



**HAL**  
open science

# Biosynthesis of new alpha-bisabolol derivatives through a synthetic biology approach

Arthur Sarrade-Loucheur

► **To cite this version:**

Arthur Sarrade-Loucheur. Biosynthesis of new alpha-bisabolol derivatives through a synthetic biology approach. Biochemistry, Molecular Biology. INSA de Toulouse, 2020. English. NNT : 2020ISAT0003 . tel-02976811

**HAL Id: tel-02976811**

**<https://theses.hal.science/tel-02976811>**

Submitted on 23 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

**En vue de l'obtention du  
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**  
Délivré par l'Institut National des Sciences Appliquées de  
Toulouse

---

**Présentée et soutenue par  
Arthur SARRADE-LOUCHEUR**

Le 30 juin 2020

**Biosynthèse de nouveaux dérivés de l' $\alpha$ -bisabolol par une  
approche de biologie synthèse**

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et  
Bioingenieries**

Spécialité : **Ingénieries microbienne et enzymatique**

Unité de recherche :  
**TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering**

Thèse dirigée par  
**Gilles TRUAN et Magali REMAUD-SIMEON**

Jury

**Mme Véronique DE BERARDINIS**, Rapporteuse, Chercheur CEA  
**M. Jean-Etienne BASSARD**, Rapporteur, Chargé de Recherche, CNRS  
**Mme Danièle WERCK-REICHHART**, Examinatrice, Directeur de Recherche Emérite, CNRS  
**M. Gilles TRUAN**, Directeur de thèse, Directeur de Recherche, CNRS  
**Mme Magali REMAUD-SIMEON**, Co-directrice de thèse, Professeur des Universités, INSA  
**Mme Fayza DABOUSSI**, Présidente, Directeur de Recherche INRAE



## Biosynthesis of new $\alpha$ -bisabolol derivatives through a synthetic biology approach

The rise of synthetic biology now enables the production of new to nature molecules. In the frame of this thesis we focused on the diversification of the (+)-*epi*- $\alpha$ -bisabolol scaffold. This molecule coming from the plant *Lippia dulcis* belongs to the vast family of sesquiterpenes. While sesquiterpenes possess diverse biological activities, (+)-*epi*- $\alpha$ -bisabolol is the precursor of hernandulcin, an intense sweetener. However, the last oxidative step(s) of the hernandulcin biosynthetic pathway remain elusive.

Rather than seeking the native oxidase responsible for hernandulcin synthesis among *L. dulcis* enzymes we turned our choice towards oxidative enzymes known to be promiscuous and that could functionalize (+)-*epi*- $\alpha$ -bisabolol in order to i) generate diversity from (+)-*epi*- $\alpha$ -bisabolol; ii) hopefully identify an oxidative enzyme catalyzing hernandulcin synthesis. First, a yeast chassis strain enabling the *in vivo* screening of cytochromes P450 (CYPs) was constructed by coexpressing two key enzymes: the (+)-*epi*- $\alpha$ -bisabolol synthase (BBS) and the NADPH cytochrome P450 reductase. Then, the *in vivo* screening assays revealed that 5 CYPs out of our library of 25 animal CYPs involved in xenobiotic metabolism oxidized (+)-*epi*- $\alpha$ -bisabolol and produced new hydroxylated regioisomers. Of the oxidized products, the structure of one compound, 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, was fully elucidated by NMR while the probable structure of a second one was determined, 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol. In parallel, the production of (+)-*epi*- $\alpha$ -bisabolol derivatives was enhanced through addition of a supplementary genomic copy of BBS that augmented the final titer of hydroxylated product to 64 mg/L. We thus demonstrate that drug metabolism CYPs can be used to produce novel compounds from a sesquiterpene scaffold.

Furthermore, different products were obtained with two homologous enzymes *i.e.* CYP2B6 and CYP2B11. This prompted us to study the molecular determinants putatively responsible for enzyme regiospecificity. From chimeric enzymes composed of secondary structure elements originating from CYP2B6 and CYP2B11 we were not able to identify specific motifs that could explain the CYPs regiospecificity. This approach suggests that the molecular determinants cannot be attributed to specific structural elements of the two enzymes but are rather widespread in the protein sequences.

In order to generate a wider molecular diversity from  $\alpha$ -bisabolol or hernandulcin, and synthesize molecules with different physico-chemical and biological properties (*i.e.* solubility, sweetening power etc.), we attempted the glycosylation of these compounds using a plant glucosyltransferase (UGT93B16) or human glucuronosyltransferases. UGT93B16 was found to glucosylate both (-)- $\alpha$ -bisabolol and hernandulcin. In addition, we carried out whole cell catalysis using *E. coli* and *S. cerevisiae* as recombinant producers of UGT93B16. Comparison of the two microbial hosts showed that glycosylation using yeast cells was more efficient. In parallel, we investigated  $\alpha$ -bisabolol glucosylation by human UDP-glucuronosyltransferases involved in the xenobiotic metabolism. *In vitro* enzymatic assays demonstrated a weak activity of UGT1A9, UGT2B4 and UGT2B7 towards  $\alpha$ -bisabolol and hernandulcin. However, their introduction in yeast failed to produce detectable amounts of glucuronide products. In summary, we proved the feasibility of producing new to nature sesquiterpene glucosides using either enzyme-based assay or bioconversion in two different hosts that are widely used in biotechnology.

To conclude, the approaches used in this thesis highlight the assets of screening promiscuous enzymes for the production of new molecules from  $\alpha$ -bisabolol. We also explored the limits of our chassis strain and a tighter regulation of *S. cerevisiae* metabolism could improve the production of (+)-*epi*- $\alpha$ -bisabolol oxidized products. Finally, the coupling in yeast of cytochrome P450 and glucosyltransferase steps can now be envisioned.

**Keywords:** Synthetic biology, sesquiterpene, metabolic engineering, cytochrome P450, *Saccharomyces cerevisiae*, enzymatic promiscuity



## Biosynthèse de nouveaux dérivés de l' $\alpha$ -bisabolol par une approche de biologie de synthèse

La biologie de synthèse permet désormais la production de nouvelles molécules n'existant pas dans la nature. Au cours de cette thèse, nous nous sommes focalisés sur la diversification de produits issus du (+)-*epi*- $\alpha$ -bisabolol. Cette molécule issue de la plante *Lippia dulcis* appartient à la vaste famille des sesquiterpènes qui recèle de nombreuses activités biologiques. En particulier, le (+)-*epi*- $\alpha$ -bisabolol est le précurseur de l'hernandulcine, un édulcorant intense. Cependant, la (les) dernière(s) étape(s) d'oxydation conduisant à ce composé demeure(nt) inconnue(s).

Plutôt que de rechercher la (ou les enzymes) d'oxydation intervenant dans la voie de biosynthèse de l'hernandulcine chez *Lippia dulcis*, nous avons choisi de mettre à profit la promiscuité connue de cytochromes P450 pour diversifier les molécules issues du (+)-*epi*- $\alpha$ -bisabolol ou reproduire la voie de synthèse naturelle. Tout d'abord, une souche châssis adaptée au criblage *in vivo* a été construite en exprimant la (+)-*epi*- $\alpha$ -bisabolol synthase (BBS) et une NADPH cytochrome P450 réductase. Puis, le criblage *in vivo* chez la levure a montré que 5 des cytochromes P450 (CYPs) parmi notre banque de 25 enzymes impliquées dans le métabolisme des xénobiotiques (CYPs animaux) oxydaient le (+)-*epi*- $\alpha$ -bisabolol. Parmi ces produits, le 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol a été purifié et caractérisé par RMN tandis que la structure probable d'un second produit a été obtenue (9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol). En parallèle, la production de produits hydroxylés (+)-*epi*- $\alpha$ -bisabolol a été optimisée notamment par l'addition d'une copie génomique supplémentaire de BBS, avec un titre de produit hydroxylé de 64 mg/L. Ainsi, nous avons démontré le potentiel des CYPs du métabolisme des xénobiotiques dans la synthèse de nouvelles molécules issues des sesquiterpènes.

Parmi les enzymes identifiées, deux enzymes homologues, CYP2B6 et CYP2B11 ont catalysé la formation de produits différents à partir du (-)- $\alpha$ -bisabolol et du trans,trans-farnésol. Pour identifier les possibles déterminants moléculaires responsables de la spécificité de chacune de ces deux enzymes, des chimères échangeant divers éléments de structure secondaire de CYP2B6 dans CYP2B11 ont été comparées. Ainsi, la spécificité ne peut pas être expliquée par un seul élément de la structure secondaire mais plus probablement par des déterminants disséminés en divers endroits des séquences protéiques des deux enzymes.

Afin d'étendre la diversité de molécules produites à partir de l' $\alpha$ -bisabolol et de l'hernandulcine et produire des composés aux propriétés physico-chimiques améliorées (solubilité, pouvoir sucrant...) des essais de glycosylation par une glucosyltransférase de plante (UGT93B16) et par des glucuronosyltransférases humaines ont été menés. UGT93B16 a démontré son potentiel en glucosylant l' $\alpha$ -bisabolol et l'hernandulcine. La bioconversion de ces molécules en utilisant *E. coli* et *S. cerevisiae* a également montré une meilleure conversion chez la levure. Concernant les UDP-glucuronosyltransférases, des essais enzymatiques ont mis en évidence une activité d'UGT1A9, d'UGT2B4 et d'UGT2B7 pour l' $\alpha$ -bisabolol et l'hernandulcine. Toutefois, l'introduction de ces enzymes dans la levure n'a pas été permis la production de sesquiterpènes glucuronylés à un niveau détectable. Pour résumer, nous avons prouvé la faisabilité de la production de nouveaux sesquiterpènes glycosylés soit par voie enzymatique soit par bioconversion chez deux microorganismes couramment utilisés en biotechnologie.

En conclusion, les approches utilisées au cours de cette thèse ont montré l'intérêt du criblage d'enzymes promiscuitaires pour l' $\alpha$ -bisabolol et la production de nouvelles molécules. Nous avons exploré les limites de notre souche châssis et une régulation plus fine du métabolisme de *S. cerevisiae* pourrait améliorer les titres obtenus. Enfin, le couplage des étapes impliquant un cytochrome P450 et une glucosyltransférase est désormais envisageable afin de créer une voie encore plus orthogonale.

**Mots clés :** Biologie de synthèse, sesquiterpènes, ingénierie métabolique, cytochrome P450, *Saccharomyces cerevisiae*, promiscuité enzymatique



## Remerciements

Débuter une thèse, c'est partir pour une longue aventure, les péripéties sont nombreuses et le chemin parfois escarpé. Être bien entouré aide à dépasser ses limites. Je tiens donc à remercier tous ceux qui de près ou de loin m'ont aidé au cours de ce beau projet !

Tout d'abord, je souhaite remercier Gilles Truan, en tant que directeur de thèse de m'avoir confié ce sujet ambitieux ou tant de choses restaient à mettre au point au laboratoire. L'autonomie que tu m'as laissée sur la thèse m'a, je l'espère, permis de grandir scientifiquement et d'apprendre à conduire plus efficacement un projet. Magali Remaud-Siméon, merci pour ton implication comme co-directrice de thèse, ta rigueur et ton esprit de synthèse permettent de toujours sortir d'une discussion avec plus de confiance sur ce que l'on doit faire. Merci de m'avoir aidé à avancer quand la thèse devenait parfois tentaculaire ou que trop de chemins étaient possibles. Et je n'oublierai jamais ce congrès « Enzyme engineering » organisé à Toulouse où j'ai pu écouter une future prix Nobel !

Je remercie l'EAD11 et ses membres successifs, Yoann Brison, Romy Honorine, Marie Carquet, Florence Bonnot, Eirik Kommedal qui ont partagé leur savoir-faire et m'ont aidé à mon arrivée. Mais aussi Christel Boutonnet, Nuria Ramos, Hery Rabeharindranto, Robert Quast, Donna-Joe Bigot, et Laure Souilles que j'ai encadré en stage de M1 et tant d'autres qui se sont succédés, la liste est longue ! Merci également à Denis Pompon, votre connaissance encyclopédique est déroutante et échanger avec vous a été un véritable plaisir. Grâce à vous, je partirai avec une connaissance moins sommaire de la chromatographie liquide et vos réparations à la main m'impressionneront toujours. Philippe Urban, merci de ton aide pour la partie sur les chimères CYP2B6/CYP2B11, et ta façon de voir le verre toujours à moitié plein est bien agréable. Merci Luis pour tes conseils sur la chimie, même quand tu étais débordé tu t'es rendu disponible. Thomas Lautier, mon premier projet à l'EAD11 était avec toi, et merci pour m'avoir encouragé à m'engager dans une thèse. Une mention très spéciale à Mathieu Fournié et Nathalie Aubry sans qui la thèse m'aurait sûrement semblé (encore) plus longue ou plus difficile. Nos discussions ont égayé tant de journées, j'espère que l'on continuera à se voir autant après la thèse ! Mathieu, avoir débuté en même temps que toi m'a parfois donné l'impression de faire une thèse en binôme, on a pu se serrer les coudes dans les aléas d'une thèse et merci pour toute ton aide, et sans ton côté bricoleur je n'aurais pas pu m'en sortir sur ces damnées HPLCs ! Merci également à l'EAD15 et plus spécifiquement Anne-Laure, Gwen, Gilles, Denis pour m'avoir tenu compagnie dans ce bâtiment Bio3, un peu isolé.

L'EAD1 vous êtes si nombreux, et je vous remercie pour votre accueil à chaque fois que je suis venu. En particulier, merci Julien Robin pour ton aide avec la chromatographie gazeuse. Merci à la fameuse « ligue », Yassim, Maher, Mounir, Manon et Catherine, se croiser au LISBP c'était bien, se voir en dehors, c'était encore mieux ! Merci également à Emeline et Emilie, pour ces bons moments et anecdotes partagées, et ce mariage en terre mosellane !

Je tiens à remercier également Véronique de Berardinis et Jean-Etienne Bassard pour leur évaluation de ce manuscrit de thèse en tant que rapporteurs ainsi que Danièle Werck-Reichhart et Fayza Daboussi pour avoir accepté de participer à mon jury de thèse.

Merci à ceux que je n'ai pas pu citer car la liste aurait été trop longue. Et pour finir merci à ma famille et mes amis toujours présents quand j'ai eu besoin de leur soutien.

“Pour ce qui est de l'avenir, il ne s'agit pas de le prévoir, mais de le rendre possible.”

Antoine de Saint-Exupéry

Papa, je te dédie cette thèse, tu m'avais toujours soutenu dans ma curiosité scientifique et *in fine* vers la recherche et cette thèse.



## PUBLICATIONS

---

- Synthetic Derivatives of (+)-epi- $\alpha$ -Bisabolol Are Formed by Mammalian Cytochromes P450 Expressed in a Yeast Reconstituted Pathway. Sarrade-Loucheur A, Ro DK, Faure R, Remaud-Siméon M, Truan G. *ACS Synth. Biol.* 2020, 9, 2, 368–380
- Recombinant bioconversion using UGT93B16 leads to new glucosides from (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin. Sarrade-Loucheur A, Remaud-Siméon M, Truan G. In preparation

## ORAL COMMUNICATIONS

---

- Sarrade-Loucheur A, Remaud-Simeon M, Truan G. Biosynthesis of hernandulcin derivatives, a natural sweetener, through a synthetic biology approach. Biosynsys 2017, 18-20 October 2017, La Grande Motte, France

## POSTER COMMUNICATIONS

---

- Sarrade-Loucheur A, Remaud-Simeon M, Truan G. Biosynthesis of hernandulcin derivatives, a natural sweetener, through a synthetic biology approach. Biosynsys 2016, 27-29 June 2016, Bordeaux, France
- Sarrade-Loucheur A, Remaud-Simeon M, Truan G. Biosynthesis of hernandulcin derivatives, a natural sweetener, through a synthetic biology approach. Microbiocccitanie 2017, 24 April 2017, Toulouse, France
- Sarrade-Loucheur A, Remaud-Simeon M, Truan G. Biosynthesis of hernandulcin derivatives, a natural sweetener, through a synthetic biology approach. Metabolic engineering 12, June 24-28 2018, Munich, Germany

## Other activities

---

- **Teaching** chemical course “Aqueous solutions” first year at INSA in 2016 and 2017
- **IGEM 2016**: practical courses to teach molecular biology to IGEM team of Toulouse
- **Science communication**: « La nuit des chercheurs » 2017 and INRA stand at « Salon de l’agriculture » 2018

## Table of content

<b>TABLE OF CONTENT .....</b>	<b>12</b>
<b>CHAPTER I. BIBLIOGRAPHIC INTRODUCTION.....</b>	<b>17</b>
<b>THESIS CONTEXT.....</b>	<b>18</b>
A. THE BIOSYNTHESIS OF TERPENES IN PLANTS .....	21
1. <i>The place of terpenes in plant metabolism</i> .....	21
2. <i>A focus at the sesquiterpene pathways</i> .....	28
a) The terpene synthase family is responsible for the scaffold synthesis.....	28
b) CYP enzymes.....	31
c) Other decorative enzymes.....	34
3. <i>The choice of hernandulcin as target</i> .....	34
B. THE RECOMBINANT PRODUCTION OF SESQUITERPENES .....	37
1. <i>Common strategies applied in metabolic engineering</i> .....	37
a) Control at the nucleic acid level .....	39
b) Control at the protein level .....	40
c) Other strategies.....	42
d) The shift induced by synthetic biology .....	43
2. <i>The rise of metabolic engineering in <i>S. cerevisiae</i> for the production of sesquiterpenes</i> .....	46
a) The native sterol pathway in <i>S. cerevisiae</i> .....	47
b) The case study of hydrocortisone metabolic engineering, derived from the sterol pathway .....	48
c) A second case study applied to sesquiterpene production, the artemisinin semi-biosynthetic production.....	49
d) Recent improvements in the engineering strategies for sesquiterpene production in <i>S. cerevisiae</i> .....	51
e) Metabolic engineering and synthetic biology access to new to nature products .....	55
<b>CHAPTER II. SYNTHETIC DERIVATIVES OF THE (+)-EPI-A-BISABOLOL ARE FORMED BY MAMMALIAN CYTOCHROMES P450 EXPRESSED IN A YEAST RECONSTITUTED PATHWAY. ....</b>	<b>59</b>
A. INTRODUCTION .....	61
B. MATERIALS AND METHODS.....	64
1. <i>Materials</i> .....	64
2. <i>Microbial strains</i> .....	64
3. <i>CYP library</i> .....	66
4. <i>Plasmid construction</i> .....	66
5. <i>Plasmid transformation and integration of DNA cassettes in yeast</i> .....	67
6. <i>Strain culture conditions and sampling</i> .....	67
7. <i>Metabolite extraction and analytical detection of the sesquiterpenes</i> .....	68
8. <i>Enzymatic assays with CYP microsomal fractions</i> .....	69
9. <i>Isolation and characterization of the bisabolol oxidation products</i> .....	69
C. RESULTS .....	70

1.	<i>Building an efficient chassis strain for screening (+)-epi-<math>\alpha</math>-bisabolol hydroxylation by CYPs</i>	70
2.	<i>In vivo screening of the CYP library in the engineered strain</i>	71
3.	<i>In vitro activities of the human CYPs with (+)-epi-<math>\alpha</math>-bisabolol, farnesol and nerolidol</i>	73
4.	<i>Structural characterization of some (+)-epi-<math>\alpha</math>-bisabolol derivatives</i>	74
a)	Improving (+)-epi- $\alpha$ -bisabolol production in chassis strain	74
b)	(+)-epi- $\alpha$ -Bisabolol <i>in vivo</i> concentration influences the oxidized metabolite production	77
c)	Product purification and structure determination	78
D.	DISCUSSION	80
E.	CONCLUSION	82
F.	SUPPORTING INFORMATION	84
G.	ADDITIONAL INVESTIGATIONS AROUND THE SYNTHESIS OF (+)-EPI-A-BISABOLOL DERIVATIVES FORMED BY CYPs EXPRESSED IN AN OPTIMIZED YEAST CHASSIS STRAIN	100
1.	<i>Material and methods</i>	100
2.	<i>Results and discussion</i>	101
a)	Further characterization of oxidized (+)-epi- $\alpha$ -bisabolol products and attempts to purify them	101
b)	Towards a finer regulation of the heterologous (+)-epi- $\alpha$ -bisabolol production in yeast	104
c)	Supplementary inputs of enzymatic assays	106
d)	Further diversification of $\alpha$ -bisabolol derivatives generated in <i>S. cerevisiae</i>	107
1)	Using plant CYPs involved in sesquiterpene metabolism	107
2)	ADHs	110
3)	The use of BBS generating other $\alpha$ -bisabolol isomers	113
3.	<i>Conclusion</i>	115
<b>CHAPTER III. CHARACTERIZATION OF CYP2B6 / CYP2B11 CHIMERAS WITH (-)-A-BISABOLOL AND TRANS,TRANS-FARNESOL 117</b>		
A.	INTRODUCTION	119
B.	MATERIALS AND METHODS	121
C.	RESULTS AND DISCUSSION	121
<b>CHAPTER IV. GLYCOSYLATION OF A-BISABOLOL AND HERNANDULCIN 127</b>		
A.	OVERVIEW OF UDP-GLYCOSYLTRANSFERASE DIVERSITY	129
1.	<i>Plant UDP-glycosyltransferases (UGTs)</i>	130
2.	<i>Human UDP-glucuronosyltransferases</i>	137
B.	MATERIAL AND METHODS	139
1.	<i>Strains and plasmids</i>	139
2.	<i>Culture conditions and sampling</i>	141
a)	<i>A. strigosa</i> UGT93B16 production	141
b)	Human UDP-glucuronosyltransferases	142
3.	<i>In vitro glycosylation assays</i>	143
C.	RESULTS AND DISCUSSION	144

1.	<i>UGT93B16: a catalyst for the glucosylation of sesquiterpenes</i>	144
a)	UGT93B16 assays from <i>E. coli</i> expression	144
b)	Bioconversion of sesquiterpenes	152
1)	In <i>E. coli</i>	152
2)	In <i>S. cerevisiae</i>	153
c)	UGT93B16 expression in an engineered yeast producing (+)- <i>epi</i> - $\alpha$ -bisabolol	156
d)	Perspectives	156
2.	<i>Human UDP-glucuronosyltransferases are able to convert (-)-<math>\alpha</math>-bisabolol and (<math>\pm</math>)-hernandulcin</i>	159
a)	<i>In vitro</i> assay with Supersomes™ microsomal fractions	159
b)	Attempts to reconstruct a functional pathway in yeast	162
c)	Perspectives	164
	CONCLUSION	166
	<b>CONCLUSION AND PERSPECTIVES</b>	<b>169</b>
1.	<i>Towards in vitro cascades to prototype new pathways</i>	171
2.	<i>Possible extension of the approach to other enzyme classes</i>	171
3.	<i>The coupling of CYP and UGT steps</i>	172
4.	<i>The biological activities of the new molecules</i>	173
	REFERENCES	177
	APPENDIX	201
A.	SUPPLEMENTARY FIGURES	202
B.	SUPPLEMENTARY TABLES	208
C.	DNA SEQUENCES	211
1.	<i>Synthetic fragments of the additional CYPs</i>	211
2.	<i>Synthetic fragments corresponding to the tested ADHs</i>	213
3.	<i>Synthetic fragments of the supplementary BBS</i>	215
4.	<i>Coding sequence of UGT93B16</i>	216
5.	<i>Synthetic fragments corresponding to TEF1 promoter and coding sequence of Rattus norvegicus UGDH</i>	216
6.	<i>Synthetic fragments of the human UGTs active with (<math>\pm</math>)-hernandulcin and (-)-<math>\alpha</math>-bisabolol</i>	217
7.	<i>Synthetic fragments of UDP-GlcA transporters</i>	219
D.	LIST OF FIGURES	221
E.	LIST OF TABLES	225
F.	ABBREVIATIONS	226



## Chapter I. Bibliographic introduction

## Thesis context

*Lippia dulcis* is a plant present in Central America whose leaves have a strong sweet taste. The molecule involved in this valuable feature was isolated and named hernandulcin. This sesquiterpene is a thousand times sweeter than sucrose and only the (+)-*epi* isomer is perceived as sweet<sup>1,2</sup>. However, its use in food industry is hindered by several characteristics: a low content from its natural source<sup>3,4</sup>, a bitterness aftertaste and a weak solubility. Metabolic engineering and microbial production could solve the low yield and the stereoselectivity issues. Indeed, metabolic engineering is emerging as an elegant and sustainable alternative to natural product extraction and/or chemical synthesis to obtain commodities from renewable carbon sources. Parallely to metabolic engineering, synthetic biology also offers new approaches, notably, redesigning metabolic pathways using combinatorial libraries of enzymes, thus enabling the production of new compounds not present in natural pathways<sup>5</sup>. Hence, obtaining such new molecules could improve the solubility and / or get a better gustative profile for hernandulcin derivatives.

In *L. dulcis*, the biosynthesis of hernandulcin branches on the mevalonate pathway which is ubiquitous in eukaryotes. (+)-*epi*- $\alpha$ -bisabolol, the direct precursor of hernandulcin, is synthesized from farnesyl pyrophosphate (FPP) by the (+)-*epi*- $\alpha$ -bisabolol synthase (BBS) (Figure 1). Then, some oxidative steps lead to the insertion of a keto group at the C-2. Thus far, the enzyme(s) responsible for this (these) modification(s) are not discovered. Owing to the similarity with other sesquiterpene pathways and the preponderance of cytochromes P450 (CYPs) in plant secondary metabolism, the reaction may be catalyzed by an uncharacterized CYP<sup>6</sup> followed (or not) by an alcohol dehydrogenase (ADH) that may convert the alcohol group into a keto one<sup>7</sup>.

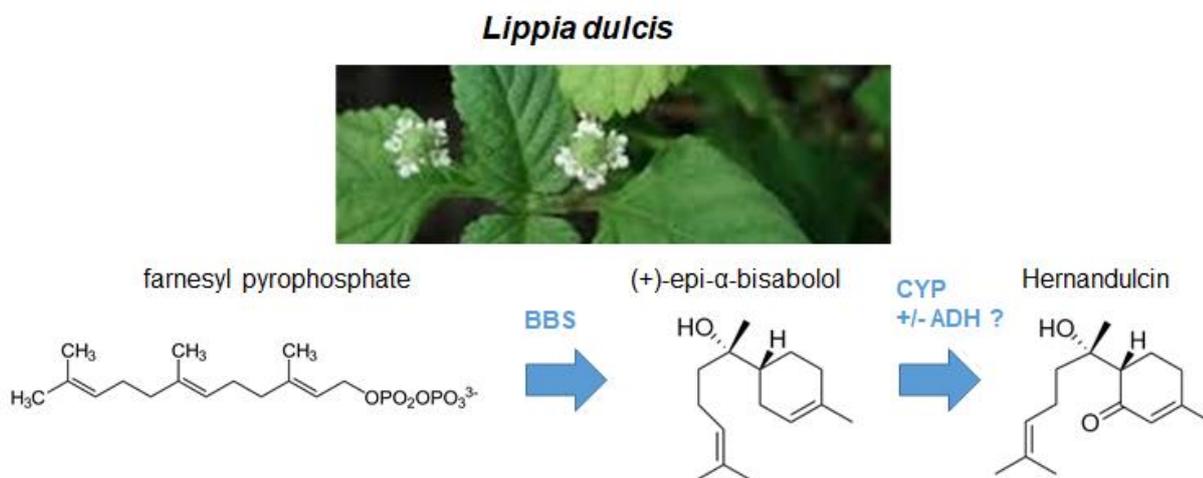


Figure 1: The biosynthesis of hernandulcin, the sweetening molecule isolated from *Lippia dulcis* with the putative presence of a CYP and/or a CYP and an ADH.

Plants possess a great diversity of CYPs. For example, *Arabidopsis thaliana* possesses 245 CYPs and *Oriza sativa*, 344. CYPs are amongst the most prominent gene families and are spread all over their genome<sup>8</sup>. This multiplicity comes from their implication in the secondary metabolism where they build a large repertoire of molecules. Owing to their broad occurrence, deciphering orphan biosynthetic pathway requires specific studies. Currently, the most used tools for enzyme discovery include genomics and transcriptomics. They are efficient in obtaining relevant data about the genes involved in the biosynthesis and tissue specific information can lead to a subset of gene candidates<sup>9,10</sup>. Biochemical studies can also be complementary for non-model organisms<sup>11-13</sup>. Still, screening of the obtained gene candidates is cumbersome and many pathways are far from being resolved. An alternative and powerful strategy could consist in screening enzymes from alternative sources and assessing their potential in fulfilling the missing reactions. In that design, the enzymes should be promiscuous enough to recognize the “foreign” molecules.

Interestingly, plants developed their chemical secondary metabolite arsenal in order to fight herbivores and other organisms<sup>14</sup>. In response, animals developed xenobiotic metabolism to cope with these threats. Consequently, detoxification enzymes like human CYPs can recognize a tremendous number of molecules; including 70-80% of clinical drugs<sup>15</sup>. As plant CYPs and animal CYPs catalyze the same type of reactions, substituting them could replace some of the gaps existing in natural plant pathways. We decided to test this hypothesis on the hernandulcin pathway for which the final oxidative enzyme(s) are unknown. Furthermore, terpenes are versatile and their diversity is linked with many biological activities<sup>16</sup> while synthetic biology approaches can now be used to discover new to nature molecules<sup>17,18</sup>. Therefore, the approach we envisioned in this thesis has the potential to generate previously inaccessible compounds with similar or new biological activities in addition to our primary goal, which was to reconstitute the natural pathway.

As one of the objectives of this thesis is to get access to an efficient microbial production of these new molecules, the screening of the oxidative enzymes was implemented *in vivo*. We chose *S. cerevisiae* as microbial host, as it possesses several advantages over other organisms. As CYPs are membrane-bound enzymes that require targeting to the endoplasmic reticulum for optimal functioning, the use of a eukaryote organism is preferable over bacteria. In addition, *S. cerevisiae* is a model organism for which many available genetic tools for recombinant protein expression are available, and it has a long history of expression of the redox partners of mammalian CYPs. Metabolic engineering massively used yeasts for biofuel and for high value added compounds productions<sup>19-21</sup>. It is especially the case for sesquiterpenes with the well-known examples of artemisinic acid and farnesene<sup>22,23</sup>. Due to its anti-malarial use, artemisinin caught a massive interest (its discovery was also rewarded by a Nobel prize in 2015 to Tu Youyou). Our work benefits from this flourishing

and historical context with previous studies that paved the way for easier manipulation of the mevalonate pathway in yeast <sup>24,25</sup>. Lastly, some molecules produced in recombinant yeast have successfully reached industrialization, like artemisinin with a semi biosynthetic process, validating our microorganism choice.

In order to improve the solubility of the molecules originating from (+)-*epi*- $\alpha$ -bisabolol and to further diversify the new molecules we thought to introduce a glycosylation step. This would mimic some existing plant pathways where hydrophobic terpenes are finally glycosylated for storage or protection from toxicity of the aglycone <sup>26</sup>. Because the full pathway leading to hernandulcin is currently unknown we had to consider the glycosylation in a separate approach. Either using enzymatic *in vitro* assays that could establish a proof of concept for possible biocatalysis or using bioconversion by adding exogenously the molecule to a strain expressing glycosyltransferases able to convert sesquiterpenes. We selected two different families of enzymes for this objective. On the one hand, plant UDP-glycosyltransferases that are cytosoluble enzymes already used by plants for glucosylation of secondary metabolites. Some of their representatives are promiscuous and permits their use as novel biocatalysts <sup>27</sup>. On the other hand, we considered human UDP-glucuronosyltransferases that are the detoxifying partners of human CYPs in the xenobiotic clearance. These membrane-bound enzymes are involved in the glucuronidation of xenobiotic compounds <sup>28</sup>. The final perspective of this work would then be to couple the oxidation and glycosylation steps *in vivo*.

## A. The biosynthesis of terpenes in plants

### 1. The place of terpenes in plant metabolism

The number of molecules synthesized by plants is arduous to appreciate; some studies mention 100,000 to 1,000,000 representatives, most of them being linked to the secondary metabolism<sup>29,30</sup>. Secondary metabolism of plants comprises the biosynthetic pathways not devoted to growth. The astonishing capability of plant biosynthetic pathways that would dazzle any chemist is partly due to the sessility of plants. Investigating plant secondary metabolism is of crucial importance to understand ecological relationships between plants and other organisms. To fulfill these accessory functions such as pathogen and herbivore defense, communication to other plants, attraction of commensal organism (pollinators, soil microbes...) plants have built up an extensive array of secondary metabolites<sup>29</sup>. The number of estimated plant metabolites varies due to the fact that only a small portion of the described plant species were sampled and an unknown reservoir of the plant metabolome has probably escaped to purification and characterization<sup>29</sup>. From that secondary plant metabolism, some compounds are volatile and dispersed in the environment reaching the plant neighbors, enabling interaction<sup>31</sup>. This includes several independent pathways: terpenes, phenylpropanoids, methyl jasmonate, etc. (Figure 2). The involvement of multiple classes of molecules favors the chemical diversity. In that introductory part, we will focus on the terpene family as  $\alpha$ -bisabolol and hernandulcin belong to this class of molecules.

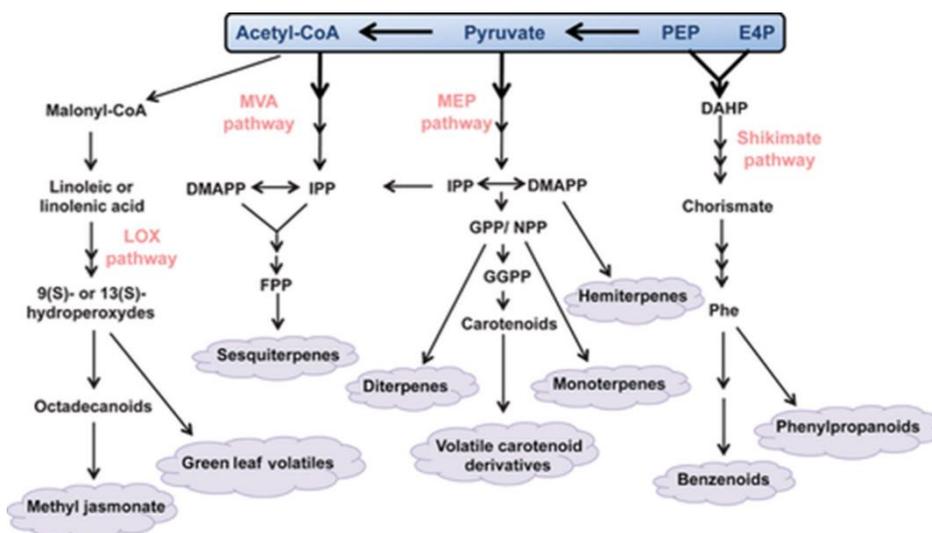


Figure 2: Overview of biosynthetic pathways leading to the emission of plant volatile organic compounds from the plant secondary metabolism (reproduced from Dudareva et al<sup>31</sup>)

Precursors from primary metabolism are represented in the blue box. (MVA) mevalonic acid, (MEP) methylerythritol phosphate, (LOX) lipoxygenase pathways. DAHP, 3-deoxy-d-arabinoheptulosonate-7 phosphate; DMAPP, dimethylallyl pyrophosphate; E4P, erythrose 4-phosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; NPP, neryl pyrophosphate; PEP, phosphoenolpyruvate; Phe, phenylalanine. Stacked arrows involve of multiple enzymatic reactions. Volatile compounds are highlighted with a purple cloud as background.

In nature, the class of terpenes has an estimated number of 50,000 original structures<sup>32</sup>. In plants, the main contributor to the terpene diversity is the secondary metabolism, but some of them also contribute to the primary metabolism such as carotenes (which are part of the photosynthesis process<sup>33,34</sup>). Terpenes are synthesized through the addition of 5 carbons units deriving from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) as theorized by the isoprene rule in the 1950's<sup>35</sup>. While terpenes were initially defined as pure hydrocarbon made of different numbers of isoprene units, these scaffolds can then be further processed by additional enzymes (CYPs, ADHs etc.). This introduces heteroatoms onto the hydrocarbon skeleton. Alternatively, sometimes the cyclization process itself leads to the introduction of alcohol function<sup>6,36</sup>. Terpenes containing heteroatoms are named terpenoids. However, more and more studies use the term terpene, or the suffix terpene to molecules that should originally be classified as terpenoids<sup>3,37-40</sup>. In this manuscript, we adopted that relaxed view of terpene definition and the reader should consider that we used terpenes interchangeably to "terpenoids" and that terpenes represent the full family. Owing to the number of isoprene units, compounds are classified in the following classes:

- Monoterpenes are C<sub>10</sub> composed of 2 isoprene units and includes for example the acyclic molecule linalol or the cyclic menthol,
- Sesquiterpenes are C<sub>15</sub> made of 3 isoprene units such as the acyclic farnesene or nerolidol and cyclic sesquiterpenes such as artemisinin or nootkatone,

- Diterpenes consist of C<sub>20</sub> originating from 4 units like the acyclic geranylgeraniol or cyclic molecules like abietic acid or ambroxide,
- Sesterterpenes are less abundant scaffolds built from 5 isoprene units <sup>41</sup>,
- Triterpene scaffolds come from 6 isoprene units (C<sub>30</sub>), as betulinic acid, or glycyrrhizin,
- Tetraterpenes (more often termed carotenes or carotenoids when linked to oxygen groups) derive from the condensation of 8 isoprene units (for instance beta-carotene, astaxanthin are C<sub>40</sub>). This class of compounds also forms products of reduced size due to carotenoid cleavage enzymes that yield to ionone (C<sub>13</sub>) or to safranal (C<sub>10</sub>), for instance.
- Some additional pathways interconnect the terpenes to other pathways like for meroterpenes <sup>42</sup>, cannabinoid biosynthesis or bitter acid pathway in hop <sup>43,44</sup>. Recently the enzymes catalyzing the critical prenylation steps in the metabolic pathway of cannabinoid and bitter acid were discovered, grafting geranyl pyrophosphate onto olivetolic acid and DMAPP or geranyl pyrophosphate onto phlorisovalerophenone /phlorisobutyrophenone respectively <sup>43,44</sup>. This interconnection between terpenes and other pathways is also evidenced in alkaloid biosynthesis where some building blocks are terpenoids <sup>45</sup>.

The tremendous diversity of terpenes is the source of a large range of applications as commodities for human use. Most of the smaller molecules are volatile and constitute natural fragrances making them highly attractive to perfumery (this includes nootkatone, santalols, linalool and many more). Food industry has an interest in some carotenes as colorants and antioxidants (lycopene, astaxanthin...), or in some diterpenes like steviol glycosides as sweeteners. Pharmaceutical and cosmetic industries identified active molecules like artemisinin, taxol,  $\alpha$ -bisabolol or squalane (derived from chemical hydrogenation of squalene). Some desaturated terpenes are also promising as fuel replacement products. A series of well-known examples along with their chemical structures is shown in Figure 3.

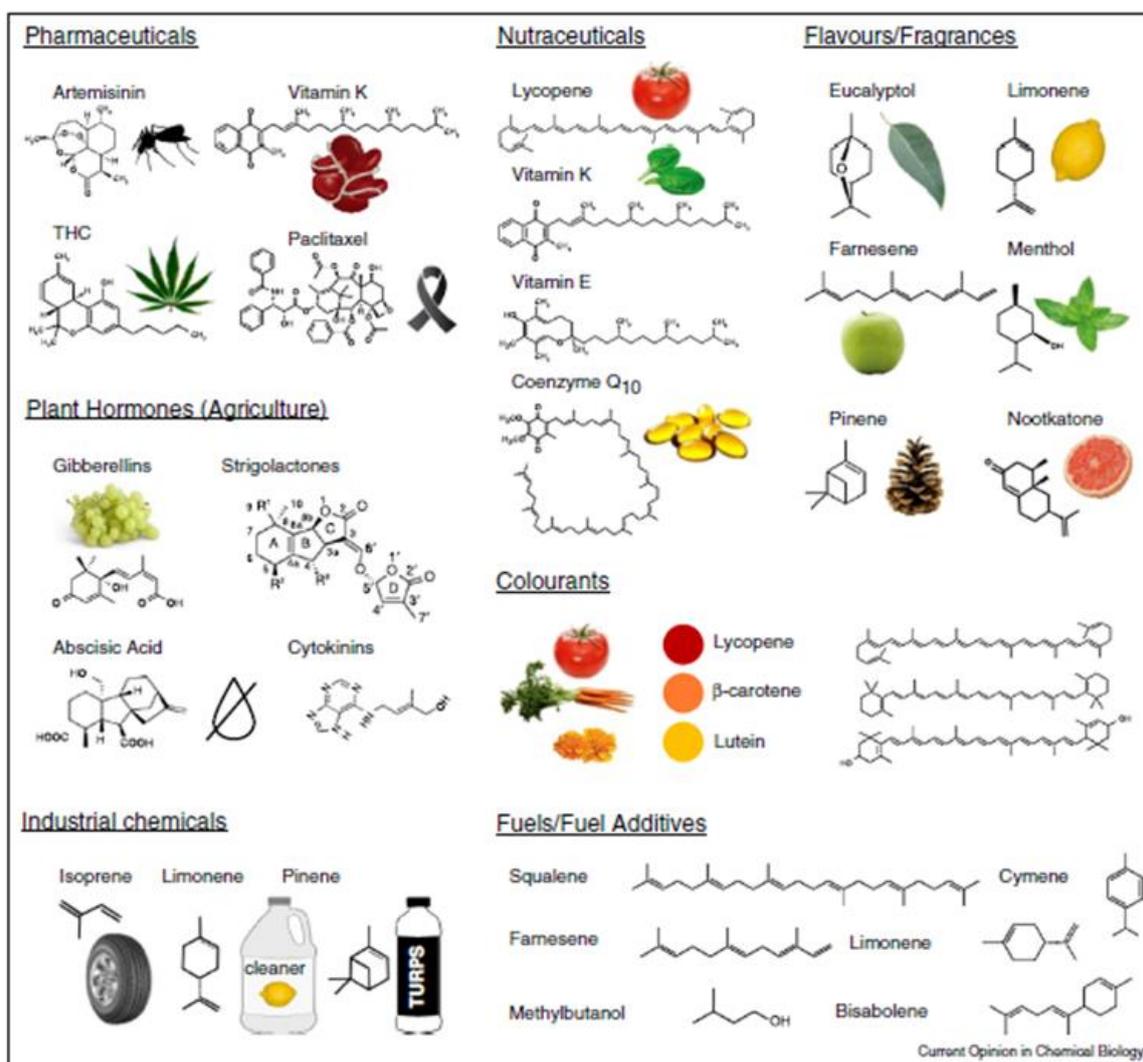


Figure 3: Examples of terpenoid products with current or potential industrial applications. This figure was reproduced from Vickers et al<sup>16</sup>

In plants, the biosynthesis of terpene precursors (IPP and DMAPP) is mediated by two separate pathways, the mevalonic acid pathway (MVA) and methylerythritol 4-phosphate pathway (MEP), respectively (Figure 4)<sup>46,47</sup>. MVA pathway is ubiquitous in eukaryotes, starts from acetyl-CoA and involves 6 enzymatic steps to produce IPP. IPP is then isomerized to DMAPP by isopentenyl diphosphate isomerase (Figure 4). While the MVA pathway is cytosolic in yeast<sup>38</sup> (Saccharomyces Genome Database, <https://www.yeastgenome.org/>, also indicates cytosolic localization to ERG8, ERG19, ERG20), recent localization studies showed that, in plants, some steps take place in the peroxisome<sup>46,48,49</sup>. In Figure 4, only one acetoacetyl-CoA thiolase is displayed. However, several isoforms exist in some plants and some of them are targeted to peroxisome<sup>48,50</sup>. Due to the complexity of plant compartmentalization and redundancy of enzyme isoforms, the localization of some enzyme components of the MVA pathway may vary depending on the organism.

The MEP pathway is localized in the plastid of plant cells and starts from pyruvate and glyceraldehyde-3-phosphate. The MEP pathway is present in most eubacteria but not in archaeobacteria, fungi, and animals, thus reflecting the endosymbiosis history of plastid formation<sup>51</sup>. This pathway with 7 enzymatic steps, also leads to the formation of IPP and DMAPP building blocks. In addition, the isopentenyl diphosphate isomerase can convert IPP to DMAPP, and vice versa. Condensation of IPP onto DMAPP generates geranyl pyrophosphate (GPP, C<sub>10</sub> molecule). Further condensation, by prenyltransferases, of IPP onto GPP produce farnesyl pyrophosphate (FPP, C<sub>15</sub> molecule), then geranylgeranyl pyrophosphate (GGPP, C<sub>20</sub> molecule). Right after, terpene synthases catalyze the cyclization or the cleavage of the pyrophosphate moiety to generate monoterpenes, sesquiterpenes and diterpenes. For triterpene formation, two FPP are condensed head to head by a squalene synthase, oxidized to 2,3-oxidosqualene before being converted by oxidosqualene cyclases<sup>36,52</sup>. For the synthesis of carotenes, two GGPP are fused in the plastid and further converted to a wide range of products. Additionally, in the monoterpene pathway, alternative precursors such as neryl diphosphate can be used. Once the terpene scaffold is formed, a battery of decorative enzymes including CYPs, ADHs, methyltransferases, epoxidases, glycosyltransferases can functionalize the core skeleton leading to a broad panel of molecules like in the artemisinin or in the mogrosides biosynthetic pathways<sup>22,53</sup>.

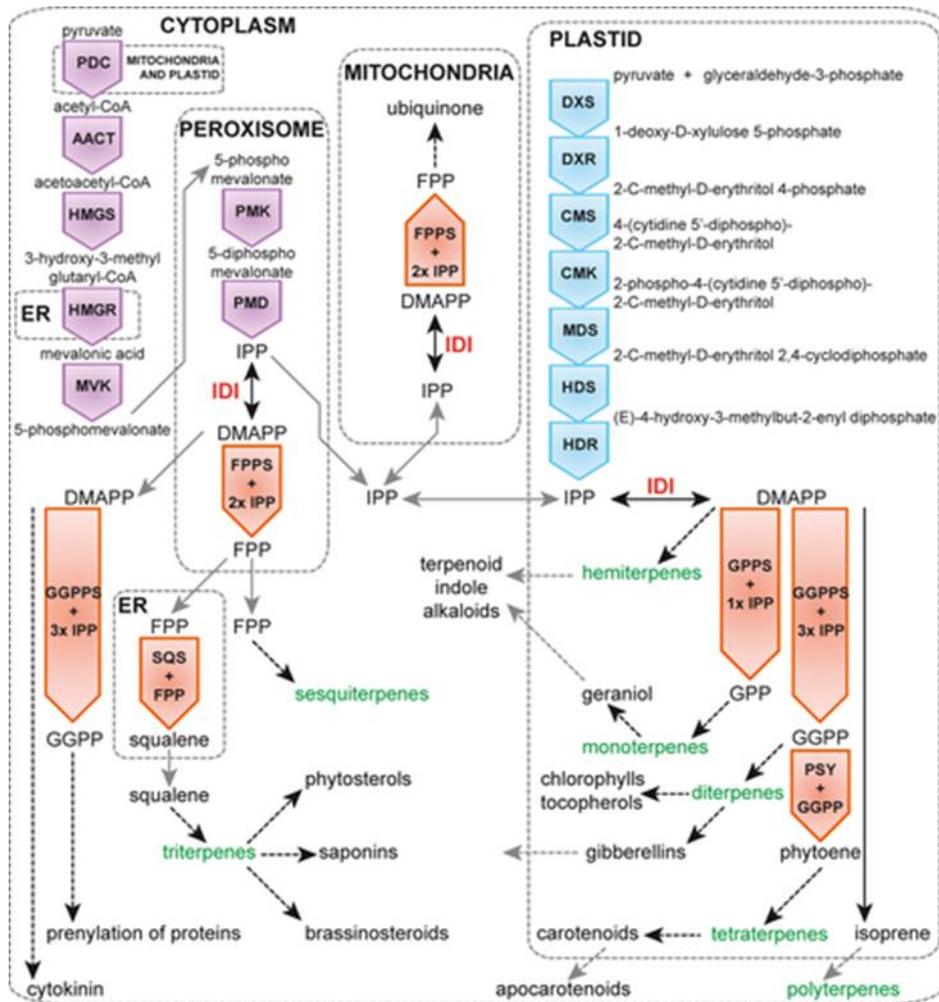


Figure 4: Plant biosynthesis of terpenes (This figure was reproduced from Moses et al<sup>47</sup>).

MVA pathway is shown in purple while MEP pathway in blue. The prenyltransferases in orange boxes generate the immediate precursors for the different terpenoid classes depicted in green. Dotted arrows indicate multiple reactions. Dotted grey boxes indicate the subcellular localization of the pathway. Grey arrows indicate metabolites that are transported between subcellular compartments. AACT, acetoacetyl-CoA thiolase; CMK, 4-diphosphocytidyl-methylerythritol kinase; CMS, 4-diphosphocytidyl-methylerythritol synthase; DMAPP, dimethylallyl pyrophosphate; DXR, deoxyxylulose 5-phosphate reductoisomerase; DXS, deoxyxylulose 5-phosphate synthase; FPP, farnesyl pyrophosphate; FPPS, FPP synthase; GGPP, geranylgeranyl pyrophosphate; GGPPS, GGPP synthase; GPP, geranyl pyrophosphate; GPPS, GPP synthase; HDR, hydroxymethylbutenyl 4-diphosphate reductase; HDS, hydroxymethylbutenyl 4-diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl pyrophosphate; MDS, methylerythritol 2,4-cyclodiphosphate synthase; MVK, mevalonate kinase; PDC, pyruvate dehydrogenase complex; PMD, 5-diphosphomevalonate decarboxylase; PMK, 5-phosphomevalonate kinase; PSY, phytoene synthase; SQS, squalene synthase.

The regulation of the terpene biosynthetic pathways is complex, varies a lot and is often not totally elucidated. Some pathways respond to environmental signals like mechanical wounding<sup>54</sup> involving effectors such as methyl jasmonate (a pleiotropic plant hormone)<sup>55,56</sup>. Other studies confirmed the inducible synthesis of several volatile terpenes including  $\alpha$ -bisabolene in *Abies grandis*<sup>57</sup> or farnesene in tea (*Camellia sinensis*)<sup>58</sup>. As the diversity of terpenes is huge, the regulators of terpenes biosynthesis may vary across terpenes that are expressed constitutively compared to the ones that are not. As an example, different regulations exist across

the sesquiterpene synthase family of *Arabidopsis thaliana*. Indeed some terpene synthases are expressed in root tissues or in an inducible manner from wound damage (At4g13280, At4g13280) while others are expressed in flowers (At5g23960 and At5g44630)<sup>59,60</sup>.

Furthermore, only a few transcription factors involved in the regulation of terpene synthesis have been characterized in details. A few noticeable exceptions are the regulators governing gossypol and artemisinin biosynthesis. Gossypol biosynthesis was shown to be regulated by GaWRKY1 (it controls the (+)- $\delta$ -Cadinene Synthase-A)<sup>61,62</sup>. Regarding artemisinin, its complex regulation network was probed in a series of studies. Several key transcription factors were unveiled including AaWRKY1, AaERF1 and AaERF2, AabHLH1, AabZIP1<sup>63-66</sup>, or recently the YABBY5 transcription factor<sup>67</sup>. Some more generic factors like jasmonate and abscisic hormones, have also an impact on artemisinin induction<sup>64,65</sup>. From those thorough analyses, it is clear that many interacting factors can govern artemisinin biosynthesis. This also underlines the tight control that plants exert on secondary metabolism. The regulation of a more complete set of terpenes by other transcriptional factors is still pending further characterization.

At the metabolic level, the existence of new metabolites originating from IPP and DMAPP dephosphorylation by hydrolases and their rephosphorylation by kinases have shed light on a more complex metabolic network than originally thought<sup>68</sup>. Additionally, the communication between cellular pools of IPP between cytosol and plastid is theorized and could assist partitioning between MVA and MEP pathways<sup>68</sup>. Focusing on the MVA pathway in plants, a similar regulation to that of other eukaryotes was noted as the enzymatic reaction catalyzed by 3-hydroxy-methylglutaryl coenzyme A reductase (HMGR) is rate limiting in the overall precursor synthesis<sup>69</sup>. Increasing the rate of this step can enhance both sesquiterpene and triterpene production<sup>40</sup>. Even though the feedback regulation of HMGR activity is similar from other eukaryotes, in plants additional control in HMGR protein degradation exists and is linked to triterpene biosynthesis as in *Medicago truncatula*<sup>70</sup>. And last but not least, another layer of metabolic control, termed metabolon, was revealed. It consists of an association of enzymes participating in the same metabolic pathway<sup>71,72</sup>. Forming metabolons enables “channeling” of the pathway intermediates, limits leakages and also prevent toxicity<sup>72,73</sup>. This was evidenced for the cucurbitacin biosynthesis, for which the interaction between a squalene epoxidase from *Cucurbita pepo* (CpSE2) and cucurbitadienol synthase improved the production of these triterpenes. The assembly of protein partners involved in a given pathway helps the plants to cope with the huge variety of enzymes, notably hundreds of CYPs, devoted to secondary pathways. Moreover, these enzymatic complexes could constitute a way to rapidly and transiently respond to environmental induction signals, with post-translational control favoring the assembly of such metabolons.

Finally, at the tissue level, specialized structures like trichomes exist in some plants in order to store and concentrate these secondary metabolites<sup>74,75</sup>. This demonstrates the sophistication of secondary

metabolism in plants where compartmentation does not limit to the inside of cells but also at the tissue scale implying cell specialization and transport systems as well. Even various trichome subtypes, differing in structures and in gene expression, occur within the same plant <sup>76</sup>. From the enzymatic discovery point of view, these dedicated structures linked to a differential expression of the biosynthetic pathway represent an opportunity to find enzyme candidates to unravel new biosynthetic pathways <sup>77</sup>.

In summary, a huge diversity of terpenes exists. Researchers did not capture yet a complete picture of the landscape as some plants produce low level of terpenes, some being restricted to specific plant tissue (such as flower, trichomes...), some being inducible or transient.

Nevertheless, the revolution induced by the Next-Generation Sequencing technologies has stunningly assisted in mapping the enzymes of more and more biosynthetic pathways. In 2013, a consortium sequencing project provided transcriptomic data of 75 non-model plants that constituted a massive input of genomic resources <sup>78</sup>. The speed increased dramatically last year when one thousand transcriptomes were released <sup>79</sup> and led to the discovery of several new terpene pathways <sup>6,9,53</sup>. Finally, the function of this arsenal of molecules is not yet fully assessed as their roles are extremely diverse and the studies to attribute their functions in plants are still quite challenging.

## 2. [A focus at the sesquiterpene pathways](#)

As  $\alpha$ -bisabolol and hernandulcin, our products of interest, belong to the sesquiterpene family a more detailed view of the enzymes involved in these biosynthetic pathways will be given.

### a) [The terpene synthase family is responsible for the scaffold synthesis](#)

Sesquiterpenes derive from the cyclization of FPP produced by terpenes synthases (TPS). In plants, TPS are coded by 20 to 150 genes (with the notable exception of moss like *Physcomitrella patens*, which encodes a single full length terpene synthase) <sup>80</sup>. Overall, the TPS family groups enzymes that use GPP, FPP or GGPP to convert them into monoterpenes, sesquiterpenes or diterpenes, respectively. The original ancestor is proposed to have arisen from a bifunctional kaurene synthase gene <sup>80</sup>. Based on phylogenetic analysis of seven plant genomes, the TPS family was divided in seven subfamilies (a, b, c, d, e/f fused to a same subfamily, g and h, Figure 5) <sup>80</sup>.

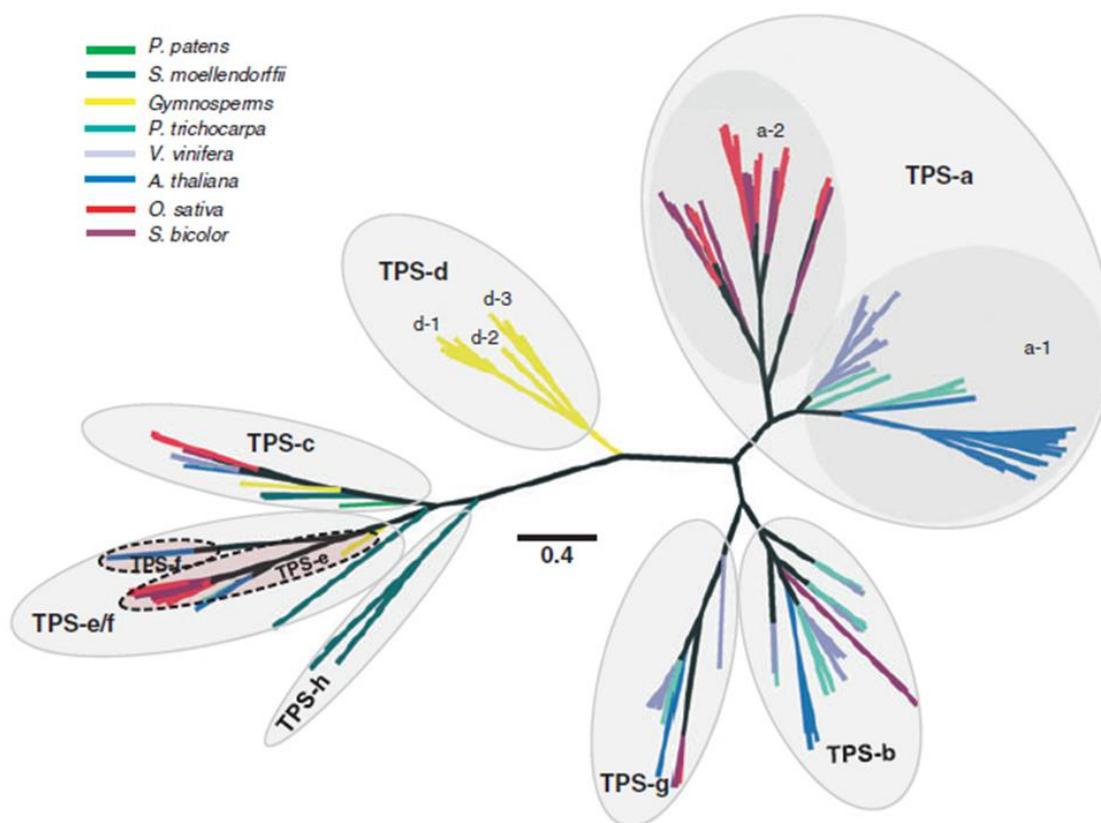


Figure 5: Phylogeny of putative full-length TPSs from seven sequenced plant genomes and representative characterized TPSs from gymnosperms. This figure was originally published by Chen et al<sup>80</sup>

Based on the phylogeny and functions of known TPSs, seven subfamilies of TPSs are recognized. These include subfamily TPS-c (most conserved among land plants), subfamily TPS-e/f (conserved among vascular plants), subfamily TPS-h (*Selaginella moellendorffii* specific), subfamily TPS-d (gymnosperm specific), and three angiosperm-specific subfamilies TPS-b, TPS-g and TPS-a. The TPS-a subfamily is further divided into two groups, a-1 being dicot-specific and a-2 being monocot-specific. The TPS-d subfamily is further divided into three groups, d-1, d-2 and d-3, which show distinction in function of TPSs in each group. The TPS-e/f subfamily is merged from the previously separate TPS-e and TPS-f subfamilies, which are also shown on the phylogenetic tree.

Typically, TPS lengths vary from ~600 to ~900 amino acids due to the loss of one domain in some of the enzyme subfamilies (Figure 6). Most plant monoterpene synthases and diterpene synthases are targeted to the plastid compartment through the presence of a N-ter signal peptide (via a RRX<sub>8</sub>W motif) while sesquiterpene synthases are generally cytosolic enzymes. Based on their reaction mechanism, TPS are further separated between class I and class II enzymes. In class I enzymes, the prenyl pyrophosphate is cleaved and a carbocation rearrangement results in the terpene cyclisation. To complete their catalysis, class I enzymes possess a DDXD and a NSE/DTE motif which are responsible for the Mg<sup>2+</sup> binding that assists in the pyrophosphate cleavage. When no water intervenes in the catalytic process of the ring closure from the carbocation intermediate, a pure hydrocarbon molecule is formed such as α-bisabolene from a bisabolyl cation intermediate<sup>81</sup>. On the contrary, when a water molecule attacks the carbocation intermediate, a sesquiterpene with an alcohol group such as α-bisabolol is synthesized<sup>82</sup>. In class II enzymes, the catalysis is

mediated by a rearrangement of GGPP through a copalyl pyrophosphate intermediate that ultimately leads to diterpene products, and the enzymes exhibit a characteristic DXDD motif<sup>80</sup>.

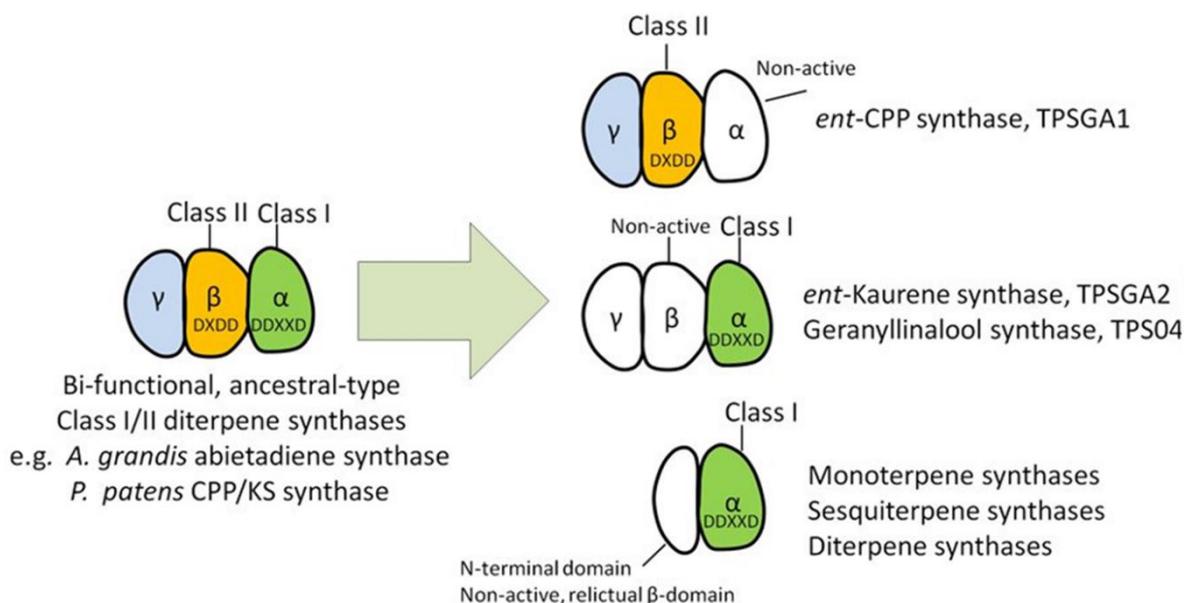


Figure 6: Protein domain structure of bi-functional class I/II TPSs and class I and class II TPSs. DXDD and DDXXD motifs are conserved in functionally active class II ( $\beta$ ) and class I ( $\alpha$ ) domains, respectively. This figure is reproduced from Tholl et al<sup>83</sup>

32 TPS were predicted from the annotation of *Arabidopsis thaliana* genome. Their characterization by several teams gives an overview of the family catalytic abilities (as summarized in the table 2 of the following review<sup>83</sup>). The current knowledge of the TPS family can be extended by focusing on specific transcriptomic approaches<sup>84</sup> or network analysis to tap particular relationships between TPS and metabolites<sup>85</sup>. Alternatively, a powerful strategy consists in an unbiased full characterization of the family at the genome level of plants, as demonstrated by the pioneering study established for *Vitis vitifera* TPS enzymes. This study predicted the function of 69 enzymes, 39 of which were assayed and active after a systematic *E. coli* based expression<sup>86</sup>. In tomato, where a similar pipeline was recently applied, the 34 TPS had their function biochemically validated<sup>87</sup>. With the availability of much more genomic and transcriptomic datasets, one should expect more and more complementary reports covering both model and non-model organisms.

At the enzymatic level, it is interesting to note that some TPS are specific to the formation of one single product<sup>6,88</sup> while others are incredibly promiscuous generating up to 52 products<sup>89</sup>. The molecular determinants associated with these differences or with enzyme specificity can be understood using classical structure/function relationships, such as directed mutagenesis studies as exemplified for amorphadiene

synthase or  $\beta$  sesquiphellandrene synthase<sup>88,90,91</sup>. Finding enzymes with high specificity (i.e. single product) is an advantage as it can avoid tedious purification steps when applied in recombinant production while working with plant extracts mean coping with complex mixtures of molecules and batch to batch heterogeneity due to seasonal variation, environmental factors and so on. Finally, characterization of a large array of TPS and discovery of efficient catalysts can serve as a base for metabolic engineering purpose driven by potential industrial applications such as jet fuel replacement in the case of sesquiterpenes derived from TPS<sup>81</sup>.

## b) CYP enzymes

Cytochromes P450 (CYPs) are proteins named after the observation of a characteristic absorption peak of these proteins at 450 nm when reduced and bound to carbon monoxide (and the letter “P” standing for pigment)<sup>92</sup>. CYPs are versatile oxidases; widespread in all life’s kingdoms even though their occurrence varies largely in specific organisms (none for *E. coli*, 3 in *S. cerevisiae*, 57 in human, much more frequent in plants for instance 245 in *A. thaliana*<sup>93</sup>). In plants, the huge diversity of encoded CYPs is correlated with the secondary metabolism and defense compounds. Besides the CYPs involved in biosynthetic pathways, animals have developed a set of CYPs in response to plant defensive compounds. These CYPs are highly promiscuous and involved in xenobiotic metabolism; in humans these CYPs are also key players in drug metabolism<sup>94</sup>.

CYPs can be directed to different cellular compartments. While most of eukaryotic CYPs are targeted to the endoplasmic reticulum, via a hydrophobic N-terminal segment, a notable exception consists in the ones involved in steroid biosynthesis which are found in mitochondria. Contrarily, prokaryote CYPs are cytosoluble enzymes. The main reaction catalyzed by CYPs is monooxygenation through the use of dioxygen. Nonetheless, they can go to oxidations up to carboxylic acid formation such as in artemisinic acid biosynthesis, perform carbon bond cleavage like in sterol demethylation, or many others<sup>95</sup>. The catalytic site of CYPs contains a heme prosthetic group bound to a cysteinyl residue of the protein with the following conserved motif:

[FW]-[SGNH]-x-[GD]-{F}-[RKHPT]-{P}-C-[LIVMFAP]-[GAD], (Prosite signature)<sup>1</sup>

Additionally, CYPs require electron donors. In the case of ER addressed CYPs, the two potential donors are the NADPH cytochrome P450 reductase (CPR) and cytochrome *b5*<sup>22,95,96</sup>. Only CPR can give the two necessary electrons for the catalytic cycle, cytochrome *b5* can only give the second one (and most CYPs even function without cytochrome *b5*). For mitochondrial CYPs, electrons are given by adrenodoxin and the recycling of electrons is carried out by adrenodoxin reductase. For bacterial enzymes, the redox partner is

---

<sup>1</sup> <https://prosite.expasy.org/PDOC00081>

ferredoxin and the recycling of electrons is mediated by a ferredoxin reductase (with a few notable exceptions in prokaryotes such as the case of BM3 that possess a CPR-type reductase naturally fused to the CYP<sup>95</sup>). The catalytic cycle of the monooxygenation carried out by CYPs is presented in Figure 7.

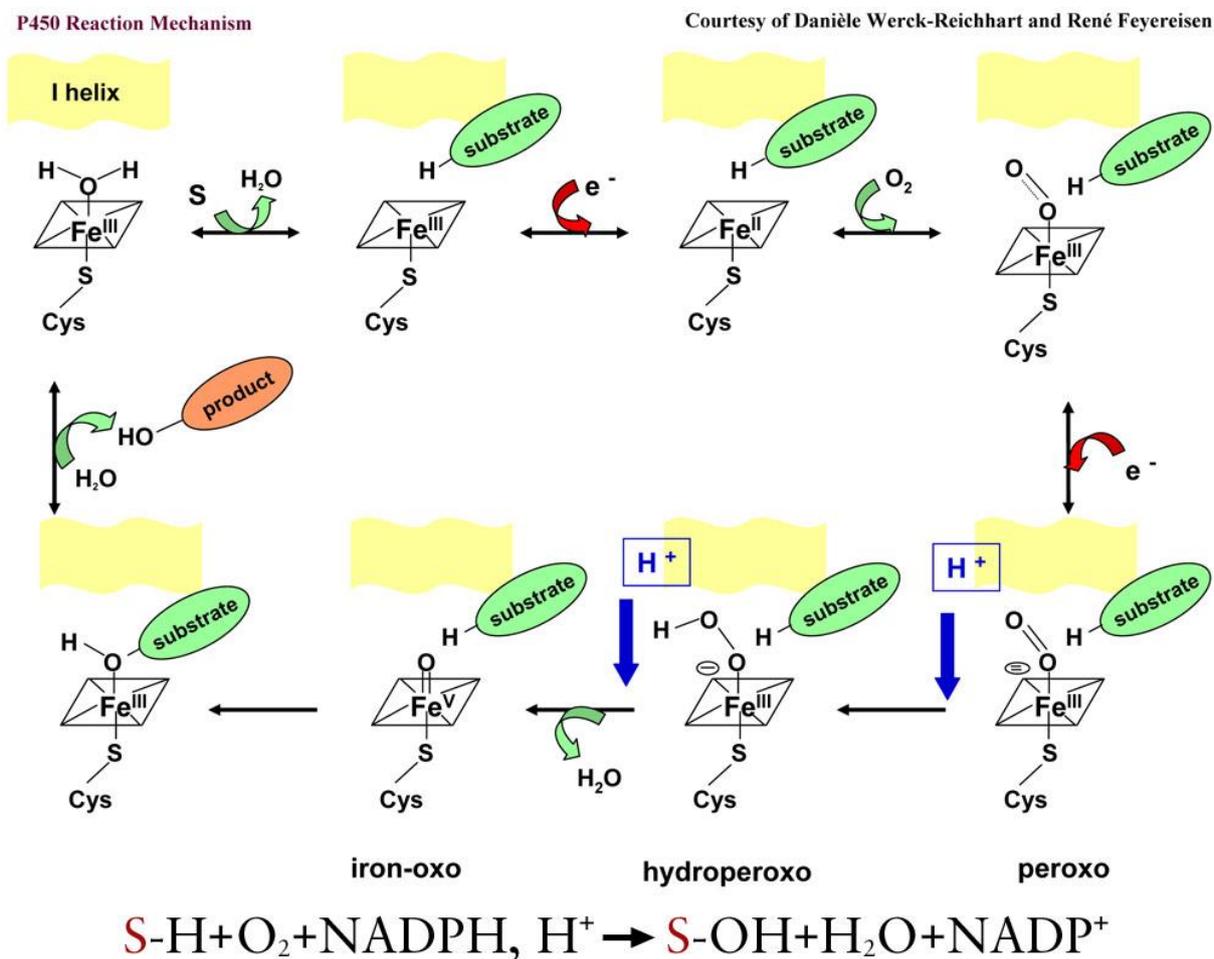


Figure 7: CYP monooxygenation mechanism. Figure originally presented by Werck-Reichhart et al<sup>97</sup>. The catalysis leads to the insertion of one of the atoms of molecular dioxygen into the substrate, the second atom of oxygen being reduced to water.

The first crystal structure of a CYP was obtained with a bacterial, cytosoluble CYP originating from *Pseudomonas putida* (P450cam, CYP101A1)<sup>98</sup>. Later on, the X-ray structures of many mammalian CYPs were solved, using a N-ter truncation strategy to obtain soluble enzymes<sup>99</sup>. While the sequence identity between distant CYP can drop to ~15%, the 3D structures obtained by crystallography showed an overall good conservation of the CYP fold. To classify the wide CYP enzyme superfamily, more than 300 000 sequences reported in 2018<sup>93</sup>, a defined nomenclature was set up. The names of the enzymes possess the prefix “CYP”, followed by a number referring to their family (enzymes sharing more than 40 % protein identity), a letter

indicating their subfamily (enzymes sharing more than >55 % protein identity), and a specific number for the isoform (for instance, the human CYP, CYP2D6)<sup>97</sup>. Regarding the CYP engineering, several strategies showed the potential to improve these enzymes, rational ones like for the bacterial CYP102A1 (formerly called BM3)<sup>100</sup> or the plant CYP71D55<sup>101</sup> as well as directed evolution strategies<sup>102</sup>.

The first characterized CYPs involved in the hydroxylation of sesquiterpenes are CYP71D20 and CYP706B1, converting 5-*epi*-aristolochene and (+)- $\delta$ -cadinene respectively<sup>103,104</sup>. Building on that, a bloom on the identification of CYP71 enzymes dedicated to sesquiterpene metabolism occurred, including as an example the renowned CYP71AV1 that catalyzes the oxidation of amorphadiene into artemisinic acid<sup>13,24,105</sup>. To discover enzymes, candidate genes are usually mined on sequence similarity approaches (i.e. belonging to CYP71 family) of transcriptomic data on specific tissues. More often a combination of these two approaches is used. Reciprocally, being a member of the CYP71 family does not imply that terpenes will be substrates, for instance CYP71B15 recognizes camalexin (indole-derived phytoalexin) as substrate<sup>106</sup> while CYP71D12 and CYP71D351 are 16-tabersonine hydroxylases (alkaloid pathway)<sup>107</sup>. Interestingly, Diaz-Chavez et al. pointed the involvement of additional alternative CYP family (CYP76) in sesquiterpene oxidation by solving santalol and bergamotol biosynthesis in sandalwood via a transcriptomic survey<sup>9</sup>. To date, the CYPs involved in sesquiterpene biosynthesis that were characterized are mainly those involved in the synthesis of natural products having industrial relevance: artemisinic acid (anti-malaria activity, CYP71AV1<sup>24</sup>), nootkatone (fragrance compound, CYP706M1<sup>108</sup>, and the promiscuous CYP71AV8<sup>13</sup>), santalol and bergamotol (fragrance molecules, CYP76Fs subfamily<sup>9</sup>), rotundone (wine aroma, CYP71BE5<sup>105</sup>), zerumbone (natural active ingredient, CYP71BA1<sup>109</sup>).

With the revolution generated by Next-Generation Sequencing technologies, enzyme identification was rewired at fast pace and most of the gene functions were attributed by transcriptomic data mining in terpene biosynthesis. Moreover, an interesting feature is the tendency, in plant genomes, for the genes involved in terpene pathways to cluster. This is observed for momilactone (diterpene)<sup>110</sup> or avenacins (triterpene)<sup>111</sup> synthesis. Bioinformatic inspection of over plant genomes by Boutanaev et al. indicates that this is a rather common feature<sup>112</sup>. In sesquiterpene metabolism, gene colocalisation was first evidenced by the vicinity of *TPS11* and *CYP706A3* in *Arabidopsis thaliana* genome, which are similarly regulated in floral tissues and ultimately contributing to the floral protection<sup>113</sup>. The presence of genomic clusters could streamline the function validation of some decorative enzymes like CYPs, not only in sesquiterpene pathways, but for all terpene scaffolds and eventually other enzymes including ADHs and UGTs. Furthermore, all the mentioned enzymes are usually hard to characterize due to the multiplicity of genes within these families and to the lack of precursors of the pathway to identify enzyme via classical *in vitro* screening assays. The recent building of several yeast chassis strains used for monoterpene production<sup>114</sup>, sesquiterpene production<sup>25</sup>,

diterpene production <sup>115</sup>, triterpene production <sup>18</sup> could help to screen large combinations of TPS/CYP tandems and assign functions at higher throughput approaches.

### c) Other decorative enzymes

Other sesquiterpene decorative enzymes include ADHs, for which a few representatives were characterized <sup>7,22</sup>, aldehyde dehydrogenases that convert keto group to carboxylic one <sup>22</sup>, reductases such as DBR2 <sup>116</sup>. Some UGTs were found to glycosylate sesquiterpenes such as farnesol or artemisinic acid <sup>117</sup> using *in vitro* assays but their role in planta was not examined. Recently, UGT91Q2A was shown to glucosylate the sesquiterpene nerolidol and play a physiological role in cold stress response of tea plants <sup>118</sup>.

A limited number of studies focused on these other decorative enzymes that may lag behind TPS and CYPs due to the fact that:

- not all terpenes are functionalized with decorative enzymes, some products generated by TPS are the end products of the sesquiterpene metabolism and are not further processed;
- decorative enzymes may act on the end product of a pathway or on intermediates that are not available from commercial sources and/or hard to synthesize chemically for *in vitro* assay.

Finally, spontaneous reactions - as exemplified by the lactonization of some sesquiterpenes (costunolide ones) <sup>119</sup> and by some photocatalytic steps occurring in artemisinin biosynthesis – undoubtedly add a degree of complexity <sup>120</sup>.

## 3. The choice of hernandulcin as target

Hernandulcin has a singular history. The Aztecs were the first to describe the sweetness of the leaves of *Tzonpelic xihuitl* centuries ago <sup>1</sup>, a plant quoted in a monograph of the Spanish physician Francisco Hernández between 1570 and 1576 and later named *Lippia dulcis*. Nonetheless, the isolation of the molecule responsible for the sweetness of *Lippia dulcis* awaited 1985 and the contribution of Compadre et al. The sesquiterpene was named hernandulcin in reference to Francisco Hernández <sup>1</sup>. The first experiments of Compadre et al showed no mutagenicity or toxicity of the compound, and revealed a 1,000 sweeter taste of hernandulcin compared to sucrose, but with “off- and aftertastes as well as some bitterness”. Moreover, only the natural isomer is perceived as sweet as proved by Mori and Kato when they established that hernandulcin

adopts a (6S,1'S) configuration, and corresponds to the (+)-*epi*- $\alpha$ -bisabolol scaffold with an additional keto group (Figure 1)<sup>2,6</sup>.

While hernandulcin accumulates in leaves of *Lippia dulcis*, botanical studies described large variations of its content in plant. Compadre et al. mentioned a low content of 0.004% w/w of the dried plant material (leaves and flowers). A much higher content was found by Adams et al. with a yield of 0.196 % from dry weight analysis (in their best extraction, using pentane condition, 2.13 % of essential oil is collected, in this oil 9.2 % is hernandulcin, so the overall yield is  $0.0213 \times 0.092 = 0.196$  %) <sup>4</sup>, and an even way higher one was reported by Souto-Bachiller et al. with around 2 % yield (4-6% of dry weight is obtained as an oil and 36% is composed of (+)-hernandulcin) <sup>121</sup>. Potential technical and biological biases may have played a role in these variations. Indeed, Compadre et al. mention a modified protocol for GC measurement of hernandulcin (due to thermal lability) <sup>3</sup>. Other studies report supplementary parameters including injection temperature, liner type of the GC equipment, split injection as possibly affecting quantitation <sup>3,4,121</sup>. Furthermore, the extraction protocol itself may affect the recovery of hernandulcin which is thermolabile, steam extraction is definitively counter indicated and leaf drying prior extraction should not be done at high temperature. In addition, Souto-Bachiller et al. proved the presence of different plant chemotypes with dramatic changes in oil composition, and the presence of the (-)-*epi*- and (+)-hernandulcin isomers in mixture as well <sup>121</sup>.

Several close relatives of hernandulcin were revealed from additional plant analyses. With another *Lippia dulcis* cultivar, Kaneda et al. isolated (+)-4 $\beta$ -hydroxy-hernandulcin, which also retained some sweetening potential. Later, some hernandulcin related metabolites were analyzed but their sweetness strength was not tested due to their presence as traces compounds <sup>122,123</sup>. Overall, this shed light on the occurrence of more complex biosynthetic pathways in some *Lippia dulcis* cultivars, not only leading to (+)-hernandulcin. At the beginning of this thesis, a team synthesized and characterized these challenging molecules (namely peroxylippidulcine A, lippidulcines A, B and C) present as traces in *Lippia dulcis* <sup>124</sup>. These new molecules do not have sweetening properties but lippidulcine A showed an interesting cooling taste.

In this context, microbial based production of hernandulcin via metabolic engineering and synthetic pathway reconstitution is an attractive alternative for several reasons. As previously seen, extraction from *Lippia dulcis* is cumbersome due to the low content and complexity of obtained oils, (presence of hernandulcin (-)-*epi* isomer, additional hernandulcin derivatives in some cultivars, camphor presence which is deleterious...). Chemical synthesis is also tedious and requires 6 steps starting from (+)-neoisopulegol <sup>124</sup>. In addition, (+)-*epi*- $\alpha$ -bisabolol and hernandulcin derivatives can display interesting properties and tastes. This is an incentive argument for the diversification of these backbones via a synthetic biology approach using CYPs or other oxygenases and for the introduction of hydroxyl groups in (+)-*epi*- $\alpha$ -bisabolol on positions never found in natural products. In addition, no glycosylated forms of hernandulcin or (+)-*epi*- $\alpha$ -bisabolol have been identified

in nature. Grafting those scaffolds with glucosyl moieties would also be valuable to generate diversity, but also to modify the hydrophilic/lipophilic balance of these molecules, enhance their solubility, and facilitate their formulation and use, which is a challenge for the poorly soluble hernandulcin (as mentioned by Rigamonti et al.<sup>124</sup>).

At the biosynthesis level, Attia et al. recently identified the (+)-*epi*- $\alpha$ -bisabolol synthase (BBS), the first enzyme of the hernandulcin synthesis pathway<sup>6</sup>. Interestingly BBS is an enzyme catalyzing the conversion of FPP to (+)-*epi*- $\alpha$ -bisabolol only, without side products, a strong advantage for metabolic engineering purpose. The possible steps leading to the introduction of the keto group on the cycle of (+)-*epi*- $\alpha$ -bisabolol have not yet been described even if transcriptomic data is available since 2012 (the one from which BBS function was attributed). As most of biosynthetic pathways showed a coexpression of the TPS and CYPs dedicated to a specific terpene<sup>9,39,113</sup> and considering the known emphasis on the CYP71/76/706 families in sesquiterpene modification, one could expect to narrow down a rather small number of candidate genes. To date, no report isolated the putative CYP, meaning that the target CYP could be either poorly transcribed or its expression pattern is independent of the BBS one. Alternatively, additional oxidase families not linked to CYP should be considered too.

From the plant physiology point of view, it would be of great interest to obtain a better knowledge at the tissue level of *Lippia dulcis* regarding hernandulcin biosynthesis. *Lippia scaberrima* (same genus as *Lippia dulcis*) has trichomes structures<sup>125</sup> checking if it is the case for *Lippia dulcis* would be a fast step (from visual inspection of the leaves, it seems likely, data not shown). To identify the enzymes involved in the conversion of (+)-*epi*- $\alpha$ -bisabolol to hernandulcin, a more precise picture of the biosynthetic pathway in *Lippia dulcis* is needed and could be obtained by i) testing if trichomes are specific tissues for hernandulcin production and assessing other plant parts ii) acquiring transcriptomic data on the productive tissues (and non-productive ones as control eventually) or more rapidly amplify the most prominent expressed CYP by degenerate primers. In the past, such strategy of trichome specific inspection greatly helped enzyme identification when biosynthesis is active at this particular localization<sup>24</sup>. Noteworthy, the physiological role of hernandulcin was neglected and no study was published on this specific question. Is it inducible by methyl jasmonate? Do environmental conditions influence its accumulation? Is it further conjugated to glycosides as other terpenes? It would definitely be interesting to uncover its potential role in accordance with pollinators or conversely with herbivores.

## B. The recombinant production of sesquiterpenes

Metabolic engineering emerged in the 1990's and aims at optimally adjusting a biological system to maximize the production of a target compound (metabolites, protein...) <sup>126,127</sup>. To that end, simply overexpressing the biosynthetic pathway is often not the most efficient strategy, due to the existence of feedback regulations <sup>128</sup>, possible toxicity of intermediates or final products <sup>129</sup>, side reactions occurrence <sup>19</sup>, metabolic crosstalks, etc.. In order to overcome the aforementioned issues, metabolic engineering benefits from its pivotal position in biotechnology with inputs of molecular biology tools (to perform selected modifications in the genome and alter pathway regulation), enzymology and biochemistry (to identify limiting steps and overcome the limits using enzyme engineering, design of protein complexes, etc.), system biology (to acquire a broader view at the cell level), physiology, modelling. Solutions can then be proposed using the best suited host among microbial organisms, human cell lines or plants, cell free and by concerted control at the nucleic and/or at the protein levels.

First, the generic tools existing in metabolic engineering will be described, then a focus will present the specific literature describing the production of sesquiterpenes in *S. cerevisiae*.

