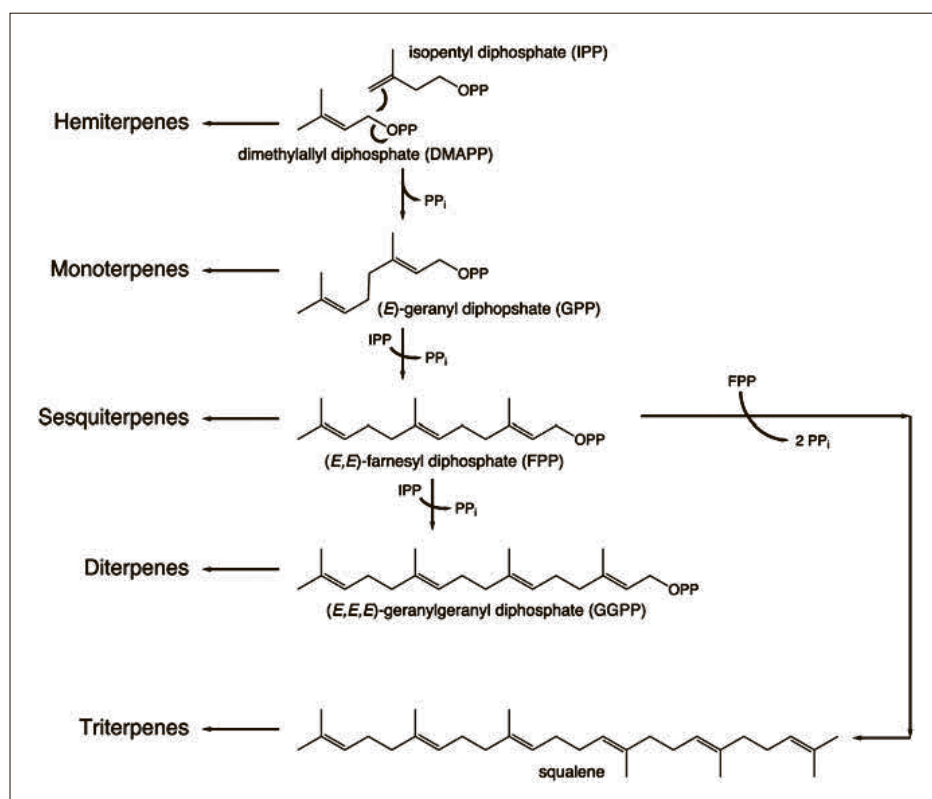

Fig 3. Scheme of the general terpenoid classes and corresponding precursors (OPP, diphosphate; PPi, inorganic pyrophosphate). Figure adapted with permission from Y. Gao, R. B. Honzatko, R. J. Peters, *Nat. Prod. Rep.* **2012**, *29*, 1153. © Royal Society of Chemistry
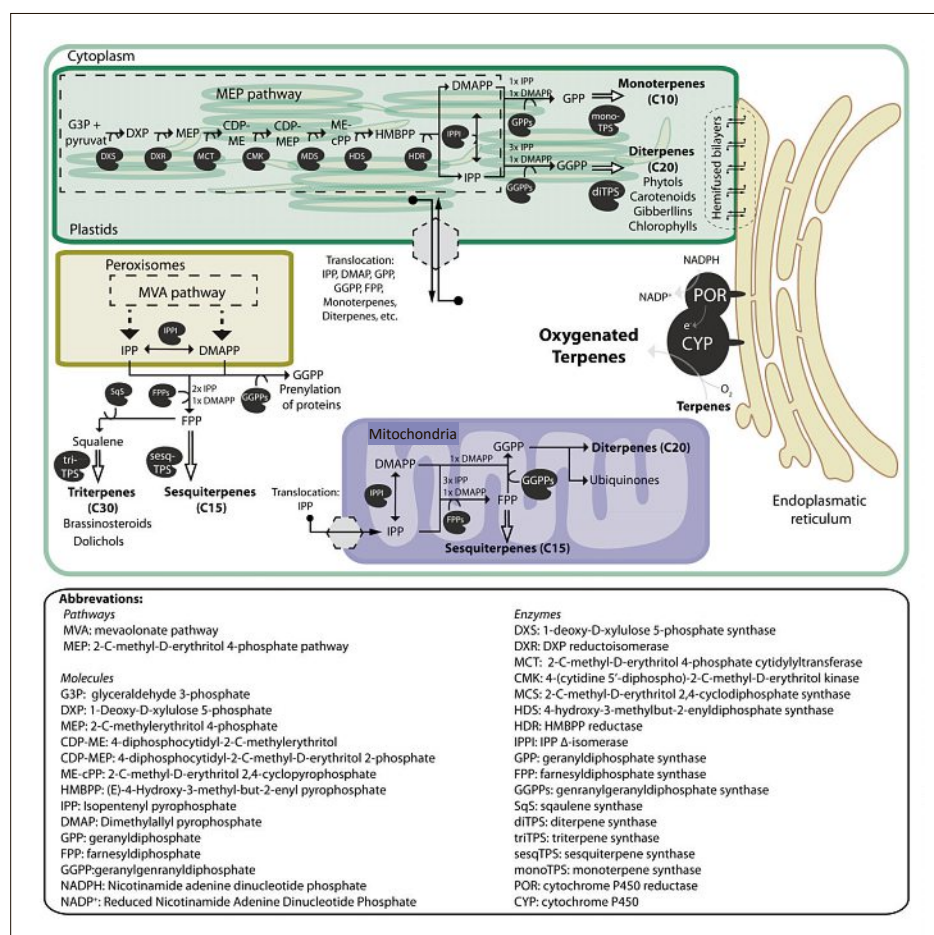
Fig. 4. Schematic representation of a plant cell and subcellular localization of terpenoid metabolism. Hemifused bilayers consisting by plastid and ER membrane, have been proposed as a translocation mechanism for nonpolar terpene compounds. Figure adapted from Johan Andersen-Ranberg's PhD thesis.

synthase. Structural features of the diTPSs also prevent a particular reaction series from premature termination.[23,24] Through mastering these reaction channels, diTPSs are capable to catalyze the formation of various classes of diterpenoids. According to different cyclization levels of their backbones, diterpenoids can be classified into three groups: linear diterpenes, macrocyclic diterpenes and polycyclic diterpenes (Fig. 5).[25–29] The biosynthesis of macrocyclic diterpenes has been of special interest in our laboratory. DiTPSs involved in the biosynthesis of macrocyclic diterpenes, such as casbene synthase (CBS) and taxadiene synthase, are mostly ionization-dependent cyclases.[28,29] These synthases share common features in their catalytic mechanisms: an initial ionization of the diphosphate group of GGPP, subsequent attack of the electrons of an internal double bond on the resulting carbocation followed by rearrangements, and finally stabilization of the intermediate by proton abstraction.[30] The catalytic mechanism of casbene synthase is analogous to those of angiosperm sesquiterpene synthases in the TPS-a subfamily, which is consistent with their sequence relatedness and phylogenies.[24]
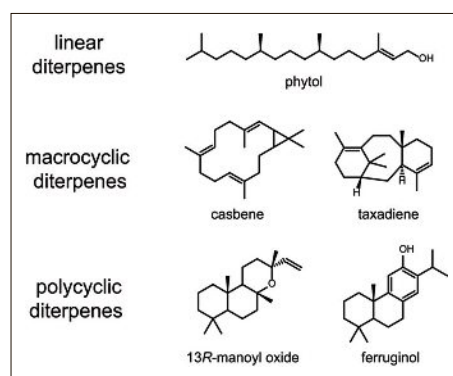

Fig. 5. Examples of diterpene backbones with different cyclization levels.

