UNIVERSITÉ DE TOURS

ÉCOLE DOCTORALE : SSBCV

EA2106 « Biomolécules et Biotechnologies Végétales »

THÈSE présentée par :

Jennifer PERRIN

Soutenue le : **31 janvier 2023**

pour obtenir le grade de : **Docteur de l'université de Tours**

Discipline/ Spécialité : Sciences de la Vie et de la Santé

Approches d'ingénierie métabolique en levure pour la sécurisation de la production de l'anticancéreux Etoposide en Région Centre-Val de Loire

THÈSE dirigée par :

M. COURDAVAULT Vincent	Maître de Conférences, Université de Tours
Mme. GUIVARC'H Nathalie	Professeure des Universités, Université de Tours

Co-encadrée par :

M. BESSEAU Sébastien	Maître de Conférences, Université de Tours

RAPPORTEURS :

M. LARBAT Romain	Chargé de recherche, INRAE d'Angers
M. PAPON Nicolas	Maître de Conférences, Université d'Angers

JURY :

Mme AUGE-GOUILLOU Corinne	Professeure des Universités, Université de Tours, *Présidente du jury*
M. BESSEAU Sébastien	Maître de Conférences, Université de Tours
M. COURDAVAULT Vincent	Maître de Conférences (HDR), Université de Tours
Mme. FLINIAUX Ophélie	Maître de Conférences, Université de Picardie
Mme. GUIVARC'H Nathalie	Professeure des Universités, Université de Tours
M. LARBAT Romain	Chargé de recherche (HDR), INRAE d'Angers
M. PAPON Nicolas	Professeur des Universités, Université d'Angers
Mme. REMAUD-SIMEON Magali	Professeure des Universités, INSA de Toulouse
M. JEANNEAU Rémi (invité)	Chef de Laboratoire Axyntis, Pithiviers

REMERCIEMENTS

La thèse est une étape importante dans la vie d'un chercheur, elle lui permet de trouver sa voie, d'évoluer dans le monde de la recherche, et de se spécialiser. Mais cette épreuve peut difficilement être réalisable sans accompagnement professionnel et psychologique. C'est pourquoi, je tiens à remercier du fond du cœur toutes les personnes qui ont été à mes côtés pour franchir chaque étape, que ce soit dans la vie professionnelle ou dans la vie personnelle.

Particulièrement, je tiens à remercier le Dr. Nathalie Guivarc'h, professeure des Universités et directrice de l'EA2106 « Biomolécule et Biotechnologies Végétales » d'avoir cru en moi et de m'avoir offert l'opportunité de réaliser ce projet au sein de son équipe, ainsi que Sébastien Rose, Vice-Président de la Recherche et du Développement au sein du groupe Axyntis d'avoir suffisamment cru en moi pour me permettre de réaliser cette thèse CIFRE en collaboration avec son entreprise.

Je voudrais également adresser toute ma gratitude au Dr. Vincent Courdavault et au Dr. Sébastien Besseau, maîtres de conférence de l'EA2106 « Biomolécule et Biotechnologies Végétales » pour leur confiance et les connaissances qu'ils ont su partager avec moi. Leur disponibilité ainsi que la qualité de leur encadrement m'ont permis d'évoluer et de prendre en autonomie tout au long de ce projet. Leur patience, leur résilience et leur passion pour le travail ont été, et resteront, pour moi, une énorme source d'inspiration. Ce travail n'aurait pas été possible sans leur aide et leur soutien constant.

Je remercie chaleureusement Dr. Romain Larbat, chargé de recherche à l'INRAE et Dr. Nicolas Papon, professeur des Universités, tous deux rapporteurs, ainsi que Dr. Magali Remaud-Siméon, professeure des Universités, examinatrice, Dr. Corinne Auge-Gouillou, professeure des universités et Dr. Ophélie Fliniaux, maître de conférences, d'avoir accepté d'évaluer le fruit de ces trois années de recherche.