### 1. Common strategies applied in metabolic engineering

To optimize the production of a product naturally synthesized by a given organism, a first usual step is to vary the culture conditions and to examine what are the parameters affecting the accumulation of products (temperature, aeration, carbon sources...). In parallel, it is usual to test several strains to check whether some supplemental abilities preexist in specific genetic backgrounds of the same specie <sup>130,131</sup>. With this data, a first general view of the metabolic fluxes can be performed. Nodes are the key players in understanding how metabolic fluxes are distributed between several branches, and, as an example, changing the carbon source from glucose to gluconate can reshape the primary metabolism in *Corynebacterium glutamicum* and provide data about the possible partition of the carbon fluxes within the different pathway branches <sup>127</sup>.

Using the knowledge of the enzymatic processes present in the host, metabolic fluxes can then be established by quantifying the diverse metabolites. Using specific carbon labeling with <sup>13</sup>C can assess the weight of different pathways and enables flux balance analysis (or other chemical elements depending on the path to follow) <sup>132</sup>. From this valuable picture, rational choices can be made to redirect the flux towards the

expected products. In addition, several computational tools can guide the strategies due to the complexity of metabolic systems and the influence of distant branches that could interact with a specific target <sup>133,134</sup>. Approaches can be based on the simulation of host metabolism and adjusting the obtained network as in Optflux tool <sup>134,135</sup>, or use chemistry inspired strategy relying on a retrosynthetic approach as developed in the Retropath software <sup>136</sup>.

Another development regarding carbon sources also consists in expanding the abilities of a microorganism, to use alternative ones. Such endeavor particularly progressed with *Saccharomyces cerevisiae* with biomass valorization towards ethanol production. While *Saccharomyces cerevisiae* is able to produce large amounts of ethanol from glucose, it does not naturally degrade xylose nor arabinose, so better carbon utilization can be implemented through metabolic engineering <sup>137</sup>.

Approaches to improve the production of a target compound can be separated into two main classes. The unbiased ones that are supported by a physiological screen <sup>138</sup> or that uses omics techniques to obtain a global picture of the system <sup>38</sup> and the rational ones that directly target a specific step or a few ones <sup>128</sup>. Adaptive laboratory evolution is one of the most used method that allows to work without *a priori* hypotheses. In such experiments, the strain is maintained under a controlled selection pressure allowing the new phenotype of interest to emerge. In a very elegant paper, Caspeta et al. demonstrated the strength of this technique by growing yeast cells over 90 days (>300 generations) and selecting cells with increased thermotolerance (>40°C) <sup>138</sup>. Ultimately, the objective is also to determine the underlying mechanism that gave rise to the new strain phenotype. In the report of Caspeta et al., the thermotolerance could be attributed to a modification of the sterol content of the new strain. While in the wild type strain the most prominent sterol is ergosterol, fecosterol was the major one in the adapted strain. Similarly, other phenotypes can be screened such as better resistance to solvents, increased tolerance to product toxicity, inhibitors, etc <sup>139</sup>. Besides adaptive laboratory evolution, another possibility is to use a rapid screen and probe the full collection of *S. cerevisiae* single deletion mutants <sup>140</sup>. Doing so, Triikka et al. took advantage of the color pattern from carotenoids to visually identify the improvements in heterozygous strains having single gene deletions. They further pushed their analysis through several iterative cycles to obtain cumulative effects when possible. A last expansion of their system was to cross validate the strategy with sclareol, a diterpene that grafts similarly to carotenes onto yeast metabolism (at the GGPP node of MVA pathway).

Optimization of production yields of a specific target can also be based on previous knowledge or from computational tools <sup>135</sup>. For instance in mevalonic acid pathway, it is widely accepted that HMGR is a limiting step <sup>69,128</sup>. If no data is available, the strategy is to try to shed light on the existing bottleneck(s) and to identify the so-called limiting enzymatic step(s) of the pathway. Controlled expression of the enzyme(s) catalyzing the corresponding reaction(s) is then attempted to enhance product accumulation. Noteworthy, in many

pathways, the regulation does not depend on a single enzyme but rather on a continuum of enzymes as brightly evidenced for tryptophan biosynthesis in yeast. By up- or down-regulating the biosynthetic enzymes TRP1, TRP2, TRP3, TRP4 and TRP5, the potential limitation(s) were uncovered <sup>141</sup>. One can then build strains to alleviate the constraints by modifying the control either at the gene regulation level or, alternatively, at the activity level of the concerned enzymes.

#### a) Control at the nucleic acid level

To control the enzyme expression of the desired pathway the most common way is to change the native promoter of a gene. Accordingly, it is possible to increase the production of an enzyme by putting its encoding gene under a strong constitutive or an inducible promoter <sup>142</sup>. On the other side, if a competitive reaction is present and decreases the yield of product formation, two options can be drawn. First, if the gene encoding the enzyme catalyzing the competitive reaction is not vital, its deletion is the simplest solution. However, if the gene is essential and its deletion cannot be alleviated by feeding the microorganism with the missing compound, then a downregulation of its expression can be envisaged, by employing a weaker or a repressible promoter <sup>143</sup>. For downregulation the tools might differ depending on the host. RNA silencing can be an option in plant or in mammalian cells <sup>144</sup>, but such mechanism of silencing is not ubiquitous <sup>145</sup>. Nonetheless, the system can be encoded in *S. cerevisiae* and serve as a new regulation tool <sup>145,146</sup>.

An emerging way to regulate efficiently the growth of a host which is impacted by the burden of metabolite production is to separate phases for biomass production and the heterologous pathway induction. Such dynamic regulations can be achieved by the use of promoters that respond to carbon sources like glucose in yeast as demonstrated for  $\alpha$ -santalene <sup>147</sup>. In the first stage, the biomass is produced from glucose consumption (when glucose is present, the HXT1 promoter is on). When glucose is exhausted, the second stage begins and at the same time the competitive pathway is turned off (no glucose anymore, the HXT1 promoter is off). This can be further extended by reprogramming the galactose regulation too. Galactose inducible promoters provide a 1,000-fold induction and are widely used due to their unrivaled strength in *S. cerevisiae* <sup>148</sup>. However, the cost of adding galactose to fermentation is too high to be industrially relevant. In the galactose regulation network, GAL80 is a repressor whose deletion leaves the galactose transcription system activated. Due to carbon preference of yeast, the GAL network is then turned on when glucose is exhausted <sup>149</sup>. Xie et al. elegantly applied this metabolic shift to the carotene overproduction, putting all the heterologous pathway under the control of galactose responsive promoters <sup>149,150</sup>. Such concept of dynamic control is sometimes referred as metabolic valve, with the underlying idea of flux control and engineering capabilities to intervene.

Another long-lasting objective is the implementation of new regulatory elements like biosensors that could respond directly to any chemical of interest (particularly the metabolites of interest), enable the monitoring of any metabolite, and, consequently, regulate the entire pathway accordingly. While some reports proved the feasibility in yeast, with tetracycline and estradiol responsive transcription factors for instance <sup>151</sup>, metabolic engineers clearly lack transcription factors that are sensitive to any desired intermediate or product and not limited to a few case examples. Efforts are spent on the discovery of more transcription factors to enlarge the array of sensed compounds and adapt it to chassis strains <sup>152,153</sup>. Alternatively, some groups turned to the molecular engineering side to reprogram known ones to other functionalities <sup>154</sup>. As nature is incredibly versatile, transcription factors are not the sole leverage to control gene expression. Chromatin structure (in eukaryotes, via histone regulation), non-coding regulatory RNA, riboswitches represent other possible control of gene expression. For example, riboswitches are particular RNA structures that can sense a compound and modulate transcription, they represent an interesting counterpart to transcription factors and an evolvable tool for new functions as well <sup>155,156</sup>.

Finally, at the nucleic acid level, it is also possible to improve heterologous expression through a few other guidelines by:

- Adjusting the codon bias. The codon usage of foreign genes may differ significantly from the well expressed genes of the host organism, leading to poor expression <sup>157,158</sup>.
- Adjusting the Kozak or Shine-Dalgarno sequences in eukaryotes or prokaryotes, respectively, as both influence protein expression <sup>159,160</sup>.
- Increasing the copy number of the gene. A high copy number strategy may end up with population heterogeneity and eventually plasmid loss when toxic effects occur <sup>142,161</sup>, and genomic integration based strategies can overcome such issues. Reciprocally, when single copy integration at the genomic level leads to low expression levels, integrating supplementary copies or random integration strategy can be a turn around <sup>159</sup>.

## b) Control at the protein level

There are several ways to control a pathway at the protein level. When introducing a heterologous pathway, one is often exposed to protein expression in an insoluble and inactive form. To solve that, testing expression of several protein homologs to choose the best expressed in the host in an empirical manner is a first possibility. Alternatively, protein solubility can sometimes be improved by fusing it to a protein partner <sup>25,162</sup>, co-expressing chaperones or using protein engineering <sup>163</sup>. Rational engineering or directed evolution

may be used to improve the catalytic efficiency of the enzyme(s) that catalyze the rate(s)-limiting step(s). After, these enzymes can be inserted in the chassis and fuel the metabolic engineering field. Mutant proteins can also be engineered to disrupt post-translational regulation such as HMG1 whose truncated form remains constantly active <sup>128,164</sup>.

At the biosynthetic pathway level, when two consecutive steps are not efficient or draw a metabolic node towards other routes, a way to directly gather two steps is to use fusion proteins. This sets a microenvironment in order to favor efficient conversion of the second step where the substrate of the second enzyme benefits from the spatial proximity of the first one <sup>165</sup>. Moreover, targeting a metabolic node could be particularly relevant when natural pathway is expected to consume significantly the competing substrate of the heterologous pathway. When natural protein fusions occur like in the carotene biosynthesis, engineering synthetic systems can reorient the pathway intermediates and increase the efficiency of the original enzymatic cascade in a heterologous host <sup>166</sup>.

Spatial confinement can also be built in a looser manner, not by physically constraining the interaction via protein fusion but using domains triggering enzyme interactions as shown by Dueber et al. <sup>167</sup>. By employing a modular scaffold dedicated to the optimization of 3 enzymatic steps in the mevalonate synthesis, the authors boosted by 77-fold the native pathway in *E. coli*. This pioneering study constructed a scaffold module where three interactor domains (GBD / SH3 / PDZ) can recruit the enzymes of the pathway. Adding more interactor copies on the scaffold allowed to vary the stoichiometry of the association and further optimize the system towards the best combination. Similar strategies were successfully applied to farnesene production or polyhydroxybutyrate biosynthesis using affibodies <sup>168</sup>, or to the mevalonate pathway using TALEN based scaffold <sup>169</sup>. The large variety of natural protein interactors and the *de novo* design of new ones <sup>170</sup> could lead to much larger library of modules to complete efficient enzyme colocalization. A valuable information would be the comparison of the different existing modules with one or two pathways as case examples to better figure out the pros and cons of the various interaction modules (distance between interaction module and enzymes, affinity ...).

These colocalization modules are directly inspired from Nature to mimic metabolons and protein complexes <sup>71,72,171</sup>. Taking advantage of natural prototypes seems logical when implementing a new pathway (Dong et al. <sup>73</sup>). However, the use of enzymes from different origins may disrupt normal and/or potential protein-protein interactions. Therefore, the presence and necessity of protein complexes need to be systematically questioned when combining enzymes from scattered origins. Finally, in a heterologous context what would be the most efficient framework, respecting natural ones <sup>73</sup> or building synthetic ones with artificial interacting domains <sup>166,167</sup>?

Totally distinct approaches recently emerged to control protein level in a predictable aspect by exploiting the host machinery that governs protein stability. For instance, Tian et al. characterized the N-ter stability of small N-ter tags by fusing it to a fluorescent protein (GFP) in *Bacillus subtilis*. They further validated such design by optimizing the expression of the enzymes leading to N-acetylneuraminic acid synthesis <sup>172</sup>. Another way to program down regulation at the protein level is to monitor protein stability inside the cells by modulating protein degradation. In eukaryote, ubiquitination directs proteasome degradation and by intentionally placing degradation tags on the protein of interest, a fine-tuning of a competing pathway can be designed. This was shown for sesquiterpene and monoterpene productions in *S. cerevisiae* and allowed to narrow down the flux towards sterol for the benefit of terpenes <sup>114,173</sup>. Similarly, with the assistance of proteases, a controllable switch was implemented in *E. coli* and permitted a regulatable protein down regulation applied to shikimate and xylonate as case examples, but this could be extend to other pathway <sup>174</sup>. Finally, a last prospect is to be able to draw from scratch post-translational control such as allosteric factors. This constitutes a grail, opening up tremendous perspectives regarding the optimization of synthetic pathways and the engineering of such devices is advancing impressively <sup>175</sup>.

While transcription control was easier to set in early designs, the recent developments and successes in the regulation at the protein level prove that multiple strategies can be applied and combined to improve chassis strains. Moreover, different response time exists at the cellular level and being controlled at the protein level may be faster than having to be processed by DNA/RNA intermediates, underlining the interest for an increased number of protein tools.

### c) Other strategies

The metabolic engineering field is evolving on an incredibly fast pace, and numerous strategies and concepts are also applied including:

- the use of consortia/co-culture to distribute heterologous pathways in different hosts or subpopulations <sup>176,177</sup>.
- the more and more systematic combination of metabolic engineering and enzyme engineering <sup>178,179</sup>.
- the development of systems allowing to alleviate product toxicity, or to pull the reaction to increase yields. There is a more frequent use of *in situ* extraction of products using i) biphasic cultures with dodecane layer to extract hydrophobic molecules such as monoterpenes and sesquiterpenes <sup>25,180</sup>; ii)

cultures in the presence of cyclodextrins to form complexes with hydrophobic products (like with triterpenes) <sup>18</sup>. Another striking example to protect from compound toxicity was the biomimetic approach used to produce indigo dye and mimic glucosylation with the plant glucosyltransferase to strongly improve microbial production <sup>11</sup>.

- The question of metabolite transport or excretion is also getting more and more attention, as in many cases, it would dramatically improve the production capabilities of microbial platforms. Uptake can be improved as shown with arabinose in *S. cerevisiae* <sup>181</sup>. However, the most critical part is product excretion as an increasing number of microbial cell factories are reaching the grams or over the >10 g/L scale of products. Using transporters for product secretion allows pulling the reaction equilibrium towards synthesis, reducing possible toxicity and facilitating the downstream processes as recently shown for fatty alcohol production in *S. cerevisiae* <sup>182</sup>.
- Targeting metabolites to cellular compartments or remodeling of organelles for metabolic engineering purpose is blooming in eukaryotes. So is the reprogramming of bacterial microcompartments from prokaryotes <sup>183,184</sup>. Spatial compartmentation could raise the productivity by creating a physical barrier for crucial intermediates and preventing their degradation as well as increasing their local concentrations. For instance, a reshaping of the endoplasmic reticulum was efficiently obtained by two different groups. PAH1 or INO2 deregulations provoked a dramatical expansion of the endoplasmic reticulum in *S. cerevisiae* <sup>185,186</sup>. While both teams aimed at triterpene biosynthesis, their plan could be extended to any pathway expressing either endoplasmic reticulum proteins or hydrophobic molecules that would “stick” to endoplasmic reticulum membranes. Arendt et al. cross-validated this hypothesis with artemisinic acid, a sesquiterpene and but also showed a negative effect when applied to carotenes <sup>185</sup>.

#### d) The shift induced by synthetic biology

In close vicinity to metabolic engineering, the field of synthetic biology recently came out <sup>187</sup>. The objective of synthetic biology is basically an extension of biotechnology and metabolic engineering but with a particular emphasis on standardization of biological elements (promoters, terminators, plasmids...). The endeavor is to achieve predictability of the different parts in order to be able to build biological devices from scratch. A major obstacle when facing biological systems via rational design is that some of the generated constructs do not yield expected outputs. Characterization as well as setting up the rules of such phenomena

is critical<sup>188</sup>. One should also be aware that when addressing to multiple enzymatic systems some overlooked factors may perturb the conclusions and complicate the analysis (such as transcription interference)<sup>188</sup>.

In synthetic biology, two angles are envisioned. “Top down” strategies intend to minimize to the known elements to their simplest forms. For instance, Craig Venter’s team quest towards synthetic minimal organism targets the definition of the simplest genome for an autonomous living microorganism and question philosophically about how life has appeared<sup>189,190</sup>. Concerning eukaryotes, the attempts were spent on rebuilding *S. cerevisiae* genome by generating synthetic chromosomes that were sequentially replaced into yeast<sup>191,192</sup> (along with massive suppression of intergenic DNA, addition of recombination sequences to favor genome plasticity, suppression of accessory genes ...). Reciprocally, Shao et al. took a different approach and completely reshaped the yeast genome and made the first eukaryote having a single circular chromosome<sup>193</sup>. Such genome redefinitions and the fact that most of the generated yeast strains grow almost normally is a massive tour de force. Last but not least, functional screening of yeast with synthetic genomes is starting with increasing ethanol tolerance<sup>194</sup>. And even more complicated phenotypes like increased metabolite productivities (carotenoids and the betulinic acid triterpene) are now under strong scrutiny<sup>195,196</sup>.

“Bottom up” strategies aim at constructing new biological bricks or modules. Such elements could divert from the original dogma of biology with orthogonal systems like orthogonal nucleic acids (also termed XNA) or the use of non-canonical amino acids that can be incorporated into proteins. Bottom up approaches also built new regulating networks as evidenced by the building of an artificial toggle switch in *E. coli* or by synthetic transcriptional regulators<sup>197,198</sup>. In this thesis, by combining enzymes from different origin (plant, and animals), we expect to follow a bottom up approach that would lead to synthetic and new sesquiterpenes that are not produced naturally.

Overall, the synthetic biology field could sustain its growth from several influences and recent changes listed below:

- Standardization of biological elements, definition of a language similar to informatic code, as proposed by Wilson et al<sup>199</sup>, following specific rules such as the IGEM guidelines used to define “biobricks”. Notably, many studies try to better characterize the biological parts such as promoters, terminators in several systems and context<sup>200,201</sup>.
- Modular cloning advances improved transposability and speeded the assembly of genes into vectors. Several new protocols emerged such Gibson assembly method, golden gate, In Fusion, etc.<sup>202,203</sup>. In parallel, vectors are getting standardized to ease their spreading<sup>204</sup>.

- The evolution of both DNA sequencing and DNA synthesis for which prices dropped dramatically. While the first human genome took decades to solve, it is now much easier to access such information. Transcriptomic and genomic dataset are literally flooding. Similarly, ordering synthetic genes is becoming common while it was unconceivable before. Highlighting this, in a so called “pressure test” of a joined synthetic biology consortium, 1.2 mega base of synthetic DNA was purchased, showing the tremendous scale up on how fast genes are synthesized, cloned and tested <sup>205</sup>.
- The rise of CRISPR technology making genetic engineering going faster and allowing the use of less conventional organisms. In addition, CRISPR can be repurposed to allow not only gene editing but also gene activation and silencing <sup>206</sup>.

*The theory of iterative improvements through a DBTL cycle standing for Design, Build, Test, Learn (*

- *Figure 8).* Iterative cycles logically appeal to common sense when dealing with engineering practices. Such principle was already present in directed evolution (rising in the 90’s) and also proposed about metabolic engineering too by Bailey in 1991 <sup>126</sup>. With synthetic biology it has become more accessible and applicable in microbial strains and will definitely benefit from biofoundries.
- The establishment of “biofoundries” that correspond to high throughput facilities dedicated to the application of the DBTL cycles <sup>207</sup>. Automation and enlarging screening capabilities in combination with AI - artificial intelligence - will probably change the habits of biotechnological developments while companies are also investing largely in screening platforms such as Amyris, Evolva, or Gingko Bioworks.

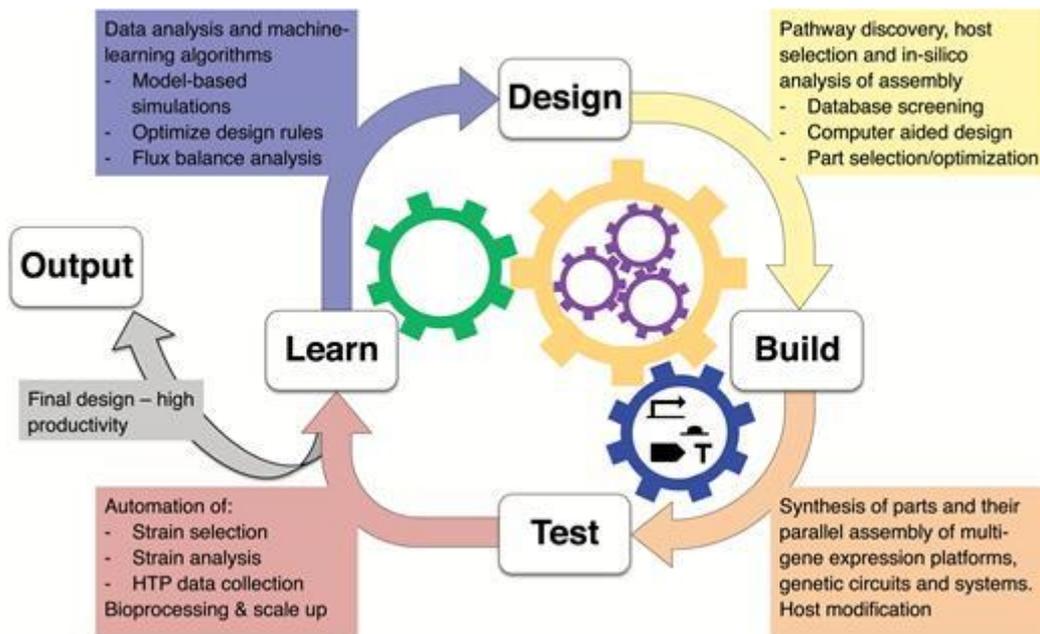


Figure 8: The design / build / test / learn cycle applied in synthetic biology (figure originally presented by Moses et al <sup>208</sup>) The workflow is standardized regardless of the DNA components or the host organism, and progresses iteratively to achieve predictable Outputs. In the Design phase, the engineer defines the goal, identifies the pathway(s) and gene(s) of interest, selects host-specific regulatory components (parts) that enable the desired expression of gene(s), and opts for an assembly approach. The design is implemented in the Build phase and introduced into the host. Evaluating the engineered strains for performance of parts (and production of target molecule(s) generates Test data, which are used during the Learn phase to analyze the components, providing an opportunity to deconstruct complex part combinations and generate knowledge, which can be used for production of tangible and predictable Output(s).

This flourishing environment then converts into massive investment in companies' sectors that may transform into the production of chemicals and natural products from microbial platforms, with around 4 billion dollars raised in 2018 according to Forbes, and 715 million dollars to Gingko Bioworks <sup>209</sup>.

## 2. The rise of metabolic engineering in *S. cerevisiae* for the production of sesquiterpenes

The use of *S. cerevisiae* started thousands of years ago with its domestication for fermented food and beverages (beer, wine, cheese, bread...) thanks to *S. cerevisiae* is innocuousness for human (*S. cerevisiae* is listed in the GRAS organisms, i.e. Generally Regarded As Safe). Besides these assets, *S. cerevisiae* is a historical model organism in genetics due to the ease of its molecular biology tools (its incredible ability to carry out homologous recombination and its transformation efficiency). With 16 chromosomes, *S. cerevisiae* genome was the first one to be sequenced <sup>210</sup>. For all these reasons, *S. cerevisiae* is a well-suitable microbial chassis for

metabolic engineering as evidenced in bio-ethanol production, artemisinic acid production etc.<sup>211</sup>. To produce sesquiterpenes, the MVA and sterol pathways of *S. cerevisiae* are the targeted pathways.

#### a) The native sterol pathway in *S. cerevisiae*

The MVA pathway of yeast and plants are similar. Starting from acetyl-CoA precursor, the major node for producing sesquiterpenes is the one occupied by the FPP (Figure 9). The pathway elucidation was performed in the 80's and 90's with enzymes being unraveled by multiple technics ranging, from homology based identification for ERG20<sup>212</sup>, or genetic complementation approach for ERG9<sup>213,214</sup>. The enzymes of the upper MVA pathway are cytosolic with the notable exception of HMG1 and HMG2 (two isoforms of HMGR are present in *S. cerevisiae*). After condensation of FPP into squalene (mediated by ERG9), the ERG enzymes are mainly distributed on the cytosolic side of the endoplasmic reticulum (ERG11, a CYP-mediated step, ERG9, ERG1...) or on the surface of the lipid bodies (ERG1, ERG7, ERG27, ERG6..) <sup>215</sup>. At the regulatory level, UPC2 and Ecm22 are global transcription factors of ergosterol in *S. cerevisiae* (involved in ERG2, ERG3 regulation, and other ERG genes targets) <sup>216</sup>. Interestingly, a particular UPC2 mutant, namely UPC2-1, showed a deregulated pattern leading to activation of the sterol pathway at a global scale, which has a clear application for increasing the flux in the mevalonate pathway <sup>217</sup>. In addition, at the enzymatic level, HMGR is a first bottleneck <sup>128</sup>, ERG20 is a second one and its overexpression also raised the sterol content <sup>218</sup>, both limitations being evidenced more than 20 years ago. Accordingly, in metabolic engineering, an overexpression of the two genes is often performed <sup>24</sup>. Other steps are targeted depending on the type of expected terpenes, downregulation of ERG9 for sesquiterpenes and diterpenes, ERG7 for triterpenes, use of ERG20 mutant changing the GPP / FPP balance in yeast for monoterpenes <sup>219</sup> ...

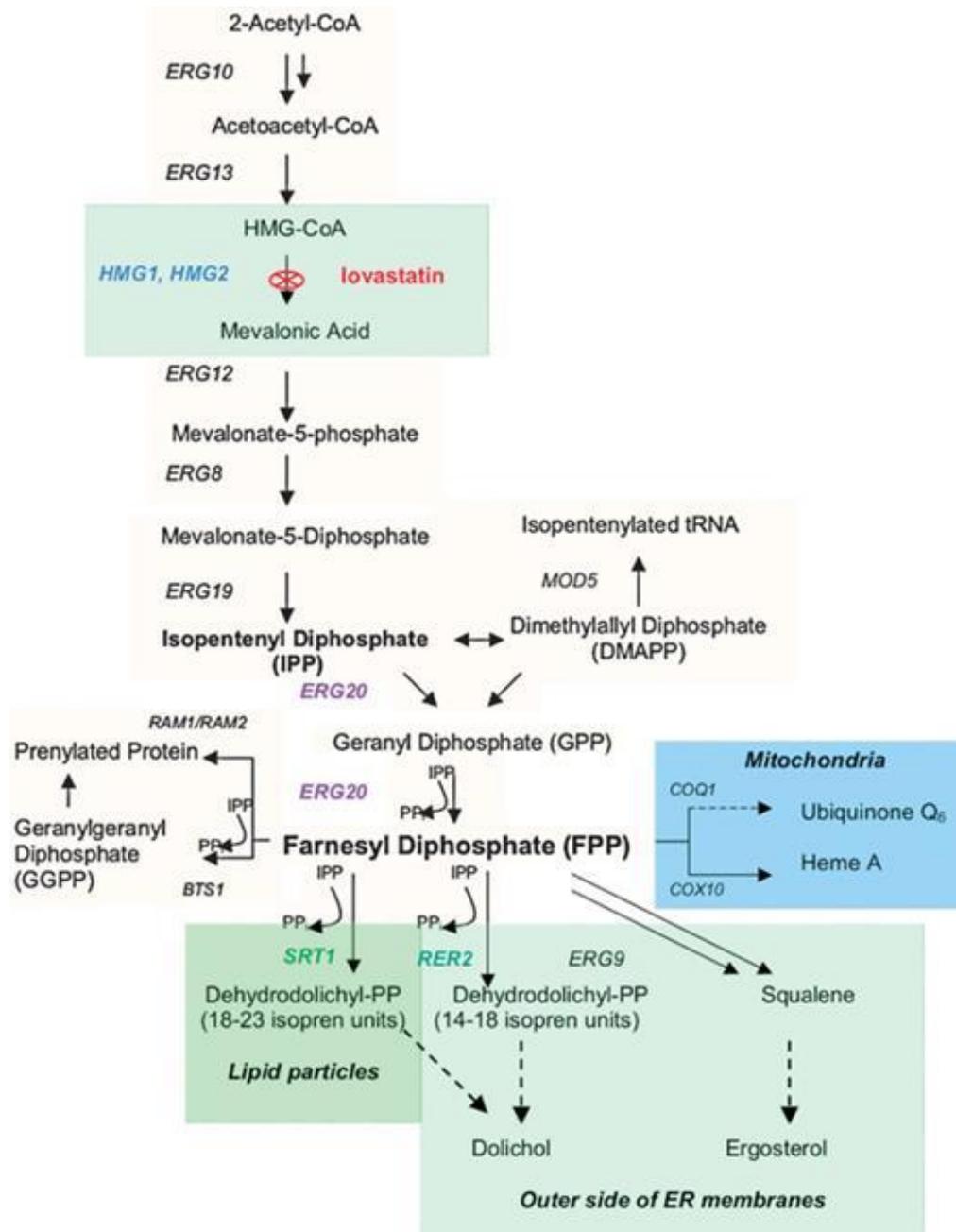


Figure 9: The MVA pathway of *S. cerevisiae* (as described by Grabińska et al <sup>220</sup>)

b) The case study of hydrocortisone metabolic engineering, derived from the sterol pathway

Described in 2003, the production of hydrocortisone from simple carbon source (glucose and ethanol) was a major breakthrough in yeast metabolic engineering <sup>19</sup>. Overall, 13 genes were engineered in the yeast, which is still quite challenging. This paved the way for future engineering strategies in yeast and made a clear

proof of concept on how metabolic engineering and synthetic biology can compete with organic chemistry, even though at the time of the study the term “synthetic biology” did not exist. This work was particularly pioneering as it tackled several aspects that are still crucial in metabolic engineering studies i. being able to drive enzymes in dedicated organelles (mitochondria for CYP11B1, CYP17A1; CYP21A1), ii. take into account accessory proteins (redox partners of the various CYPs adrenodoxin and adrenodoxin reductase, iii. Map the competitive pathways that emerge from expressing a heterologous pathway, and that can lead to unknown and unexpected side reactions (ATF2, GCY1, and YPR1 endogenous genes being found responsible).

### c) A second case study applied to sesquiterpene production, the artemisinin semi-biosynthetic production

While hydrocortisone production in yeast targeted late stages of the sterol pathway, another striking work was the production of artemisinic acid in *S. cerevisiae*<sup>22,24</sup>. This example is more in line with this thesis as it addressed the plant sesquiterpene production. The milestones reached in this project and the major coverage of the results (artemisinin treatment is a key drug to fight malaria while facing shortage for its supply) boosted the whole metabolic engineering and synthetic biology sectors.

In a first paper, Ro et al. described the construction of a chassis strain that can produce up to 100 mg/L of artemisinic acid, the artemisinin precursor<sup>24</sup>. By reaching artemisinic acid through metabolic engineering and fermentation, a semi synthetic process with chemistry completing the conversion of the last steps of the anti-malarial precursor became achievable. The key results of this work comprise i. finding the step converting amorphadiene to artemisinic acid (CYP71A1 enzyme, and the redox NADPH cytochrome P450 reductase (CPR)) ii. designing an improved strain for increased sesquiterpene accumulation. The basis of their metabolic engineering is now the standard in terpene production: integration of tHMG1 to boost the limiting step of mevalonate pathway (2 copies as only 1 can still be limiting), UPC2-1 mutant transcription factor that possesses pleiotropic effects on the engineered pathway, additional copy of ERG20, downregulation of ERG9 by MET3 repressible promoter. The overall upregulation strategy is under galactose inducible promoters, the scheme of the implemented pathway is presented in Figure 10 and resulted in the construction of the EPY300. This strain, widely known in the sesquiterpene field, served as a suitable platform strain for TPS enzyme identification or sesquiterpene production<sup>6,25</sup> and will be used in chapter II of this thesis<sup>24</sup>.

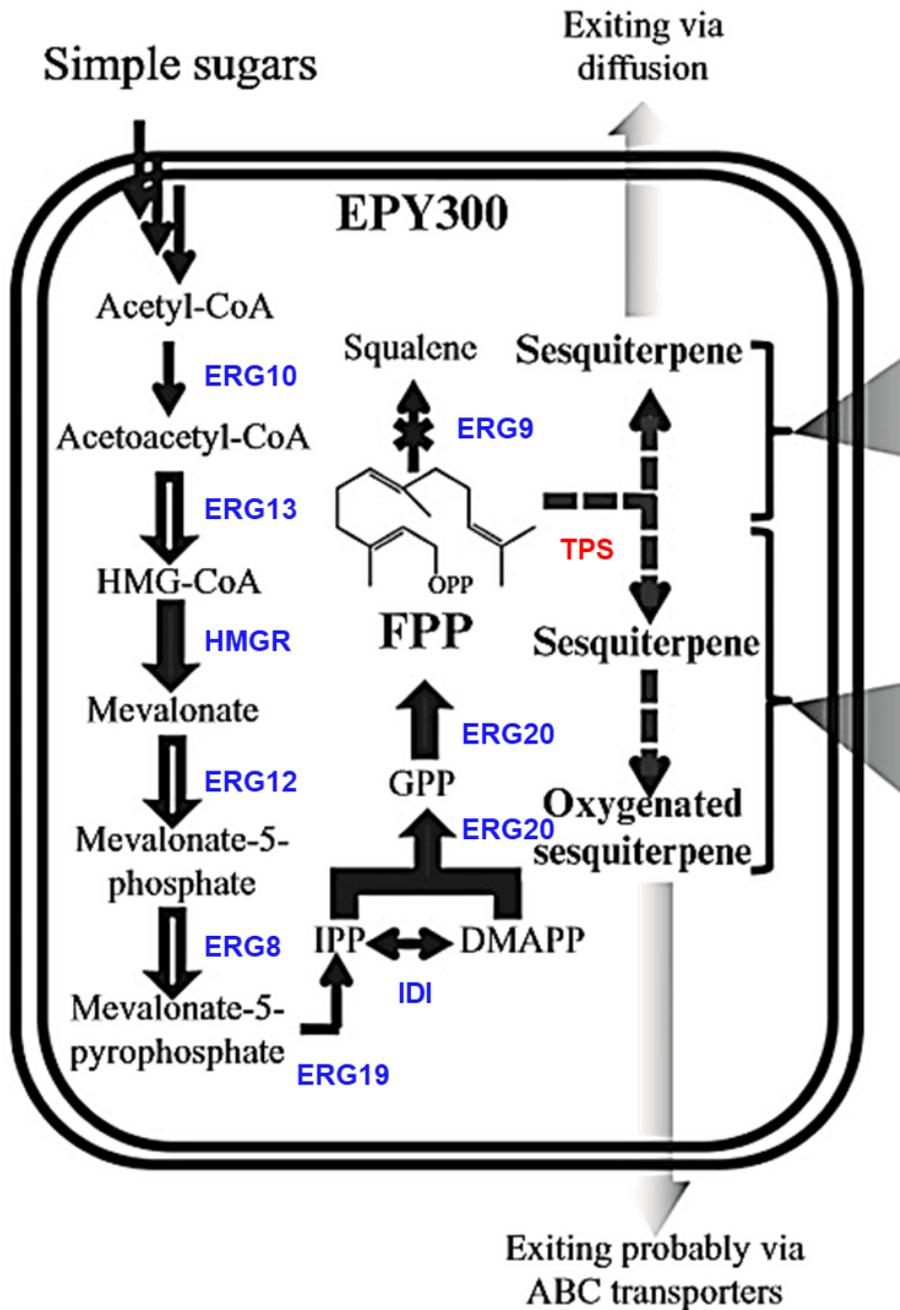


Figure 10: The genetic engineering of EPY300 *S. cerevisiae* strain suitable for de novo production of sesquiterpenoids, Figure adapted from Nguyen et al.<sup>25</sup>.

The metabolic optimization aims at augmenting farnesyl FPP pool, for its diversion to sesquiterpene production. Thick, solid arrows indicate the genes directly upregulated by GAL1-driven overexpression (HMGR truncated version of HMG1, tHMG1 and FPP synthase namely ERG20). Empty arrows are the indirectly upregulated genes by the mutant transcription factor *upc2-1*, including ERG13 (HMG-CoA synthase), ERG12 (mevalonate kinase), and ERG8 (phosphomevalonate kinase). The cross indicates the squalene synthase gene controlled under the methionine-repressible promoter. The various yeast enzymes are indicated in blue; the heterologous pathway can branch onto FPP (TPS enzyme in red).

Further efforts, overexpressing all enzymes from the mevalonate pathway down to ERG20, testing other genetic backgrounds (initially S288C vs CEN.PK2), deleting GAL80 (for derepressed galactose induction) significantly raised the production with a bottleneck from amorphaadiene conversion to artemisinic acid. Still, an impressive titer of 40 g/L of amorphaadiene was obtained <sup>221</sup>. Lastly, in a supplemental study, the discovery of missing links in the natural biosynthetic pathway from *Artemisia annua* completed the overall framework of artemisinic acid <sup>22</sup>. Three enzymes involved in the conversion of amorphaadiene to artemisinic acid escaped prior characterization, an alcohol dehydrogenase, an aldehyde dehydrogenase and a cytochrome *b5*. Adding these enzymes along with a slight reshaping of the metabolic engineering design ended up in a spectacular enhancement with 25 g/L of artemisinic yielded from their most engineered yeast (a 15-fold improvement to their last report). Most notably, adjusting CPR / CYP expression ratio was a previously overlooked parameter involved in the growth defect of the strains.

In summary, there exist similarities between the two aforementioned examples (hydrocortisone and artemisinin). In both cases, important investments were made towards the correct heterologous expression of the pathways: unlocking some biosynthetic steps that were missing in literature and both projects were long term efforts with intermediate steps in the chassis building <sup>22,129,222</sup>. Deep engineering of the yeast host (catching the side reactions in hydrocortisone, debottlenecking the flux in the artemisinic acid to reach higher yields) were decisive to adapt the biosynthesis of foreign molecules. This knowledge acquisition about the limits of initial systems opened up the perspective of commercial applications with the two processes. It should be noted that both belong to the same pharmaceutical company, Sanofi.

#### d) Recent improvements in the engineering strategies for sesquiterpene production in *S. cerevisiae*

Beyond the example of the artemisinic project, other chassis strains were designed to overproduce sesquiterpenes including farnesene and nerolidol <sup>23,223</sup>. As the mevalonate pathway deregulation in yeast to accumulate sesquiterpene is well mastered, an emerging interest is to act at the level of the yeast central metabolism to enhance the acetyl-CoA pool available for terpene biosynthesis <sup>16,23</sup>. The central metabolism of *S. cerevisiae*, the pivotal position of acetyl-CoA and the main strategies to enhance the acetyl-CoA pool are shown in Figure 11. At the pyruvate dehydrogenase level, these approaches included the set-up of a pyruvate dehydrogenase by-pass with ALD6 and ACS1, ACS2 overexpression <sup>224</sup> or alternatively employ a bacterial pyruvate dehydrogenase, which is supposed to be more efficient <sup>225</sup>.

In a groundbreaking study, Meadows et al. applied such extensive central metabolism reconstruction towards the farnesene production, attaining the astonishing titer of 130 g/L (and a mass yield of 0.298 g/g of glucose). To do so, multiple modifications were introduced in yeast (Figure 11), i. a pyruvate bypass that substituted ALD6 and ACS1/ACS2 step by an acetaldehyde dehydrogenase (ADA), this dramatically saves adenosine triphosphate, ii. the pentose phosphate pathway was redirected to acetyl-CoA formation too with the intervention of xylulose-5-phosphate phosphoketolase (xPK) and phosphate acetyltransferase (PTA), iii. the acetyl-phosphate consumption due to a promiscuous activity of glycerol-3-phosphate phosphatase (RHR2) was prevented through gene deletion, iv. a correct balance of the redox cofactors NADH/NADPH was achieved by complementing yeast endogenous HMGR enzymes (HMG1, HMG2, NADPH dependent) with the HMGR enzyme from *Silicibacter pomeroyi*, which is NADH dependent. Overall, this cutting-edge work underlines how deep the central metabolism can be tuned to reach higher yields in metabolic engineering.

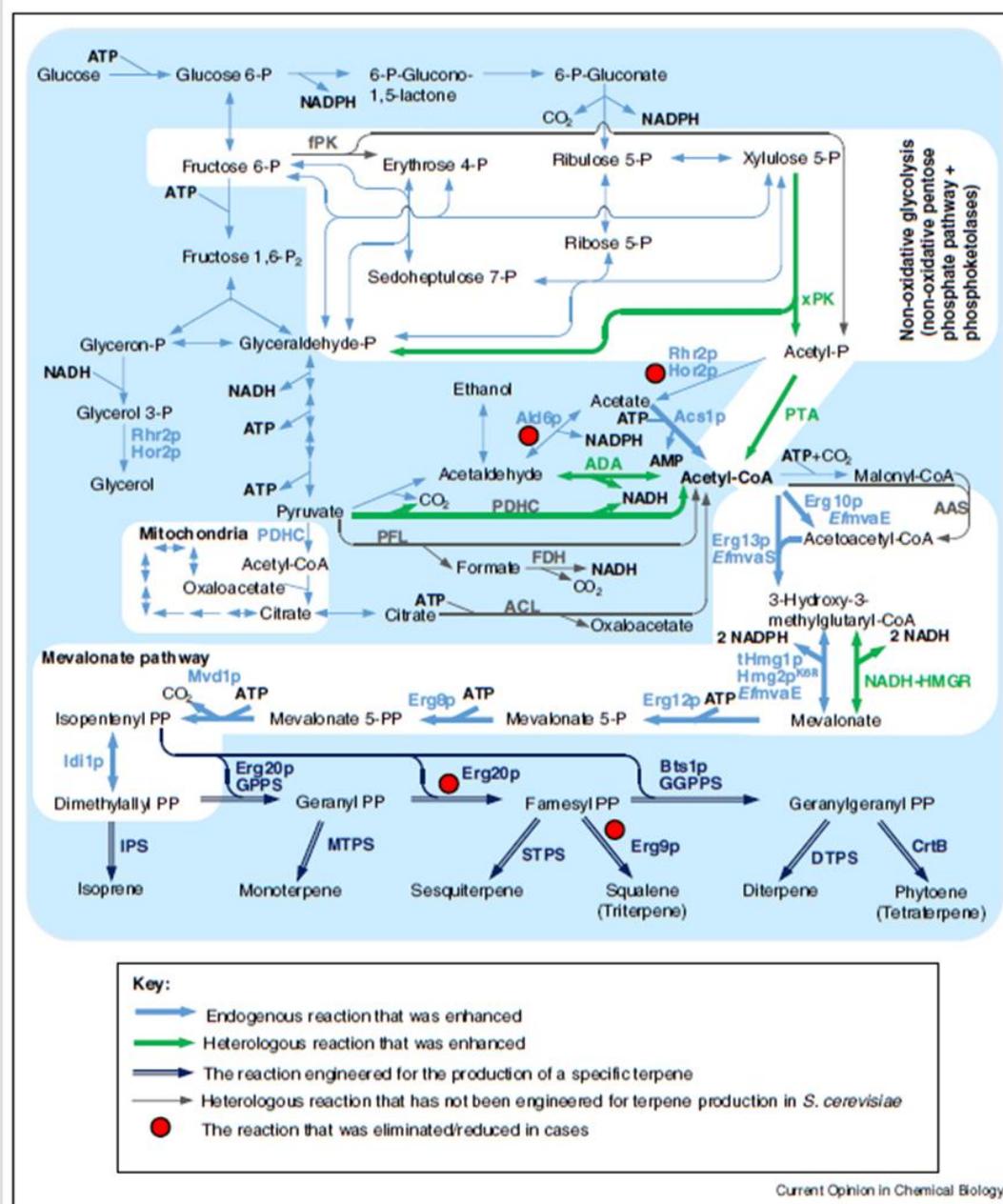


Figure 11: Metabolic network design for terpene production in yeast. This figure was originally presented by Vickers et al 16.

(1) Central metabolic pathway: fPK, fructose-6-phosphate phosphoketolase; xPK, xylulose-5-phosphate phosphoketolase; Rhr2p/Hor2p, glycerol-3-phosphate phosphatase; Ald6p, aldehyde dehydrogenase; Acs1p, acetyl-CoA synthase; PTA, phosphate acetyltransferase; ADA, acetaldehyde dehydrogenase (acylating); PDHC, pyruvate dehydrogenase complex; PFL, pyruvate-formate lyase; FDH, formate dehydrogenase; ACL, ATP citrate lyase. (2) Mevalonate pathway: Erg10p, acetoacetyl-CoA thiolase; AAS, acetoacetyl-CoA synthase; Erg13p, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase; EfmvaE, acetoacetyl-CoA thiolase/HMG-CoA reductase; EfmvaS, *Enterococcus faecalis* HMG-CoA synthase; tHmg1p, truncated HMG-CoA reductase 1; Hmg2pK6R, HMG-CoA reductase 2 K6R mutant; NADH-HMGR, NADH-specific HMG-CoA reductase; Erg12p, mevalonate kinase; Erg8p, phosphomevalonate kinase; MVD1p, mevalonate diphosphate decarboxylase; IDI1p, isopentenyl diphosphate:dimethylallyl diphosphate isomerase. (3) Specific terpene synthetic pathway: Erg20p, farnesyl diphosphate synthase; GPPS, geranyl diphosphate synthase; Bts1p/GGPPS, geranylgeranyl diphosphate synthase; IPS, isoprene synthase; MTPS, monoterpene synthase; STPS, sesquiterpene synthase; Erg9p, squalene synthase; DTPS, diterpene synthase; CrtB, phytoene synthase.

As mentioned earlier, accumulating intermediates of the mevalonate proved to be toxic to cells. Better regulation of the pathway proved that dynamic regulation are favorable<sup>150,173</sup>. An elegant solution can be the use of stress-response promoters as demonstrated in *E. coli* where amorphaadiene production was enhanced<sup>226</sup>. Promoters responding to FPP stress were first identified and further transformed into expression tools to induce amorphaadiene biosynthesis. Shift completely the usual system and use cell free pathways is one of the other developments that could overcome toxicity of some terpenes. It has the advantage of escaping complicated stress responses that are hard to identify. To debottleneck some steps, one can test individually the enzymes to assess which step(s) is/are affected or inhibited at some stages<sup>227</sup>.

Another developing biosynthetic route is to use the alcohols of DMAPP and IPP, namely, dimethylallyl alcohol and isopentenol as terpene precursors in a shortcut version of the mevalonate pathway. Several groups reported success in the building of such pathways<sup>228,229</sup>. Chatzivasileiou et al. proposed a two-step pathway from dimethylallyl alcohol and isopentenol to the IPP and DMAPP building blocks. Their team screened for promiscuous kinases able to build isopentenyl monophosphate and dimethylallyl monophosphate (*S. cerevisiae* choline kinase being the most efficient described in their study) and then looked for isopentenyl phosphate kinase of high activities (*Arabidopsis thaliana* enzyme). Finally, this minimal pathway was introduced in *E. coli* as exclusive IPP/DMAPP biosynthetic module by deleting the endogenous MEP pathway validating elegantly their design<sup>228</sup>. Rico et al. also developed an alternative to drive dimethylallyl alcohol and isopentenol utilization for DMAPP and IPP synthesis with a simplification of the pathway to its minimal. Indeed, they focused on the identification of kinases that could perform both phosphorylation steps, simplifying the system. They probed the natural diversity of isopentenyl phosphate kinases and their colorimetric screen based on carotene accumulation could serve as a base for either enzyme identification or pathway optimization<sup>229</sup>. One can now expect the use of these new pathways into other microbial chassis or in cell free approaches.

A last vast and emerging field for terpene production in yeast is the targeting of mevalonate pathway or products into dedicated cellular compartments. As mentioned earlier in I.B.1.c., reshaping and affecting endoplasmic reticulum morphology is one of the possibility<sup>185,186</sup>. Nonetheless, additional compartments were recently targeted such as mitochondria, peroxisomes, lipid droplets, remarkably showing the thriving interest in distributing pathway burden to several localizations. Lipid droplets were efficiently engineered as hubs for lycopene accumulation<sup>230</sup>. Ma et al. exploited i. the overexpression of the lipid droplets biosynthetic genes (PAH1, DGA1, mutated version of ACC1), ii. the changes in lipid composition (OLE1 overexpression), iii the increase in the lipid droplets size (deletion of FLD1). These three parameters positively affected lycopene production, and so did their combination. Furthermore, microscopy revealed that targeted molecules localized in the lipid droplets. Two teams also reported the retargeting of the mevalonate pathway in mitochondria for monoterpene production and obtained higher production titers than when the enzymes were expressed in

the cytosol<sup>231,232</sup>. Full reconstitution was achieved by targeting up to 9 enzymes into yeast mitochondria and exhibited the best monoterpenes accumulation to date for linalool and 8-hydroxygeraniol in *S. cerevisiae*. One advantage of mitochondria targeting is that the pool of acetyl-CoA is independent from the cytosolic one and may represent a valuable strategy for other classes of terpene as well (for sesquiterpene, diterpene...). Concerning peroxisome targeting, the proof of concept was recently established with the overproduction of squalene in this organelle<sup>233</sup>. Similarly to mitochondria engineering, the whole mevalonate pathway was introduced into peroxisome (but with 10 genes being retargeted as their design goes up to squalene). During their metabolic engineering Liu et al. also shed light on the fact that when squalene was overproduced in *S. cerevisiae*, peroxisomes were an overlooked storage depot. Indeed, yeast naturally accumulates squalene in this specific location, which was not reported before and their strategy pushed the limits of the native system. With the diversity of the various compartments that were engineered, one can now design heterologous pathways with many more options in hands, using even combination of the different targeting. Conversely, one could wonder whether the aforementioned examples would benefit to all terpenes or if there are cases where new targeting would be deleterious.

#### e) [Metabolic engineering and synthetic biology access to new to nature products](#)

Historically, chemistry tackled the production of new molecules with multiple functionalities such as plastics, glue, drugs and the building of incredibly large chemolibraries. Until recently, biologists and biochemists, mostly involved in the production of natural products, did not conceive that they had the tools to compete in such field. This vision changed with the breakthrough of enzyme directed evolution and synthetic biology. Using directed evolution of enzymes (Nobel prize awarded to Frances Arnold in 2018), it became clear that one could build catalysts for almost any reaction, even making a carbon silicon covalent bond for example<sup>234</sup>. Synthetic biology also stimulated the implementation of orthogonal pathways into chassis strain to produce new molecules from a biological design<sup>235</sup> (Figure 12). Following this approach, an orthogonal pathway comprising three engineered and reprogrammed enzymes was inserted in *E. coli* to produce 2,4-dihydroxybutyric acid, a precursor of a methionine analog, for which no natural pathway had ever been described<sup>178</sup>.

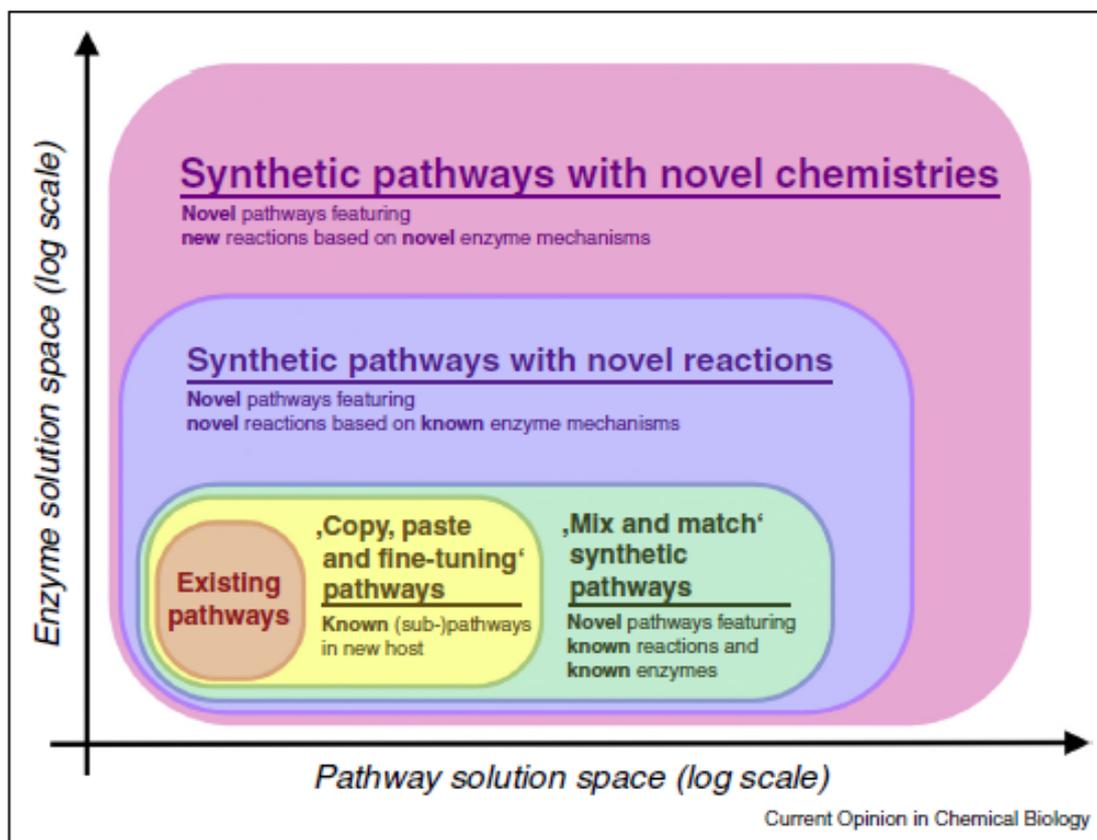


Figure 12: Increasing complexity in new to nature products (figure originally presented by Erb et al <sup>235</sup>)

The access to new to nature products from terpenes was recently reviewed by Arendt et al. <sup>5</sup> and mainly based on the combinatorial association of natural enzymes from diverse plant origins and/or engineered enzymes. Such approaches enabled the synthesis of new carotenoids from odd C<sub>35</sub> scaffolds instead of C<sub>40</sub> usually found in carotenes in a pioneering work <sup>236</sup> and, later, the diversification of triterpene and diterpene scaffolds <sup>17,18,237,238</sup> as well as new non-canonical 11-carbon terpenes <sup>239</sup>. For this latter case, a methyltransferase acting on GPP was expressed in yeast while TPS specificity was modified to make it accept preferentially methyl-GPP rather than GPP. Concerning sesquiterpenes, no reports describe the “on purpose” synthesis of new to nature products. When, new architectures and unexpected molecules were reported, they were essentially generated from “silent metabolism” and endogenous side reactions as observed for the glycosylated forms of artemisinic acid synthesized by *Nicotiana benthamiana* or the glutathione conjugate of costunolide obtained in *Nicotiana tabacum*.

\*\*\*

Most if not all model organisms have been engineered for terpene production due, in part, to the fact that MVA or MEP pathways are present in all living kingdoms. Due to the economic interest for recombinant production of the various terpenes (Figure 3<sup>16</sup>), the potentials of the model chassis were rapidly compared<sup>22,129</sup>. For the biosynthesis of terpenes devoid of membrane bound enzymes (like CYPs for example), *E. coli* performs efficiently with high titers reported for amorphadiene or (-)- $\alpha$ -bisabolol synthesis (27.4 g/L and 9 g/L, respectively)<sup>240,241</sup>. However, the expression of CYP from eukaryotes in prokaryotes is challenging and fails to sustain high activity levels<sup>242</sup>. Other prokaryotes such as *Bacillus subtilis* or *Corynebacterium glutamicum* were evaluated. Much lower titers were obtained. This is probably due to a lack of knowledge about the steps limiting the expression of the heterologous terpene pathways in these organisms<sup>243,244</sup>.

Other chassis such as cyanobacteria and eukaryotic algae may be interesting and valuable in the future due to their photosynthetic capabilities but are not mature yet. Pathway regulations and genes are much less known as these organisms have been far less experimented<sup>245</sup>. In addition, model plants like *Nicotiana benthamiana* could be more relevant in terms of physiology and cellular targeting for testing enzymes and pathways from plant origin. In Europe, public opinion is hostile to GMO cultivation in open fields, which is submitted to strong directives for authorization due to horizontal gene transfer issues and risk of biodiversity decline. Nonetheless, Reed et al. reported that gram scale of triterpenes can be recovered from *Nicotiana benthamiana* and assist in pathway prototyping and enzyme identification<sup>246</sup>.

In light of this data, *S. cerevisiae* stands as an attractive chassis organism for sesquiterpene production. The mevalonate pathway and its limitations have been historically unveiled and genetic tools facilitate the rapid design of *S. cerevisiae* to the desired pathway. Moreover, expressing plant CYPs may not represent a limit in *S. cerevisiae* as gram scale of oxidized terpenes were synthesized in the most compelling studies<sup>22</sup>.

Our work will focus on (+)-*epi*- $\alpha$ -bisabolol oxidation as it is involved in the hernandulcin pathway and will not be limited to the production of this molecule. Indeed, the production of new sesquiterpenes derivatives via metabolic engineering was never demonstrated. Therefore, implementing pathways possibly leading to new oxidized molecules from (+)-*epi*- $\alpha$ -bisabolol is of great interest owing to the multiple properties of sesquiterpenes. In that perspective, deviating from the canonical pathway and using promiscuous CYPs from xenobiotic metabolism may be promising. In a parallel approach, glucosylation of  $\alpha$ -bisabolol, (+)-*epi*- $\alpha$ -bisabolol derivatives, and hernandulcin will be attempted by using plant or human of glycosyltransferases to obtain other derivatives with enhanced solubility. To that end, both *E. coli* and *S. cerevisiae* were considered since plant glycosyltransferases are not membrane bound.



## Chapter II.

Synthetic derivatives of the (+)-*epi*- $\alpha$ -bisabolol are formed by mammalian cytochromes P450 expressed in a yeast reconstituted pathway.

Arthur Sarrade-Loucheur,<sup>†</sup> Dae-Kyun Ro,<sup>‡</sup> Régis Fauré,<sup>†</sup> Magali Remaud-Siméon,<sup>\*,†</sup> Gilles Truan<sup>\*,†</sup>

<sup>†</sup>TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

<sup>‡</sup>Department of Biological Sciences, University of Calgary, Calgary, Canada

## **Abstract**

Identification of the enzyme(s) involved in complex biosynthetic pathways can be challenging. An alternative approach might be to deliberately diverge from the original natural enzyme source and use promiscuous enzymes from other organisms. In this paper, we have tested the ability of a series of human and animal cytochromes P450 involved in xenobiotic detoxification to generate derivatives of (+)-*epi*- $\alpha$ -bisabolol and attempt to produce the direct precursor of hernandulcin, a sweetener from *Lippia dulcis* for which the last enzymatic steps are unknown.

Screening steps were implemented *in vivo* in *S. cerevisiae* optimized for the biosynthesis of oxidized derivatives of (+)-*epi*- $\alpha$ -bisabolol by coexpressing two key enzymes: the (+)-*epi*- $\alpha$ -bisabolol synthase and the NADPH cytochrome P450 reductase. Five out of 25 cytochromes P450 were capable of producing new hydroxylated regioisomers of (+)-*epi*- $\alpha$ -bisabolol. Of the new oxidized bisabolol products, the structure of one compound, 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, was fully elucidated by NMR while the probable structure of the second product was determined. In parallel, the production of (+)-*epi*- $\alpha$ -bisabolol derivatives was enhanced through addition of a supplementary genomic copy of (+)-*epi*- $\alpha$ -bisabolol synthase that augmented the final titer of hydroxylated product to 64 mg/L.

We thus demonstrate that promiscuous drug metabolism cytochromes P450 can be used to produce novel compounds from a terpene scaffold.

## A. Introduction

Terpenes represent one of the most diverse class of secondary metabolites comprising more than 50,000 molecules to date.<sup>32</sup> Present in all living kingdoms, their roles range from communication (hormone, abscisic acid), insect attraction, deterrent and defense molecules in plants, hormones and signaling compounds in insects,<sup>247,248</sup> and mainly as defense molecules in bacteria and fungi.<sup>249</sup> Terpenes also possess interesting pharmacological activities such as antimalarial (artemisinin), anti-cancer (taxol) and anti-inflammatory ((-)- $\alpha$ -bisabolol).<sup>16,250</sup> Many of them also display aromatic properties of major interest for industrial uses (ambroxide, nootkatone, etc.). The vast diversity of terpenes originates from the modularity of their chemical structures. They are synthesized by head to tail condensation of two building blocks composed of five carbons linked to a pyrophosphate group: isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), leading to geranyl-pyrophosphate (GPP, C<sub>10</sub>). Condensation of IPP onto GPP leads to farnesyl pyrophosphate (FPP, C<sub>15</sub>), and further to geranylgeranyl pyrophosphate (GGPP, C<sub>20</sub>) using another IPP molecule. Terpene synthases can use any of these precursors, by removing the pyrophosphate group while rearranging the carbon backbone and ultimately producing an extremely large diversity of terpene molecules. Most of them can be further functionalized via enzymatic reactions involving oxidative enzymes such as cytochromes P450 (CYP), alcohol dehydrogenases, epoxidases and other enzyme classes such as glycosyltransferases.<sup>53,251</sup>

Owing to their broad biological and physico-chemical properties, terpenes are widely used as commodities despite a low content often found in natural sources. Chemical synthesis of terpenes was thus largely explored but remains challenging due to the presence of chiral centers and/or stereospecific oxidation positions in natural terpenes.<sup>252</sup> To bypass these hurdles, synthetic biologists alternatively implemented metabolic engineering strategies, notably in industrial workhorse microorganisms such as *Saccharomyces cerevisiae* and *Escherichia coli*. Production of terpenes in *S. cerevisiae* and *E. coli* relies on the possibility to branch synthetic terpene pathways onto the MEV or MEP pathways, respectively. In *S. cerevisiae*, optimization of the metabolic flux towards terpene molecules was investigated using the overexpression of a truncated version of the HMG1 gene involved in sterol feedback regulation,<sup>253</sup> overexpression of key enzymes such as ERG20 or downregulation of native competing pathways.<sup>22,173</sup> Building on these strategies, sesquiterpenes including farnesene, nerolidol and artemisinic acid were produced at grams per liter scale in both prokaryotes and eukaryotes.<sup>22,223,254</sup>

Identification of natural enzymes of terpene biosynthetic pathways are essentially based on genomics and transcriptomics data mining. As terpene precursors are reasonably well produced in engineered microbial

cell factories, the discovery of enzymes involved in terpene pathway through direct *in vivo* screening has become a viable strategy.<sup>84,255</sup> Furthermore, engineered chassis microorganisms offer the possibility to enlarge the repertoire of natural products by producing new to nature compounds.<sup>5</sup> For example, terpenes absent in their native plant species were obtained from *in vivo* coexpression and/or combination of heterologous enzymes originating from multiple plants<sup>18,237</sup> or obtained via enzyme engineering,<sup>17,256</sup> thus highlighting the potential of these approaches to diversify terpene products.

Hernandulcin is a sesquiterpene derived from (+)-*epi*- $\alpha$ -bisabolol, which was originally isolated from *Lippia dulcis*, a plant from Central America (Figure 13). The pure molecule has attracted interest due to its sweetening properties; 1000 times sweeter than sucrose. However, its production level *in planta* is rather low<sup>3</sup> and its chemical synthesis involves multiple tedious steps due to the presence of two chiral centers whereas only one isomer is perceived as sweet.<sup>2</sup> The recent discovery of the (+)-*epi*- $\alpha$ -bisabolol synthase (BBS) converting FPP to (+)-*epi*- $\alpha$ -bisabolol unraveled a critical step in the synthesis of the terpene backbone of hernandulcin and its activity was validated in *S. cerevisiae*.<sup>6</sup> However, the final oxidation step(s) leading to hernandulcin in *L. dulcis* is (are) still to be discovered although enzymes such as CYPs are considered as serious candidates.

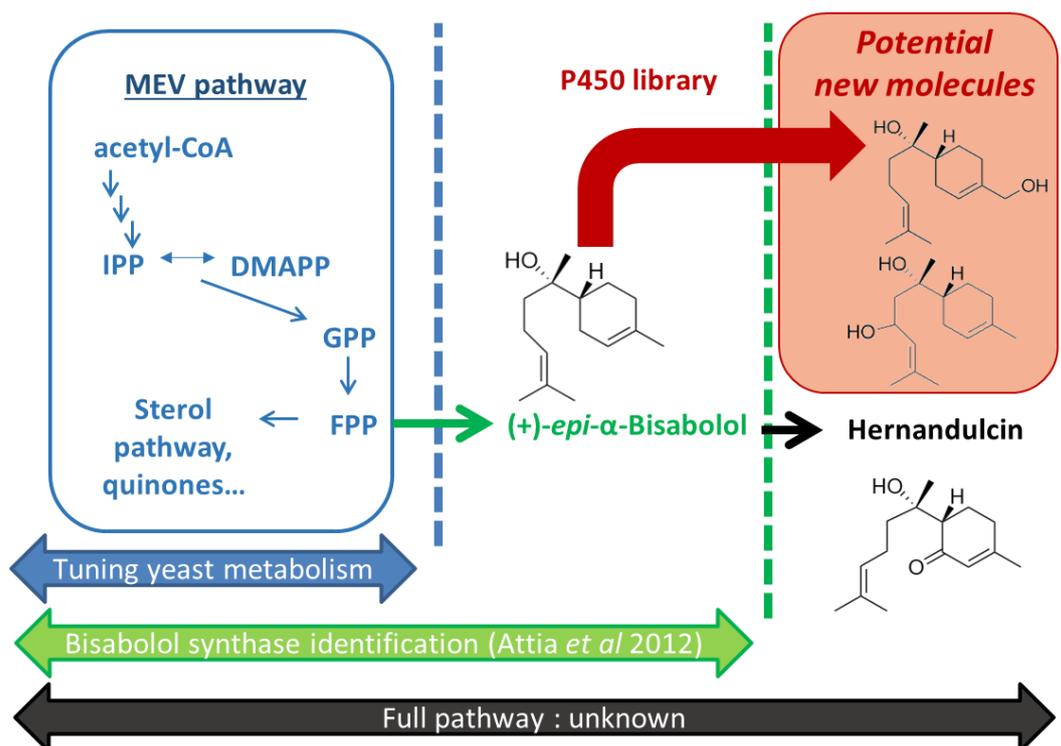


Figure 13: Strategy to produce derivatives of (+)-epi- $\alpha$ -bisabolol (including hernandulcin) using the yeast optimized MEV pathway, (+)-epi- $\alpha$ -bisabolol synthase and an *in vivo* screen to unravel oxidative enzymes.

IPP, isopentenyl pyrophosphate; DMAPP, dimethylallylpyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate.

Most plant CYPs involved in secondary metabolism have a rather narrow substrate specificity, rendering the identification of the natural oxidative enzyme(s) from *L. dulcis* highly challenging due to the multiplicity of CYP encoding genes. In contrast, animals possess a limited set of highly promiscuous CYP enzymes responsible for xenobiotic detoxification.<sup>15</sup> Surprisingly, there is no report describing the use of their remarkable oxidative properties to generate natural or new to nature molecules related to the plant secondary metabolism.

Knowing that expression of functional CYPs in engineered yeast is well mastered in our group<sup>257</sup> and that (+)-epi- $\alpha$ -bisabolol can be directly produced by *S. cerevisiae*, we assessed the ability of animal CYPs to convert (+)-epi- $\alpha$ -bisabolol into hernandulcin or related molecules (Figure 13). Furthermore, as  $\alpha$ -bisabolol has anti-inflammatory properties and can potentiate some antibiotic effects, we hypothesized that (+)-epi- $\alpha$ -bisabolol oxidation by various drug metabolism CYPs could generate other derivatives of interest, not only regarding their sweetness potency as sole biological feature but also for potential pharmaceutical properties.<sup>250,258</sup> Indeed, some human CYPs are competitively inhibited by  $\alpha$ -bisabolol suggesting that the latter can enter the active site and potentially act as an undescribed substrate<sup>259</sup> and we hypothesized the same behavior for (+)-epi- $\alpha$ -bisabolol. Our in house-generated library of CYPs was thus used together with a suitable cytochrome P450 reductase (CPR) in the genome of *S. cerevisiae* to assess their ability to produce new

derivatives. In parallel we pointed out an imbalance in the designed pathway and improved the production of the desired molecules by adding an extra copy of BBS.

## B. Materials and methods

### 1. Materials

Diclofenac, (-)- $\alpha$ -bisabolol, *trans,trans*-farnesol, *trans*-nerolidol, nootkatone were purchased from Sigma Aldrich (Munich, Germany), 4-hydroxy-diclofenac from Bertin Bioreagent (Montigny le Bretonneux, France). The racemic ( $\pm$ )-hernandulcin chemical standard (*R/R, S/S*) was purchased from Synphabase (Pratteln, Switzerland). DNA polymerases and restriction enzymes were supplied from New England Biolabs (Ipswich, MA, US) and Thermo Fischer Scientific (Waltham, MA, US). Primers were ordered at IDT (Coralville, IA, US) and sequencing was performed using Eurofins (Luxembourg, Luxembourg). For microbial cultures, peptone was from MP Biomedicals (Santa Ana, CA, US), tryptone, yeast extract, yeast nitrogen base and amino acid mixtures were from Formedium (Hunstanton, UK). Yeast microsomal fractions coexpressing CPR and the human CYPs in yeast were supplied by Bertin Pharma (Montigny le Bretonneux, France) while CYP2B11 microsomal fractions were prepared as previously described.<sup>260</sup>

### 2. Microbial strains

The yeast strains used in this study are based on S288C genetic background<sup>261</sup> and EPY300 is a chassis strain of *S. cerevisiae* engineered to trigger FPP accumulation and optimize the mevalonate pathway flux.<sup>24,25</sup> Additional genomic integrations were performed using the pMRI31-32-34 vectors listed in Table 1 to generate the strains listed in Table 2.

Table 1: Plasmids used in this study

Plasmid Name	Cloned enzymes	Origin of replication	Promoter	terminator	Yeast marker	Source
pESC-Leu2d + CPR + BBS	BBS + Aa CPR	2 $\mu$	GAL1/GAL10	CYC1	Leu2	Described in <sup>6</sup>
pMRI31	-	Integrative in HO locus	GAL1/GAL10	CYC1 / ADH1	KanMX	Described in <sup>149</sup>
pMRI32	-	Integrative in Ty4 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	Described in <sup>149</sup>
pMRI34	-	Integrative in DPP1 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	Described in <sup>150</sup>
pMRI31H+BBS	BBS + Hs CPR	Integrative in HO locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pMRI31S+BBS	BBS + Sc CPR	Integrative in HO locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pMRI32H+BBS	BBS + Hs CPR	Integrative in Ty4 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pMRI32S+BBS	BBS + Sc CPR	Integrative in Ty4 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pMRI34S-BBS	BBS + Sc CPR	Integrative in DPP1 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pMRI34S+BBS	BBS + Sc CPR	Integrative in DPP1 locus	GAL1/GAL10	CYC1 / ADH1	KanMX	This study
pSH47	Cre	ARSH4-CEN6	GAL1	CYC1	Ura3	Euroscarf (Scientific R&D GmbH, Oberursel, Germany)
pYEDP60	-	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	Ura3	Described in <sup>262</sup>
pYEDP60 + CYP	CYP library Table 4	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	Ura3	In house

Table 2: Strains used in this study

Strain Name	Genotype	Origin
Top10	F- mcrA $\Delta$ ( mrr-hsdRMS-mcrBC) $\Phi$ 80lacZ $\Delta$ M15 $\Delta$ lacX74 recA1 araD139 $\Delta$ ( ara-leu)7697 galU galK rpsL (StrR) endA1 nupG	Thermo Fischer Scientific (Waltham, MA, US)
BY4741	MATa his3 $\Delta$ 1 leu2 $\Delta$ 0 met15 $\Delta$ 0 ura3 $\Delta$ 0	Euroscarf
EPY300	MAT $\alpha$ his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 ura3 $\Delta$ 0 PGAL1-tHMG1:: $\delta$ 1 PGAL1-upc2-1:: $\delta$ 2 erg9::PMET3-ERG9::HIS3 PGAL1-ERG20:: $\delta$ 3 PGAL1-tHMG1:: $\delta$ 4	Described in <sup>25</sup>
YBIS_01	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1	This study
YBIS_02	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1-HsCPR-TCYC1	This study
YBIS_03	EPY300, $\Delta$ Ty4::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1	This study
YBIS_04	EPY300, $\Delta$ Ty4::TADH1-BBS-PGAL10-PGAL1-HsCPR-TCYC1	This study
YBIS_05	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1, $\Delta$ DPP1:: TADH1- -PGAL10-PGAL1-ScCPR-TCYC1	This study
YBIS_06	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1, $\Delta$ DPP1:: TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1	This study

### 3. [CYP library](#)

The CYPs encoding genes were inserted in the multi cloning site of pYEDP60 under the control of GAL10-CYC1 hybrid promoter.<sup>262</sup> The list of the cloned CYPs is described in Table 4.

### 4. [Plasmid construction](#)

All cloning and plasmid amplification steps were performed in TOP10 *E. coli* cells. The coding sequence of BBS was amplified using primer 1 and 2 with Phusion polymerase and pESC-Leu2d + CPR + BBS as template for 30 cycles (primers in Table 5). The resulting PCR product was digested by EcoRI / BglII and inserted in GAL10 promoter of pMRI31 and pMRI32. Then *S. cerevisiae* or *H. sapiens* CPR (UniprotKB numbers P16603 and P16435 respectively) coding sequences were amplified by PCR using primers 3-4 and 5-6 and inserted in the GAL1 promoter of pMRI31 + BBS or pMRI32 + BBS using BamHI /Sall restriction sites, leading to the plasmids pMRI31H + BBS, pMRI31S + BBS, pMRI32H + BBS, pMRI32S+BBS (Table 1). For pMRI34S + BBS, the cassette comprising BBS and CPR coding genes was digested from pMRI31S + BBS using PacI and Sall restriction enzymes and cloned in pMRI34 backbone. The plasmid pMRI34S - BBS was constructed from the plasmid

pMRI34S+BBS by PCR amplification and using primers 7 and 8, which allowed the removal of the BBS coding sequence.

## 5. [Plasmid transformation and integration of DNA cassettes in yeast](#)

The lithium acetate method was applied for yeast transformation with replicative plasmids.<sup>263</sup> When genomic integrations were carried out, the integrative vectors were linearized and transformed.<sup>264</sup> After obtaining clones resistant to genitacin (400 µg/mL), integration at the desired locus was checked by colony PCR. After transformation or integration, each yeast strain respiration ability was checked by streaking the strain on a non-fermentable growth media N3 (2% peptone (w/v), 1% yeast extract (w/v), and 3% glycerol (w/v)).

When successive rounds of integrations were needed, recycling of the KanMX cassette was performed by transforming the pSH47 plasmid expressing Cre recombinase.<sup>265</sup> The Cre recombination event was selected by screening genitacin sensible clones and the curation of pSH47 plasmid was selected by the recovery of Ura<sup>-</sup> clones. Before the second integration of the pMRI34 plasmids, the (+)-*epi*-α-bisabolol productivity was verified to ensure that potential deleterious genetic events did not occur during marker excision.

## 6. [Strain culture conditions and sampling](#)

For production of (+)-*epi*-α-bisabolol with the different strains, precultures were inoculated in SD medium (synthetic complete drop-out medium with 2% glucose (w/v), yeast nitrogen base + ammonium sulfate 6.9 g/L) and a mixture of all amino acids except those used for plasmid selection. In addition, EPY300 and derivative strains were grown without histidine to ensure genetic stability and without methionine to avoid repression of ERG9 by MET3 promoter. Precultures were grown overnight at 30°C under 200 rpm stirring. Using these seed precultures, cultures were then inoculated at 0.1 OD<sub>600nm</sub> either in SD medium or in YPA medium containing (peptone 20 g/L, yeast extract 10 g/L, adenine 80 mg/L). Both media were supplemented with 0.4% glucose (w/v) and 1.6% galactose (w/v) for induction, and 2 mM methionine (ERG9 repression). A dodecane overlay representing 10% (v/v) of the culture was used to enhance (+)-*epi*-α-bisabolol secretion when mentioned.

To produce hydroxylated (+)-*epi*-α-bisabolol metabolites, a preculture was grown overnight in SD medium -His-Met-Ura at 200 rpm and 30°C, and used to inoculate 0.1 OD<sub>600nm</sub> YPA medium supplemented with 0.4% glucose (w/v), 1.6% galactose (w/v) and 2 mM methionine. The cultures were grown at 160 rpm and 30°C.

Samples of 5 mL culture broth and cells were taken after 120 hours of incubation and kept at -20°C until extraction.

To check that chassis strains are suitable for CYPs screening, a diclofenac bioconversion experiment was conducted using 1 mM diclofenac. Diclofenac was added at the beginning of the culture done in the same conditions as those used for hydroxylated (+)-*epi*- $\alpha$ -bisabolol products.

Cell dry weights were determined from 5 mL of culture broth samples that were filtered, washed with 0.1 N HCl under a vacuum pump, and dried in a desiccator chamber (at 60°C under vacuum) before weighing.

## 7. Metabolite extraction and analytical detection of the sesquiterpenes

When a dodecane overlay was used during cultivation, cultures were centrifuged at 8,000 *g* for 10 min at ambient temperature to separate dodecane from culture broth. The absence of any effect upon the addition of the internal standard addition to the dodecane overlay while inducing the culture was verified (data not shown). The dodecane fraction was directly injected in GC equipped with FID and MS detection after dilution when required. GC-FID-MS module Trace1310 and ISQ LT (Thermo Scientific) were used to quantify (+)-*epi*- $\alpha$ -bisabolol, farnesol and nerolidol, and nootkatone as internal standard using N<sub>2</sub> as carrier gas. The inlet was fixed at 200°C and samples were injected on TG-5MS column (Thermo Scientific) and then eluted using a linear gradient starting at 40°C and ramping up to 250°C in 21 minutes with a final step at 250°C held 9 min.

In the absence of dodecane overlay, 5 mL samples of culture broth were taken and mixed with 1 g of glass beads ( $\emptyset$ =0.4  $\mu$ m-0.6  $\mu$ m, acid washed beads Sigma Aldrich), 2.5 mL of hexane. This mix was vortexed 1 minute in order to disrupt yeast cell wall and membrane and ensure the transfer of the hydrophobic molecules to the solvent phase. Then, the mixture was centrifuged at 10,000 *g* for 10 minutes at ambient temperature. The solvent fraction was collected and the extraction was repeated a second time. After solvent evaporation under airflow, the samples were suspended in 0.5 mL pure ethanol and injected in HPLC-UV-MS or in GC-FID-MS. HPLC was performed using a Waters HPLC 2790 separation module equipped with a C<sub>18</sub> precolumn and a ProntoSIL 120-5-C18-SH 5.0  $\mu$ m 250\*4.0 mm column (Bischoff chromatography). An elution gradient was applied at a constant flow rate of 0.75 mL/min using solvent A (water + 0.03% formic acid (v/v)) and B (acetonitrile + 0.03% formic acid (v/v)). Gradient started at 85% A (v/v) and 15% B (v/v) for 3 min, ramping up to 100% B over 24 min then held for 5.5 min at 100% B. (+)-*epi*- $\alpha$ -Bisabolol and (+)-*epi*- $\alpha$ -bisabolol derivatives were detected at 205 nm, using a Waters 996 photodiode array detector (Waters). For mass spectrometry

detection, a Micromass ZQ module (Waters) was used using ESI<sup>+</sup> ionization and a capillary voltage of 3.5 kV and a cone set at 30 V, masses were scanned in a range spanning 150 to 350 Da.

For diclofenac bioconversion experiment, 0.5 mL of the culture medium were acidified with 5  $\mu$ L of 1 M HCl and then mixed with 0.5 mL acetonitrile and glass beads. The mix was vortexed 1 minute, centrifuged at 16,000 g for 5 minutes and directly injected in HPLC. The gradient started at 85% A (v/v) and 15% B (v/v) for 3 min, ramping up to 100% B over 24 min then held for 5.5 min at 100% B. Diclofenac was quantified at 276 nm and 4-hydroxy-diclofenac at 267 nm.

## 8. [Enzymatic assays with CYP microsomal fractions](#)

Reactions were set up with microsomal fractions in 50  $\mu$ L volume containing: 1.2 mM NADPH, 400  $\mu$ M (-)- $\alpha$ -bisabolol or farnesol or nerolidol (10 mM solubilized in DMSO), 50 nM CYP microsomal fractions in 50 mM Tris-HCl pH 7.5, 1 mM EDTA. The reactions were stopped after 5 minutes or one hour incubation at 30°C by adding one volume of acetonitrile and vortexing. Finally, the assay was centrifuged 16,000 g 5 minutes at ambient temperature before HPLC-UV-MS analysis. The HPLC method was optimized for the analysis of the (-)- $\alpha$ -bisabolol oxidized products in order to improve their separation even if the products still were not completely separated using this gradient. Gradient started at 85% A (v/v) and 15% B (v/v) for 3 min, ramped up to 65% B in 4 min, reached 90% B over a 20 min gradient and then held for 5.5 min at 100% B.

## 9. [Isolation and characterization of the bisabolol oxidation products](#)

The strains YBIS\_06 + CYP2B6 or CYP2B11, CYP2C9, CYP2D6 showing accumulation of (+)-*epi*- $\alpha$ -bisabolol oxidation products were grown in 5 L Erlenmeyer flask using 1 L to 2 L culture volumes. Firstly, a seed culture of 5 mL in SD-His-Met-Ura was inoculated at 30°C, 200 rpm overnight and served to start a second pre-culture of 50 mL in same conditions. Then a culture of 1 to 2 L was inoculated at 0.1 OD<sub>600nm</sub> YPA medium supplemented with 0.2% glucose (w/v), 1.8% galactose (w/v) and 2 mM methionine. The culture was harvested after five days, centrifuged at 6,000 g for 25 min at ambient temperature. The whole supernatant was then applied on a preparative column filled with 20 g of octadecyl-functionalized silica gel (Sigma Aldrich). The column was washed with 400 mL of water containing 10% ethanol (v/v). (+)-*epi*- $\alpha$ -Bisabolol oxidized derivatives were then eluted with 100 mL pure ethanol, filtered on a Buchner funnel and concentrated under vacuum. The oily fraction was purified on a silica cartridge (HP silica 20  $\mu$ m) using automated Reveleris flash chromatography System (Grace) with a gradient of ethyl acetate in hexane. The recovered fractions were injected on HPLC-UV-MS and analyzed in the same conditions as those described above to verify the purity of each compound.

The hydroxylated metabolites of (+)-*epi*- $\alpha$ -bisabolol were then characterized by NMR. NMR spectra were recorded at 298 K on a Bruker Avance II spectrometer at 500 and 125 MHz for  $^1\text{H}$  and  $^{13}\text{C}$  respectively. Coupling constants ( $J$ ) are reported in Hz, and chemical shifts ( $\delta$ ) are given in ppm with residual solvents signal as internal reference. Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, m = multiplet, and br = broad. Analysis and assignments were performed using 1D ( $^1\text{H}$ ,  $^{13}\text{C}$ , Jmod and selective TOCSY) and 2D (COSY, TOCSY, HSQC and HMBC) experiments.

Trimethylsilylation of the (+)-*epi*- $\alpha$ -bisabolol hydroxylated products was carried out using bis(trimethylsilyl)acetamide/trimethylchlorosilane (5/1, v/v) (Sigma Aldrich) by adding one volume of the products extracted from the culture broth in hexane to one volume of silylation agent. The reaction was incubated 30 minutes at 50°C and analyzed by GC/MS as described above.

## C. Results

### 1. Building an efficient chassis strain for screening (+)-*epi*- $\alpha$ -bisabolol hydroxylation by CYPs

*In vivo* screening of CYPs capable of (+)-*epi*- $\alpha$ -bisabolol oxidation in yeast requires a convenient strain. Hence, we compared the wild-type strain BY4741 to EPY300,<sup>24</sup> a previously optimized strain for terpene production relying on the same S288C genetic background. BBS was expressed from the high copy number plasmid pESC leu2D BBS + CPR. After growth in a biphasic culture with a dodecane overlay, (+)-*epi*- $\alpha$ -bisabolol was only detected in the engineered EPY300 strain, indicating that a strain optimized for sesquiterpene production is indeed a prerequisite for *in vivo* screening purposes (Figure 19). In order to select animal CYPs enzymes active on (+)-*epi*- $\alpha$ -bisabolol, further modifications of the engineered strain EPY300 were required, notably the overexpression of yeast or human CPR, as previously described.<sup>257</sup> In order to ease subsequent screening of the plasmid expressing the CYP, the two CPR genes were separately integrated, under the control of a strong, inducible promoter (GAL10), in the genome of *S. cerevisiae*.