4. Cytochrome P450 Monooxygenases

Oxygenation is the most common modification to extend the structural diversity and complexity of plant diterpenoids.[4,31] Cytochrome P450 monooxygenases (P450s) are the major enzymes to generate a variety of regiospecific and stereospecific oxidative modifications at physiological conditions and ambient temperatures, including hydroxylation, ketonization, epoxidation, oxidative C–C bond cleavage

and isomerization.[32] As one of the largest super-families of enzymes, genes encoding P450s are found in the genomes of all plants. In angiosperms, the number of P450s within each species has exploded, with an average of 300 members per species, which may be divided into 50 families and subfamilies based on their sequence homology. The sequence identity of P450s from different families, even present in the same species, is often quite low and even less than 20%.[33] Despite their variability, the structural fold of P450 proteins is highly conserved. As heme-thiolate proteins, a conserved structural feature is the cysteine residue in the core of the protein serving as fifth (axial) ligand to the heme iron and thereby mediating heme-binding and catalytic properties. The thiolate function of the conserved cysteine residue, together with the iron-protoporphyrin IX (heme) constitutes the active center for catalysis. The catalytic cycle of a P450 requires electron transfer from NADPH carried out by an FMN/FAD containing NADPH-dependent P450 oxidoreductase (POR). *In vivo*, plant P450s and PORs are membrane-anchored to the surface of the endoplasmic reticulum (ER) (Fig. 4).[34] The less conserved regions of the P450s are associated with the substrate specificity, type of catalytic reactions and membrane localization. The regions related to substrate binding and recognition are flexible, adjusting conformation upon binding of substrate to allow the catalytic reaction. The abundance of variations of these regions in the P450s family reflects their high diversity of substrate selectivity and catalytic capability.[31,33] As their functions are extremely diverse, P450s are involved in almost all metabolic pathways in plants, for example, the biosynthesis of alkaloids, phenylpropanoids, cyanogenic glucosides, glucosinolates and terpenoids.[4] Despite the importance of this class of enzymes, functional characterization of newly discovered P450 remains an open field of investigation due to the large number of P450 encoding genes within a single plant species.

5. Short-chain Alcohol Dehydrogenases in Terpenoid Biosynthesis

Alcohol dehydrogenases (ADHs) are enzymes catalyzing oxidation of alcohols and reduction of aldehydes/ketones in the presence of coenzymes NAD(H)/NADP(H).[35] ADHs can be classified into long-, medium- and short-chain dehydrogenases/reductases according to their polypeptide chain length, mechanistic features and distinct sequence motifs.[36] In terpenoid metabolism, short-chain ADHs or SDRs, in addition to P450s, are an

important family of enzymes for catalyzing oxidation/reduction reactions for the modification or functionalization of terpenoid molecules.[37–42] Enzymes in this family, with the typical mass of 25/30 kDa and a polypeptide chain composed of approximately 250 amino acid residues, are characterized by several common properties: a conserved 3D structure consisting of 'Rossmann-fold' β-sheet with α-helices on both sides, an N-terminal dinucleotide cofactor binding motif and an active site with a catalytic residue motif YxxxK.[43] Despite their conserved structure, these short-chain ADHs exhibit large sequence divergences. The low sequence similarities among the plant short-chain ADHs renders functional prediction through phylogenetic analysis challenging. Functional characterization of these short-chain ADHs should therefore be based on experimental data regarding their bioactivity, like in the case of P450s. In terpenoid biosynthesis, short-chain ADHs cover a wide range of substrate spectra from monoterpenes, sesquiterpenes, diterpenes to triterpenes. The majority of the short chain ADHs known to be involved in terpenoid biosynthesis are members of the SDR110C family,[43,44] including zerumbone synthase in *Zingiber zerumbet*;[37] xanthoxin dehydrogenase in *Arabidopsis thaliana*;[38] borneol dehydrogenase in *Lavandula x intermedia*;[39] momilactone A synthase in *Oryza sativa*;[40] aldehyde dehydrogenase 1 in *Artemisia annua*;[41] (–)-isopiperitenol dehydrogenase in *Mentha x piperita*.[42] In terpenoid biosynthesis, their catalytic activities are mainly reported as oxidative rather than reductive.

6. Why are Diterpenoids Favored for Biological Activities?

Throughout the human history, indigenous people from different areas of the planet have used diterpenoid-producing plants and their extracts as traditional medicines sharing many of the same properties with present day therapeutics.[45] Within the plant-derived terpenoids, diterpenoids are especially rich in structures which are biologically active against a wide range of human diseases, including cancer, cardiovascular diseases, malaria, inflammations, diabetes and viral and bacterial infections.[46–51] Studies on structure–activity relationships (SAR) of diterpenoids document that the diverse bioactivity and high efficiencies of this class of compounds arise from their structurally unique features.[52–56] From a mechanistic point of view, the structural conformation of these small molecules is crucial as it determines whether they bind and interact with targets like receptors or transporters. In humans, drugs administered to cure a disease or

ailment are mostly targeting large biological molecules such as proteins, DNA and RNA. Inhibition, activation or other interactions with those targets elicit a biological response in the human body, resulting in a positive therapeutic effect following administration of the right drug.[57] With the huge number of potential structures already known to be produced by plants, diterpenoids represent a tremendous structural repository, not only with respect to the numerous patterns of their backbone structures, but also by the suite of incorporated functional groups (*e.g.* hydroxyl groups, keto groups, carboxyl groups, alkyl groups, sugar residues, *etc.*). In addition, diterpenoids contain an extremely high fraction of carbons in tetrahedral sp^3 hybridization, which offers bountiful opportunities for incorporation of chiral centers within the diterpenoid structures. These chemical features result in an unprecedented number of unique structural conformations, providing diterpenoid molecules with a strong potential for properly matching the shape of the binding or active sites of drug-target proteins. Building on a number of comprehensive experimental analyses, a small structural change is capable to lead to a significant alteration in the biological response, especially when these changes affect the energy of interaction of a ligand with its receptor.[58] Dramatic changes in their therapeutic activities ranging from none to inhibitory or excitatory might be induced by the replacement of a single functional group.[59,60] The diversity of functional groups in diterpenoid structures thus contributes greatly to their ability to interact with the very same target in different ways, resulting in completely different biological effects. This augments the opportunity for diterpenoids to be involved and actively targeting a wide range of signaling pathways in the human body.