Je saisis cette occasion pour remercier tous les membres de l'EA2106 « Biomolécule et Biotechnologies Végétales » pour leur soutien et le partage de leurs connaissances. En particulier, je remercie Dr. Natalja Kulagina d'avoir été la première à m'accueillir et à m'aiguiller sur le fonctionnement du laboratoire. Je remercie également Dr. Louis Valentin Meteignier, Dr. Ines Carqueijeiro, Dr. Gregory Guirimand et Dr. Konstantinos Koudounas pour l'entre-aide et leur bonne humeur au laboratoire. Je suis reconnaissante envers toutes les

personnes du laboratoire qui ont su être à l'écoute et m'aiguiller, notamment Dr. Arnaud Lanoue pour les analyses UPLC/MS et le soutien lors des tests de phytochimie, Dr. Johan De Craene et ses connaissances inébranlables des levures, Dr. Marc Clastre et ses superbes blagues, Emeline Marais qui trouve toujours la verrerie dont j'ai besoin. Merci à Dr. Emily Stander d'avoir initié les recherches bio-informatiques sur le projet.

Je suis reconnaissante envers Dr. Jillian Marc et Céline Melin de m'avoir partagé leurs connaissances en bioprocédés et Rémi Jeanneau pour son accueil au sein du laboratoire de Recherche et Développement d'Axyntis.

Et puisque la thèse est aussi une aventure humaine, je remercie les « jeunes » de l'EA2106 qui ont su me soutenir dans les beaux jours comme dans les plus mauvais. Mon premier colocataire de bureau Antoine Audéon, la maman du bureau toujours à l'écoute et de bons conseils Manon Ferrier, la superbe danseuse Chaïma Sbai, Dr. Clément Cuello et son rire communicatif, ses bonbons abracadabrants et ses précieuses connaissances en bioinformatique, Dr. Pamela Cruz et sa gentillesse à toute épreuve et son petit accent brésilien, Nicolas Gautron et ses Kinder, Enzo Lezin le petit bébé thésard, et finalement Dr. Duchesse Lacours pour sa bienveillance. Je tiens également à remercier Cécile Abdallah, le soleil de la bande, Dr. Marin-Pierre Gémin de nous partager sa passion pour la bière et Dr. Valentin Vergès d'être toujours lui-même, un peu sataniste, mais empathique. Je remercie aussi Dr. Léo Bellenger et Cassandre Maillard d'être toujours prêts à aller boire un verre pour se vider la tête.

Finalement, je suis profondément reconnaissante envers mes proches. Mes parents et ma meilleure amie d'enfance Valessa qui m'ont soutenu tout au long de mon parcours d'étudiante, bien que mon domaine d'études leur soit flou, Silène pour son soutien sans failles depuis 10 ans et nos appels indispensables à ma santé mentale, Thibault de m'avoir toujours encouragé à garder la tête haute et à ne pas baisser les bras, Marc et Alice pour avoir maintenu à flot le bureau des plaintes et toutes les autres personnes qui m'ont suivi de près ou de loin. Un grand merci à Lady et Django pour leurs câlins, leurs visites sur le clavier de l'ordinateur pendant la rédaction et la joie de trouver des poils dans mon café le matin. Et bien évidemment, je remercie du fond du cœur Corentin pour son soutien inconditionnel et pour son amour.

“ La gratitude c’est de savoir que toutes les personnes qui croisent notre chemin sont nécessaires à notre évolution.” Lume

TABLE DES MATIERES

PREAMBULE……1

INTRODUCTION ……6

Chapitre I : Ingénierie métabolique……..7

I.1 Généralités sur l'ingénierie métabolique et les biotechnologies……..……7

I.2 L'ingénierie métabolique pour la production de molécules biosynthétiques végétales…….12

a) Elucidation des voies de biosynthèse végétales…….……...……13

1. Approches bio-informatiques pour l'identification des gènes……...……13

2. Validation fonctionnelle des gènes candidats……...……...….15

3. Réactions enzymatiques des voies de biosynthèse végétales……....18

Réactions d'oxydoréduction et alcool déshydrogénases……..19

Réactions catalysées par les cytochromes P450……..21

Réactions d'oxydation et dioxygénases……24

Méthylation par les O-méthyltransférases……...25

Acylation par les acyltransférases……...26

b) Production de métabolites spécialisés végétaux par ingénierie métabolique.....28