The genome integration of BBS and CPR genes under the control of GAL10 and GAL1 was performed in EPY300, at locus HO (strains YBIS\_01 (yCPR)/ YBIS\_02 (hCPR)) or Ty4 (YBIS\_03 (yCPR)/ YBIS\_04 (hCPR)) (

Table 2). In YBIS\_01, the (+)-*epi*- $\alpha$ -bisabolol production reached 0.176 mM in the dodecane layer and a similar production level was obtained in YBIS\_03 (data not shown). Hence, only the YBIS\_01 and YBIS\_02 strains (integration at HO) were further used as chassis strains. Additionally, the overexpression of both CPR genes did not cause any measurable toxicity as the growth of both strains was equivalent to that of EPY300.

Overexpression of both CPR was evidenced using bioconversion of diclofenac to 4-hydroxy-diclofenac by CYP2C9.<sup>266</sup> The basal level of expression of the endogenous yeast CPR of EPY300 strain sustained a very low activity of CYP2C9, which converted 0.33% of substrate in 120 hours. On the contrary, strains YBIS\_01 ( $\gamma$ CPR) and YBIS\_02 (hCPR), efficiently transformed diclofenac with conversion rates of 98% and 44 % respectively (Figure 20). As already reported with human CYPs the overexpression of *S. cerevisiae* CPR resulted in a higher specific activity of CYP2C9 compared than that obtained using the *H. sapiens* enzyme.<sup>257</sup>

Thus, strain YBIS\_01, producing significant amounts of (+)-*epi*- $\alpha$ -bisabolol while harboring an optimized redox environment compatible with millimolar conversion of human CYP substrate, was retained as a competent chassis strain for the *in vivo* screening of (+)-*epi*- $\alpha$ -bisabolol oxidation activities by drug metabolism CYPs.

## 2. *In vivo* screening of the CYP library in the engineered strain

We then implemented the screening of our CYP library, which contains 25 different CYPs, 13 from *H. sapiens*, 8 from other animals, and 4 additional mutants generated from human CYPs (Table 4). After individual transformation of the YBIS\_01 strain, each strain was analyzed for the potential production of hydroxylated forms of (+)-*epi*- $\alpha$ -bisabolol. Detection of hernandulcin using GC remains tedious due to its thermal lability.<sup>3,4</sup> As hydroxylated products of (+)-*epi*- $\alpha$ -bisabolol may also be prone to instability during GC analysis, we relied on a HPLC-UV-MS protocol for screening the CYPs library.

As there is no commercially available (+)-*epi*- $\alpha$ -bisabolol standard, (-)- $\alpha$ -bisabolol (molecular weight: 222.3 g/mol) was used as standard for quantitation, assuming that both isomers possess a similar molar extinction coefficient at 205 nm. (-)- $\alpha$ -bisabolol gave a strong MS signal at  $m/z=205.3$  corresponding to a dehydrated rearrangement ( $M+H^+-H_2O = 205.3$ ), which is often observed with hydroxylated products.<sup>267</sup>

HPLC-UV-MS analysis of the culture supernatant from individual clones of our library was then undertaken. We assumed that monohydroxylated or epoxy compounds generated from (+)-*epi*- $\alpha$ -bisabolol (molecular weight of 238.3 g/mol) should produce ions at  $m/z$  equivalent to  $M+H^+-H_2O = 221.3$  or  $M+H^+-2H_2O = 203.3$ . Indeed, we identified such ions for the strains expressing human CYP1A1, CYP2B6, CYP2B11, CYP2C9, CYP2C18 or CYP2D6 albeit with low levels for the latter. Conversely, these two ions were never detected in the strain transformed by the void vector pointing that such metabolites are produced by CYP activities and therefore their molecular weight corresponds to the mass of (+)-*epi*- $\alpha$ -bisabolol plus one oxygen atom. Overall, the generated molecules showed UV spectral properties similar to the (+)-*epi*- $\alpha$ -bisabolol precursor with an absorbance maximum at 205 nm. The chromatograms of the various strains are shown in Figure 14A for UV and in Figure 14B for  $m/z$  ( $M+H^+-2H_2O$ ) = 203.3. This ion was the dominantly detected one for the new metabolites using our MS settings while the signal at  $m/z$  ( $M+H^+-H_2O = 221.3$ ) was weaker. Regarding the low production of new metabolites with CYP2D6, we attempted to improve the production level of oxidized (+)-*epi*- $\alpha$ -bisabolol product by placing the open reading frame of the CYP2D6 gene in different codon initiation environments. The best level of oxidized compounds was obtained using an optimized Kozak environment (CYP2D6 H1, Table 4 and Figure 21), which was then systematically used throughout the study. In addition to monohydroxylated products, we also searched for the potential formation of a keto group ( $MW_{\text{bisabolol}} + 14$ ) as occurring in hernandulcin. The racemic form of ( $\pm$ )-hernandulcin (*R/R*, *S/S*) yielded a peak at  $m/z = 219.3$  corresponding to a dehydrated rearrangement. A thorough analysis of the extracts of the various strains did not show any peak with such a  $m/z$  value, indicating that neither hernandulcin nor any keto derivatives of (+)-*epi*- $\alpha$ -bisabolol were produced or stable enough to be detected. We also did not detect either any dihydroxylated products with  $m/z = 254.3$  ( $MW_{\text{bisabolol}} + 32$ ) or giving potential rearrangement products ( $M+H^+-H_2O = 237.3$  or  $M+H^+-2 H_2O = 219.3$  or  $M+H^+-3 H_2O = 201.3$ ).

In the different strains, (+)-*epi*- $\alpha$ -bisabolol was detected within a region containing several non-well resolved peaks. It is therefore possible that endogenous molecules of close retention times to that of (+)-*epi*- $\alpha$ -bisabolol are present in the cell extracts. If such molecules possess the same MW and are hydroxylated by our CYP library, they would be counted as false positive hits and would probably elute in the same region as (+)-*epi*- $\alpha$ -bisabolol products. Indeed, in cases where sesquiterpene synthases could not totally convert FPP into sesquiterpenes, dephosphorylation of FPP led to farnesol, a (+)-*epi*- $\alpha$ -bisabolol isomer.<sup>268</sup> In our conditions, the resolution of our HPLC-UV-MS method was not sufficient to allow a separation between (-)- $\alpha$ -bisabolol chemical standard, farnesol and nerolidol, thus leading to a probable coelution of their hydroxylated products if they would be produced in our strains (Figure 14C). A further analysis of the sesquiterpenes from YBIS\_01 by GC/MS indeed revealed the presence of farnesol and nerolidol (Figure 22).

In the absence of chemical standards for either hydroxylated (+)-*epi*- $\alpha$ -bisabolol, farnesol or nerolidol derivatives, solid conclusions on the specificity of the tested CYP towards the sesquiterpenes solely based on

our primary *in vivo* screening experiments could not be drawn (Figure 14D). This prompted us to verify that the identified CYP enzymes are consistently active with  $\alpha$ -bisabolol, farnesol and nerolidol as substrates via an enzymatic *in vitro* reaction.

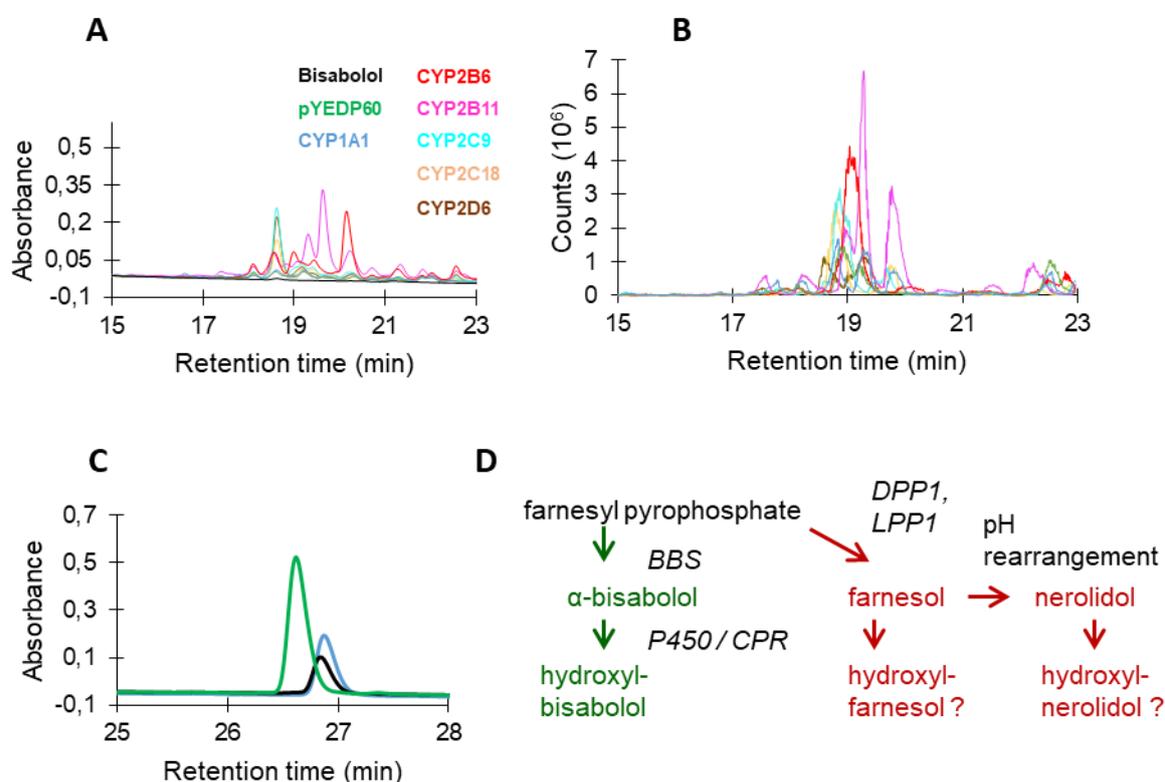


Figure 14: Detection of new molecules generated in YBIS\_01 strain transformed by the empty plasmid pYEDP60, or by some CYPs.

(A) UV chromatogram (205 nm). (B) Mass spectrometry signal showing signal at  $m/z = 203.3$  corresponding to hydroxylated (+)-*epi*- $\alpha$ -bisabolol molecules in accordance with Figure 13A chromatograms. (C) HPLC-UV chromatograms of (-)- $\alpha$ -bisabolol blue trace, farnesol green trace and nerolidol black trace. (D) Potential off target oxidation routes derived from (+)-*epi*- $\alpha$ -bisabolol isomers that may exist *in vivo* in the chassis strain.

### 3. *In vitro* activities of the human CYPs with (+)-*epi*- $\alpha$ -bisabolol, farnesol and nerolidol

To verify whether the selected CYPs had genuine activities on (+)-*epi*- $\alpha$ -bisabolol rather than farnesol or nerolidol, *in vitro* enzyme assays were conducted. Due to the absence of commercial source of (+)-*epi*- $\alpha$ -bisabolol, the diastereoisomer (-)- $\alpha$ -bisabolol was used as a surrogate substrate, assuming that both molecules would have comparable binding properties. Enzymatic tests were performed at the same concentration of substrate (0.4 mM). Enzymatic assays of the identified CYPs confirmed that all selected enzymes were active

with either some or all sesquiterpenes (Table 3). For each of the tested molecules, the number of detected products and the corresponding conversion yields are listed in Table 3, while the original chromatograms of the enzymatic assays are presented in the Figure 23, Figure 24, Figure 25, Figure 26.

Table 3: *In vitro* activities of the CYPs towards (-)- $\alpha$ -bisabolol, farnesol and nerolidol

Substrate \ CYP	CYP1A1		CYP2B6		CYP2B11		CYP2C9		CYP2C18		CYP2D6	
	P	C	P	C	P	C	P	C	P	C	P	C
Bisabolol	+/- <sup>a</sup>	0	2	0.64	+++ <sup>b</sup>	0.72	3	0.65	2	0.24	+/- <sup>a</sup>	0
Farnesol	1	0.05	2	0.15	+++ <sup>b</sup>	0.57	1	0.01	1	0.05	+/- <sup>a</sup>	0
Nerolidol	-	-	-	-	1	0.55	2	0.10	1	0.10	+/- <sup>a</sup>	0

P: Number of detected products. C: Conversion after 1 hour of reaction. Conversion was calculated as the ratio between the area of product(s) and the sum of the areas of substrate and product(s) according to UV quantitation. <sup>a</sup> Trace activities corresponding to a product formation only detectable in MS (see Figure 23, Figure 24, Figure 25). <sup>b</sup> Multiple products are evidenced by HPLC but counting single products is not feasible (see Figure 22, Figure 23, Figure 24, Figure 25). In the absence of available standards, we assumed that substrates and products have similar molar extinction coefficient at 205 nm.

CYP1A1 converts farnesol more efficiently than (-)- $\alpha$ -bisabolol as evidenced by the very low levels of hydroxylated compounds with (-)- $\alpha$ -bisabolol, indicating that *in vivo*, farnesol is preferentially oxidized. CYP2D6 shows only weak activities towards the three substrates in the tested conditions even if the *in vivo* accumulation sesquiterpene oxidized products was clearly shown in our chassis strain. CYP2B6, CYP2C9 and CYP2C18 oxidized (-)- $\alpha$ -bisabolol 2.4 to 6-fold more efficiently than farnesol or nerolidol. Therefore, one can expect that, in the chassis strain, the formation of (+)-*epi*- $\alpha$ -bisabolol oxidized products is favored. With CYP2B11, conversion of the three sesquiterpenes is in the same range, and thus, *in vivo*, oxidized derivatives may originate from (+)-*epi*- $\alpha$ -bisabolol, farnesol and nerolidol. Interestingly, the different CYPs exhibited distinctive patterns of oxidized products from (-)- $\alpha$ -bisabolol, CYP2C18 and CYP2B6 led to two compounds while CYP2B11 and CYP2C9 catalyze the formation of a more diverse range of products.

#### 4. Structural characterization of some (+)-*epi*- $\alpha$ -bisabolol derivatives

##### a) Improving (+)-*epi*- $\alpha$ -bisabolol production in chassis strain

Sesquiterpene production can be modulated by several factors, including biphasic culture with dodecane,<sup>25</sup> pH of the medium,<sup>269</sup> the presence of endogenous phosphatases<sup>147,268</sup> and finally the copy number of the sesquiterpene synthase encoding genes. We thus assessed potential improvements of the (+)-*epi*- $\alpha$ -bisabolol production by modifying the balance of the FPP flux towards (+)-*epi*- $\alpha$ -bisabolol, farnesol and nerolidol.

(+)-*epi*- $\alpha$ -bisabolol, farnesol and nerolidol were quantified by GC and normalized according to the cell dry weight of the cultures (Figure 15A). Addition of a dodecane overlay enhanced the accumulation of sesquiterpenes as previously reported.<sup>25</sup> Additionally, we also demonstrated that it does not alter the distribution of the different products (Figure 15A). Cultures conducted with buffered (pH 7.4) or non-buffered media reveal that the pH exerts a significant effect on the pattern of sesquiterpenes accumulated in the dodecane overlay: there is a concomitant reduction of nerolidol synthesis and a greater accumulation of farnesol at pH 7.4. With a non-buffered minimal medium (SD), the final pH dropped to *ca* 3 after 120 hours of culture. In such an acidic condition, Cope rearrangement can occur and convert farnesol to nerolidol.<sup>269</sup> In contrast, with a non-buffered rich medium (YPA) the pH dropped to *ca* 5 at the end of the culture and the accumulation of nerolidol was limited, confirming the pH-dependent conversion of farnesol to nerolidol (Figure 15A). Surprisingly, the overall production of sesquiterpenes, including (+)-*epi*- $\alpha$ -bisabolol, per cell is lower in a complete medium (YPA) compared to minimal medium (SD), which might be attributed to a possible difference in the way methionine represses the expression of *ERG9* by the *MET3* promoter, leading to a change in the flux to FPP. However, when comparing the total amount of (+)-*epi*- $\alpha$ -bisabolol produced per liter of culture (Figure 15B), there is a 34% improvement of the (+)-*epi*- $\alpha$ -bisabolol quantity produced in YPA medium compared to the SD medium due to a higher cell density obtained in rich medium.

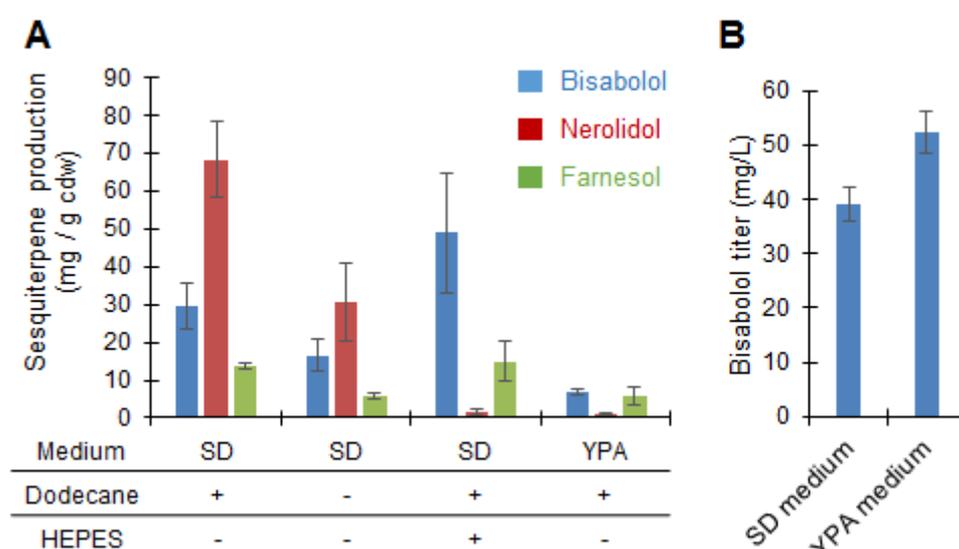


Figure 15: The production of sesquiterpenes by YBIS\_01 strain cultured in different conditions. All cultures were performed in triplicate.

(A) YBIS\_01 sesquiterpene production with or without dodecane overlay, in non-buffered conditions or buffered conditions, and SD or YPA media. The sesquiterpene production was normalized according to the cell dry weight. (B) (+)-*epi*- $\alpha$ -bisabolol titer in SD and YPA media.

Competing reactions by endogenous phosphatases such as DPP1 and LPP1 may decrease the production of (+)-*epi*- $\alpha$ -bisabolol by transforming part of the FPP to farnesol (Figure 14D).<sup>270</sup> To resolve this issue, we deleted DPP1 with or without adding a second copy of BBS, generating YBIS\_05 (one BBS copy) and YBIS\_06, (2 copies of BBS) from YBIS\_01. The comparison between YBIS\_01 and YBIS\_05 indicates that the deletion of DPP1 gene has negligible effect on (+)-*epi*- $\alpha$ -bisabolol production (Figure 16). On the contrary, adding a second copy of the BBS gene significantly enhanced (+)-*epi*- $\alpha$ -bisabolol production (~ 70% increase) showing that further potential improvements, either at the level of quantity or specific activity of BBS, are still feasible.

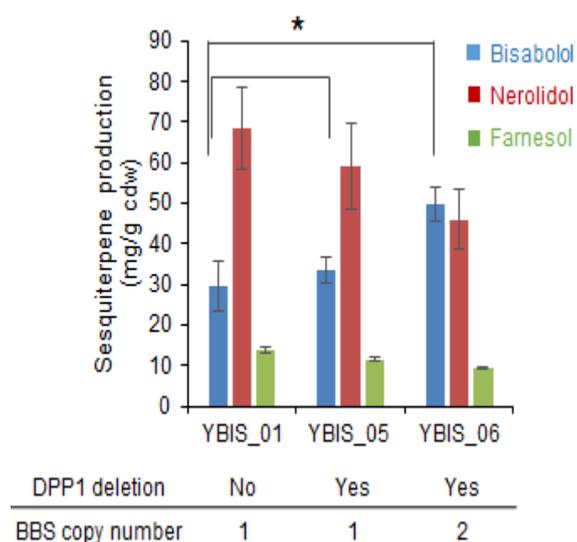


Figure 16: Comparison of the sesquiterpene production by YBIS\_01, YBIS\_05 (DPP1 gene deletion) and YBIS\_06 strains (DPP1 gene deletion and one supplementary copy of the BBS encoding gene).

All cultures were performed in triplicate in SD medium with a dodecane overlay spiked with nootkatone as internal standard. Statistical comparison of (+)-*epi*- $\alpha$ -bisabolol production using a t-test between YBIS\_01 and YBIS\_05 p-value = 0.17, YBIS\_01 and YBIS\_06 p-value = 0.0056.

## b) (+)-*epi*- $\alpha$ -Bisabolol *in vivo* concentration influences the oxidized metabolite production

While the addition of dodecane increased the production of (+)-*epi*- $\alpha$ -bisabolol and isomers, we did not use this pulling effect for the production of the oxidized derivatives as it may decrease the presence of intracellular (+)-*epi*- $\alpha$ -bisabolol available for the CYP. The comparison between YBIS\_05 and YBIS\_06 for the production of hydroxylated derivatives by CYP2C9 and CYP2D6 is presented in Figure 17. Interestingly, in the absence of the pulling effect of dodecane, YBIS\_06 produced only 32% more (+)-*epi*- $\alpha$ -bisabolol than YBIS\_05 (compared to the 70% increase with dodecane, Figure 16), suggesting a partial inhibition of BBS by the product of the reaction, namely (+)-*epi*- $\alpha$ -bisabolol.

While *in vitro* assays with CYP2C9 gave 3 different products, only two of them were detected in the cultured cells. In the strain YBIS\_05 + CYP2C9, the remaining level of (+)-*epi*- $\alpha$ -bisabolol indicates an 87% conversion of this precursor to its oxidized derivatives. This conversion yield dropped slightly to 77% in YBIS\_06 + CYP2C9. Interestingly, an extra copy of the BBS encoding gene (YBIS\_06 + CYP2C9) raised the quantity of oxidized products compared to YBIS\_05 + CYP2C9 with a concomitant accumulation of (+)-*epi*- $\alpha$ -bisabolol. Indeed, 36 mg/L (+)-*epi*- $\alpha$ -bisabolol oxidized product 1 and 5.7 mg/L of product 2 are produced in YBIS\_05 + CYP2C9 while 64 mg/L of (+)-*epi*- $\alpha$ -bisabolol oxidized product 1 and 13 mg/L of product 2 were synthesized YBIS\_05 + CYP2C9. This demonstrates that the reaction promoted by CYP2C9 in YBIS\_05 + CYP2C9 is not limiting. The conversion yield in YBIS\_05 + CYP2D6 and YBIS\_06 + CYP2D6 were respectively, 59 and 51%, while the final titer of the (+)-*epi*- $\alpha$ -bisabolol oxidized product 1 reached 37 and 50 mg/L in YBIS\_05 + CYP2D6 and YBIS\_06 + CYP2D6, respectively. Hence, with CYP2D6, the difference of production of oxidized products suggests that the enzymatic step catalyzed by CYP2D6 was not completely limiting in YBIS\_05 + CYP2D6.

The comparison between the strains YBIS\_05 and YBIS\_06 with both CYP2C9 and CYP2D6 also points that although none of the reaction catalyzed by the two CYPS are limiting, the *in vivo* efficiency of conversion of CYP2C9 is greater than that of CYP2D6. Alternatively, the diminution of the farnesol pool due to the second copy of BBS (Figure 16) could also be beneficial due to the possible change in the competition between the two substrates for the CYP. Thus, having a second copy of BBS is still beneficial for the accumulation of oxidized product by CYP2D6.

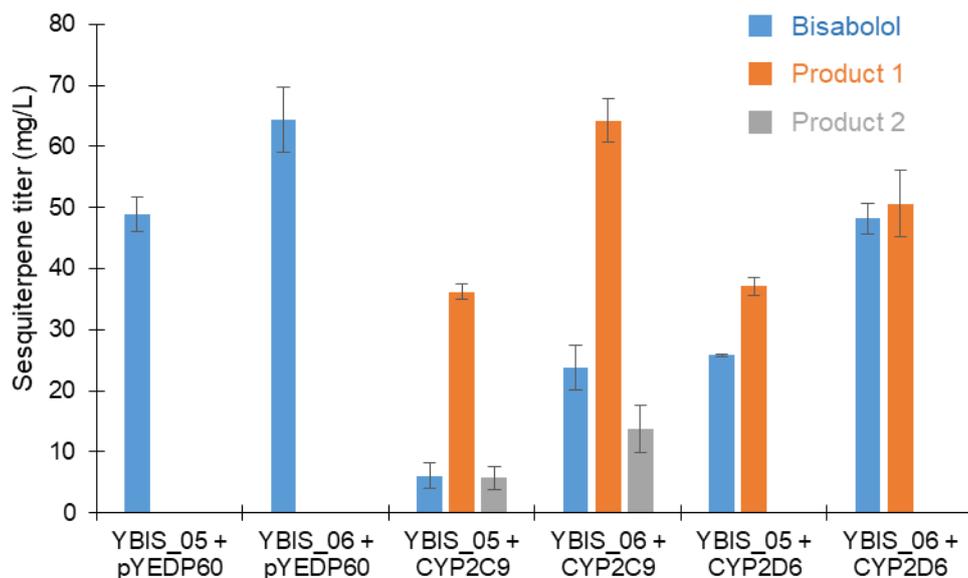


Figure 17: Effect of chassis improvement on accumulation of CYP2C9 and CYP2D6 products. The level of (+)-*epi*- $\alpha$ -bisabolol was varied in the strains by comparing YBIS\_05 (1 integrated BBS copy) to YBIS\_06 (2 integrated BBS copies).

### c) Product purification and structure determination

For structural characterization of the oxidized molecules produced by CYP2B6, CYP2B11, CYP2C9, CYP2C18 and CYP2D6 and to circumvent the limit of resolution of the HPLC method, we turned to GC analyses with and without performing a trimethylsilylation derivatization reaction. (-)- $\alpha$ -bisabolol was used as control. According to HPLC analysis, CYP2C18 and CYP2D6 yielded only one molecule and the same pattern was found in GC-MS. Furthermore, their mass spectra (native product as well as the trimethylsilylated derivative) are identical (Figure 27). We can then assume that the products synthesized by CYP2C18 and CYP2D6 are identical. With CYP2C9, two products are also detected by HPLC and GC. The major compound produced by CYP2C9 is identical to the one synthesized by CYP2C18 and CYP2D6. For CYP2B6 and CYP2B11, the analysis of the oxidized products in GC and HPLC was difficult due to multiplicity of the products formed. However, the product profiles were distinguishable from those obtained with the other CYPs, indicating the presence of different products. The fractions isolated from YBIS\_06 + CYP2B6 and YBIS\_06 + CYP2B11 culture supernatant contained a complex mixture of oxidized sesquiterpene products potentially originating from both (+)-*epi*- $\alpha$ -bisabolol and/or farnesol and additional endogenous contaminants, which could not be separated prior NMR characterization. We thus pursued the purification of the products synthesized by CYP2C9 and CYP2D6 using the strain YBIS\_06 transformed by the corresponding high copy plasmids. CYP2C18, which probably produces the same oxidized (+)-*epi*- $\alpha$ -bisabolol derivative as CYP2D6 (and in lower yield), was not further analyzed. We

noted that, even in the absence of a dodecane layer, more than 80% of the oxidized products of (+)-*epi*- $\alpha$ -bisabolol accumulated in the culture broth while most of the sesquiterpenes remained trapped within the cell (Figure 28). The different cell/medium partition behavior between (+)-*epi*- $\alpha$ -bisabolol and its oxidized products may be in part due to the presence of an endogenous yeast transporter that could excrete the oxidized products but not (+)-*epi*- $\alpha$ -bisabolol.

After isolation of the oxidized products recovered from the culture supernatants of YBIS\_06 + CYP2D6 and YBIS\_06 + CYP2C9, sufficient amounts of oxidized (+)-*epi*- $\alpha$ -bisabolol was purified and recovered for NMR analysis. One single oxidized product was found in the purified fraction obtained from YBIS\_06 + CYP2D6 culture, and two in the fraction generated with YBIS\_06 + CYP2C9. NMR analysis confirmed that both CYP2D6 and CYP2C9 produced new to nature molecules derived from (+)-*epi*- $\alpha$ -bisabolol (Figure 18, Figure 29 and Table 6). The <sup>1</sup>H spectrum of 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol produced by CYP2D6 displays only three signals attributed to methyl groups (H-13 at 1.69, H-12 at 1.63 and H-15 at 1.15 ppm). It is noteworthy that the disappearance of the H-14 methyl singlet of (+)-*epi*- $\alpha$ -bisabolol at 1.64 ppm is combined with the appearance of the H-14 methylene signal at 4.01 ppm of 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol. Similarly, the <sup>13</sup>C spectrum of 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol reveals only three methyl signals (28.7, 24.0 and 17.7 ppm corresponding to C-13, C-15 and C-12, respectively) combined with the additional C-14 methylene signal at 67.2 ppm. In addition to 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, NMR spectra of the mixture produced by CYP2C9 also highlighted some traces of a second metabolite. It is characterized by the presence of two new signals at 5.41 (triplet, H-3) and 5.38 ppm (broad singlet, H-10; unblinded compared to the H-10 signals of (+)-*epi*- $\alpha$ -bisabolol and 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol) correlating with two methyne signal (CH=) at 126.1 and 120.6 ppm, respectively, and combined with at least two putative methyl singlets (1.63/23.3 and 1.11/24.1 ppm, H/C-14 and 15 pairs respectively) (Table 6). This led us to presume that the minor product is 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol but cannot exclude completely that it corresponds to 8-hydroxy-(+)-*epi*- $\alpha$ -bisabolol.

It is worth noting that at least three CYPs (CYP2C9, CYP2C18 and CYP2D6) produce 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol. Considering their known difference in substrate specificity,<sup>271</sup> it may be tempting to hypothesize that, while various CYPs have different binding sites and may promote various positioning of the substrate within the active site, the hydroxylation at C-14 is thermodynamically favored over the other positions. The minor product formed by CYP2C9, most probably 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, shows an attack of the substrate at a completely opposite position compared to the 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, which probably signs a different mode of binding or, alternatively, the use of a second substrate channel for entering CYP2C9 that is known to alter oxidation specificity.<sup>272</sup> Overall, for the products of CYP2D6, CYP2C18 and CYP2C9 the hydroxyl group was not introduced on the C-2carbon, the one that possess the carbonyl function in hernandulcin. However, due to the fact that other, minor, products are formed with CYP2B6 or CYP2B11, we cannot exclude that this hydroxy derivative is indeed present in the complex mixture. Further studies, lowering

the farnesol content and increasing the amount of (+)-*epi*- $\alpha$ -bisabolol formed should help in the identification of these new metabolites.

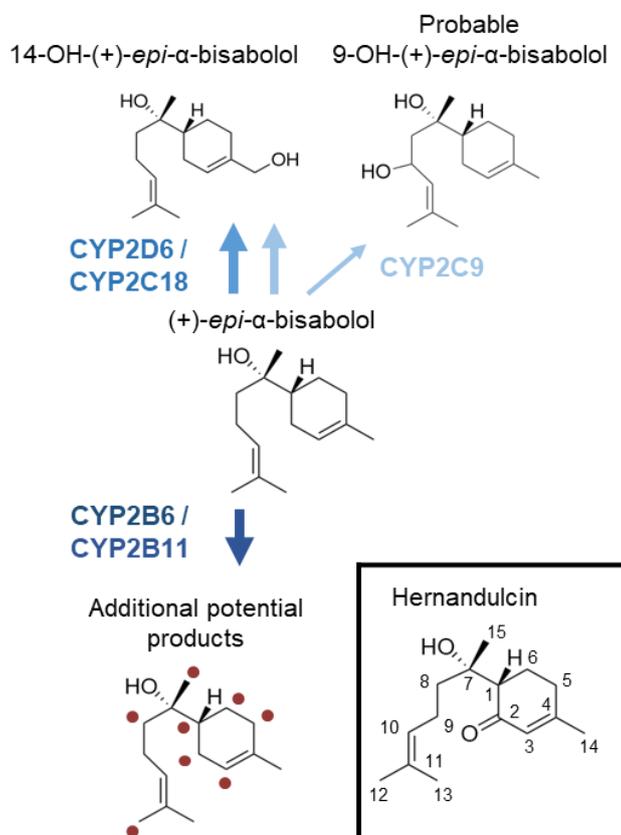


Figure 18: Chemical structures of the newly generated molecules. CYP2D6 produces 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol while CYP2C9 produces 80% of 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol and 20% of a second additional hydroxylated (+)-*epi*- $\alpha$ -bisabolol derivative, most probably 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol. The red dots on the (+)-*epi*- $\alpha$ -bisabolol molecule indicates the other positions that could be hydroxylated in CYP2B6 and CYP2B11 although the original molecules could not be fully characterized by NMR.

## D. Discussion

In this paper, we developed an *in vivo* screening strategy to discover oxidized derivatives of (+)-*epi*- $\alpha$ -bisabolol such as hernandulcin (Figure 18), a powerful sweetener naturally produced by *Lippia dulcis*. Although human CYPs have already been efficiently used for metabolic engineering purposes such as the complete hydrocortisone biosynthesis,<sup>19</sup> there is no demonstration, to our knowledge, of the use of human CYPs in diverting plant terpene metabolism. Therefore, rather than searching for the right, specific, CYP candidate in *Lippia dulcis*, we deliberately screened an animal drug metabolism CYP enzyme library. Our rationale was based on the facts that i) the oxidizing enzymes from *L. dulcis* were not yet identified and no CYPs were

characterized with (+)-*epi*- $\alpha$ -bisabolol as substrate, ii) drug metabolism CYP isoforms are constantly exposed to terpenes from natural sources via food supply and could well have evolved to oxidize these compounds prior to their excretion. Furthermore, comparative studies of plant and animal CYP metabolism, which focused on inhibition/competition experiments identified inhibitor molecules from plants but rarely considered them as potential substrates.

Out of the 25 screened CYPs, we identified 5 CYPs that produced hydroxylated derivatives of (+)-*epi*- $\alpha$ -bisabolol from the CYP2 family (CYP2B6, CYP2B11, CYP2C9, CYP2C18 and CYP2D6) and 1 from the CYP1 family (CYP1A1), which was actually more active on farnesol. The fact that mostly CYP2 family isoforms were identified corroborates existing data on their role in oxidizing monoterpenes<sup>273</sup> and sesquiterpenes.<sup>274</sup> We can thus hypothesize that structural elements conserved in the CYP2 family favor the recognition of this class of C<sub>15</sub> molecules. Wider structure/activity relationship studies may help to decipher which structural motifs are responsible for this recognition in the CYP2 family, especially considering the quite strong regioselectivity of CYP2D6, CYP2C18 and CYP2C9 for the position C-14.

Some available bioinformatics tools seek the prediction of human CYPs specificities due to their main contribution in drug metabolism.<sup>275</sup> We then assessed two different tools, namely FAME3 and Xenosite to evaluate their accuracy in predicting if (+)-*epi*- $\alpha$ -bisabolol could be substrate of the CYP tested.<sup>276,277</sup> Their results are presented in Figure 30. Fame3 output is the overall metabolism from human CYPs, and the 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol that is the major product of CYP2D6, CYP2C18 and CYP2C9 has a score that predicts it at the third probable metabolite. The 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol produced by CYP2C9 is even less successfully considered. Regarding Xenosite tool, it aims at predicting the metabolites from each individual human CYP. Its prediction with CYP2D6 shows that 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol is indeed the favored product. However additional metabolites are proposed whereas from the yeast *in vivo* production we do not detect such products. For CYP2C9, 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol is found as well, but the 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol production is not identified and multiple additional metabolites are expected. Additionally, some CYPs like CYP3A4 or CYP2C8 are proposed to convert (+)-*epi*- $\alpha$ -bisabolol to a wide range of new molecules while we do not observe this pattern in our *in vivo* screening. Overall, this demonstrates that validating the human CYPs activities and specificities is a necessary step.

Although the screening was implemented in a strain adequate for FPP and (+)-*epi*- $\alpha$ -bisabolol synthesis, both (+)-*epi*- $\alpha$ -bisabolol and farnesol were produced in similar amounts. *In vitro* testing of the positive hits from our library with both substrates revealed that all identified CYPs accepted both sesquiterpenes as substrates. However, when product purification on a subset of molecules was carried out from *in vivo* experiments, only (+)-*epi*- $\alpha$ -bisabolol oxidized products were purified. Although CYP2E1 and CYP2C19 were also shown to oxidize farnesol *in vitro*,<sup>274</sup> we could not find any evidence for the production of

such molecules with both isoforms in our *in vivo* experiments. The low farnesol hydroxylation activities in the selected CYP2 isoforms could be due to a weaker affinity or lower turnover number for farnesol than (+)-*epi*- $\alpha$ -bisabolol or, in the case of CYP2E1, potential ethanol inhibition of the CYP activity.<sup>274</sup> Alternatively, additional hurdles such as poor substrate accessibility inside yeast cells or a different compartmentalization of the two substrates might also account for the lack of farnesol metabolism.

Remarkably, most of the (+)-*epi*- $\alpha$ -bisabolol metabolites were secreted into the culture medium. Understanding whether passive or active transport or free diffusion cause this phenomenon would be valuable for further improvement of our chassis strain and would be also of interest to potentially excrete other types of molecules. The recent discovery of the positive role of heterologous expression of human transporters in excreting fatty alcohols from yeast is a strong advocacy for the search of potential transporter gene candidates that would increase product efflux to the culture medium.<sup>182</sup> Such efflux pumps not only could help metabolic engineering reach better titers by finding highly active transporters but may also, if inactivated, avoid the escape of intermediates prior to their reaction with specific enzymes in synthetic pathways.

Our data show accumulation of oxidized products of (+)-*epi*- $\alpha$ -bisabolol up to 64 mg/L without any enzyme engineering refinement strategy. Further improvements of these new metabolic pathways may be achievable by a finer tuning of the fluxes toward the side reactions,<sup>278</sup> or, alternatively, by CYP engineering which are not fully dedicated and evolved for this type of reaction. Our metabolic engineering strategy can now be extended to additional monoterpenes, sesquiterpenes or diterpenes backbones or even other classes of molecules such as flavonoids as some human CYPs are also known to metabolize them.<sup>279</sup> Plant secondary metabolism includes thousands of chemical backbones while human CYPs metabolize 70-80% of clinical drugs.<sup>15</sup> Combining these two characteristic features could open up a tremendous chemical space of active compounds. The remaining challenge of such studies would be the adjustment of relevant screens for biological activities of the newly generated molecules.

## E. Conclusion

The approach consisting in screening potentially new enzyme activities straightly *in vivo*, in an optimized chassis, possesses several advantages when it comes to metabolic engineering. On the one hand, all enzymes identified in first place are expressed in conditions that are relevant for product accumulation inside biologically active cells. While sometimes discrepancies between the enzymatic reactions and the production in whole cells may exist.<sup>280</sup> *In vivo* screening may allow a much higher throughput when dealing with cumbersome enzymes like CYPs that are harder to evaluate from the *in vitro* point of view due to their

endoplasmic reticulum targeting and tedious methods to isolate microsomal fractions. Even if our strategy, different from the classical enzyme mining from natural host, did not succeed in hernandulcin production, we clearly demonstrate the validity of our design in expanding the scope of known chemicals and exemplified, using (+)-*epi*- $\alpha$ -bisabolol, the generation of completely new molecules. Given the biological activities of bisabolol (anti-inflammatory, potentiating antibiotics) and hernandulcin (sweetener), these novel molecules may exhibit interesting properties that have now to be explored. To further expand the diversity of accessible molecules after hydroxylation, the addition of plant alcohol dehydrogenases to the system may enable the formation of keto groups in the hydroxylated sesquiterpene, mimicking potential biosynthetic pathways existing in plants<sup>7,251</sup> and potentially allowing access to hernandulcin when a CYP targeting the carbon at the C-2 position will be discovered.

### **Conflict of interest**

The authors declare no commercial or financial conflict of interest.

### **Author contributions**

MRS and GT conceived the study. ASL, MRS and GT designed the experiments. ASL performed all biological and analytical experiments with the help of RF for the purification of products; RF performed the NMR acquisition and analysis. DKR participated in the design of the early stages of the project and provided the EPY300 strain. ASL wrote the first draft of the manuscript, which was then revised by all authors.

### **Fundings**

ASL received a PhD fellowship from the French minister of education research. Other than ASL salary, this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### **Acknowledgments**

We acknowledge Thomas Lautier for the supply of microsomal fractions of CYP2B11. NMR analyses were performed using facilities at MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France, [www.metatoul.fr](http://www.metatoul.fr)), which is part of the national infrastructure MetaboHUB (The French National infrastructure for metabolomics and fluxomics, [www.metabohub.fr](http://www.metabohub.fr)) and is supported by grants from the Région Midi-

Pyrénées, the European Regional Development Fund, SICOVAL, IBiSa-France, CNRS, and INRA. Edern Cahoreau, staff member of the MetaToul platform, is gratefully acknowledged for the technical support. And we would like to thank the ICEO facility dedicated to enzyme screening and discovery, part of the Integrated Screening Platform of Toulouse (PICT, IBiSA), for providing access to GC-FID-MS.

## F. Supporting information

Table 4: CYPs library cloned in the multi cloning site of pYEDP60 and screened for the oxidation of (+)-*epi- $\alpha$ -bisabolol*. Human CYPs are depicted in red, animals CYPs are depicted in light red.

CYP nomenclature name	Related organism	UniprotKB accession number	Additional features
<b>CYP1A1</b>	<i>Homo sapiens</i>	P04798	
<b>CYP1A2</b>	<i>Homo sapiens</i>	P05177	
<b>CYP1B1</b>	<i>Homo sapiens</i>	Q16678	
<b>CYP1B1 3A</b>	<i>Homo sapiens</i>	Q16678	Kozak sequence optimization: with 3 adenines before ATG start codon
<b>CYP2D6</b>	<i>Homo sapiens</i>	P10635	
<b>CYP2D6 3A</b>	<i>Homo sapiens</i>	P10635	Kozak sequence optimization: with 3 adenines before ATG
<b>CYP2D6 F1</b>	<i>Homo sapiens</i>	P10635	Kozak sequence optimization <sup>a</sup>
<b>CYP2D6 H1</b>	<i>Homo sapiens</i>	P10635	Kozak sequence optimization <sup>a</sup>
<b>CYP2A6</b>	<i>Homo sapiens</i>	P11509	
<b>CYP2A6 3A</b>	<i>Homo sapiens</i>	P11509	Kozak sequence optimization: with 3 adenines before ATG
<b>CYP2A6 H1</b>	<i>Homo sapiens</i>	P11509	Kozak sequence optimization <sup>a</sup>
<b>CYP2E1</b>	<i>Homo sapiens</i>	P05181	
<b>CYP3A4</b>	<i>Homo sapiens</i>	P08684	

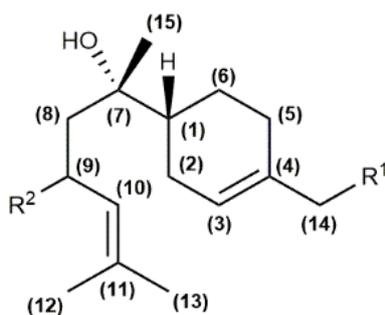
CYP3A5	<i>Homo sapiens</i>	P20815	
CYP2B6	<i>Homo sapiens</i>	P20813	
CYP2C8	<i>Homo sapiens</i>	P10632	
CYP2C9	<i>Homo sapiens</i>	P11712	
CYP2C18	<i>Homo sapiens</i>	P33260	
CYP2C19	<i>Homo sapiens</i>	P33261	
CYP1A2 I117T	<i>Homo sapiens</i>	Described in <sup>281</sup>	
CYP1A2 D348N	<i>Homo sapiens</i>	Described in <sup>282</sup>	Natural polymorphism
CYP1A2 C406Y	<i>Homo sapiens</i>	Described in <sup>282</sup>	Natural polymorphism
CYP1A2 R431W	<i>Homo sapiens</i>	Described in <sup>282</sup>	Natural polymorphism
CYP3A4 /b5	<i>Homo sapiens</i>	P08684 + P00167	Coexpression of human cytochrome b5
CYP1A1	<i>Mus musculus</i>	P00185	
CYP1A2	<i>Mus musculus</i>	P00186	
pLM46-1 1A1	<i>Oryctolagus cuniculus</i>	Chimera (20% CYP1A2 N-term and 80% CYP1A1)	Described in <sup>283</sup>
CYP1A2	<i>Oryctolagus cuniculus</i>	P00187	
CYP2B11	<i>Canis lupus familiaris</i>	P24460	
CYP1C1	<i>Danio rerio</i>	Q503H0	
CYP1C2	<i>Danio rerio</i>	BOJZN5	
CYP1D1	<i>Danio rerio</i>	Q5U396	

<sup>a</sup>Kozak sequence optimization: “F1” sequence adds AAA before ATG while first codon is changed to GGT, “H1” sequence adds AAA before ATG while first codon is changed to GAT in both cases after the stop an additional A was also added.

Table 5: Primers used in this study, restriction sites are underlined.

Primer number	Sequence	Amplicon	Purpose
1	AGTAGAATT <u>C</u> AAAAATGAATTCCACATC CAGGAGATCAGCCAAC	BBS flanked with EcoRI	Cloning in pMRI31 – PMRI32
2	AGTCAGATCTTCATGGAAGAGGAATGG GTTC	BBS flanked with BgIII	Cloning in pMRI31 – PMRI32
3	TATAAAGATCTAAAAATGCCGTTTGGGA ATAGACAAC	<i>S. cerevisiae</i> CPR flanked with BgIII	Cloning in pMRI31 – PMRI32
4	TATAAGTCGACTTACCAGACATCTTCTT GGTATC	<i>S. cerevisiae</i> CPR flanked with Sall	Cloning in pMRI31 – PMRI32
5	TATAAGGATCCAAAAATGGGAGACAGT CACGTGGACACCAGCTCC	<i>H. sapiens</i> CPR flanked with BamHI	Cloning in pMRI31 – PMRI32
6	TATAAGTCGACCTAGCTCCACACGTCCA GG	<i>H. sapiens</i> CPR flanked with Sall	Cloning in pMRI31 – PMRI32
7	AAGGGAATTCGAATTTTCAAAAATTCTT ACTTTT	pMRI plasmid	Removing BBS from plasmid
8	AGATCTGAGCTCTTAATTAACAATTCTT CC	pMRI plasmid	Removing BBS from plasmid

Table 6: Summary of NMR data of (+)-*epi*- $\alpha$ -bisabolol<sup>6,284,285</sup> 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol produced by CYP2D6, and presumed 9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol produced by CYP2C9.<sup>286</sup>



(+)-*epi*- $\alpha$ -bisabolol: R<sup>1</sup> = H, R<sup>2</sup> = H

14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol: R<sup>1</sup> = OH, R<sup>2</sup> = H

9-hydroxy-(+)-*epi*- $\alpha$ -bisabolol: R<sup>1</sup> = H, R<sup>2</sup> = OH

	(+)- <i>epi</i> - $\alpha$ -bisabolol	14-Hydroxy-(+)- <i>epi</i> - $\alpha$ -bisabolol	Presumed 9-hydroxy-(+)- <i>epi</i> - $\alpha$ -bisabolol <sup>1</sup> H
Carbon number	$\delta$ (ppm) <sup>1</sup> H / <sup>13</sup> C (C <sub>6</sub> D <sub>6</sub> )	$\delta$ (ppm) <sup>1</sup> H / <sup>13</sup> C (CDCl <sub>3</sub> )	$\delta$ (ppm) <sup>1</sup> H / <sup>13</sup> C (CDCl <sub>3</sub> )
1	43.4	1.63-1.58 / 43.4	
2	26.1	2.16-1.85 / 26.6	
3	5.44 / 120.9	5.71 / 122.5	5.41 / 126.1
4	- / 133.8	- / 137.3	
5	31.1	2.16-1.85 / 25.8	
6	24.1	1.31-1.23 / 29.7	
7	- / 74.3	- / 74.3	
8	1.45 / 39.4	1.52 / 39.4	
9	22.3	2.16-1.85 / 22.3	
10	5.22 / 124.7	5.13 / 124.5	5.38 / 120.6
11	- / 131.7	- / 131.9	
12	1.58 / 17.7	1.63 / 17.7	
13	1.68 / 25.7	1.69 / 28.7	
14	1.64 / 24.1	4.01 / 67.2	1.63 / 23.3
15	0.99 / 23.4	1.15 / 24.0	1.11 / 24.1

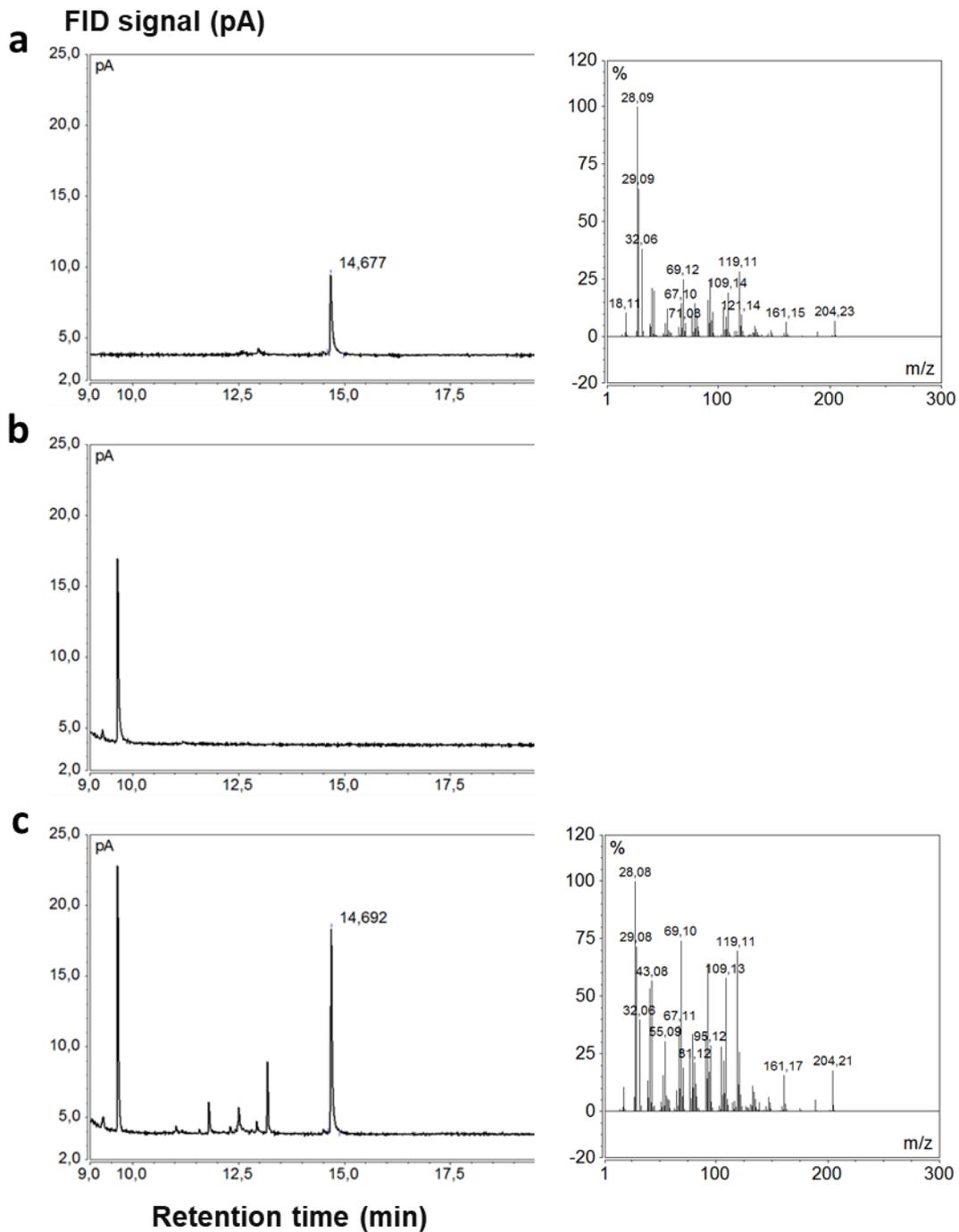


Figure 19: In vivo production of (+)-epi- $\alpha$ -bisabolol in BY4741 in SD-Leu and EPY300 in SD-His-Met-Leu after expression of the BBS on a high copy plasmid. GC-FID chromatograms of (a) (-)- $\alpha$ -bisabolol standard along with its MS spectrum; (b) BY4741 + pESC leu2d BBS; (c) EPY300 + pESC leu2d BBS along with the MS spectrum of (+)-epi- $\alpha$ -bisabolol. For this preliminary set of injections, the GC was run with a constant pressure of 75 kPa while the further experiments were performed with a constant flow of 5 ml/min of  $N_2$ .

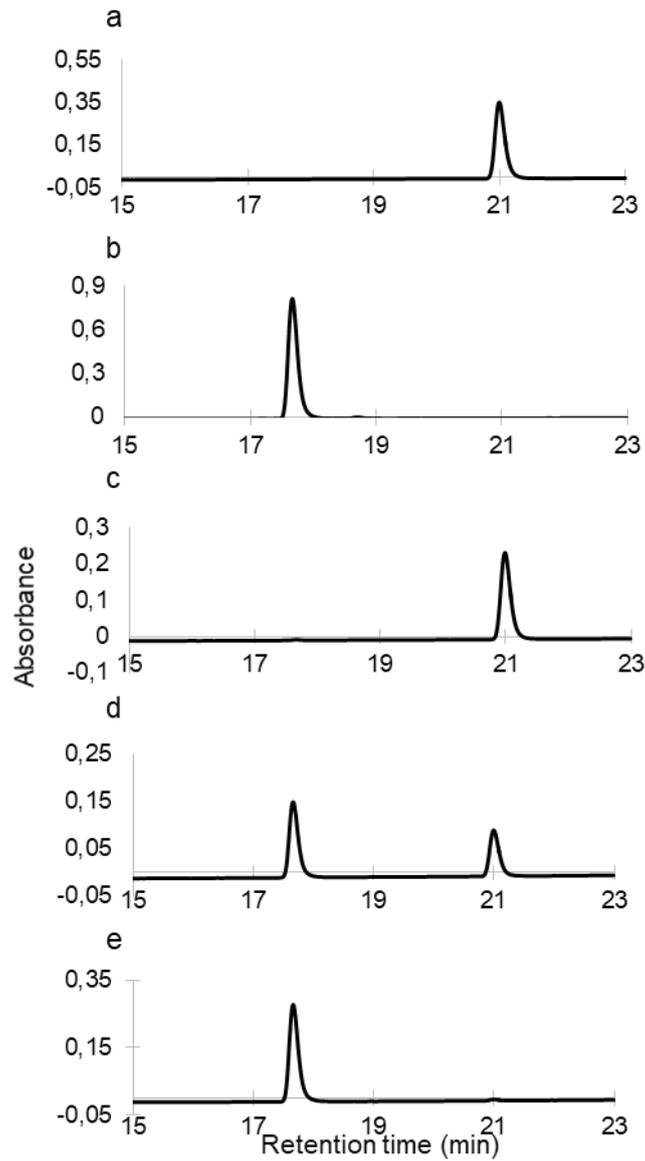


Figure 20: Bioconversion of diclofenac to 4-hydroxy-diclofenac by strains YBIS\_01 and YBIS\_02. Extracted molecules were separated by HPLC-UV and metabolites quantified at 276 and 267 nm respectively. (a) Diclofenac standard; (b) 4-hydroxy-diclofenac standard; (c) EPY300 + CYP2C9; (d) YBIS\_01 + CYP2C9; (e) YBIS\_02 + CYP2C9.

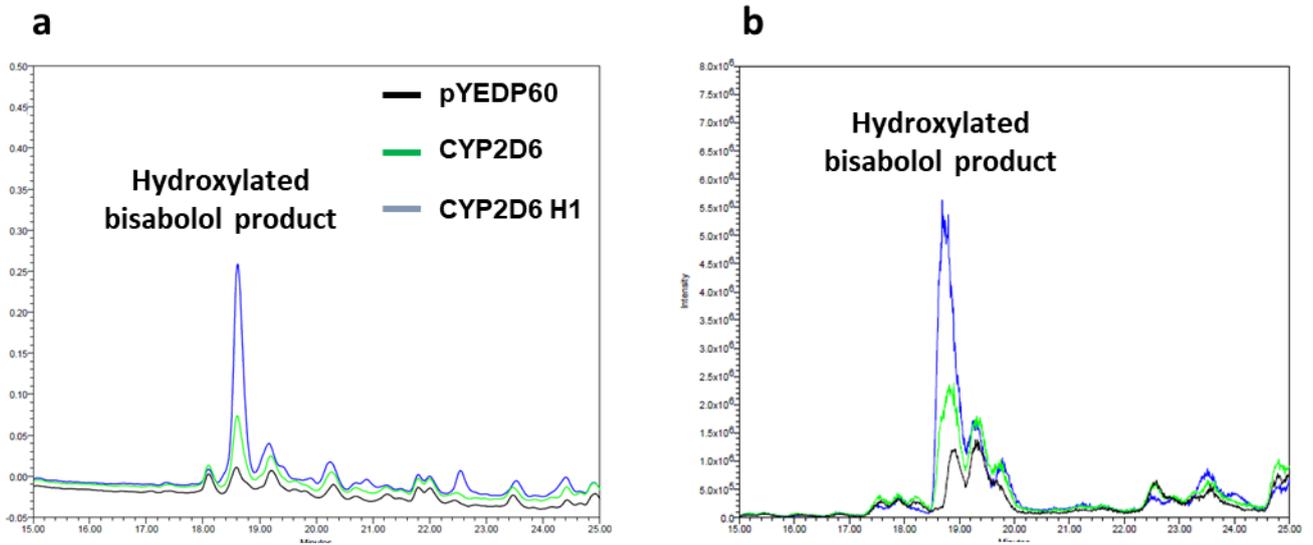


Figure 21: An optimized Kozak environment for the expression of CYP2D6 enzyme enhances the (+)-*epi*- $\alpha$ -bisabolol oxidized product accumulation.  
 (a) UV chromatogram at 205 nm; (b) MS chromatogram at  $m/z = 203.3$  corresponding to hydroxylated (+)-*epi*- $\alpha$ -bisabolol molecules in accordance with (a).

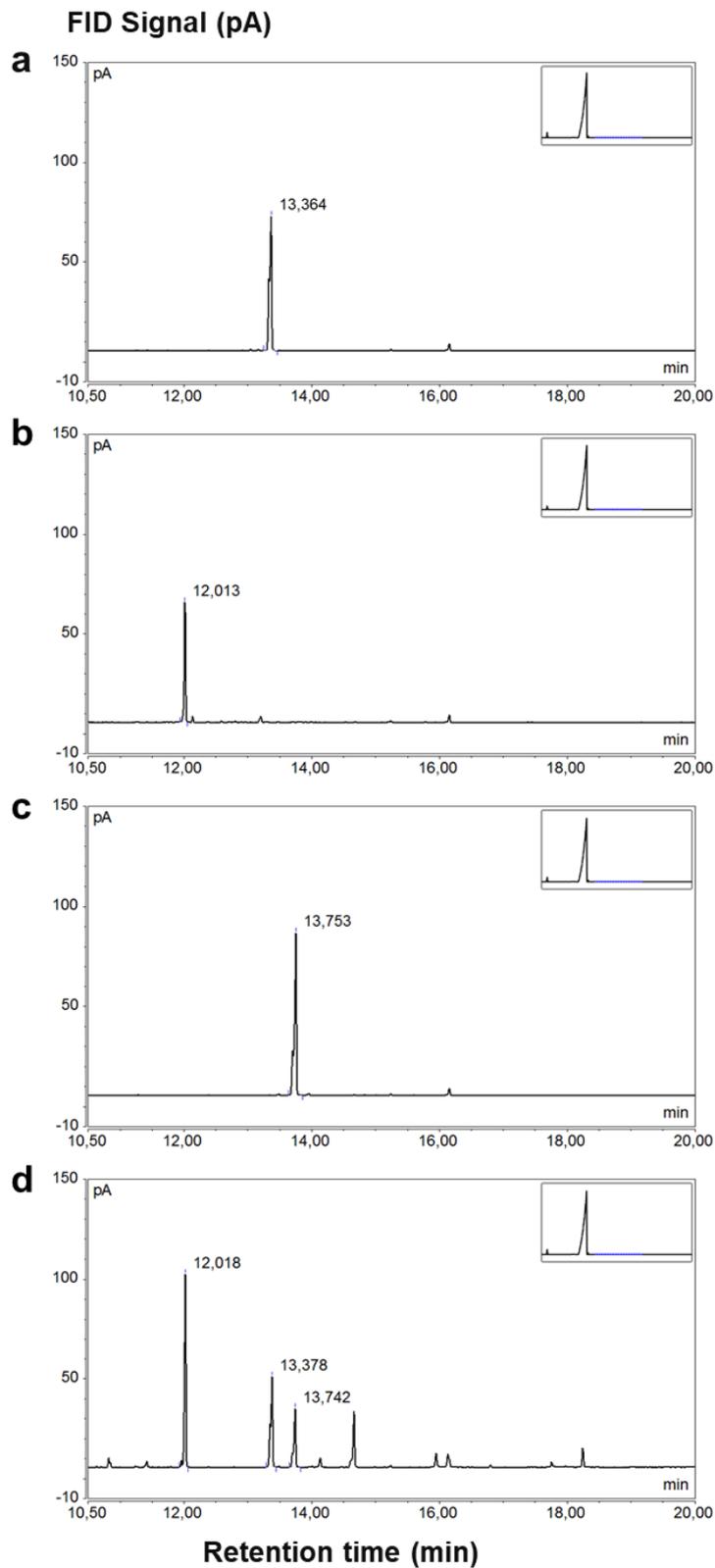


Figure 22: Chromatograms of GC-FID signal showing the sesquiterpene content in the chassis strain YBIS\_01 cultured in SD medium with a dodecane overlay. (a) (-)- $\alpha$ -bisabolol, (b) nerolidol, (c) farnesol, (d) YBIS\_01.

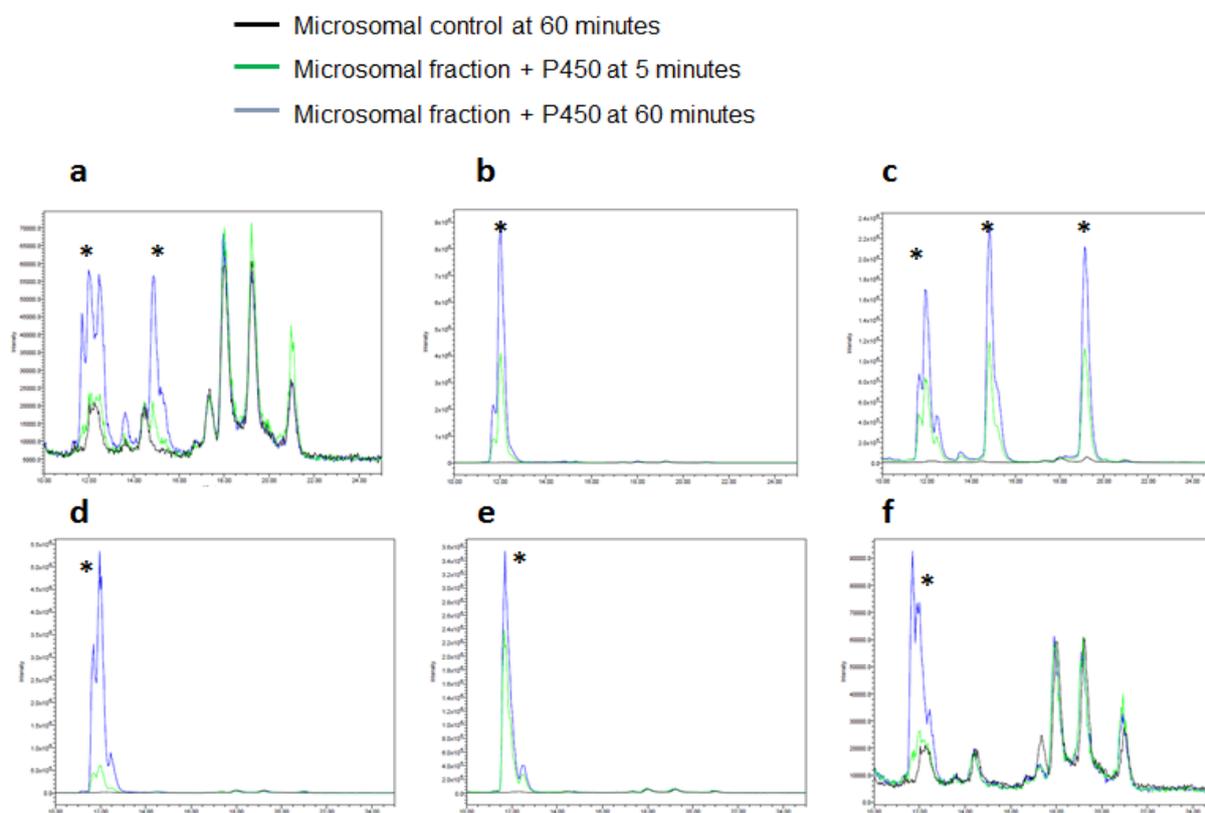


Figure 23: MS chromatograms at  $m/z = 203.3$  of the enzymatic reactions with  $(-)\text{-}\alpha\text{-bisabolol}$  and demonstrating the *in vitro* activities of the CYPs.

(a) CYP1A1; (b) CYP2B6; (c) CYP2B11; (d) CYP2C9; (e) CYP2C18; (f) CYP2D6. The oxidized products of  $(-)\text{-}\alpha\text{-bisabolol}$  that were detected are indicated with asterisks. Enzymatic reactions were carried out as described in materials and methods; the microsomal control corresponds to microsomal fractions containing the yeast CPR but no CYP.

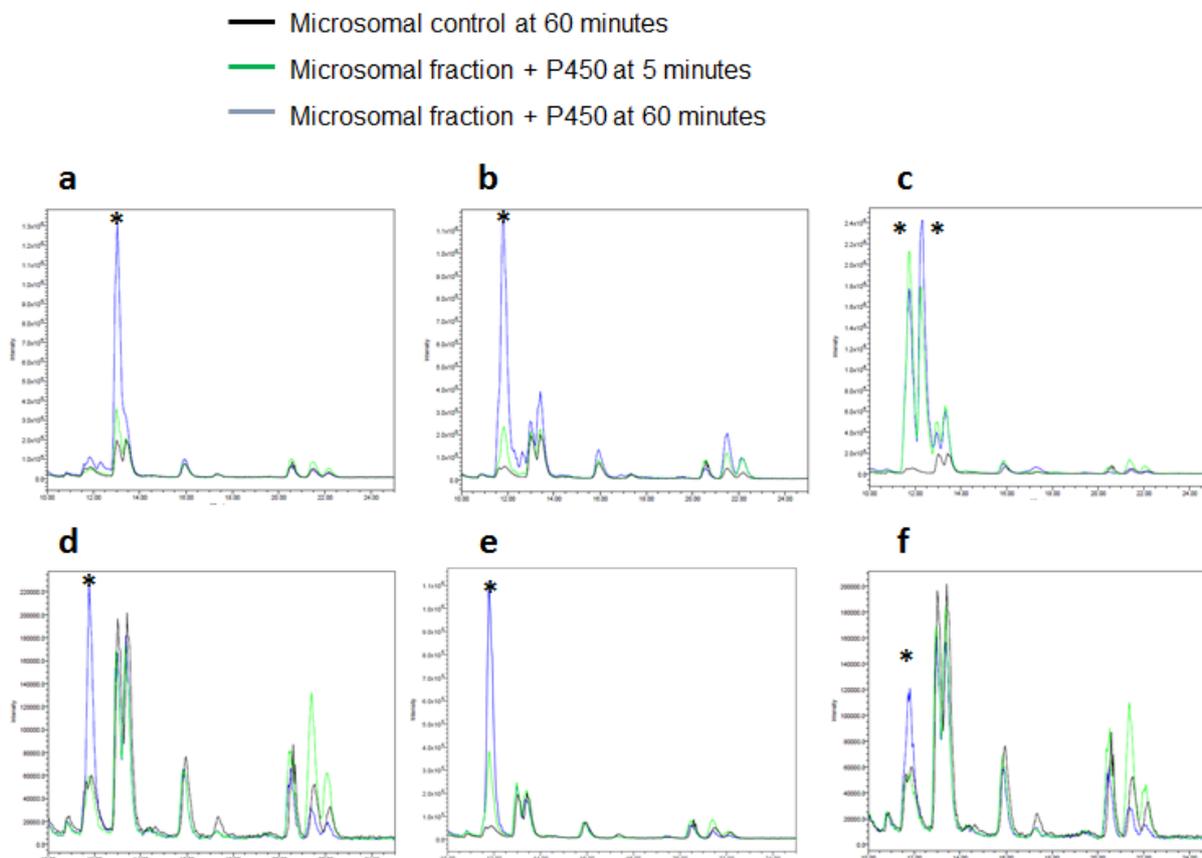


Figure 24: MS chromatograms at  $m/z = 203.3$  of the enzymatic reactions with farnesol and demonstrating the *in vitro* activities of the CYPs. (a) CYP1A1; (b) CYP2B6; (c) CYP2B11; (d) CYP2C9; (e) CYP2C18; (f) CYP2D6. The oxidized products of (-)- $\alpha$ -bisabolol that were detected are indicated with asterisks.

— Microsomal control at 60 minutes  
 — Microsomal fraction + P450 at 5 minutes  
 — Microsomal fraction + P450 at 60 minutes

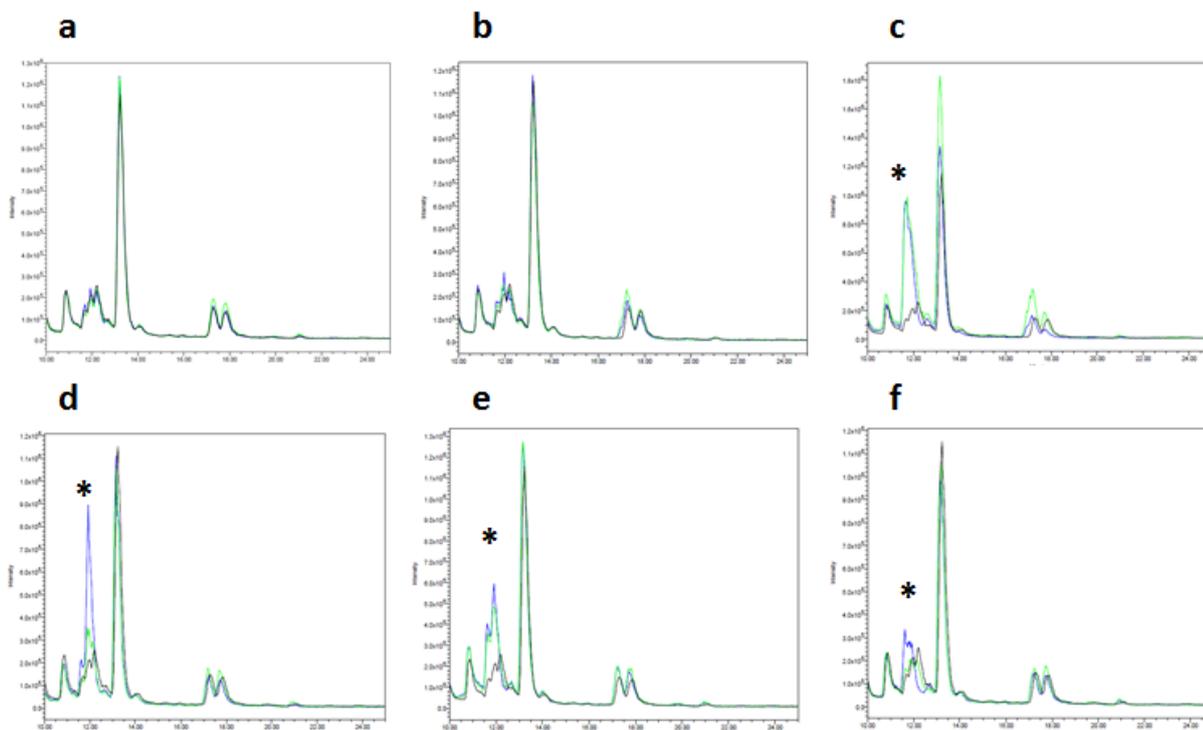


Figure 25: MS chromatograms at  $m/z = 203.3$  of the enzymatic reactions with nerolidol and demonstrating the *in vitro* activities of the CYPs.

(a) CYP1A1; (b) CYP2B6; (c) CYP2B11; (d) CYP2C9; (e) CYP2C18; (f) CYP2D6. The oxidized products of nerolidol that were detected are indicated with asterisks.

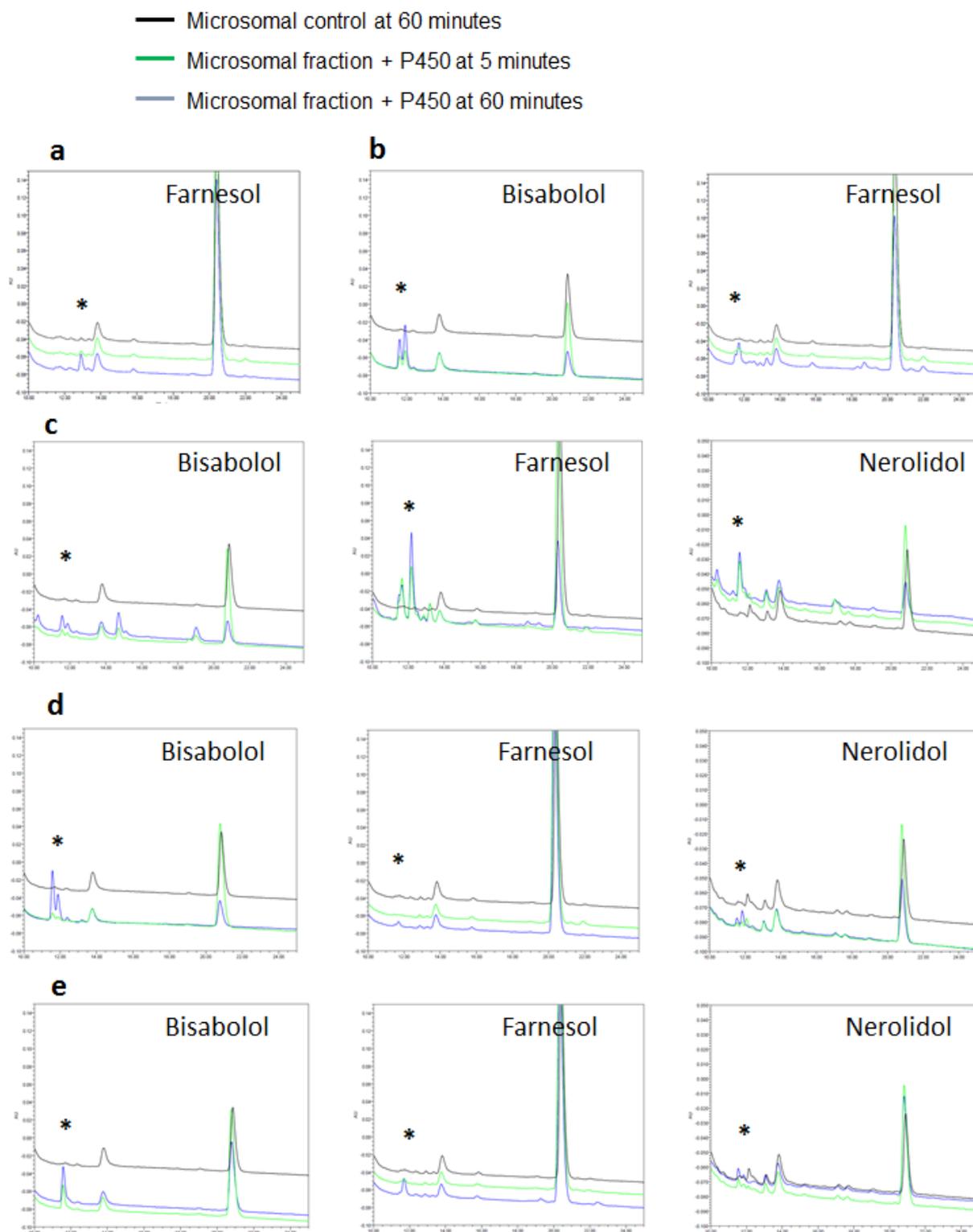


Figure 26: UV chromatograms at 205 nm validating *in vitro* activities of the CYPs towards (-)- $\alpha$ -bisabolol, farnesol and nerolidol.

(a) CYP1A1; (b) CYP2B6; (c) CYP2B11; (d) CYP2C9; (e) CYP2C18. The oxidized products are indicated with asterisks. Contrarily to Figure 23, Figure 24, Figure 25, only reactions where quantifiable products were detected are shown for the ease of reading.

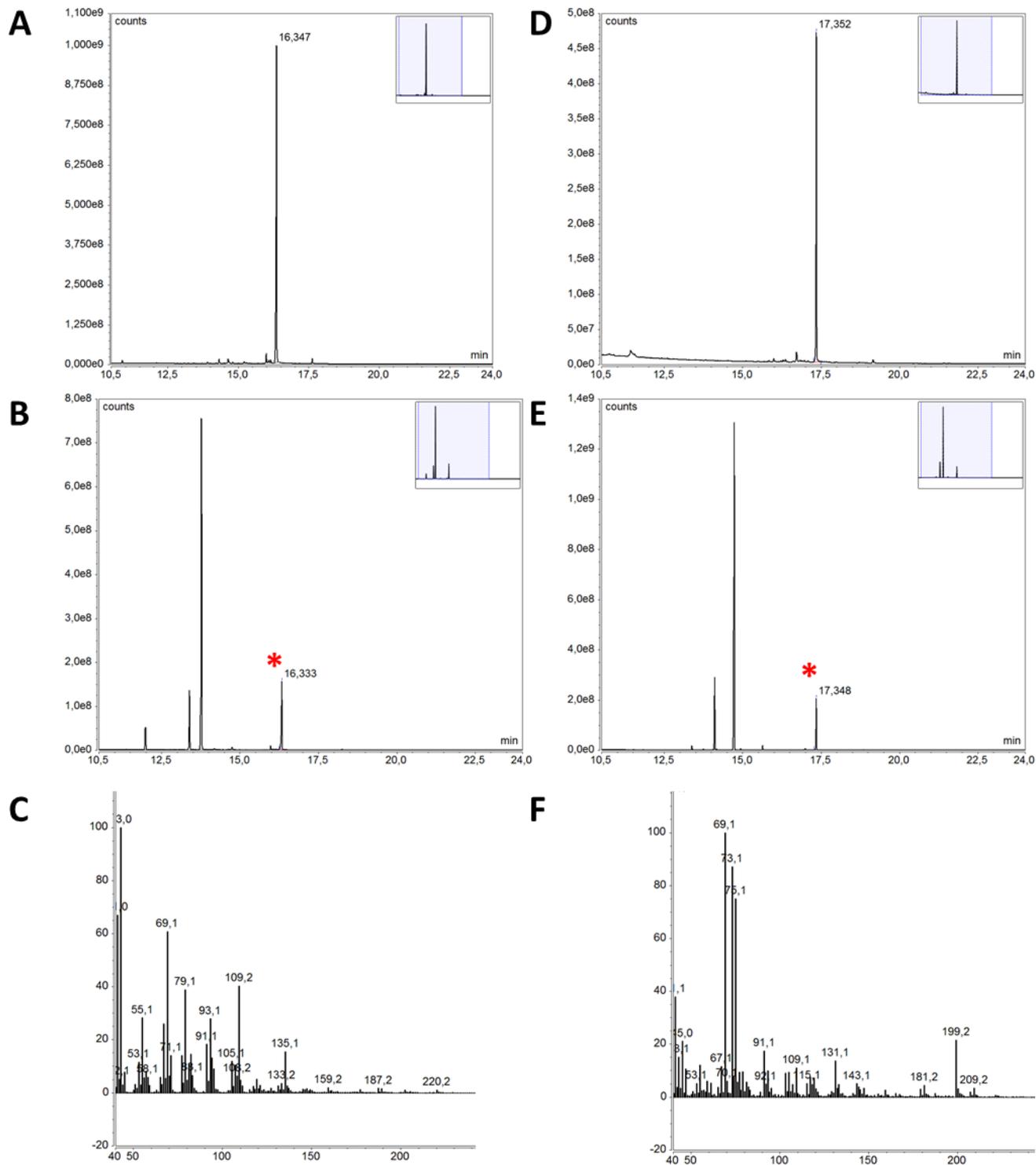


Figure 27: Total ion current chromatograms of GC-MS showing identical product formation between the CYP2D6 and CYP2C18 enzymes.

(A) (+)-*epi*- $\alpha$ -bisabolol oxidized product purified from YBIS\_06 + CYP2D6 (B) YBIS\_06 + CYP2C18 metabolite extraction (C) MS spectra of (+)-*epi*- $\alpha$ -bisabolol oxidized product purified from YBIS\_06 + CYP2D6 (D) (+)-*epi*- $\alpha$ -bisabolol oxidized product purified YBIS\_06 + CYP2D6 and trimethylsilylated using BSA TMCS (E) YBIS\_06 + CYP2C18 metabolite extraction and trimethylsilylated (F) MS spectra of (+)-*epi*- $\alpha$ -bisabolol oxidized product and trimethylsilylated. The asterisk highlights the product from CYP2C18 which has same retention time and same MS spectrum as CYP2D6 product.

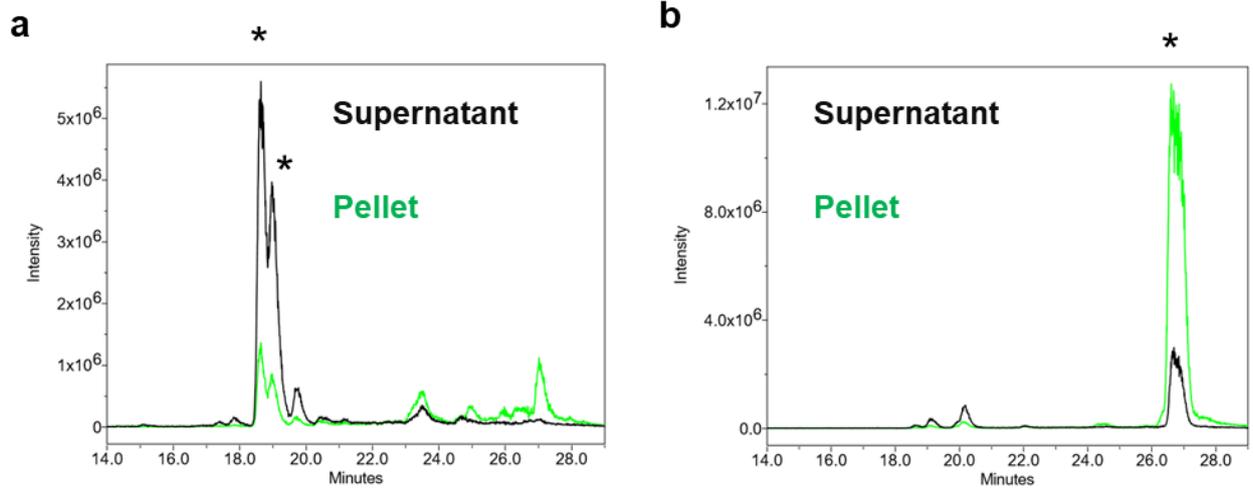
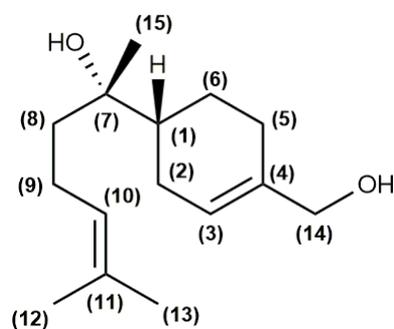


Figure 28: Metabolite partition in the cell pellet and in the culture broth of strain YBIS\_06 + CYP2C9. 5 ml of cells and culture broth were separated by centrifugation and extracted by hexane. After extraction the fractions were resuspended in same volume of ethanol. (a) MS chromatogram at  $m/z = 203.3$  indicating the (+)-*epi*- $\alpha$ -bisabolol hydroxylated products location (b) MS chromatogram at  $m/z = 205.3$  indicating the overall sesquiterpene location. The  $m/z = 205.3$  detects (+)-*epi*- $\alpha$ -bisabolol, *trans,trans*-farnesol and nerolidol and their retention times overlaps in the presented condition. \* The followed metabolites are indicating with the stars.