The structural complexity and diversity of diterpenoids reflects their biosynthetic pathways. In comparison to the biosynthetic pathways for other classes of natural products such as the alkaloids and phenylpropanoids, the biosynthetic pathway for diterpenoids are built based on modules which are used in many different combinations. The plethora of core structures, which results from the ability of the soluble diTPSs to cyclize GGPP according to different model schemes, is truly impressive. While diTPSs mainly produce hydrocarbon backbones, they are also able to generate some oxygenated diterpenes directly without the involvement of other enzymes. For example, a pair of diterpene synthases, CfTPS2 and CfTPS3, from the Indian plant *Coleus forskohlii*, was found to catalyze the formation of 13R-manoyl oxide.[61] This catalytic character of some diTPSs enriches the elements involved in

the early intermediates of diterpenoid synthesis. Beyond that, a geometrically different alternative structure of the C20 precursor further serves to extend the diversity of diterpenoids discovered in nature. This was recently discovered, when some novel diTPSs from plants in the *Solanaceae* family were found to use the *cis*-prenyl diphosphate nerylneryl diphosphate (NNPP) as a substrate, rather than the *trans*-prenyl diphosphate *E,E,E*-geranylgeranyl diphosphate (GGPP), which as described above is the substrate of most known diTPSs.[62] After formation of the different core structures, membrane-bound P450s are subsequently recruited as versatile catalysts to enhance the structural diversity by regio- and stereospecific introduction of oxygenated functional groups. Apart from diTPSs and P450s, enzymes from other families, such as iridoid synthase-like reductases, have recently been reported to be able to use primary backbones constructed by diTPSs as substrates for multi-step modifications and rearrangements to generate cyclized structures of increased complexity.[63] Last but not least, soluble enzymes like acyltransferases and glucosyltransferases, typically complete the biosynthetic pathways by adding structural motifs derived from fatty acid and/or carbohydrate biosynthesis. Together these sets of soluble and membrane bound enzymes are able to provide the catalytic abilities required for the formation of the multitude of diterpenoids present in the plant kingdom.

7. Macrocyclic Diterpenoids and their Pharmacological Properties

Some diterpenoids, such as the gibberellin-type phytohormones and phytol, play important roles in general metabolism and are therefore present across the plant kingdom.[64,65] But the large majority of plant diterpenoids are specialized metabolites with a restricted taxonomic distribution. Macrocyclic diterpenes and their polyester derivatives are mainly isolated from the *Thymelaeaceae* and *Euphorbiaceae* plants, which could be considered as taxonomic markers or signature molecules.[66] *Euphorbiaceae*, the spurge family, is one of the largest families of flowing plants, consisting of approximately 7,500 species subdivided into 300 genera.[67] Members of this family produce a vast array of phytotoxins, mainly diterpenoid esters, alkaloids and glycosides.[68] The genus *Euphorbia* is the largest genera in the spurge family with about 2,160 species. Plants in this genus share a common feature, the production of a poisonous, milky and irritant latex-like sap. The skin-irritating and caustic effects of the latex sap are attributed to the presence of specific classes of macrocyclic

diterpenoids.[69] According to different stages of cyclization of their C_{20} backbones, they are classified into two groups: simple bicyclic casbene type and further cyclized types. The latter includes jatrophanes, lathyranes, tiglianes, daphnanes, and ingenanes ranked according to their increased structural complexity. In addition to the high structural variety of the backbones, the immense structural diversity occurring among macrocyclic diterpenoids results from the monooxygenation reactions catalyzed by P450s and the derived options to introduce esterification with different aliphatic (*e.g.* acetyl, n-butanoyl, isobutanoyl, methylbutanoyl, tiglyl, angeloyl, isovaleroyl, *etc.*) and aromatic (benzoyl and nicotinoyl) carboxylic acids. As the best characterized metabolites of this genus, these macrocyclic diterpenoids are also called 'Euphorbia diterpenoids'. Their specific taxonomic distribution and unique structural features have inspired chemists to search for the discovery of potential drug leads. Thus many macrocyclic diterpenoids have been tested for a wide range of therapeutically relevant biological activities, including antitumor, multidrug-resistance-reversal, various vascular effects, antiviral and anti-inflammatory activity.[69,70] *Euphorbia* factor L10 and euphodendroidin D show powerful inhibition of the transport activity of P-glycoprotein, a multidrug transporter overexpressed in cancer cell plasma membranes as an efflux pump conferring cellular resistance to anticancer chemotherapy.[55,71] Prostratin is in phase I human clinical trials for the treatment of HIV, as it was shown to activate viral reservoirs in latently infected T cells, as well as to inhibit viral replication.[72] Resiniferatoxin, now in phase II and III clinical trials, activates the transient vanilloid receptor 1 (TRPV1) in a subpopulation of primary afferent sensory neurons, blocking nerves that transmit physiological pain.[73] These biological investigations provide natural science-based evidence for the beneficial and health promoting uses of many *Euphorbia* plants in traditional medicines.[74] With the improved analytical technologies now available, numerous new macrocyclic diterpenoids will be identified and made available for determination of their pharmacological activities.