I.3 La levure, un châssis cellulaire……..……31

a) Modalités de bioproduction en levure…….....31

b) Mécanismes de transfert de gènes en levure……33

c) Outils et stratégies d'optimisation des cellules « usines » de levure……...37

1. Développement d'outils d'édition du génome et d'assemblage de fragments d'ADN multiples ……38

2. Outils de régulation transcriptionnelle de l'expression génique……...….39

3. Outils d'optimisation de la traduction……...……..41

4. Régulation du métabolisme primaire et recyclage des co-facteurs ..……...42

5. Reconfiguration spatiale……43

6. Contournement des fuites métaboliques……..….45

7. Stratégies adaptatives d'évolution……...46

8. Perspectives d'optimisation complémentaires des cellules « usines » de levure……..47

d) Optimisation de la conduite de culture de levures par le génie des procédés……49

Chapitre II : Les vertus pharmaceutiques de la podophyllotoxine ……..……..51

II.1 Généralités sur la podophyllotoxine……..……51

II. 2 Utilisation de la podophyllotoxine en santé humaine ……..53

a) Propriétés pharmacologiques de la podophyllotoxine……....53

b) L'étoposide……..……...…….54

II.3 Approvisionnement en podophylloxotine……55

a) Podophyllum hexandrum, une espèce menacée d'extinction……...55

b) Productions alternatives de podophyllotoxine……57

II.4 Découverte de la voie de biosynthèse de la podophyllotoxine *in planta*……...58

Chapitre III : Exploration du potentiel pharmaceutique des tri-hydroxycinnamoyl spermidines……61

III.1 Généralités sur les tri-hydroxycinnamoyl spermidines……..……..61

a) Les phénolamides……61

b) Diversité des THCSpd……63

III.2 Les tri-hydroxycinnamoyl spermidines et leurs bienfaits……...65

III.3 Voie de biosynthèse……66

III.4 Limites à l'étude des THCSpd……68

OBJECTIFS ET ORGANISATION DE LA THESE……...……..70

RESULTATS…...……..74

Partie 1 : Production de phénolamides par ingénierie métabolique…….……75

Partie 2 : Test de faisabilité d'un système de production de 4'dEPT en levure: levée de verrous techniques……...……..…114

2.1 Résultats ……..……...114

2.1.1 Evaluation de la toxicité des dérivés de PPT sur les levures ……114

2.1.2 Evaluation de l'efficacité catalytique des enzymes de la voie de biosynthèse du 4'dEPT en levure ……..…….. 115

a) Activité de la 2ODD en levure ……...… 116

b) Activité de CYP71BE54 en levure……...…… 117

c) Activité de CYP82D61 en levure…… 118

2.1.3 Alternatives à l'utilisation de CYP82D61 pour catalyser l'étape finale de bioproduction de 4'dEPT 119

a) Recherche d'isoformes de CYP82D61 plus efficaces pour la synthèse de 4'dEPT 120

b) Comparaison des activités catalytiques de CYP82D61 entre espèces du genre Podophyllum 120

c) Exploitation de la CYP3A4 humaine.....122

2.1.4 Optimisation de réduction de CYP82D61 dans la levure.....123

a) Substitution du domaine transmembranaire d'ancrage à la membrane du réticulum endoplasmique.....124

b) Fusion de CYP82D61 à une CPR 125

c) Criblage de CPR végétales pour leur efficacité de réduction de CYP82D61.....126

2.2 Conclusion 127

2.3 Matériel et méthodes.....129

2.3.1 Test de toxicité.....129

2.3.2 Gènes d'intérêt 129

2.3.3 Construction des plasmides.....130

2.3.4 Fusions de fragments d'ADN.....131

2.3.5 Souches de levure 131

2.3.6 Transformation des levures.....132

2.3.7 Test de conversion enzymatique.....132

2.3.8 Analyses métaboliques 133

2.3.9 Western Blot.....134

Partie 3 : Ingénierie métabolique de la levure pour la bioconversion de YAT en 4'dEPT : construction d'une souche industrialisable.....135