$^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  5.71 (1H, br m, H-3), 5.13 (1H, br t,  $J=7.1$ , H-10), 4.01 (2H, m, H-14), 2.16-1.85 (6H, m, H-2<sub>a</sub>/H-5<sub>a</sub>, H-9<sub>a</sub> and b, H-2<sub>b</sub> and H-5<sub>b</sub>), 1.69 (3H, br d,  $J=0.9$ , H-13), 1.63 (3H, s, H-12), 1.63-1.58 (1H, m, H-1), 1.52 (2H, t,  $J=8.6$ , H-8), 1.31-1.23 (2H, m, H-6), 1.15 (3H, s, H-15);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  137.3 (C-4), 131.9 (C-11), 124.5 (C-10), 122.5 (C-3), 74.3 (C-7), 67.2 (C-14), 43.4 (C-1), 39.4 (C-8), 29.7 (C-6), 26.6 (C-2), 25.8 (C-5), 28.7 (C-13), 24.0 (C-15), 22.3 (C-9), 17.7 (C-12).

Figure 29: NMR analysis of 14-hydroxy-(+)-epi- $\alpha$ -bisabolol produced by CYP2D6.

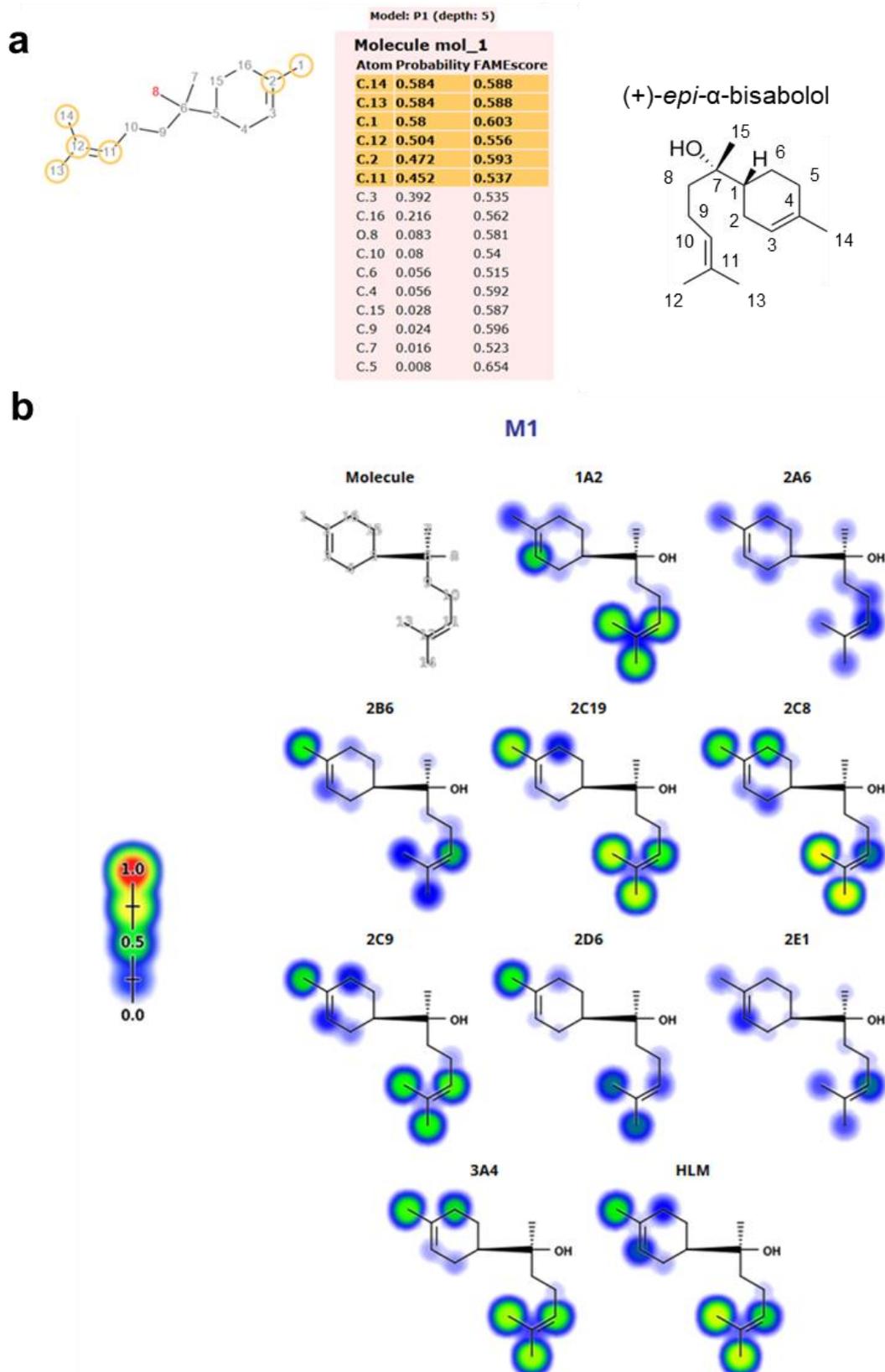


Figure 30: The predicted metabolism of human CYPs with (a) FAME3 algorithm, (b) Xenosite algorithm. Carbon numbering of (+)-*epi*- $\alpha$ -bisabolol used both in literature and in this study is recalled as it differs from the FAME3.

## G. Additional investigations around the synthesis of (+)-*epi*- $\alpha$ -bisabolol derivatives formed by CYPs expressed in an optimized yeast chassis strain

### 1. Material and methods

The separation of hydroxy-(+)-*epi*- $\alpha$ -bisabolol derivatives was optimized with a Prontosil 120-3-C18-AQ 3.0  $\mu$ m 125\*4.0 mm column (Bischoff chromatography) equipped with a C18 precolumn. The elution conditions on HPLC relied on an isocratic method composed of 60 % (water + 0.03% (v/v) formic acid) and 40 % (acetonitrile + 0.03% (v/v) formic acid) at 0.75 mL/min.

Reduction of ( $\pm$ )-hernandulcin was attempted with sodium borohydride as reductant. A first reaction was set using ( $\pm$ )-hernandulcin 5  $\mu$ L, 0.5 molar equivalent of NaBH<sub>4</sub> in ethanol and let react 1 hour at room temperature. The reaction was stopped by addition of acetic acid to a final concentration of 10 mM and a tenfold dilution was injected in HPLC to analyze the products. In a second experiment, the reaction was incubated on ice and then stopped with HCl instead of acetic acid.

To introduce a different regulation of the heterologous pathway, a supplementary plasmid was constructed by inserting the PDC1 promoter into pMRI34S+BBS in order to achieve a strong constitutive expression of BBS instead of the galactose inducible one. For that purpose, the PDC1 promoter from *S. cerevisiae* was amplified using the primers 2.1/2.2 (described in appendix Table 18) and the pMRI34S + BBS using the primers 2.3/2.4. Subsequently, the two fragments were assembled by Gibson cloning to generate the plasmid pMRI34S + PDC1 BBS. This plasmid was then integrated in YBIS\_01 whose selection marker was previously recycled to build the strain YBIS\_07 (see appendix Table 19).

In order to test plant CYPs, a chassis strain expressing a plant CPR was designed. Similarly to the plasmids pMRI31H+ BBS and pMRI31S+ BBS, pMRI31A + BBS was built by amplifying *Arabidopsis thaliana* NADPH-cytochrome P450 reductase 1 using the primers 2.5/2.6. This plasmid was then integrated in the strain EPY300 leading to the strain YBIS\_08 (appendix Table 19). Of the tested plant CYPs, all were already cloned in pYEDP60 except CYP71BE5, CYP71D495, CYP706M1, and CYP726A20. These genes were ordered as synthetic DNA fragments from Twist Biosciences (San Francisco, USA) after codon optimization for *S. cerevisiae* (IDT, codon optimization webtool) and cloned in the pYEDP60 using BamHI / EcoRI sites (Appendix DNA sequences 1).

The *in vivo* assay of ADHs relied on their expression from a low copy plasmid compatible with the pYEDP60 ones expressing the different mammalian CYPs. The coding sequences of the ADHs were codon optimized for *S. cerevisiae* expression and ordered as synthetic DNA fragments from Twist Biosciences (San Francisco, USA), Appendix DNA sequences 2. The backbone plasmid pFPP13 that bears Leu2 marker and an ARS CEN6 replication origin was amplified using 2.7/2.8 primers. The synthetic DNA fragments were used straightly for Gibson cloning in between TEF1 promoter and PGK terminator. The resulting library of plant ADHs (7 ADHs and control vector) and mammals CYPs (pYEPD60 control vector, CYP2B6, CYP2B11, CYP2C9, CYP2D6) was then transformed in YBIS\_01 using the protocol described in <sup>264</sup>, and selected on SD-His-Met-Ura-Leu plates.

For the *in vivo* production of the additional  $\alpha$ -bisabolol isomers, the coding sequences of the *A. kurramensis*, *M. recutita* and *S. citricolor* BBS were purchased from Baseclear B.V (Leiden, Netherlands) and cloned in EcoRI/BglII sites of pMRI31 plasmids (Appendix DNA sequences 3). These integrative vectors were transformed into EPY300 strain leading to YBIS\_09, YBIS\_10, and YBIS\_11 respectively (Appendix Table 19).

Concerning all the above experiments, metabolites were analyzed as described previously in HPLC-UV-MS after yeast expression cultures. The only exception involved the evaluation of the effect with the supplementary copy BBS under PDC1 or Gal10 promoters and the additional BBS enzymes evaluation for which cultures and metabolite analysis was carried out as already described with the addition of a dodecane overlay (10% (v/v) of the culture) and examined in GC-FID-MS.

## 2. Results and discussion

### a) Further characterization of oxidized (+)-*epi*- $\alpha$ -bisabolol products and attempts to purify them

In order to characterize the (+)-*epi*- $\alpha$ -bisabolol oxidized products obtained from the strains YBIS\_06 + CYP2B6, CYP2B11 and CYP2C9, the separation method had to be improved and scaled up. For that purpose, a different C18 column (Prontosil 120-3-C18-AQ instead of Prontosil 120-5-C18-SH) was tested and the solvent and gradient optimized. First a linear gradient was applied and two solvents, acetonitrile and methanol were compared for the elution. The use of methanol showed a decreased efficiency in the separation; hence, acetonitrile was used in further refinements of the separation conditions. A method relying on optimized gradients of acetonitrile and several isocratic conditions were evaluated. Finally, an isocratic method consisting of 60/40 (acetonitrile/water) supplemented with 0.03% (v/v) formic acid was the most efficient

condition as shown in Figure 31. Separation of the two YBIS\_06 + CYP2C9 was achieved with the new method (Figure 31A). For the products originating from YBIS\_06 + CYP2B6 and YBIS\_06 + CYP2B, the separation of the products was improved. Even though an exact determination of the product number and the clear separation of the various species could not be reached (Figure 31B, C).

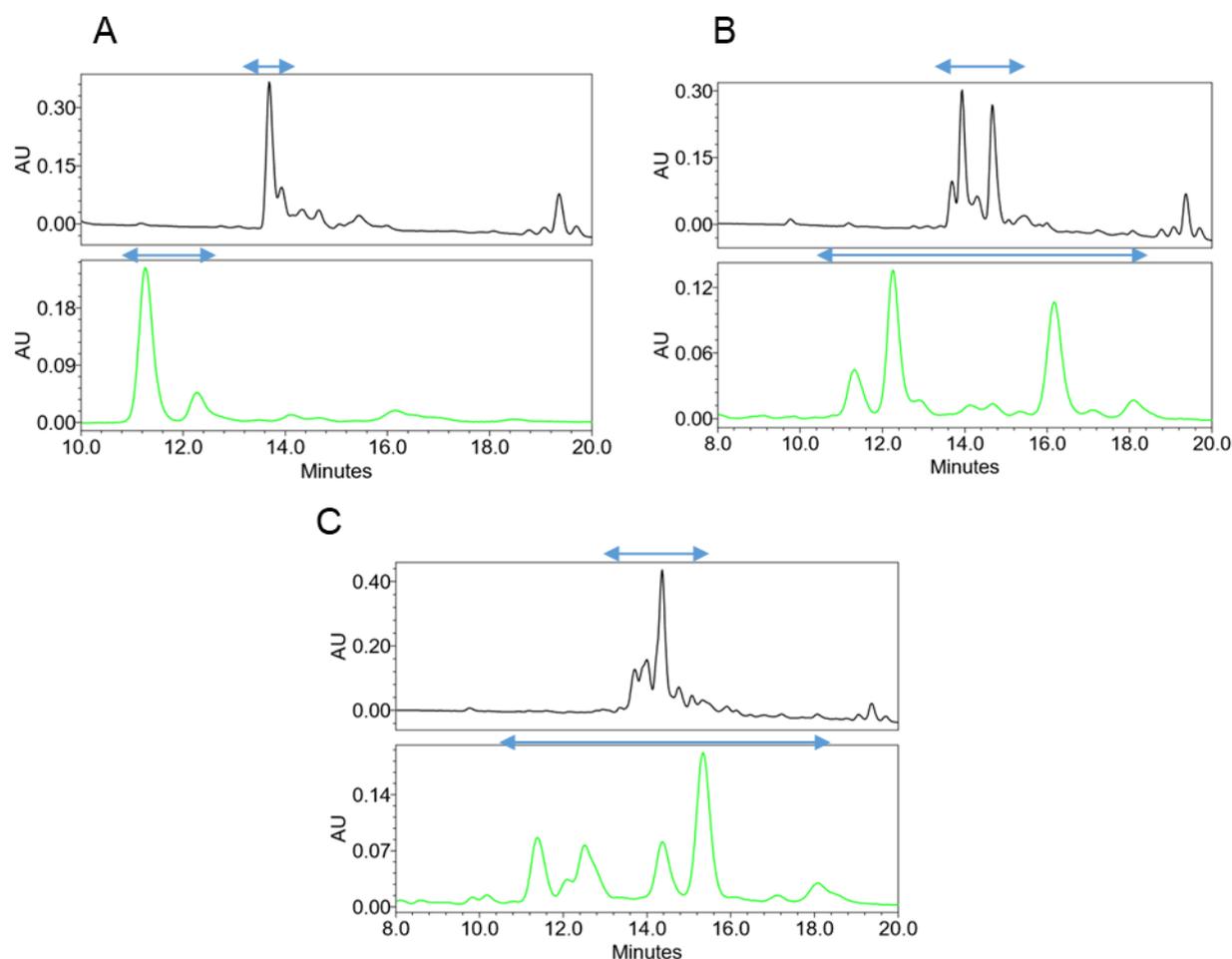


Figure 31: Optimization of the separation of the oxidized sesquiterpene products from (A) YBIS\_06 + CYP2C9 (B) YBIS\_06 + CYP2B6 (C) YBIS\_06 + CYP2B11.

The blue arrows depict the windows where oxidized (+)-epi- $\alpha$ -bisabolol and/or farnesol products are eluted in the different samples according to a consistent MS signal. The upper black chromatograms correspond to the initial separation with a linear gradient over the column from 15% acetonitrile to 100% for 17 minutes. The lower green chromatograms to the injection of the same sample using the optimized isocratic separation. The absorbance of the products was followed at 205 nm.

As described in Chapter II, the fractions of the strains YBIS\_06 + CYP2B6 and CYP2B11 were first purified using an octadecyl-functionalized silica gel and refined using a silica cartridge. As the attempt to scale up the separation using a C18-AQ column with a larger diameter section did not succeed, the Prontosil 120-3-C18-AQ column supposed to be used for analytical purpose was also used for collecting the fractions of the products to purify. Several fractions were collected over time on a HPLC-UV module Agilent infinity 1260

equipped with a UV module ultimate VWD 3000 coupled with Dionex Ultimate 3000 automated fraction collector (see appendix Figure 67) for the corresponding chromatograms. The material was repetitively injected over the column by loading  $\sim 1/20$  of the semi purified products solubilized in DMSO. The recovered fractions are now pending validation of the purity of each sample before a potential NMR analysis.

In a supplementary approach, we investigated the possibility to chemically synthesize the putative precursor of hernandulcin, the 2-hydroxy-(+)-*epi*- $\alpha$ -bisabolol. Indeed, having access to this chemical standard would allow us to check if the potential hernandulcin precursor could be detected amongst the products for which structural information was not obtained with the NMR data. This would allow a comparison of the MS spectra in LC and GC (with and without silylation) with the other oxidized (+)-*epi*- $\alpha$ -bisabolol products. The reduction of ( $\pm$ )-hernandulcin was attempted using sodium borohydride and the products characterized in LC-UV/MS (Figure 32). ( $\pm$ )-hernandulcin limitedly reacted in the two conditions tested (Figure 32B, C). However, multiple products were obtained, and no signal consistent with the 2-hydroxy-(+)-*epi*- $\alpha$ -bisabolol synthesis was identified. Instead, in both conditions, the reaction led to the concomitant reduction of conjugated double bond even at low temperature. Compadre et al. previously reported the reduction of hernandulcin<sup>3</sup>, their catalysis relied on the use of an additional lanthanide salt that protects the conjugated double bond as evidenced by Luche<sup>287</sup>. We carried out another synthesis with the addition of the samarium salt and products obtained from the reaction did not show the presence 2-hydroxy-(+)-*epi*- $\alpha$ -bisabolol either (data not shown). Further optimization of the chemical synthesis is then necessary to reach better conditions as Compadre et al. indicated the yield of reduction to be 40%.

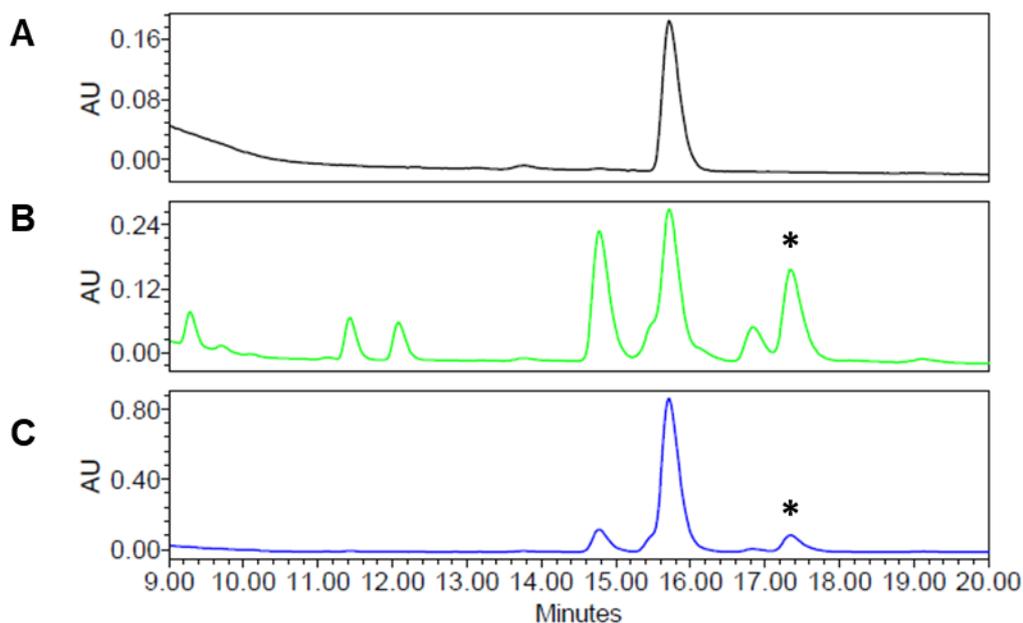


Figure 32: UV chromatogram (205 nm) of the chemical reaction aiming at reducing ( $\pm$ )-hernandulcin (A) ( $\pm$ )-hernandulcin standard, (B) 1-hour reaction at room temperature stopped by acetic acid, (C) 1-hour reaction incubated on ice stopped by HCl.

\* Highlights the product with concomitant reduction of conjugated double (this compound did not absorb at 238 nm like hernandulcin. Hence, the ketone was reduced, but the  $m/z$  is consistent with reduction of the conjugated double bond).

b) Towards a finer regulation of the heterologous (+)-*epi*- $\alpha$ -bisabolol production in yeast

As evidenced earlier in Figure 15, the strains in which we introduced the BBS did not exclusively accumulated (+)-*epi*- $\alpha$ -bisabolol but also side products such as farnesol and nerolidol probably due to a too weak expression of BBS. Thus, a tighter regulation of the heterologous pathway is required. The time course of the (+)-*epi*- $\alpha$ -bisabolol production was investigated using YBIS\_08 strain (see appendix Table 19 for the strain genotype) and we noticed that (+)-*epi*- $\alpha$ -bisabolol production occurred in late stages of the cultures (Figure 33).

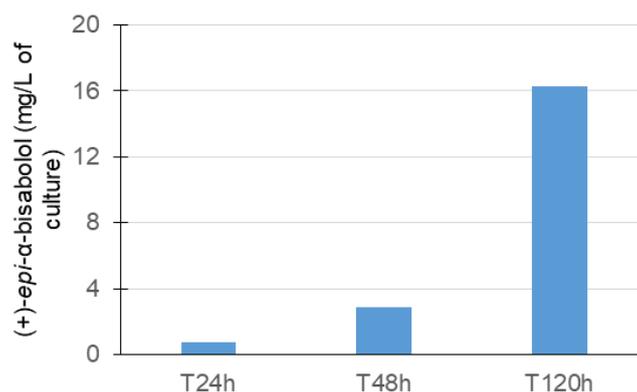


Figure 33: Time course of (+)-epi-α-bisabolol production in the YBIS\_08 strain according to GC-FID quantitation. YBIS\_08 possesses a single copy of BBS along with the integration of *Arabidopsis thaliana* reductase 1 integrated in EPY300 chassis. This strain will be used subsequently in section G.2.d. In term of BBS regulation, the expression is similar to YBIS\_01 with the use of GAL10 promoter in both cases. No dodecane overlay was included in this time course, quantitation was performed after extraction on whole cells and culture broth.

The addition of a second copy of BBS increased (+)-epi-α-bisabolol production (Figure 16). We then questioned if the addition of a second copy of the BBS gene should follow the same regulation of gene expression (galactose inducible). Indeed, having a preexisting pool of the BBS enzyme expressed in yeast might be favorable to cope with the excess of FPP generated in the chassis strain. We thus built the strain YBIS\_07 starting from the parental strain YBIS\_01 by integrating the plasmid pMRI34S + PDC1 BBS. In YBIS\_07, the second copy of BBS is integrated in the genome under a strong constitutive promoter (PDC1) instead of being regulated through galactose induction like in YBIS\_06. The sesquiterpene production of YBIS\_06 and YBIS\_07 was then compared (Figure 34). (+)-epi-α-bisabolol accumulation is enhanced by 38% in YBIS\_07 compared to the YBIS\_06 strain. This points that having a dual regulation of the BBS expression is favorable in our case. To monitor more precisely the level of BBS expressed, further constructions with tagged enzyme and western blotting could be applied. Fusion of a fluorescent protein was also shown to be efficient to follow protein expression as evidenced in the production of nerolidol<sup>173</sup>. In addition, the regulation of ERG9 is the crossroad of the FPP consumption to sterol pathway. This node could be regulated using a control at the protein level, which may be more subtle than the regulation at promoter level we used with MET3 regulation<sup>173</sup>. Another perspective could also be the use of an ERG20-BBS fusion protein to channel the FPP to the heterologous pathway as this proved to be adequate for other sesquiterpenes<sup>165,278</sup>.

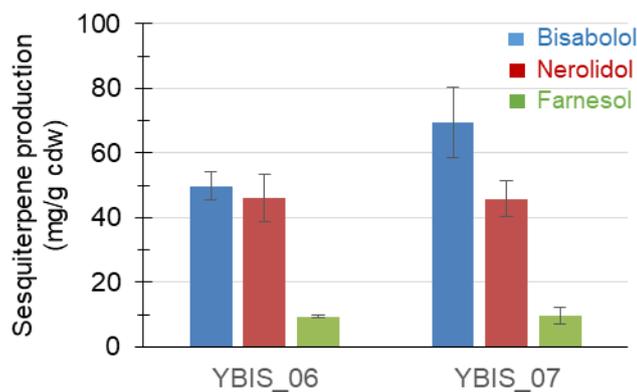


Figure 34: Sesquiterpene distribution in YBIS\_01, YBIS\_06 strains (with respectively the second BBS copy either under GAL10 promoter or PDC1). Cultures were performed in triplicate in SD medium with a dodecane overlay spiked with nootkatone as internal standard.

### c) Supplementary inputs of enzymatic assays

The use of enzymatic assays evidenced that (+)-*epi*- $\alpha$ -bisabolol is a substrate of human and mammal CYPs, yet, hernandulcin may be recognized too. To this end, we assessed *in vitro* the activities of CYP1A1, CYP2B6, CYP2B11, CYP2C9 and CYP2D6 with ( $\pm$ )-hernandulcin. While CYP1A1 and CYP2D6 did not exhibit any oxidation activity with ( $\pm$ )-hernandulcin, CYP2B6, CYP2B11, CYP2C9 and CYP2C18 were able to form new oxidized molecules (Figure 35 for UV detection, and appendix Figure 68 for the MS). While CYP2B6, CYP2C9 and CYP2C18 yielded two major products (but with different ratio), CYP2B11 had an even more relaxed oxidation pattern (at least four different products formed). The MS signal observed with these new molecules is consistent with monohydroxylated products with a  $m/z$  corresponding to  $(M+H^+-H_2O) = 235.3$ . Further characterization could be envisioned via purification of the various new metabolites and structure determination using NMR. This could give valuable information about the position of hydroxylation, and notably it would be of interest to know if the hydroxylation regioselectivity is the same for ( $\pm$ )-hernandulcin as it is with (+)-*epi*- $\alpha$ -bisabolol and CYP2C9 and CYP2C18. If the position of hydroxylation is the same, this would imply that building a synthetic pathway to produce hydroxylated hernandulcin analogs would still be conceivable when the enzyme(s) producing hernandulcin will be discovered. In parallel, the properties of hernandulcin analogs could be established to consider if they have a sweetener profile of interest. With CYP2B6 and CYP2B11, the multiplicity of products formed with ( $\pm$ )-hernandulcin and/or (+)-*epi*- $\alpha$ -bisabolol would probably limit this opportunity. But, CYP2C9 and CYP2C18 have narrower specificity making this strategy more feasible.

- Microsomal control at 60 minutes
- Microsomal fraction + CYP at 5 minutes
- Microsomal fraction + CYP at 60 minutes

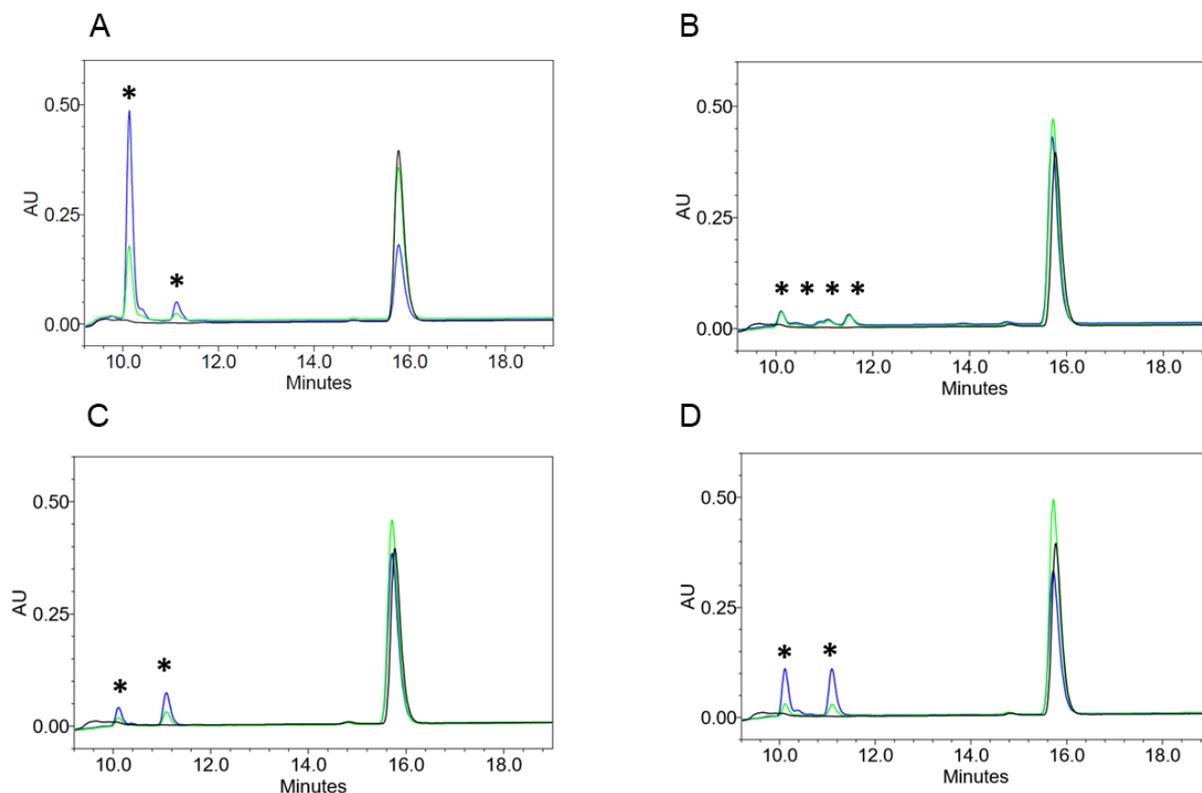


Figure 35: Enzymatic assays showing ( $\pm$ )-hernandulcin oxidizing activities with (A) CYP2B6 (B) CYP2B11 (C) CYP2C9 (D) CYP2C18, the results present the UV chromatograms at 238nm.

\* Highlights the oxidized products

d) Further diversification of  $\alpha$ -bisabolol derivatives generated in *S. cerevisiae*

#### 1) Using plant CYPs involved in sesquiterpene metabolism

In our search of enzymes able of hydroxylation on (+)-*epi*- $\alpha$ -bisabolol, we first considered xenobiotic metabolism due to the novelty of this strategy and the possibility to access new compounds. Nonetheless, some plants CYPs devoted to sesquiterpene biosynthetic pathways were previously characterized, some of them being weakly promiscuous<sup>9,101,105</sup>. Thus, we investigated the use of plant CYPs for the diversification of the (+)-*epi*- $\alpha$ -bisabolol precursor. To this end, we characterized a subset of CYPs comprising i. an in-house library of several CYPs not described for sesquiterpene activities but neither tested on such molecules, ii. some active CYPs active on sesquiterpenes gathered through the contact of several research groups implicated in

enzyme discovery, iii. a few CYPs of interest that were added due to their capability of oxidizing terpenes into keto or aldehydes derivatives, namely CYP71BE5, CYP71D495, CYP706M1, and CYP726A20 that were obtained through DNA synthesis. The Table 7 summarizes the CYPs library we employed. The most promising enzyme of this collection is CYP76F39v1, as Diaz-Chavez et al. described its promiscuous activity towards  $\alpha$ -bisabolol (even though the stereoisomer of  $\alpha$ -bisabolol was not specified in the article)<sup>9</sup>. The CYP76F39v1 activity with  $\alpha$ -bisabolol measured from *in vitro* assay using several sesquiterpenes was found to be 9.4% of the one using  $\beta$ -santalene.

Table 7 Plant CYPs investigated for (+)-*epi*- $\alpha$ -bisabolol activities

CYP name	Biosynthetic pathway	Related organism	Reference	CYP acquisition
CYP73A1	Phenylpropanoid	<i>Helianthus tuberosus</i>	288	In-house library
CYP73A5	Phenylpropanoid	<i>Arabidopsis thaliana</i>	289	In-house library
CYP94A1	Fatty acids	<i>Vicia sativa</i>	290	In-house library
CYP94A2	Fatty acids	<i>Vicia sativa</i>	291	In-house library
CYP94A3	Fatty acids	<i>Vicia sativa</i>	Unpublished isoform	In-house library
CYP94A4	Fatty acids	<i>Vicia sativa</i>	Unpublished isoform	In-house library
CYP94A5	Fatty acids	<i>Vicia sativa</i>	Unpublished isoform	In-house library
CYP72A68v2	Triterpene	<i>Medicago truncatula</i>	237	In-house library
CYP76F39v1	Sesquiterpene	<i>Santalum album</i>	9	Obtained from MTA
CYP76F39v2	Sesquiterpene	<i>Santalum album</i>	9	Obtained from MTA
CYP76F40	Sesquiterpene	<i>Santalum album</i>	9	Obtained from MTA

<b>CYP76F342</b>	Sesquiterpene	<i>Santalum album</i>	9	Obtained from MTA
<b>CYP76F343</b>	Sesquiterpene	<i>Santalum album</i>	9	Obtained from MTA
<b>CYP71AV8</b>	Sesquiterpene	<i>Cichorium intybus</i>	13	Obtained from MTA
<b>CYP71D55</b>	Sesquiterpene	<i>Hyoscyamus muticus</i>	101	Obtained from MTA
<b>CYP71D55 V482I/A484I</b>	Sesquiterpene	<i>Hyoscyamus muticus</i>	101	Obtained from MTA
<b>CYP71BE5</b>	Sesquiterpene	<i>Vitis vinifera</i>	105	Gene synthesis
<b>CYP71D495</b>	Sesquiterpene	<i>Jatropha curcas</i>	292	Gene synthesis
<b>CYP706M1</b>	Sesquiterpene	<i>Callitropsis nootkatensis</i>	108	Gene synthesis
<b>CYP726A20</b>	Sesquiterpene	<i>Jatropha curcas</i>	293	Gene synthesis

For the test of the plant CYPs, an additional strain expressing an adequate CPR was built. *Arabidopsis thaliana* NADPH-cytochrome P450 reductase 1 (AtCPR1, UniprotKB number Q9SB48), was chosen as it accommodates well with CYPs of multiple origins<sup>103,108,289</sup>. After integration of the plasmid pMRI31A + BBS in EPY300 leading to YBIS\_08 (see appendix Table 19 for genotype), the (+)-*epi*- $\alpha$ -bisabolol production was verified and found to be similar to the other chassis strains YBIS\_01 / YBIS\_02 described previously.

After transforming the plant CYPs library and analysis of the metabolites, the strains YBIS\_08 + CYP76F39v2, YBIS\_08 + CYP76F40, YBIS\_08 + CYP76F342, YBIS\_08 + CYP76F343 contained supplementary molecules having a signal in MS consistent with monohydroxylated or epoxy compounds generated from (+)-*epi*- $\alpha$ -bisabolol or trans,trans-farnesol (molecular weight of 238.3 g/mol and the detected *m/z* corresponding to  $M+H^+-2H_2O = 203.3$ , Figure 36). We did not detect such activities with the other plant CYPs in our chassis strain. With our *in vivo* approach, we did not observe the promiscuous activity of CYP76F39v1 towards (+)-*epi*- $\alpha$ -bisabolol. This may be due too a possible weak expression of this CYP in our strain, or, alternatively, CYP76F39v1 is not active with the (+)-*epi*- $\alpha$  isomer (but maybe with other ones as the isomer used for *in vitro* was not mentioned by Diaz-Chavez et al). Interestingly, from the limited library we screened, two products are recovered, one produced by CYP76F39v2, CYP76F40, CYP76F342 in different amounts between the strains while CYP76F43 is able to yield a different one. The obtained products from the plant CYPs characterization now require a deeper analysis to accurately attribute the activities towards the relevant substrate i.e (+)-*epi*-

$\alpha$ -bisabolol or trans,trans-farnesol, using both enzymatic assays and purification of the molecules to carry out NMR structure determination as it was conducted with mammal CYPs. This structure determination would then conclude whether we reached molecules that are different from the ones produced by mammalian CYPs.

We showed that exploring the promiscuity of animal and plant CYPs enabled the production of new to nature molecules originating from (+)-*epi*- $\alpha$ -bisabolol. A last reservoir of CYPs lies in the bacterial world. In this bacterial diversity, some versatile catalysts were studied and demonstrated their potential to oxidize sesquiterpenes or monoterpenes<sup>294,295</sup>. Conversely, enzyme engineering efficiently broadened the substrate range of some bacterial CYPs towards sesquiterpenes<sup>296</sup>. Recently, the promiscuity of the bacterial CYP260B1 was used in combination with 2 sesquiterpene cyclases to diversify the obtained sesquiterpenoids in *E. coli*<sup>297</sup>. Again, this highlights the vast potential of CYPs and the fact that building larger combinatorial studies could provide new products with relevant strain design and analytical tools. Hence, an even larger repertoire of (+)-*epi*- $\alpha$ -bisabolol oxidized products may be attained. Eventually, hernandulcin and / or its hydroxylated precursor may be synthesized by some of the enzymes within this tremendous collection of mammalian, plant or bacterial enzymes.

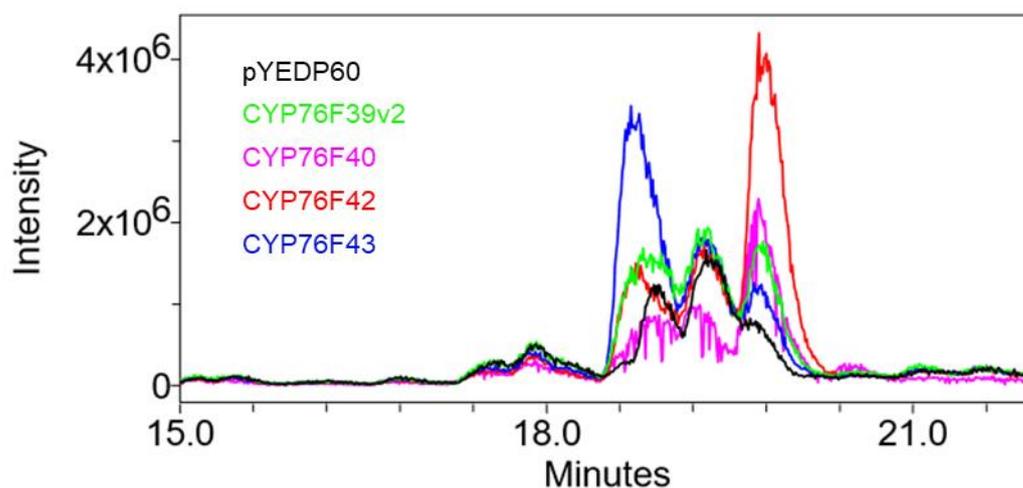


Figure 36: CYPs oxidation activities in YBIS\_08 strains. The overlaid chromatograms depict MS signal at  $m/z = 203.3^*$

## 2) ADHs

In terpene biosynthesis, the introduction of keto groups is accomplished either by CYPs themselves<sup>105,108</sup>, or by a CYP that synthesizes an hydroxylated compound which is further oxidized by an alcohol dehydrogenase (ADH)<sup>7</sup>. As the identified mammalian CYPs stopped after a single hydroxylation step with (+)-

*epi- $\alpha$ -bisabolol*, we sought to add an ADH enzyme in order to obtain keto derivatives. Moreover, Wong et al. have shown that plant ADHs can sometimes be promiscuous and accept molecules similar to their native substrates<sup>298</sup>. In their case, *Artemisia annua* ADH could replace and even outperform the native *Valeriana officinalis* ADH for valerenic acid production in a yeast chassis. Thus, we retained a small subset of enzymes to assay their potential with (+)-*epi- $\alpha$ -bisabolol*. We restricted to biochemically characterized ADH that can accept monoterpenes, sesquiterpenes or diterpenes that are cyclic. The selection we accomplished is shown in Table 8.

Table 8 : Selected ADHs for the oxidation of (+)-*epi- $\alpha$ -bisabolol*

ADH name	Related organism	Reference
AaADH	<i>Artemisia annua</i>	22
EIADH	<i>Euphorbia lathyris</i>	293
JcADH	<i>Jatropha curcas</i>	293
LiADH	<i>Lavandula intermedia</i>	299
MpADH	<i>Mentha piperita</i>	251
VoADH	<i>Valeriana officinalis</i>	298
ZzADH	<i>Zingiber zerumbet</i>	7

After cloning on a compatible plasmid with the one expressing mammalian CYPs, cotransformation of the library composed of i. the 7 ADHs and void plasmid (pFPP13) and ii. CYP2B6, CYP2B11, CYP2C9, CYP2D6 and void vector (pYEDP60) covering the spectrum of the different products obtained from (+)-*epi- $\alpha$ -bisabolol* was carried out in the strain YBIS\_01. Four out 39 strains exhibited signals that could suggest a low formation of products (Table 9). However, among these strains, YBIS\_01 + CYP2B11 + EIADH displayed the formation of two products at 238 nm (Figure 37) but the MS signal could not correlate to expected adducts for keto-(+)-*epi- $\alpha$ -bisabolol* products. Regarding the putative product from YBIS\_01 + CYP2B11 + AaADH, YBIS\_01 + CYP2B11 + LiADH, YBIS\_01 + CYP2B11 + VoADH (Figure 38), a signal at  $m/z$  equivalent to  $M+H^+-H_2O = 219.3$  was observed but with a low intensity and no UV spectra corresponding to keto group was detected at this retention time<sup>300</sup>. Due to a lack of time, a deeper survey has to be performed on these preliminary results. First, as the strain YBIS\_01 + CYP2B11 could produce both hydroxy-(+)-*epi- $\alpha$ -bisabolol* and hydroxy-trans,trans-farnesol derivatives, an *in vitro* assay on purified molecules should confirm which molecules are substrates. In addition, we looked whether the conversion of farnesol to farnesal occurred in our strains, and no such signal was detected with YBIS\_01 + pYEDP60 + ADHs. Indeed, in these strains hydroxy-(+)-*epi- $\alpha$ -bisabolol* and trans,trans-farnesol are not converted to their hydroxy derivatives while (+)-*epi- $\alpha$ -bisabolol* cannot be converted to a keto

group as it is a tertiary alcohol. No signal consistent with farnesal was noted. Finally, more ADHs might be added to the selection such as the one from *Bacillus megaterium* which was recently found to convert the sesquiterpene nootkatol to (+)-nootkatone <sup>301</sup>.

Table 9 Preliminary results of the ADH yeast in vivo assay

CYP ADH	pYEPD60	CYP2B6	CYP2B11	CYP2C9	CYP2D6
pFPP13					
AaADH			+/-		
EIADH			+/-		
JcADH					
LiADH	∅		+/-		
MpADH					
VoADH			+/-		
ZzADH					

\*∅: not assayed

+/-: corresponds to trace/weak activity detected only in UV or only in MS signal

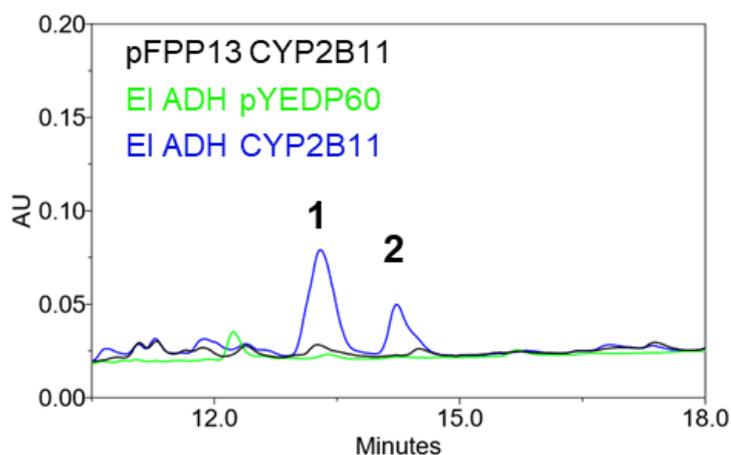


Figure 37: Chromatograms at 238 nm showing the new products observed in YBIS\_01 + CYP2B11 + EIADH

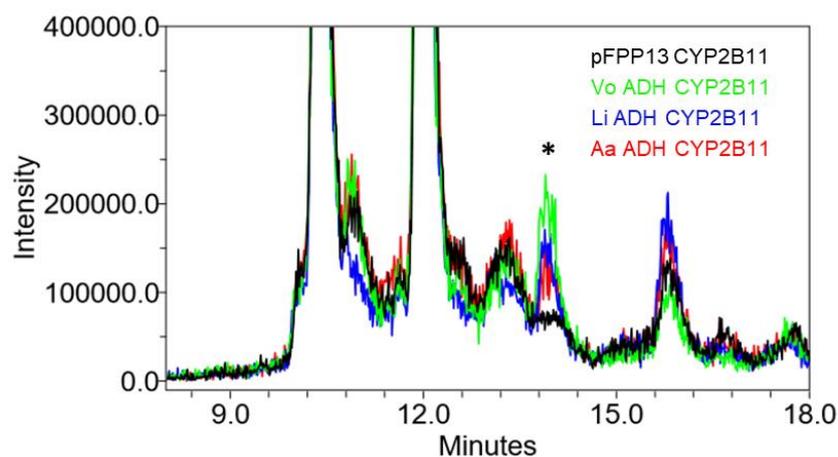


Figure 38: Chromatograms at  $m/z = 219.3$  showing the new products observed in YBIS\_01 + CYP2B11 + AaADH, YBIS\_01 + CYP2B11 + LiADH, YBIS\_01 + CYP2B11 + VoADH

### 3) The use of BBS generating other $\alpha$ -bisabolol isomers

The occurrence of other isomers of  $\alpha$ -bisabolol are described in nature and the enzymes responsible for the synthesis of each specific isomer was recently unveiled (Figure 39).

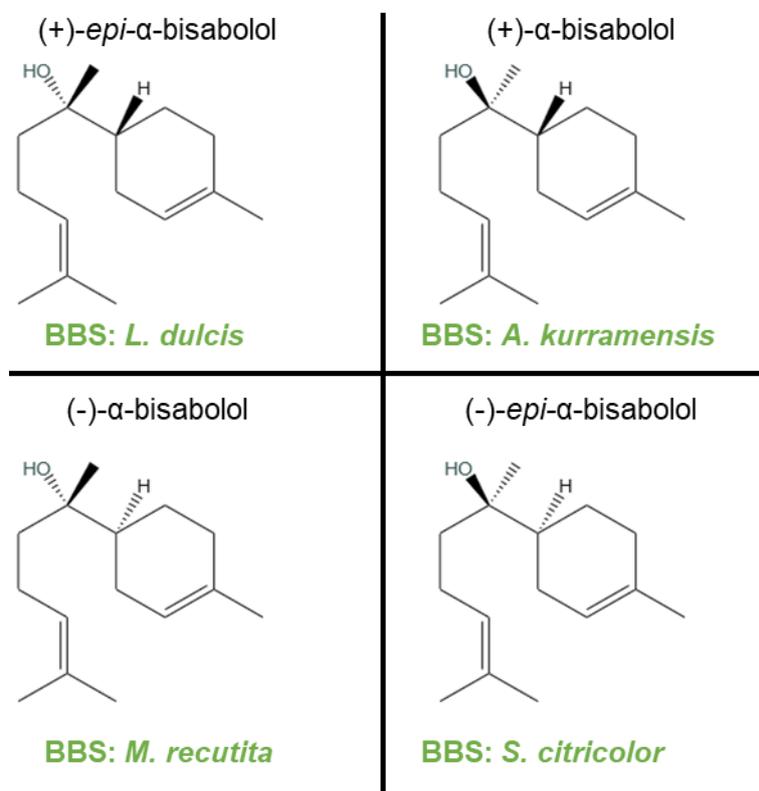


Figure 39: The bisabolol isomers and the corresponding BBS that were biochemically characterized <sup>82,302,303</sup>

We integrated the 3 other BBS genes from *A. kurramensis*, *M. recutita* and *S. citricolor*, into EPY300, leading to YBIS\_09, YBIS\_10, YBIS\_11 (appendix Table 19). The production of the different isomers was then assayed for each one (Figure 40). Even though all BBS were functional in yeast, the  $\alpha$ -bisabolol production was notably different. As the integration locus and the promoters are identical between the strains, the reason for the distinctive patterns originates from inefficient protein production or variability in kinetic parameters of the BBS. *Lippia dulcis* BBS was by far the most efficient enzyme. Improvements of the other BBS productivity could rely on testing several codon bias adjustments, adding a protein fusion to favor soluble expression<sup>25</sup>, or increasing copy number as we did with the strains YBIS\_06/YBIS\_07. Once the production of the 3 additional  $\alpha$ -bisabolol isomers will be adjusted, a perspective would be to transform the strains with the mammalian CYPs. It would allow to understand if the identified CYP are sensitive to different stereoisomerism of  $\alpha$ -bisabolol. Do the CYPs are affected similarly? Are their spectra of products differing between  $\alpha$ -bisabolol stereoisomers?

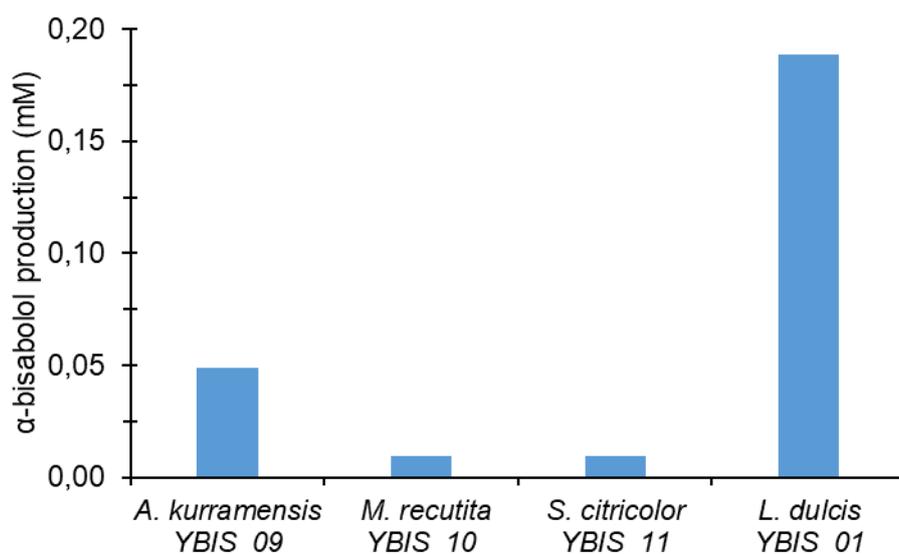


Figure 40:  $\alpha$ -bisabolol production with the different BBS integrated in EPY300 background

### 3. Conclusion

In summary, we acquired some information about the finer control of *S. cerevisiae* showing that the regulation can be improved with a better timing of the  $\alpha$ -bisabolol production. The screening of plant CYPs efficiently demonstrated the versatility of our chassis strain as potentially new molecules were generated from our *in vivo* experiment. Ultimately, by adding a supplementary component like an ADH enzyme, further progress could be made to complete a full biosynthetic pathway towards keto derivatives of (+)-*epi*- $\alpha$ -bisabolol. Finally, the versatility of the yeast chassis was enlarged to the various  $\alpha$ -bisabolol isomers and could expand the scope of this study beyond the synthesis of hernandulcin analogs and (+)-*epi*- $\alpha$ -bisabolol derivatives.



## Chapter III.

### Characterization of CYP2B6 / CYP2B11 chimeras with (-)- $\alpha$ -bisabolol and trans,trans-farnesol

## **Abstract**

Ordered chimeras constructed from CYP2B6 and CYP2B11 genes were previously studied using cyclophosphamide and 7-ethoxy-4-trifluoromethyl-coumarin in order to delineate sequences from the two CYPs that are responsible for substrate specificity and affinity <sup>260</sup>. The original design was based on the sequential swap of CYP2B6 structural elements into a CYP2B11 core. The construction of this set of engineered enzymes showed that 11 out of 15 of the generated chimeras were functional. We identified previously (chapter II. C) that both CYP2B6 and CYP2B11 oxidized trans,trans-farnesol and (-)- $\alpha$ -bisabolol and that their product profiles were different. We then thought to assess this set of chimeras and evaluate whether the exchange of the segment from CYP2B6 into CYP2B11 chassis can switch the product pattern formation. In other words, can we explain the stereospecificity using the exchange of single element of the secondary structure of CYP2B6?

We focused on the eleven active chimeras to investigate if punctual structural elements could correlate with a swap towards CYP2B6 product specificity. Metabolic profiles on chimeras and wild-type parental enzymes expressed in yeast were determined using (-)- $\alpha$ -bisabolol and trans,trans-farnesol as sesquiterpene substrates. Our results suggest that multiple structural elements are involved in the control of regioselectivity of sesquiterpene metabolism.

## A. Introduction

CYPs are heme-thiolate enzymes that are found across all kingdoms of life and represent more than 300,000 protein sequences<sup>93</sup>. Their catalyzed reaction is generally a mono-oxygenation through cleavage of a dioxygen molecule, introducing an oxygen atom in the substrate (so-called functionalization reaction) and reducing the other to water<sup>304</sup>. In humans, the oxidative drug metabolism depends predominantly on microsomal CYP enzymes that are known to present a high level of promiscuity in their specificity, tolerating substrates with extremely different sizes and chemical nature<sup>305</sup>.

The numerous crystal structures available evidenced that CYPs have similar three-dimensional structures<sup>306</sup>. However, this similarity hardly accounts for their relative substrate specificity. The explanation of how such similar 3D structures yield so different and broad substrate specificities could lead to both clinical and biotechnological applications<sup>307,308</sup>. Mammalian CYPs of the CYP2B subfamily have been extensively studied both biochemically and structurally, the substrate specificity of CYP2B enzymes differs between isoforms from different mammals<sup>309</sup>. To date, 13 crystal structures of CYP2B6 have been described and 35 X-ray structures have been solved if we extend to the CYP2B enzyme subfamily (PDB, <https://www.rcsb.org/>). Both ligand-free structures and structure complexed with a series of ligands of various sizes were obtained. The crystal structure of CYP2B6 in complex with  $\alpha$ -pinene was recently deciphered<sup>310</sup>. Later, structures of CYP2B6 in complex with two other bicyclic monoterpenes, sabinene and 3-carene, were obtained by X-ray crystallography and the thermodynamics of their binding was assessed by isothermal titration calorimetry<sup>273</sup>. These crystal structures evidenced several architectures of the active site and some secondary structural elements, leading to the hypothesis that there may be critical motions of these elements in order to favor the adjustment of CYP2B6 with different ligands. Moreover, multiple channels could be involved in the access to the catalytic site of CYPs adding more complexity to the binding of molecules in the active site<sup>272,311</sup>.

Regarding the terpene oxidation by human CYPs, this implication was already evidenced about twenty years ago in the drug metabolism of artemisinin<sup>312</sup>. Additional studies on the drug metabolism of terpenes also pointed out the redundancy of the CYPs and the overlapping products generated, as in the example of thujone for which CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4 are involved<sup>313</sup>. While the role in human drug metabolism of CYP2B6 is recognized to participate in 4% of all drugs; on the contrary, only scarce information is available on the canine CYP2B11's involvement in the metabolism of active substances<sup>94</sup>. Notably, the two enzymes differ in term of substrate recognition like in the case of the prodrug cyclophosphamide, which requires drug activation by human CYP2B6 enzyme to exert its anti-tumor function. Cyclophosphamide exhibits a lower affinity for CYP2B6 than for CYP2B11<sup>314,315</sup>. We have previously described

activities of CYP2B6 and CYP2B11 towards (-)- $\alpha$ -bisabolol and trans,trans-farnesol in Chapter II and their specificity noticeably differed in term of the variety of products they generated. In that perspective, studying such similar enzymes could give valuable information about structure/activity relationship, by deciphering the implication of differences existing between these two CYPs.

Therefore, we compared the two wild type enzymes CYP2B6 and CYP2B11 to the chimeras possessing swapped structural elements. CYP2B6 and CYP2B11 share 78% of identity after protein alignment making them close homologs. This high homology makes them amenable for chimera studies as one could expect that switching structural elements between homologs may lead to a conserved functionality of the new chimeras. The protein sequence of CYP2B6 was previously split into 15 short structural segments of 17 to 44 amino acid length and replaced into CYP2B11 sequence. The scheme of this chimera design spanning the whole CYP sequence is represented in Figure 41.

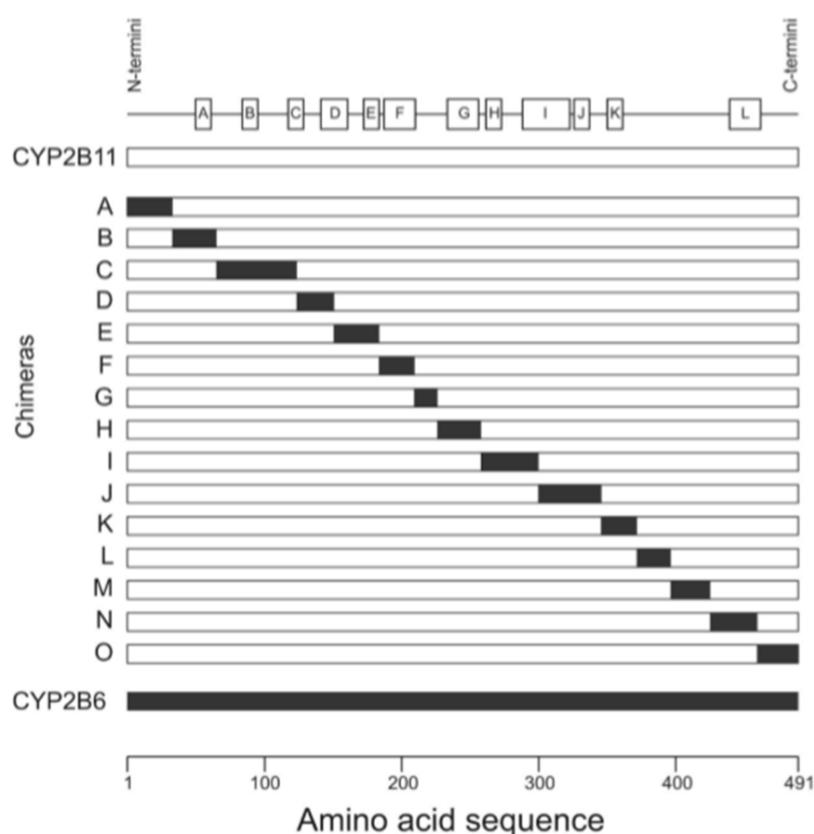


Figure 41: Comparison of parental and chimeric CYP2B6-CYP2B11 primary structures. Solid bars represent human CYP2B6 sequence elements, gray-filled bars represent canine CYP2B11 sequence elements. This figure was reproduced from Lautier et al <sup>260</sup>.

The position of CYP2B helices is shown along the sequence. For sake of clarity, only main helices (from A to L) have been indicated.

Moreover, as evidenced by Li et al. in a pioneering study focusing on the generation of chimeras from bacterial CYPs, the screening of the novel enzymes not only gave rise to thermostable enzymes but extended dramatically the ability of the new enzymes. In the diversity of the library new specificities appeared and molecules that could not be recognized by the parental enzymes became substrates for some of the chimeric enzymes<sup>316</sup>. Similarly, the use of the chimeric enzymes from CYP2B6 and CYP2B11 could give rise to new specificity not existing in any parental enzymes and broaden the set of hydroxy- $\alpha$ -bisabolol or hydroxy-trans,trans-farnesol derivatives.

## B. Materials and methods

Microsomal fractions were prepared as described in Lautier et al<sup>260</sup>. All chimeras from A to O (Figure 41) were assayed except chimeras C, D, I and L that were previously described as inactive<sup>260</sup>. A control reaction was also included with microsomal fraction of the W(R) strain transformed with the void plasmid pYEDP60 in order to obtain the background signal due to endogenous yeast components.

Chemicals, enzymatic assays conditions of experiments and LC/MS detection were identical to those described in Chapter II.

## C. Results and discussion

*In vitro* oxidation activities of CYP2B6 and CYP2B11 using (-)- $\alpha$ -bisabolol and trans,trans-farnesol were conducted and the products were separated by HPLC-MS (Figure 42A and B). No peak at  $m/z=203.3$ , corresponding to oxidized forms of (-)- $\alpha$ -bisabolol, were detected when microsomal fractions devoid of CYP were used. However, with trans,trans-farnesol a low interfering signal trans,trans-farnesol around 13 minutes (Figure 42A and B) was detected. The signal at  $m/z (M+H^+-2H_2O) = 203.3$  corresponds to the dehydrated rearrangement of the various hydroxy derivatives, as already seen in Chapter II. No derivatives with different oxidation products were observed (for instance products with a keto group or two hydroxy introduced). With (-)- $\alpha$ -bisabolol, CYP2B6 formed two products while CYP2B11 produced 6 different molecules with two of them common with CYP2B6. With trans,trans-farnesol, CYP2B6 formed a metabolite while CYP2B11 produced two main metabolites. Clearly, CYP2B6 and CYP2B11 do not perform the same number of oxidations on a cyclic sesquiterpene ((-)- $\alpha$ -bisabolol) or an acyclic sesquiterpene (trans,trans-farnesol) and the molecular determinants involved in substrate positioning within the catalytic site differ between the two enzymes.

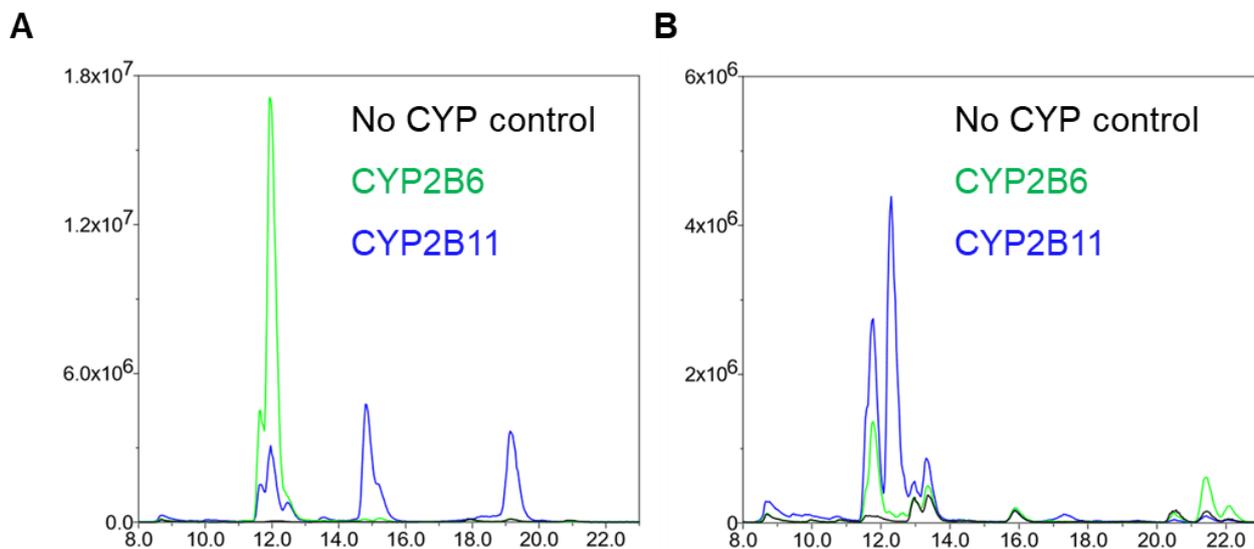


Figure 42: CYP2B6 and CYP2B11 *in vitro* activities towards (A) (-)- $\alpha$ -bisabolol (B) trans,trans-farnesol. The chromatograms depict the MS signal at  $m/z = 203.3$ .

The enzymatic activities of these chimeras were assessed using both (-)- $\alpha$ -bisabolol and trans,trans-farnesol (Figure 43, Figure 44). With (-)- $\alpha$ -bisabolol, all chimeras displayed a pattern of products that is similar to the CYP2B11 parental enzyme except for chimera G that seems to have a higher accumulation of the product P4 while P5 seems to be absent compared to CYP2B11. However, the change remains minor. Interestingly, chimera G was previously shown to affect also cyclophosphamide affinity<sup>260</sup>. Thus, this portion of CYP2B6 is critical for substrate discrimination. Further kinetic analysis may indicate whether affinity with (-)- $\alpha$ -bisabolol is affected in the chimera G compared to CYP2B6 and CYP2B11. Regarding trans,trans-farnesol (Figure 44), the chimeras do not strongly shifted towards a CYP2B6 pattern. All of them displayed a pattern of products that is similar to the parental CYP2B11 with the notable exceptions of chimeras H and M that present a potential new peak at ~10 minutes. Further experiments with time course of the reaction will have to be conducted to corroborate the formation of the potential new product. Contrarily to what was observed with (-)- $\alpha$ -bisabolol, the chimera G was not affected with trans,trans-farnesol. As chimera G is supposed to be involved in an access channel<sup>260</sup>, this could explain why the two substrates are affected differentially.

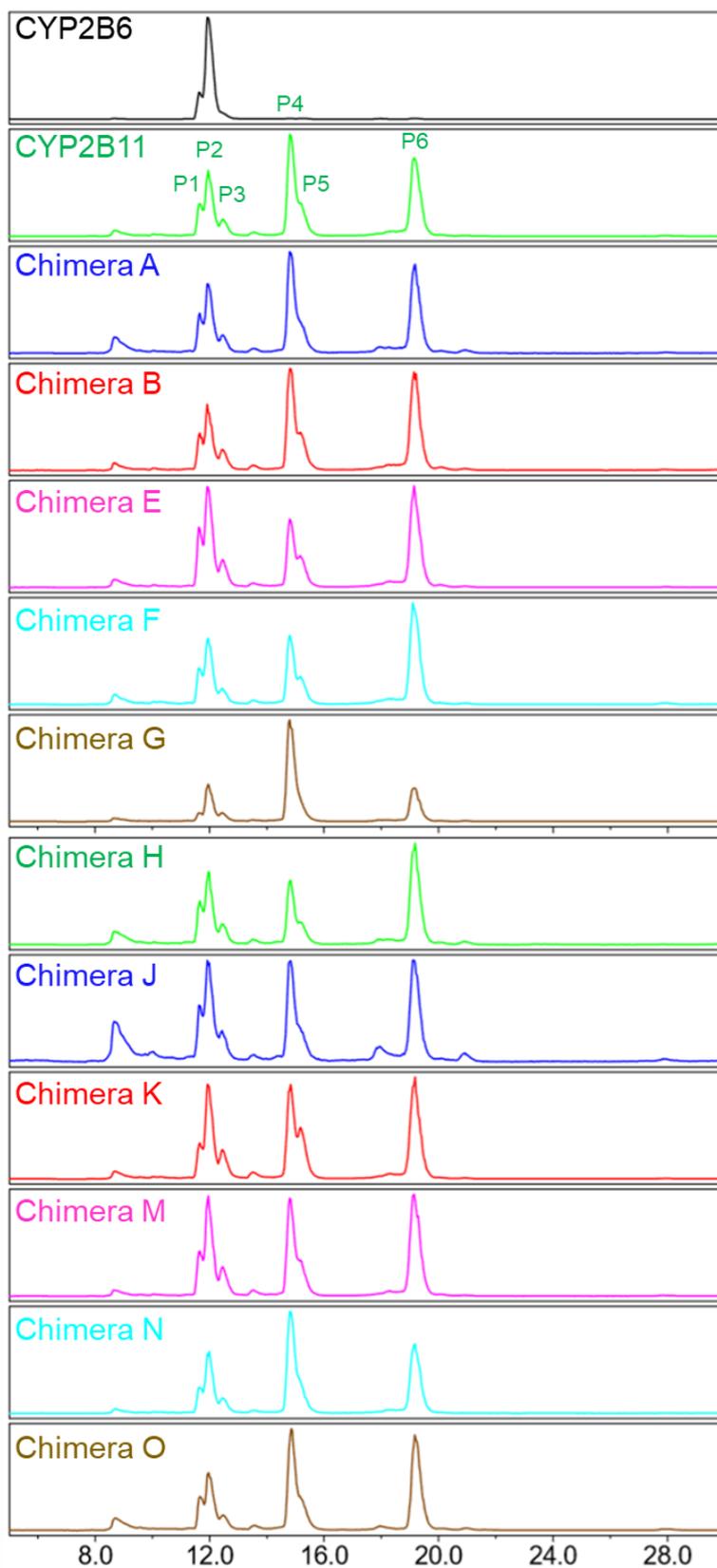


Figure 43: The behavior of the chimeras using (-)- $\alpha$ -bisabolol as substrate  
 The chromatograms depict the MS signal at  $m/z = 203.3$ . Signal was scaled to the highest peak in order to ease visual comparison of the product diversity generated by each enzyme. The products from CYP2B11 were numbered from P1 to P6 to ease comparison between chimeras.

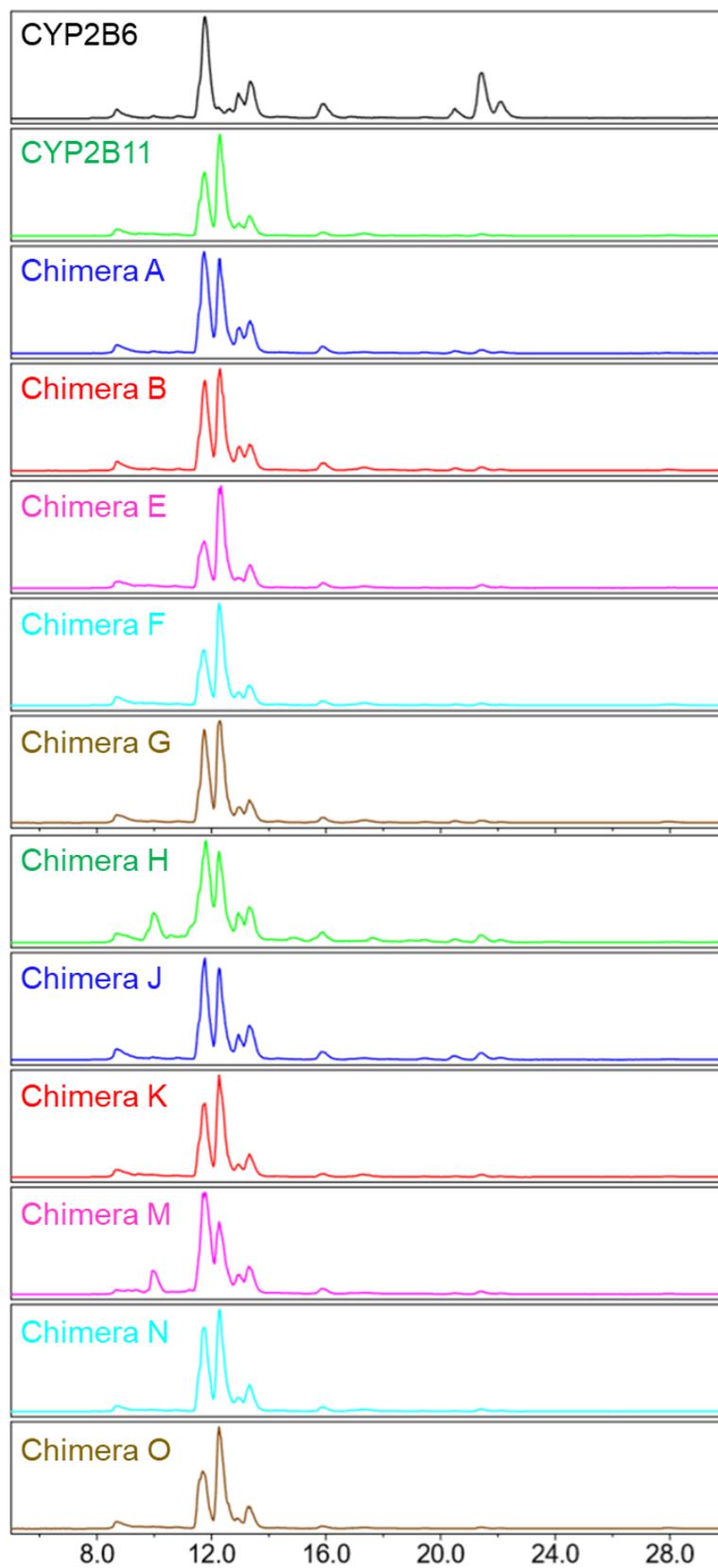


Figure 44: The behavior of the chimeras using *trans,trans*-farnesol as substrate  
 The chromatograms depict the MS signal at  $m/z = 203.3$ . Signal was scaled to the highest peak in order to ease visual comparison of the product diversity generated by each enzyme.

Thus, all chimeras (except for chimera G) exhibit a parental pattern corresponding to CYP2B11. Therefore, it seems likely that multiple concurring determinants are involved in sesquiterpene recognition. Alternatively, these structural elements may also be located in chimeras which could not be properly evaluated due to the absence of detectable hydroxylation activities (chimeras C, D, I and L). Supplementary assistance of computational tools and docking experiments of substrate and enzymes might eventually give support in the understanding of sesquiterpene recognition and the distinctive properties of these enzymes. To study further the structure/activity relationships of CYP2B6-CYP2B11 with trans,trans-farnesol and (-)- $\alpha$ -bisabolol, one could also use the reciprocal approach that we used by exchanging CYP2B11 elements into CYP2B6 chassis to obtain the symmetrical chimeras we did. Additionally, if multiple secondary elements are involved, an alternative solution could be to build chimeras with the exchange of larger segments than the ones employed. Finally, in order to have a more complete framework for chimera design, the approach of chimera generation could eventually be applied to CYP2C9, CYP2C18 and CYP2D6 that are homologs too and that generate a different range of products from (-)- $\alpha$ -bisabolol. However, increasing complexity may not ease the interpretation of the results.

## **Acknowledgments**

Thomas Lautier for the supply of yeast microsomal fractions.

Philippe Urban and Denis Pompon for fruitful discussions about the CYP chimeras.



## Chapter IV.

### Glycosylation of $\alpha$ -bisabolol and hernandulcin

## **Abstract**

Glycosylation is often encountered in multiple biosynthetic pathways and in particular in secondary metabolism such as antibiotics (vancomycin, erythromycin...), sweeteners such as plant steviol glycosides, antioxidants (flavonoids...), etc. Chemistry hardly tackles the challenge of glycosylation due to the difficulty coming from stereoselectivity, regioselectivity and multiple protection / deprotection steps that are often required to obtain the target product<sup>317</sup>. Moreover, glycosylation by biological means allows to modulate the physicochemical properties of secondary metabolites, while avoiding the use of polluting catalysts and / or solvents. However, natural glycosylated secondary metabolites are often produced in low yields. To overcome the obstacle of their low abundance, setting up recombinant production represents a promising alternative. To do so, researchers are tapping the natural diversity of glycosylating enzymes, which is huge but far from being well characterized, although it can open access to a tremendous chemical repertoire of molecules. With the emergence of synthetic biology, the frontiers are pushed even further to shift from existing pathways towards new ones and conceive undescribed new-to-nature routes for glycosylated products, especially secondary metabolites.

$\alpha$ -bisabolol and hernandulcin are two sesquiterpenes of interest for human use but their highly hydrophobic nature may restrict their formulation and bioavailability. In this context, investigating their enzymatic glycosylation could enhance their solubility and avoid possible toxicity in recombinant production while opening routes to molecular diversification. We first selected the UGT93B16, an oat UDP-glucosyltransferase, presumed to glucosylate (-)- $\alpha$ -bisabolol and showed that, indeed, this enzyme converts the two scaffolds in *in vitro* enzymatic assays. We next carried out whole cell catalysis of the reaction using *E. coli* and *S. cerevisiae* as recombinant producers of UGT93B16. Comparison of the two microbial hosts pointed out the better ability of yeast cells to catalyze this sesquiterpene glucosylation. The full synthetic pathway enabling the *in vivo* glucoside synthesis was thus reconstituted yielding a final titer of several mg/L of sesquiterpene glucoside in yeast ((+)-*epi*- $\alpha$ -bisabolol and/or trans,trans farnesol derivatives). In a separate approach, we tested bisabolol glucosylation using human UDP-glucuronosyltransferases involved in the xenobiotic metabolism. *In vitro* enzymatic assays demonstrated a weak activity of UGT1A9, UGT2B4 and UGT2B7 towards  $\alpha$ -bisabolol and hernandulcin. However, their implementation in yeast failed to produce detectable amounts of glucuronide products. In summary, we could prove the feasibility of producing new to nature sesquiterpene glucosides using either enzyme-based assay, bioconversion or biosynthesis in two different hosts that are widely used in biotechnology.

## A. Overview of UDP-glycosyltransferase diversity

In the tremendous diversity of plant natural products, glycosylation is widely distributed and occurs in a broad variety of biosynthetic pathways leading to flavonoids, terpenes, alkaloids, etc. Glycosylation may have different roles, including the aglycon stabilization, the control of the transport to the vacuole for anthocyanins<sup>318</sup>, the homeostasis of some plant hormones such as abscisic acid<sup>319,320</sup>, the protection against endogenous defense molecules such as those encountered in glucosinolate pathway<sup>321</sup>. Interestingly, glycosylation can also serve to protect plants from their own defense molecules, their biological activities being triggered upon glycoside hydrolysis<sup>321</sup> (Figure 45B). In other cases, glycosylation of a given scaffold is required for a complete product activity (Figure 45A).

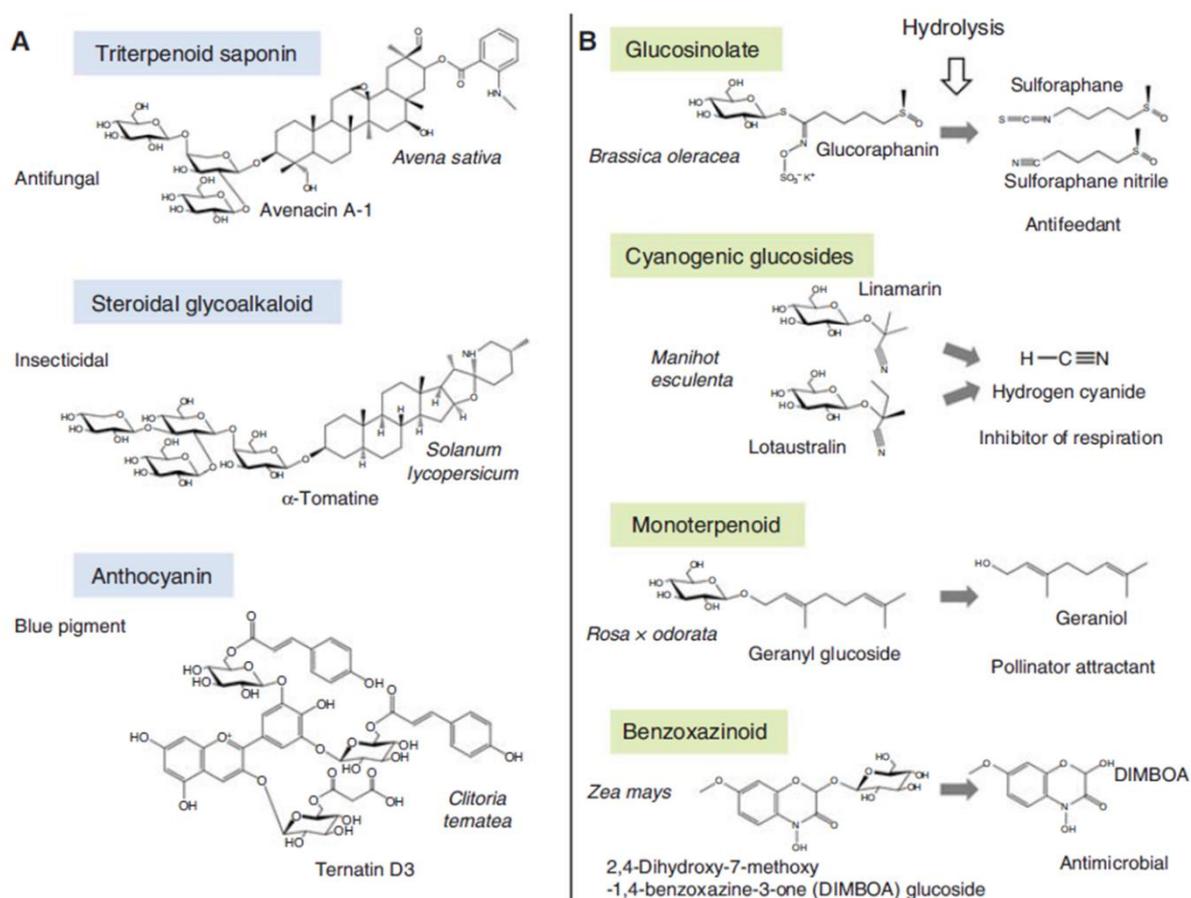


Figure 45: Examples of glycosylated plant-specialized metabolites. Compounds that are active (A) in their fully glycosylated forms or (B) upon hydrolysis are shown. Figure reproduced from Louveau et al.<sup>26</sup>

Enzymes from the UGT superfamily are recognized as the main contributors to the glycosylation of plant natural products<sup>26</sup>. We will briefly present this family in the following chapter.

## 1. Plant UDP-glycosyltransferases (UGTs)

Based on sequence alignments, donor substrate used and mechanistic data, the CAZY database (Carbohydrate-Active enZymes, <http://www.cazy.org/>) classifies plant UDP-glycosyltransferases in the GT1 family. In December 2019, the CAZY database gathers enzymes from all domains and kingdoms of life with more than 20,000 GT1 sequences of which 7,000 are from eukaryotes<sup>322</sup>. To further distinguish the enzymes from this vast superfamily, a complementary nomenclature, based on sequence identity, was proposed by Ross et al. for plant UGTs (Figure 46). In addition, advances in genomic data acquisition in the last decades enabled a thorough analysis of this UGTs superfamily. In 2001, genome analysis of the model organism *A. thaliana* revealed 107 sequences encoding UGTs and their subclassification in 14 distinct clusters within the superfamily<sup>323</sup>. More recently, a much wider analysis over 65 plant genomes extended the numbers of UGT clusters to 18 with additional “divergent groups”<sup>324</sup>. Of that tremendous amount of putative UGTs, only a few hundreds have been functionally characterized highlighting the current bottleneck of functional annotation. The experimental validation of UGT activities and the determination of their roles remain difficult due to the complexity of *in planta* studies and a limited access to their native substrates for *in vitro* biochemical characterization.

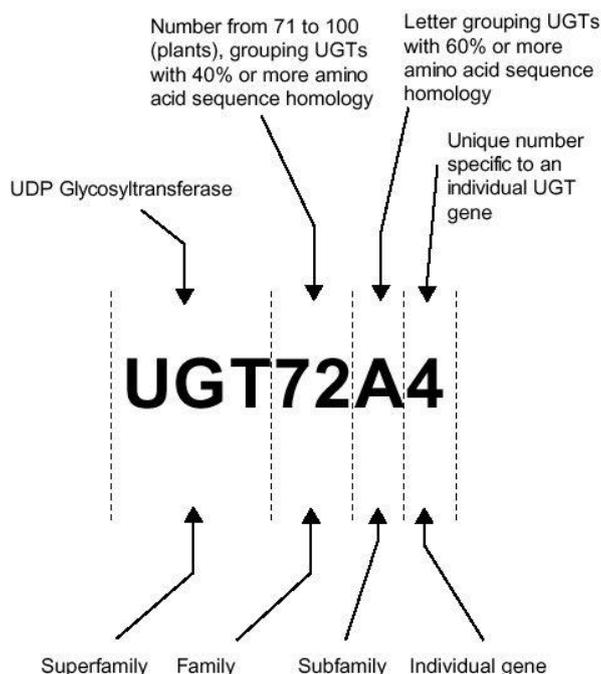


Figure 46: The current UGT superfamily nomenclature system. The diagram illustrates the system used to name plant UDP glycosyltransferases (adapted from Ross et al <sup>323</sup>)

UGTs transfer a sugar moiety from a uridine 5'-diphosphate sugar (donor molecule) to an acceptor molecule. An impressively diverse range of molecules of the plant kingdom can act as acceptors including flavonoids, terpenes, auxin, cytokinin, salicylic acid and many others <sup>324</sup>. Depending on the enzyme, the acceptor specificity can be broad or in contrast extremely stringent <sup>325,326</sup>.

The first X-ray crystal structure of a plant UGT from *Medicago truncatula* was obtained in 2005 and revealed that the protein adopts a GT-B fold structure consisting in two N- and C-terminal domains displaying a Rossmann-type fold <sup>37</sup> (Figure 47). At the interface of the two domains lies a cavity where both the acceptor and the UDP-sugar encounter. In this architecture, the recognition of the UDP-sugar is mainly mediated by a conserved region in the C-ter of UGT protein sequences containing the "PSPG" motif (plant secondary product glycosyltransferase) <sup>327</sup>. In family GT1, all enzymes employ an inverting mechanism, which is proposed to involve a catalytic His residue assisted by an Asp residue of the N-terminal domain (referred to as H17 and D114 in UGT78K6 <sup>328</sup>). The catalytic histidine first deprotonates the acceptor, which in turn exerts a nucleophilic attack at the C1 atom of the sugar donor, forming the glycosylated product while releasing the nucleotide. Meanwhile, the aspartate residue stabilizes the protonated histidine <sup>37</sup>. In the course of the reaction, the configuration of the anomeric carbon is inverted, leading for example from a UDP- $\alpha$ -glucose to a  $\beta$ -glucoside product (Figure 48). With a notable exception of sterol UGTs, these enzymes are generally cytosoluble <sup>323,329</sup>.