8. Ingenol Mebutate

Ingenol mebutate (ingenol-3-angelate or PEP005) is an ingenane-type macrocyclic diterpenoid ester with anti-leukemic and anti-tumor activity. Ingenol mebutate was first isolated from *Euphorbia peplus* (Fig. 6).[75] The latex of *E. peplus* is a traditional remedy for warts, skin keratosis and basal cell carcinoma.[76] Activity-
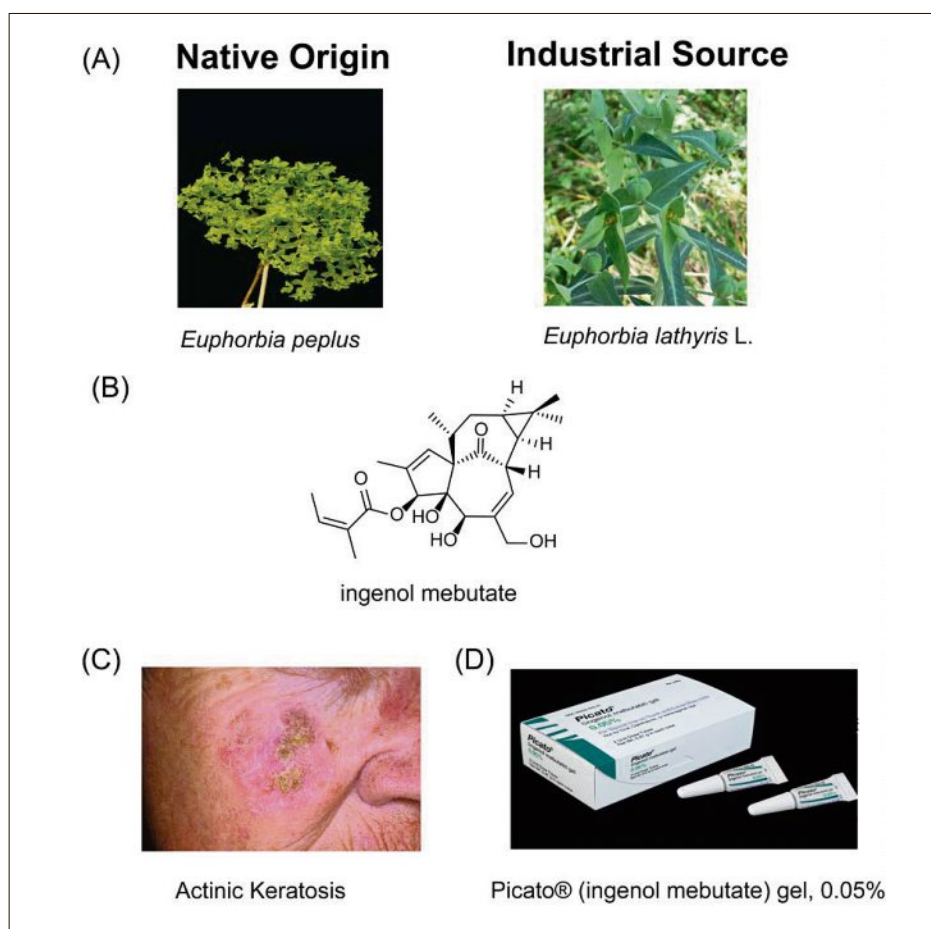

Fig. 6. A. *Euphorbia peplus* is the natural source of ingenol mebutate, but due to it's low concentration in the specific plant and the low availability of plant tissue, LEO pharma is using *Euphorbia lathyris* L. for the extraction of an ingenol mebutate derivative which is easily chemically converted to ingenol mebutate. B. Structure of the diterpenoid ingenol mebutate. C. Typical presentation of a patient with actinic keratosis lesions (Photo courtesy of Dermnet.com). D. Picato Gel produced by LEO Pharma (Photo from *http://www.qldsun.com.au/picato*).

guided fractionation and identification of the diterpenoids present in the latex sap revealed that the active constituent of the sap was ingenol mebutate. In 2012, a gel containing ingenol mebutate (trade name Picato) was approved by the US Food and Drug Administration (FDA) and by the European Medicines Agency (EMA) for topical treatment of actinic keratosis, a kind of common skin lesions with the potential to transform into squamous cell carcinoma (SCC). The mechanism of action for ingenol-3-angelate is described as dual, combining rapid induction of primary necrosis followed by neutrophil-mediated, antibody-dependent cellular cytotoxicity towards residual diseased cells. The latter activity is mediated by the action of ingenol-3-angelate towards a broad range of protein kinase C isoforms, leading to anti-cancer CD8 T cell-based immunotherapies.[77,78] Apart from actinic keratosis, ingenol-3-angelate was found to induce apoptosis in acute myeloid leukemia cells by activating the protein kinase C isoform PKCδ.[79] In other studies, ingenol mebutate was found to inhibit T cell apoptosis by activating PKCθ.[80]

9. Synthetic Biology-based Platforms for Sustainable Production of Structurally Complex Diterpenoids

Most bioactive diterpenoids have been isolated from non-cultivated medicinal plants or from endangered or red-listed species. A shortage of available wild plant material represents a serious obstacle for therapeutic uses. Because of their structural complexity, chemical synthesis of plant-derived diterpenoids, especially particular stereoisomers, remains uneconomic and often inadequate for large-scale production. Additionally, in chemical synthesis, the use of toxic catalysts and solvents has raised public concerns about the environment. The lack of a reliable and environmentally benign and stable supply chain thus remains a major obstacle when the desire arises to develop a specific diterpenoid into a medicinal drug.

Different strategies to improve the production yield of pharmaceutically interesting diterpenoids have been exploited. In the early stages of taxol production, the compound paclitaxel was directly

harvested from the bark of Pacific yew. This production process killed numerous Pacific yew trees, seriously vitiating the wild population of this *Taxus* species and the ecology of the forests in the Northwest Pacific. To protect the Pacific yew and its natural habitats, two alternative approaches were developed to obtain paclitaxel.[81] Early in the 1990s, cell cultures of *Taxus* species producing paclitaxel were used as starting materials.[82] Total synthesis of paclitaxel was accomplished in 1994 based on extensive efforts of synthetic chemists in the US.[83] The chemical synthesis afforded new general knowledge on structurally complex diterpenoids, but due to the low overall process yield, it did not form the basis for commercial production. A third production approach was based on a short semisynthetic route using 10-deacetylbaccatin as starting material and was shown feasible for industrial production because the compound 10-deacetylbaccatin used as starting material could be isolated in high yields from the European yew, *Taxus baccata*.[84] After that, semi-synthetic production of paclitaxel was further improved using combinations of precursors isolated from foliage of cultivated *Taxus* species and plant cell cultures.[85,86] The same situation also applied to ingenol mebutate. Direct isolation from the original plant source *E. peplus* only yielded 1.1 mg/kg from the aerial tissue, which hampered commercial production. Current production of ingenol mebutate proceeds as a semi-synthetic approach using hydrolysis products from the seed oil of an alternative *Euphorbia* species, *E. lathyris* L.[87] Due

to natural fluctuations and a biennial life cycle, the supply of *E. lathyris* L. seeds is still quite limited. The increasing demand for pharmaceutical-grade ingenol-3-angelate motivated us to embark into development of a synthetic biology-based approach towards economic and environmentally benign production.[88]