3.1 Résultats 135

3.1.1 Première souche stable complète pour la biosynthèse de 4'dEPT.....135

3.1.2 Optimisation de la réduction de CYP82D61 dans la souche stable.....136

3.1.3 Optimisation de la traduction de CYP82D61.....137

3.1.4 Multiplication du nombre de copies de CYP82D61.....139

3.1.5 Recherche de fuites métaboliques.....141

3.1.6 Relocalisation de la 2ODD au peroxysome.....142

3.1.7 Multiplication du nombre de copies de 2ODD et CYP71BE54.....143

3.1.8 Altération de la croissance cellulaire.....145

3.2 Conclusion.....146

3.3 Matériel et méthodes……..147
3.3.1 Construction des plasmides……147
3.3.2 Souche de levure ……148
3.3.3 Transformation des levures……....149
3.3.4 Test de conversion enzymatique……149
3.3.5 Courbes de calibration……...…150
3.3.6 Analyses statistiques……...…151

Partie 4 : Biosourcing et purification de YAT pour une montée de la production de 4'dEPT à l'échelle d'un bioréacteur……...152

4.1 Résultats ……...……152
4.1.1 Identification d'espèces végétales productrices de YAT…..……152
4.1.2 Mise en place d'un protocole d'extraction et de purification de la YAT…….154
4.1.3 Première montée en échelle de la bioproduction……156
4.2 Conclusion ……..157
4.3 Matériel et méthodes…….……158
4.3.1 Extraction de la YAT à partir de matière végétale……158
4.3.2 Purification de la YAT……....158

CONCLUSIONS GENERALES ……...……...……..159

PERSPECTIVES ……...……...……..169

BIBLIOGRAPHIE ……...……..……..175

ANNEXES ……...…...……...……...208

Annexe I: Efficient Terpene Production by Marine Thraustochytrids : Shedding Light on the Thermodynamic Driving Force……..……...……..209

Annexe II: Boosting lignan-precursor synthesis in yeast cell factories through co-factor supply optimization ……..……...……..214

Annexe III: Tonoplast and Peroxisome Targeting of γ-tocopherol *N*-methyltransferase Homologs Involved in the Synthesis of Monoterpene Indole Alkaloids……...……...……..221

Annexe IV: A Rapid and Efficient Vacuum-Based Agroinfiltration Protocol for Transient Gene Overexpression in Leaves of *Catharanthus roseus*……...……...……..247

DONNEES SUPPLEMENTAIRES……..……..265

TABLE DES ILLUSTRATIONS

Figure 1 : Familles de métabolites spécialisés d'intérêt retrouvées dans le règne végétal…….2

Figure 2 : Cycle de la biologie synthétique……...9

Figure 3 : Reconstitution de la voie de biosynthèse de l'hydrocortisone chez *S. cerevisiae*.……..10

Figure 4 : Produits biopharmaceutiques approuvés par la FDA produits dans des cellules de levure…..12

Figure 5 : Représentation schématique du fonctionnement de deux technologies de séquençage de troisième génération……13

Figure 6 : Représentation schématique du fonctionnement de l'outil Trinity…….13

Figure 7 : Représentation schématique de l'utilisation de données transcriptomiques à haut débit pour la découverte de gènes biosynthétiques MIA……..14

Figure 8 : Histoire évolutive du groupe de gènes BIA……15

Figure 9 : Représentation schématique de la technique de répression d'expression génique induite par un virus (VIGS)……...17

Figure 10 : Regroupement physique (cluster) au sein du génome de *C. roseus* des gènes connus de la voie MIA……19

Figure 11 : Les voies de biosynthèse de la tyrosine et de la bétalaïne……..20

Figure 12 : Voie phénylpropanoïde menant aux coumarines……..20

Figure 13 : Analyse phylogénétique des P450 connues pour catalyser les modifications structurelles sur les squelettes stéroïdiens et triterpéniques……...……21