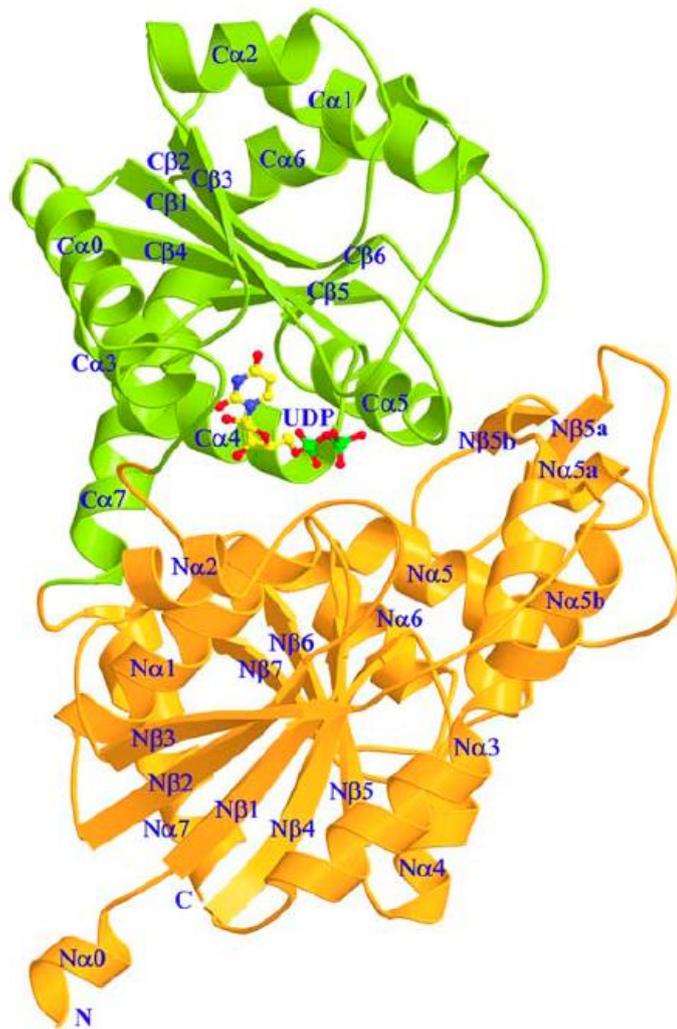


Figure 47: Ribbon Diagram of the Structure of UGT71G1 with Bound UDP (reproduced from Shao et al <sup>37</sup>)  
 The N- and C-terminal domains are shown in orange and green, respectively. The  $\alpha$  helices and  $\beta$  strands of the N- and C-terminal domains are numbered separately. N-terminal domain comprises a central seven-stranded parallel  $\beta$  sheet flanked by eight  $\alpha$  helices on both sides and a small two-stranded  $\beta$  sheet. The C-term contains a six-stranded  $\beta$  sheet flanked by eight  $\alpha$  helices. The UDP molecule is shown as a ball-and-stick model colored by atom type (nitrogen, blue; carbon, yellow; oxygen, red; phosphorus, green).

As previously mentioned, UGTs generally have a rather strict specificity for their individual sugar donor substrate. This does not limit the repertoire of UDP-sugars that can be used by these enzymes. The most commonly transferred sugar is glucose (around 80% characterized UGTs family use UDP-Glc <sup>26</sup>) but alternative donors comprise UDP-glucuronic acid, UDP-galactose, UDP-rhamnose or even C5 sugars like UDP-xylose or UDP-arabinose. The donor substrate specificity is not flagrantly related to the phylogenetic clusters indicating that the donor specificity evolved independently in the different clusters. For instance, this was observed in family UGT93 where a zeatin glucosyltransferase and xylosyltransferase were characterized <sup>330,331</sup> or in family UGT73 where there is a multiplicity of sugar donors including UDP-Ara, UDP-Gal, UDP-Xyl, UDP-Glc used by

UGT73P10<sup>332</sup>, UGT73P2<sup>333</sup>, UGT73F4<sup>334</sup>, UGT73K1, respectively<sup>39</sup>. Notably, only a few mutations can change the preference for any of the sugar nucleotides. Structural and mutagenesis studies also revealed that the determinants of substrate specificity do not strictly cluster in the PSPG motif<sup>26,335,336</sup>.

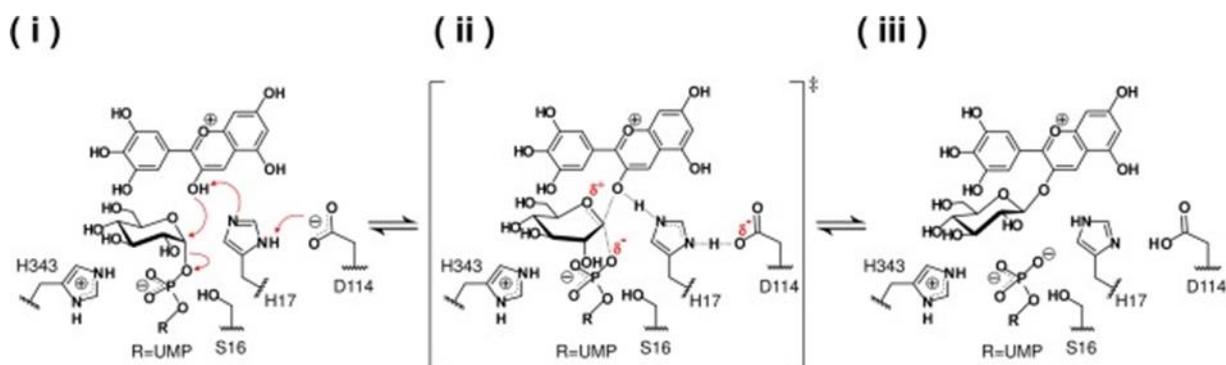


Figure 48: Proposed catalytic mechanism for the conversion of delphinidin to delphinidin 3-O-glucoside. This figure was presented by Hiromoto et al<sup>328</sup>.

(i) The His17-Asp114 catalytic dyad serves as a general base that deprotonates the nucleophile hydroxyl group of the acceptor substrate. (ii) The generated negative charge on the departing phosphate can be stabilized by the positive amino acid, His343, or the hydroxyl group of Ser16, which is adjacent to both the  $\beta$ -phosphate of UDP and the O5 oxygen of glucose. (iii) The resulting glucoside product and UDP are released from the active site.

Regarding the acceptor accommodation, the key determinants for molecular recognition are mostly scattered in the N-ter domain of UGTs. Highly divergent sequences can take up similar substrates. Neofunctionalization of UGTs during evolution was supported by the fast expansion of the superfamily and evolution trajectories of plants through multiple gene duplication or even whole genome duplication<sup>324</sup>. For illustration, UGT74 family can cope with flavonoids, benzoic acid derivatives, lignans, hydroxy coumarins, phenylethanoids, hydroquinones, diterpenes, triterpenes, glucosinolates, sesquiterpenoids, auxins and xenobiotics as listed by<sup>324</sup>. Conversely, a defined class of molecules like flavonoids can be recognized by UGTs belonging to different phylogenic clusters<sup>324</sup>. Furthermore, *in vitro* enzymatic assays can be misleading and divergent from *in vivo* assays regarding acceptor specificity. For instance, UGT71G1 converts more efficiently flavonoids than triterpenes *in vitro*. Nevertheless, *in planta* study underlines its involvement in triterpene biosynthesis<sup>39</sup>. Noticeably, some plant UGTs are highly specific for their acceptor substrates and are even able to discriminate enantiomers<sup>326</sup>, others are much more permissive: 27 out of 33 molecules assayed being converted by UGT71A35, for instance<sup>337</sup>. Promiscuous reactions can also be catalyzed with some UGTs as pointed out by Xie et al. who showed that UGT73AE1 from *Carthamus tinctorius* can graft glucose onto O-, S- or NH<sub>2</sub> groups catalyzing the formation of C-O, C-S or C-N linkage<sup>27</sup>. Peculiar UGTs were even shown to be able to yield C-glucosides<sup>338</sup>. In addition, due to the reversibility of the reaction and the substrate flexibility of

some enzymes more uncommon reactions can also be accessible such as sugar exchange from a glycosylated moiety or transglycosylation albeit with lower efficiencies than the canonical reaction<sup>339</sup>.

Predicting acceptor recognition from UGT sequence remains difficult. Different research groups undertook systematic studies to identify the substrates of *A. thaliana* UGTs. In a series of publications, the whole superfamily of *A. thaliana* UGTs was tested *in vitro* for glycosylation of benzoates<sup>340</sup>, hydroxycoumarins<sup>341</sup>, and terpenoids<sup>117</sup>. Other studies focused more on the *in planta* role of some UGTs in specific biosynthetic pathways<sup>342</sup>. A recent investigation applied a machine learning approach to predict the donor and acceptor specificities of UGTs from a data set established from an activity screening of the 107 UGTs towards 13 sugar donors and 91 potential nucleophiles. The results showed that functional properties were not related to phylogeny whereas the training set revealed interesting clusters based on donor or acceptor recognition and physico-chemical responses. The algorithmic prediction, named “GT-predict”, showed good correlation and a positive agreement between predicted and experimental results. This approach paves the way for predicting donor and acceptor recognition<sup>343</sup> and could enhance the discovery of new UGTs activities based on *in silico* approach.

Regarding the use of plant UGTs for the glycosylation of molecules dedicated to biotechnological applications, *in vitro* and *in vivo* approaches can be considered. *In vitro* approaches (enzyme-based processes of glycosylation) are limited by the cost and low availability of UDP-sugars. This can be overcome using coupling reactions as originally evidenced with sucrose synthase to recycle UDP-Glc from inexpensive sucrose source<sup>344</sup>. Recently, cascade reactions including donor substrate regeneration systems were also successfully used to produce UDP-galactose and establish an enzyme-based process leading to gram per liter product synthesis<sup>345</sup>. Further optimization of such cascades enabled a ~220 times efficient recycling of the UDP-Glc donor. Combination with optimized downstream purification steps afforded gram scale production of the purified glucosylated product<sup>346</sup>.

To avoid the problem of donor substrate regeneration, alternatives are based on *in vivo* glycosylation *via* bioconversion or cellular production. As *E. coli* possesses a pool of UDP-Glc, recombinant expression of plant UGTs was described to enable glucoside production. When the microbial host do not synthesize the targeted acceptor, the latter was added to the culture medium and glycosylated *via* bioconversion also termed “whole cell catalysis”<sup>347,348</sup>. Upregulation and/or deletions of targeted genes involved in the metabolism of UDP-Glc improved nucleotide sugar accumulation in cells and favored glycoside formation<sup>349</sup>. Further genetic engineering allowed the use of UDP-sugars, different from UDP-Glc, by complementing the missing biosynthetic gaps in *E. coli* leading to UDP-rhamnose or UDP-xylose<sup>349</sup>. The synthesis of new glycosides using glycosyltransferases specific for UDP-GlcNac donor, a nucleotide sugar naturally present in *E. coli*, was also performed to produce quercetin 3-O-N-acetylglucosamine<sup>349,350</sup>. Overall, these improvements in UDP-sugar

pathway regulation remarkably boosted the bioconversion abilities of *E. coli*. Notably, the first report of flavonoid glucoside production using recombinant *E. coli* reached 99.3 mg/L in fermentor<sup>348</sup> while recent studies mention bioconversion titers of 3.6 g/L to 3.9 g/L using different strategies of metabolic engineering targeting UDP-Glc metabolism<sup>351,352</sup>. The risk of degradation of the glycoside derivatives due to the action endogenous glycosidases has also to be taken into account too, as revealed by the pioneering work dealing with the conversion of indican<sup>11</sup>.

Ultimately, metabolic engineering tools can support an efficient supply of the aglycone without having to feed the strain with the aglycone or its precursor. A complete recombinant production of the glycosides is then achievable<sup>353</sup>. Furthermore, the balance between cell toxicity of the aglycone as well as its price and availability dictate the choice of the chassis strain and the strategy to be applied between full synthesis and bioconversion. With the advent of metabolic engineering and synthetic biology, one can expect that more examples will use the implementation of complete biosynthetic *in vivo* pathways. Again, the selection of the most appropriate host will be guided by the target pathway to be implemented. For example, if it involves membrane bound CYPs, a eukaryotic chassis may be more efficient<sup>354</sup>. This is evidenced with the case of steviol glycosides that are natural sweeteners having a diterpene scaffold with diverse glycosylation patterns. The tremendous industrial interest in these molecules has attracted two synthetic biology companies, namely, Amyris and Evolva, to develop fermentation processes based on recombinant production in yeast<sup>355</sup>. Alternatively, plant expression systems like *N. benthamiana* can also be used to evaluate the enzyme candidates in a more “native” condition and can help the correct assignment of their biochemical function<sup>356,357</sup>. However, this requires that endogenous UGTs do not interfere with the heterologous pathway.

Our use of the plant UGTs aimed at glucosylating sesquiterpene scaffolds and more precisely bisabolol and hernandulcin. Noteworthy, among the numerous studies on UGTs, only a few ones described glycosylation of sesquiterpenes<sup>324</sup>. To date, few plant UGTs have been identified as relevant for glycosylation of sesquiterpenes. They are limited to those described for abscisic acid, deoxynivalenol, and recently nerolidol glycosylation<sup>118,319,320,358</sup>. In addition, a systematic study of *A. thaliana* UGT superfamily revealed the promiscuity of some UGTs for sesquiterpenes such as farnesol and artemisinic acid<sup>117</sup>. A possible explanation for this rare occurrence of articles related to sesquiterpene glycosylation may be due to the overrepresentation of studies focusing on UGTs using flavonoid as acceptors or to the fact that promiscuous activities towards sesquiterpenes have been less investigated. Indeed, the occurrence of natural sesquiterpene glycosides is poorly described compared to glycosides derived from diterpene, triterpene or flavonoid scaffold. Sesquiterpenes and monoterpenes are usually studied as volatile fractions of natural products in GC while glycosylated molecules may not be detected because it would require LC analysis<sup>359</sup>. Hence the importance of such compounds may be underestimated and/or their occurrence may be more limited in plants.

In the frame of the thesis of Thomas Louveau, entitled “Investigation of glycosyltransferases from oat”, a set of UGT candidates from *Avena strigosa* were tested to identify their role in the biosynthesis of the defense molecule avenacin<sup>360</sup>. Their ability to glycosylate different acceptor substrates among which (-)- $\alpha$ -bisabolol was also assessed. The preliminary results reported in Thomas Louveau’s PhD thesis suggested that AsGT05827 enzyme was able to use (-)- $\alpha$ -bisabolol as an acceptor substrate (Table 10). However, some discrepancies between radioactivity-based assays and TLC detection of the products were noticed. It is therefore, difficult to draw definitive conclusions on the ability of AsGT05827 to glucosylate (-)- $\alpha$ -bisabolol (Table 10). As a follow-up to their initial studies, this research group investigated the role of AsGT05827 in the saponin pathway and renamed it UGT93B16 according to the UDP-glycosyltransferase nomenclature. They showed that this enzyme is not taking part in the avenacin biosynthesis<sup>361</sup>. Accordingly, AsGT05827 will be termed UGT93B16 in the rest of the manuscript.

Table 10: The AsGT05287 known acceptors from Thomas Louveau's thesis

Tested acceptors	Relative activity in radioassay (%)*	Product formation detected in TLC
2,4,5-Trichlorophenol	100	Not determined
Quercetin	-	Not determined
Kaempferol	84	Trace activity
Tricin	95	-
Capsidiol	-	-
(-)- $\alpha$ -Bisabolol	29	+
Lupeol	-	Not determined
$\beta$ -Amyrin	-	Not determined
$\beta$ -Amyrin-Ara	-	-

\* Relative activities deduced from radioactivity-based assays enzymatic reactions using UDP- $\beta$ -D-[<sup>3</sup>H]-glucose

In the present study, our goal is to identify enzymes catalyzing glycosylation of  $\alpha$ -bisabolol and hernandulcin to diversify the panel of molecules obtained from these scaffolds and extend their range of applications. Indeed, bisabolol glycosides showing anticancer activities were synthesized via classical chemistry<sup>362</sup>. Similarly, accessing to new glycosides of hernandulcin could help to enhance its solubility and also lower its bitter aftertaste while keeping its sweetening property. This was observed for steviol for which the different glycoside derivatives retained a sweetening power and a lower bitterness<sup>355,363</sup>.

A step-by-step approach was applied to attempt the glucosylation of  $\alpha$ -bisabolol and hernandulcin. First, we produced recombinant UGT in *E. coli* and performed *in vitro* enzymatic assays. Second, bioconversion using *E. coli* and *S. cerevisiae* cells was tested. Finally, the strain producing bisabolol (developed in chapter II) was used to implement a full biosynthetic pathway dedicated to the production of sesquiterpene glucoside.

## 2. Human UDP-glucuronosyltransferases

Human UDP-glucuronosyltransferases also share the same evolutionary origin as plant UGTs and belong to GT1 family (<http://www.cazy.org/><sup>322</sup>). Logically, they possess the same fold as plant enzymes although only one crystal structure could be obtained on a truncated version of the human UGT2B7<sup>364</sup>. Their reaction mechanism is supposed to be identical, with a catalytic histidine subtracting a proton that is assisted by an aspartic acid residue. Like plant UGTs their N-terminal domain predominantly mediates the acceptor recognition and the C-terminal domain is mainly devoted to UDP-sugar recognition. The C-terminal domain also shares with plant UGTs the following signature registered in Prosite database as:

[FW]-x(2)-[QL]-x(2)-[LIVMYA]-[LIMV]-x(4,6)-[LVGAC]-[LVFYAHM]-[LIVMF]-[STAGCM]-[HNQ]-[STAGC]-G-x(2)-[STAG]-x(3)-[STAGL]-[LIVMFA]-x(4,5)-[PQR]-[LIVMTA]-x(3)-[PA]-x(2,3)-[DES]-[QEHR]

Nonetheless, they exhibit important differences from plant UGTs. Compared to plants that have many representatives in UGT superfamily, the number of human UDP-glucuronosyltransferases is rather limited, humans having 4 families UGT1, UGT2, UGT3 and UGT8<sup>365</sup>. The xenobiotic metabolism is mainly relying on the 19 enzymes distributed in UGT1 and UGT2 families. These two families catalyze the transfer of glucuronic acid to their substrates using UDP-glucuronic acid as donor substrate. In humans, these enzymes are major contributors of the metabolism of xenobiotics and drugs. While phase I of drug metabolism involve CYPs, phase II includes the human UDP-glucuronosyltransferases that further helps in detoxification of exogenous compounds by making them more polar, and facilitates their clearance<sup>28</sup>. Hence, human UDP-glucuronosyltransferases are often studied with their generated metabolites for drug and toxicity issues.

UGT1A isoforms originate from alternative splicing (sharing the same C-ter domain with the same protein sequence from the 289 amino acid to 533<sup>366</sup>). Regarding UGT2, the family members derive from both gene duplication and alternative splicing. A schematic representation of two representative protein sequences of these families is depicted in Figure 49<sup>366</sup>. Another major difference from plant UGTs is that human UDP-glucuronosyltransferases are not cytosoluble but membrane bound. They both contain a N-ter signal peptide and are mainly located in the lumen side of the endoplasmic reticulum. As they are not cytosoluble, the recombinant expression of functional enzyme is trickier than that of plant UGTs.

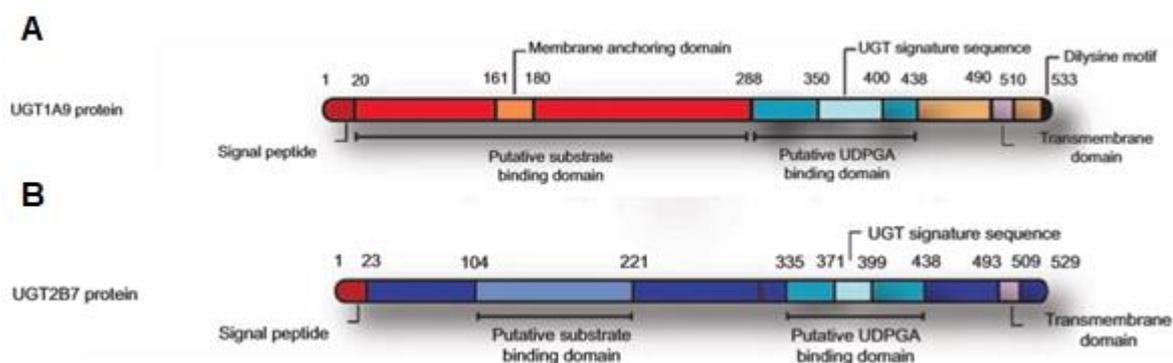


Figure 49: Protein primary structures of (A) UGT1A9 and (B) UGT2B7, figure reproduced from Guillemette et al.<sup>366</sup>

Determination of the metabolites made by human UDP-glucuronosyltransferases are usually investigated using human liver tissue from donors<sup>367–371</sup>, recombinant expression of individual enzymes in insect cells<sup>369–371</sup>, human cell lines<sup>371</sup> or ultimately combination of these different approaches. To date, two groups mentioned the use of yeasts to produce medicinal glucuronides using metabolic engineering. Overall their strategy was similar even though the host differed, the first one chose *S. pombe*<sup>372</sup> while the second employed *S. cerevisiae*<sup>373</sup>. Due to the addressing of the enzymes to the endoplasmic reticulum, a eukaryotic system is preferable. Donor substrate supply was established by the addition of a UDP-glucose 6-dehydrogenase (UGDH) encoding gene that converts UDP-Glc to UDP-GlcA. In *S. pombe*, rat or human UGDH were expressed from an autosomal plasmid whereas genomic integration of rat UGDH and glucuronosyltransferases were performed in *S. cerevisiae*. Despite the N-ter signal peptide which may not be properly processed by yeast and potential hurdle for UDP-GlcA access to the ER lumen or catalytic site of human UDP-glucuronosyltransferases both teams obtained glucuronides formation in their engineered strains. The highest conversion leading to 4-methylumbelliferone glucuronidation (24.8  $\mu\text{mol} / \text{day} / \text{g}$  cell dry weight) was obtained by Ikushiro et al. using *S. cerevisiae* and UGT1A6.

As glucuronides of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin have never been described in the literature, we decided to test these compounds as acceptor of human UDP-glucuronosyltransferases via *in vitro* enzymatic assays. We also integrated UGDH in the genome of *S. cerevisiae* to produce UDP-GlcA, the substrate donor of human UDP-glucuronosyltransferases. The engineered strains were then evaluated for glucuronide production either by bioconversion or by biosynthesis of sesquiterpene glucuronide (this would be the first example of a biosynthetic pathway in yeast for glucuronide production using human UDP-glucuronosyltransferases without having to supply exogenous substrate<sup>373</sup>).

## B. Material and methods

### 1. Strains and plasmids

The different plasmids and microbial strains used in this section are described in Table 11 and Table 12, respectively. All the cloning steps were carried out in *E. coli* Top10 cells (Thermo Fischer Scientific (Waltham, MA, US) and all the newly constructed plasmids were verified using Sanger sequencing from Eurofins (Luxembourg, Luxembourg).

The pH9-GW-UGT93B16 plasmid containing the *A. strigosa* UDP-glucosyltransferase UGT93B16 encoding gene and allowing its expression in *E. coli* was kindly provided by the John Innes Centre. For the construction of pYEDP60 + UGT93B16 and pYEDP60 + FlagUGT93B16, the coding sequence of the enzyme (Appendix “DNA sequence 4”) was amplified from pH9-GW-UGT93B16 template using primers 4.1, 4.3 and 4.2, 4.3 respectively (Table 18). The PCR products were then digested by EcoRI / BamHI and inserted in pYEDP60.

The pMRI34 + UGDH was obtained by Gibson cloning of the two synthetic DNA fragments purchased from Twist Biosciences (San Francisco, USA) encoding TEF1 promoter and the coding sequence of the *Rattus norvegicus* UGDH (Appendix “DNA sequence 5”). The two synthetic DNA fragments were amplified by PCR using the primers 4.4/ 4.5 and 4.6/4.7, respectively. pMRI34 was digested by PacI / Sall enzymes and the fragments were cloned in digested vector to yield pMRI34 + UGDH.

UGT1A9 / UGT2B4 were ordered as synthetic genes with an optimized codon usage for *S. cerevisiae* and received in pTwistAmp vector, while UGT2B7 was recoded, received as linear fragment and re-amplified by the primers 4.8/ 4.9 (Appendix “DNA sequence 6”). The digestion of the UGT1A9, UGT2B4 and UGT2B7 inserts was carried out with SmaI / BglII and cloned into pYEDP60 cut by SmaI / BamHI, yielding pYEDP60 + UGT1A9, pYEDP60 + UGT2B4 and pYEDP60 + UGT2B7.

The UDP-GlcA transporters FRC / UGTReI7 were ordered from Twist Bioscience as synthetic DNA fragments and amplified by PCR using the primers 4.10/4.11 and 4.12/4.13 while pFPP13 plasmid was obtained by amplification using the primers 2.7/2.8. The inserts and backbone were then joined by Gibson cloning.

Yeast cells were transformed as described previously using frozen cells for single plasmid transformation<sup>263</sup> or a higher efficiency protocol for plasmid cotransformation and integration<sup>264</sup>.

Table 11 Plasmids used in this chapter

Plasmid Name	Cloned enzyme	Origin of replication	Promoter	Terminator	Tag	Selection marker	Host expression
pET28b	-	pBR322	T7	T7	His	Kan	<i>E. Coli</i>
pH9-GW-UGT93B16	UGT93B16	pBR322	T7	T7	His	Kan	<i>E. Coli</i>
pYEDP60	-	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	-	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pYEDP60 + UGT93B16	UGT93B16	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	-	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pYEDP60 + FlagUGT93B16	UGT93B16	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	Flag	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pYEDP60 + UGT1A9	UGT1A9	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	-	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pYEDP60 + UGT2B4	UGT2B4	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	-	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pYEDP60 + UGT2B7	UGT2B7	2 $\mu$	GAL10-CYC1 hybrid promoter	PGK	-	Amp / Ura3 / Ade2	<i>S. cerevisiae</i>
pMRI34 + UGDH	UGDH	Integrative in DPP1 locus	TEF1	ADH1	-	Kan	<i>S. cerevisiae</i>
pFPP13 + FRC	FRC	ARS CEN6	TEF1	PGK	-	Amp / Leu2	<i>S. cerevisiae</i>
pFPP13 + UGTRel7	UGTRel7	ARS CEN6	TEF1	PGK	-	Amp / Leu2	<i>S. cerevisiae</i>

Table 12 : Strains used in this chapter

Strain Name	Organism	Genotype
Top10	<i>E. coli</i>	F- mcrA Δ( mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ( ara-leu)7697 galU galK rpsL (StrR) endA1 nupG
BL21DE3	<i>E. coli</i>	F' ompT hsdS <sub>B</sub> (r <sub>B</sub> -m <sub>B</sub> -) gal dcm (DE3)
BY4742	<i>S. cerevisiae</i>	MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0
EPY300	<i>S. cerevisiae</i>	MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 PGAL1-tHMG1::δ1 PGAL1-upc2-1::δ2 erg9::PMET3-ERG9::HIS3 PGAL1-ERG20::δ3 PGAL1-tHMG1::δ4
YBIS_02	<i>S. cerevisiae</i>	EPY300, Δho::TADH1-BBS-PGAL10-PGAL1-HsCPR-TCYC1
YBIS_01 + UGDH	<i>S. cerevisiae</i>	EPY300, Δho::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1 ΔDPP1:: PTEF1- UGDH-TADH1
BY4742+ UGDH	<i>S. cerevisiae</i>	BY4742, ΔDPP1:: PTEF1- UGDH-TADH1

## 2. Culture conditions and sampling

### a) *A. strigosa* UGT93B16 production

The *E. coli* production of N-Ter His tagged UGT93B16 was performed in BL21DE3 strain transformed by pH9-GW-UGT93B16. A 5 ml preculture in LB supplemented with Kanamycin (40 µg/mL) was grown overnight. Then, 50 mL of ZYM5052 medium (prepared as in reference <sup>374</sup>) supplemented with Kanamycin (40 µg/mL) were inoculated at 0.05 OD with the preculture and the cells were grown under 200 rpm agitation at either 28°C or 18°C. At the end of UGT93B16 expression, the cells were pelleted at 10 000 g for 5 minutes and stored at -20°C. Finally, the production of the heterologous enzyme was verified on SDS PAGE gel in reducing conditions.

BL21DE3 strains producing UGT93B16 were also used for bioconversion of (-)-α-bisabolol or (±)-hernandulcin. To do so, cultures were carried out at 18°C and 200 rpm using ZYM5052 medium supplemented with Kanamycin (40 µg/mL). (-)-α-bisabolol or (±)-hernandulcin (dissolved in DMSO) were added at a final concentration of 0.1 mM at the beginning of the bioconversion. 5 mL of culture medium were taken after 26 hours and 48 hours of cultures for HPLC analysis of the glucosylation products and stored at -20°C until further manipulation. The glucosylated products were extracted by adding 2.5 mL of ethyl acetate to the 5 mL culture aliquot. After 1 minute vortexing, the organic phase was separated by centrifugation at 10,000 g for 10

minutes at ambient temperature. The organic layer was removed and a second extraction was performed. The ethyl acetate fraction was dried under airflow and the metabolites were resuspended in 0.5 ml ethanol. These samples were analyzed in HPLC-UV-MS using a C18 precolumn and a Prontosil 120-5-C18-SH 5.0  $\mu\text{m}$  250\*4.0 mm column (Bischoff chromatography) using the separation method described previously (in chapter II) except that the scanned range of the MS detector was extended to a window covering 150 Da to 450 Da.

For expression of UGT93B16 in *S. cerevisiae* YBIS\_02 strain, a preculture was inoculated in SD -His -Met -Ura overnight at 30°C, 200 rpm. The preculture was used to inoculate at 0.1 OD<sub>600nm</sub> 50 mL of expression culture medium YPA (peptone 20 g/L, yeast extract 10 g/L, adenine 80 mg/L) supplemented with galactose 18 g/L, glucose 2 g/L and 2 mM methionine. Cells were grown for 120 hours at 30°C, 200 rpm. For metabolite analysis of the glucosylated products, 5 mL aliquots of culture medium and cells were sampled and kept at -20°C until further processing. To these 5 ml, 1 g of glass beads ( $\varnothing=0.4 \mu\text{m}$ - $0.6 \mu\text{m}$ , acid washed beads Sigma Aldrich) was added and extraction of glucosides was performed as described for *E. coli*.

Bioconversion experiment in *S. cerevisiae* relied on BY4742 strain transformed by pYEDP60, pYEDP60 + UGT93B16 or pYEDP60 + FlagUGT93B16. An overnight preculture in SD -Ura (at 30°C, 200 rpm) was used to inoculate, at 0.1 OD<sub>600nm</sub>, 50 mL of YP (peptone 20 g/L, yeast extract 10 g/L) supplemented with galactose 18 g/L, glucose 2 g/L and 0.1 mM of (-)- $\alpha$ -bisabolol or ( $\pm$ )-hernandulcin. 5 mL were sampled at 24 hours, 48 hours and 120 hours to evaluate the glucosylated product formation by HPLC after chemical extraction.

## b) Human UDP-glucuronosyltransferases

Human UDP-glucuronosyltransferases were expressed in the strain YBIS\_01 + UGDH in the same conditions as the strain YBIS\_02 transformed by UGT93B16. In the strain YBIS\_01 + UGDH, when UDP-glucuronosyltransferases were coexpressed with UDP-GlcA transporters (UGTRel7 or FRC), the culture steps were identical except for the preculture, which was performed using SD -His -Met -Ura -Leu medium.

Bioconversion assay of (-)- $\alpha$ -bisabolol using BY4742 + UGDH and UDP-glucuronosyltransferases was performed as described for BY4742 + UGT93B16. When cotransformation with UDP-GlcA transporters encoding plasmids (pFPP13 + FRC or pFPP13 + UGTRel7) were tested, the preculture was performed with an appropriate selection medium (SD -Ura -Leu).

In all chemical extractions of yeast cultures for UDP-glucuronosyltransferase assay, 50  $\mu\text{L}$  of HCl 5N were added to favor extraction of potential glucuronides (pH dropped to  $\sim$ 3-3.5). Ethyl acetate extraction and HPLC analysis was subsequently performed as previously described.

### 3. *In vitro* glycosylation assays

The frozen pellets of UGT93B16 expression in *E. coli* cells were suspended to a final concentration of 80 OD in 50 mM Tris HCl pH 7.5 supplemented with 0.7 mg/mL lysozyme, 2.5 mg/L DNaseI, 50 mM NaCl and incubated 1 hour on ice for lysozyme digestion. Then, the solution was sonicated at 30% power for 30 seconds and let on ice for 2 minutes (Sonics, Vibra-cell VCX 130 sonicator), this cycle was repeated 4 times. The cell debris were pelleted for 5 minutes at 10 000 *g*. For SDS-PAGE analysis, the supernatant was loaded on gels and considered as the soluble fraction, while the pellet was resuspended with the same initial volume and termed as insoluble fraction.

Enzymatic activity was assayed at 30°C in 50 mM Tris HCl pH 7.5, 1 mM UDP-Glc, 0.2 mM of acceptor sesquiterpene dissolved in ethanol (10 mM stock solution) and using crude lysate (10 % of the final volume of the assay). For product analysis, samples were taken at 5 minutes, 60 minutes and 18h reactions for each tested molecule. The reactions were stopped using 1 volume of acetonitrile and vortexing. After a 5-minute centrifugation step at 13 000 *g*, 40 µL of reactions were injected in HPLC-UV-MS using the same column and gradient as described previously.

To prepare hydroxy derivatives of (+)-*epi*- $\alpha$ -bisabolol and test them as potential acceptor substrates of UGT93B16, we performed hexane extractions from cultures of the strains YBIS\_01 + CYP2B6, YBIS\_01 + CYP2B11, YBIS\_01 + CYP2C9 and YBIS\_01 + CYP2D6 as described previously (see chapter II), except that after drying under airflow, the extract were suspended in 25 µL ethanol and analysed using HPLC to control their content. From this solution, 2 µL were transferred in the enzymatic reactions in a final volume of 100 µL.

Regarding the human UDP-glucuronosyltransferases, semi-purified enzymes corresponding to microsomal fractions expressed from insect cells were purchased from Corning as Supersomes™ (Corning, US). The enzymatic assays were performed with 0.5 mg/mL final protein concentration of microsomal fractions, 2 mM UDP-GlcA, 50 mM Tris HCl pH 7.5, 8 mM MgCl<sub>2</sub>, 25 µg/ml alamethicin, 0.2 mM ( $\pm$ )-hernandulcin or (-)- $\alpha$ -bisabolol added from ethanol stock solution. The reaction was incubated overnight at 37°C and stopped by adding one volume of acetonitrile. After a 3 minute centrifugation step at 10 000 *g*, 40 µL were analyzed by HPLC as described previously, with the extended window covering 150 Da to 450 Da for MS detection.

## C. Results and discussion

### 1. UGT93B16: a catalyst for the glucosylation of sesquiterpenes

#### a) UGT93B16 assays from *E. coli* expression

UGT93B16 production was first monitored by SDS PAGE analysis of whole cell extract taken at 4, 7.5 and 24 hours of cultures in ZYM5052 medium. The BL21DE3 + UGT93B16 strain was compared to the control strain transformed with pET28b vector to visualize the protein background of BL21DE3 (supplementary Figure 69). We observed the presence of a protein of around 50 kDa at 24 hours, which corresponds to the molar mass of UGT93B16 (51 kDa) and attests of the recombinant production of the enzyme. Cultures at two different temperatures (18°C and 28°C) were then performed and we verified whether UGT93B16 was expressed as a soluble protein or recovered in the insoluble fraction. As shown in Figure 50 and Figure 51, the total level of production of UGT93B16 (considering soluble and insoluble fraction) is higher at 28°C than 18°C. However, growing the cells at 18°C was more favorable to produce soluble enzyme. Lysozyme digestion enabled *E. coli* lysis, but additional sonication steps enhanced the enzyme recovery. We thus decided to grow *E. coli* at 18°C and systematically sonicated the cells for enzyme recovery.

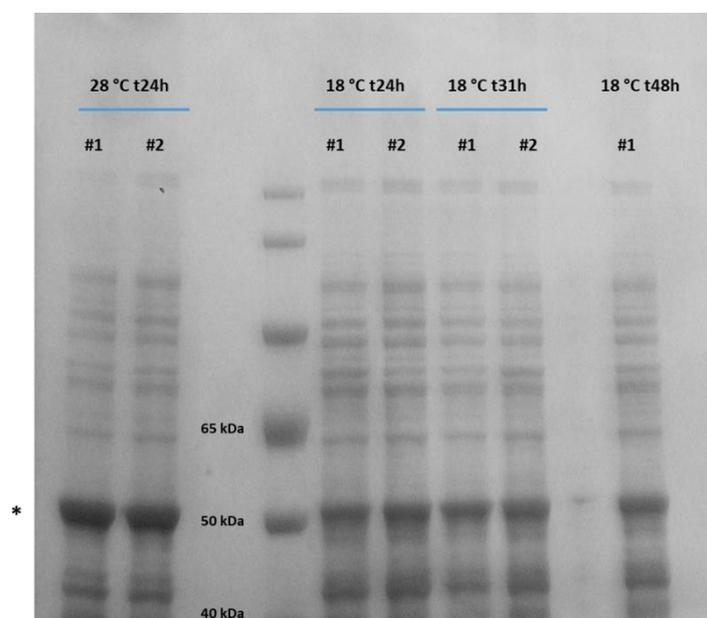


Figure 50: NuPAGE Tris-acetate 7% protein gel on whole cells to assess UGT93B16 expression at different temperatures #1 and #2 correspond to duplicated cultures  
\* Highlights the presence of UGT93B16

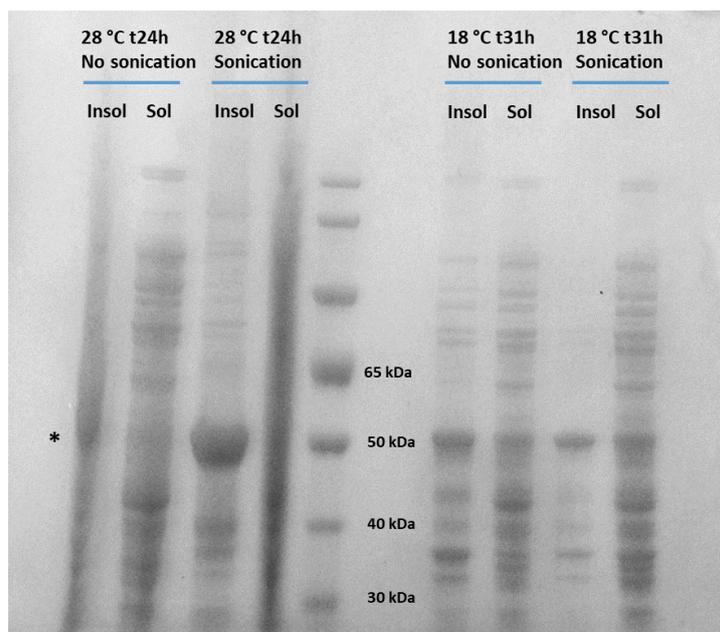


Figure 51: NuPAGE Tris-acetate 7% protein gel assessing the UGT93B16 soluble expression.  
 Insol: insoluble fraction, Sol: soluble fraction  
 \* Highlights the presence of UGT93B16

We first tested the ability of UGT93B16 to glucosylate ( $\pm$ )-hernandulcin and (-)- $\alpha$ -bisabolol in enzymatic assay with a crude lysate from the BL21DE3 + UGT93B16 strain. We also included trans,trans-farnesol and nerolidol in our acceptor list to assess the promiscuity of UGT93B16 with acyclic sesquiterpenes as well as the new hydroxy derivatives of (+)-*epi*- $\alpha$ -bisabolol obtained from cultures of the strains YBIS\_01 + CYP2B6, YBIS\_01 + CYP2B11, YBIS\_01 + CYP2C9 and YBIS\_01 + CYP2D6 (see Chapter II).

LC/MS analysis evidenced conversion of ( $\pm$ )-hernandulcin (Figure 52) and proved that the obtained products correspond to glucosides. Indeed, the product obtained from ( $\pm$ )-hernandulcin eluted at  $\sim$ 10 min (Figure 52A, B) and MS scan detected a  $m/z = 399.51, 421.42$  corresponding to the mass of  $[M_{\text{hernandulcin monoglucoside}} + H]^+$  and  $[M_{\text{hernandulcin monoglucoside}} + Na]^+$ , respectively, a  $m/z = 237.4$  corresponding to the aglycone release  $[M_{\text{hernandulcin}} + H]^+$  and a  $m/z = 219.47$ , corresponding to a rearrangement of the dehydrated aglycone  $[M_{\text{hernandulcin}} + H - H_2O]^+$ , Figure 53. This is in agreement with the literature, which reports the formation of sodium adducts<sup>375</sup>, aglycone release and dehydrated rearrangements<sup>267</sup> from glycoconjugates.

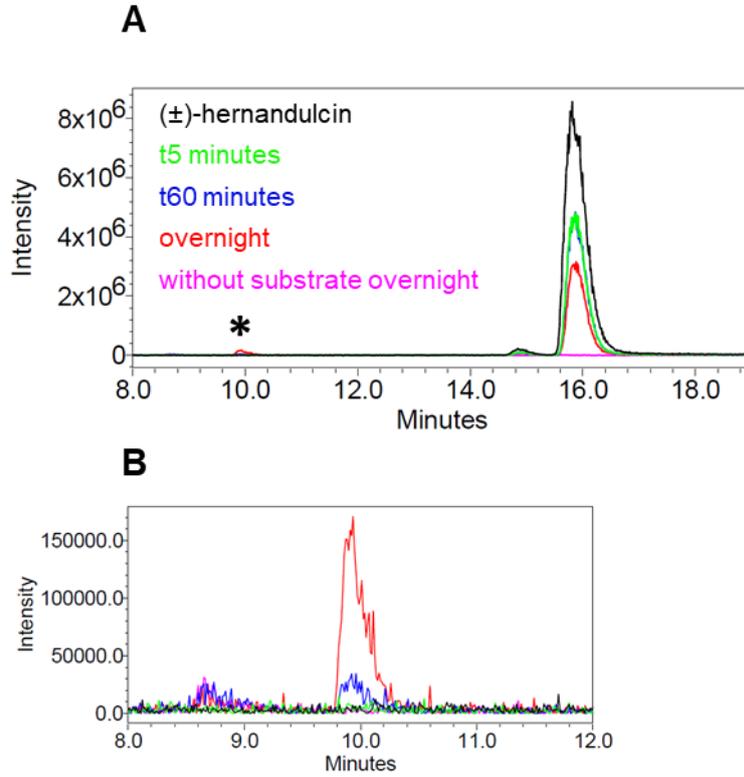


Figure 52: MS signal at  $m/z = 219.3$  showing the time course of  $(\pm)$ -hernandulcin-Glc formation by UGT93B16  
 (A) MS signal at  $m/z = 219.3$  showing the time course of  $(\pm)$ -hernandulcin-Glc formation,  
 \* Highlights the presence of glucosylated product, (B) Zoom in the region of  $(\pm)$ -hernandulcin-Glc formation

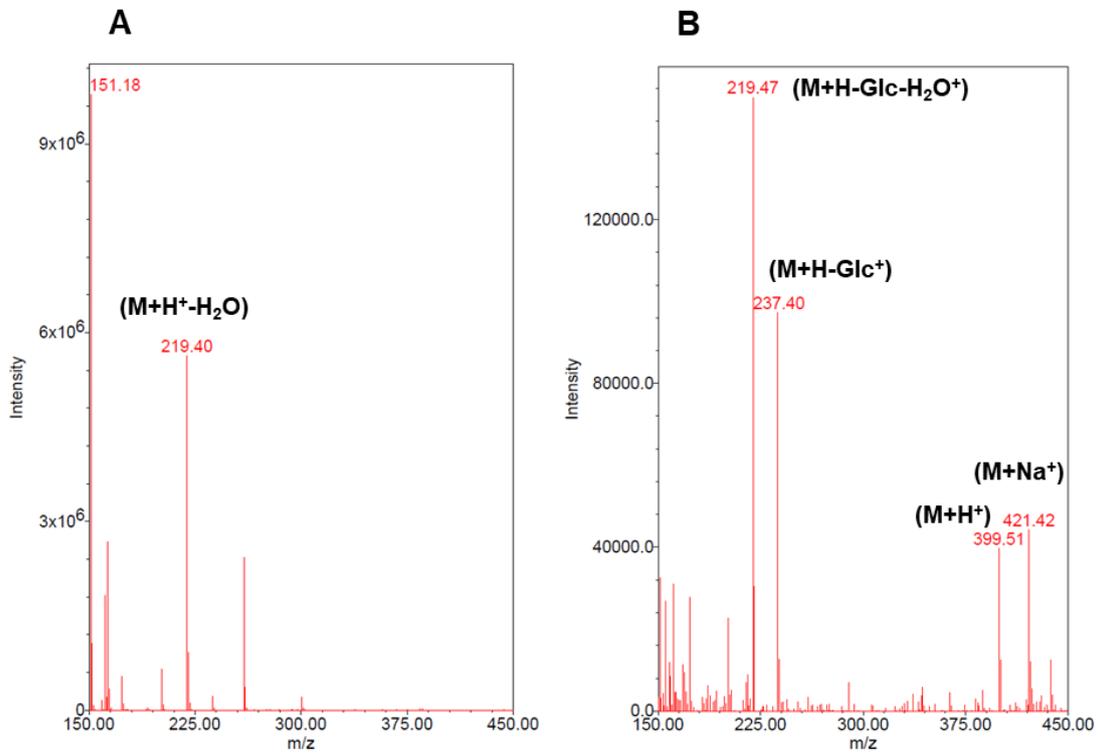


Figure 53: MS spectra of (A)  $(\pm)$ -hernandulcin and (B)  $(\pm)$ -hernandulcin-Glc after UGT93B16 incubation.

Conversion of (-)- $\alpha$ -bisabolol and trans,trans-farnesol was more efficient and also concomitant to the production of a more polar compound corresponding to the sesquiterpene glucoside (Figure 54, Figure 55). For both cases, the mass signal corresponding to the aglycone release with a dehydrated rearrangement  $[M_{\text{sesquiterpene}} + \text{H} - \text{Glc} - \text{H}_2\text{O}]^+$  was the main one detected for all the other sesquiterpenes (Figure 54C, Figure 55C). In order to ease the reading of the results, the time course of this specific ion signal from mass spectrometry is shown in appendix Figure 70, Figure 71, Figure 72 for the other substrates. Only traces of nerolidol were glucosylated (Figure 72, UV cannot quantify the conversion for nerolidol). The conversion values are summarized in Table 13 according to UV signal quantitation at  $\lambda_{\text{max}}$ .

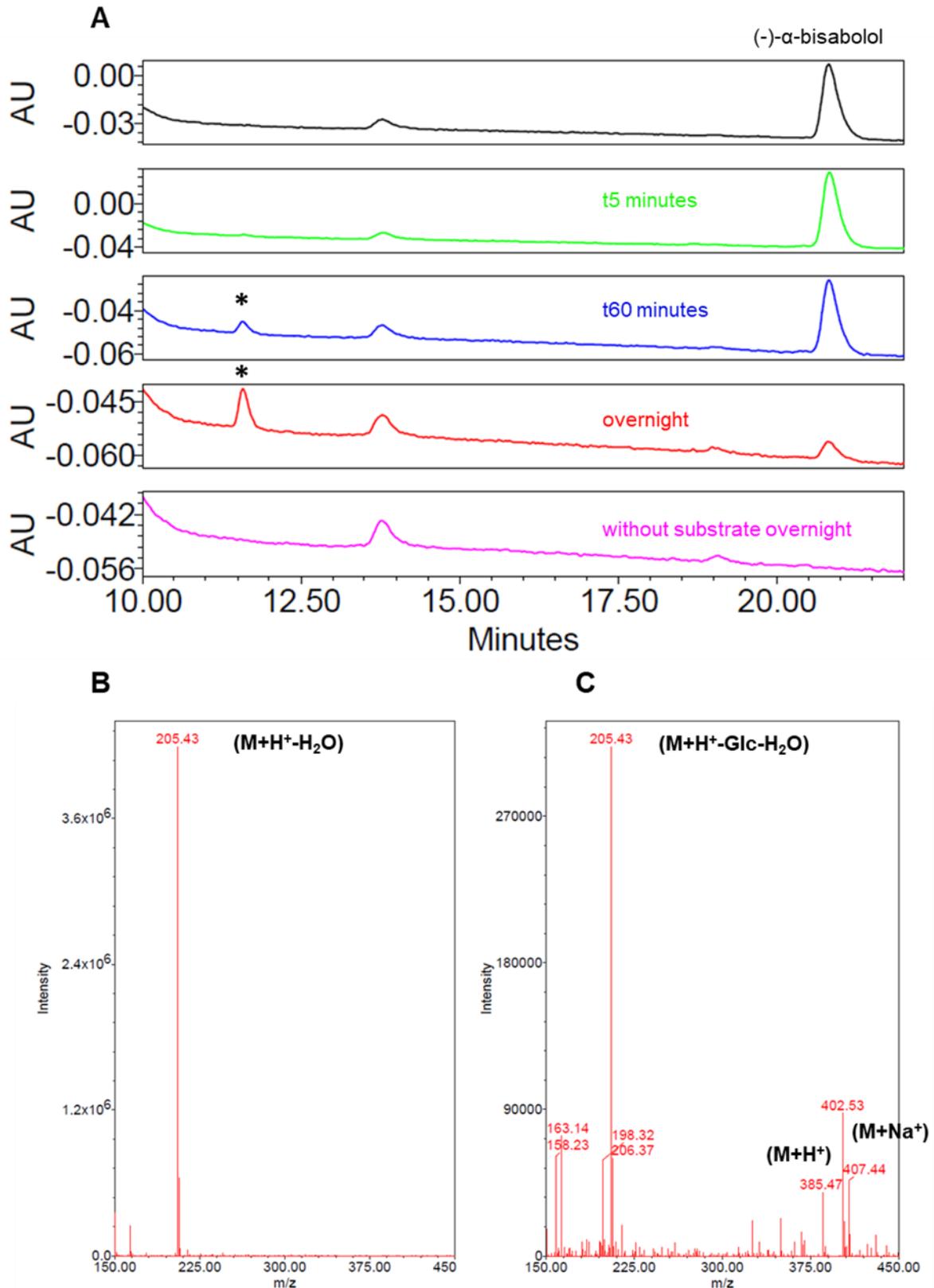


Figure 54: Time course conversion of (-)- $\alpha$ -bisabolol by UGT93B16  
 (A) UV chromatograms following the reaction (205 nm). \* corresponds to (-)- $\alpha$ -bisabolol-Glc formation, (B) MS spectra of (-)- $\alpha$ -bisabolol, (C) MS spectra of (-)- $\alpha$ -bisabolol-Glc

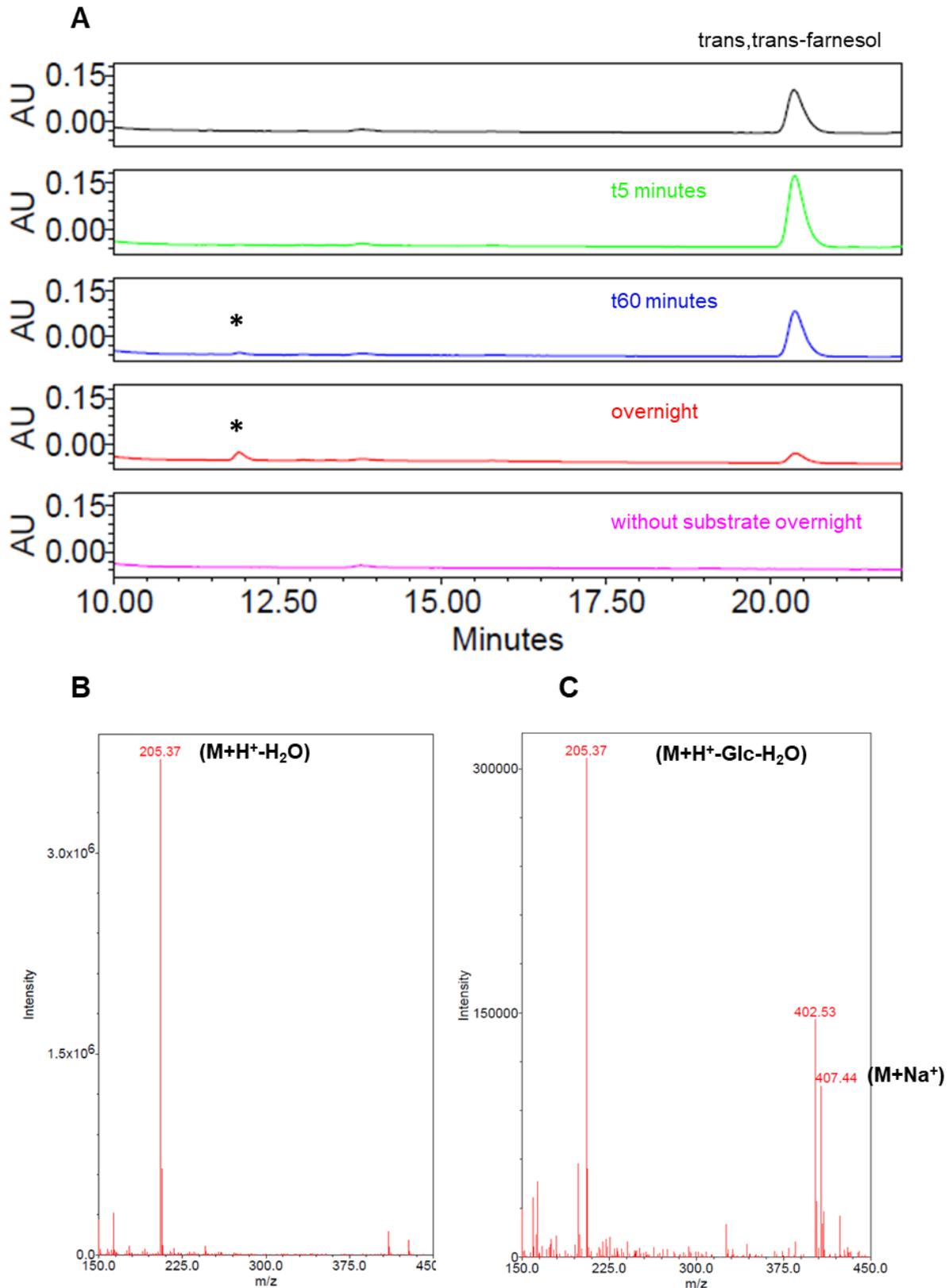


Figure 55: Time course conversion of trans,trans-farnesol by UGT93B16  
 (A) UV chromatograms following the reaction (205 nm), \* corresponds to trans,trans-farnesol -Glc formation, (B) MS spectra of trans,trans-farnesol, (C) MS spectra of trans,trans-farnesol -Glc

We showed that ( $\pm$ )-hernandulcin, (-)- $\alpha$ -bisabolol, trans,trans-farnesol, and nerolidol are converted with UGT93B16, thus enlarging its array of substrates. Our LC/MS analysis unequivocally demonstrates that (-)- $\alpha$ -bisabolol is an acceptor substrate of UGT93B16 and confirms the results of Thomas Louveau obtained by TLC analysis. Moreover, LC/MS analyses further revealed that hydroxy-(+)-*epi*- $\alpha$ -bisabolol produced from the engineered strains of *S. cerevisiae* (YBIS\_01 + CYP2B6, CYP2B11, CYP2C9 and CYP2D6) are also acceptor substrates (Figure 56). It was not possible to elucidate the structures of all the products obtained using the various CYPs due to the difficulty to purify them (chapter II). In the case of 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol acceptor produced by YBIS\_01 + CYP2D6, we confirm that the molecule is glucosylated by UGT93B16 (Figure 56 and the corresponding MS spectra are presented Figure 57). To conclude on the precise nature of the new molecules, purification and NMR characterization of the glucosides are required to determine the regioselectivity of the O- glucosylation (the most probable one) or C- glucosylation which cannot be excluded. Indeed, most plant UGTs are active on hydroxy groups. Only few of them, clustered in a specific monophyletic group were reported to catalyze C-glycosylation and UGT93B16 does not belong to this group <sup>26</sup>.

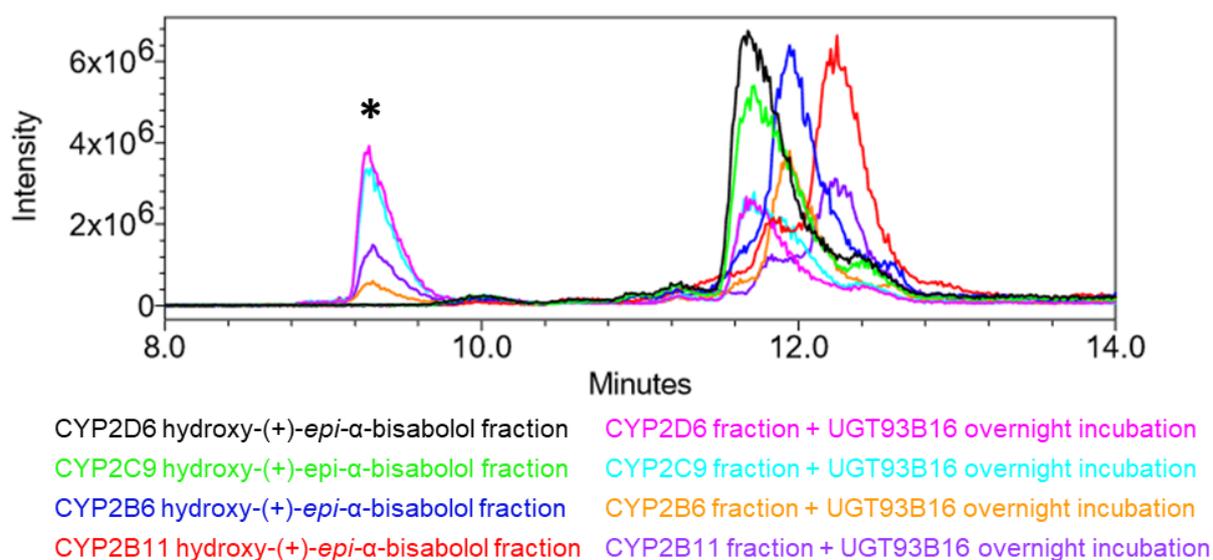


Figure 56: MS signal at  $m/z = 203.3$  showing the time course of hydroxy-(+)-*epi*- $\alpha$ -bisabolol-Glc formation by UGT93B16 from YBIS\_01 + CYP extractions.

\* Highlights the presence of glucosylated product

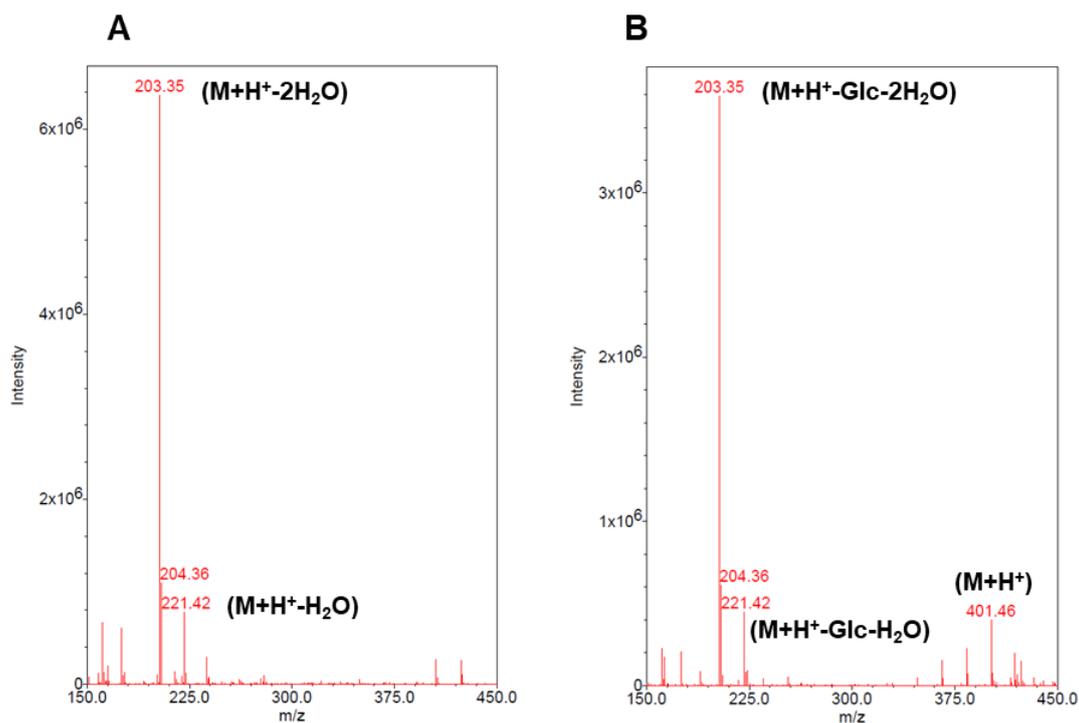


Figure 57: MS spectra of (A) 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol from hexane extraction of YBIS\_01+CYP2D6 strain and (B) 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol-Glc spectra after UGT93B16 incubation.

A more detailed biochemical study could be performed in order to determine the kinetic parameters ( $k_{cat}$  and  $K_m$ ) for the glucosylation of ( $\pm$ )-hernandulcin, (-)- $\alpha$ -bisabolol, trans,trans-farnesol and nerolidol and further assess the substrate preference of UGT93B16. Additionally, structural analysis or molecular modelling of UGT93B16 combined with docking experiments of the various substrates could help understanding how UGT93B16 is able to handle molecules as different as (-)- $\alpha$ -bisabolol (a cyclic sesquiterpene with a tertiary alcohol), farnesol (an acyclic sesquiterpene with a primary alcohol) or ( $\pm$ )-hernandulcin. This could provide useful information to engineer the enzyme for a better recognition of the sesquiterpenes and for instance adapt the enzyme to the keto group of hernandulcin that seems to prevent efficient glucosylation compared to (-)- $\alpha$ -bisabolol.

The data showing that UGT93B16 is active on hydroxy-(+)-*epi*- $\alpha$ -bisabolol also opens up to the design of orthogonal pathways leading to new sesquiterpene glucosides. This could greatly expand the potential of diversification of the plant (+)-*epi*- $\alpha$ -bisabolol scaffold. Besides, it raises the question on how to organize such new metabolic routes because competition between oxidative and glucosylation enzymes could exist for (+)-*epi*-bisabolol substrate in a microbial host coexpressing the CYPs and the UGT. Would a tight control of the enzyme expression be sufficient? Or should we envision a compartmentalization strategy?

Table 13: Sesquiterpene conversion with cell lysate containing UGT93B16.  
Conversion is expressed as percentage transformation of the exogenous substrate (0.2 mM)

Substrate	Conversion % **
(-)- $\alpha$ -bisabolol	61
trans,trans-farnesol	33
( $\pm$ )-hernandulcin	1.3
nerolidol	Trace activity in MS
YBIS_01 + CYP2B6 fraction*	+***
YBIS_01 + CYP2B11 fraction*	+***
YBIS_01 + CYP2C9 fraction*	+***
YBIS_01 + CYP2D6 fraction*	+***

\* Fractions containing hydroxylated (+)-epi- $\alpha$ -bisabolol as shown in Chapter II

\*\* Conversion was calculated as the ratio between the area of product(s) and the sum of the areas of substrate and product(s) according to UV quantitation from O/N incubation with the UGT93B16. In the absence of available standards, we assumed that substrates and products have similar molar extinction coefficient at 205 nm.

\*\*\* Glucosylated product(s) observed in UV and MS detection which could not be quantified due to the presence of additional contaminant molecules in the cell extract

## b) Bioconversion of sesquiterpenes

### 1) In *E. coli*

As we confirmed the *in vitro* activity of UGT93B16 on several sesquiterpenes, we tested whole cell glucosylation. As reported by Caputi et al <sup>117</sup>, whole cell biotransformation using engineered *E. coli* can sustain the production of glucosylated terpenoids. This was evidenced for two sesquiterpenes namely farnesol and artemisinic acid although with low conversion yields of 4.6 % and 18.6 %, respectively. Nonetheless, this work validated the fact that certain sesquiterpenes can diffuse through the cells and become accessible to cytosolic UGT. Therefore, one can expect that (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin could also be internalized in *E. coli* and converted.

To test this hypothesis, we carried out (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin glucosylation using whole cells grown at 18°C to favor soluble expression in the presence of 0.1 mM of acceptor, added at the beginning of the culture. The glucoside formation was monitored during the culture (Table 14).

Table 14: Conversion of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin bioconversion in *E. coli*. Conversion is expressed as percentage transformation of the exogenous substrate (0.1 mM)

Bioconversion condition	Conversion %
t = 26h (-)- $\alpha$ -bisabolol	5.3
t = 48h (-)- $\alpha$ -bisabolol	27
t = 26h ( $\pm$ )-hernandulcin	-
t = 48h ( $\pm$ )-hernandulcin	+*

\*Not quantifiable, activity could be detected only in MS

In agreement with *in vitro* results, UGT93B16 glucosylated more efficiently (-)- $\alpha$ -bisabolol than ( $\pm$ )-hernandulcin. For more accuracy, comparison with *in vitro* assay should be performed using the same initial concentration of sesquiterpenes. The values obtained for (-)- $\alpha$ -bisabolol is higher than those reported by Caputi et al. for other sesquiterpene conjugation<sup>117</sup>. We obtained lower conversion values with whole cells than with enzymatic assay. This may originate from a poor uptake or diffusion of the sesquiterpenes in *E. coli*. One direction to test the uptake issue could be the addition of a cell permeabilizer, such as DMSO or detergents, during bioconversion.

## 2) In *S. cerevisiae*

Due to the low conversion observed in *E. coli*, we considered to evaluate *S. cerevisiae* as an alternative microbial workhorse. UDP-Glc, the donor substrate of UGTs, is present in *S. cerevisiae* indicating that it could be a relevant host. Notably, *S. cerevisiae* was used to select and identify a UGT acting on solanidine thanks to a positive selection screen<sup>376</sup>. More recently, *S. cerevisiae* was used either for the whole cell bioconversion of zearalenone or for the production of high value added saponins in fully biosynthetic approaches<sup>18,377</sup>. So, we prepared two plasmids expressing either UGT93B16 or its Flag tagged version. Both were used to transform *S. cerevisiae* BY4742 strain and assess the ability of the strains to catalyze (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin glucosylation. Indeed, the tag could be deleterious to the activity while its presence could also be useful to probe expression using Western blot for instance.

The conversion of (-)- $\alpha$ -bisabolol or ( $\pm$ )-hernandulcin obtained after 5 days of culture is reported in Table 15 and growth of the yeasts producing the tagged or the untagged UGT was equivalent. As previously seen for *in vitro* assay or *E. coli* bioconversion, (-)- $\alpha$ -bisabolol is more efficiently converted than ( $\pm$ )-hernandulcin. In addition, conversion with the Flag tagged enzyme is much higher than that observed with the

native enzyme, which could be due to a higher expression and/or catalytic efficiency of the tagged version. This could be verified by checking the level of production of the two enzymes and by comparing the activities of the pure native UGT93B16 and its tagged version.

Notably, the conversion yield is significantly higher in *S. cerevisiae* than in *E. coli*. (-)- $\alpha$ -bisabolol conversion increased from 27% in *E. coli* to 98 % in yeast while ( $\pm$ )-hernandulcin conversion reached 28% with the Flag tagged enzyme in yeast whereas only traces of conversion were obtained using *E. coli*. This could be attributed to multiple factors such as a better soluble expression of UGT93B16, a better uptake of acceptor substrates or a better availability of UDP-Glc in *S. cerevisiae* as compared to *E. coli*. It is worth noting that, for all our assays, addition of the sesquiterpenes in culture broth did not affect *E. coli* and *S. cerevisiae* growth. This underlines that there was no detectable toxicity of the sesquiterpenes and their glucoside derivatives in the range of tested concentrations.

Table 15: Bioconversion yield of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin in *S. cerevisiae*  
Yields are expressed as percentage conversion of the exogenous substrate (0.1 mM) and indicated for the 120 hour incubation

Strain	Conversion	Conversion
	(-)- $\alpha$ -bisabolol %	( $\pm$ )-hernandulcin %
BY4742 + pYEDP60 UGT93B16	41.9	3.2
BY4742 + pYEDP60 FlagUGT93B16	98.1	28.2

The accumulation of glucosides was monitored during the bioconversion using BY4742 strain + FlagUGT93B16. The results are presented in Figure 58. The (-)- $\alpha$ -bisabolol glucoside formation rapidly reached a plateau within 48 hours. On the contrary, the ( $\pm$ )-hernandulcin glucoside formation steadily increased during the 120 hours. This reflects that a higher substrate supplementation could be applied for (-)- $\alpha$ -bisabolol. For ( $\pm$ )-hernandulcin, the slower conversion highlights the weaker catalytic efficiency of UGT93B16 toward this substrate and/or may also result from a poorer cell uptake. A summary of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin glucoside yield using *in vitro* assay and conversion values is presented in Table 16.

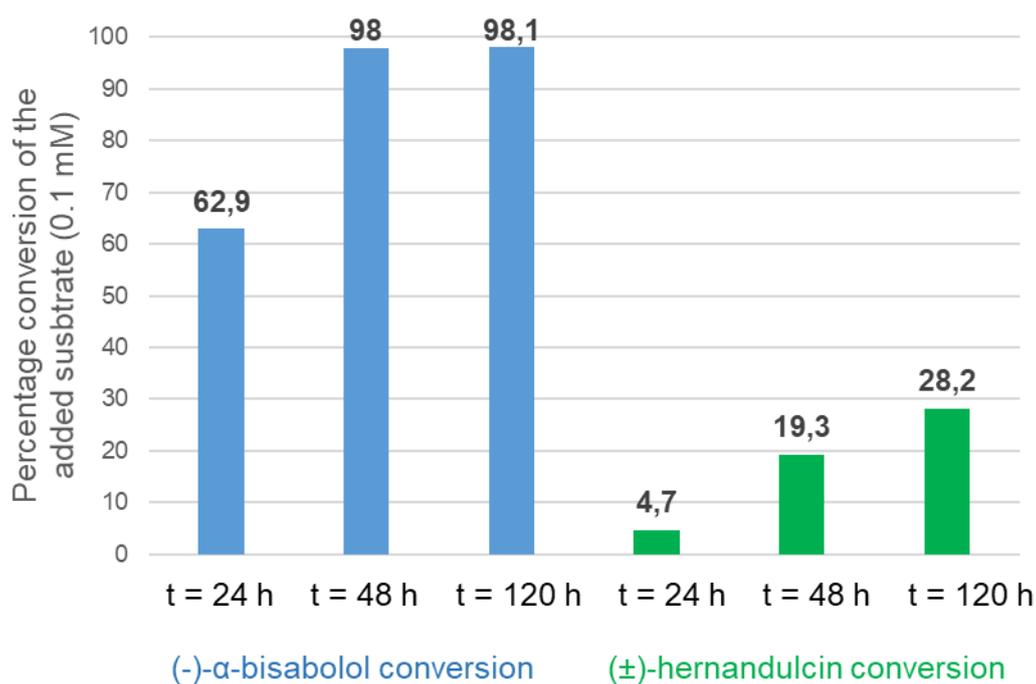


Figure 58: Time course of the (-)-α-bisabolol and (±)-hernandulcin glucosylation by FlagUGT93B16 showed as percentage conversion of the added substrate.

Table 16: (-)-α-bisabolol and (±)-hernandulcin conversion obtained in various conditions of glucosylation

	Enzymatic <i>in vitro</i> assay <sup>1</sup>		Bioconversion in <i>E. coli</i> BL21DE3 strain <sup>2</sup>		Bioconversion in <i>S. cerevisiae</i> BY4742 strain <sup>3</sup>	
	(-)-α-bisabolol (0.2 mM)	(±)-hernandulcin (0.2 mM)	(-)-α-bisabolol (0.1 mM)	(±)-hernandulcin (0.1 mM)	(-)-α-bisabolol (0.1 mM)	(±)-hernandulcin (0.1 mM)
Conversion <sup>4</sup>	61	33	27	0	98	28

<sup>1</sup> Crude extract of N-ter His tagged UGT93B16, <sup>2</sup> N-ter His tagged UGT93B16, <sup>3</sup> N-ter Flag tagged UGT93B16

<sup>4</sup> Conversion is expressed as percentage transformation of the exogenous substrate

c) UGT93B16 expression in an engineered yeast producing (+)-*epi*- $\alpha$ -bisabolol

Finally, we attempted to produce glucoside derivatives of (+)-*epi*- $\alpha$ -bisabolol *in vivo* by building a full biosynthetic pathway comprising BBS, and UGT93B16. To this end, we transformed YBIS\_02, which produces (+)-*epi*- $\alpha$ -bisabolol with the UGT93B16 encoding plasmids. LC-MS analysis revealed the presence of a glucoside product in both YBIS\_02 + pYEDP60 UGT93B16 and YBIS\_02 + pYEDP60 FlagUGT93B16 extracts (and absence in the control strain YBIS\_02 + pYEDP60), Figure 59. Considering that YBIS\_02 produces a mixture of sesquiterpenes (+)-*epi*- $\alpha$ -bisabolol and trans,trans-farnesol, two glucosylated compounds could be produced. The glucoside product should be purified to determine its structure with accuracy. Again, the strain producing the tagged enzyme had superior conversion abilities.

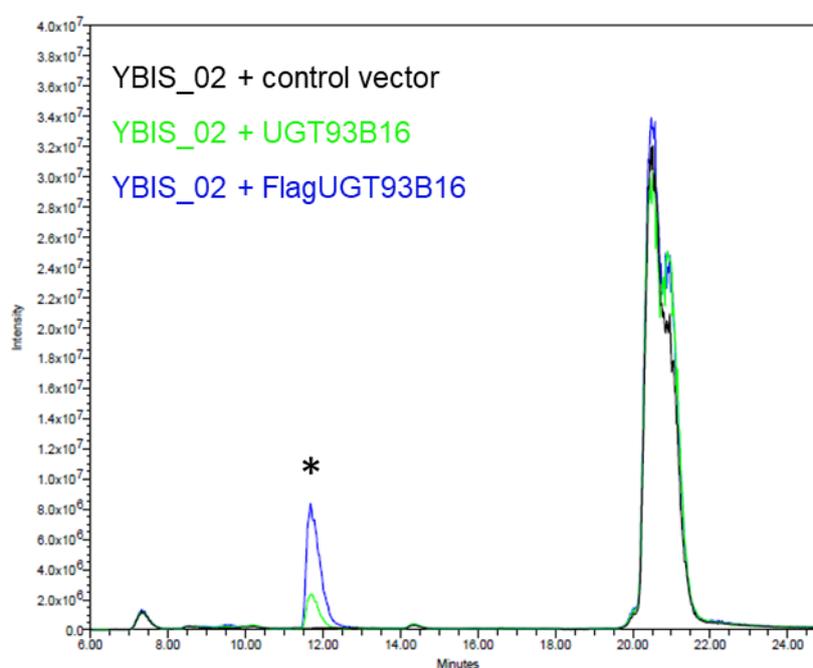


Figure 59: MS signal at  $m/z = 205.3$  showing YBIS\_02 *In vivo* sesquiterpenoid glucosylation mediated by UGT93B16  
\* corresponds to (+)-*epi*- $\alpha$ -bisabolol-Glc or farnesol-Glc formation as both sesquiterpenoids are produced in YBIS\_02. In addition to the ion at 205.3, the same peak yielded the 385.3 ion specific to glucosylated product (as shown in Figure 54)

d) Perspectives

The present study highlights the astonishing acceptor promiscuity of UGT93B16 towards sesquiterpene acceptors. Our results show that this enzyme can generate new glucoside derivatives of ( $\pm$ )-hernandulcin, (-)- $\alpha$ -bisabolol, trans,trans-farnesol, hydroxy-(+)-*epi*- $\alpha$ -bisabolol, nerolidol, and (+)-*epi*- $\alpha$ -

bisabolol. Sesquiterpene glycosides may possess physico-chemical and biological properties which could be of interest for different types of applications. For example,  $\alpha$ -bisabolol is widely used in cosmetic creams due to its anti-inflammatory effects, although several studies report on allergenic side effects that may be encountered<sup>250,378,379</sup>. Glucosylation of this molecule could limit its allergenicity and, at the same time, could improve its solubility thus facilitating formulation. Another study also pointed that glucosylated  $\alpha$ -bisabolol as pharmaceutical potential as drug candidate<sup>362</sup>. Similarly, producing hernandulcin glucoside could allow the evaluation of its sweetening power compared to the aglycon and lead to a new type of sweetener.

We also showed that bioconversion of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin in their corresponding glucosides is higher in *S. cerevisiae* than in *E. coli*. Expression of the enzymes could probably be enhanced in both organisms, notably the Flag tagged UGT could be expressed in *E. coli*. Different types of promoters could also be used. We could look at the UDP-Glc donor supply and investigate the substrate uptake as well as the presence of competing pathways for the UDP-Glc pool to re-orientate the pool towards glucoside formation. In addition, the possible degradation of glucoside derivatives by endogenous glycoside hydrolases should be checked and deletion of the corresponding genes could be envisioned to favor the accumulation of the glucoside as previously evidenced for *in vivo* production of flavonoid glucoside in yeast<sup>380</sup>.

Focusing on the bioconversion of (-)- $\alpha$ -bisabolol in yeast, we have seen that with a FLAG tagged protein, we reached almost quantitative glucosylation of (-)- $\alpha$ -bisabolol (98.1%) and that was not the case with the native enzyme (41.9%). Looking at the expression of these enzymes at both the RNA and protein levels and comparing their catalytic efficiency ( $k_{cat}/K_m$ ) would be informative to conclude on the differences observed and on the role of the FLAG tag. In addition, *in vitro* enzymatic assays did not allow an efficient production of hernandulcin glucoside, and we have seen that, *in vivo*, ( $\pm$ )-hernandulcin could be glucosylated with a conversion of 28%. These results are surprising and could reflect a poor stability of the enzyme mainly impacting *in vitro* assay. Conversion yields obtained with higher amounts of UGT93B16 and enzyme stability measurements should help to test this hypothesis. Notably, *in vitro* assays with the tagged protein should also be carried out and compared with the results obtained by bioconversion.

Investigating the promiscuity of UGT93B16 on a broader array of sesquiterpenes and other potential acceptors namely zeatin, terpenes, polyphenols could also further demonstrate the potentialities of UGT93B16 compared to other UGTs. Solving and/or modelling the enzyme structure in complex with various acceptors could also be conducted to identify determinants possibly involved in the acceptor specificity. This could open the way to rational enzyme engineering to better control and improve the acceptor specificity as it was done for UGTs catalyzing steviol glycoside synthesis<sup>355</sup>. In parallel, testing the potential of other enzymes of family UGT93 could also help to identify other glycosyltransferases of interest for sesquiterpene glucosylation.

Finally, if UGT93B16 was shown to be highly promiscuous, the role of this enzyme *in planta* remains unknown. Surprisingly, no exhaustive analysis of the terpene variety present in oat is available. Thus, the acceptors tested in our study may or may not be natural substrates. Furthermore and to date, UGT93 family was exclusively characterized as a family comprising zeatin-O-glycosyltransferases<sup>330,331,381</sup>, zeatin being a cytokinin derived from adenine. A protein alignment of the characterized enzymes with UGT93B16 is shown in Figure 60. UGT93B16 is sharing a high similarity with the other representatives, with 67% identity to CZog1 (also termed as UGT93B1). An investigation assessing its activity *in vitro* using zeatin while using *in planta* cross validation could help elucidate its role *in vivo*. Alternatively, if zeatin is not a substrate this would be the first evidence of a neofunctionalization in UGT93 family and could highlight multiple roles of this UGT family in plants.

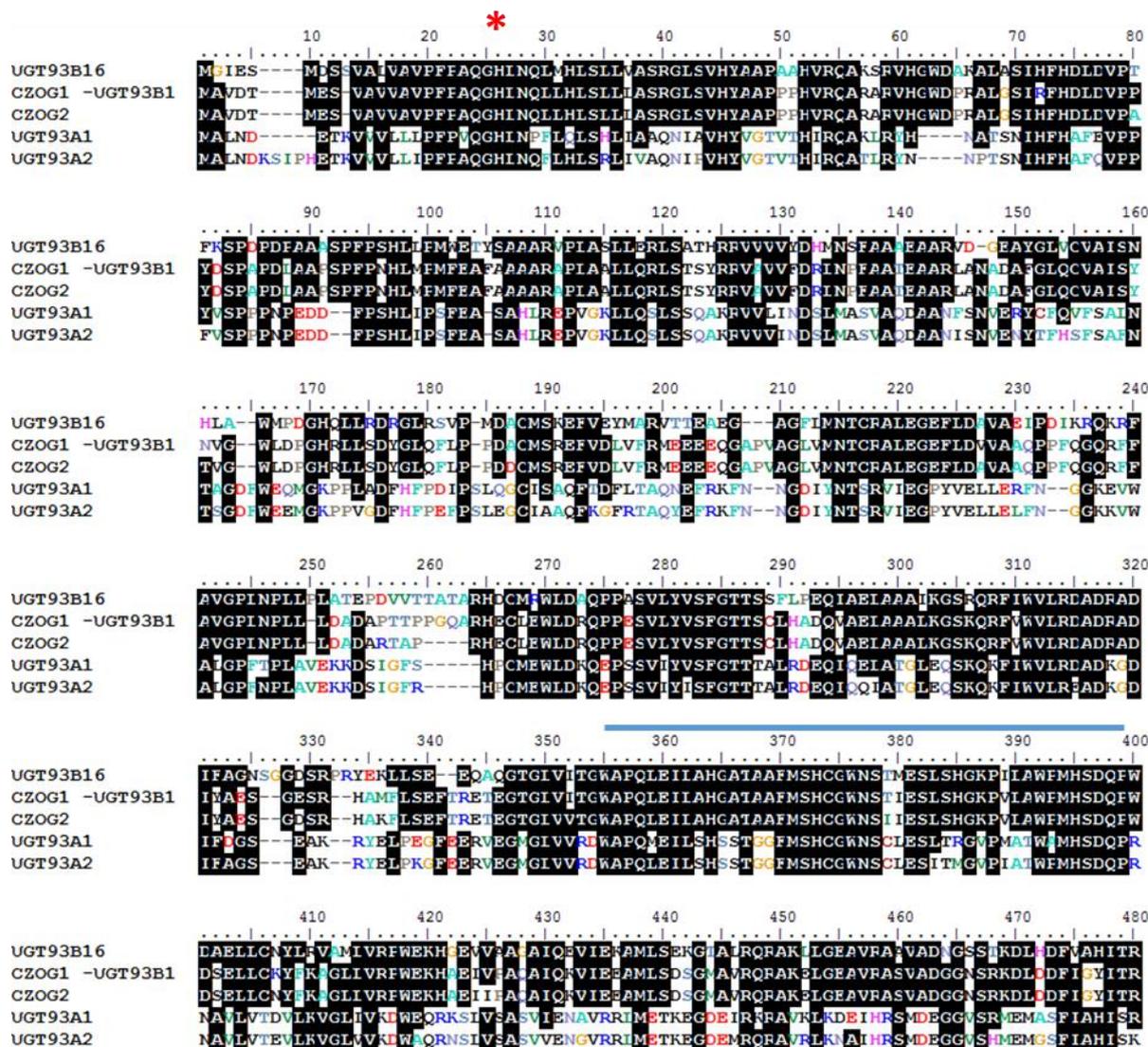


Figure 60: Protein sequence alignment of UGT93B16 and the other UGT93 currently characterized in literature. Sequence alignment was generated using ClustalW, from Bioedit. The catalytic histidine is marked with a red asterisk, and PSPG motif spans the region indicated with a blue line.