Microbial production using synthetic biology approaches has attracted wide attention as an alternative strategy for sustainable and environmentally friendly production of high-value terpenoids. Synthetic biology involves metabolic engineering of the modules of nature into new combinations to produce biological components, systems, cells and organisms that address society's needs. The modules are introduced in combinations where each module is functionally active receiving an input in the form of a substrate which the module transforms into an output in the form of a product serving as substrate for the next module. A basic principle is to work with nature-based knowledge on how modules may work and interact. In synthetic biology, the modules used are shared for use in bottom-up approaches to reach sustainable solutions.[89] Reconstruction of specific terpenoid pathways in microbial host organisms has been achieved, including production of the antimalarial sesquiterpenoid artemisinin that in nature is produced by the plant sweet wormwood (*Artemisia annua*).[90] Likewise, microbial routes towards production of the fragrances sclareol and santalol have been designed.[91,92] Recently, the complete biosynthetic pathway for the cyclic AMP booster forskolin was stably reconstructed

in yeast and afforded forskolin titers of 40 mg/L.[93] Forskolin, a structurally complex diterpenoid, demonstrates the potential of synthetic biology for sustainable production of complex diterpenoids with important bioactivities. In parallel with the synthetic biology-based approach, targeted molecular breeding efforts of the plant species producing the desired compound may raise the production level of the target compounds significantly and thus could be used as an alternative commercial source, as observed in the successful work in *Artemisia annua*.[94–96]

Light-driven biotechnological production (Fig. 7), which uses 'green' photosynthetic cells as production hosts, is a possible future ideal and ultimate way to obtain structurally complex plant diterpenoids.[97] Different from traditional microbial production systems such as *E. coli* and yeast, 'green' cell hosts, including cyanobacteria algae and moss, are highly 'self-sufficient'. No external sugar supply is required during the growth of these cells. Instead, they provide themselves with energy, reducing power, and carbon building blocks through the reactions of photosynthesis, channeling the energy of sunlight and CO_2 from the atmosphere into metabolic pathways. While the cultivation of genetically modified plants in the field is controversial, algae and cyanobacteria can easily be grown environmentally contained in photo-bioreactors (Fig. 8). Therefore, light-driven production has the potential to be economic feasible and sustainable for large-scale production, since sunlight and CO_2 are inexhaustible energy and carbon sources provided by nature.[98,99]
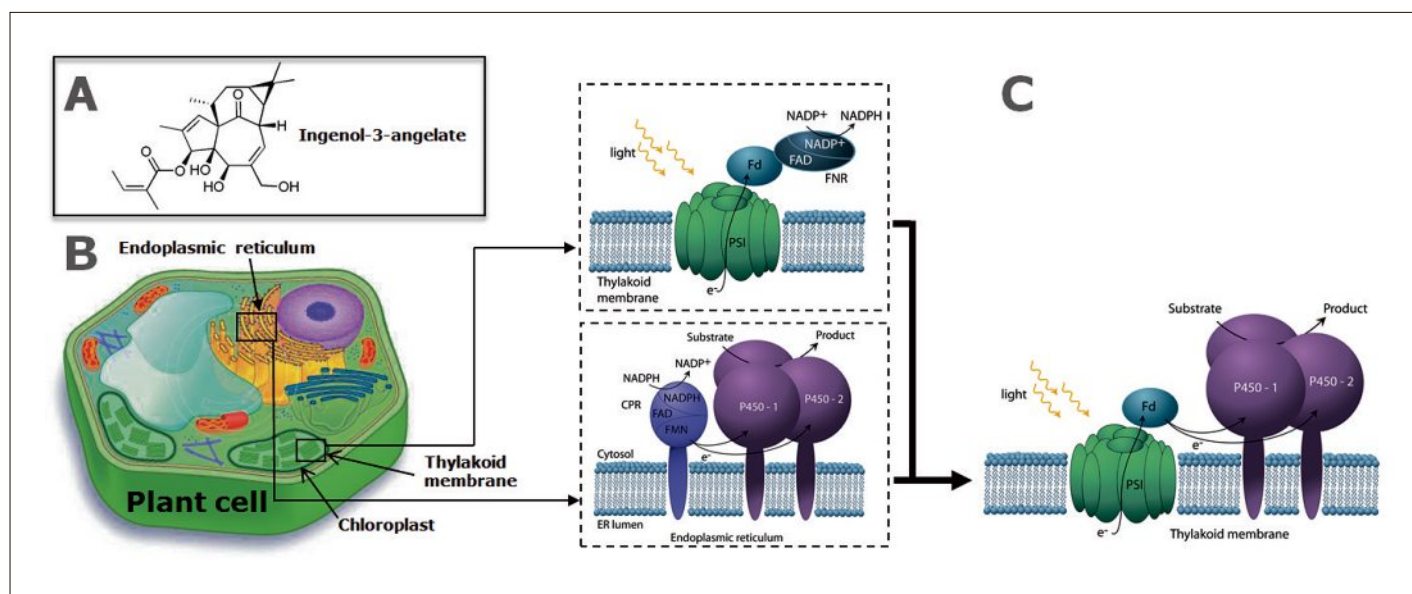

Fig. 7. Strategy for light-driven biotechnological production of ingenol mebutate (ingenol-3-angelate) A. Structure of the diterpenoid. B. Photosynthesis and biosynthesis of specialized bioactive compounds are separated at the cellular level taking place in the chloroplast and endoplasmic reticulum, respectively. C. Co-localization of the biosynthetic pathway of ingenol-3-angelate together with the photosynthetic complexes in the chloroplasts of moss and green algae *via* bioengineering. Note that the same principles can be used to target P450s into the cyanobacterial thylakoid membrane and likewise drive the reactions using solar energy. Figure adapted from *http://plantpower.ku.dk/about/*

Fig. 8. Microalgae are cultured in photo-bioreactors at Novagreen using natural light.

10. Strategies used for Gene Discovery

To enable heterologous production of high-value diterpenoids in non-photosynthetic as well as photosynthetic microbial cells as mentioned above, knowledge of the genes encoding the enzymes involved in the biosynthetic pathways giving rise to compound formation is essential. Fortunately, new technologies facilitating and guiding plant pathway discovery have been developed.[100]

The genomes of numerous medicinal plants are currently being sequenced but most of them are not yet publicly available. This also applies for *C. forskohlii*, *E. peblus* and *E. lathyris*. As a fast and cheap alternative, RNA sequencing (RNAseq) based transcriptome libraries have been shown very useful in the field of pathway discovery in these non-model plants, offering the opportunity to explore unknown enzyme and functional space involved in a biosynthetic pathway. Additionally, these RNAseq libraries can provide relative transcript abundances of the cDNAs expressed in the selected plant tissue or cell type, which can lead the selection of candidate cDNAs.[100,101] Biosynthesis and accumulation of specialized diterpenoids is typically limited to specific cell types.[61] Investigation of the formation of a specific diterpenoid in the course of plant ontogeny and within a specific cell type using tracer studies, metabolite profiling or mass spectrometry identifies the most suitable tissues for pathway elucidation, where genes encoding enzymes involved in the target pathway are probably highly expressed. Next-generation sequencing technology is then performed to obtain the transcriptomes of the chosen tissues. To select candidate genes represented in the transcriptomes, homology-based screening is carried out using previously characterized diTPS and P450 encoding gene sequences from other species involved in terpenoid metabolism.