Figure 14 : Schéma des complexes CPR-P450 chez les plantes…….22

Figure 15 : Représentation schématique d'une cytochrome P450……..22

Figure 16 : Cladogramme circulaire de 246 P450 pleine longueur d'*A. thaliana*……22

Figure 17 : Phylogénie du clade CYP71…….....23

Figure 18 : Mécanisme catalytique des enzymes P450……...24

Figure 19 : Exemples de réaction d'une dioxygénase et d'une monooxygénase…….24

Figure 20 : Classification des 2OGB végétales……...25

Figure 21 : Voie de synthèse de la morphine et de la berbérine chez *P. somniferus*……26

Figure 22 : Utilisation de la tabersonine dans les différents organes de *C. roseus*……...27

Figure 23 : Chronologie de la production de métabolites spécialisés végétaux dans des organismes hétérologues……28

Figure 24 : Voie de biosynthèse vers la synthèse de l'acide artémisinique en levure……..28

Figure 25 : Voies de biosynthèses vers la synthèse de resvératrol en levure……...29

Figure 26 : Voie de biosynthèse reconstituée de la stricosidine chez *S. cerevisiae*…….29

Figure 27 : Voie biosynthétique complète pour la production de vinblastine dans la levure……...30

Figure 28 : Biosynthèse de novo de la noscapine chez *S. cerevisiae*……..……...31

Figure 29 : Mécanisme de fonctionnement du système CRISPR/Cas9……...36

Figure 30 : Loci d'intégration dans le génome de la levure…….36

Figure 31 : Illustration des différents outils de biologie de synthèse utilisés pour la construction et l'optimisation des cellules de levure usines…….38

Figure 32 : Fonctionnement de l'assemblage de Gibson pour la construction de plasmides multigéniques……..39

Figure 33 : Codons utilisés préférentiellement chez les plantes et la levure……41

Figure 34 : Exemple de comparaison des séquences nucléotidiques du cytochrome P450 CYP82D61 natif de *Podophyllum hexandrum* et de sa version optimisée pour la levure……....41

Figure 35 : Voie de biosynthèse du squalène chez *S. cerevisiae*…….42

Figure 36 : Biosynthèse de S-Adénosyl-L-méthionine (SAM) dans une souche *S. cerevisiae* recombinante……...43

Figure 37 : Utilisation du peroxysome pour limiter la concurrence entre les substrats……...44

Figure 38 : lllustration de la conception de sélection couplée à la croissance chez *S. cerevisiae*……47

Figure 39 : Couplage de l'ingénierie métabolique classique et d'un stratégie adaptative d'évolution pour la synthèse d'isobutanol chez *S. cerevisiae*…….47

Figure 40 : Schéma d'un bioréacteur pour l'expérimentation en laboratoire……..49

Figure 41 : Principales familles de composés dérivant du métabolisme des phénylpropanoïdes……51

Figure 42 : Base de la nomenclature des lignanes issues du couplage de deux alcool coniféryliques par une liaison β-β…….51

Figure 43 : Structure de la (-)-podophyllotoxine……52

Figure 44 : Sources fongiques signalées comme source de podophyllotoxine……52

Figure 45 : Structure de l'étoposide……....54

Figure 46 : Schéma d'une fourche de réplication les contraintes topologiques qui y sont associées…...54

Figure 47 : Etoposide sous forme liquide utilisée dans les protocoles de chimiothérapie……...54

Figure 48 : Synthèse chimique de l'étoposide à partir de la podophyllotoxine……...55

Figure 49 : *Podophyllum hexandrum* et ses différents organes……...56

Figure 50 : Durée des ruptures de stock des médicaments contenant de l'étoposide de novembre 2016 à novembre 2018…….57

Figure 51 : Voie de biosynthèse de la podophyllotoxine……58

Figure 52 : Exemples de la diversité des phénolamides……..61

Figure 53 : Structures des différents acides hydroxycinnamiques et de la spermidine……64

Figure 54 : Structure de la tri-hydroxycinnamoyl spermidine (THCSpd)……...64

Figure 55 : Phylogénie des N-hydroxycinnamoyltransférases caractérisées……..67

Figure 56 : Synthèse de la tri-hydroxycinnamoyl spermidine (THCSpd)……...68

Figure 57 : Schéma de la stratégie de bioconversion de 4'dEPT à partir de YAT en levure……71