## 2. Human UDP-glucuronosyltransferases are able to convert (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin

### a) *In vitro* assay with Supersomes™ microsomal fractions

LC/MS analysis of the reactions performed with human UDP-glucuronosyltransferases incubated with (-)- $\alpha$ -bisabolol showed no detectable product formation in the UV signal, which maybe due to a poor extraction of  $\alpha$ -bisabolol. This may originate from the hydrophobicity of (-)- $\alpha$ -bisabolol that partly remained trapped in the Supersomes™ microsomal fractions. However, MS detection yielded a more sensitive detection of (-)- $\alpha$ -bisabolol, which was evidenced in control and in all the enzymatic assays. As previously indicated, (-)- $\alpha$ -bisabolol signal in MS was detected at  $m/z = 205.3$ , at a retention time of 20.8 minutes. Additionally, a more polar compound displaying signal at  $m/z = 205.3$  and eluting at 11.75 minutes was detected in the enzymatic assays containing UGT1A3, UGT1A4, UGT1A9, UGT2B4 or UGT2B7. This new compound could correspond to a rearrangement product of (-)- $\alpha$ -bisabolol-GlcA [ $M_{\text{bisabolol glucuronide}} - \text{GlcA} - \text{H}_2\text{O} + \text{H}$ ]<sup>+</sup> = 205.3. Indeed, a supplementary ion at  $m/z = 421.3$  corresponding to [ $M_{\text{bisabolol glucuronide}} + \text{Na}$ ]<sup>+</sup> = 421.3 was detected, thus strengthening the fact that it originates from (-)- $\alpha$ -bisabolol-GlcA fragmentation<sup>375</sup>. The corresponding MS chromatograms are shown in Figure 61.

HPLC analysis of the reaction performed with ( $\pm$ )-hernandulcin and UGT1A9, UGT2B4 or UGT2B7 indicates the presence of a new compound eluting at 10.42 min with a UV response comparable to that of hernandulcin with two maxima at 240 and 239 nm, respectively (Figure 62A, B, C). MS analysis further revealed a sodium adduct  $m/z = 435.3$  [ $M_{\text{hernandulcin glucuronide}} + \text{Na}$ ]<sup>+</sup> corresponding to the glucosylated form ( $\pm$ )-hernandulcin (Supplementary Figure 73). The results obtained with the various human UGTs are summarized in the Table 17. Interestingly, both human UDP-glucuronosyltransferases from family 1 and 2 can recognize (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin showing the flexibility of the phase II conjugation enzymes. Our next goal was to implement these enzymes in a yeast chassis for *in vivo* production of glucuronides. We pursued only with UGT1A9, UGT2B4 and UGT2B7, which glucosylated the two substrates.

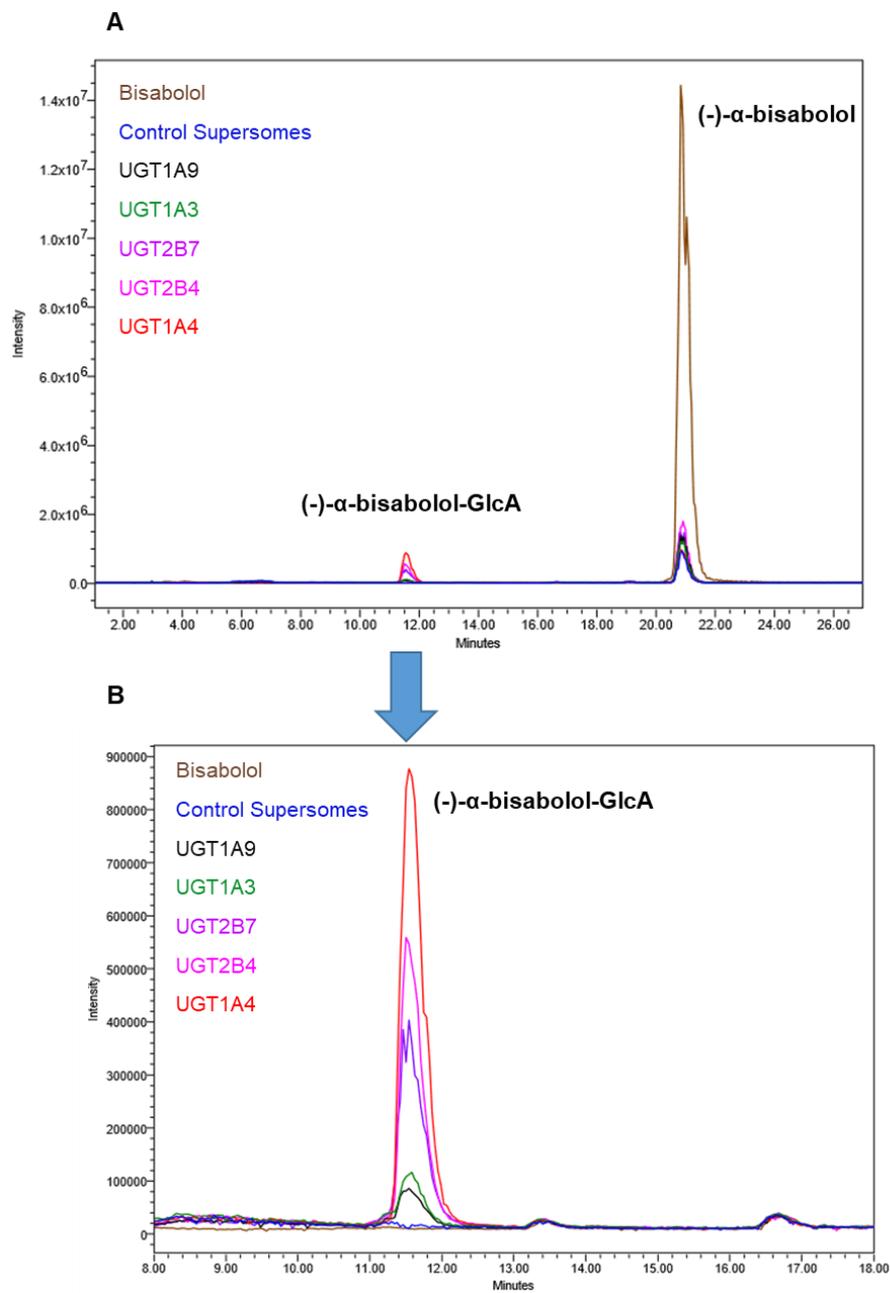


Figure 61: Human UDP-glucuronosyltransferase activities with (-)- $\alpha$ -bisabolol  
 (A) MS signal at  $m/z=205.3$  showing (-)- $\alpha$ -bisabolol substrate detection and the appearance of a more polar product: (-)- $\alpha$ -bisabolol-GlcA  
 (B) Zoom in the region of (-)- $\alpha$ -bisabolol-GlcA

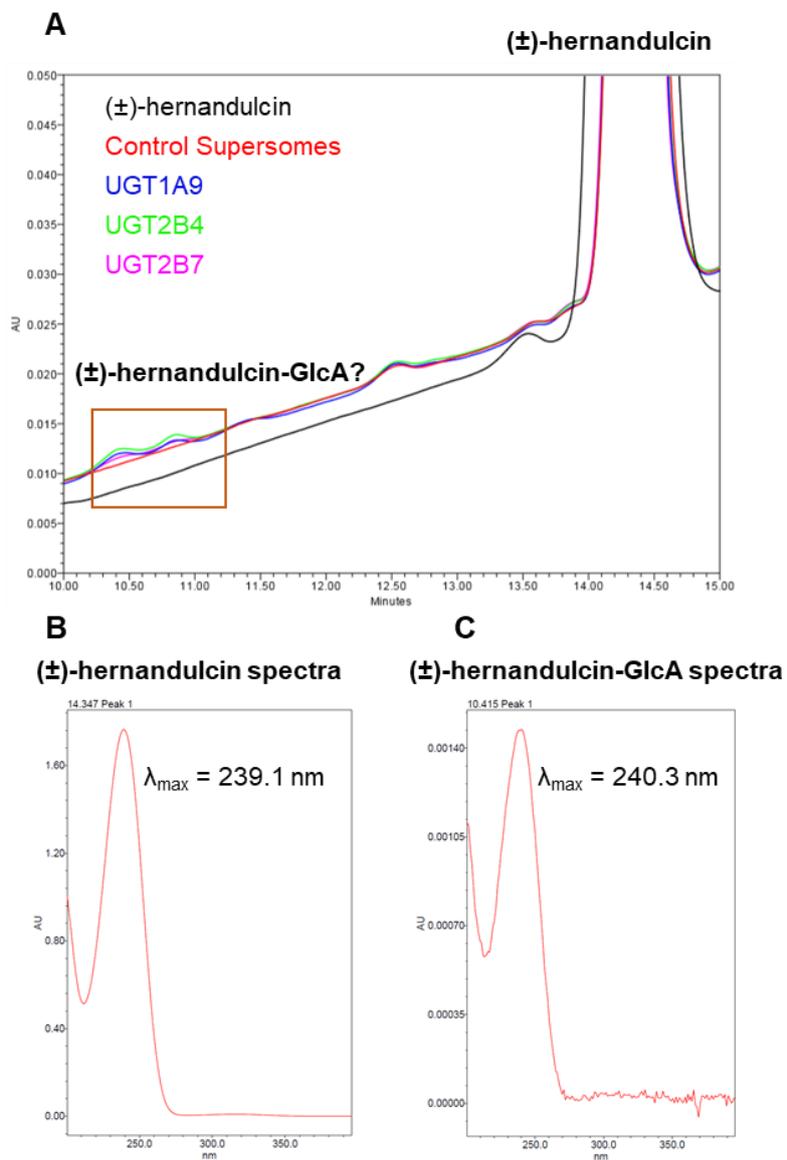


Figure 62: Human UDP-glucuronosyltransferase activities with (±)-hernandulcin  
 (A) UV signal at 234 nm showing (±)-hernandulcin substrate detection and the appearance of a more polar product (±)-hernandulcin-GlcA  
 (B) (±)-hernandulcin UV spectra from 200 to 400 nm  
 (C) (±)-hernandulcin-GlcA UV spectra from 200 to 400 nm

Table 17: Human UDP-glucuronosyltransferase activities detected with (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin

Substrate Enzyme	(-)- $\alpha$ -bisabolol	( $\pm$ )-hernandulcin
Control microsomes		
UGT1A1		
UGT1A3	+	
UGT1A4	+	
UGT1A6		
UGT1A7		
UGT1A8		
UGT1A9	+	+
UGT1A10		
UGT2B4	+	+
UGT2B7	+	+
UGT2B10		
UGT2B15		
UGT2B17		

(+) corresponds to a detected activity in MS detection

#### b) Attempts to reconstruct a functional pathway in yeast

We based our design on previous reports describing the expression of human UDP-glucuronosyltransferases in yeast<sup>372,373</sup>. We integrated the rat UGDH in *S. cerevisiae* genomes of BY4742 and YBIS\_01 leading to BY4742 + UGDH and YBIS\_01 + UGDH. The plasmids coding for UGT1A9 / UGT2B4 / UGT2B7 and the control plasmid (pYEDP60) were transformed in the two chassis strains. Bioconversion cultures were carried out with (-)- $\alpha$ -bisabolol or ( $\pm$ )-hernandulcin using the strains BY4742 + UGDH + human UDP-glucuronosyltransferases and the strains YBIS\_01 + UGDH + human UDP-glucuronosyltransferases. No signal corresponding to glucuronide products were identified using LC/MS analysis.

There is still a debate regarding the need of a UDP-GlcA transporter in human cell for optimal activity of the phase II conjugation system<sup>382</sup>. From the currently available literature, this transporter is supposed to be UGTRel7 and biochemical studies in *S. cerevisiae* indicate that it controls active transport of UDP-GlcA in microsomal fractions<sup>382,383</sup>. Interestingly, another group used UGTRel7 along with a UDP-glucuronosyltransferase from *A. thaliana* in *S. cerevisiae*. The addition of transporters UGTRel7 was crucial to

detect UDP-GlcA activity from the plant glucuronosyltransferase (Figure S1 of the reference)<sup>384</sup>. Based on this information, we add the human transporter, UGTRel7, or separately the homolog from drosophila, FRC (a good alternative as it was also characterized in *S. cerevisiae*<sup>385</sup>) in our constructs. Plasmids encoding each transporter (pFPP13 + UGTRel7 and pFPP13 + FRC) and those encoding human UDP-glucuronosyltransferases or control plasmid (pYEDP60, pYEDP60 + UGT1A9, pYEDP60 + UGT2B4, pYEDP60 + UGT2B7) were cotransformed in BY4742 + UGDH and YBIS\_01 + UGDH affording 16 new strains. To summarize, the tested conditions are given in Figure 63. Despite our efforts, we could not detect any traces of  $\alpha$ -bisabolol glucuronide.

### Strategies used for the *in vivo* assay of human glucuronosyltransferases

#### Bioconversion approach



#### Biosynthesis approach



 **No glucuronides detected in HPLC-UV-MS**

Figure 63: The approaches for *in vivo* production of glucuronides in *S. cerevisiae*. Human UDP-glucuronosyltransferases included UGT1A9 / UGT2B4 / UGT2B7 and control vector pYEDP60. For the first bioconversion experiment consisting in BY4742 + UGDH + human UDP-glucuronosyltransferase (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin were tested as exogenous added substrates. For the second set of experiments with BY4742 + UGDH + transporters + human UDP-glucuronosyltransferase only (-)- $\alpha$ -bisabolol was assayed.

### c) Perspectives

Despite many efforts and using strategies which were described as successful in the literature, we could not observe the production of glucuronide using bioconversion. In a first place, we have to check our genomic integrations in term of activity. The presence of UDP-GlcA donor must be verified in the strains BY4742 + UGDH and YBIS\_01 + UGDH. Similarly, the production of the human UGTs and UGTReI7 / FRC should also be checked by western blot. If the various enzymes and transporters are present and no glucuronide products are formed, the discrepancy could originate from the yeast genetic background AH22 vs BY4742, the expression promoter for human UDP-glucuronosyltransferases "GAPDH" vs GAL10-CYC1 hybrid promoter or the culture conditions <sup>373</sup>.

Otherwise, further adjustments of the system could include a truncation of the human UDP-glucuronosyltransferases on their N-ter to prevent their anchor to the lumen of endoplasmic reticulum by removing the final dilysine motif (Figure 49) <sup>386</sup>. Another alternative would be to try to "humanize" more the transport of glucuronide (Figure 64 I). In such endeavor, introducing an efflux system excreting the glucuronide from the cytosol might be advantageous as the glucuronide path from endoplasmic reticulum lumen towards cytosol remains unclear. Adding a last step of glucuronide efflux with a MRP or another transporter from the cytosol to the outside of yeast cells may be favorable (Figure 64 II) <sup>387</sup>. However, bringing more complexity to the system should be considered only after a careful control of our various constructs.

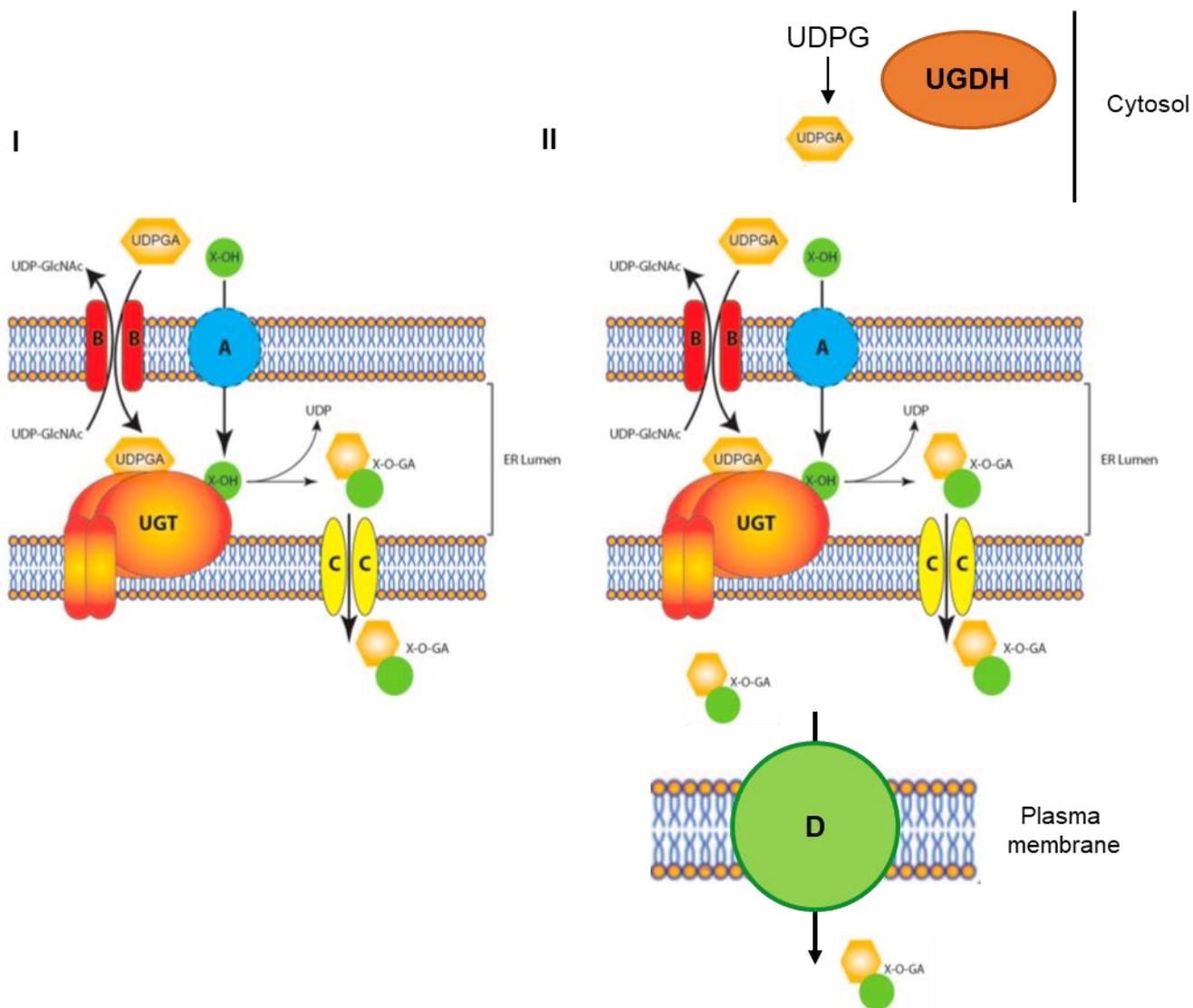


Figure 64: The glucuronidation system in the endoplasmic reticulum in human cells and a proposal for an improved system for glucuronide production in yeast based on a humanized efflux. Figure adapted from Liu et al<sup>382</sup>

(I) The glucuronidation system in the endoplasmic reticulum. (A) Substrates (X-OH) enter the lumen by diffusion. (B) UDPGA is transported via the UDPGA/UDP-GlcNAc antiporter (UGTrel7). (C) Following the conjugation reaction, the glucuronide products (X-O-GA) are removed from the lumen by glucuronide transporter(s). UDPGA: Uridine diphosphate-glucuronic acid; UDP-GlcNAc: UDP-N-acetylglucosamine.

(II) Towards a humanized efflux of glucuronides in yeast. UGDH is absent in the yeast system and has to be included. (A) Substrates (X-OH) diffusion in the lumen. (B) UGTrel7 or FRC expression. (C) The unknown or hypothetical efflux transporter in human could be absent in yeast and should be added when it will be unraveled. (D) Efflux transporter of MRP family could finally excrete the glucuronides outside of the cells.

## Conclusion

As a summary of our glycodiversification approach, we could produce 6 new sesquiterpene glycosides and 2 known products using both a plant UGT and human UDP-glucuronosyltransferases. The probable structures of the molecules are recapitulated in Figure 65. Plant UGT proved to be a versatile tool for such approach, easy to express in *E. coli* for enzymatic assays or *in vivo* bioconversion in *E. coli* or yeast. In contrast, human UDP-glucuronosyltransferases were less efficient *in vitro* and no traces of glycosylation could be detected *in vivo* with the engineered yeast. A way to enable sesquiterpene glucuronide production may lie on finding a plant UGT catalyst that accepts UDP-GlcA as sugar donor as proposed for the flavonoid breviscapine glucuronide<sup>388</sup>. Enzyme engineering could also be used to reprogram a plant UGT like UGT93B16 that already recognize  $\alpha$ -bisabolol and change its donor specificity as point mutations were shown to govern the UDP-sugar specificity from UDP-GlcA towards UDP-Glc in some UGT, and *vice versa*<sup>335</sup>.

## Acknowledgments

Laure Souilles is gratefully acknowledged for technical support in the bioconversion experiments with UGTs and for technical support for yeast strain construction.

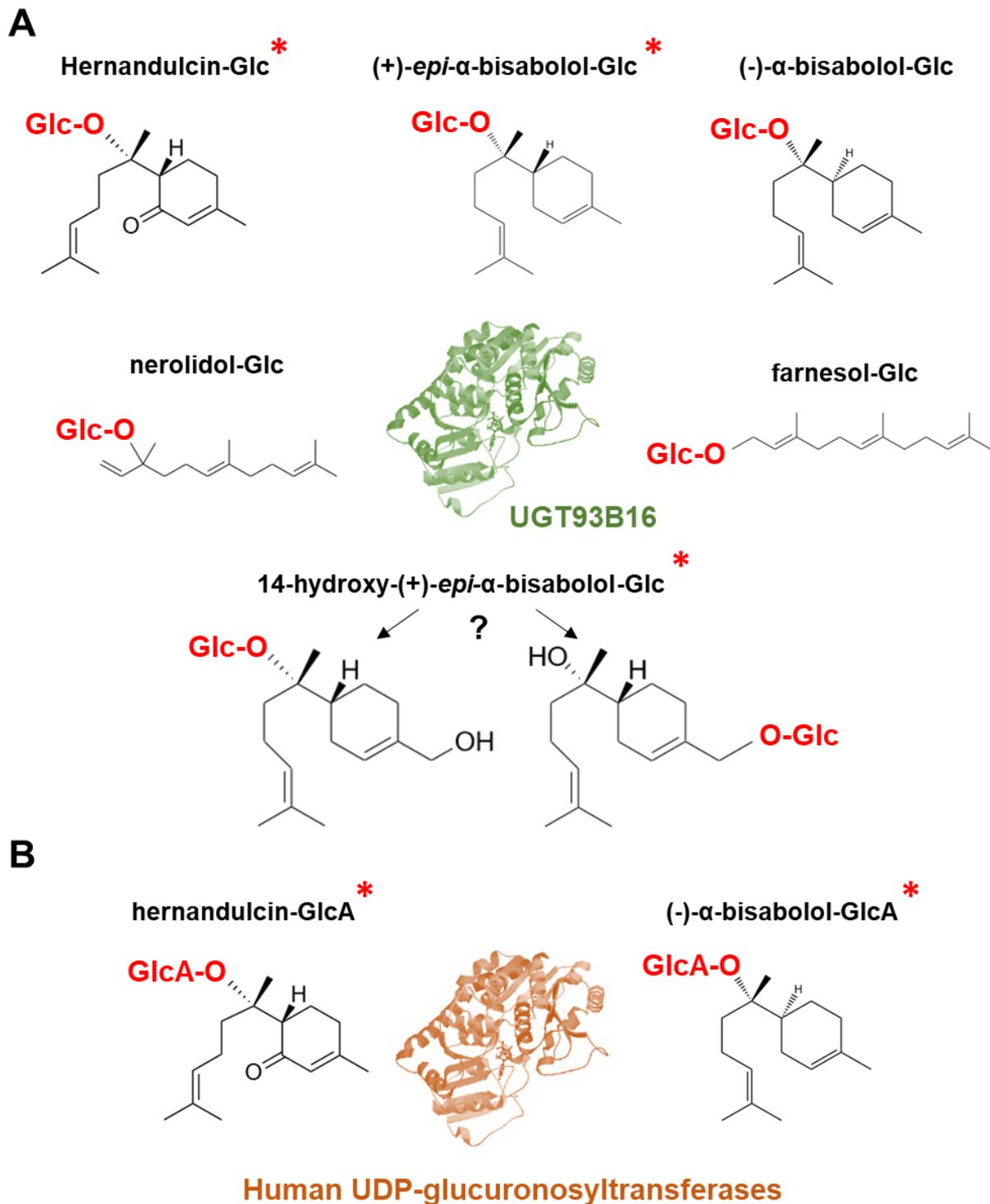


Figure 65: Summary of the generated products from glycosylation diversification

(A) The glycosides produced from UGT93B16

(B) The glycosides obtained from human UDP-glucuronosyltransferases

Most probably the glycosyltransferases activities are directed on OH groups of the aglycone. Products are represented using this hypothesis. For 14-hydroxy-(+)-*epi*- $\alpha$ -bisabolol two positions are possible leading to potentially two products while the glucosylation position requires additional characterization. The products for which no glycosyltransferase was characterized to date are indicated with a red asterisk. Farnesol-Glc was described in <sup>117</sup>, (-)- $\alpha$ -bisabolol-Glc preliminary results originate from Thomas Louveau's thesis, UGT91Q2 was recently shown to convert nerolidol <sup>118</sup>.



## Conclusion and Perspectives

We conducted approaches to enrich the known chemical repertoire of  $\alpha$ -bisabolol derivatives using enzyme promiscuity and screen small enzymes libraries. Our choice of enzymes was guided by the following reasoning: i) finding routes to oxidize (+)-*epi*- $\alpha$ -bisabolol using promiscuous oxidative enzymes and getting closer to the biosynthesis of hernandulcin or its analogs ii) mimicking the plant and human natural secondary metabolism to modify the hydrophobicity of these molecules using glycosyltransferases also described to be promiscuous. Our goal of modifying sesquiterpenes to synthesize new derivatives using CYPs was not previously reported and animal CYPs were overlooked in term of versatility. Similarly, only few enzymes were described to perform sesquiterpene glycosylation.

For (+)-*epi*- $\alpha$ -bisabolol oxidation, we followed a bottom up synthetic biology approach where a chassis strain previously built to overproduce sesquiterpenes in yeast served as a basis for the generation of a new platform. In this new strain, BBS and CPR were integrated in *S. cerevisiae* genome to produce (+)-*epi*- $\alpha$ -bisabolol *in vivo* and provide the redox partner of CYPs. From the validation of this new chassis, we were able to graft and screen more efficiently human CYPs for the synthesis of oxidized compounds. As CYPs are membrane bound enzymes, such strategy is facilitated compared to enzymatic assays where microsomal fractions are prepared at a lower throughput. We deliberately targeted CYPs from xenobiotic metabolism as no report on the repurpose of these enzymes in metabolic engineering towards the production of natural compound existed.

Overall, the main advances during this thesis includes:

- A critical concern was first the search for metabolic engineering strategies allowing to increase the bisabolol pool. We mapped side reaction that led to farnesol (and nerolidol when pH is acidic). We also pointed that increasing BBS expression lowers the farnesol side reaction. But, a tighter control of the MVA pathway to fully deplete farnesol from our strain is required. In this direction, BBS activity could be enhanced by additional genomic copies, or by engineering BBS for higher specific activity. A better regulation of the FPP node of the MVA may be critical too, a more dynamic regulation that would prevent FPP overflow could be set by changing the MET3 repressible promoter, and/or by regulating ERG9 at the protein level by degron as did Peng et al <sup>173</sup>.
- 5 out of 25 animal CYPs converted (+)-*epi*- $\alpha$ -bisabolol to at least 3 different products. We further considered plant CYPs, out of a small library of 20 CYPs (of which 12 were previously reported to accept sesquiterpenes as substrates), 4 exhibited *in vivo* activities never reported before towards (+)-*epi*- $\alpha$ -

bisabolol and/or farnesol. Structural characterization of the products obtained by plant CYPs awaits product purification and NMR analysis. Even though we did not find a CYP acting on the C-2 (+)-*epi*- $\alpha$ -bisabolol, screening a larger library of promiscuous animal CYPs could provide such catalyst. Alternatively, targeted enzyme engineering performed on the identified CYPs could be considered as a different strategy. Conversely, one could turn to enzyme discovery and screen *L. dulcis* CYPs, but this was not the direction of our work and enzyme discovery remains challenging in numerous natural pathways. Therefore, bringing another possibility by assaying xenobiotic metabolism could facilitate future reconstruction of natural pathways whose missing steps remain elusive.

- Being able to oxidize at least 3 positions of (+)-*epi*- $\alpha$ -bisabolol with a small set of animal CYPs underlines the potential of promiscuous CYPs for the production of new molecules. Moreover, the *in vitro* assays carried out with the 5 CYPs oxidizing (+)-*epi*- $\alpha$ -bisabolol, also demonstrated activities with farnesol, nerolidol and hernandulcin too. This evidences that a generalization of our strategy could be pursued with more diverse terpene scaffolds.
- Reciprocally, one could use farnesol as a substrate in itself from our chassis strain. Indeed, we showed that some CYPs recognize it as substrate *in vitro*. Farnesol is a precursor in some juvenile hormones from insects and these hormones represent a target to control insect populations<sup>389,390</sup>. Analogues of such hormones may alter some growth stages of various insects. Then, such new molecules could be assayed with respect to putative activities for insect biocontrol.
- By focusing on UGT93B16 for glucosylation (whose identification by Louveau et al. showed preliminary results with  $\alpha$ -bisabolol), we confirmed that this UGT is active with  $\alpha$ -bisabolol. We also proved the promiscuity of UGT93B16 on sesquiterpenes by finding that hernandulcin, farnesol, nerolidol, and hydroxylated (+)-*epi*- $\alpha$ -bisabolol are also substrates. Comparison of *E. coli* and *S. cerevisiae* as bioconversion hosts for UGT93B16 showed a better suitability of *S. cerevisiae*. *In vivo*, comparison of UGT93B16 with a flag-tagged version revealed that conversion with the tagged enzyme largely outcompeted the native one, which may be attributed to improved expression, catalytic efficiency or stability due to the presence of the tag. *In vitro* biochemical characterization of both forms is needed to conclude between these various assumptions.
- For glucuronidation, human glucuronosyltransferases were considered. Out of 13 human isoforms, 3 converted both (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin. However, yeast expression of these human glucuronosyltransferases failed to produce glucuronides in bioconversion experiments.

- During the search of other relevant enzymes able to yield supplemental (+)-*epi*- $\alpha$ -bisabolol derivatives some may fail to function into a chassis strain. This was evidenced with human glucuronosyltransferases. As a consequence, we may lack the proper understanding of the enzyme biochemical requirements to carry out the expected activity *in vivo*. Therefore, one can argue that basic knowledge on the enzyme requirements for a good functioning *in vivo* - in terms of cofactor, membrane anchoring, efflux transporters - is a prerequisite for efficient metabolic engineering.

## 1. Towards *in vitro* cascades to prototype new pathways

Preliminary testing of 7 ADHs suggests that low *in vivo* activity with hydroxylated (+)-*epi*- $\alpha$ -bisabolol occurs. Further *in vitro* enzymatic assays could confirm these preliminary hints. Screening additional ADHs with a chemical standard of 2-hydroxy-(+)-*epi*- $\alpha$ -bisabolol, the hernandulcin possible precursor, could represent a new route when a CYP hydroxylating the C-2 of (+)-*epi*- $\alpha$ -bisabolol will be discovered.

While *in vivo* screening in *S. cerevisiae* of single CYPs was fruitful, an alternative and complementary approach would be to set up *in vitro* enzymatic cascades for coupling multiple steps. Hence, testing enzymatic cascades (CYPs / ADHs / UGTs) *in vitro* with commercially available sesquiterpenes could be envisioned prior to their integration into yeast chassis. This could avoid the risk of false positive results when *in vivo* side reactions occur and the need of testing multiple control strains. When CYPs + ADHs were directly screened *in vivo*, controls with strains devoted of each ADH and CYPs has to be done for the corresponding combination. Moreover, the modularity of *in vitro* enzymatic cascades could also question whether human glucuronosyltransferases could act before a CYP step on sesquiterpenes. This could seem counterintuitive for plant enzymes, but human ones are able to deal with substrates showing such differences in size and hydrophobicity that this could hypothetically happen.

## 2. Possible extension of the approach to other enzyme classes

Other catalyst classes can also be envisioned as we better understand the production of (+)-*epi*- $\alpha$ -bisabolol derivatives in yeast. We showed the existence of side reactions and established metabolite extraction and analysis protocols that allow structure determination of new metabolites. This extension towards other classes of enzymes might be limited by several factors. The throughput of screening which is

applicable may represent the major one. If LC/GC is used the throughput is limited compared to methods involving reporter assays, absorbance or fluorescence outputs. If higher throughput can be obtained, the next barrier may be the intrinsic promiscuity of an enzyme class. We targeted notoriously promiscuous enzymes or enzymes known to accept some sesquiterpenes. If less versatile catalysts are screened, the outcomes may be limited.

Regarding the choice of oxidative enzymes, one could imagine to probe different classes of oxidases and not only CYPs and ADHs. For instance, a limited number of examples have proved that laccases, peroxidases or dioxygenases can be promiscuous too and oxidize some sesquiterpenes such as valencene to nootkatone, a compound of interest for fragrance and flavor industries <sup>391,392</sup>. Concerning the enzymes of interest to balance hydrophobicity, again, completely different classes of enzymes such as sulfotransferases could be envisioned <sup>393</sup>. Modification through sulfation was inspired by natural pathways in which phenolic derivatives are sulfated. As evidenced by Jendersen et al, sulfation had a dual effect in their microbial cell factories increasing the solubility of their phenolic products and alleviating toxicity of the aglycone.

### 3. The coupling of CYP and UGT steps

As mammalian CYPs are active using (+)-*epi*- $\alpha$ -bisabolol and UGT93B16 demonstrates activity with hydroxylated (+)-*epi*- $\alpha$ -bisabolol derivatives, the *in vivo* coupling of the two steps is feasible. However, as UGT93B16 and CYPs are active using (+)-*epi*- $\alpha$ -bisabolol, a competition for the substrate would probably exist between the two enzymes and UGT93B16 would not necessarily act on hydroxylated (+)-*epi*- $\alpha$ -bisabolol derivatives. In addition, we observed that the hydroxylated (+)-*epi*- $\alpha$ -bisabolol derivatives were excreted from yeast cells (Figure 28). This sets a very interesting situation for a compartmentation strategy where UGT93B16 could be addressed to the periplasmic space of yeast where a UDP-Glc pool exists for cell wall synthesis <sup>394,395</sup>. This could compartmentalize CYP oxidation (inside the yeast cell) from UGT glucosylation of the hydroxylated (+)-*epi*- $\alpha$ -bisabolol derivatives to the outside. Such compartmentalization could be complementary to a condition in absence of particular targeting or to more usual strategies consisting in enzyme fusions or scaffolding. Interestingly, there are no reports of CYP-UGT fusions even if the steps are consecutive in some plant pathways. This could assist in metabolite channeling when such enzymes are used in heterologous host like *S. cerevisiae*. Alternatively, a protein fusion consisting in BBS-UGT93B16 to produce (+)-*epi*- $\alpha$ -bisabolol could favor higher titers of the end product by avoiding possible escape of (+)-*epi*- $\alpha$ -bisabolol intermediate to membranes (farnesyl-pyrophosphate, the BBS substrate is cytosolic / amphiphilic, the (+)-*epi*- $\alpha$ -bisabolol intermediate is hydrophobic and the final product using a UGT is of intermediate amphiphilicity).

#### 4. The biological activities of the new molecules

Having reached the production of new molecules, another challenge lies in the evaluation of the new derivatives potential. As hernandulcin is reported to be a potent sweetener a thousand times sweeter than sucrose, we could first investigate the features of the new hydroxylated and glycosylated derivatives of  $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin. This ability can be probed without having to physically taste the molecule - whose unknown toxicity may hinder fast evaluation of the new compounds. Human sweet taste perception is mediated by an heterodimeric G protein-coupled receptor heterodimer (T1R2-T1R3) Figure 66 A <sup>396</sup>. The signaling cascade of T1R2-T1R3 heterodimer leads to a rise of intracellular  $[Ca^{2+}]$ . And calcium imaging is a technical tool that utilizes a fluorescent probe which specifically responds to  $[Ca^{2+}]$ . Building on that, multiple groups reported cellular based experiments that enabled the affinity measurement of known sweeteners. Some teams even pushed further the biochemical characterization by attributing which domain of T1R2-T1R3 interacts with some sweeteners <sup>396-400</sup> (Figure 66 B, summarizes the current knowledge about the interacting sites of various molecules).

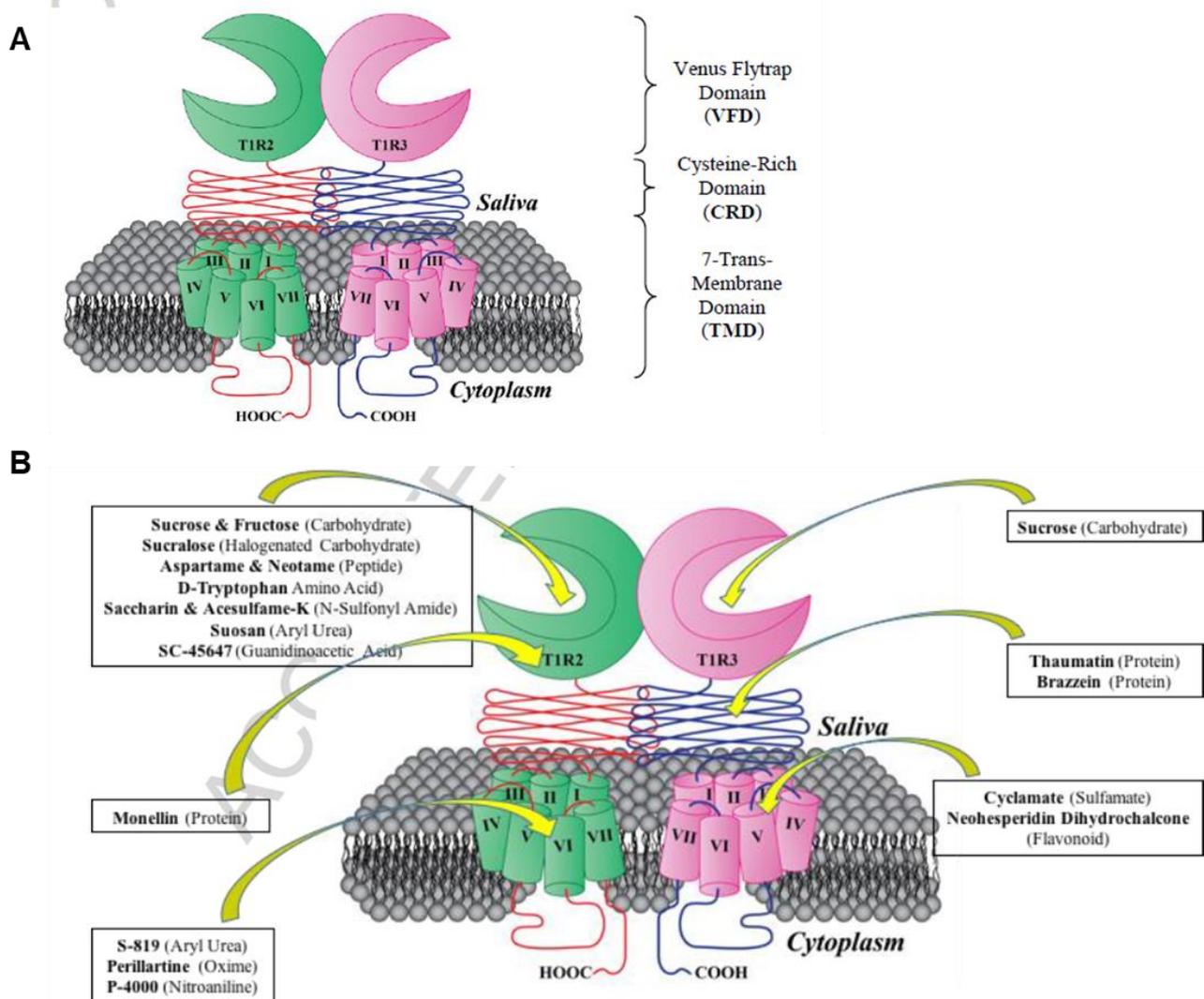


Figure 66: The human sweet taste receptor heterodimer T1R2-T1R3 topology (A) and the binding loci of various sweeteners (B), this figure is reproduced from Dubois et al <sup>401</sup>

Thus, we initiated a collaboration with the team of Loïc Briand in established in the “Centre des Sciences du Goût et de l’Alimentation” an INRAE Institute that studies the human sweet taste receptor. We first undertook a comparison of the response of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin. Doing so, we aimed at verifying that ( $\pm$ )-hernandulcin activates the human sweet taste receptor while (-)- $\alpha$ -bisabolol is not reported to be sweet and would constitute a negative control. The experiments were performed as described in appendix Figure 74 by Christine Belloir in the team of Loïc Briand. Preliminary results are shown in Figure 75. Sucralose, a well-known sweetener, was included as a positive control and demonstrated a half maximal effective concentration ( $EC_{50}$ ) of 49  $\mu$ M, in the range of literature values <sup>397–399</sup>. ( $\pm$ )-hernandulcin exhibited an activation of T1R2-T1R3 receptor with a high affinity ( $EC_{50}$  of 2.7  $\mu$ M), validating the described sweetening property of the molecule. In contrast, (-)- $\alpha$ -bisabolol did not activate the receptor and even had an artefactual response with a diminished fluorescence response compared to the control strain at concentrations higher

than 100  $\mu\text{M}$  (Figure 75C). Therefore, the various metabolites corresponding to hydroxylated and glycosylated derivatives of  $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin can now be compared in the same conditions and reveal whether they are able to activate T1R2-T1R3 receptor. To speed up the process and eventually avoid tedious purification steps of the molecules, some compounds could be tested as mixture (for example the CYP2B6 / CYP2B11 hydroxylated metabolites of (+)-*epi*- $\alpha$ -bisabolol). In addition, studying which site of T1R2-T1R3 receptor interacts with hernandulcin and/or derivatives could be envisioned as this sweetener was not linked to a specific target yet (Figure 75B).

Another output could be the study of structure/activity relationships between the close derivatives of (+)-*epi*- $\alpha$ -bisabolol and hernandulcin and T1R2-T1R3 receptor. A recent model also addressed this question and even proposed the prediction of the sweetening potential of a given molecule <sup>402</sup>. Benchmarking the results from biological assays to the predicted values could provide information regarding the prediction sweetness and the validity for sesquiterpene derivatives. Ultimately, if some molecules have an interesting behavior with T1R2-T1R3 receptor *in vitro*, their taste profile could be addressed in human trials (once the molecules proved to be safe), and the bitter aftertaste evaluated *in vitro* too <sup>403</sup>.

Sweetness potency constitutes only one out of the possible biological activities of our new molecules. In the literature,  $\alpha$ -bisabolol possesses anti-inflammatory effects and potentiates antibiotic efficacy, complementary studies on these leads are promising <sup>250,258</sup>. Additionally, nothing restricts the fact that other abilities emerged as sesquiterpene scaffolds hold a tremendous amount of biological activities. The major hurdle for the discovery of biological activities is the implementation of relevant screening assays.

A pioneering work in this direction was accomplished within a project led by Evolva in a series of groundbreaking achievements <sup>404,405</sup>. At first, a combinatorial tool relying on yeast artificial chromosomes was set <sup>404</sup>. They focused on the flavonoid pathway of plants and mixed 5 to 7 individual steps randomly on a yeast artificial chromosome. Doing so, large combinatorial libraries can be obtained. Then, implementation of a functional screen in yeast based on the growth inhibition of Brome Mosaic Virus was coupled to their combinatorial libraries <sup>405</sup>. Brome Mosaic Virus was used as a reporter for inhibition of human infectious viruses like dengue or hepatitis C. This allows at the same time to: i). explore wide range of gene combinations (8 plants pathway were considered including alkaloid, flavonoid and terpene pathways) ii). select only the strains that produces compounds able to inactivate Brome Mosaic Virus. They reported 74 novel compounds originating from 35 strains using this strategy and 28 molecules inhibited Brome Mosaic Virus in secondary screening assay.

Similar strategies could be built towards sesquiterpenes whose biosynthetic genes have been largely characterized to date. A design could be to split the various enzyme classes into different modules such as scaffold synthesis by sesquiterpene synthases, CYPs decoration by a large collection of plant / animal /

bacterial enzymes and to express multiple CYPs per sesquiterpene synthase to rise the chance of hydroxylation and favor multiple oxidations on a same scaffold, a last module comprising decorative enzymes involving ADHs, UGTs, sulfotransferases... Functional screening would then refine the selection of valuable compounds. This would speed considerably the process, but with risk of missing the information about the production of new molecules if they do not exhibit the desired feature. This represents the trade-off between larger throughput / combinations versus lower throughput HPLC / GC means of detection (but HPLC / GC can assess more precisely the range of new molecules synthesized in the process).

In conclusion to this thesis, both a deeper functional analysis of the newly generated molecules from the  $\alpha$ -bisabolol scaffold is considered along with the development of upgraded chassis strains whose regulation and production of the multiple derivatives will be more efficiently mastered.

## References

- (1) Compadre, C. M., Pezzuto, J. M., Kinghorn, A. D., and Kamath, S. K. (1985) Hernandulcin: an intensely sweet compound discovered by review of ancient literature. *Science* 227, 417–419.
- (2) Mori, K., and Kato, M. (1986) Synthesis of (6S, 1'S)-(+)-hernandulcin, a sweetener, and its stereoisomers. *Tetrahedron* 42, 5895–5900.
- (3) Compadre, C. M., Hussain, R. A., Lopez de Compadre, R. L., Pezzuto, J. M., and Kinghorn, A. D. (1987) The intensely sweet sesquiterpene hernandulcin: isolation, synthesis, characterization, and preliminary safety evaluation. *J. Agric. Food Chem.* 35, 273–279.
- (4) Adams, R. P., Weerasooriya, A., and Gao, M. (2014) Comparison of volatile leaf terpenoids from *Lippia dulcis* (Verbenaceae) obtained by steam distillation and pentane liquid extraction. *Phytologia* 96, 252–259.
- (5) Arendt, P., Pollier, J., Callewaert, N., and Goossens, A. (2016) Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. *Plant J.* 87, 16–37.
- (6) Attia, M., Kim, S.-U., and Ro, D.-K. (2012) Molecular cloning and characterization of (+)-epi- $\alpha$ -bisabolol synthase, catalyzing the first step in the biosynthesis of the natural sweetener, hernandulcin, in *Lippia dulcis*. *Arch. Biochem. Biophys.* 527, 37–44.
- (7) Okamoto, S., Yu, F., Harada, H., Okajima, T., Hattan, J., Misawa, N., and Utsumi, R. (2011) A short-chain dehydrogenase involved in terpene metabolism from *Zingiber zerumbet*. *FEBS J.* 278, 2892–2900.
- (8) Schuler, M. A., Duan, H., Bilgin, M., and Ali, S. (2006) Arabidopsis cytochrome P450s through the looking glass: a window on plant biochemistry. *Phytochem. Rev.* 5, 205–237.
- (9) Diaz-Chavez, M. L., Moniodis, J., Madilao, L. L., Jancsik, S., Keeling, C. I., Barbour, E. L., Ghisalberti, E. L., Plummer, J. A., Jones, C. G., and Bohlmann, J. (2013) Biosynthesis of Sandalwood Oil: *Santalum album* CYP76F Cytochromes P450 Produce Santalols and Bergamotol. *PLoS ONE* (Cirino, P. C., Ed.) 8, e75053.
- (10) Orme, A., Louveau, T., Stephenson, M. J., Appelhagen, I., Melton, R., Cheema, J., Li, Y., Zhao, Q., Zhang, L., Fan, D., Tian, Q., Vickerstaff, R. J., Langdon, T., Han, B., and Osbourn, A. (2019) A noncanonical vacuolar sugar transferase required for biosynthesis of antimicrobial defense compounds in oat. *Proc. Natl. Acad. Sci. U. S. A.*
- (11) Hsu, T. M., Welner, D. H., Russ, Z. N., Cervantes, B., Prathuri, R. L., Adams, P. D., and Dueber, J. E. (2018) Employing a biochemical protecting group for a sustainable indigo dyeing strategy. *Nat. Chem. Biol.* 14, 256–261.
- (12) de Kraker, J.-W., Schurink, M., Franssen, M. C. R., König, W. A., de Groot, A., and Bouwmeester, H. J. (2003) Hydroxylation of sesquiterpenes by enzymes from chicory (*Cichorium intybus* L.) roots. *Tetrahedron* 59, 409–418.
- (13) Cankar, K., van Houwelingen, A., Bosch, D., Sonke, T., Bouwmeester, H., and Beekwilder, J. (2011) A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS Lett.* 585, 178–182.
- (14) Speed, M. P., Fenton, A., Jones, M. G., Ruxton, G. D., and Brockhurst, M. A. (2015) Coevolution can explain defensive secondary metabolite diversity in plants. *New Phytol.* 208, 1251–1263.
- (15) Zanger, U. M., and Schwab, M. (2013) Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* 138, 103–141.
- (16) Vickers, C. E., Williams, T. C., Peng, B., and Cherry, J. (2017) Recent advances in synthetic biology for engineering isoprenoid production in yeast. *Curr. Opin. Chem. Biol.* 40, 47–56.
- (17) Andersen-Ranberg, J., Kongstad, K. T., Nielsen, M. T., Jensen, N. B., Pateraki, I., Bach, S. S., Hamberger, B., Zerbe, P., Staerk, D., Bohlmann, J., Møller, B. L., and Hamberger, B. (2016) Expanding the Landscape of Diterpene Structural Diversity through Stereochemically Controlled Combinatorial Biosynthesis. *Angew. Chem. Int. Ed.* 55, 2142–2146.
- (18) Moses, T., Pollier, J., Almagro, L., Buyst, D., Van Montagu, M., Pedreño, M. A., Martins, J. C., Thevelein, J. M., and Goossens, A. (2014) Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16 $\alpha$  hydroxylase from *Bupleurum falcatum*. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1634–1639.
- (19) Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Dupont, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R.,

- Pompon, D., and Dumas, B. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat. Biotechnol.* *21*, 143–149.
- (20) Ko, J. K., Um, Y., Woo, H. M., Kim, K. H., and Lee, S.-M. (2016) Ethanol production from lignocellulosic hydrolysates using engineered *Saccharomyces cerevisiae* harboring xylose isomerase-based pathway. *Bioresour. Technol.* *209*, 290–296.
- (21) Zhang, Y., Nielsen, J., and Liu, Z. (2017) Engineering yeast metabolism for production of terpenoids for use as perfume ingredients, pharmaceuticals and biofuels. *FEMS Yeast Res.* *17*.
- (22) Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievens, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., and Newman, J. D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* *496*, 528–532.
- (23) Meadows, A. L., Hawkins, K. M., Tsegaye, Y., Antipov, E., Kim, Y., Raetz, L., Dahl, R. H., Tai, A., Mahatdejkul-Meadows, T., Xu, L., Zhao, L., Dasika, M. S., Murarka, A., Lenihan, J., Eng, D., Leng, J. S., Liu, C.-L., Wenger, J. W., Jiang, H., Chao, L., Westfall, P., Lai, J., Ganesan, S., Jackson, P., Mans, R., Platt, D., Reeves, C. D., Saija, P. R., Wichmann, G., Holmes, V. F., Benjamin, K., Hill, P. W., Gardner, T. S., and Tsong, A. E. (2016) Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* *537*, 694–697.
- (24) Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* *440*, 940–943.
- (25) Nguyen, T.-D., MacNevin, G., and Ro, D.-K. (2012) De Novo Synthesis of High-Value Plant Sesquiterpenoids in Yeast, in *Methods in Enzymology*, pp 261–278. Elsevier.
- (26) Louveau, T., and Osbourn, A. (2019) The Sweet Side of Plant-Specialized Metabolism. *Cold Spring Harb. Perspect. Biol.* a034744.
- (27) Xie, K., Chen, R., Li, J., Wang, R., Chen, D., Dou, X., and Dai, J. (2014) Exploring the Catalytic Promiscuity of a New Glycosyltransferase from *Carthamus tinctorius*. *Org. Lett.* *16*, 4874–4877.
- (28) Almazroo, O. A., Miah, M. K., and Venkataramanan, R. (2017) Drug Metabolism in the Liver. *Clin. Liver Dis.* *21*, 1–20.
- (29) Kessler, A., and Kalske, A. (2018) Plant Secondary Metabolite Diversity and Species Interactions. *Annu. Rev. Ecol. Evol. Syst.* *49*, 115–138.
- (30) Alseekh, S., and Fernie, A. R. (2018) Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant J.* *94*, 933–942.
- (31) Dudareva, N., Klempien, A., Muhlemann, J. K., and Kaplan, I. (2013) Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* *198*, 16–32.
- (32) Liao, P., Hemmerlin, A., Bach, T. J., and Chye, M.-L. (2016) The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* *34*, 697–713.
- (33) Platt, J. R. (1959) Carotene-Donor-Acceptor Complexes in Photosynthesis: The predicted lowering of the excited states of carotenoids may offer a new photosynthetic pathway. *Science* *129*, 372–374.
- (34) Ruban, A. V., Berera, R., Illoiaia, C., Stokkum, I. H. M. van, Kennis, J. T. M., Pascal, A. A., Amerongen, H. van, Robert, B., Horton, P., and Grondelle, R. van. (2007) Identification of a mechanism of photoprotective energy dissipation in higher plants. *Nature* *450*, 575–578.
- (35) Ruzicka, L. (1953) The isoprene rule and the biogenesis of terpenic compounds. *Experientia* *9*, 357–367.
- (36) Kushiro, T., Shibuya, M., and Ebizuka, Y. (1998) Beta-amyrin synthase—cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur. J. Biochem. FEBS* *256*, 238–44.
- (37) Shao, H., He, X., Achnine, L., Blount, J. W., Dixon, R. A., and Wang, X. (2005) Crystal Structures of a Multifunctional Triterpene/Flavonoid Glycosyltransferase from *Medicago truncatula*. *Plant Cell* *17*, 3141–3154.

- (38) Madsen, K. M., Udatha, G. D. B. R. K., Semba, S., Otero, J. M., Koetter, P., Nielsen, J., Ebizuka, Y., Kushiro, T., and Panagiotou, G. (2011) Linking Genotype and Phenotype of *Saccharomyces cerevisiae* Strains Reveals Metabolic Engineering Targets and Leads to Triterpene Hyper-Producers. *PLoS ONE* 6.
- (39) Achnine, L., Huhman, D. V., Farag, M. A., Sumner, L. W., Blount, J. W., and Dixon, R. A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J. Cell Mol. Biol.* 41, 875–887.
- (40) Lee, A.-R., Kwon, M., Kang, M.-K., Kim, J., Kim, S.-U., and Ro, D.-K. (2019) Increased sesqui- and triterpene production by co-expression of HMG-CoA reductase and biotin carboxyl carrier protein in tobacco (*Nicotiana benthamiana*). *Metab. Eng.* 52, 20–28.
- (41) Huang, A. C., Hong, Y. J., Bond, A. D., Tantillo, D. J., and Osbourn, A. (2018) Diverged Plant Terpene Synthases Reroute the Carbocation Cyclization Path towards the Formation of Unprecedented 6/11/5 and 6/6/7/5 Sesterterpene Scaffolds. *Angew. Chem. Int. Ed.* 57, 1291–1295.
- (42) Long, C., Marcourt, L., Raux, R., David, B., Gau, C., Menendez, C., Gao, M., Laroche, M.-F., Schambel, P., Delaude, C., Ausseil, F., Lavaud, C., and Massiot, G. (2009) Meroterpenes from *Dichrostachys cinerea* Inhibit Protein Farnesyl Transferase Activity. *J. Nat. Prod.* 72, 1804–1815.
- (43) Luo, X., Reiter, M. A., d’Espaux, L., Wong, J., Denby, C. M., Lechner, A., Zhang, Y., Grzybowski, A. T., Harth, S., Lin, W., Lee, H., Yu, C., Shin, J., Deng, K., Benites, V. T., Wang, G., Baidoo, E. E. K., Chen, Y., Dev, I., Petzold, C. J., and Keasling, J. D. (2019) Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 567, 123–126.
- (44) Li, H., Ban, Z., Qin, H., Ma, L., King, A. J., and Wang, G. (2015) A Heteromeric Membrane-Bound Prenyltransferase Complex from Hop Catalyzes Three Sequential Aromatic Prenylations in the Bitter Acid Pathway. *Plant Physiol.* 167, 650–659.
- (45) Brown, S., Clastre, M., Courdavault, V., and O’Connor, S. E. (2015) De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci.* 112, 3205–3210.
- (46) Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D., and Eyal, Y. (2008) Peroxisomal Localization of Arabidopsis Isopentenyl Diphosphate Isomerases Suggests That Part of the Plant Isoprenoid Mevalonic Acid Pathway Is Compartmentalized to Peroxisomes. *Plant Physiol.* 148, 1219–1228.
- (47) Moses, T., Pollier, J., Thevelein, J. M., and Goossens, A. (2013) Bioengineering of plant (tri)terpenoids: from metabolic engineering of plants to synthetic biology in vivo and in vitro. *New Phytol.*
- (48) Simkin, A. J., Guirimand, G., Papon, N., Courdavault, V., Thabet, I., Ginis, O., Bouzid, S., Giglioli-Guivarc’h, N., and Clastre, M. (2011) Peroxisomal localisation of the final steps of the mevalonic acid pathway in planta. *Planta* 234, 903–914.
- (49) Guirimand, G., Guihur, A., Phillips, M. A., Oudin, A., Glévarec, G., Melin, C., Papon, N., Clastre, M., St-Pierre, B., Rodríguez-Concepción, M., Burlat, V., and Courdavault, V. (2012) A single gene encodes isopentenyl diphosphate isomerase isoforms targeted to plastids, mitochondria and peroxisomes in *Catharanthus roseus*. *Plant Mol. Biol.* 79, 443–459.
- (50) Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G. E., Rasche, N., Lüder, F., Weckwerth, W., and Jahn, O. (2007) Proteome Analysis of Arabidopsis Leaf Peroxisomes Reveals Novel Targeting Peptides, Metabolic Pathways, and Defense Mechanisms. *Plant Cell* 19, 3170–3193.
- (51) McFadden, G. I. (2014) Origin and Evolution of Plastids and Photosynthesis in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6.
- (52) R. Herrera, J. B., Bartel, B., Wilson, W. K., and Matsuda, S. P. T. (1998) Cloning and characterization of the Arabidopsis thaliana lupeol synthase gene. *Phytochemistry* 49, 1905–1911.
- (53) Itkin, M., Davidovich-Rikanati, R., Cohen, S., Portnoy, V., Doron-Faigenboim, A., Oren, E., Freilich, S., Tzuri, G., Baranes, N., Shen, S., Petreikov, M., Sertchook, R., Ben-Dor, S., Gottlieb, H., Hernandez, A., Nelson, D. R., Paris, H. S., Tadmor, Y., Burger, Y., Lewinsohn, E., Katzir, N., and Schaffer, A. (2016) The biosynthetic pathway of the nonsugar, high-intensity sweetener mogroside V from *Siraitia grosvenorii*. *Proc. Natl. Acad. Sci.* 113, E7619–E7628.
- (54) Tomlin, E. S., Antonejevic, E., Alfaro, R. I., and Borden, J. H. (2000) Changes in volatile terpene and diterpene resin acid composition of resistant and susceptible white spruce leaders exposed to simulated white pine weevil damage. *Tree Physiol.* 20, 1087–1095.

- (55) Martin, D., Tholl, D., Gershenzon, J., and Bohlmann, J. (2002) Methyl Jasmonate Induces Traumatic Resin Ducts, Terpenoid Resin Biosynthesis, and Terpenoid Accumulation in Developing Xylem of Norway Spruce Stems. *Plant Physiol.* **129**, 1003–1018.
- (56) Lee, G. W., Chung, M.-S., Lee, S. S., Chung, B. Y., and Lee, S. (2019) Transcriptome-guided identification and functional characterization of key terpene synthases involved in constitutive and methyl jasmonate-inducible volatile terpene formation in *Eremochloa ophiuroides* (Munro) Hack. *Plant Physiol. Biochem.* **141**, 193–201.
- (57) Bohlmann, J., Crock, J., Jetter, R., and Croteau, R. (1998) Terpenoid-based defenses in conifers: cDNA cloning, characterization, and functional expression of wound-inducible (E)- $\alpha$ -bisabolene synthase from grand fir (*Abies grandis*). *Proc. Natl. Acad. Sci.* **95**, 6756–6761.
- (58) Wang, X., Zeng, L., Liao, Y., Li, J., Tang, J., and Yang, Z. (2019) Formation of  $\alpha$ -Farnesene in Tea (*Camellia sinensis*) Leaves Induced by Herbivore-Derived Wounding and Its Effect on Neighboring Tea Plants. *Int. J. Mol. Sci.* **20**.
- (59) Tholl, D., Chen, F., Petri, J., Gershenzon, J., and Pichersky, E. (2005) Two sesquiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from *Arabidopsis* flowers. *Plant J. Cell Mol. Biol.* **42**, 757–771.
- (60) Ro, D.-K., Ehltling, J., Keeling, C. I., Lin, R., Mattheus, N., and Bohlmann, J. (2006) Microarray expression profiling and functional characterization of AtTPS genes: Duplicated *Arabidopsis thaliana* sesquiterpene synthase genes At4g13280 and At4g13300 encode root-specific and wound-inducible (Z)- $\gamma$ -bisabolene synthases. *Arch. Biochem. Biophys.* **448**, 104–116.
- (61) De Geyter, N., Gholami, A., Goormachtig, S., and Goossens, A. (2012) Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci.* **17**, 349–359.
- (62) Xu, Y.-H., Wang, J.-W., Wang, S., Wang, J.-Y., and Chen, X.-Y. (2004) Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)- $\delta$ -cadinene synthase-A. *Plant Physiol.* **135**, 507–515.
- (63) Ma, D., Pu, G., Lei, C., Ma, L., Wang, H., Guo, Y., Chen, J., Du, Z., Wang, H., Li, G., Ye, H., and Liu, B. (2009) Isolation and Characterization of AaWRKY1, an *Artemisia annua* Transcription Factor that Regulates the Amorpho-4,11-diene Synthase Gene, a Key Gene of Artemisinin Biosynthesis. *Plant Cell Physiol.* **50**, 2146–2161.
- (64) Yu, Z.-X., Li, J.-X., Yang, C.-Q., Hu, W.-L., Wang, L.-J., and Chen, X.-Y. (2012) The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. *Mol. Plant* **5**, 353–365.
- (65) Ji, Y., Xiao, J., Shen, Y., Ma, D., Li, Z., Pu, G., Li, X., Huang, L., Liu, B., Ye, H., and Wang, H. (2014) Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annua*. *Plant Cell Physiol.* **55**, 1592–1604.
- (66) Zhang, F., Fu, X., Lv, Z., Lu, X., Shen, Q., Zhang, L., Zhu, M., Wang, G., Sun, X., Liao, Z., and Tang, K. (2015) A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* **8**, 163–175.
- (67) Kayani, S.-I., Shen, Q., Ma, Y., Fu, X., Xie, L., Zhong, Y., Tiantian, C., Pan, Q., Li, L., Rahman, S., Sun, X., and Tang, K. (2019) The YABBY Family Transcription Factor AaYABBY5 Directly Targets Cytochrome P450 Monooxygenase (CYP71AV1) and Double-Bond Reductase 2 (DBR2) Involved in Artemisinin Biosynthesis in *Artemisia Annua*. *Front. Plant Sci.* **10**.
- (68) Henry, L. K., Thomas, S. T., Widhalm, J. R., Lynch, J. H., Davis, T. C., Kessler, S. A., Bohlmann, J., Noel, J. P., and Dudareva, N. (2018) Contribution of isopentenyl phosphate to plant terpenoid metabolism. *Nat. Plants* **4**, 721–729.
- (69) Chappell, J., Wolf, F., Proulx, J., Cuellar, R., and Saunders, C. (1995) Is the Reaction Catalyzed by 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase a Rate-Limiting Step for Isoprenoid Biosynthesis in Plants? *Plant Physiol.* **109**, 1337–1343.
- (70) Pollier, J., Moses, T., González-Guzmán, M., De Geyter, N., Lippens, S., Bossche, R. V., Marhavý, P., Kremer, A., Morreel, K., Guérin, C. J., Tava, A., Oleszek, W., Thevelein, J. M., Campos, N., Goormachtig, S., and Goossens,

- A. (2013) The protein quality control system manages plant defence compound synthesis. *Nature* 504, 148–152.
- (71) Laursen, T., Møller, B. L., and Bassard, J.-E. (2015) Plasticity of specialized metabolism as mediated by dynamic metabolons. *Trends Plant Sci.* 20, 20–32.
- (72) Laursen, T., Borch, J., Knudsen, C., Bavishi, K., Torta, F., Martens, H. J., Silvestro, D., Hatzakis, N. S., Wenk, M. R., Dafforn, T. R., Olsen, C. E., Motawia, M. S., Hamberger, B., Møller, B. L., and Bassard, J.-E. (2016) Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. *Science* 354, 890–893.
- (73) Dong, L., Pollier, J., Bassard, J.-E., Ntallas, G., Almeida, A., Lazaridi, E., Khakimov, B., Arendt, P., de Oliveira, L. S., Lota, F., Goossens, A., Michoux, F., and Bak, S. (2018) Co-expression of squalene epoxidases with triterpene cyclases boosts production of triterpenoids in plants and yeast. *Metab. Eng.* 49, 1–12.
- (74) Tissier, A., Morgan, J. A., and Dudareva, N. (2017) Plant Volatiles: Going ‘In’ but not ‘Out’ of Trichome Cavities. *Trends Plant Sci.* 22, 930–938.
- (75) Spring, O., Pfannstiel, J., Klaiber, I., Conrad, J., Beifuß, U., Apel, L., Aschenbrenner, A.-K., and Zipper, R. (2015) The nonvolatile metabolome of sunflower linear glandular trichomes. *Phytochemistry* 119, 83–89.
- (76) Soetaert, S. S., Van Neste, C. M., Vandewoestyne, M. L., Head, S. R., Goossens, A., Van Nieuwerburgh, F. C., and Deforce, D. L. (2013) Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biol.* 13, 220.
- (77) Wang, G., Tian, L., Aziz, N., Broun, P., Dai, X., He, J., King, A., Zhao, P. X., and Dixon, R. A. (2008) Terpene Biosynthesis in Glandular Trichomes of Hop. *Plant Physiol.* 148, 1254–1266.
- (78) Xiao, M., Zhang, Y., Chen, X., Lee, E.-J., Barber, C. J. S., Chakrabarty, R., Desgagné-Penix, I., Haslam, T. M., Kim, Y.-B., Liu, E., MacNevin, G., Masada-Atsumi, S., Reed, D. W., Stout, J. M., Zerbe, P., Zhang, Y., Bohlmann, J., Covello, P. S., De Luca, V., Page, J. E., Ro, D.-K., Martin, V. J. J., Facchini, P. J., and Sensen, C. W. (2013) Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.* 166, 122–134.
- (79) (2019) One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685.
- (80) Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E. (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom: Terpene synthase family. *Plant J.* 66, 212–229.
- (81) McAndrew, R. P., Peralta-Yahya, P. P., DeGiovanni, A., Pereira, J. H., Hadi, M. Z., Keasling, J. D., and Adams, P. D. (2011) Structure of a Three-Domain Sesquiterpene Synthase: A Prospective Target for Advanced Biofuels Production. *Structure* 19, 1876–1884.
- (82) Son, Y.-J., Kwon, M., Ro, D.-K., and Kim, S.-U. (2014) Enantioselective microbial synthesis of the indigenous natural product (-)- $\alpha$ -bisabolol by a sesquiterpene synthase from chamomile (*Matricaria recutita*). *Biochem. J.* 463, 239–248.
- (83) Tholl, D., and Lee, S. (2011) Terpene Specialized Metabolism in *Arabidopsis thaliana*. *Arab. Book* 9, e0143.
- (84) Aschenbrenner, A.-K., Kwon, M., Conrad, J., Ro, D.-K., and Spring, O. (2016) Identification and characterization of two bisabolene synthases from linear glandular trichomes of sunflower (*Helianthus annuus* L., Asteraceae). *Phytochemistry* 124, 29–37.
- (85) Jin, J., Zhang, S., Zhao, M., Jing, T., Zhang, N., Wang, J., Wu, B., and Song, C. (2020) Scenarios of Genes-to-Terpenoids Network Led to the Identification of a Novel  $\alpha/\beta$ -Farnesene/ $\beta$ -Ocimene Synthase in *Camellia sinensis*. *Int. J. Mol. Sci.* 21, 655.
- (86) Martin, D. M., Aubourg, S., Schouwey, M. B., Daviet, L., Schalk, M., Toub, O., Lund, S. T., and Bohlmann, J. (2010) Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FcDNA cloning, and enzyme assays. *BMC Plant Biol.* 10, 226.
- (87) Zhou, F., and Pichersky, E. (2020) The complete functional characterisation of the terpene synthase family in tomato. *New Phytol.*
- (88) Loizzi, M., Miller, D. J., and Allemann, R. K. (2019) Silent catalytic promiscuity in the high-fidelity terpene cyclase  $\delta$ -cadinene synthase. *Org. Biomol. Chem.* 17, 1206–1214.

- (89) Steele, C. L., Crock, J., Bohlmann, J., and Croteau, R. (1998) Sesquiterpene Synthases from Grand Fir (*Abies grandis*): comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and bacterial expression of  $\delta$ -selinene synthase and  $\gamma$ -humulene synthase. *J. Biol. Chem.* **273**, 2078–2089.
- (90) Abdallah, I. I., van Merkerk, R., Klumpenaar, E., and Quax, W. J. (2018) Catalysis of amorpho-4,11-diene synthase unraveled and improved by mutability landscape guided engineering. *Sci. Rep.* **8**, 9961.
- (91) Ker, D.-S., Chan, K. G., Othman, R., Hassan, M., and Ng, C. L. (2020) Site-directed mutagenesis of  $\beta$  sesquiphellandrene synthase enhances enzyme promiscuity. *Phytochemistry* **173**, 112286.
- (92) Omura, T., and Sato, R. (1964) The carbon monoxide-binding pigment of liver microsomes. Part I. Evidence for its hemoprotein nature. *J. Biol. Chem.* **239**, 2370–2378.
- (93) Nelson, D. R. (2018) Cytochrome P450 diversity in the tree of life. *Biochim. Biophys. Acta* **1866**, 141–154.
- (94) Rendic, S., and Guengerich, F. P. (2015) Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals. *Chem. Res. Toxicol.* **28**, 38–42.
- (95) Hannemann, F., Bichet, A., Ewen, K. M., and Bernhardt, R. (2007) Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1770**, 330–344.
- (96) Lu, A. Y. H., Junk, K. W., and Coon, M. J. (1969) Resolution of the Cytochrome P-450-containing  $\omega$ -Hydroxylation System of Liver Microsomes into Three Components. *J. Biol. Chem.* **244**, 3714–3721.
- (97) Werck-Reichhart, D., and Feyereisen, R. (2000) Cytochromes P450: a success story. *Genome Biol.* **1**, REVIEWS3003.
- (98) Poulos, T. L., Finzel, B. C., and Howard, A. J. (1986) Crystal structure of substrate-free *Pseudomonas putida* cytochrome P-450. *Biochemistry* **25**, 5314–5322.
- (99) Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F., and McRee, D. E. (2000) Mammalian Microsomal Cytochrome P450 Monooxygenase: Structural Adaptations for Membrane Binding and Functional Diversity. *Mol. Cell* **5**, 121–131.
- (100) Graham-Lorence, S., Truan, G., Peterson, J. A., Falck, J. R., Wei, S., Helvig, C., and Capdevila, J. H. (1997) An active site substitution, F87V, converts cytochrome P450 BM-3 into a regio- and stereoselective (14S,15R)-arachidonic acid epoxygenase. *J. Biol. Chem.* **272**, 1127–1135.
- (101) Takahashi, S., Yeo, Y.-S., Zhao, Y., O'Maille, P. E., Greenhagen, B. T., Noel, J. P., Coates, R. M., and Chappell, J. (2007) Functional Characterization of Premnaspirodiene Oxygenase, a Cytochrome P450 Catalyzing Regio- and Stereo-specific Hydroxylations of Diverse Sesquiterpene Substrates. *J. Biol. Chem.* **282**, 31744–31754.
- (102) Li, H., Mei, L., Urlacher, V. B., and Schmid, R. D. (2008) Cytochrome P450 BM-3 Evolved by Random and Saturation Mutagenesis as an Effective Indole-Hydroxylating Catalyst. *Appl. Biochem. Biotechnol.* **144**, 27–36.
- (103) Ralston, L., Kwon, S. T., Schoenbeck, M., Ralston, J., Schenk, D. J., Coates, R. M., and Chappell, J. (2001) Cloning, Heterologous Expression, and Functional Characterization of 5-epi-Aristolochene-1,3-Dihydroxylase from Tobacco (*Nicotiana tabacum*). *Arch. Biochem. Biophys.* **393**, 222–235.
- (104) Luo, P., Wang, Y. H., Wang, G. D., Essenberg, M., and Chen, X. Y. (2001) Molecular cloning and functional identification of (+)-delta-cadinene-8-hydroxylase, a cytochrome P450 mono-oxygenase (CYP706B1) of cotton sesquiterpene biosynthesis. *Plant J. Cell Mol. Biol.* **28**, 95–104.
- (105) Takase, H., Sasaki, K., Shinmori, H., Shinohara, A., Mochizuki, C., Kobayashi, H., Ikoma, G., Saito, H., Matsuo, H., Suzuki, S., and Takata, R. (2016) Cytochrome P450 CYP71BE5 in grapevine (*Vitis vinifera*) catalyzes the formation of the spicy aroma compound (-)-rotundone. *J. Exp. Bot.* **67**, 787–798.
- (106) Schuегger, R., Nafisi, M., Mansourova, M., Petersen, B. L., Olsen, C. E., Svatoš, A., Halkier, B. A., and Glawischnig, E. (2006) CYP71B15 (PAD3) Catalyzes the Final Step in Camalexin Biosynthesis. *Plant Physiol.* **141**, 1248–1254.
- (107) Besseau, S., Kellner, F., Lanoue, A., Thamm, A. M. K., Salim, V., Schneider, B., Geu-Flores, F., Höfer, R., Guirimand, G., Guihur, A., Oudin, A., Glevarec, G., Foureau, E., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., Werck-Reichhart, D., Burlat, V., De Luca, V., O'Connor, S. E., and Courdavault, V. (2013) A pair of tabersonine 16-hydroxylases initiates the synthesis of vindoline in an organ-dependent manner in *Catharanthus roseus*. *Plant Physiol.* **163**, 1792–1803.

- (108) Cankar, K., van Houwelingen, A., Goedbloed, M., Renirie, R., de Jong, R. M., Bouwmeester, H., Bosch, D., Sonke, T., and Beekwilder, J. (2014) Valencene oxidase CYP706M1 from Alaska cedar (*Callitropsis nootkatensis*). *FEBS Lett.* 588, 1001–1007.
- (109) Yu, F., Okamoto, S., Harada, H., Yamasaki, K., Misawa, N., and Utsumi, R. (2011) Zingiber zerumbet CYP71BA1 catalyzes the conversion of  $\alpha$ -humulene to 8-hydroxy- $\alpha$ -humulene in zerumbone biosynthesis. *Cell. Mol. Life Sci. CMLS* 68, 1033–1040.
- (110) Wang, Q., Hillwig, M. L., and Peters, R. J. (2011) CYP99A3: functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. *Plant J.* 65, 87–95.
- (111) Mugford, S. T., Louveau, T., Melton, R., Qi, X., Bakht, S., Hill, L., Tsurushima, T., Honkanen, S., Rosser, S. J., Lomonosoff, G. P., and Osbourn, A. (2013) Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. *Plant Cell* 25, 1078–1092.
- (112) Boutanaev, A. M., Moses, T., Zi, J., Nelson, D. R., Mugford, S. T., Peters, R. J., and Osbourn, A. (2015) Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. U. S. A.* 112, E81–88.
- (113) Boachon, B., Burdloff, Y., Ruan, J.-X., Rojo, R., Junker, R. R., Vincent, B., Nicolè, F., Bringel, F., Lesot, A., Henry, L., Bassard, J.-E., Mathieu, S., Allouche, L., Kaplan, I., Dudareva, N., Vuilleumier, S., Miesch, L., André, F., Navrot, N., Chen, X.-Y., and Werck-Reichhart, D. (2019) A Promiscuous CYP706A3 Reduces Terpene Volatile Emission from Arabidopsis Flowers, Affecting Florivores and the Floral Microbiome. *Plant Cell* 31, 2947–2972.
- (114) Peng, B., Nielsen, L. K., Kampranis, S. C., and Vickers, C. E. (2018) Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in *Saccharomyces cerevisiae*. *Metab. Eng.* 47, 83–93.
- (115) Ignea, C., Trikka, F. A., Nikolaidis, A. K., Georgantea, P., Ioannou, E., Loupassaki, S., Kefalas, P., Kanellis, A. K., Roussis, V., Makris, A. M., and Kampranis, S. C. (2015) Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. *Metab. Eng.* 27, 65–75.
- (116) Zhang, Y., Teoh, K. H., Reed, D. W., Maes, L., Goossens, A., Olson, D. J. H., Ross, A. R. S., and Covello, P. S. (2008) The molecular cloning of artemisinic aldehyde Delta11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J. Biol. Chem.* 283, 21501–21508.
- (117) Caputi, L., Lim, E.-K., and Bowles, D. J. (2008) Discovery of New Biocatalysts for the Glycosylation of Terpenoid Scaffolds. *Chem. – Eur. J.* 14, 6656–6662.
- (118) Zhao, M., Zhang, N., Gao, T., Jin, J., Jing, T., Wang, J., Wu, Y., Wan, X., Schwab, W., and Song, C. (2019) Sesquiterpene glucosylation mediated by glucosyltransferase UGT91Q2 is involved in the modulation of cold stress tolerance in tea plants. *New Phytol.*
- (119) Frey, M., Schmauder, K., Pateraki, I., and Spring, O. (2018) Biosynthesis of Eupatolide—A Metabolic Route for Sesquiterpene Lactone Formation Involving the P450 Enzyme CYP71DD6. *ACS Chem. Biol.* 13, 1536–1543.
- (120) Brown, G. D., and Sy, L.-K. (2007) In vivo transformations of artemisinic acid in *Artemisia annua* plants. *Tetrahedron* 63, 9548–9566.
- (121) Souto-Bachiller, F. A., De Jesus-Echevarría, M., Cárdenas-González, O. E., Acuña-Rodríguez, M. F., Meléndez, P. A., and Romero-Ramsey, L. (1997) Terpenoid composition of *Lippia dulcis*. *Phytochemistry* 44, 1077–1086.
- (122) Ono, M., Morinaga, H., Masuoka, C., Ikeda, T., Okawa, M., Kinjo, J., and Nohara, T. (2005) New Bisabolane-Type Sesquiterpenes from the Aerial Parts of *Lippia dulcis* 3.
- (123) Ono, M., Tsuru, T., Abe, H., Eto, M., Okawa, M., Abe, F., Kinjo, J., Ikeda, T., and Nohara, T. (2006) Bisabolane-Type Sesquiterpenes from the Aerial Parts of *Lippia dulcis*. *J. Nat. Prod.* 69, 1417–1420.
- (124) Rigamonti, M. G., and Gatti, F. G. (2015) Stereoselective synthesis of hernandulcin, peroxylippidulcine A, lippidulcines A, B and C and taste evaluation. *Beilstein J. Org. Chem.* 11, 2117–2124.
- (125) Combrinck, S., Du Plooy, G. W., McCrindle, R. I., and Botha, B. M. (2007) Morphology and histochemistry of the glandular trichomes of *Lippia scaberrima* (Verbenaceae). *Ann. Bot.* 99, 1111–1119.
- (126) Bailey, J. (1991) Toward a science of metabolic engineering. *Science* 252, 1668–1675.
- (127) Stephanopoulos, G., and Vallino, J. J. (1991) Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Sci. New Ser.* 252, 1675–1681.

- (128) Donald, K. A., Hampton, R. Y., and Fritz, I. B. (1997) Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* *63*, 3341–3344.
- (129) Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* *21*, 796–802.
- (130) Liu, Y., van Bennekom, E. O., Zhang, Y., Abee, T., and Smid, E. J. (2019) Long-chain vitamin K2 production in *Lactococcus lactis* is influenced by temperature, carbon source, aeration and mode of energy metabolism. *Microb. Cell Factories* *18*.
- (131) Brice, C., Cubillos, F. A., Dequin, S., Camarasa, C., and Martínez, C. (2018) Adaptability of the *Saccharomyces cerevisiae* yeasts to wine fermentation conditions relies on their strong ability to consume nitrogen. *PLoS One* *13*, e0192383.
- (132) Dai, Z., and Locasale, J. W. (2017) Understanding metabolism with flux analysis: from theory to application. *Metab. Eng.* *43*, 94–102.
- (133) Wang, L., Dash, S., Ng, C. Y., and Maranas, C. D. (2017) A review of computational tools for design and reconstruction of metabolic pathways. *Synth. Syst. Biotechnol.* *2*, 243–252.
- (134) Rocha, I., Maia, P., Evangelista, P., Vilaça, P., Soares, S., Pinto, J. P., Nielsen, J., Patil, K. R., Ferreira, E. C., and Rocha, M. (2010) OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Syst. Biol.* *4*, 45.
- (135) Shabestary, K., and Hudson, E. P. (2016) Computational metabolic engineering strategies for growth-coupled biofuel production by *Synechocystis*. *Metab. Eng. Commun.* *3*, 216–226.
- (136) Delépine, B., Duigou, T., Carbonell, P., and Faulon, J.-L. (2018) RetroPath2.0: A retrosynthesis workflow for metabolic engineers. *Metab. Eng.* *45*, 158–170.
- (137) Garcia Sanchez, R., Karhumaa, K., Fonseca, C., Sánchez Nogué, V., Almeida, J. R., Larsson, C. U., Bengtsson, O., Bettiga, M., Hahn-Hägerdal, B., and Gorwa-Grauslund, M. F. (2010) Improved xylose and arabinose utilization by an industrial recombinant *Saccharomyces cerevisiae* strain using evolutionary engineering. *Biotechnol. Biofuels* *3*, 13.
- (138) Caspeta, L., Chen, Y., Ghiaci, P., Feizi, A., Buskov, S., Hallström, B. M., Petranovic, D., and Nielsen, J. (2014) Altered sterol composition renders yeast thermotolerant. *Science* *346*, 75–78.
- (139) Pereira, R., Wei, Y., Mohamed, E., Radi, M., Malina, C., Herrgård, M. J., Feist, A. M., Nielsen, J., and Chen, Y. (2019) Adaptive laboratory evolution of tolerance to dicarboxylic acids in *Saccharomyces cerevisiae*. *Metab. Eng.* *56*, 130–141.
- (140) Trikka, F. A., Nikolaidis, A., Athanasakoglou, A., Andreadelli, A., Ignea, C., Kotta, K., Argiriou, A., Kampranis, S. C., and Makris, A. M. (2015) Iterative carotenogenic screens identify combinations of yeast gene deletions that enhance sclareol production. *Microb. Cell Factories* *14*.
- (141) Niederberger, P., Prasad, R., Miozzari, G., and Kacser, H. (1992) A strategy for increasing an in vivo flux by genetic manipulations. The tryptophan system of yeast. *Biochem. J.* *287 ( Pt 2)*, 473–479.
- (142) Verwaal, R., Wang, J., Meijnen, J.-P., Visser, H., Sandmann, G., van den Berg, J. A., and van Ooyen, A. J. J. (2007) High-Level Production of Beta-Carotene in *Saccharomyces cerevisiae* by Successive Transformation with Carotenogenic Genes from *Xanthophyllomyces dendrorhous*. *Appl. Environ. Microbiol.* *73*, 4342–4350.
- (143) Asadollahi, M. A., Maury, J., Møller, K., Nielsen, K. F., Schalk, M., Clark, A., and Nielsen, J. (2008) Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: Effect of ERG9 repression on sesquiterpene biosynthesis. *Biotechnol. Bioeng.* *99*, 666–677.
- (144) Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., and Jacobsen, S. E. (2006) Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat. Genet.* *38*, 721–725.
- (145) Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., and Bartel, D. P. (2009) RNAi in budding yeast. *Science* *326*, 544–550.
- (146) Crook, N. C., Schmitz, A. C., and Alper, H. S. (2014) Optimization of a Yeast RNA Interference System for Controlling Gene Expression and Enabling Rapid Metabolic Engineering. *ACS Synth. Biol.* *3*, 307–313.