Next, to determine which of the selected cDNAs participate in the target biosynthetic pathway, functional characterization of candidate genes has to be performed; for this step different screening and expression systems may be used. These include *in vivo* experiments using the *Nicotiana benthamiana*/agrobacterium transient expression system, engineered *Saccharomyces cerevisiae* (yeast) strains and *in vitro* assays with recombinant enzymes purified from *Escherichia coli* strains or coupled in transcription/translation assays. The *N. benthamiana*/agrobacterium transient expression system offers fast functional screening of candidate genes.[102] This plant-based expression system has great advantage in heterologous expression of plant enzymes owing to the 'close to native' internal environment in which the proteins will be produced and functionally embedded. Because *N. benthamiana* plants produce diterpenoids for their own growth and self-defense, the machinery for GGPP synthesis exists naturally in this model-plant host. An additional benefit of this 'green' host is that it favors the expression and functionality of heterologous plant P450s.[103] Due to issues such as membrane localization of the protein and availability of reducing power, it can be quite challenging to functionally express plant P450s in *E.coli* and yeast.[104] Despite these benefits, the *Nicotiana* expression system also has its limits. Intermediates and products of interest synthesized from heterologously introduced pathways might be further metabolized by endogenous enzymes and other detoxification mechanisms in *N. benthamiana*, since this plant species also produce oxygenated diterpenoids formed by endogenous diTPSs and P450s.[105] To confirm the results observed using *N. benthamiana* as host, it is therefore advisable to carry out parallel experiments in engineered yeast strains or *in vitro* enzyme assays (where possible) as controls.

11. Conclusion

The global challenges that humankind is facing worldwide related to the use of fossil fuels, climate change, food security, environmental preservation, mass migration and financial inequality are all well known. In order to address these challenges, we should first shift our main production systems away from petrochemicals. This transition towards a truly bio-based society requires multidisciplinary research and the development and successful introduction of new products manufactured using renewable resources and by the use of yet to be developed novel green technologies that are economically viable and have transformative power. In our research, we are trying to contribute to this transition by developing sunlight-driven production platforms for high value products. Because the photosynthetic cells optimized for production of high value compounds are genetically engineered, we decided not to work with plants grown in the field or in greenhouses. Instead all production will be environmentally contained using cyanobacteria, algae or moss cells grown in suspension cultures in transparent plastic bags. The major challenge we currently face is to optimize the photosynthetic production systems to efficiently compete with classical production systems based on fermentation processes like yeast or *E. coli* cultures, whose production is based on direct use of petrochemicals. We are confident that within the next decade or two, the growing demand for sunlight-driven environmental benign production systems will attract the necessary huge investments required to make this vision a reality.

[1] H. P. Makkar, P. Siddhuraju, K. Becker, *Meth. Mol. Biol.* **2007**, *393*, 1.
[2] K. Springob, T. M. Kutchan, 'Introduction to the Different Classes of Natural Products', in 'Plant-derived Natural Products, Synthesis, Function and Applications', Eds. A. E. Osbourn, V. Lanzotti, Springer-Verlag New York, **2009**, 3–50.
[3] B. Singh, R. A. Sharma, *3 Biotech* **2015**, *5*, 129.
[4] B. Hamberger, S. Bak, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, 20120426.
[5] S. Mafu, P. Zerbe, *Phytochem. Rev.* **2017**, DOI: 10.1007/s11101-017-9513-5.
[6] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggon, A. T. McPhail, *J. Am. Chem. Soc.* **1971**, *93*, 2325.
[7] G. M. Cragg, *Med. Res. Rev.* **1998**, *18*, 315.
[8] S. Doseyici, I. Mehmetoglu, A. Toker, F. H. Yerlikaya, E. Erbay, *Biotech. Histochem.* **2014**, *89*, 388.
[9] K. B. Seamon, W. Padgett, J. W. Daly, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 3363.
[10] K. Stromgaard, K. Nakanishi, *Angew. Chem. Int. Ed.* **2004**, *43*, 1640.

[11] K. M. Nash, Z. A. Shah, *Integr. Med. Insights* **2015**, *10*, 1.

[12] G. Siller, K. Gebauer, P. Welburn, J. Katsamas, S. M. Ogbourne, *Aust. J. Dermatol.* **2009**, *50*, 16.

[13] Y. Gao, R. B. Honzatko, R. J. Peters, *Nat. Prod. Rep.* **2012**, *29*, 1153.

[14] T. W. Goodwin, *Annu. Rev. Plant Physiol.* **1979**, *30*, 369.

[15] A. X. Cheng, Y. G. Lou, Y. B. Mao, S. Lu, L. J. Wang, X. Y. Chen, *J. Integ. Plant Biol.* **2007**, *49*, 179.

[16] E. Vranova, D. Coman, W. Gruissem, *Mol. Plant* **2012**, *5*, 318,.

[17] D. Tholl, *Curr. Op. Plant Biol.* **2006**, *9*, 297.

[18] E. Vranova, D. Coman, W. Gruissem, *Annu. Rev. Plant Biol.* **2013**, *64*, 665.

[19] D. Tholl, S. Lee, *The Arabidopsis Book* **2011**, *9*, e0143.

[20] O. Laule, A. Fürholz, H.-S. Chang, T. Zhu, X. Wang, P. B. Heifetz, W. Gruissem, M. Lange, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6866.

[21] A. Hemmerlin, J.-F. Hoeffler, O. Meyer, D. Tritsch, I. A. Kagan, C. Grosdemange-Billiard, M. Rohmer, T. J. Bach, *J. Biol. Chem.* **2003**, *278*, 26666.

[22] P. Zerbe, J. Bohlmann, *Trends Biotechnol.* **2015**, *33*, 419.

[23] E. Davis, R. Croteau, *Biosynthesis* **2000**, 53.

[24] J. Bohlmann, G. Meyer-Gauen, R. Croteau, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4126.

[25] D. M. Martin, J. Faldt, J. Bohlmann, *Plant Physiol.* **2004**, *135*, 1908.

[26] M. Herde, K. Gartner, T. G. Kollner, B. Fode, W. Boland, J. Gershenzon, C. Gatz, D. Tholl, *Plant Cell* **2008**, *20*, 1152.

[27] C. I. Keeling, H. K. Dullat, M. Yuen, S. G. Ralph, S. Jancsik, J. Bohlmann, *Plant Physiol.* **2010**, *152*, 1197.

[28] C. Mau, C. A. West, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8497.

[29] M. Hezari, N. G. Lewis, R. Croteau, *Arch. Biochem. Biophys.* **1995**, *322*, 437.

[30] R. Croteau, E. Davis, T. Hartmann, T. Hemscheidt, J. Sanz-Cervera, B. Shen, E. Stocking, R. Williams, 'Biosynthesis: Aromatic Polyketides, Isoprenoids, Alkaloids', Vol. 209, Springer, **2003**.