Figure 58 : Test de toxicité de dérivés de lignanes……115

Figure 59 : Activité catalytique de la 2ODD en levure…….116

Figure 60 : Activité catalytique de CYP71BE54 en levure……...117

Figure 61 : Activité catalytique de CYP82D61 en levure…….118

Figure 62 : Comparaison de l'accumulation de CYP71BE54 et CYP82D61 en levure par Western Blot……119

Figure 63 : Comparaison de l'accumulation des dérivés de 4'dDPT hydroxylés en C-7 dans les feuilles et le rhizome de *Podophyllum hexandrum*, *Podophyllum peltatum* et *Podophyllum spotty dotty*……...120

Figure 64 : Comparaison des structures protéiques de PhCYP82D61 et PsdCYP82D61…….121

Figure 65 : Comparaison des activités catalytiques des orthologues de CYP82D61 en levure…….122

Figure 66 : Alternative d'hydroxylation de la DPT et de la 4'dDPT par CYP3A4……123

Figure 67 : Prévision des DTM de CYP82D61 et CYP71BE54 par le logiciel TMHMM……124

Figure 68 : Représentation schématique des protéines chimériques construites à partir de CYP82D61……124

Figure 69 : Effet de la substitution du DTM de CYP82D61 sur son activité catalytique……...125

Figure 70 : Activité enzymatique des CYP82D61 fusionnées à la CPR……...125

Figure 71 : Impact de l'origine des CPR co-exprimées sur l'efficacité catalytique de CYP71BE54 et CYP82D61 en levure……...….126

Figure 72 : Exemple de carte d'un plasmide pESC utilisé pour les clonages……130

Figure 73 : Schémas des méthodes utilisées dans ce projet pour la fusion entre deux fragments d'ADN……..132

Figure 74 : Activités catalytiques de la 2ODD, CYP71BE54 et CYP82D61 dans la souche stable…..135

Figure 75 : Impact de l'expression de CrCPR2 (Eto2.1) ou PhCPR2 (Eto2.2) sur l'efficacité catalytique de CYP82D61 dans la souche stable……137

Figure 76 : Comparaison des activités catalytiques des versions optimisées de CYP82D61 dans la souche stable……138

Figure 77 : Production de 4'dEPT en fonction du nombre de copies de CYP82D61 et ses versions optimisées dans la souche stable……139

Figure 78 : Evolution au cours du temps de la quantité relative de chaque intermédiaire en fonction du nombre de copies de CYP82D61 et ses versions optimisées dans la souche stable……140

Figure 79 : Production de PPT/EPT par la levure en absence de CYP71BE54……141

Figure 80 : Production de DPT par les souches comportant la 2ODD native et la 2ODD orientée au peroxysome……...142

Figure 81 : Production de 4'dEPT en fonction du nombre de copies de 2ODD et CYP71BE54 dans la souche stable……143

Figure 82 : Evolution au cours du temps de la quantité relative de chaque intermédiaire en fonction du nombre de copies de 2ODD et CYP71BE54 dans la souche stable……144

Figure 83 : Exemple de carte d'un plasmide intégratif utilisé pour les clonages……...146

Figure 84 : Gels d'électrophorèse obtenus après amplification par PCR sur l'ADNg de chaque souche……...149

Figure 85 : Recherche d'espèces végétales productrices de YAT……152

Figure 86 : Recherche de l'espèce de Juniperus accumulant le plus de YAT……153

Figure 87 : Evolution de la purification de YAT à partir d'extrait de *J. scopulorum*……154

Figure 88 : Première montée en échelle de la culture de la souche EtoA……..155

Figure 89 : Analyses par RMN de la structure de la 4'dEPT produite dans le milieu de culture……156

LISTE DES TABLES

Tableau 1 : Taux de production en levure des métabolites spécialisés présentés……32

Tableau 2 : Liste des amorces utilisées pour l'amplification des gènes d'intérêt……..129