- (147) Scalcinati, G., Knuf, C., Partow, S., Chen, Y., Maury, J., Schalk, M., Daviet, L., Nielsen, J., and Siewers, V. (2012) Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene  $\alpha$ -santalene in a fed-batch mode. *Metab. Eng.* *14*, 91–103.
- (148) Lohr, D., Venkov, P., and Zlatanova, J. (1995) Transcriptional regulation in the yeast GAL gene family: a complex genetic network. *FASEB J.* *9*, 777–787.
- (149) Xie, W., Liu, M., Lv, X., Lu, W., Gu, J., and Yu, H. (2014) Construction of a controllable  $\beta$ -carotene biosynthetic pathway by decentralized assembly strategy in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* *111*, 125–133.
- (150) Xie, W., Ye, L., Lv, X., Xu, H., and Yu, H. (2015) Sequential control of biosynthetic pathways for balanced utilization of metabolic intermediates in *Saccharomyces cerevisiae*. *Metab. Eng.* *28*, 8–18.
- (151) Garí, E., Piedrafita, L., Aldea, M., and Herrero, E. (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast Chichester Engl.* *13*, 837–848.
- (152) Kunjapur, A. M., and Prather, K. L. J. (2019) Development of a Vanillate Biosensor for the Vanillin Biosynthesis Pathway in *E. coli*. *ACS Synth. Biol.* *8*, 1958–1967.
- (153) Dabirian, Y., Gonçalves Teixeira, P., Nielsen, J., Siewers, V., and David, F. (2019) FadR-Based Biosensor-Assisted Screening for Genes Enhancing Fatty Acyl-CoA Pools in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* *8*, 1788–1800.
- (154) Snoek, T., Chaberski, E. K., Ambri, F., Kol, S., Bjørn, S. P., Pang, B., Barajas, J. F., Welner, D. H., Jensen, M. K., and Keasling, J. D. (2020) Evolution-guided engineering of small-molecule biosensors. *Nucleic Acids Res.* *48*, e3.
- (155) Epshtein, V., Mironov, A. S., and Nudler, E. (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 5052–5056.
- (156) Pang, Q., Han, H., Liu, X., Wang, Z., Liang, Q., Hou, J., Qi, Q., and Wang, Q. (2020) In vivo evolutionary engineering of riboswitch with high-threshold for N-acetylneuraminic acid production. *Metab. Eng.* *59*, 36–43.
- (157) Kaishima, M., Ishii, J., Matsuno, T., Fukuda, N., and Kondo, A. (2016) Expression of varied GFPs in *Saccharomyces cerevisiae*: codon optimization yields stronger than expected expression and fluorescence intensity. *Sci. Rep.* *6*, 1–15.
- (158) Lanza, A. M., Curran, K. A., Rey, L. G., and Alper, H. S. (2014) A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. *BMC Syst. Biol.* *8*, 33.
- (159) Park, S.-H., and Hahn, J.-S. (2019) Development of an efficient cytosolic isobutanol production pathway in *Saccharomyces cerevisiae* by optimizing copy numbers and expression of the pathway genes based on the toxic effect of  $\alpha$ -acetolactate. *Sci. Rep.* *9*.
- (160) Hamilton, R., Watanabe, C. K., and de Boer, H. A. (1987) Compilation and comparison of the sequence context around the AUG startcodons in *Saccharomyces cerevisiae* mRNAs. *Nucleic Acids Res.* *15*, 3581–3593.
- (161) Ro, D.-K., Ouellet, M., Paradise, E. M., Burd, H., Eng, D., Paddon, C. J., Newman, J. D., and Keasling, J. D. (2008) Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol.* *8*, 83.
- (162) Rabhi-Essafi, I., Sadok, A., Khalaf, N., and Fathallah, D. M. (2007) A strategy for high-level expression of soluble and functional human interferon  $\alpha$  as a GST-fusion protein in *E. coli*. *Protein Eng. Des. Sel.* *20*, 201–209.
- (163) Sørensen, H. P., and Mortensen, K. K. (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb. Cell Factories* *4*, 1.
- (164) Sever, N., Yang, T., Brown, M. S., Goldstein, J. L., and DeBose-Boyd, R. A. (2003) Accelerated Degradation of HMG CoA Reductase Mediated by Binding of Insig-1 to Its Sterol-Sensing Domain. *Mol. Cell* *11*, 25–33.
- (165) Brodelius, M., Lundgren, A., Mercke, P., and Brodelius, P. E. (2002) Fusion of farnesyl diphosphate synthase and epi-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis in *Nicotiana tabacum*. *Eur. J. Biochem.* *269*, 3570–3577.
- (166) Rabeharindranto, H., Castaño-Cerezo, S., Lautier, T., Garcia-Alles, L. F., Treitz, C., Tholey, A., and Truan, G. (2019) Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in *S. cerevisiae*. *Metab. Eng. Commun.* *8*.

- (167) Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L. J., and Keasling, J. D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* *27*, 753–759.
- (168) Tippmann, S., Anfelt, J., David, F., Rand, J. M., Siewers, V., Uhlén, M., Nielsen, J., and Hudson, E. P. (2017) Affibody Scaffolds Improve Sesquiterpene Production in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* *6*, 19–28.
- (169) Xie, S., Qiu, X., Zhu, L., Zhu, C., Liu, C., Wu, X., Zhu, L., and Zhang, D. (2019) Assembly of TALE-based DNA scaffold for the enhancement of exogenous multi-enzymatic pathway. *J. Biotechnol.* *296*, 69–74.
- (170) Edgell, C. L., Smith, A. J., Beesley, J. L., Savery, N. J., and Woolfson, D. N. (2020) De Novo Designed Protein-Interaction Modules for In-Cell Applications. *ACS Synth. Biol.* *9*, 427–436.
- (171) An, S., Kumar, R., Sheets, E. D., and Benkovic, S. J. (2008) Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells. *Science* *320*, 103–106.
- (172) Tian, R., Liu, Y., Chen, J., Li, J., Liu, L., Du, G., and Chen, J. (2019) Synthetic N-terminal coding sequences for fine-tuning gene expression and metabolic engineering in *Bacillus subtilis*. *Metab. Eng.* *55*, 131–141.
- (173) Peng, B., Plan, M. R., Chrysanthopoulos, P., Hodson, M. P., Nielsen, L. K., and Vickers, C. E. (2017) A squalene synthase protein degradation method for improved sesquiterpene production in *Saccharomyces cerevisiae*. *Metab. Eng.* *39*, 209–219.
- (174) Gao, C., Hou, J., Xu, P., Guo, L., Chen, X., Hu, G., Ye, C., Edwards, H., Chen, J., Chen, W., and Liu, L. (2019) Programmable biomolecular switches for rewiring flux in *Escherichia coli*. *Nat. Commun.* *10*, 1–12.
- (175) Guo, Z., Johnston, W. A., Whitfield, J., Walden, P., Cui, Z., Wijker, E., Edwardraja, S., Retamal Lantadilla, I., Ely, F., Vickers, C., Ungerer, J. P. J., and Alexandrov, K. (2019) Generalizable Protein Biosensors Based on Synthetic Switch Modules. *J. Am. Chem. Soc.* *141*, 8128–8135.
- (176) Zhou, K., Qiao, K., Edgar, S., and Stephanopoulos, G. (2015) Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* *33*, 377–383.
- (177) Fedeson, D. T., Saake, P., Calero, P., Nikel, P. I., and Ducat, D. C. (2020) Biotransformation of 2,4-dinitrotoluene in a phototrophic co-culture of engineered *Synechococcus elongatus* and *Pseudomonas putida*. *Microb. Biotechnol.*
- (178) Walther, T., Topham, C. M., Irague, R., Auriol, C., Baylac, A., Cordier, H., Dressaire, C., Lozano-Huguet, L., Tarrat, N., Martineau, N., Stodel, M., Malbert, Y., Maestracci, M., Huet, R., André, I., Remaud-Siméon, M., and François, J. M. (2017) Construction of a synthetic metabolic pathway for biosynthesis of the non-natural methionine precursor 2,4-dihydroxybutyric acid. *Nat. Commun.* *8*, 15828.
- (179) Xie, W., Lv, X., Ye, L., Zhou, P., and Yu, H. (2015) Construction of lycopene-overproducing *Saccharomyces cerevisiae* by combining directed evolution and metabolic engineering. *Metab. Eng.* *30*, 69–78.
- (180) Brennan, T. C. R., Turner, C. D., Krömer, J. O., and Nielsen, L. K. (2012) Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* *109*, 2513–2522.
- (181) Li, J., Xu, J., Cai, P., Wang, B., Ma, Y., Benz, J. P., and Tian, C. (2015) Functional Analysis of Two l-Arabinose Transporters from Filamentous Fungi Reveals Promising Characteristics for Improved Pentose Utilization in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* *81*, 4062–4070.
- (182) Hu, Y., Zhu, Z., Nielsen, J., and Siewers, V. (2018) Heterologous transporter expression for improved fatty alcohol secretion in yeast. *Metab. Eng.* *45*, 51–58.
- (183) Lee, M. J., Brown, I. R., Juodeikis, R., Frank, S., and Warren, M. J. (2016) Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in *Escherichia coli*. *Metab. Eng.* *36*, 48–56.
- (184) Ferlez, B., Sutter, M., and Kerfeld, C. A. (2019) A designed bacterial microcompartment shell with tunable composition and precision cargo loading. *Metab. Eng.* *54*, 286–291.
- (185) Arendt, P., Miettinen, K., Pollier, J., De Rycke, R., Callewaert, N., and Goossens, A. (2017) An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. *Metab. Eng.* *40*, 165–175.
- (186) Kim, J.-E., Jang, I.-S., Son, S.-H., Ko, Y.-J., Cho, B.-K., Kim, S. C., and Lee, J. Y. (2019) Tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metab. Eng.* *56*, 50–59.

- (187) Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014) A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390.
- (188) Carquet, M., Pompon, D., and Truan, G. (2015) Transcription Interference and ORF Nature Strongly Affect Promoter Strength in a Reconstituted Metabolic Pathway. *Front. Bioeng. Biotechnol.* **3**.
- (189) Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2010) Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **329**, 52–56.
- (190) Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z.-Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., and Venter, J. C. (2016) Design and synthesis of a minimal bacterial genome. *Science* **351**.
- (191) Annaluru, N., Muller, H., Mitchell, L. A., Ramalingam, S., Stracquadanio, G., Richardson, S. M., Dymond, J. S., Kuang, Z., Scheifele, L. Z., Cooper, E. M., Cai, Y., Zeller, K., Agmon, N., Han, J. S., Hadjithomas, M., Tullman, J., Caravelli, K., Cirelli, K., Guo, Z., London, V., Yeluru, A., Murugan, S., Kandavelou, K., Agier, N., Fischer, G., Yang, K., Martin, J. A., Bilgel, M., Bohutski, P., Boulter, K. M., Capaldo, B. J., Chang, J., Charoen, K., Choi, W. J., Deng, P., DiCarlo, J. E., Doong, J., Dunn, J., Feinberg, J. I., Fernandez, C., Floria, C. E., Gladowski, D., Hadidi, P., Ishizuka, I., Jabbari, J., Lau, C. Y. L., Lee, P. A., Li, S., Lin, D., Linder, M. E., Ling, J., Liu, J., Liu, J., London, M., Ma, H., Mao, J., McDade, J. E., McMillan, A., Moore, A. M., Oh, W. C., Ouyang, Y., Patel, R., Paul, M., Paulsen, L. C., Qiu, J., Rhee, A., Rubashkin, M. G., Soh, I. Y., Sotuyo, N. E., Srinivas, V., Suarez, A., Wong, A., Wong, R., Xie, W. R., Xu, Y., Yu, A. T., Koszul, R., Bader, J. S., Boeke, J. D., and Chandrasegaran, S. (2014) Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55–58.
- (192) Richardson, S. M., Mitchell, L. A., Stracquadanio, G., Yang, K., Dymond, J. S., DiCarlo, J. E., Lee, D., Huang, C. L. V., Chandrasegaran, S., Cai, Y., Boeke, J. D., and Bader, J. S. (2017) Design of a synthetic yeast genome. *Science* **355**, 1040–1044.
- (193) Shao, Y., Lu, N., Wu, Z., Cai, C., Wang, S., Zhang, L.-L., Zhou, F., Xiao, S., Liu, L., Zeng, X., Zheng, H., Yang, C., Zhao, Z., Zhao, G., Zhou, J.-Q., Xue, X., and Qin, Z. (2018) Creating a functional single-chromosome yeast. *Nature* **560**, 331–335.
- (194) Luo, Z., Wang, L., Wang, Y., Zhang, W., Guo, Y., Shen, Y., Jiang, L., Wu, Q., Zhang, C., Cai, Y., and Dai, J. (2018) Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. *Nat. Commun.* **9**, 1–10.
- (195) Jia, B., Wu, Y., Li, B.-Z., Mitchell, L. A., Liu, H., Pan, S., Wang, J., Zhang, H.-R., Jia, N., Li, B., Shen, M., Xie, Z.-X., Liu, D., Cao, Y.-X., Li, X., Zhou, X., Qi, H., Boeke, J. D., and Yuan, Y.-J. (2018) Precise control of SCRaMbLE in synthetic haploid and diploid yeast. *Nat. Commun.* **9**, 1933.
- (196) Gowers, G.-O. F., Chee, S. M., Bell, D., Suckling, L., Kern, M., Tew, D., McClymont, D. W., and Ellis, T. (2020) Improved betulonic acid biosynthesis using synthetic yeast chromosome recombination and semi-automated rapid LC-MS screening. *Nat. Commun.* **11**, 868.
- (197) Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342.
- (198) Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338.
- (199) Wilson, E. H., Sagawa, S., Weis, J. W., Schubert, M. G., Bissell, M., Hawthorne, B., Reeves, C. D., Dean, J., and Platt, D. (2016) Genotype Specification Language. *ACS Synth. Biol.* **5**, 471–478.
- (200) Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., Moriya, H., and Matsuyama, T. (2013) A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a “terminatome” toolbox. *ACS Synth. Biol.* **2**, 337–347.
- (201) Redden, H., Morse, N., and Alper, H. S. (2015) The synthetic biology toolbox for tuning gene expression in yeast. *FEMS Yeast Res.* **15**, 1–10.
- (202) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345.

- (203) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6, e16765.
- (204) Rantasalo, A., Kuivanen, J., Penttilä, M., Jäntti, J., and Mojzita, D. (2018) Synthetic Toolkit for Complex Genetic Circuit Engineering in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 7, 1573–1587.
- (205) Casini, A., Chang, F.-Y., Eluere, R., King, A. M., Young, E. M., Dudley, Q. M., Karim, A., Pratt, K., Bristol, C., Forget, A., Ghodasara, A., Warden-Rothman, R., Gan, R., Cristofaro, A., Borujeni, A. E., Ryu, M.-H., Li, J., Kwon, Y.-C., Wang, H., Tatsis, E., Rodriguez-Lopez, C., O'Connor, S., Medema, M. H., Fischbach, M. A., Jewett, M. C., Voigt, C., and Gordon, D. B. (2018) A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology. *J. Am. Chem. Soc.* 140, 4302–4316.
- (206) Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017) Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8, 1688.
- (207) Hillson, N., Caddick, M., Cai, Y., Carrasco, J. A., Chang, M. W., Curach, N. C., Bell, D. J., Feuvre, R. L., Friedman, D. C., Fu, X., Gold, N. D., Herrgård, M. J., Holowko, M. B., Johnson, J. R., Johnson, R. A., Keasling, J. D., Kitney, R. I., Kondo, A., Liu, C., Martin, V. J. J., Menolascina, F., Ogino, C., Patron, N. J., Pavan, M., Poh, C. L., Pretorius, I. S., Rosser, S. J., Scrutton, N. S., Storch, M., Tekotte, H., Travník, E., Vickers, C. E., Yew, W. S., Yuan, Y., Zhao, H., and Freemont, P. S. (2019) Building a global alliance of biofoundries. *Nat. Commun.* 10, 1–4.
- (208) Moses, T., Mehrshahi, P., Smith, A. G., and Goossens, A. (2017) Synthetic biology approaches for the production of plant metabolites in unicellular organisms. *J. Exp. Bot.* 68, 4057–4074.
- (209) Cumbers, J. Synthetic Biology Unicorn Ginkgo Bioworks Raises An Additional \$290 Million. *Forbes*.
- (210) Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996) Life with 6000 genes. *Science* 274, 546, 563–567.
- (211) Borodina, I., and Nielsen, J. (2014) Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. *Biotechnol. J.* 9, 609–620.
- (212) Anderson, M. S., Yarger, J. G., Burck, C. L., and Poulter, C. D. (1989) Farnesyl diphosphate synthetase. Molecular cloning, sequence, and expression of an essential gene from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 264, 19176–19184.
- (213) Fegueur, M., Richard, L., Charles, A. D., and Karst, F. (1991) Isolation and primary structure of the ERG9 gene of *Saccharomyces cerevisiae* encoding squalene synthetase. *Curr. Genet.* 20, 365–372.
- (214) Lees, N. D., Skaggs, B., Kirsch, D. R., and Bard, M. (1995) Cloning of the late genes in the ergosterol biosynthetic pathway of *Saccharomyces cerevisiae*--a review. *Lipids* 30, 221–226.
- (215) Ott, R. G., Athenstaedt, K., Hrstnik, C., Leitner, E., Bergler, H., and Daum, G. (2005) Flux of sterol intermediates in a yeast strain deleted of the lanosterol C-14 demethylase Erg11p. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* 1735, 111–118.
- (216) Vik, Å., and Rine, J. (2001) Upc2p and Ecm22p, Dual Regulators of Sterol Biosynthesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 21, 6395–6405.
- (217) Davies, B. S. J., Wang, H. S., and Rine, J. (2005) Dual Activators of the Sterol Biosynthetic Pathway of *Saccharomyces cerevisiae*: Similar Activation/Regulatory Domains but Different Response Mechanisms. *Mol. Cell. Biol.* 25, 7375–7385.
- (218) Chambon, C., Ladeveze, V., Servouse, M., Blanchard, L., Javelof, C., Vladescu, B., and Karst, F. (1991) Sterol pathway in yeast. Identification and properties of mutant strains defective in mevalonate diphosphate decarboxylase and farnesyl diphosphate synthetase. *Lipids* 26, 633–636.
- (219) Blanchard, L., and Karst, F. (1993) Characterization of a lysine-to-glutamic acid mutation in a conservative sequence of farnesyl diphosphate synthase from *Saccharomyces cerevisiae*. *Gene* 125, 185–189.
- (220) Grabińska, K., and Palamarczyk, G. (2002) Dolichol biosynthesis in the yeast *Saccharomyces cerevisiae* : an insight into the regulatory role of farnesyl diphosphate synthase. *FEMS Yeast Res.* 2, 259–265.
- (221) Westfall, P. J., Pitera, D. J., Lenihan, J. R., Eng, D., Woolard, F. X., Regentin, R., Horning, T., Tsuruta, H., Melis, D. J., Owens, A., Fickes, S., Diola, D., Benjamin, K. R., Keasling, J. D., Leavell, M. D., McPhee, D. J., Renninger, N. S., Newman, J. D., and Paddon, C. J. (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinin acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U. S. A.* 109, E111-118.

- (222) Duport, C., Spagnoli, R., Degryse, E., and Pompon, D. (1998) Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nat. Biotechnol.* *16*, 186–189.
- (223) Peng, B., Plan, M. R., Carpenter, A., Nielsen, L. K., and Vickers, C. E. (2017) Coupling gene regulatory patterns to bioprocess conditions to optimize synthetic metabolic modules for improved sesquiterpene production in yeast. *Biotechnol. Biofuels* *10*.
- (224) Shiba, Y., Paradise, E. M., Kirby, J., Ro, D.-K., and Keasling, J. D. (2007) Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab. Eng.* *9*, 160–168.
- (225) Kozak, B. U., Rossum, H. M. van, Luttik, M. A. H., Akeroyd, M., Benjamin, K. R., Wu, L., Vries, S. de, Daran, J.-M., Pronk, J. T., and Maris, A. J. A. van. (2014) Engineering Acetyl Coenzyme A Supply: Functional Expression of a Bacterial Pyruvate Dehydrogenase Complex in the Cytosol of *Saccharomyces cerevisiae*. *mBio* *5*.
- (226) Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., Petzold, C. J., Mukhopadhyay, A., Lee, T. S., Adams, P. D., and Keasling, J. D. (2013) Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* *31*, 1039–1046.
- (227) Korman, T. P., Opgenorth, P. H., and Bowie, J. U. (2017) A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat. Commun.* *8*.
- (228) Chatzivasileiou, A. O., Ward, V., Edgar, S. M., and Stephanopoulos, G. (2019) Two-step pathway for isoprenoid synthesis. *Proc. Natl. Acad. Sci.* *116*, 506–511.
- (229) Rico, J., Duquesne, K., Petit, J.-L., Mariage, A., Darii, E., Peruch, F., de Berardinis, V., and Iacazio, G. (2019) Exploring natural biodiversity to expand access to microbial terpene synthesis. *Microb. Cell Factories* *18*.
- (230) Ma, T., Shi, B., Ye, Z., Li, X., Liu, M., Chen, Y., Xia, J., Nielsen, J., Deng, Z., and Liu, T. (2019) Lipid engineering combined with systematic metabolic engineering of *Saccharomyces cerevisiae* for high-yield production of lycopene. *Metab. Eng.* *52*, 134–142.
- (231) Yee, D. A., DeNicola, A. B., Billingsley, J. M., Creso, J. G., Subrahmanyam, V., and Tang, Y. (2019) Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metab. Eng.* *55*, 76–84.
- (232) Zhang, Y., Wang, J., Cao, X., Liu, W., Yu, H., and Ye, L. (2020) High-level production of linalool by engineered *Saccharomyces cerevisiae* harboring dual mevalonate pathways in mitochondria and cytoplasm. *Enzyme Microb. Technol.* *134*, 109462.
- (233) Liu, G.-S., Li, T., Zhou, W., Jiang, M., Tao, X.-Y., Liu, M., Zhao, M., Ren, Y.-H., Gao, B., Wang, F.-Q., and Wei, D.-Z. (2019) The yeast peroxisome: A dynamic storage depot and subcellular factory for squalene overproduction. *Metab. Eng.*
- (234) Kan, S. B. J., Lewis, R. D., Chen, K., and Arnold, F. H. (2016) Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science* *354*, 1048–1051.
- (235) Erb, T. J., Jones, P. R., and Bar-Even, A. (2017) Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.* *37*, 56–62.
- (236) Umeno, D., and Arnold, F. H. (2003) A C35 carotenoid biosynthetic pathway. *Appl. Environ. Microbiol.* *69*, 3573–3579.
- (237) Fukushima, E. O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., and Muranaka, T. (2013) Combinatorial Biosynthesis of Legume Natural and Rare Triterpenoids in Engineered Yeast. *Plant Cell Physiol.* *54*, 740–749.
- (238) Mafu, S., Jia, M., Zi, J., Morrone, D., Wu, Y., Xu, M., Hillwig, M. L., and Peters, R. J. (2016) Probing the promiscuity of ent-kaurene oxidases via combinatorial biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 2526–2531.
- (239) Ignea, C., Pontini, M., Motawia, M. S., Maffei, M. E., Makris, A. M., and Kampranis, S. C. (2018) Synthesis of 11-carbon terpenoids in yeast using protein and metabolic engineering. *Nat. Chem. Biol.* *14*, 1090–1098.
- (240) Han, G. H., Kim, S. K., Yoon, P. K.-S., Kang, Y., Kim, B. S., Fu, Y., Sung, B. H., Jung, H. C., Lee, D.-H., Kim, S.-W., and Lee, S.-G. (2016) Fermentative production and direct extraction of (-)- $\alpha$ -bisabolol in metabolically engineered *Escherichia coli*. *Microb. Cell Factories* *15*, 185.
- (241) Tsuruta, H., Paddon, C. J., Eng, D., Lenihan, J. R., Horning, T., Anthony, L. C., Regentin, R., Keasling, J. D., Renninger, N. S., and Newman, J. D. (2009) High-level production of amorpho-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* *4*, e4489.

- (242) Christensen, U., Vazquez-Albacete, D., Sjøgaard, K. M., Hobel, T., Nielsen, M. T., Harrison, S. J., Hansen, A. H., Møller, B. L., Seppälä, S., and Nørholm, M. H. H. (2017) De-bugging and maximizing plant cytochrome P450 production in *Escherichia coli* with C-terminal GFP fusions. *Appl. Microbiol. Biotechnol.* **101**, 4103–4113.
- (243) Henke, N. A., Wichmann, J., Baier, T., Frohwitter, J., Lauersen, K. J., Risse, J. M., Peters-Wendisch, P., Kruse, O., and Wendisch, V. F. (2018) Patchoulol Production with Metabolically Engineered *Corynebacterium glutamicum*. *Genes* **9**.
- (244) Guan, Z., Xue, D., Abdallah, I. I., Dijkshoorn, L., Setroikromo, R., Lv, G., and Quax, W. J. (2015) Metabolic engineering of *Bacillus subtilis* for terpenoid production. *Appl. Microbiol. Biotechnol.* **99**, 9395–9406.
- (245) Vavitsas, K., Fabris, M., and Vickers, C. E. (2018) Terpenoid Metabolic Engineering in Photosynthetic Microorganisms. *Genes* **9**.
- (246) Reed, J., Stephenson, M. J., Miettinen, K., Brouwer, B., Leveau, A., Brett, P., Goss, R. J. M., Goossens, A., O'Connell, M. A., and Osbourn, A. (2017) A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules. *Metab. Eng.* **42**, 185–193.
- (247) Chiu, C. C., Keeling, C. I., and Bohlmann, J. (2018) Monoterpenyl esters in juvenile mountain pine beetle and sex-specific release of the aggregation pheromone trans-verbenol. *Proc. Natl. Acad. Sci.* **115**, 3652–3657.
- (248) Umehara, M., Hanada, A., Yoshida, S., Akiyama, K., Arite, T., Takeda-Kamiya, N., Magome, H., Kamiya, Y., Shirasu, K., Yoneyama, K., Kyozuka, J., and Yamaguchi, S. (2008) Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455**, 195–200.
- (249) Kramer, R., and Abraham, W.-R. (2012) Volatile sesquiterpenes from fungi: what are they good for? *Phytochem. Rev.* **11**, 15–37.
- (250) Kim, S., Jung, E., Kim, J.-H., Park, Y.-H., Lee, J., and Park, D. (2011) Inhibitory effects of (–)- $\alpha$ -bisabolol on LPS-induced inflammatory response in RAW264.7 macrophages. *Food Chem. Toxicol.* **49**, 2580–2585.
- (251) Ringer, K. L., Davis, E. M., and Croteau, R. (2005) Monoterpene Metabolism. Cloning, Expression, and Characterization of (–)-Isopiperitenol/(–)-Carveol Dehydrogenase of Peppermint and Spearmint. *Plant Physiol.* **137**, 863–872.
- (252) Maimone, T. J., and Baran, P. S. (2007) Modern synthetic efforts toward biologically active terpenes. *Nat. Chem. Biol.* **3**, 396.
- (253) Polakowski, T., Stahl, U., and Lang, C. (1998) Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl. Microbiol. Biotechnol.* **49**, 66–71.
- (254) Zhu, F., Zhong, X., Hu, M., Lu, L., Deng, Z., and Liu, T. (2014) In vitro reconstitution of mevalonate pathway and targeted engineering of farnesene overproduction in *Escherichia coli*. *Biotechnol. Bioeng.* **111**, 1396–1405.
- (255) Moses, T., Pollier, J., Faizal, A., Apers, S., Pieters, L., Thevelein, J. M., Geelen, D., and Goossens, A. (2015) Unraveling the Triterpenoid Saponin Biosynthesis of the African Shrub *Maesa lanceolata*. *Mol. Plant* **8**, 122–135.
- (256) Bian, G., Han, Y., Hou, A., Yuan, Y., Liu, X., Deng, Z., and Liu, T. (2017) Releasing the potential power of terpene synthases by a robust precursor supply platform. *Metab. Eng.* **42**, 1–8.
- (257) Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64.
- (258) Brehm-Stecher, B. F., and Johnson, E. A. (2003) Sensitization of *Staphylococcus aureus* and *Escherichia coli* to Antibiotics by the Sesquiterpenoids Nerolidol, Farnesol, Bisabolol, and Apritone. *Antimicrob. Agents Chemother.* **47**, 3357–3360.
- (259) Ganzera, M., Schneider, P., and Stuppner, H. (2006) Inhibitory effects of the essential oil of chamomile (*Matricaria recutita* L.) and its major constituents on human cytochrome P450 enzymes. *Life Sci.* **78**, 856–861.
- (260) Lautier, T., Urban, P., Loeper, J., Jezequel, L., Pompon, D., and Truan, G. (2016) Ordered chimerogenesis applied to CYP2B P450 enzymes. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1860**, 1395–1403.
- (261) Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115–132.
- (262) Urban, P., Cullin, C., and Pompon, D. (1990) Maximizing the expression of mammalian cytochrome P-450 monooxygenase activities in yeast cells. *Biochimie* **72**, 463–472.

- (263) Gietz, R. D., and Schiestl, R. H. (2007) Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2, 1.
- (264) Gietz, D., St Jean, A., Woods, R. A., and Schiestl, R. H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* 20, 1425.
- (265) Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D., and Hegemann, J. H. (2002) A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* 30, e23.
- (266) Othman, S., Mansuy-Mouries, V., Bensoussan, C., Battioni, P., and Mansuy, D. (2000) Hydroxylation of diclofenac: an illustration of the complementary roles of biomimetic metalloporphyrin catalysts and yeasts expressing human cytochromes P450 in drug metabolism studies. *Comptes Rendus Académie Sci. - Ser. IIC - Chem.* 3, 751–755.
- (267) Demarque, D. P., Crotti, A. E., Vessecchi, R., Lopes, J. L., and Lopes, N. P. (2016) Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat. Prod. Rep.* 33, 432–455.
- (268) Kampranis, S. C., and Makris, A. M. (2012) Developing a yeast cell factory for the production of terpenoids. *Comput. Struct. Biotechnol. J.* 3, e201210006.
- (269) Hu, Y., Zhou, Y. J., Bao, J., Huang, L., Nielsen, J., and Krivoruchko, A. (2017) Metabolic engineering of *Saccharomyces cerevisiae* for production of germacrene A, a precursor of beta-elemene. *J. Ind. Microbiol. Biotechnol.* 44, 1065–1072.
- (270) Faulkner, A., Chen, X., Rush, J., Horazdovsky, B., Waechter, C. J., Carman, G. M., and Sternweis, P. C. (1999) The LPP1 and DPP1 Gene Products Account for Most of the Isoprenoid Phosphate Phosphatase Activities in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 274, 14831–14837.
- (271) Tornio, A., and Backman, J. T. (2018) Chapter One - Cytochrome P450 in Pharmacogenetics: An Update, in *Advances in Pharmacology* (Brøsen, K., and Damkier, P., Eds.), pp 3–32. Academic Press.
- (272) Urban, P., Lautier, T., Pompon, D., and Truan, G. (2018) Ligand Access Channels in Cytochrome P450 Enzymes: A Review. *Int. J. Mol. Sci.* 19, 1617.
- (273) Shah, M. B., Wilderman, P. R., Liu, J., Jang, H.-H., Zhang, Q., Stout, C. D., and Halpert, J. R. (2015) Structural and Biophysical Characterization of Human Cytochromes P450 2B6 and 2A6 Bound to Volatile Hydrocarbons: Analysis and Comparison. *Mol. Pharmacol.* 87, 649–659.
- (274) DeBarber, A. E., Bleye, L. A., Roullet, J.-B. O., and Koop, D. R. (2004)  $\omega$ -Hydroxylation of farnesol by mammalian cytochromes P450. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* 1682, 18–27.
- (275) Tyzack, J. D., and Kirchmair, J. (2019) Computational methods and tools to predict cytochrome P450 metabolism for drug discovery. *Chem. Biol. Drug Des.* 93, 377–386.
- (276) Šícho, M., Stork, C., Mazzolari, A., de Bruyn Kops, C., Pedretti, A., Testa, B., Vistoli, G., Svozil, D., and Kirchmair, J. (2019) FAME 3: Predicting the Sites of Metabolism in Synthetic Compounds and Natural Products for Phase 1 and Phase 2 Metabolic Enzymes. *J. Chem. Inf. Model.* 59, 3400–3412.
- (277) Zaretski, J., Matlock, M., and Swamidass, S. J. (2013) XenoSite: Accurately Predicting CYP-Mediated Sites of Metabolism with Neural Networks. *J. Chem. Inf. Model.* 53, 3373–3383.
- (278) Albertsen, L., Chen, Y., Bach, L. S., Rattleff, S., Maury, J., Brix, S., Nielsen, J., and Mortensen, U. H. (2011) Diversion of Flux toward Sesquiterpene Production in *Saccharomyces cerevisiae* by Fusion of Host and Heterologous Enzymes. *Appl. Environ. Microbiol.* 77, 1033–1040.
- (279) Androutsopoulos, V. P., Papakyriakou, A., Vourloumis, D., and Spandidos, D. A. (2011) Comparative CYP1A1 and CYP1B1 substrate and inhibitor profile of dietary flavonoids. *Bioorg. Med. Chem.* 19, 2842–2849.
- (280) Schaub, P., Yu, Q., Gemmecker, S., Poussin-Courmontagne, P., Mailliot, J., McEwen, A. G., Ghisla, S., Al-Babili, S., Cavarelli, J., and Beyer, P. (2012) On the Structure and Function of the Phytoene Desaturase CRTI from *Pantoea ananatis*, a Membrane-Peripheral and FAD-Dependent Oxidase/Isomerase. *PLOS ONE* 7, e39550.
- (281) Laine, R. (1996, January 1) Modulation du stress oxydatif associe aux cytochromes p450 1a1 : de la regulation transcriptionnelle dependante de gstm1 au controle des voies abortives du cycle catalytique. thesis, Paris 6.
- (282) Zhou, H., Josephy, P. D., Kim, D., and Guengerich, F. P. (2004) Functional characterization of four allelic variants of human cytochrome P450 1A2. *Arch. Biochem. Biophys.* 422, 23–30.

- (283) Pompon, D. (1988) cDNA cloning and functional expression in yeast *Saccharomyces cerevisiae* of  $\beta$ -naphthoflavone-induced rabbit liver P-450 LM4 and LM6. *Eur. J. Biochem.* 177, 285–293.
- (284) rugosa, R., Hashidoko, Y., Tahara, S., and Mizutani, J. (1994) Six sesquiterpenoids from glandular trichome exudates of. *Phytochemistry* 35, 325–329.
- (285) Guimarães, L. G. de L., Cardoso, M. das G., Lucas, E. M. F., Freitas, M. P. de, Francisco, W., and Nelson, D. L. (2015) Structural Elucidation of a New Sesquiterpene Alcohol by Comparative NMR Studies. *Rec. Nat. Prod.* 9, 201–207.
- (286) Hashidoko, Y., Tahara, S., and Mizutani, J. (1992) Bisabolane sesquiterpenes and a 2-phenoxychromone from *Rosa Woodsii* leaves. *Phytochemistry* 31, 2148–2149.
- (287) Luche, J. L. (1978) Lanthanides in organic chemistry. 1. Selective 1,2 reductions of conjugated ketones. *J. Am. Chem. Soc.* 100, 2226–2227.
- (288) Urban, P., Werck-Reichhart, D., Teutsch, H. G., Durst, F., Regnier, S., Kazmaier, M., and Pompon, D. (1994) Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. *Eur. J. Biochem.* 222, 843–850.
- (289) Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., and Pompon, D. (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J. Biol. Chem.* 272, 19176–19186.
- (290) Tijet, N., Helvig, C., Pinot, F., Le Bouquin, R., Lesot, A., Durst, F., Salaün, J. P., and Benveniste, I. (1998) Functional expression in yeast and characterization of a clofibrate-inducible plant cytochrome P-450 (CYP94A1) involved in cutin monomers synthesis. *Biochem. J.* 332 ( Pt 2), 583–589.
- (291) Le Bouquin, R., Pinot, F., Benveniste, I., Salaün, J. P., and Durst, F. (1999) Cloning and functional characterization of CYP94A2, a medium chain fatty acid hydroxylase from *Vicia sativa*. *Biochem. Biophys. Res. Commun.* 261, 156–162.
- (292) King, A. J., Brown, G. D., Gilday, A. D., Forestier, E., Larson, T. R., and Graham, I. A. (2016) A Cytochrome P450-Mediated Intramolecular Carbon–Carbon Ring Closure in the Biosynthesis of Multidrug-Resistance-Reversing Lathyrane Diterpenoids. *Chembiochem* 17, 1593–1597.
- (293) Wong, J., de Rond, T., d’Espaux, L., van der Horst, C., Dev, I., Rios-Solis, L., Kirby, J., Scheller, H., and Keasling, J. (2018) High-titer production of lathyrane diterpenoids from sugar by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* 45, 142–148.
- (294) Girhard, M., Machida, K., Itoh, M., Schmid, R. D., Arisawa, A., and Urlacher, V. B. (2009) Regioselective biooxidation of (+)-valencene by recombinant *E. coli* expressing CYP109B1 from *Bacillus subtilis* in a two-liquid-phase system. *Microb. Cell Factories* 8, 36.
- (295) Hernandez-Ortega, A., Vinaixa, M., Zebec, Z., Takano, E., and Scrutton, N. S. (2018) A Toolbox for Diverse Oxyfunctionalisation of Monoterpenes. *Sci. Rep.* 8.
- (296) Seifert, A., Vomund, S., Grohmann, K., Kriening, S., Urlacher, V. B., Laschat, S., and Pleiss, J. (2009) Rational Design of a Minimal and Highly Enriched CYP102A1 Mutant Library with Improved Regio-, Stereo- and Chemoselectivity. *ChemBioChem* 10, 853–861.
- (297) Yuan, Y., Litzenburger, M., Cheng, S., Bian, G., Hu, B., Yan, P., Cai, Y., Deng, Z., Bernhardt, R., and Liu, T. (2019) Sesquiterpenoids Produced by Combining Two Sesquiterpene Cyclases with Promiscuous Myxobacterial CYP260B1. *ChemBioChem* 20, 677–682.
- (298) Wong, J., d’Espaux, L., Dev, I., van der Horst, C., and Keasling, J. (2018) De novo synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*. *Metab. Eng.* 47, 94–101.
- (299) Sarker, L. S., Galata, M., Demissie, Z. A., and Mahmoud, S. S. (2012) Molecular cloning and functional characterization of borneol dehydrogenase from the glandular trichomes of *Lavandula x intermedia*. *Arch. Biochem. Biophys.* 528, 163–170.
- (300) Al-Khafaji, J. A., and Shanshal, M. (1975) UV Absorption Spectra of Methyl, Cyclopropyl Ketones. *Z. Für Naturforschung A* 30.
- (301) Milhim, M., Hartz, P., Gerber, A., and Bernhardt, R. (2019) A novel short chain dehydrogenase from *Bacillus megaterium* for the conversion of the sesquiterpene nootkatol to (+)-nootkatone. *J. Biotechnol.* 301, 52–55.

- (302) Muangphrom, P., Seki, H., Suzuki, M., Komori, A., Nishiwaki, M., Mikawa, R., Fukushima, E. O., and Muranaka, T. (2016) Functional Analysis of Amorpha-4,11-Diene Synthase (ADS) Homologs from Non-Artemisinin-Producing *Artemisia* Species: The Discovery of Novel Koidzumiol and (+)- $\alpha$ -Bisabolol Synthases. *Plant Cell Physiol.* *57*, 1678–1688.
- (303) Nakano, C., Kudo, F., Eguchi, T., and Ohnishi, Y. (2011) Genome mining reveals two novel bacterial sesquiterpene cyclases: (-)-germacradien-4-ol and (-)-epi- $\alpha$ -bisabolol synthases from *Streptomyces citricolor*. *Chembiochem Eur. J. Chem. Biol.* *12*, 2271–2275.
- (304) Denisov, I. G., Makris, T. M., Sligar, S. G., and Schlichting, I. (2005) Structure and Chemistry of Cytochrome P450. *Chem. Rev.* *105*, 2253–2278.
- (305) Johnson, E. F., and Stout, C. D. (2013) Structural Diversity of Eukaryotic Membrane Cytochrome P450s. *J. Biol. Chem.* *288*, 17082–17090.
- (306) Pochapsky, T. C., Kazanis, S., and Dang, M. (2010) Conformational plasticity and structure/function relationships in cytochromes P450. *Antioxid. Redox Signal.* *13*, 1273–1296.
- (307) Guengerich, F. P., Waterman, M. R., and Egli, M. (2016) Recent Structural Insights into Cytochrome P450 Function. *Trends Pharmacol. Sci.* *37*, 625–640.
- (308) Girvan, H. M., and Munro, A. W. (2016) Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Curr. Opin. Chem. Biol.* *31*, 136–145.
- (309) Zhao, Y., and Halpert, J. R. (2007) Structure-function analysis of cytochromes P450 2B. *Biochim. Biophys. Acta* *1770*, 402–412.
- (310) Wilderman, P. R., Shah, M. B., Jang, H.-H., Stout, C. D., and Halpert, J. R. (2013) Structural and thermodynamic basis of (+)- $\alpha$ -pinene binding to human cytochrome P450 2B6. *J. Am. Chem. Soc.* *135*, 10433–10440.
- (311) Ebert, M. C. C. J. C., Dürr, S. L., A. Houle, A., Lamoureux, G., and Pelletier, J. N. (2016) Evolution of P450 Monooxygenases toward Formation of Transient Channels and Exclusion of Nonproductive Gases. *ACS Catal.* *6*, 7426–7437.
- (312) Svensson, U. S. H., and Ashton, M. (1999) Identification of the human cytochrome P450 enzymes involved in the in vitro metabolism of artemisinin. *Br. J. Clin. Pharmacol.* *48*, 528–535.
- (313) Pelkonen, O., Abass, K., and Wiesner, J. (2013) Thujone and thujone-containing herbal medicinal and botanical products: Toxicological assessment. *Regul. Toxicol. Pharmacol.* *65*, 100–107.
- (314) May-Manke, A., Kroemer, H., Hempel, G., Bohnenstengel, F., Hohenlöchter, B., Blaschke, G., and Boos, J. (1999) Investigation of the major human hepatic cytochrome P450 involved in 4-hydroxylation and N-dechloroethylation of trofosfamide. *Cancer Chemother. Pharmacol.* *44*, 327–334.
- (315) Chen, C.-S., Lin, J. T., Goss, K. A., He, Y., Halpert, J. R., and Waxman, D. J. (2004) Activation of the Anticancer Prodrugs Cyclophosphamide and Ifosfamide: Identification of Cytochrome P450 2B Enzymes and Site-Specific Mutants with Improved Enzyme Kinetics. *Mol. Pharmacol.* *65*, 1278–1285.
- (316) Li, Y., Drummond, D. A., Sawayama, A. M., Snow, C. D., Bloom, J. D., and Arnold, F. H. (2007) A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments. *Nat. Biotechnol.* *25*, 1051–1056.
- (317) Bāti, G., He, J.-X., Pal, K. B., and Liu, X.-W. (2019) Stereo- and regioselective glycosylation with protection-less sugar derivatives: an alluring strategy to access glycans and natural products. *Chem. Soc. Rev.* *48*, 4006–4018.
- (318) Kovinich, N., Saleem, A., Arnason, J. T., and Miki, B. (2010) Functional characterization of a UDP-glucose:flavonoid 3-O-glucosyltransferase from the seed coat of black soybean (*Glycine max* (L.) Merr.). *Phytochemistry* *71*, 1253–1263.
- (319) Priest, D. M., Ambrose, S. J., Vaistij, F. E., Elias, L., Higgins, G. S., Ross, A. R. S., Abrams, S. R., and Bowles, D. J. (2006) Use of the glucosyltransferase UGT71B6 to disturb abscisic acid homeostasis in *Arabidopsis thaliana*. *Plant J.* *46*, 492–502.
- (320) Dong, T., Xu, Z.-Y., Park, Y., Kim, D. H., Lee, Y., and Hwang, I. (2014) Abscisic acid uridine diphosphate glucosyltransferases play a crucial role in abscisic acid homeostasis in *Arabidopsis*. *Plant Physiol.* *165*, 277–289.

- (321) Morant, A. V., Jørgensen, K., Jørgensen, C., Paquette, S. M., Sánchez-Pérez, R., Møller, B. L., and Bak, S. (2008)  $\beta$ -Glucosidases as detonators of plant chemical defense. *Phytochemistry* 69, 1795–1813.
- (322) Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490-495.
- (323) Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001) Higher plant glycosyltransferases. *Genome Biol.* 2, REVIEWS3004.
- (324) Wilson, A. E., and Tian, L. (2019) Phylogenomic analysis of UDP-dependent glycosyltransferases provides insights into the evolutionary landscape of glycosylation in plant metabolism. *Plant J.* 100, 1273–1288.
- (325) Taguchi, G., Imura, H., Maeda, Y., Kodaira, R., Hayashida, N., Shimosaka, M., and Okazaki, M. (2000) Purification and characterization of UDP-glucose: hydroxycoumarin 7-O-glycosyltransferase, with broad substrate specificity from tobacco cultured cells. *Plant Sci.* 157, 105–112.
- (326) Lim, E.-K., Doucet, C. J., Hou, B., Jackson, R. G., Abrams, S. R., and Bowles, D. J. (2005) Resolution of (+)-abscisic acid using an Arabidopsis glycosyltransferase. *Tetrahedron Asymmetry* 16, 143–147.
- (327) Hughes, J., and Hughes, M. A. (1994) Multiple secondary plant product UDP-glucose glycosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Seq.* 5, 41–49.
- (328) Hiromoto, T., Honjo, E., Noda, N., Tamada, T., Kazuma, K., Suzuki, M., Blaber, M., and Kuroki, R. (2015) Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-O-glycosyltransferase from *Clitoria ternatea*. *Protein Sci. Publ. Protein Soc.* 24, 395–407.
- (329) Warnecke, D. C., Baltrusch, M., Buck, F., Wolter, F. P., and Heinz, E. (1997) UDP-glucose:sterol glucosyltransferase: cloning and functional expression in *Escherichia coli*. *Plant Mol. Biol.* 35, 597–603.
- (330) Martin, R. C., Mok, M. C., Habben, J. E., and Mok, D. W. S. (2001) A maize cytokinin gene encoding an O-glucosyltransferase specific to cis-zeatin. *Proc. Natl. Acad. Sci.* 98, 5922–5926.
- (331) Martin, R. C., Mok, M. C., and Mok, D. W. (1999) A gene encoding the cytokinin enzyme zeatin O-xylosyltransferase of *Phaseolus vulgaris*. *Plant Physiol.* 120, 553–558.
- (332) Takagi, K., Yano, R., Tochigi, S., Fujisawa, Y., Tsuchinaga, H., Takahashi, Y., Takada, Y., Kaga, A., Anai, T., Tsukamoto, C., Seki, H., Muranaka, T., and Ishimoto, M. (2018) Genetic and functional characterization of Sg-4 glycosyltransferase involved in the formation of sugar chain structure at the C-3 position of soybean saponins. *Phytochemistry* 156, 96–105.
- (333) Shibuya, M., Nishimura, K., Yasuyama, N., and Ebizuka, Y. (2010) Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*. *FEBS Lett.* 584, 2258–2264.
- (334) Sayama, T., Ono, E., Takagi, K., Takada, Y., Horikawa, M., Nakamoto, Y., Hirose, A., Sasama, H., Ohashi, M., Hasegawa, H., Terakawa, T., Kikuchi, A., Kato, S., Tatsuzaki, N., Tsukamoto, C., and Ishimoto, M. (2012) The Sg-1 Glycosyltransferase Locus Regulates Structural Diversity of Triterpenoid Saponins of Soybean[W][OA]. *Plant Cell* 24, 2123–2138.
- (335) Noguchi, A., Horikawa, M., Fukui, Y., Fukuchi-Mizutani, M., Iuchi-Okada, A., Ishiguro, M., Kiso, Y., Nakayama, T., and Ono, E. (2009) Local Differentiation of Sugar Donor Specificity of Flavonoid Glycosyltransferase in Lamiales. *Plant Cell* 21, 1556–1572.
- (336) Chen, H.-Y., and Li, X. (2017) Identification of a residue responsible for UDP-sugar donor selectivity of a dihydroxybenzoic acid glycosyltransferase from *Arabidopsis* natural accessions. *Plant J.* 89, 195–203.
- (337) Song, C., Gu, L., Liu, J., Zhao, S., Hong, X., Schulenburg, K., and Schwab, W. (2015) Functional Characterization and Substrate Promiscuity of UGT71 Glycosyltransferases from Strawberry (*Fragaria × ananassa*). *Plant Cell Physiol.* 56, 2478–2493.
- (338) Brazier-Hicks, M., Evans, K. M., Gershater, M. C., Puschmann, H., Steel, P. G., and Edwards, R. (2009) The C-Glycosylation of Flavonoids in Cereals. *J. Biol. Chem.* 284, 17926–17934.
- (339) Yuan, S., Yin, S., Liu, M., and Kong, J.-Q. (2018) Isolation and characterization of a multifunctional flavonoid glycosyltransferase from *Ornithogalum caudatum* with glycosidase activity. *Sci. Rep.* 8, 5886.
- (340) Lim, E.-K., Doucet, C. J., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Ross, J., and Bowles, D. J. (2002) The Activity of *Arabidopsis* Glycosyltransferases toward Salicylic Acid, 4-Hydroxybenzoic Acid, and Other Benzoates. *J. Biol. Chem.* 277, 586–592.

- (341) Lim, E.-K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Jackson, R. G., Taguchi, G., Ross, J., and Bowles, D. J. (2003) Evolution of substrate recognition across a multigene family of glycosyltransferases in *Arabidopsis*. *Glycobiology* *13*, 139–145.
- (342) Jackson, R. G., Kowalczyk, M., Li, Y., Higgins, G., Ross, J., Sandberg, G., and Bowles, D. J. (2002) Over-expression of an *Arabidopsis* gene encoding a glucosyltransferase of indole-3-acetic acid: phenotypic characterisation of transgenic lines. *Plant J.* *32*, 573–583.
- (343) Yang, M., Fehel, C., Lees, K. V., Lim, E.-K., Offen, W. A., Davies, G. J., Bowles, D. J., Davidson, M. G., Roberts, S. J., and Davis, B. G. (2018) Functional and informatics analysis enables glycosyltransferase activity prediction. *Nat. Chem. Biol.* *14*, 1109–1117.
- (344) Masada, S., Kawase, Y., Nagatoshi, M., Oguchi, Y., Terasaka, K., and Mizukami, H. (2007) An efficient chemoenzymatic production of small molecule glucosides with in situ UDP-glucose recycling. *FEBS Lett.* *581*, 2562–2566.
- (345) Pei, J., Chen, A., Zhao, L., Cao, F., Ding, G., and Xiao, W. (2017) One-Pot Synthesis of Hyperoside by a Three-Enzyme Cascade Using a UDP-Galactose Regeneration System. *J. Agric. Food Chem.* *65*, 6042–6048.
- (346) Schmölder, K., Lemmerer, M., and Nidetzky, B. (2018) Glycosyltransferase cascades made fit for chemical production: Integrated biocatalytic process for the natural polyphenol C-glucoside nothofagin. *Biotechnol. Bioeng.* *115*, 545–556.
- (347) Arend, J., Warzecha, H., Hefner, T., and Stöckigt, J. (2001) Utilizing genetically engineered bacteria to produce plant-specific glucosides: Engineered Bacteria to Produce Plant-Specific Glucosides. *Biotechnol. Bioeng.* *76*, 126–131.
- (348) Lim, E.-K., Ashford, D. A., Hou, B., Jackson, R. G., and Bowles, D. J. (2004) *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnol. Bioeng.* *87*, 623–631.
- (349) Kim, B. G., Yang, S. M., Kim, S. Y., Cha, M. N., and Ahn, J.-H. (2015) Biosynthesis and production of glycosylated flavonoids in *Escherichia coli*: current state and perspectives. *Appl. Microbiol. Biotechnol.* *99*, 2979–2988.
- (350) Kim, B.-G., Sung, S. H., and Ahn, J.-H. (2012) Biological synthesis of quercetin 3-O-N-acetylglucosamine conjugate using engineered *Escherichia coli* expressing UGT78D2. *Appl. Microbiol. Biotechnol.* *93*, 2447–2453.
- (351) Pei, J., Dong, P., Wu, T., Zhao, L., Fang, X., Cao, F., Tang, F., and Yue, Y. (2016) Metabolic Engineering of *Escherichia coli* for Astragalin Biosynthesis. *J. Agric. Food Chem.* *64*, 7966–7972.
- (352) Xia, T., and Eiteman, M. A. (2017) Quercetin Glucoside Production by Engineered *Escherichia coli*. *Appl. Biochem. Biotechnol.* *182*, 1358–1370.
- (353) Zhou, W., Bi, H., Zhuang, Y., He, Q., Yin, H., Liu, T., and Ma, Y. (2017) Production of Cinnamyl Alcohol Glucoside from Glucose in *Escherichia coli*. *J. Agric. Food Chem.* *65*, 2129–2135.
- (354) Vanegas, K. G., Larsen, A. B., Eichenberger, M., Fischer, D., Mortensen, U. H., and Naesby, M. (2018) Indirect and direct routes to C-glycosylated flavones in *Saccharomyces cerevisiae*. *Microb. Cell Factories* *17*, 107.
- (355) Olsson, K., Carlsen, S., Semmler, A., Simón, E., Mikkelsen, M. D., and Møller, B. L. (2016) Microbial production of next-generation stevia sweeteners. *Microb. Cell Factories* *15*, 207.
- (356) Irmisch, S., Jo, S., Roach, C. R., Jancsik, S., Yuen, M. M. S., Madilao, L. L., O’Neil-Johnson, M., Williams, R., Withers, S. G., and Bohlmann, J. (2018) Discovery of UDP-Glycosyltransferases and BAHD-Acyltransferases Involved in the Biosynthesis of the Antidiabetic Plant Metabolite Montbretin A. *Plant Cell* *30*, 1864–1886.
- (357) Irmisch, S., Jancsik, S., Yuen, M. M. S., Madilao, L. L., and Bohlmann, J. (2019) Biosynthesis of the anti-diabetic metabolite montbretin A: glucosylation of the central intermediate mini-MbA. *Plant J.* *100*, 879–891.
- (358) Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glössl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* Mycotoxin Deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*. *J. Biol. Chem.* *278*, 47905–47914.
- (359) Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., and Schnitzler, J.-P. (2006) Practical approaches to plant volatile analysis. *Plant J.* *45*, 540–560.
- (360) Louveau, T. (2013, October) Investigation of glycosyltransferases from oat. doctoral, University of East Anglia.

- (361) Louveau, T., Orme, A., Pfalzgraf, H., Stephenson, M. J., Melton, R., Saalbach, G., Hemmings, A. M., Leveau, A., Rejzek, M., Vickerstaff, R. J., Langdon, T., Field, R. A., and Osbourn, A. (2018) Analysis of Two New Arabinosyltransferases Belonging to the Carbohydrate-Active Enzyme (CAZY) Glycosyl Transferase Family1 Provides Insights into Disease Resistance and Sugar Donor Specificity[OPEN]. *Plant Cell* 30, 3038–3057.
- (362) Piochon, M., Legault, J., Gauthier, C., and Pichette, A. (2009) Synthesis and cytotoxicity evaluation of natural  $\alpha$ -bisabolol  $\beta$ -d-fucopyranoside and analogues. *Phytochemistry* 70, 228–236.
- (363) Prakash, I., Markosyan, A., and Bunders, C. (2014) Development of Next Generation Stevia Sweetener: Rebaudioside M. *Foods* 3, 162–175.
- (364) Miley, M. J., Zielinska, A. K., Keenan, J. E., Bratton, S. M., Radomska-Pandya, A., and Redinbo, M. R. (2007) Crystal Structure of the Cofactor-Binding Domain of the Human Phase II Drug-Metabolism Enzyme UDP-Glucuronosyltransferase 2B7. *J. Mol. Biol.* 369, 498–511.
- (365) Meech, R., Mubarakah, N., Shivasami, A., Rogers, A., Nair, P. C., Hu, D. G., McKinnon, R. A., and Mackenzie, P. I. (2015) A Novel Function for UDP Glycosyltransferase 8: Galactosidation of Bile Acids. *Mol. Pharmacol.* 87, 442–450.
- (366) Guillemette, C., Lévesque, E., Harvey, M., Bellemare, J., and Menard, V. (2010) UGT genomic diversity: beyond gene duplication. *Drug Metab. Rev.* 42, 24–44.
- (367) Fang, J.-B., Nikolić, D., Lankin, D. C., Simmler, C., Chen, S.-N., Ramos Alvarenga, R. F., Liu, Y., Pauli, G. F., and van Breemen, R. B. (2019) Formation of (2 R)- and (2 S)-8-Prenylnaringenin Glucuronides by Human UDP-Glucuronosyltransferases. *J. Agric. Food Chem.* 67, 11650–11656.
- (368) Matin, B., Sherbini, A. A., Alam, N., Harmatz, J. S., and Greenblatt, D. J. (2019) Resveratrol glucuronidation in vitro: potential implications of inhibition by probenecid. *J. Pharm. Pharmacol.* 71, 371–378.
- (369) Zhou, D., Kong, L., Jiang, Y., Wang, C., Ni, Y., Wang, Y., Zhang, H., and Ruan, J. (2019) UGT-dependent regioselective glucuronidation of ursodeoxycholic acid and obeticholic acid and selective transport of the consequent acyl glucuronides by OATP1B1 and 1B3. *Chem. Biol. Interact.* 310, 108745.
- (370) Kim, Y., Shrestha, R., Kim, S., Kim, J. A., Lee, J., Jeong, T. C., Kim, J.-H., and Lee, S. (2019) *In vitro* characterization of glycyrol metabolites in human liver microsomes using HR-resolution MS spectrometer coupled with tandem mass spectrometry. *Xenobiotica* 1–9.
- (371) Qi, C., Fu, J., Zhao, H., Xing, H., Dong, D., and Wu, B. (2019) Identification of UGTs and BCRP as potential pharmacokinetic determinants of the natural flavonoid alpinetin. *Xenobiotica* 49, 276–283.
- (372) Drăgan, C.-A., Buchheit, D., Bischoff, D., Ebner, T., and Bureik, M. (2010) Glucuronide Production by Whole-Cell Biotransformation Using Genetically Engineered Fission Yeast *Schizosaccharomyces pombe*. *Drug Metab. Dispos.* 38, 509–515.
- (373) Ikushiro, S., Nishikawa, M., Masuyama, Y., Shouji, T., Fujii, M., Hamada, M., Nakajima, N., Finel, M., Yasuda, K., Kamakura, M., and Sakaki, T. (2016) Biosynthesis of Drug Glucuronide Metabolites in the Budding Yeast *Saccharomyces cerevisiae*. *Mol. Pharm.* 13, 2274–2282.
- (374) Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- (375) Dwivedi, P., Bendiak, B., Clowers, B. H., and Hill, H. H. (2007) Rapid Resolution of Carbohydrate Isomers by Electrospray Ionization Ambient Pressure Ion Mobility Spectrometry-Time-of-Flight Mass Spectrometry (ESI-APIMS-TOFMS). *J. Am. Soc. Mass Spectrom.* 18, 1163–1175.
- (376) Moehs, C. P., Allen, P. V., Friedman, M., and Belknap, W. R. (1997) Cloning and expression of solanidine UDP-glucose glycosyltransferase from potato. *Plant J.* 11, 227–236.
- (377) Poppenberger, B., Berthiller, F., Bachmann, H., Lucyshyn, D., Peterbauer, C., Mitterbauer, R., Schuhmacher, R., Krska, R., Glossl, J., and Adam, G. (2006) Heterologous Expression of Arabidopsis UDP-Glycosyltransferases in *Saccharomyces cerevisiae* for Production of Zearalenone-4-O-Glucoside. *Appl. Environ. Microbiol.* 72, 4404–4410.
- (378) Kamatou, G. P. P., and Viljoen, A. M. (2010) A Review of the Application and Pharmacological Properties of  $\alpha$ -Bisabolol and  $\alpha$ -Bisabolol-Rich Oils. *J. Am. Oil Chem. Soc.* 87, 1–7.
- (379) Russell, K., and Jacob, S. E. (2010) Bisabolol: *Dermatitis* 21, 57–58.

- (380) Wang, H., Yang, Y., Lin, L., Zhou, W., Liu, M., Cheng, K., and Wang, W. (2016) Engineering *Saccharomyces cerevisiae* with the deletion of endogenous glucosidases for the production of flavonoid glucosides. *Microb. Cell Factories* 15, 134.
- (381) Veach, Y. K., Martin, R. C., Mok, D. W. S., Malbeck, J., Vankova, R., and Mok, M. C. (2003) O-glucosylation of cis-zeatin in maize. Characterization of genes, enzymes, and endogenous cytokinins. *Plant Physiol.* 131, 1374–1380.
- (382) Yuejian Liu, and Michael Coughtrie. (2017) Revisiting the Latency of Uridine Diphosphate-Glucuronosyltransferases (UGTs)—How Does the Endoplasmic Reticulum Membrane Influence Their Function? *Pharmaceutics* 9, 32.
- (383) Muraoka, M., Kawakita, M., and Ishida, N. (2001) Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity. *FEBS Lett.* 495, 87–93.
- (384) Rennie, E. A., Ebert, B., Miles, G. P., Cahoon, R. E., Christiansen, K. M., Stonebloom, S., Khatab, H., Twell, D., Petzold, C. J., Adams, P. D., Dupree, P., Heazlewood, J. L., Cahoon, E. B., and Scheller, H. V. (2014) Identification of a Sphingolipid  $\alpha$ -Glucuronosyltransferase That Is Essential for Pollen Function in *Arabidopsis*. *Plant Cell* 26, 3314–3325.
- (385) Goto, S., Taniguchi, M., Muraoka, M., Toyoda, H., Sado, Y., Kawakita, M., and Hayashi, S. (2001) UDP-sugar transporter implicated in glycosylation and processing of Notch. *Nat. Cell Biol.* 3, 816–822.
- (386) Kinosaki, M., Masuko, T., Sogawa, K., Iyanagi, T., Yamamoto, T., Hashimoto, Y., and Fujii-Kuriyama, Y. (1993) Intracellular Localization of UDP-Glucuronosyltransferase Expressed from the Transfected cDNA in Cultured Cells. *Cell Struct. Funct.* 18, 41–51.
- (387) Yang, N., Sun, R., Liao, X., Aa, J., and Wang, G. (2017) UDP-glucuronosyltransferases (UGTs) and their related metabolic cross-talk with internal homeostasis: A systematic review of UGT isoforms for precision medicine. *Pharmacol. Res.* 121, 169–183.
- (388) Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., Kui, L., Cheng, X., Ruan, J., Fan, W., Chen, J., Long, G., Zhao, Y., Cai, J., Wang, W., Ma, Y., Dong, Y., Yang, S., and Jiang, H. (2018) Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. *Nat. Commun.* 9, 448.
- (389) Mayoral, J. G., Nouzova, M., Navare, A., and Noriega, F. G. (2009) NADP<sup>+</sup>-dependent farnesol dehydrogenase, a *corpora allata* enzyme involved in juvenile hormone synthesis. *Proc. Natl. Acad. Sci.* 106, 21091–21096.
- (390) Jindra, M., Palli, S. R., and Riddiford, L. M. (2013) The Juvenile Hormone Signaling Pathway in Insect Development. *Annu. Rev. Entomol.* 58, 181–204.
- (391) Kolwek, J., Behrens, C., Linke, D., Krings, U., and Berger, R. G. (2018) Cell-free one-pot conversion of (+)-valencene to (+)-nootkatone by a unique dye-decolorizing peroxidase combined with a laccase from *Funalia trogii*. *J. Ind. Microbiol. Biotechnol.* 45, 89–101.
- (392) Zelena, K., Krings, U., and Berger, R. G. (2012) Functional expression of a valencene dioxygenase from *Pleurotus sapidus* in *E. coli*. *Bioresour. Technol.* 108, 231–239.
- (393) Jendresen, C. B., and Nielsen, A. T. (2019) Production of zosteric acid and other sulfated phenolic biochemicals in microbial cell factories. *Nat. Commun.* 10.
- (394) Daran, J. M., Bell, W., and François, J. (2006) Physiological and morphological effects of genetic alterations leading to a reduced synthesis of UDP-glucose in *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* 153, 89–96.
- (395) Orlean, P. (2012) Architecture and Biosynthesis of the *Saccharomyces cerevisiae* Cell Wall. *Genetics* 192, 775–818.
- (396) Nelson, G., Hoon, M. A., Chandrashekar, J., Zhang, Y., Ryba, N. J. P., and Zuker, C. S. (2001) Mammalian Sweet Taste Receptors. *Cell* 106, 381–390.
- (397) Servant, G., Tachdjian, C., Tang, X.-Q., Werner, S., Zhang, F., Li, X., Kamdar, P., Petrovic, G., Ditschun, T., Java, A., Brust, P., Brune, N., DuBois, G. E., Zoller, M., and Karanewsky, D. S. (2010) Positive allosteric modulators of the human sweet taste receptor enhance sweet taste. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4746–4751.

- (398) Masuda, K., Koizumi, A., Nakajima, K., Tanaka, T., Abe, K., Misaka, T., and Ishiguro, M. (2012) Characterization of the Modes of Binding between Human Sweet Taste Receptor and Low-Molecular-Weight Sweet Compounds. *PLoS ONE* 7.
- (399) Assadi-Porter, F. M., Maillet, E. L., Radek, J. T., Quijada, J., Markley, J. L., and Max, M. (2010) Key amino acid residues involved in multi-point binding interactions between brazzein, a sweet protein, and the T1R2-T1R3 human sweet receptor. *J. Mol. Biol.* 398, 584–599.
- (400) Xu, H., Staszewski, L., Tang, H., Adler, E., Zoller, M., and Li, X. (2004) Different functional roles of T1R subunits in the heteromeric taste receptors. *Proc. Natl. Acad. Sci. U. S. A.* 101, 14258–14263.
- (401) DuBois, G. E. (2016) Molecular mechanism of sweetness sensation. *Physiol. Behav.* 164, 453–463.
- (402) Chéron, J.-B., Casciuc, I., Golebiowski, J., Antonczak, S., and Fiorucci, S. (2017) Sweetness prediction of natural compounds. *Food Chem.* 221, 1421–1425.
- (403) Hellfritsch, C., Brockhoff, A., Stähler, F., Meyerhof, W., and Hofmann, T. (2012) Human Psychometric and Taste Receptor Responses to Steviol Glycosides. *J. Agric. Food Chem.* 60, 6782–6793.
- (404) Naesby, M., Nielsen, S. V., Nielsen, C. A., Green, T., Tange, T. Ø., Simón, E., Knechtle, P., Hansson, A., Schwab, M. S., Titiz, O., Folly, C., Archila, R. E., Maver, M., van Sint Fiet, S., Boussemerghoune, T., Janes, M., Kumar, A. S. S., Sonkar, S. P., Mitra, P. P., Benjamin, V. A. K., Korrapati, N., Suman, I., Hansen, E. H., Thybo, T., Goldsmith, N., and Sorensen, A. S. (2009) Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microb. Cell Factories* 8, 45.
- (405) Klein, J., Heal, J. R., Hamilton, W. D. O., Boussemerghoune, T., Tange, T. Ø., Delegrange, F., Jaeschke, G., Hatsch, A., and Heim, J. (2014) Yeast Synthetic Biology Platform Generates Novel Chemical Structures as Scaffolds for Drug Discovery. *ACS Synth. Biol.* 3, 314–323.



## Appendix

## A. Supplementary figures

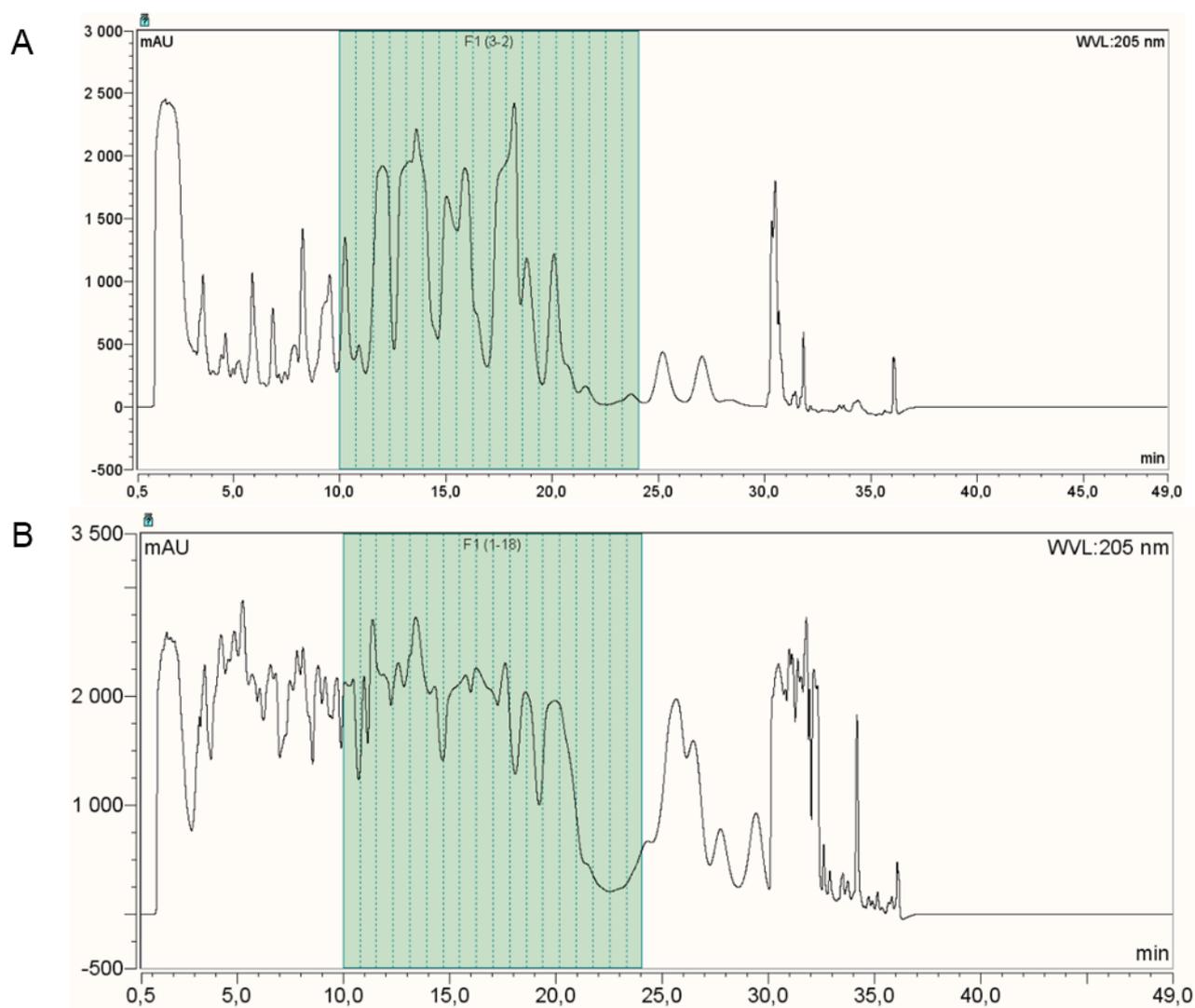


Figure 67: Fractions collected from (A) YBIS\_06 + CYP2B6, (B) YBIS\_06 + CYP2B11. Fractions were sampled from 10 min to 24.1 min with a step of 47 sec. This corresponds to a larger window of the observed separation of the sesquiterpene oxidized product at the “analytical” scale for CYPB6 and CYP2B11 on the HPLC module. The displayed injections correspond to injection of 1/20 of the material obtained after the two first chromatographic steps.

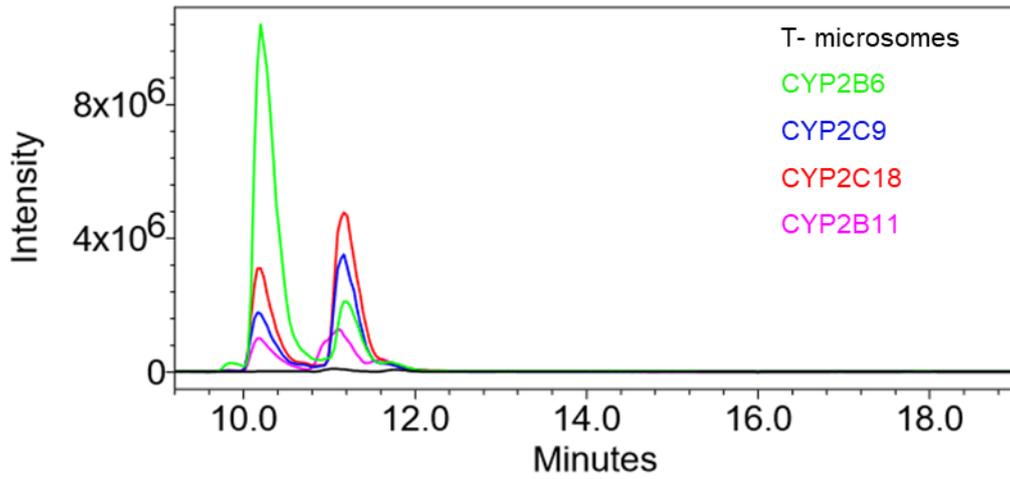


Figure 68: MS chromatograms at  $m/z = 235.3$  showing oxidation activities of the different CYPs with ( $\pm$ )-hernandulcin after 60-minute incubation.

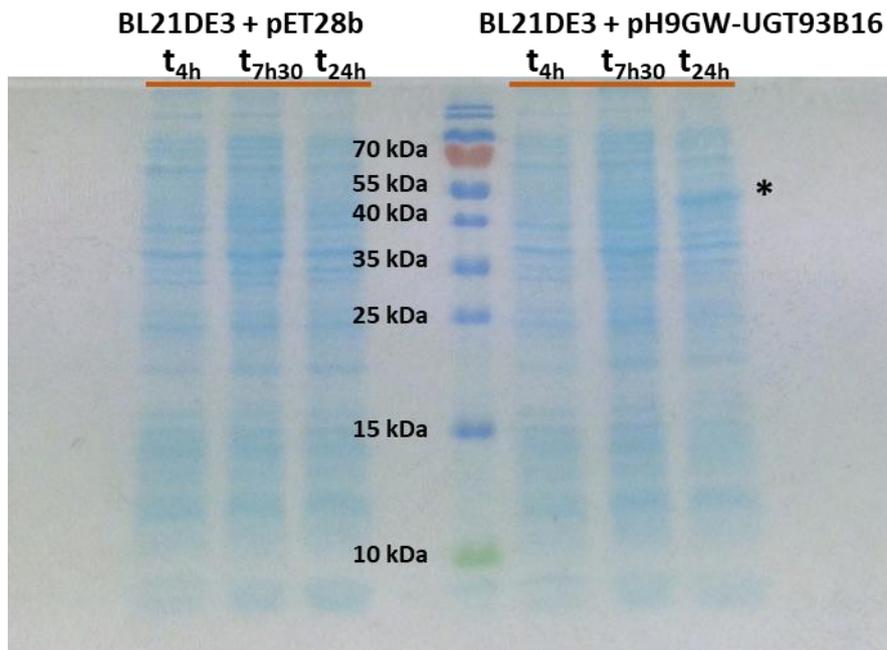


Figure 69: SDS-PAGE (15% polyacrylamide gel) analysis of the whole cell extract showing recombinant UGT93B16 production in *E. coli* after 4h, 7h30 and 24h of culture  
 \* Highlights the presence of UGT93B16.

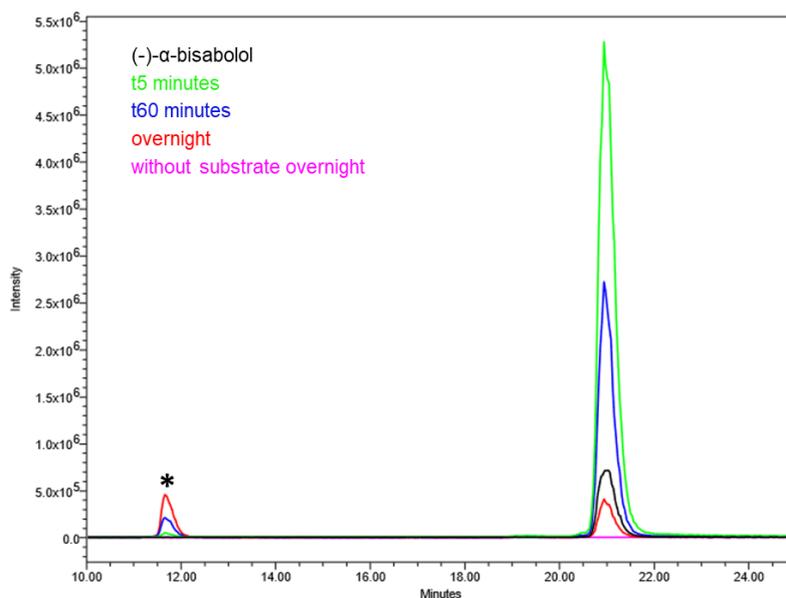


Figure 70: Time course of (-)- $\alpha$ -bisabolol-Glc formation by UGT93B16 detected by MS signal at  $m/z = 205.3$   
 \* Highlights the presence of glucosylated product.

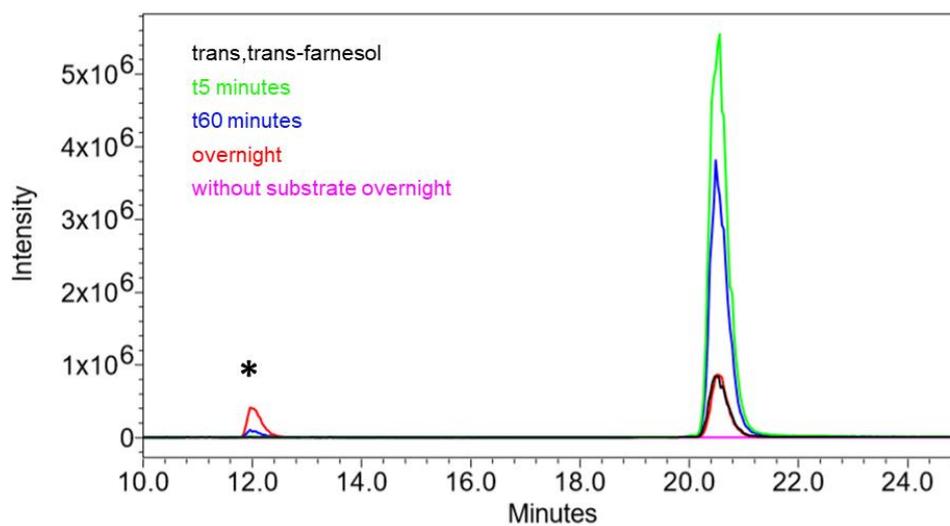


Figure 71: MS signal at  $m/z = 205.3$  showing the time course of trans,trans-farnesol-Glc formation by UGT93B16  
 \* Highlights the presence of glucosylated product.

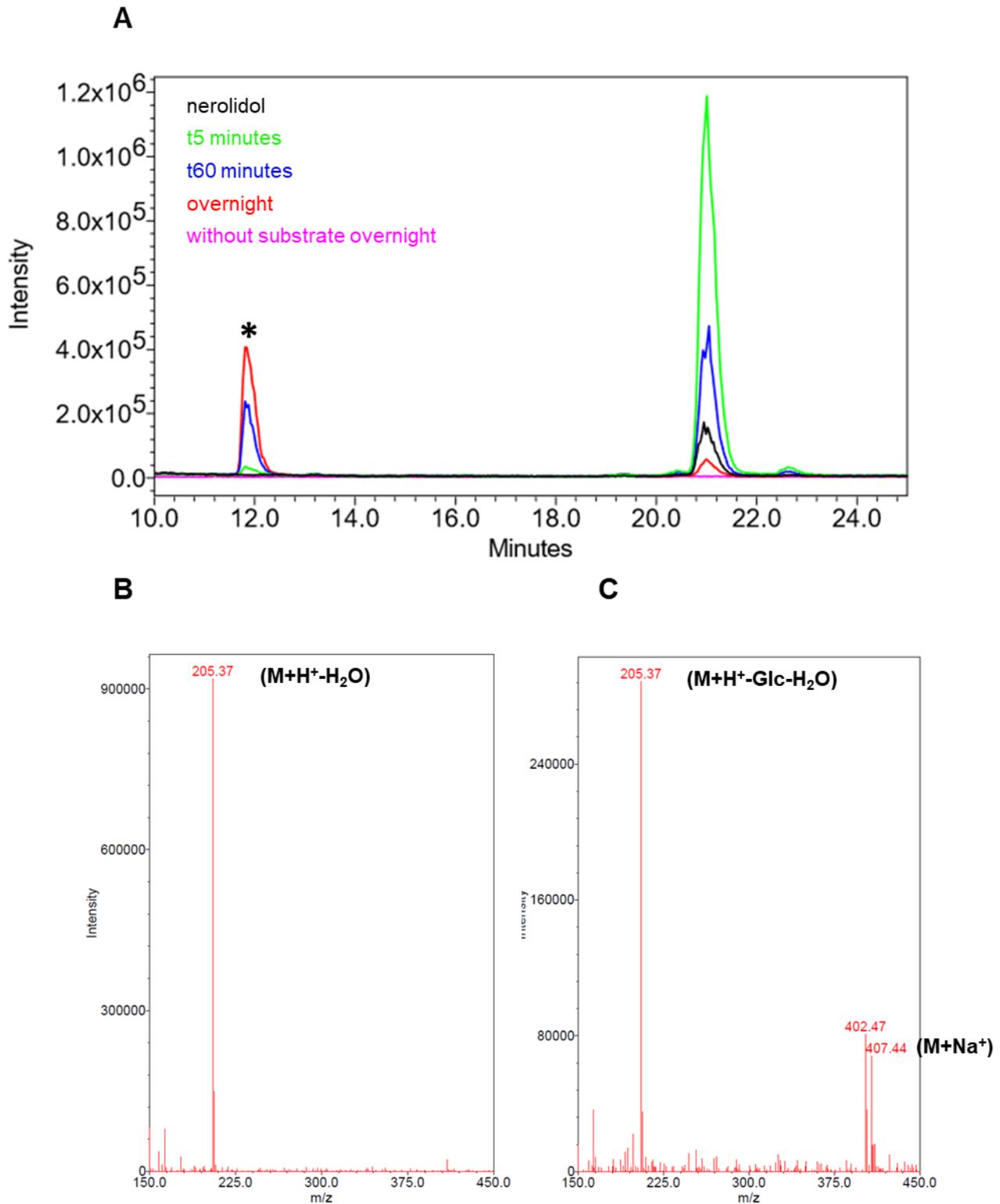


Figure 72: Time course conversion of nerolidol by UGT93B16

(A) MS chromatograms following the reaction at  $m/z = 205.3$  showing nerolidol -Glc formation by UGT93B16

\* corresponds to nerolidol -Glc formation

(B) MS spectra of nerolidol

(C) MS spectra of nerolidol -Glc.

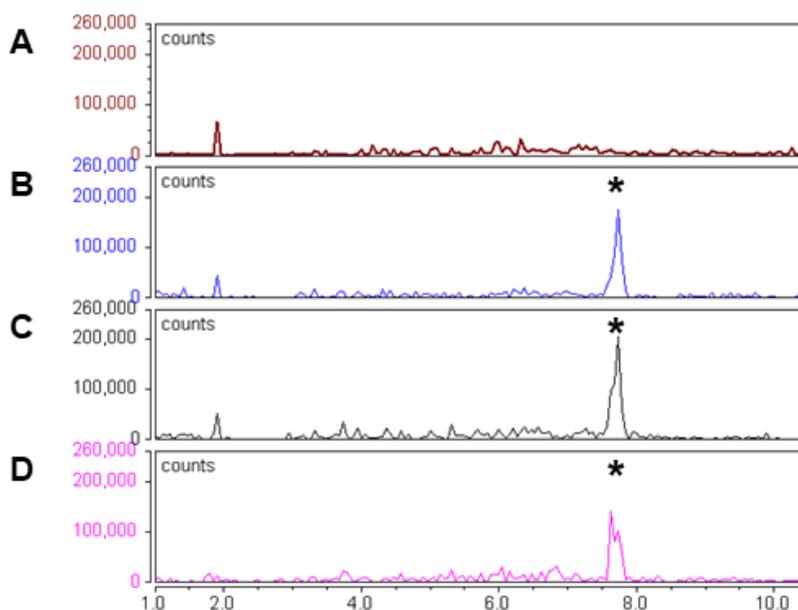


Figure 73: MS signal at  $m/z = 435.3 (M+Na^+)$  showing human UDP-glucuronosyltransferase activities with ( $\pm$ )-hernandulcin (A) Control supersomes without human UDP-glucuronosyltransferase (B) UGT1A9 supersomes (C) UGT2B4 supersomes (D) UGT2B7 supersomes

\* Highlights the presence of a glucuronylated product

This set of injections was performed using an Ultimate 300 LC module (ThermoFischer Scientific) equipped with a MSQ plus mass spectrometer (ThermoFischer Scientific) in same analytical conditions as described in material and methods.