[31] I. Pateraki, A. M. Heskes, B. Hamberger, *Adv. Biochem. Eng. Biotechnol.* **2015**, *148*, 107.

[32] F. P. Guengerich, A. W. Munro, *J. Biol. Chem.* **2013**, *288*, 17065.

[33] D. Werck-Reichhart, R. Feyereisen, *Genome Biol.* **2000**, *1*, REVIEWS3003.

[34] V. B. Urlacher, M. Girhard, *Trends Biotechnol.* **2012**, *30*, 26.

[35] Y. Kallberg, U. Oppermann, H. Jörnvall, B. Persson, *FEBS J.* **2002**, *269*, 4409.

[36] K. Kavanagh, H. Jörnvall, B. Persson, U. Oppermann, *Cell. Mol. Life Sci.* **2008**, *65*, 3895.

[37] S. Okamoto, F. Yu, H. Harada, T. Okajima, J. Hattan, N. Misawa, R. Utsumi, *FEBS J.* **2011**, *278*, 2892.

[38] M. Gonzalez-Guzman, N. Apostolova, J. M. Belles, J. M. Barrero, P. Piqueras, M. R. Ponce, J. L. Micol, R. Serrano, P. L. Rodriguez, *Plant Cell* **2002**, *14*, 1833.

[39] L. S. Sarker, M. Galata, Z. A. Demissie, S. S. Mahmoud, *Arch. Biochem. Biophys.* **2012**, *528*, 163.

[40] A. Atawong, M. Hasegawa, O. Kodama, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 566.

[41] M. Liu, P. Shi, X. Fu, P. E. Brodelius, Q. Shen, W. Jiang, Q. He, K. Tang, *Plant Cell, Tissue and Organ Culture (PCTOC)* **2016**, *126*, 469.

[42] G. W. Turner, R. Croteau, *Plant Physiol.* **2004**, *136*, 4215.

[43] H. Moummou, Y. Kallberg, L. B. Tonfack, B. Persson, B. van der Rest, *BMC Plant Biol.* **2012**, *12*, 219.

[44] S. A. McAdam, F. C. Sussmilch, T. J. Brodribb, J. J. Ross, *AoB Plants* **2015**, *7*, DOI: 10.1093/aobpla/plv091.

[45] V. Lanzotti, in 'Natural Products', Springer, **2013**, p. 3173.

[46] M. T. Islam, *Phytotherapy Res.* **2017**, *31*, 691.

[47] A. M. de Oliveira, C. R. Tirapelli, S. R. Ambrosio, F. B. da Costa, *Recent Pat. Cardiovasc. Drug Discov.* **2008**, *3*, 1.

[48] R. Batista, J. Silva Ade, Jr., A. B. de Oliveira, *Molecules* **2009**, *14*, 3037.

[49] C. Cardenas, A. R. Quesada, M. A. Medina, *PLoS One* **2011**, *6*, e23407.

[50] A. Nagarajan, P. Brindha, *J. Pharma. Res.* **2012**, *5*, 4530.

[51] A. Ulubelen, S. Öksüz, G. Topcu, A. C. Gören, W. Voelter, *J. Nat. Prod.* **2001**, *64*, 549.

[52] K. B. Seamon, J. W. Daly, H. Metzger, N. J. de Souza, J. Reden, *J. Med. Chem.* **1983**, *26*, 436.

[53] L. Moujir, A. M. Gutiérrez-Navarro, L. San Andrés, J. G. Luis, *Phytochem.* **1993**, *34*, 1493.

[54] G. Krauter, C. W. Von der Lieth, R. Schmidt, E. Hecker, *Eur. J. Biochem.* **1996**, *242*, 417.

[55] G. Corea, E. Fattorusso, V. Lanzotti, O. Taglialatela-Scafati, G. Appendino, M. Ballero, P. N. Simon, C. Dumontet, A. Di Pietro, *J. Med. Chem.* **2003**, *46*, 3395.

[56] G. Corea, E. Fattorusso, V. Lanzotti, R. Motti, P.-N. Simon, C. Dumontet, A. Di Pietro, *J. Med. Chem.* **2004**, *47*, 988.

[57] J. P. Hughes, S. Rees, S. B. Kalindjian, K. L. Philpott, *Br. J. Pharmacol.* **2011**, *162*, 1239.

[58] A. C. Anderson, *Chem. Biol.* **2003**, *10*, 787.

[59] D. Colquhoun, *Br. J. Pharmacol.* **1998**, *125*, 923.

[60] S. Souza, M. Trindade, J. R. G. d. S. Almeida, A. A. Souza Araujo, M. C. Duarte, D. P. Gelain, J. C. F. Moreira, M. R. V. Santos, L. J. Quintans Júnior, *Basic Clin. Pharmacol. Toxicol.* **2014**, *115*, 244.

[61] I. Pateraki, J. Andersen-Ranberg, B. Hamberger, A. M. Heskes, H. J. Martens, P. Zerbe, S. S. Bach, B. L. Moller, J. Bohlmann, B. Hamberger, *Plant Physiol.* **2014**, *164*, 1222.

[62] Y. Matsuba, J. Zi, A. D. Jones, R. J. Peters, E. Pichersky, *PLoS One* **2015**, *10*, e0119302.

[63] F. Geu-Flores, N. H. Sherden, V. Courdavault, V. Burlat, W. S. Glenn, C. Wu, E. Nims, Y. Cui, S. E. O'Connor, *Nature* **2012**, *492*, 138.

[64] R. J. Peters, in 'Isoprenoid synthesis in plants and microorganisms', Springer, **2012**, p. 233.

[65] T. Ischebeck, A. M. Zbierzak, M. Kanwischer, P. Dormann, *J. Biol. Chem.* **2006**, *281*, 2470.

[66] R. F. Keeler, in 'Handbook of Natural Toxins: Toxicology of Plant and Fungal Compounds', Vol. 6, CRC Press, **1991**, p. 217.

[67] A. Takhtajan, 'Flowering plants', Springer Science & Business Media, **2009**.

[68] A. F. M. Rizk, *Bot. J. Linnean Soc.* **1987**, *94*, 293.

[69] A. Vasas, J. Hohmann, *Chem. Rev.* **2014**, *114*, 8579.

[70] Q. W. Shi, X. H. Su, H. Kiyota, *Chem. Rev.* **2008**, *108*, 4295.

[71] G. Appendino, C. Della Porta, G. Conseil, O. Sterner, E. Mercalli, C. Dumontet, A. Di Pietro, *J. Nat. Prod.* **2003**, *66*, 140.

[72] H. E. Johnson, S. A. Banack, P. A. Cox, *J. Nat. Prod.* **2008**, *71*, 2041.

[73] M. Raisinghani, R. M. Pabbidi, L. S. Premkumar, *J. Physiol.* **2005**, *567*, 771.

[74] M. Ernst, O. M. Grace, C. H. Saslis-Lagoudakis, N. Nilsson, H. T. Simonsen, N. Ronsted, *J. Ethnopharmacol.* **2015**, *176*, 90.

[75] J. Hohmann, F. Evanics, L. Berta, T. Bartok, *Planta Med.* **2000**, *66*, 291.

[76] J. Ramsay, A. Suhrbier, J. Aylward, S. Ogbourne, S. J. Cozzi, M. Poulsen, K. Baumann, P. Welburn, G. Redlich, P. Parsons, *Br. J. Dermatol.* **2011**, *164*, 633.

[77] N. Kedei, D. J. Lundberg, A. Toth, P. Welburn, S. H. Garfield, P. M. Blumberg, *Cancer Res.* **2004**, *64*, 3243.

[78] R. H. Rosen, A. K. Gupta, S. K. Tyring, *J. Am. Acad. Dermatol.* **2012**, *66*, 486.

[79] P. Hampson, H. Chahal, F. Khanim, R. Hayden, A. Mulder, L. K. Assi, C. M. Bunce, J. M. Lord, *Blood* **2005**, *106*, 1362.

[80] W. Y. Lee, P. Hampson, L. Coulthard, F. Ali, M. Salmon, J. M. Lord, D. Scheel-Toellner, *J. Biol. Chem.* **2010**, *285*, 23889.

[81] J. Goodman, V. Walsh, 'The story of taxol: nature and politics in the pursuit of an anti-cancer drug', Cambridge University Press, **2001**.

[82] Y. Yukimune, H. Tabata, Y. Higashi, Y. Hara, *Nat. Biotechnol.* **1996**, *14*, 1129.

[83] K. Nicolaou, Z. Yang, J. Liu, H. Ueno, P. Nantermet, R. Guy, C. Claiborne, J. Renaud, E. Couladouros, K. Paulvannan, *Nature* **1994**, *367*, 630.

[84] P. G. Wuts, *Curr. Opin. Drug Discov. Devel.* **1998**, *1*, 329.

[85] S. A. Wilson, S. C. Roberts, *Plant Biotechnol. J.* **2012**, *10*, 249.

[86] S.-H. Pyo, H.-B. Park, B.-K. Song, B.-H. Han, J.-H. Kim, *Proc. Biochem.* **2004**, *39*, 1985.

[87] X. Liang, G. Grue-Sørensen, A. K. Petersen, T. Högberg, *Synlett* **2012**, 2647.

[88] G. Appendino, *Angew. Chem. Int. Ed.* **2014**, *53*, 927.

[89] E. Andrianantoandro, S. Basu, D. K. Karig, R. Weiss, *Mol. Syst. Biol.* **2006**, *2*, 2006 0028.

[90] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, D. R. Polichuk, K. H. Teoh, D. W. Reed, T. Treynor, J. Lenihan, M. Fleck, S. Bajad, G. Dang, D. Dengrove, D. Diola, G. Dorin, K. W. Ellens, S. Fickes, J. Galazzo, S. P. Gaucher, T. Geistlinger, R. Henry, M. Hepp, T. Horning, T. Iqbal, H. Jiang, L. Kizer, B. Lieu, D. Melis, N. Moss, R. Regentin, S. Secrest, H. Tsuruta, R. Vazquez, L. F. Westblade, L. Xu, M. Yu, Y. Zhang, L. Zhao, J. Lievense, P. S. Covello, J. D. Keasling, K. K. Reiling, N. S. Renninger, J. D. Newman, *Nature* **2013**, *496*, 528.

[91] M. Schalk, L. Pastore, M. A. Mirata, S. Khim, M. Schouwey, F. Deguerry, V. Pineda, L. Rocci, L. Daviet, *J. Am. Chem. Soc.* **2012**, *134*, 18900.

[92] G. Scalcinati, S. Partow, V. Siewers, M. Schalk, L. Daviet, J. Nielsen, *Microbial Cell Factories* **2012**, *11*, 117.

[93] I. Pateraki, J. Andersen-Ranberg, N. B. Jensen, S. G. Wubshet, A. M. Heskes, V. Forman, B. Hallstrom, B. Hamberger, M. S. Motawia, C. E. Olsen, D. Staerk, J. Hansen, B. L. Moller, B. Hamberger, *eLife* **2017**, *6*, DOI: 10.7554/eLife.23001.

[94] I. A. Graham, K. Besser, S. Blumer, C. A. Branigan, T. Czechowski, L. Elias, I. Guterman, D. Harvey, P. G. Isaac, A. M. Ikram, T. R. Larson, Y. Li, T. Pawson, T. Penfield, A. M. Rae, D. A. Rathbone, S. Reid, J. Ross, M. F. Smallwood, V. Segura, T. Townsend, D. Vyas, T. Winzer, D. Bowles, *Science* **2010**, *327*, 328.

[95] T. Townsend, V. Segura, G. Chigeza, T. Penfield, A. Rae, D. Harvey, D. Bowles, I. A. Graham, *Plos one* **2013**, *8*, e61989.

[96] M. Peplow, *Nature* **2016**, *530*, 389.

[97] K. Jensen, P. E. Jensen, B. L. Moller, *Trends Plant Sci.* **2012**, *17*, 60.

[98] A. Z. Nielsen, B. Ziersen, K. Jensen, L. M. n. Lassen, C. E. Olsen, B. L. Møller, P. E. Jensen, *ACS Syn. Biol.* **2013**, *2*, 308.

[99] L. M. n. Lassen, A. Z. Nielsen, B. Ziersen, T. Gnanasekaran, B. L. Møller, P. E. Jensen, *ACS Syn. Biol.* **2013**, *3*, 1.

[100] Y. Higashi, K. Saito, *Plant Cell Environ.* **2013**, *36*, 1597.

[101] E. Gongora-Castillo, C. R. Buell, *Nat. Prod. Rep.* **2013**, *30*, 490.

[102] S. S. Bach, J.-É. Bassard, J. Andersen-Ranberg, M. E. Møldrup, H. T. Simonsen, B. Hamberger, in 'Methods in molecular biology', Springer, **2014**, p. 245.

[103] K. Geisler, R. K. Hughes, F. Sainsbury, G. P. Lomonossoff, M. Rejzek, S. Fairhurst, C. E. Olsen, M. S. Motawia, R. E. Melton, A. M. Hemmings, S. Bak, A. Osbourn, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E3360.

[104] H. Duan, M. A. Schuler, *Phytochem. Rev.* **2006**, *5*, 507.

[105] E. Wang, J. T. Hall, G. J. Wagner, *Mol. Breeding* **2004**, *13*, 49.