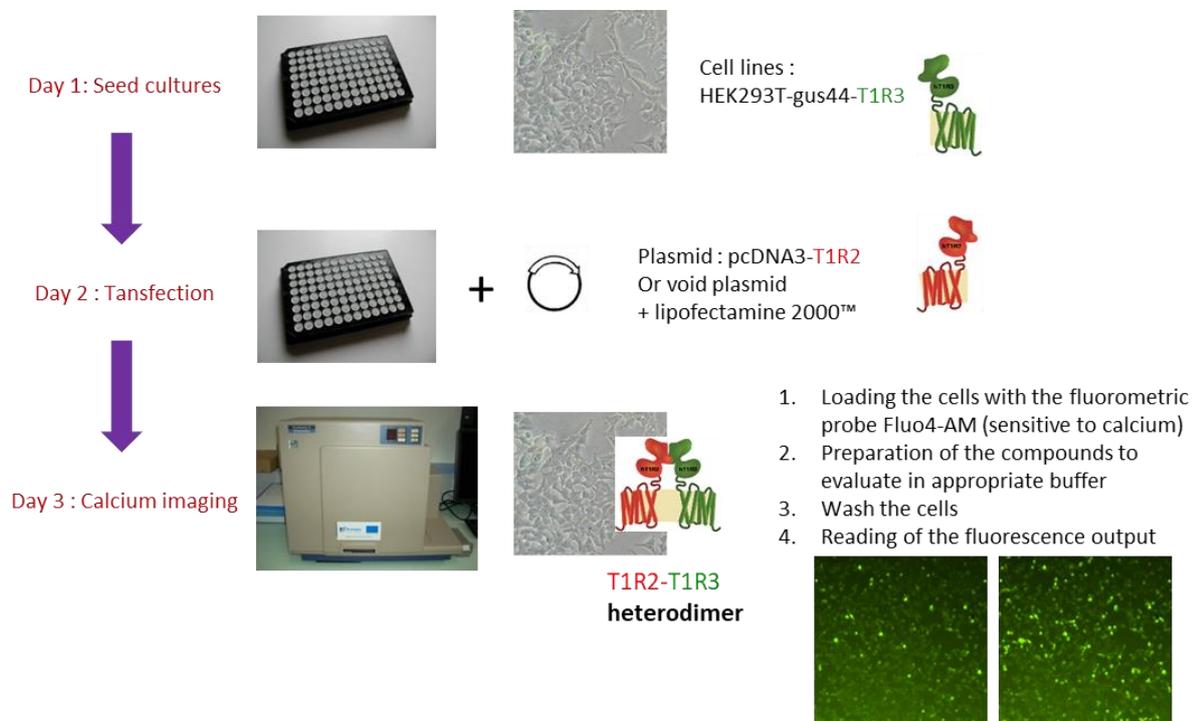
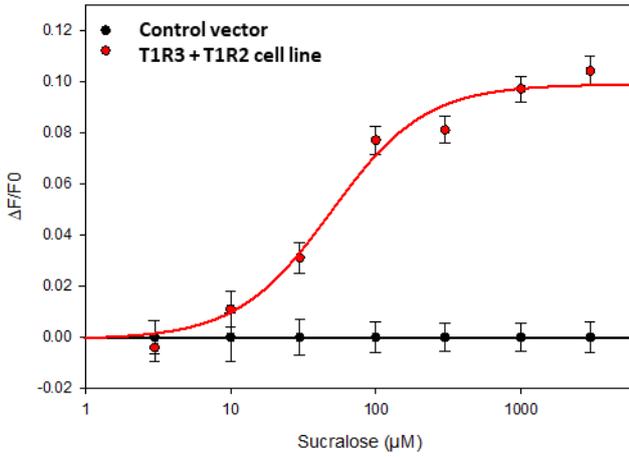


Figure 74: Sweetness potency measurement carried out by the team located at the "Centre des Sciences du Goût et de l'Alimentation"

At day 3 the heterodimer T1R2-T1R3 responsible for sweet taste perception is expressed. Loading sweetening molecules induce a rise of  $[Ca^{2+}]$  which is followed by a fluorescence output.

**A** Sucralose, mean (n=5)

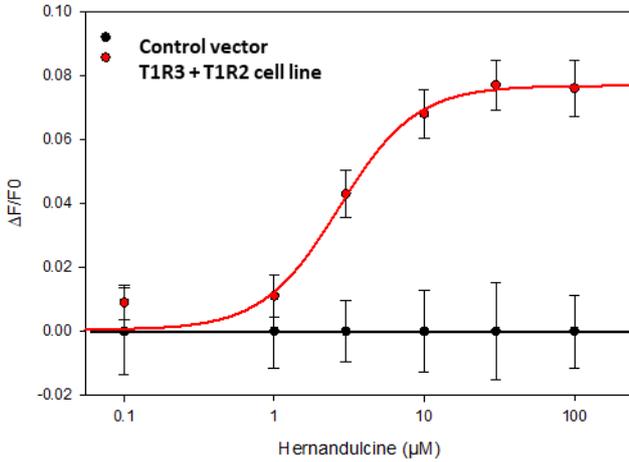


$EC_{50} = 49 \pm 12 \mu\text{M}$

In agreement with  $EC_{50}$  values reported in the literature

Servant et al. (2010)	$EC_{50} = 39 \pm 3 \mu\text{M}$
Masuda et al. (2012)	$EC_{50} = 80 \pm 20 \mu\text{M}$
Assadi-Porter et al. (2010)	$EC_{50} = 65 \pm 5 \mu\text{M}$

**B** ( $\pm$ )-hernandulcin, mean (n=11)



**Activation of hT1R2/hT1R3**  
**With an  $EC_{50} = 2.70 \pm 0.60 \mu\text{M}$**

**C** (-)- $\alpha$ -bisabolol, mean (n=7)

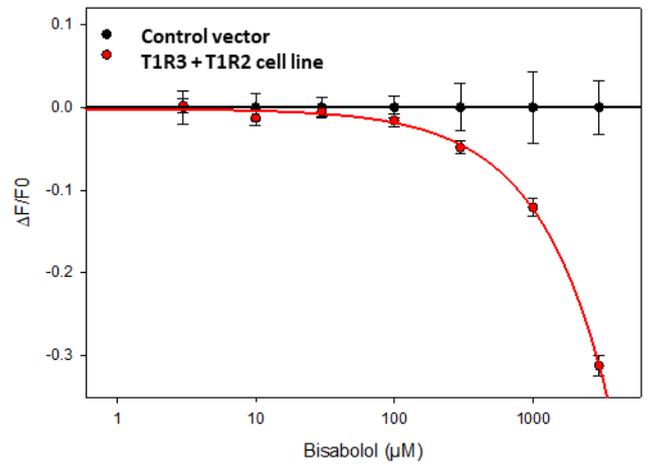


Figure 75: Sweetness potency evaluation of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin measured by the team in "Centre des Sciences du Goût et de l'Alimentation".

## B. Supplementary tables

Table 18 : Primers used in this thesis. Restriction sites are underlined, homology sequences for Gibson cloning are displayed in bold.

Primer number	Sequence	Amplicon	Purpose
2.1	<b>TGGATGTGTAATTCATTTTT</b> GATTGATT TGACTGTGTTATTTTGCCTG	PDC1 promoter	Cloning in pMRI34S+BBS
2.2	<b>ATTTTTGAAAATTCGAATTC</b> CATGCGAC TGGGTGAGCATATG	PDC1 promoter	Cloning in pMRI34S+BBS
2.3	<b>AAAAATGAATTACACATCCAGGAG</b>	pMRI34S+BBS	Cloning in pMRI34S+BBS
2.4	<b>GAATTCGAATTTTCAAAAATTCTTACTT</b> <b>TTTTTTTG</b>	pMRI34S+BBS	Cloning in pMRI34S+BBS
2.5	TATAA <u>AGATCT</u> A <del>AAAAATGACTTCTGCTTT</del> GTATGCTTCC	AtCPR1	Cloning in pMRI31
2.6	TATAA <u>AGTCTGACTC</u> ACCAGACATCTCTGA GGTATC	AtCPR1	Cloning in pMRI31
2.7	<b>GATCTCCCATGTCTCTACTGG</b>	pFPP13	Cloning ADHs in pFPP13
2.8	<b>AAAACCTAGATTAGATTGCTATGCTTTC</b> <b>TTTC</b>	pFPP13	Cloning ADHs in pFPP13
4.1	ACTAAC <u>CGGATCC</u> ATGGGGATTGAGTCG ATGGACAG	UGT93B16 flanked with BamHI site	Cloning in pYEDP60
4.2	ACTAAC <u>CGGATCC</u> ATGGACTACAAAGAC GATGACGAAAGGGGATTGAGTCGATGG ACAG	UGT93B16 flanked with a FLAG tag and BamHI site	Cloning in pYEDP60
4.3	ACTAACGAATTCTCAAATCCTTGTGATG TGAGCAACG	UGT93B16 flanked with EcoRI site	Cloning in pYEDP60

4.4	<b>CGTACACAGCACAGAGTGCCTGAACAG</b> C	Fragment 1 TEF1- UGDH	Cloning in pMRI34
4.5	<b>AGCGGATCTTAGCTAGCCGCGGATAGC</b> TTCAAATGTTTCTACTCCTTTTTTAC	Fragment 1 TEF1- UGDH	Cloning in pMRI34
4.6	<b>GGCACTCTGTGCTGTGTACGAGCACTG</b> G	Fragment 2 TEF1- UGDH	Cloning in pMRI34
4.7	<b>TCTGGCGAAGAATTGTTAATTA</b> ACTAG ACTTTGGGCTTCTTGTTAG	Fragment 2 TEF1- UGDH	Cloning in pMRI34
4.8	CTGACCTAGATCTATGTCCGTTAAGTGG	UGT2B7	Cloning in pYEDP60
4.9	TAGCCTCCCGGGTCAGTCGTTCT	UGT2B7	Cloning in pYEDP60
4.10	<b>GCATAGCAATCTAATCTAAGTTTTAATT</b> ACAAAATGTCTATGAGTAGGGGAGGCA AC	FRC	Cloning in pFPP13
4.11	<b>CCAGTAGAGACATGGGAGATCCTACAC</b> ATTTTCTCCCCTCGTAG	FRC	Cloning in pFPP13
4.12	<b>GCATAGCAATCTAATCTAAGTTTTAATT</b> ACAAAATGGCAGAAGTGACAGAAGG	UGTRel7	Cloning in pFPP13
4.13	<b>CCAGTAGAGACATGGGAGATCTCAAAC</b> TGCACCCTTCCCCTTG	UGTRel7	Cloning in pFPP13

Table 19: Additional strains generated for the (+)-*epi*- $\alpha$ -bisabolol oxidation.

Strain Name	Genotype
YBIS_07	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1-ScCPR-TCYC1, $\Delta$ DPP1:: TADH1-BBS-PDC1-PGAL1-ScCPR-TCYC1
YBIS_08	EPY300, $\Delta$ ho::TADH1-BBS-PGAL10-PGAL1- AtCPR1-TCYC1
YBIS_09	EPY300, $\Delta$ ho::TADH1-AkBBS-PGAL10-PGAL1- -TCYC1
YBIS_10	EPY300, $\Delta$ ho::TADH1-MrBBS-PGAL10-PGAL1- -TCYC1
YBIS_11	EPY300, $\Delta$ ho::TADH1-ScBBS-PGAL10-PGAL1- -TCYC1

## C. DNA sequences

### 1. Synthetic fragments of the additional CYPs

>CYP726A20 codon optimized

```
ATGGAGCATCAAATTCTATCCTTTCCCGTCTTGTTCTCATTGCTATTATTCATCCTGGTACTACTGAAAGTATCCAAAAG
CTTTACAAGCACGATTCTAAGCCTCCACCGGGTCCGTGGAAATTACCTTTTCATAGGCAATTTAATACAGCTTGTGGGTGA
CACCCCGCACAGGAGGCTGACCCGCTAGCGAAGACTTACGGCCCCGTCATGGGCGTTCAGCTTGGACAGGTCCCTTT
CCTGGTAGTTTCCAGCCCGGAGACTGCGAAAGAAGTTATGAAGATAACAAGACCCTGTCTTTCGCGAGAGGCCATTGGT
TCTTGCTGGAGAAATAGTTTTATACAACCGTAACGATATTGTATTTGGCAGTTATGGGGATCAATGGAGGCAAATGAGA
AAATTCTGTACTTTGGAATTACTTTCTACGAAGAGGGTTCAAAGCTTCAGACCTGTAAGGGAAGAGGAAGTTGCTTCAT
TTGTGAAGCTAATGCGTACGAAAAAAGGAACCCCGTGAACCTTACACACGCACTGTTTGCACCTACAAATTCAATCGT
AGCTAGAAACGCGGTGGGTCACAAGTCCAAGAACCAAGAGGCCCTGTTAGAGGTCATTGACGACATCGTGGTGAGCG
GAGGGGGGGTCTCAATAGTTGACATTTTCCCATCCCTACAGTGGTTGCCGACTGCAAAGAGAGAGAGGTCAAGGATCT
GGAAGCTACACCAAATACCGACGAGATTCTTGAAGATATACTACAAGAGCACCGTGCCAAGCGTCAGGCAACTGCCT
CCAAGAACTGGGATAGGTCTGAGGCCGATAATCTACTAGACGTGCTTTTGGATTTACAACAAAGCGGAAATCTAGATGT
CCCCCTGACTGATGTTGCCATAAAGGCCGCAATTATCGATATGTTCCGGGGCTGGGAGCGATACTAGCTCAAAGACAGC
GGAGTGGGCCATGGCCGAGCTGATGAGGAACCCTGAAGTCATGAAGAAGGCTCAAGAGGAGTTACGTAATTTCTTTG
GTGAAAACGGAAAGGTAGAAGAGGCTAAGTTGCATGAACTAAATGGATTAATTGATAATTAAGGAGACCTTGCGTC
TGCACCCGCGGTTGCTGTAATTCCTAGGGTGTGCAGGGAGAAAATAAAGTCTACGGGTACGACGTAGAGCCAGGG
ACAAGAGTTTTATAAATGTTTGGAGTATAGGTAGGGACCAAAGGTGTGGAGTGAGGCTGAAAGATTTAAGCCTGAA
CGTTTTATCGATTAGCTATAGATTACAGGGGTCTTAATTTGAGCTTATTCCTTTGGGGCGGGCAAACGTATTTGCC
GGGGATGACTCTGGGTATGGCTAATCTGAAAATATTCCTGGCGAATTTATTGTACCACTTCGATTGGAAGTTTCCAAG
GGCGTAACAGCTGAGAATTTAGATATGAACGAGGCCTTCGGCGGGGCTGTGAAGCGTAAGGTCGATCTAGAATTGATT
CCTATACCCTTTAGGCCATAA
```

>CYP71D495 codon optimized

```
ATGCTGTTCTTTATTACGGTGCTTTTTATCTTTATCGCACTTAGGATATGGAAAAAAGTAAGGCAAACCTCTACACCAA
CCTTCCCCCTGGTCCAATAAATTGCCTCTGATAGGAAATGCCACAATTTAGTTGGTGATTTGCCCTACCATAGATTGA
GGGATCTATCAAAGAAGTACGGTCCAATTATGCACCTTCAGTTAGGAGAAAACACAACCTGTGGTTATTTTCATCACC
ACTTGCCCAAGAGGTTATGAAGACGCACGACGTGAACCTTCGCACAAAGGCCTTTTGTATTAGCAGGTGACATAGTGAG
CTATAAATGTAAAGATATCGCGTTTGCTCCCTATGGTGAGTACTGGCGTCAACTTAGAAAGATGTGCTCCCTTGAGCTG
CTTACCGCGAAGCGTGACAGTCATTCAAATCCATCAGGGAAGAAGAAGTCAGTAAATGGTAGAGTCTATAAGCAGC
TCATCCGTTTCCCCATAAATTTTTCTAAAATGGCCTCCTCCTAACATATGCGATTATATCTAGAGCAGTTTGTGGGAAG
GTATCAAGAGGCGAAGAGGTTTTTGTACCAGCTGTAGAAAAGTTAGTGGAGGCTGGGAGAAAGTATCTCACTAGCAGAT
TTATACCCAAGCGTTAAACTGTTCAACGCGTTGAGTGTGTGAGGAGACGTGTCGAAAAAATTCACGGGGAAGTAGAT
AAAATCATCGAGAATATCGTTATCGAACATAGGGAGAGAAAGCGTATGGCGCATGCGGGTATTAATTCCAAGGAAGA
GGAAGATTTAGTAGACGTATTACTGAAGTTTCAAGAAAACGGAGATTTGGATTACTACTGTCAAATGATGGTATCAAA
GCAGTCATCTTAGATATGTTTCATTGCAGGTAGTGACACTAGTTCAACTACTATAGAGTGGGCGATAAGTGAAATGGTTA
AGAATCCGTCATCATGGAGAAGGCTCAGGCTGAAGTTAGAGAAGTATTCGGTTCCAAGGGTAAAGTAGATGAGGCC
GATCTGCACGAGCTTAACTACCTGAAACTAGTCATTAAGAAACGCTTAGGCTACATCCCGCCGTGCCTTTACTATTACC
CCGTCAAAGTAGAGAGGATTGTGTTATCGAAGGATATAATATCGCTACCAAGTCTACGGTTATTGTAATGCCTGGGCT
ATAGCTAGGGACCCTAAATATTGGGACGAGGCCGAACGTTTCTATCCAGAAAGGTTTATAAACAGCTCAATTGACTTTA
AGGGAACCAATTTGAGTTTATACCGTTTGGGGCAGGGCGTAGGATGTGCCCTGGTATGCTATTTGGGCTGGCCAGCG
TAGAGTTGCCTCTTGACAGTTGCTGTATCACTTCGATTGAAATTGCCGGGTGGACAGAAGCCTGAGGATCTGGACAT
GAGCGACGATTTAGACGGCACCGCAACCAGGAGGCATGCACTATACCTGACAGCTACACCTTATCTTCTTCTGCGGTC
GGCAAATTAGTCGTTAG
```

>CYP706M1 codon optimized

ATGGATATGTCTACCATTTGGTACTACTGGGTAAGCATAATACTGGGGGTCTTCATCTTTCTTATAGTCGGCATCCAAAA  
ATGGAGAAGCAAGAAACTGCCACCTGGCCCTTTTGCCTGCGTGGTGGGACACCTGCACCTGTTGGAGCCCAATGTC  
CATGAGTGCCTGAGTAAAATAAGCGAGAAGTTCGGCCCTTTGATGTCAATTAAGTTTGGTATGAAAACCTCAATTATCG  
TGTCTTCCCCTGCAATGGCAAAGGAGATACTTAGGGAGAACGATCAAATCTTTGCAAACAGAAGCATAACCCGTAGTTGC  
AAGATGTATCGCTTATGACGCCTCCGATATTCTATGGTCCCAAACGGCCCGCTTGGCGTTTACTTAGAAAAATATGTG  
TCAAGGAGCTTTTCAGTCCTAAGTCTACAGAGGCACTACAGCCATTAAGACGTGAAGAGGTGAGGAGGACAATGGGG  
AACATTTATAAAGATTCCATAAATGGTGTTCCTGATAGTGTAGGGGCCAAAGCTTTCATAACAAGCCTTAATCTAATAAC  
GAACATGATGTGGAGTACGTCTACTGAAACGGGCGAAAGAGGGGGTGAAGTCAAGGACTTAGTGGGGGAATTAGTCC  
ATGTTCTAGGGGTCCCAAACGCCTCCGATCTGTTCCATTCCCTTGAAGGTTGACGTCCAGGGCCTTTATAGACGTATG  
GAAAAAGTTTTCGTTAGGTTGATAAGATGTTTACGGCATCATAGAGGACAACTTAGTGCAAGTCTAAAGAGAAA  
GACTTTCTGCAATCCCTGCTAGACCTGGTAGAAAGAGGGGTAGATGAGCAGGACCCCGATTAGTACAACCTTACGATG  
AAAGATGTAAGAGTGTGCTGATGGACATGGTTACAGGTTCAACTGACACGACGTCCAACACAGTTGAATGGGCTATG  
GCCGAGTTGTTACAACAACCTGAGATAATGAAACGTGCCAAAAAGAAGTGAAGAGGTGCTAGGATTAGATAACATG  
GTCGAAGAATGTCACCTTTCTCAACTACCATACTTAGATATTATCGTCAAGGAGGTATTGAGATTACATCCTGCTTTGCC  
TCTACTGGCTCCTCACAGGCCGAGAGAGAGTGTGAGATAGGCGGATACATTATACCAAGGATACCAAGTACTTAT  
CAACGTTTGGTCAATCCAAAGGAATCCTAAAGTTTGAAGGAGCCTTACTGTTGACCCAGAAAGATTCTCCGATAGC  
AAATGGGACTACAACGGCCGTGATTTGACTACTTCCCCTTCGGGTCTGGTCTAGGATTGCGCGGGCTTGAGTATGG  
CGAAGATAATGGTTCATATAGCCTGGCATCCCTACTTCACTCTTTCGATTGGTCTTGCAGTGGCTGAAAAATTGAAC  
ATGGATGAGAAGTACGGTATTGTATTGAGGAAAGCAGTTCCTTGTGGCACTTCCCAAACCGAGATTGCTTTACCCCA  
ACCTATATGAGTAG

>CYP71BE5 codon optimized

ATGGAGTTGCAATTTTCTTTTTTCCCAATCCTATGTACCTTCTGTTATTTATCTATTTACTTAAAAGGTTAGGCAAACCT  
CCCGTACTAATCACCCGGCACCAAATGGCCCCCGGACCCTGGAAGCTGCCTATTATAGGCAACATGCATCAACTGGT  
CGGGTCACTACCGCACCGTTCCTGAGATCTTTCGAAAAAACACGGGCCACTAATGCACCTTCAATTAGGAGAGGTC  
AGCGCTATCGTCGTAAGCAGTAGGGAAATGGCGAAAGAGGTAATGAAGACACACGATATAATCTTTTCTCAAAGACCG  
TGTATTCTAGCCGCTTATTGTCAGCTATGATTGCACAGACATTGCGTTTGCCTCGTATGGCGGCTACTGGCGTCAGAT  
AAGAAAAATCAGCGTATTAGAGCTGTAAGCGCCAAGCGTGTTCAGAGCTTCAGGTCGTCAGGGAAGAAGAGGTATT  
GAATTTAGTCCGTTCCGTGTCACTTCAGGAGGGGGTCTAATAAACCTGACCAAGTCCATATTTTCTTTGACCTTCTCAAT  
AATAAGCCGTACAGCTTTTGGCAAAAAGTGAAGGATCAGGAGGCGTTTGTGACTTTGGATAAGTTCGCCGATTC  
AGCGGGGGGGTTTACGATCGCTGATGTTTTCCCATCCATCAAAGTGTACATGTCGTGAGTGGGATGAGAAGAAAGCT  
TGAGAAAGTCCAAAAAATTAGACAGGATCTTAGGCAACATCATTAATGAGCATAAGGCTAGATCTGCAGCGAAAAGA  
AACGTGTGAGGCGGAAGTGGATGATGATCTGGTTGATGTCTTGTGAAAGTACAGAAACAGGGCGATCTAGAATTTCC  
ATTGACAATGGATAACATAAAAGCTGTCTTGTGGACCTATTTGTTGCCGGCACTGAAACAAGTAGCACTGCGGTAGAA  
TGGGCGATGGCCGAGATGCTAAAAAACCAAGAGTGTATGGCCAAAGCTCAGGCTGAAGTTAGGGACATATTTTCCAG  
GAAAGGGAATGCAGATGAGACGGTAGTAAGAGAGCTAAAATTCCTGAAATTAGTGATAAAGGAAACCTTAAGGCTGC  
ATCCGCCCGTACCGCTATTAATTCCTCGTGTAGAGCAGGGAGAGGTGCGCAATCAACGGATATGAGATTCCAGTCAAGA  
CCCGTGTGATAATAAACCGCTGGGCTATCGCACGTGACCCGAAATATTGGACGGATGCTGAGTCTTTCAACCCGGAGA  
GGTTTCTTGATTCTCCATTGATTACCAGGGAACGAATTCGAGTATATACCCTTTGGAGCTGGGAGACGTATGTGTCCC  
GGTATCCTATTCGGTATGGCTAATGTGGAGCTGGCACTGGCCCAATTGCTGTATCATTTCGATTGGAAACTACCAATG  
GAGCACGTCATGAAGAGTTAGACATGACCGAAGGCTTTAGAAGTACTAAGCGTAACAAGACCTTTACTTGATACC  
CATCACATACAGGCCGCTTCCCGTCGAGTAA

## 2. Synthetic fragments corresponding to the tested ADHs

>AaADH codon optimized

```
ATGGCACAAAAGGCACCTGGCGTTATTACATGTAAAGCTGCTGTGGTCTGGGAATCATCTGGCCCAGTAGTTTTGGAA
GAAATAAGAGTTCGATCCTCCAAAAGCATCTGAAGTGAGAATAAAGATGTTATGCGCAAGTTTGTGTCATACAGACGTTT
TATGTACTAAGGGTTTCCCTATTCCTTTATTTCTAGAATCCCTGGTCATGAGGGCGTTGGTGTGATTGAAAGTATAGGT
AAAGATGCTAAGGGTTTGAACCTGGTGACATTGTTATGCCTTTGATTTAGGTGAATGCGGTCAATGCTTAAACTGTA
AGACTGGCAAAACAAATTTATGTCATGTATACCCACCTTCTTTCTCTGGATTGATGAATGATGGAACCTCAAGAATGTCA
ATTGCACGTACAGGAGAGTCCATTTATCACTTTGCCTCTTGTAGTACTTGGACTGAATACGCTGTTGCTGACTGTAECTA
TGTCTTGAAGATTAACCCCTAAAATTTACATACCTCATGCCTCCTTCTAAGTTGCGGCTTTACTACTGGATTGCGTGCTAC
ATGGCGTGAAACCAAGTCTCAAAGGGCAGTTCTGTAGCCGTGTTCCGAATTGGTACAGTTGGTTTAGGAGTGATTAA
AGGAGCTCAGTTGCAAGGCGCATCTAAGATTATCGGCGTTGATGTCAACCAATACAAAGCTGCAAAAGGAAAAGTTTT
TGGCATGACTGATTTTATCAATCCTAAAGATCATCCAGACAAAAGTGTTCAGAATTGGTTAAAGAGTTGACACATGGC
TTGGGCGTCGATCACTGTTTCGAATGCACAGGCGTTCCTAGTTTATTGAACGAAGCTTTGGAAGCTAGTAAGATAGGCA
TTGGTACAGTAGTCCCAATTGGAGCAGGTGGTGAAGCAAGTGTGCCATTAACAGTTTGATTTTGTCTGGTAGAAC
ATTGAAATTTACTGCATTTGGCGGAGTAAGAACACAATCTGATTTGCCTGTGATTATTGACAAATGTTTGAATAAGGAG
ATTCAGTTAGATGAATTGTTAACACATGAGATACATTTGGACAACATCCAGGAAGCCTTTGAGATATTAAGAAAACCG
ATTGCGTCAAAATATTGATTAATTTCTAA
```

>EiADH codon optimized

```
ATGAATGGTTGTTGTTCTCAGGACCCACGTCCAAAAGATTAGAAGGAAAAGTCGCGGTAAATCACGGGAGGGGCGAG
TGGAATCGGCGCGTGTACTGTTAAGCTTTTTCTGAAGCACGGGGCTAAGGTTGTTATCGCCGATGTCCAAGACGAGTT
GGGACATTCCTTGTGTAAGAAATAGGCAGCGAAGACGTAGTCACTTATGTACATTGCGACGTATCTCCGATAGCGAT
GTTAAAAATGTGGTAGACTCTGCGGTAAGTAAATACGGGAAGCTTGATATTATGTTCTCAAACGCCGGCGTTAGCGGT
GGTTTAGACCTCGTATATTAGCGACAGAAAATGACGAATTTAAGAAGGTATTTGAGGTTAATGTTTTCGGTGGGTTTT
TAGCCGCCAAACACGCTGCTAGGGTCATGATACCTGAAAAGAAAGGTTGTATATTATTCACGTCTCCAACAGCGCTGC
AATCGCCATTCTGGACCCATTACATCGTTGTGTCCAAGCACGCGCTGAATGGTTAATGAAGAACCTAAGTGCAGAG
TTAGGACAGCATGGGATAAGGGTGAATTGCGTCTCACCATTGCGCGTAGTAACGCCTATGATGGCTACGGCATTGGC
ATGAAGGATGCTGATCCTGAGGTCGTAAGGCAACCATAGAAGGTCTGTTAGCTAGTGCCGCTAACCTAAAGAAGTA
ACACTAGGCGCTGAGGATATAGCTAACGCCGCTTGTACCTGGCAAGCGACGAGGCAAAGTATGTTAGCGGCCTAAC
CTGGTGGTAGATGGTGGTTACTCCGTTACCAACCCATCCTTACGGCAACCCTACAAAAGCATTGCGAGTTGCTCATGT
GTAA
```

>JcADH codon optimized

```
ATGGCAAGCTCTTCTAGCCCGGCGCCTACCGCAAAACGTCTAGAGGGCAAGGTAGCCCTAATAACGGGGGGGGCGAG
CGGGATCGGCGAATGCACGGCGCGTCTGTTCCGACAGATGGTGCCAAAGTATTATCGCAGACGTACAATCAGAATT
GGGTAGAAGTGTGGCAGAGAAAATTGGTTCAGAAACAGGGCAGCCTGTGACCTATGTTGATTGTAATGTGACAGTGG
AGTCTGATGTGAAAATGCAGTGAACACCGCCGTTTCTTTACACGGGAAATTGGATATCATGTTCAATAACGCGGGGAT
TGCAGGCAACAACCATGATAAGATCCTTAGCACTGAGCGTGAGGATTTTATGCGTGTACTTGACATTAACATATATGGC
GGAGTCTTGGGGGCCAAACATGCAGCTAGAGTGATGATTCCGGAAAAGAAAGGGTGCATACTTTTTACTGCATCAGTT
TCATCTGCTTTGTATGGAGGACCGTACGCCCTACACGGCATCTAACATGCCGTTGTTGGGCTTGCCAAGAATTTAGCCAT
AGAGCTTGGTCAACACGGCATAACGTGTCAATTGCATTAGTCCGGGAGCTGTCCAGACTGGCCTTGCTAAACAATTGGG
ACTATCAGAACAGCAGGTGCAAGAGTGGTCTAGTGCCTGGCTAACTTAAAAGTGGTGAAGTGAAGTCAACGATAT
AGCTGAAGCGGCGTTATACTTGGCATCTGATGATAGTAAGTTTGAAGTGGATTAATCTGTTGGTTGACGGAGCTGCC
AGTCTGCCGACCACGACGAGGGCATACTAG
```

>LiADH codon optimized

```
ATGGCGTCAACGGTATTGAGGAGACTGGAGGGAAAAGTCGCATTGATAACGGGTGCCGCTAGTGAATCGGCGAATC
AGCCGCGAGGCTTTTTCTCGTACGGTGCCAAAGTCGTGATTGCTGATATACAAGACGAATTGGCCCTTAATATTTGCA
AAGATCTAGGCTCTACCTTCGTCCACTGCGACGTTACAAAAGAGTTCGATGTGAAACCGCAGTCAACACCGCGGTGAG
```

CACTTATGGGAAATTGGATATAATGTAAATAACGCCGGCATCTCCGGCGCACCTAAATACAAAATATCCAATACTCAA  
CTTAGTGATTTTAAACGTGTTGTGGACGTTAACCTAGTAGGCGTGTCTTAGGCACCAAGCACGCGGGGAGGGTAATG  
ATACCAAACCGTAGCGGCTCTATCATCAGTACCGCGTCTGCGGCAACCGCTGCAGCGGGCGGCACGCCGTACCCATAC  
ATCTGTTCCAAGCACGGAGTGGTAGGATTGACTCGTAATGCTGCAGTGGAGATGGGTGGACATGGGATACGTGTAAT  
TGCGTAAGTCCCTATTACGTAGCGACGCCGATGACTCGTGATGATGATTGGATTCAAGGATGTTTCTCCAACCTGAAGG  
GCGCCGTCCTGACCGCCGAAGATGTGGCTGAAGCCGCGCTATATTTAGCTTCCGACGAGTCAAAGTATGTTTCCGGACA  
TAATTTGCTGGTTGACGGCGGCGTTTCAATAATGAATCAGGGTTGTAATATGTTTGACTTAATGGATAGCTGA

>MpADH codon optimized

ATGGCAAGCGTAAAGAAGTTGGCCGGGAAGGTAGCCATTGTAACAGGGGGCGCATCAGGTATCGGAGAGGTGACTG  
CTAGATTGTTTGGGAGAGAGGGGCTAGAGCAGTCGTCATTGCAGATATGCAACCAGAAAAAGGTGGCACTGTTGCG  
GAGAGTATTGGCGGTCGTCGTTGCTCCTACGTGCATTGCGACATTACGGATGAACAGCAGGTCAGGTCTGTGGTTGATT  
GGACGGCGGCTACCTACGGTGGGGTCGATGTAATGTTTTGCAATGCCGGAACCGCTAGTGCTACCGCGAAACTGTTT  
TGGACCTTGATCTTGCTCAGTTTCGATAGAGTAATGCGTGTTAACGCTAGAGGGACAGCCGCTTGTGTTAAGCAGGCAG  
CGAGGAAAATGGTTGAGTTGGGCCGTGGCGGAGCTATCATATGCACCGCCTCAGCGACTGTGCATCACGCTGGACCGA  
ATCTAACTGACTACATCATGTCAAATGTGGCGTCTGGGTCTAGTCAGATCAGCCTCTTTCGAGCTGGGCGTACACGG  
GATCAGGGTGAACCTCTGTATCACCCACCGCACTGGCTACACCGCTAACTGCCACTATTGGACTTAGGACAGCGGCAGAT  
GTCGAGTCTTTCTACGGACAGGTTACAAGTCTAAAGGGAGTGGCGATAACGGCTGAACATGTCGCTGAGGCCGTCGCC  
TTTCTAGCGTCAGATGAGGCTGCCTTTGTCACAGGGCACGACCTAGCTGTCGACGGCGGGCTACAGTGCTTACCCTTTG  
TCGCTGTTGCTAAATGA

>VoADH codon optimized

ATGACCAAGAGTCCGGTGAGGTAATCAGTTGCAAAGCCGCCGTTATATACAAATCTGGAGAGCCTGCAAAGTTGAA  
GAGATTAGAGTGGACCCACCGAAATCATCAGAGGTAAGAATAAAGATGTTGTATGCATCTCTATGCCATACGGATATTC  
TATGCTGTAACGGACTACCCGTCCCTTTATTTCCGCGTATACCTGGCCATGAGGGAGTTGGCGTAGTGAATCAGCGGG  
TGAGGACGTGAAAGATGTCAAGGAGGGTGACATCGTGATGCCACTGTATCTTGGAGAATGTGGGGAGTGCCTTAACT  
GTTTCATCCGGGAAAACAATTTATGTCACAAGTATCCACTTGACTTTAGCGGTGTTCTACCCAGCGATGGGACCAGTCG  
TATGTCGCTCGCTAAGTCAGGCGAAAAAATTTCCATCATTTCAGTTGTTCCACATGGTCCGAGTACGTCGTGATAGAGT  
CTAGTTATGTCGTAAGAGTCGACTCTCGTCTGCCCTACCACACGCCTCCTTTCTGGCGTGCGGGTTTACCACCGGATAC  
GGTGCTGCATGGAAAGAAGCTGATATCCCCAAGGGAAGTACGGTAGCAGTGTTAGGGCTGGGCGCGGTGGGTTTAGG  
TGATAGTTGCTGGCGCGGTTACAAGGTGCCAGTCGTATAATTGGAGTTGACATCAACGACAAAAAAAAGCCAAAGC  
TGAGATATTTGGAGTACTGAGTTTCTTAATCCGAAACAGTTGGGTAAAAGTGCTTCCGAAAGCATCAAAGACGTTACT  
GGCGGCCTGGGCGTGGACTATTGCTTTGAGTGTACCGGAGTACCGCTTTATTAATGAGGCTGTAGACGCTTCAAAG  
ATTGGATTAGGACTATCGTCATGATCGGAGCCGGAATGGAAACCTCTGGTGTATCAATTACATTCTCTTCTATGTGG  
AAGAAAGCTGATCGGCTCTATCTATGGTGGAGTTCGTATAAGGTCCGACCTACCCTTGATTATCGAAAAATGCATAAAT  
AAAGAAATCCACTTAATGAGTTACAACTCACGAAGTGTCTTTGGAGGGCATAAACGATGCATTCGGTATGCTGAAGC  
AGCCAGACTGTGTGAAAATAGTAATCAAGTTTGAACAGAAGTGA

>ZzADH codon optimized

ATGAGGCTTGAGGGTAAGGTCGCCTTGGTCACGGGTGGGGCCTCCGGAATAGGTGAGTCTATAGCTCGTCTATTTATA  
GAGCACGGAGCCAAAATTTGCATCGTAGACGTTCAAGGATGAACCTGGTCAACAAGTCAGCCAAAGGCTTGGAGGAGAT  
CCTCACGCATGTTACTTTCACTGTGACGTGACTGTAGAGGACGATGTTTCGTAGGGCAGTGGATTTTACAGCAGAAAAGT  
ATGGAACCAATTGACATTATGGTGAACAACGCTGGCATAACTGGGGACAAAGTAATCGATATTCGTGATGCTGACTTTAA  
CGAGTTTAAAGAAGGTTTTCGACATCAATGTGAACGGTGTTTTTTTGGGGATGAAGCATGCTGCTCGTATTATGATCCCTA  
AGATGAAAGGGTCCATAGTATCTCTAGCCAGCGTTAGCAGTGTGATTGCAGGGGCTGGCCCCACGGTTACACGGGAG  
CCAAACACGCCGTTGTAGGGCTGACAAAGTCCGTAGCAGCCGAGTTAGGGAGGCATGGAATCAGGGTCAATTGTGTCT  
CTCCATACGCCGTGCCACTAGACTTTCCATGCCGTATTTACCTGAATCAGAAATGCAGGAGGACGCGCTAAGAGGCTT  
CTTAACATTTGTGCGTTCAAACGCCAACTTAAAGGGGGTGGACTTGATGCCAAATGATGTTGCGGAAGCCGTTTTATAT  
CTTGCGACAGAAGAAAGTAAATATGTAAGTGGCCTTAATCTTGTAAATAGACGGCGGGTTTAGTATTGCAAACACACAC  
TGCAAGTATTTGAGTAA

### 3. Synthetic fragments of the supplementary BBS

>BBS *S. citricolor*

aTGA ACTCCCTCCTCCACGCCGAGAGTTAGCGCCGAAGAAACGGA ACTGCAGCCCGCGCAGCCCCGAGGAGTTCGAG  
GCCCGGTACCCGCCACACCGCCTGGGCGGTGCGCCGCCACCTCCTCGCGCCGACGAGGACGTTCCGCACTACCGGCTC  
GCGCTGCCCACCTGATCGGCCACGCCTACCCCCGGGCCCGCGGGCCCGAACTCGACCTGCTGCTGGACATCCTCGGCT  
GGTTCACCATCCTCGACGACCGCTTCGACGGACCGGTGCGCCACCGCCCGAAGGACGCCACGCCCTGATCGACCCACT  
GCTCGGCATCCTCCGGTACCCCGGGCCCCCGGCCATCGCGCCGGAGGACCCGCTCGTGGCCGCGTGGCGGGACCTCTG  
GCACCGCCAGGCCGGCCGATGCCCGACACCTGGCGCCACCGGGCCGCCGCGAGTGCGAGGCCTGCCTGACCACCTT  
CCTCGCCGAGACCCACCACCGGGCCGGGGGAACCACTCCGGACCTCCCCGAGACGGCGCTGCTGCGCCGCCACGCCAG  
CTGCCTGTACCCGTTTATGAACATGCTGGAGCGGGTGC CGCGCACCGAGGCCCGCGCTGCTGCTCGCGGAGCCCCG  
CCTGTACCGGCTGCGCGCGTACACCGCCGACGCGGCAACCCTCATCAACGACCTCTGCTCGCTGCAACGCGAGGAAGG  
CCTGCCCCGGTCCAGTTCAACATGGTGATGACCTTCAGCGAACCCACGGGCTCAGCCGGAACCAGGCGGTCCAGGT  
GGTCCGCACCCGGGTGCGGCGCTGCGGGACGACAGCGAGGTGCTCCGCGGCCACCTGCTGCGCCGGCACCCGGCCG  
CCGGCTGGTACCTGAACGGGACCCGCGACATGGTCGACGGCCTCCACGTCTGGGCCGGCACGTGCGCCGCTACCACC  
CATGA

>BBS *M. recutita*

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCACGAAGCAACATG  
T TACTCGCAACAGTGT CATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATAAGGAGAAATTCAATGTAGC  
TACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGA ACTAATGATAAGAGCTTGTAAATGAAGCAA  
GCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCTTGGCCTAGCCTATCATTTTTGAAAAGGAGATCGA  
GGAATCCTTGCAACATATCTATGTTACATATGGCCATAAATGGACCAACTATAACAACATTGAAAGCCTTTCGCTGTGGT  
TTCGACTGCTACGACAAAATGGCTTCAACGTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCA  
GGAATCTTTATGTAATGATCCTCAAGGGATGCTTGTCTTATACGAAGCAGCATATATGAGGGTGAAGGAGAAATAAT  
ACTAGATAAGGCACTCGAGTTCACCAAACCTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTAAGAAC  
AGAAAATAAACAAGCTCTAAAGCAGCCGCTTCGTAGAAGGTTGCCAAGGCTAGAGGCGGTGCGCTACATAGCAATCTA  
CCAACAAAAGCTTCTCACAGTGAGGTCTTGTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAGAAATGCACAAA  
GACGAGCTTAGCCAAATCTGCAAATGGTGAAAGATTTGGACATTCGAAAACAAGTTACCATATGTTTCGAGACAGATTG  
ATTGAAGGCTACTTTTGGATATTGGGAATCTATTTGAGCCTCAACATTCTCGTACAAGAATGTTCTTAATGAAAACATG  
CATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTACTTATGAGGAACTCGAGATATTTACACAAGCTGTGCG  
AAAGATGGTCCATAACCTGCTTGGATGAGCTGCCAGAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTACCA  
AGAAATGGAGGAATCACTTGAGAAGGAGGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCA  
CACGCAGCCTTTTATTAGAAGCCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCCT  
AGTTACTTGTGGATATGCATTGATGACAGCAAGATCaTATGTTGCCCGGATGACGGTATAGTCACCGAGGATGCCTTT  
AAATGGGTGGCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTAAGACTTATGGATGATATTGCCACCCACAA  
GGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCAGAGGAGGAAG  
CATGCATGGATTTCTTAAAACAAGTCAAGATGGTTGAAAGGTTATAAATCAGGAGTCGCTCATGCCTACAGATGTACC  
ATTTCTCTCCTTATTCTGCAATCAACCTTGC GCGTGTGAGTGATACCTTATATAAAGACAATGATGGCTACAATCATGC  
TGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTACCCTATGATTGTCTAG

>BBS *A. kurramensis*

ATGTCACTTACCGAGGAGAAACCTATTCGTCCGATAGCTAATTTCTCCCCATCTATATGGGAAGATCAATTTTTAATATA  
CGCTAAGCAGGTGGAGCACGGTGTAGAGCAACGTGTTAAGGATTTGACTAAAGAAGTAAGGCAGTTGCTTAAAGAAG  
CCTTAGACATACCAATGAAACATGCTAACCTACTAAAACCTGATCGACGAGATCCAAAGACTGGGCATCTCTTATTTGTTT  
GAGCAAGAAATCGATCACGCATTACAACACATCTACGAGACATACGGGGATAATTGGAGCGGAGACAGGAGTTCACTA  
TGGTTCAGACTAATGAGGAAGCAAGGGTATTTCTGTGACATGCGACGTATTTAATAACCATAAAGACGAAAGTGGAGCC  
TTCAAACAGAGTCTAGCAAACGACGTAGAGGGTCTGTTGGAGCTGTACGAGGCAACTTCCATGAGGGTTGCCGGCGAA  
ATCATTTTGGATGATGCATTAGTGTTCACCAGATCAAACCTATCAATAATAGCTAAGGATACATTATCAACAAAATCCCGC  
CCTGTCCACTGAAATTCAGAGAGCCTTAAAGCAACCTCTTTGGAAAAGATTGCCCAGGATAGAAGCTGCACAATACATT

CCTTTCTACGAACAGCAAGACAGTCATAATATGGCCTTATTAAGCTGGCTAAGTTGGAGTTCAACTTGTTACAAAGTCT  
TCATAGAGAGGAACTGTCACAGCTATCAAAATGGTGAAGGCCCTTCGATGTGAAGAACAATGCTCCATACTCTAGGGA  
TAGGATTGTGGAGTGTTACTTCTGGGGCCTTGCCTCAAGGTTTGAACCGCAATTCAGCCGTGCTCGTATCTTTTTAGCCA  
AAGTCATCGCGCTTGTAACCCTTATCGACGACACCTACGATGCGTATGGGACATATGAGGAGTTAAAGATTTTCACCGA  
AGCAATTGAAAGATGGTCCATAACCTGCCTGGACATGATCCCGAATATATGAAACCCATTTACAACTTCTGATGGAC  
ACCTACACTGAAATGGAGGAGGTATTGGCCAAAGAAGGAAAAACAGACATATTGATTGTGGAAAGGAGTTCGTTAA  
GGATTTTCGTAAGGGTCTAATGGTTCGAGGCCCAATGGTTAAACGAAGGACATATACCGACCACGGAGGAGCTTGACTC  
TATCGCAGTAAACCTGGGAGGAGCCAACCTACTGACCACTACTTGTACTTACGATGTCTGATATTGTTACGAAAGAA  
GCGGTGCAATGGGCGGTGAGCGAACCCCTGCTGCGTTACAAGGGGATACTGGGCCGTAGGCTGAACGACTTAGC  
TGGACATAAAGAGGAACAGGAAAGGAAACATGTAAGTTCAAGCGTGGAGTCATACATGAAAGAGTATAATGTCAGCG  
AAGAATACGCGCAGAATTTGCTATAACAAGCAAGTTGAAGACTTATGGAAAGACATCAATCGTGAGTACCTGATAACCA  
AAACGATCCCTCGTCCACTTCTTGTGGCTGTAATCAACCTAGTCCATTTTCTAGAAGTATTATACGCGGAGAAGGACAAC  
TTTACTCGTATGGGCGATGAATACAAAGACTTAGTAAAATCCCTATTGGTTTACCCTATGTCAATTTGA

#### 4. [Coding sequence of UGT93B16](#)

>UGT93B16

ATGGGGATTGAGTCGATGGACAGTAGTGTGGCGTTGGTGGCGGTGCCGTTCCAGCGCAGGGCCACCTGAACCAGCT  
GATGCACTTGTCCCTGCTGGTGGCGTCGCGGGCCTCTCCGTGCACTACGCAGCGCCGGCGGCACGTCCGGCAGGC  
CAAGTCGCGCGTGCACGGCTGGGACGCCAAGGCCCTGGCCTCCATCCACTTCCACGACCTGGACGTCCCCACCTTCAAG  
TCGCCCAGCCCGACCCGCGCCGCGCATCGCCCTTCCCAGCCACCTGCTGCCCATGTGGGAGACCTACAGCGCCGCG  
CCCGCGTCCCTCTGGCGTCCCTCCTCGAGCGCCTCTCGGCCACCCACCGCCGCGTGGTGGTGTACGACCACATGAA  
CTCCTTCGCGGCGGCGGAGGCGGCGGGTTCGACGCGGAGGCGTACGGCCTGGTGTGCGTGGCCATCTCCAACCACC  
TGGCGTGGATGCCAGACGGGCACCAGCTCCTCCGCGACCGCGGCCCTCCGCTCCGTGCCCATGGATGCGTGCATGTCCA  
AGGAGTTTGTGGAGTACATGGCCCGGTCACCACGGAGGCGGAGGGCGCCGGTTCTCATGAACACCTGCCGCGCG  
CTCGAGGGCGAGTTCCTCGACGCCGTGCGCGAGATCCCCGACATCAAGCGCCAGAAGCGTTTCGCGGTGGGGCCGCTG  
AACCCGCTGCTTCCCCTCGCCACCGAGCCCGACGTGGTCAACCCGCCACGGCGCGGCACGACTGCATGCGCTGGCTC  
GACGCGCAGCCTCCGGCGTGGTGTCTACGTCTCATTGCGCACCACTCGTCTTTCCTGCCCGAGCAGATCGCGGAGC  
TCGCCCGCGATCAAGGGCAGCAGGCAGAGTTTATCTGGGTGCTGCGCGACGCCGACCGCGCCGATATTTTCGCCG  
GCAACTCCGGCGGCGACAGCCGCCGCGCTACGAGAAGCTCCTGTGCGGAGGCAAGCCCAGGGCACCGGGCTGGTG  
ATCACCGGTGGGCGCCGAGCTGGAGATCTGGCGCACGGCGCCACGGCGGCCCTTCATGAGCCACTGCGGATGGAA  
CTCCACCATGGAGAGCCTCAGCCACGGCAAGCCATCCTCGCCTGGCCCATGCACAGCGACCAGCCATGGGACGCCGA  
GCTCCTCTGCAACTACCTCCGGGTTGCCATGCTCGTCAGGCCGTGGGAGAAGCACGGCGAGGTCTGGCCGAGGGG  
CCATCCAGGAGGTGATCGAGAAGGCAATGCTCTCGGAGAAAGGACGGCGCTCCGACAGAGGGCGAAGCTGCTGGG  
TGAGGCTGTTTCGTGCAGCCGTGGCGGACAATGGCTCCTCCACCAAGGACCTCCATGACTTCTGTTGCTCACATACAAGG  
ATTTGA

#### 5. [Synthetic fragments corresponding to TEF1 promoter and coding sequence of \*Rattus norvegicus\* UGDH](#)

>fragment 1 TEF1/ UGDH

ATAGCTTCAAAATGTTTCTACTCCTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACAC  
CCAAGCACAGCATACTAAATTTCCCTCTTCTTCTAGGGTGTGTTAATTACCCGTAATAAAGTTTGGAAAAGAA  
AAAAGAGACCGCCTCGTTTCTTTTCTTCTCGTGAAGGCAATAAAAATTTTATCACGTTTCTTTTCTTGAATAATTTT  
TTTTTGAATTTTTTCTTTTCGATGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTC

ATTTTCTTGTTCTATTACAACCTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAATCTAATCTAAGTTTTAATTAC  
AAAATGGTTGAGATCAAGAAGATCTGTTGCATTGGTGCGGGCTACGTCGGCGGACCCACATGCAGTGCATTGCTCGC  
ATGTGCCCTGAAATCAGGGTAACGGTTGTGGATGTCAATGAGGCCAGGATCAATGCATGGAATTCTCCAACGCTTCTTA  
TTTATGAGCCTGGACTAAAAGAAGTAGTCGAATCCTGTCGAGGGAAAAACCTCTTTTTTCTACCAATATTGATGATGCC  
ATCAGAGAAGCCGATCTAGTGTATTTCTGTGAACACACCAACAAAAACATATGGAATGGGAAAAGGCCGGGCGGCA  
GATCTGAAGTATATCGAAGCTTGTGCTCGCCGATTGTGCAGAACTCAATGGGTACAAAATTGTGACTGAGAAAAGC  
ACAGTCCCTGTGCGGGCAGCGGAAAGCATCCGCCGATATTTGATGCCAACACAAAGCCCAACTTGAATCTACAGGTT  
TGTCCAATCCTGAGTTCTTGGCAGAGGGAACAGCCATCAAGGACCTAAAGAACCCAGACAGAGTCTGATTGGAGGGG  
ATGAGACCCAGAGGGCCAGAGAGCTGTTCAGGCACTCTGT

>fragment 2 TEF1/ UGDH

ugdhGCTGTGTACGAGCACTGGGTTCCCAAGGAAAAGATCCTCACCACCAACACTTGGTCTCAGAGCTTCCAACTG  
GCAGCCAATGCTTTTCTTGCCAGAGGATCAGCAGCATTAACCCATAAGTGCTCTGTGTGAAAGCACAGGCGCCGATG  
TGAAGAGGTGGCAACGGCTATCGGGATGGACCAAAGAATTGGAAATAAGTTTCTAAAAGCCAGCGTTGGTTTTGGT  
GGGGGCTGCTTCCAAAAGATGTTCTGAATTTGGTTTATCTCTGTGAGGCTCTGAATCTGCCGAAGTAGCTCGTTACT  
GGCAGCAGGTCATAGACATGAATGACTACCAGAGGAGGAGTTTGCATCACGGATCATAGACAGCCTGTTAATACAG  
TGACTGATAAGAAGATAGCTATCTTGGGGTTTGCCTTCAAAAAGGATACTGGTGATAACCAGGGAGTCTCCAGTATCTA  
CATTAGCAAATACCTGATGGACGAGGGTGCACACTCCACATCTACGACCCCAAAGTACCCAGGGAGCAGATAGTGGT  
GGATCTTTCTCATCCAGGCGTCTCAGCGGATGACCAAGTGCCAGACTGGTGACCATTTCCAAGGATCCATATGAAGCA  
TGTGATGGCGCCCATGCCCTCGTTATCTGCACAGAGTGGGACATGTTAAAGGAACTGGATTATGAACGGATTCATAAAA  
GAATGCTGAAGCCAGCCTTCATATTTGATGGCCGGCGTGTCTGGATGGGCTCCACAATGAGCTACAGACCATTGGCTT  
CCAGATTGAAACAATTGGCAAAAAGGTATCTTCCAAGAGAATCCATACACTCTGGTGAATTTCAAAGTTTAGTCTTC  
AGGATCCACCTAACAAAGAAGCCCAAAGTCTAG

## 6. [Synthetic fragments of the human UGTs active with \(±\)-hernandulcin and \(-\)-α-bisabolol](#)

>UGT1A9

ATGGCTTGCACAGGGTGGACCAGCCCCCTTCTCTATGTGTGTGTCTGCTGCTGACCTGTGGCTTTGCCGAGGCAGGGA  
AGCTACTGGTAGTGCCCATGGATGGGAGCCACTGGTTCACCATGAGGTCGGTGGTGGAGAAACTCATTCTCAGGGGGC  
ATGAGGTGGTTGTAGTCATGCCAGAGGTGAGTTGGCAACTGGGAAGATCACTGAATTGCACAGTGAAGACTTATTC  
CTTCATATACCCTGGAGGATCTGGACCGGGAGTTCAAGGCTTTTCCCATGCTCAATGGAAAGCACAAGTACGAAGTAT  
ATATTCTCTATTAATGGGTTTACATAATGACATTTTTGACTTATTTTTTCAAATTGCAGGAGTTTGTAAAGACAAAA  
ATTAGTAGAATACTTAAAGGAGAGTTCTTTTGTGTCAGTGTCTCGATCCTTTTGTAACTGTGGCTTAATTGTTGCCA  
AATATTTCTCCCTCCCTCCGTGGTCTTCGCCAGGGGAATACTTTGCCACTATCTTGAAGAAGGTGCACAGTGCCTGCT  
CCTCTTCTATGTCCCAGAAATTCTTAGGGTTCTCAGATGCCATGACTTTCAAGGAGAGAGTACGGAACCACATCAT  
GCACTTGGAGGAACATTTATTATGCCACCGTTTTTCAAATGCCCTAGAAATAGCCTCTGAAATTCTCCAAACACCTG  
TTACGGAGTATGATCTCTACAGCCACACATCAATTTGGTTGTTGCGAACGGACTTTGTTTTGGACTATCCCAAACCCGTG  
ATGCCAACATGATCTTCATTGGTGGTATCAACTGCCATCAGGGAAAGCCGTTGCCTATGGAATTTGAAGCCTACATTA  
ATGCTTCTGGAGAACATGGAATTGTGGTTTTCTTTGGGATCAATGGTCTCAGAAATTCAGAGAAGAAAGCTATGGC  
AATTGCTGATGCTTTGGGCAAAATCCCTCAGACAGTCTGTGGCGGTACACTGGAACCCGACCATCGAATCTTGCGAAC  
AACACGATACTTGTTAAGTGGCTACCCCAAACGATCTGCTTGGTCACCCGATGACCCGTGCCTTTATCACCCATGCTGG  
TTCCCATGGTGTATGAAAGCATATGCAATGGCGTTCCCATGGTGTGATGATGCCCTTGTGGTGTGATCAGATGGACAAT  
GCAAAGCGCATGGAGACTAAGGGAGCTGGAGTGACCCTGAATGTTCTGGAAATGACTTCTGAAGATTTAGAAAATGCT  
CTAAAAGCAGTCATCAATGACAAAAGTTACAAGGAGAACATCATGCGCTCTCCAGCCTTACAAGGACCGCCCGGTG  
GAGCCGCTGGACCTGGCCGTGTTCTGGGTGGAGTTTGTGATGAGGCACAAGGGCGCGCCACACCTGCGCCCCGAGC  
CCACGACCTCACCTGGTACCAGTACCATTCTTGGACGTGATTGGTTTTCTTGGCCGTCGTGCTGACAGTGGCCTTCA

TCACCTTTAAATGTTGTGCTTATGGCTACCGGAAATGCTTGGGGAAAAAAGGGCGAGTTAAGAAAGCCACAAATCCA  
AGACCCATTGA

>UGT2B4

ATGAGCATGAAGTGGACCTCTGCGTTGCTTCTTATCCAATTGTCTTGCTATTTCTCCAGCGGTTCTTGCGGGAAAGTCTT  
GGTCTGGCCAACTGAGTTTTCTCATTGGATGAACATTAACCATACTCGACGAGCTGGTTCAGAGGGGACACGAAGT  
AACAGTGCTGGCTTCTCTGCCAGCATCTCCTTCGACCCTAATTCTCCTAGCACCCCTCAAGTTCGAGGTCTACCCGGTGA  
GTCTGACCAAGACCGAATTCGAAGACATCATAAAACAACCTCGTGAAACGGTGGGCCGAGCTGCCAAGGATACCTTCT  
GGTCTTACTTCTCTCAGGTCCAGGAGATTATGTGGACTTTAACGATATTCTGCGCAAATTCTGCAAAGACATCGTCAGC  
AACAGAAGTTGATGAAGAAGCTGCAAGAAAGCAGGTTGACGTGGTCTGGCTGACGCAGTCTTTCTTTTCGGGGAA  
TTGCTCGCGGAACTGCTGAAGATTCCATTCTTTATAGTCTGCGTTTTCCCCGGCTATGCTATCGAGAAACACAGCGG  
CGGTCTGTTGTTTCCCCGTCATACGTTCCCGTCGTCATGAGTGAGCTGTCCGATCAGATGACATTTATCGAACGTGTTA  
AGAACATGATACGTACTGTACTTCGAGTTCTGGTTCCAAATCTTCGATATGAAGAAATGGGACCAGTTTTATTGAGA  
GGTCTTGGGGCGGCCAACAACTCTCAGCGAAACCATGGCTAAGGCCGATATCTGGCTGATACGGAATTATTGGGACTT  
CCAGTTCCACATCTCTGCTCCCAACGTGGAATTTGTCGGTGGGTTGCATTGTAAGCCAGCAAAGCCACTCCCTAAAG  
AGATGGAAGAATTCGTCCAATCATCCGGGAGAACGGCGTGGTCTTCAGTTGGGAAGTATGGTATCAAATACAA  
GCGAGGAGCGGGCTAACGTCATCGCCAGCGCTCTGGCTAAGATACCCAGAAAGTGCTTTGGCGCTTCGACGGAAACA  
AGCCGGACACCCTGGGCTTGAATACCCGCTTGATAAATGGATCCCGCAAACGACCTCTTGGGGCATCCTAAGACGCG  
AGCATTATTACCCACGGCGGCGCTAACGGGATTTACGAAGCCATTTATCACGGCATTCCAATGGTTGGAGTGCCCTG  
TTCGCCGACCAGCCAGACAATATAGCCCATATGAAAGCAAAGGTGCTGCAGTCAGCCTGGATTTTCATACGATGAGTT  
CCACCGATCTGCTGAACGTTTGAAGACCGTGATAAACGACCCCTTTACAAGGAAAACGCAATGAAGCTCTCTAGGAT  
CCACCACGACCAGCCAGTCAAACCACTGGACCCGGCCGTGTTTTGGATCGAGTTCTGTGATGCGGCACAAGGGCGCGAA  
ACATCTGCGTGTGGCCGCACATGATCTTACTTGGTTCCAATATCATAGCCTGGACGTCACAGGCTTTCTTCTTGTGCTGCG  
TAGCCACCGTAATCTTTATTATTACCAAGTGCTCTTCTGCGTATGGAAATTCGTCCGACCAGGGAAGAAAGGTAAGCG  
GGACTGA

>UGT2B7

ATGTCCGTTAAGTGGACCAGCGTGATCCTTCTCATCCAACCTGAGTTTCTGTTTCTTCCGGCAACTGCGGGAAAGTCTT  
GGTTTGGGCCGCGGAGTATCCCACTGGATGAACATTAACCATTCTTGACGAACTGATCCAAAGGGGCCACGAAGTC  
ACGGTCTCGCTCATCCGCCAGTATCCTGTTTCGACCCTAATAATAGCTCTGCCTTGAAGATCGAGATATACCTACCTC  
ACTGACAAAGACCGAACTGGAAAACCTTTATTATGCAGCAAATCAAAGGTGGTCTGATCTCCCAAGGACACTTTCTGG  
CTTTACTTCAGCCAGGTCCAAGAGATTATGTCTATCTTCGGGGATATTACAAGGAAATTCTGCAAGGACGTCGTGTCCA  
ACAAGAAGTTCATGAAGAAGGTGCAGGAAAGTCGGTTCGATGTAATCTTCGCCGACGCAATCTTCCCTTGCAGCGAACT  
TCTTGCCGAACTGTTTAATATTCCATTCTGATATTTCCCTGTCATTTAGCCCCGTTATACCTTCGAGAAACACTCCGGTGG  
CTTCATCTTTCCCCAAGCTATGTCCAGTGGTGTGTCGAGCTTACGGACCAGATGACATTTATGGAACGGGTGAAG  
AACATGATTTACGTCTTGTATTTGATTTCTGGTTCGAAATCTTCGATATGAAGAAATGGGACCAATTCTACAGCGAGGT  
GTTGGGGCGGCCTACAACCTTTGTCAGAAACCATGGGTAAGGCCGATGTTTGGCTGATACGCAATAGTTGGAACCTCCA  
ATTCCTTACCCCTGCTGCCAACGTCGACTTCGTCGGCGGCCTGCATTGTAAGCCAGCAAAGCCACTTCCCAAAGAG  
ATGGAAGATTTCTGTCAGTCTTCAGGTGAGAACGGCGTCGTAGTCTTCTCATTGGGCAGTATGGTTTCTAATATGACTG  
AGGAGAGAGCGAATGTTATCGCTCTGCACTTGTCAAATACCCAGAAAGTATTGTGGCGGTTTCGACGGCAACAAGC  
CGGACACACTGGGGCTTAACACCAGGCTGTATAAATGGATCCCTCAAACGATCTGCTGGGACACCCCAAACACGCG  
CCTTCATCACCCACGGCGGCGCAAACGGTATATATGAAGCCATTTATCACGGAATACCAATGGTCGGCATCCCTCTGTT  
GCAGACCAGCCGACAATATAGCCCATATGAAAGCACGCGGTGCGGCAGTCCGGGTCGACTTTAATACGATGAGCAGC  
ACCGATCTGCTTAACGCGCTGAAACGGGTGATCAACGACCCAAGCTACAAGGAAAACGTCATGAAGCTGTCTCGCATC  
CAGCACGACCAGCCCGTCAAACCAATTGGACCGTGCCGTGTTTTGGATAGAGTTCGTGATGAGACATAAGGGTGCCAAG  
CATCTCCGCGTCGCTGCGCATGATCTGACATGGTTCCAATATCATAGCCTGGACGTTATCGGATTTCTCCTCGTGTGCGT  
CGCTACAGTCATTTTCATAGTGACCAAGTGCTGCTTGTCTGCTTTTGGAAATTCGCCGTAAGGCTAAGAAAGGGGAAAG  
AACGACTGA

## 7. Synthetic fragments of UDP-GlcA transporters

>FRC

```
ATGTCTATGAGTAGGGGAGGCAACACCACACTAGATCTTCAACCCCTTCTAGCTGAGTCTGATGTAGGCAATAGAGAAC
TGGAGGAAAAAATGGGAGGAAGTGCTGACAGGAGCTCATTACTGGATGGTAGTGGATCCAAGGAATTATCACACCGT
GAGAGAGAGGACTCAGCGCTTTTTGTCAAAAAAATAGGGTCTGCGTTGTTCTACGGTCTAAGCTCCTTTATGATAACTG
TGGTAAATAAGACAGTGTTAACCAGCTACCATTTTCCTAGTTTTCTGTTCTGAGTCTAGGCCAACTAACGGCATCCATT
GTCGTGCTGGGTATGGGTAAACGTTTGAAACTTGTAATTTTCTCCTTTGCAGAGGAACACATTTGCGAAGATTTTTCC
GTTACCGTTGATATTTCTGGGGAACATGATGTTTCGGCTTGGGTGGCACAAAGACGTTATCATTACCAATGTTTCGCGGCG
CTGCGTAGATTTTCTATTTTGATGACGATGCTGCTTGAACCTAAGATACTGGGATTAAGACCTTCCAACGCGGTTTCAGGT
GTCTGTATATGCGATGATAGGAGGTGCGCTTCTAGCGGCCTCCGACGATCTATCATTAAACATGAGGGGCTATATATAT
GTAATGATAACTAATGCTTTAACTGCGTCCAACGGTGTCTATGTGAAGAAGAAGCTAGACACCAGTGAAATAGGTAAAT
ACGTTTAAATGTATTACAATTCAGTGTTCATGTTCTTGCCTGCCCTAGCCTTAAATTATGTTACTGGGAATCTTGACCAGG
CGCTTAACTTTGGGCAGTGGAATGATAGCTTGTTCGTTGTGCAGTTTTTGTGAGCTGCGTTATGGGTTTTATTTTAAAGC
TATAGTACAATTCTGTGTACGCAGTTCAACTCCGCGTTGACTACTACCATAGTAGGGTGCCTAAAAAACATCTGTGTGAC
TTACCTGGGCATGTTTATAGGTGGCGACTACGTGTTTTCTGGTTAAATTGTATAGGGATCAATATCTCAGTCTAGCTT
CATTATTATACACATACGTTACGTTTAGGAGAAAAAGAGCACCGGACAAGCAGGACCATCTACCCTCTACGAGGGGAG
AAAATGTGTAG
```

>UGTrel7

```
ATGGCAGAAGTGACAGAAGGCAACATGCTAGAGTAAAGGGGGAGGCACCAGCAAAGAGCAGTACGTTGCGTGATG
AGGAGGAGCTAGGTATGGCTTCTGCCGAGACATTAACAGTCTTTCTAAAGCTGCTGGCGGCTGGTTTCTATGGAGTCA
GTAGTTTCCTTATTGTTGTCGTAATAAAAGCGTATTAACCAATTATAGATTTCTTCTAGTTTGTGTGTCGGGTTGGGAC
AAATGGTAGCGACCGTTGCCGTGCTATGGGTTGGCAAAGCGCTTAGAGTGGTAAAATCCCTGACCTTGACAGGAACG
TCCCCCGTAAAACCTTCCCACTGCCACTACTGTACTTTGAAACCAGATAACTGGTTTATTTTCTACTAAAAAACTTAACC
TACCAATGTTTACGGTCCTGCGTAGATTTTCTATATTATTCACTATGTTTGCTGAAGGAGTACTGCTAAAGAAGACCTTTT
CATGGGGCATCAAGATGACGGTGTTCGCCATGATTATTGGAGCCTTCGTAGCAGCCAGCAGCGATTTAGCGTTCGATTT
GGAAGGATACGCCTTCACTTCTGATAAACGATGTATTGACCCGAGCCAATGGTGCTTATGTTAAACAAAACTGGACTCT
AAGGAGTTGGGTAAGTACGGGCTGCTTATTATAATGCACTGTTTATGATTTTGCCAACACTGGCAATTGCGTATTTTAC
CGGGGATGCTCAAAGGCGGTAGAATTTGAAGGATGGGCCGATACGCTTTTTCTTCTCAGTTTACGTTAAGCTGTGTG
ATGGGTTTTATATTAATGTATGCTACGGTGCTATGTACGCAATACAACAGTGCCTTACAACAACGATTGTGGGCTGTAT
AAAGAATATCCTTATTACCTACATCGGAATGGTTTTTGGCGGCGACTATATCTTTACTTGGACCAACTTCATTGGGTTAA
ATATATCAATTGCCGGTCTCTAGTTTACAGCTATATAACCTTTACAGAGGAGCAGTTATCTAAGCAGTCCGAAGCCAAT
AATAAGCTGGACATCAAGGGGAAGGGTGCAGTTTGA
```



## D. List of figures

Figure 1: The biosynthesis of hernandulcin, the sweetening molecule isolated from <i>Lippia dulcis</i> with the putative presence of a CYP and/or a CYP and an ADH. ....	18
Figure 2: Overview of biosynthetic pathways leading to the emission of plant volatile organic compounds from the plant secondary metabolism (reproduced from Dudareva et al <sup>31</sup> ).....	22
Figure 3: Examples of terpenoid products with current or potential industrial applications. ....	24
Figure 4: Plant biosynthesis of terpenes (This figure was reproduced from Moses et al <sup>47</sup> ). ....	26
Figure 5: Phylogeny of putative full-length TPSs from seven sequenced plant genomes and representative characterized TPSs from gymnosperms. This figure was originally published by Chen et al <sup>80</sup> .....	29
Figure 6: Protein domain structure of bi-functional class I/II TPSs and class I and class II TPSs. ....	30
Figure 7: CYP monooxygenation mechanism. Figure originally presented by Werck-Reichhart et al <sup>97</sup> . ....	32
Figure 8: The design / build / test / learn cycle applied in synthetic biology (figure originally presented by Moses et al <sup>208</sup> ) .....	46
Figure 9: The MVA pathway of <i>S. cerevisiae</i> (as described by Grabińska et al <sup>220</sup> ) .....	48
Figure 10: The genetic engineering of EPY300 <i>S. cerevisiae</i> strain suitable for de novo production of sesquiterpenoids, Figure adapted from Nguyen et al <sup>25</sup> . ....	50
Figure 11: Metabolic network design for terpene production in yeast. This figure was originally presented by Vickers et al <sup>16</sup> . ....	53
Figure 12: Increasing complexity in new to nature products (figure originally presented by Erb et al <sup>235</sup> ).....	56
Figure 13: Strategy to produce derivatives of (+)-epi- $\alpha$ -bisabolol (including hernandulcin) using the yeast optimized MEV pathway, (+)-epi- $\alpha$ -bisabolol synthase and an in vivo screen to unravel oxidative enzymes. ....	63
Figure 14: Detection of new molecules generated in YBIS_01 strain transformed by the empty plasmid pYEDP60, or by some CYPs. ....	73
Figure 15: The production of sesquiterpenes by YBIS_01 strain cultured in different conditions. All cultures were performed in triplicate. ....	76
Figure 16: Comparison of the sesquiterpene production by YBIS_01, YBIS_05 (DDP1 gene deletion) and YBIS_06 strains (DDP1 gene deletion and one supplementary copy of the BBS encoding gene). ....	76
Figure 17: Effect of chassis improvement on accumulation of CYP2C9 and CYP2D6 products. The level of (+)-epi- $\alpha$ -bisabolol was varied in the strains by comparing YBIS_05 (1 integrated BBS copy) to YBIS_06 (2 integrated BBS copies). .....	78
Figure 18: Chemical structures of the newly generated molecules. CYP2D6 produces 14-hydroxy-(+)-epi- $\alpha$ -bisabolol while CYP2C9 produces 80% of 14-hydroxy-(+)-epi- $\alpha$ -bisabolol and 20% of a second additional hydroxylated (+)-epi- $\alpha$ -bisabolol derivative, most probably 9-hydroxy-(+)-epi- $\alpha$ -bisabolol. The red dots on the (+)-epi- $\alpha$ -bisabolol molecule indicates the other positions that could be hydroxylated in CYP2B6 and CYP2B11 although the original molecules could not be fully characterized by NMR.....	80

Figure 19: In vivo production of (+)-epi- $\alpha$ -bisabolol in BY4741 in SD-Leu and EPY300 in SD-His-Met-Leu after expression of the BBS on a high copy plasmid. GC-FID chromatograms of (a) (-)- $\alpha$ -bisabolol standard along with its MS spectrum; (b) BY4741 + pESC leu2d BBS; (c) EPY300 + pESC leu2d BBS along with the MS spectrum of (+)-epi- $\alpha$ -bisabolol.....	88
Figure 20: Bioconversion of diclofenac to 4-hydroxy-diclofenac by strains YBIS_01 and YBIS_02.....	89
Figure 21: An optimized Kozak environment for the expression of CYP2D6 enzyme enhances the (+)-epi- $\alpha$ -bisabolol oxidized product accumulation. ....	90
Figure 22: Chromatograms of GC-FID signal showing the sesquiterpene content in the chassis strain YBIS_01 cultured in SD medium with a dodecane overlay. (a) (-)- $\alpha$ -bisabolol, (b) nerolidol, (c) farnesol, (d) YBIS_01. ....	91
Figure 23: MS chromatograms at $m/z = 203.3$ of the enzymatic reactions with (-)- $\alpha$ -bisabolol and demonstrating the in vitro activities of the CYPs.....	92
Figure 24: MS chromatograms at $m/z = 203.3$ of the enzymatic reactions with farnesol and demonstrating the in vitro activities of the CYPs. ....	93
Figure 25: MS chromatograms at $m/z = 203.3$ of the enzymatic reactions with nerolidol and demonstrating the in vitro activities of the CYPs. ....	94
Figure 26: UV chromatograms at 205 nm validating in vitro activities of the CYPs towards (-)- $\alpha$ -bisabolol, farnesol and nerolidol. ....	95
Figure 27: Total ion current chromatograms of GC-MS showing identical product formation between the CYP2D6 and CYP2C18 enzymes. ....	96
Figure 28: Metabolite partition in the cell pellet and in the culture broth of strain YBIS_06 + CYP2C9. 5 ml of cells and culture broth were separated by centrifugation and extracted by hexane. ....	97
Figure 29: NMR analysis of 14-hydroxy-(+)-epi- $\alpha$ -bisabolol produced by CYP2D6. ....	98
Figure 30: The predicted metabolism of human CYPs with (a) FAME3 algorithm, (b) Xenosite algorithm. ....	99
Figure 31: Optimization of the separation of the oxidized sesquiterpene products from (A) YBIS_06 + CYP2C9 (B) YBIS_06 + CYP2B6 (C) YBIS_06 + CYP2B11.....	102
Figure 32: UV chromatogram (205 nm) of the chemical reaction aiming at reducing ( $\pm$ )-hernandulcin.....	104
Figure 33: Time course of (+)-epi- $\alpha$ -bisabolol production in the YBIS_08 strain according to GC-FID quantitation. ....	105
Figure 34: Sesquiterpene distribution in YBIS_01, YBIS_06 strains (with respectively the second BBS copy either under GAL10 promoter or PDC1).....	106
Figure 35: Enzymatic assays showing ( $\pm$ )-hernandulcin oxidizing activities with (A) CYP2B6 (B) CYP2B11 (B) CYP2C9 (B) CYP2C18, the results present the UV chromatograms at 238nm. ....	107
Figure 36: CYPs oxidation activities in YBIS_08 strains. The overlaid chromatograms depict MS signal at $m/z = 203.3$ * .....	110
Figure 37: Chromatograms at 238 nm showing the new products observed in YBIS_01 + CYP2B11 + EIADH.....	112
Figure 38: Chromatograms at $m/z = 219.3$ showing the new products observed in YBIS_01 + CYP2B11 + AaADH, YBIS_01 + CYP2B11 + LiADH, YBIS_01 + CYP2B11 + VoADH .....	113
Figure 39: The bisabolol isomers and the corresponding BBS that were biochemically characterized <sup>82,302,303</sup> .....	113
Figure 40: $\alpha$ -bisabolol production with the different BBS integrated in EPY300 background .....	114

Figure 41: Comparison of parental and chimeric CYP2B6-CYP2B11 primary structures. Solid bars represent human CYP2B6 sequence elements, gray-filled bars represent canine CYP2B11 sequence elements. This figure was reproduced from Lautier et al <sup>260</sup> .	120
Figure 42: CYP2B6 and CYP2B11 in vitro activities towards (A) (-)- $\alpha$ -bisabolol (B) trans,trans-farnesol. The chromatograms depict the MS signal at m/z = 203.3.	122
Figure 43: The behavior of the chimeras using (-)- $\alpha$ -bisabolol as substrate	123
Figure 44: The behavior of the chimeras using trans,trans-farnesol as substrate	124
Figure 45: Examples of glycosylated plant-specialized metabolites. Compounds that are active (A) in their fully glycosylated forms or (B) upon hydrolysis are shown. Figure reproduced from Louveau et al. <sup>26</sup>	129
Figure 46: The current UGT superfamily nomenclature system. The diagram illustrates the system used to name plant UDP glycosyltransferases (adapted from Ross et al <sup>323</sup> )	131
Figure 47: Ribbon Diagram of the Structure of UGT71G1 with Bound UDP (reproduced from Shao et al <sup>37</sup> )	132
Figure 48: Proposed catalytic mechanism for the conversion of delphinidin to delphinidin 3-O-glucoside. This figure was presented by Hiromoto et al <sup>328</sup> .	133
Figure 49: Protein primary structures of (A) UGT1A9 and (B) UGT2B7, figure reproduced from Guillemette et al <sup>366</sup> .	138
Figure 50: NuPAGE Tris-acetate 7% protein gel on whole cells to assess UGT93B16 expression at different temperatures	144
Figure 51: NuPAGE Tris-acetate 7% protein gel assessing the UGT93B16 soluble expression.	145
Figure 52: MS signal at m/z = 219.3 showing the time course of ( $\pm$ )-hernandulcin-Glc formation by UGT93B16	146
Figure 53: MS spectra of (A) ( $\pm$ )-hernandulcin and (B) ( $\pm$ )-hernandulcin-Glc after UGT93B16 incubation.	146
Figure 54: Time course conversion of (-)- $\alpha$ -bisabolol by UGT93B16	148
Figure 55: Time course conversion of trans,trans-farnesol by UGT93B16	149
Figure 56: MS signal at m/z = 203.3 showing the time course of hydroxy-(+)-epi- $\alpha$ -bisabolol-Glc formation by UGT93B16 from YBIS_01 + CYP extractions.	150
Figure 57: MS spectra of (A) 14-hydroxy-(+)-epi- $\alpha$ -bisabolol from hexane extraction of YBIS_01+CYP2D6 strain and (B) 14-hydroxy-(+)-epi- $\alpha$ -bisabolol-Glc spectra after UGT93B16 incubation.	151
Figure 58: Time course of the (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin glucosylation by FlagUGT93B16 showed as percentage conversion of the added substrate.	155
Figure 59: MS signal at m/z = 205.3 showing YBIS_02 In vivo sesquiterpenoid glucosylation mediated by UGT93B16 ..	156
Figure 60: Protein sequence alignment of UGT93B16 and the other UGT93 currently characterized in literature	158
Figure 61: Human UDP-glucuronosyltransferase activities with (-)- $\alpha$ -bisabolol	160
Figure 62: Human UDP-glucuronosyltransferase activities with ( $\pm$ )-hernandulcin	161
Figure 63: The approaches for in vivo production of glucuronides in <i>S. cerevisiae</i>	163
Figure 64: The glucuronidation system in the endoplasmic reticulum in human cells and a proposal for an improved system for glucuronide production in yeast based on a humanized efflux. Figure adapted from Liu et al <sup>382</sup>	165
Figure 65: Summary of the generated products from glycosylation diversification	167
Figure 66: The human sweet taste receptor heterodimer T1R2-T1R3 topology (A) and the binding loci of various sweeteners (B), this figure is reproduced from Dubois et al <sup>401</sup>	174

Figure 67: Fractions collected from (A) YBIS_06 + CYP2B6, (B) YBIS_06 + CYP2B11. ....	202
Figure 68: MS chromatograms at $m/z = 235.3$ showing oxidation activities of the different CYPs with ( $\pm$ )-hernandulcin after 60-minute incubation. ....	203
Figure 69 : SDS-PAGE (15% polyacrylamide gel) analysis of the whole cell extract showing recombinant UGT93B16 production in <i>E. coli</i> after 4h, 7h30 and 24h of culture.....	203
Figure 70: Time course of (-)- $\alpha$ -bisabolol-Glc formation by UGT93B16 detected by MS signal at $m/z = 205.3$ .....	204
Figure 71: MS signal at $m/z = 205.3$ showing the time course of trans,trans-farnesol-Glc formation by UGT93B16 .....	204
Figure 72: Time course conversion of nerolidol by UGT93B16.....	205
Figure 73: MS signal at $m/z = 435.3$ ( $M+Na^+$ ) showing human UDP-glucuronosyltransferase activities with ( $\pm$ )-hernandulcin .....	206
Figure 74: Sweetness potency measurement carried out by the team located at the “Centre des Sciences du Goût et de l’Alimentation” .....	206
Figure 75: Sweetness potency evaluation of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin measured by the team in “Centre des Sciences du Goût et de l’Alimentation” .....	207

## E. List of tables

Table 1: Plasmids used in this study.....	65
Table 2: Strains used in this study.....	66
Table 3: In vitro activities of the CYPs towards (-)- $\alpha$ -bisabolol, farnesol and nerolidol .....	74
Table 4: CYPs library cloned in the multi cloning site of pYEDP60 and screened for the oxidation of (+)- <i>epi</i> - $\alpha$ -bisabolol..	84
Table 5: Primers used in this study, restriction sites are underlined.....	86
Table 6: Summary of NMR data of (+)- <i>epi</i> - $\alpha$ -bisabolol <sup>6,284,285</sup> 14-hydroxy-(+)- <i>epi</i> - $\alpha$ -bisabolol produced by CYP2D6, and presumed 9-hydroxy-(+)- <i>epi</i> - $\alpha$ -bisabolol produced by CYP2C9. <sup>286</sup> .....	87
Table 7 Plant CYPs investigated for (+)- <i>epi</i> - $\alpha$ -bisabolol activities.....	108
Table 8 : Selected ADHs for the oxidation of (+)- <i>epi</i> - $\alpha$ -bisabolol .....	111
Table 9 Preliminary results of the ADH yeast in vivo assay.....	112
Table 10: The AsGT05287 known acceptors from Thomas Louveau's thesis .....	136
Table 11 Plasmids used in this chapter .....	140
Table 12 : Strains used in this chapter .....	141
Table 13: Sesquiterpene conversion with cell lysate containing UGT93B16. ....	152
Table 14: Conversion of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin bioconversion in <i>E. coli</i> . ....	153
Table 15: Bioconversion yield of (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin in <i>S. cerevisiae</i> .....	154
Table 16: (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin conversion obtained in various conditions of glucosylation.....	155
Table 17: Human UDP-glucuronosyltransferase activities detected with (-)- $\alpha$ -bisabolol and ( $\pm$ )-hernandulcin.....	162
Table 18 : Primers used in this thesis. Restriction sites are underlined, homology sequences for Gibson cloning are displayed in bold. ....	208
Table 19: Additional strains generated for the (+)- <i>epi</i> - $\alpha$ -bisabolol oxidation.....	210

## F. Abbreviations

**BBS:** (+)-*epi*- $\alpha$ -bisabolol synthase

**CPR:** NADPH cytochrome P450 reductase

**CYP:** cytochrome P450

**DNA:** deoxyribonucleic acid

**FID:** flame ionization detector

**FPP:** farnesyl pyrophosphate

**GC:** gas chromatography

**HMGR:** 3-hydroxy-methylglutaryl coenzyme A reductase

**HPLC:** high-performance liquid chromatography

**MS:** mass spectrometry

**MW:** molecular weight

**NADPH:** nicotinamide adenine dinucleotide 2'-phosphate reduced

**NMR:** nuclear magnetic resonance

**OD:** optical density

**PCR:** polymerase chain reaction

**RNA:** ribonucleic acid

**SDS-PAGE:** sodium dodecyl sulfate polyacrylamide gel electrophoresis

**TPS:** terpene synthase

**Tris:** 2-amino-2-hydroxymethyl-1,3-propanediol

**UDP:** uridine 5'-diphosphate

**UDP-Glc:** UDP-glucose

**UDP-GlcA:** UDP-glucuronic acid

**UGDH:** UDP-glucose 6-dehydrogenase

**UGT:** UDP-glycosyltransferase

**UV:** ultraviolet