Tableau 3 : Liste des différents plasmides construits pour réplication autonome en levure……..……...131

Tableau 4 : Tests de croissance réalisés sur les souches CEN.PK113-7D contenant un nombre variable de copies de gènes codant des P450……...145

Tableau 5 : Liste des amorces utilisées pour l'amplification des gènes d'intérêt……..146

Tableau 6 : Liste des différents plasmides construits et des loci d'intégration ciblés associés……..147

Tableau 7 : Liste des ARNg correspondant à chaque locus d'intégration……148

Tableau 8 : Amorces utilisées pour les PCR de vérification……148

Tableau 9 : Liste des souches construites et le génotype associé……..150

ABREVIATIONS

2ODD : désoxypodophyllotoxine synthase

2OGB : dioxygénases 2-oxoglutarate-dépendantes

4'dDPT : (-)-4'-déméthyl-DéoxyPodophylloToxine

4'dEPT : (-)-4'- déméthyl-EpiPodophylloToxine

ADN : Acide DésoxyriboNucléique

ADNc : ADN complémentaire

ADNg : AND génomique

AFLP : Amplified Fragment Lenght Polymorphism

AHCT: Anthocyanin *O*-hydroxycinnamoyltrasnferase

ARN : Acide Ribonucléique

ARNg : ARN guide

ARNt : ARN de transfert

BAHD: BEAT AHCT HCBT DAT

BEAT: Benzylalcohol *O*-acetyltransferase

BIA : Alcaloïde de benzylisoquinoline

CaA: Acide Caféique

Cas9 : CRISPR-associated protein 9

CBD : cannabidiol

CCM : Chromatographie sur Couche Mince

COOL : Codon Optimization OnLine

CPR : Cytochrome P450 Réductase

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

DAT: Deacetylvindoline 4-*O*-acetyltransferase

DGAT : Acyl-CoA-DAG acyltransférase

DPT : (-)-DéoxyPodophylloToxine

DTM : Domaine TransMembranaire

EPT : épipodophyllotoxine

FA : Acide Férulique

FAD : Flavine Adénine Dinucléotide

FMN : Flavine MonoNucléotide

GFP : Green Fluorescent Protein

GPAT : glycéraldéhyde-3-phosphate acyltransférase

HCAA : HydroxyCinnamic Acid Amides

HCBT: N-hydroxycinnamoyl/benzoyltransferase

HIV : Virus de l'Immunodéficience Humaine

HPLC : High Performance Liquid Chromatography

HPV : PapillomaVirus Humain

LB : Luria-Bertani

LPAAT : acides lysophosphatidique acyltransférase

MIA : alcaloïde indolo-monoterpénique

MS : Métabolites Secondaires/Spécialisés

NADPH : Nicotinamide Adénine Dinucléotide Phosphate

NGS : Next Generation Sequencing

NHEJ : Non-Homologous End-Joining

OMT : *O*-méthyltransférase

ORF : Open Reading Frame

P450 : Cytochrome P450

Pbs : paire de bases

***p*CA :** acide *p*-coumarique

PCR : Polymerase Chain Reaction

PPT : Podophyllotoxine

RE : Réticulum Endoplasmique

RMN : Résonance Magnétique Nucléaire

Rpm : Rotation par minute

SAH : S-adénosyl-L-Homocystéine

SAM : S-adénosyl-L-méthionine

SDS-PAGE : Sodium Dodecyl Sulfate PolyacrylAmide Gel Electrophoresis

SERT : Transporteur de la Sérotonine

SHT : Spermidine N-HydroxycinnamoylTransférase

SIM : Single Ion Monitoring

SSH : Hybridation par Suppression Soustractive

TALENs : Transcription activator-like effector nucleases

TGS : Third Generation Sequencing

THC : Δ9-TetraHydroCannabinol

THCSpd : Tri-HydroxyCinnamoyl Spermidines

THIC : TétraHydroIsoQuinoline

TVR : Virus du Rattle du Tabac

SAM : S-Adénosyl-l-Méthionine

UV: Ultra-Violets

VIGS : Virus-Induced Gene Silencing

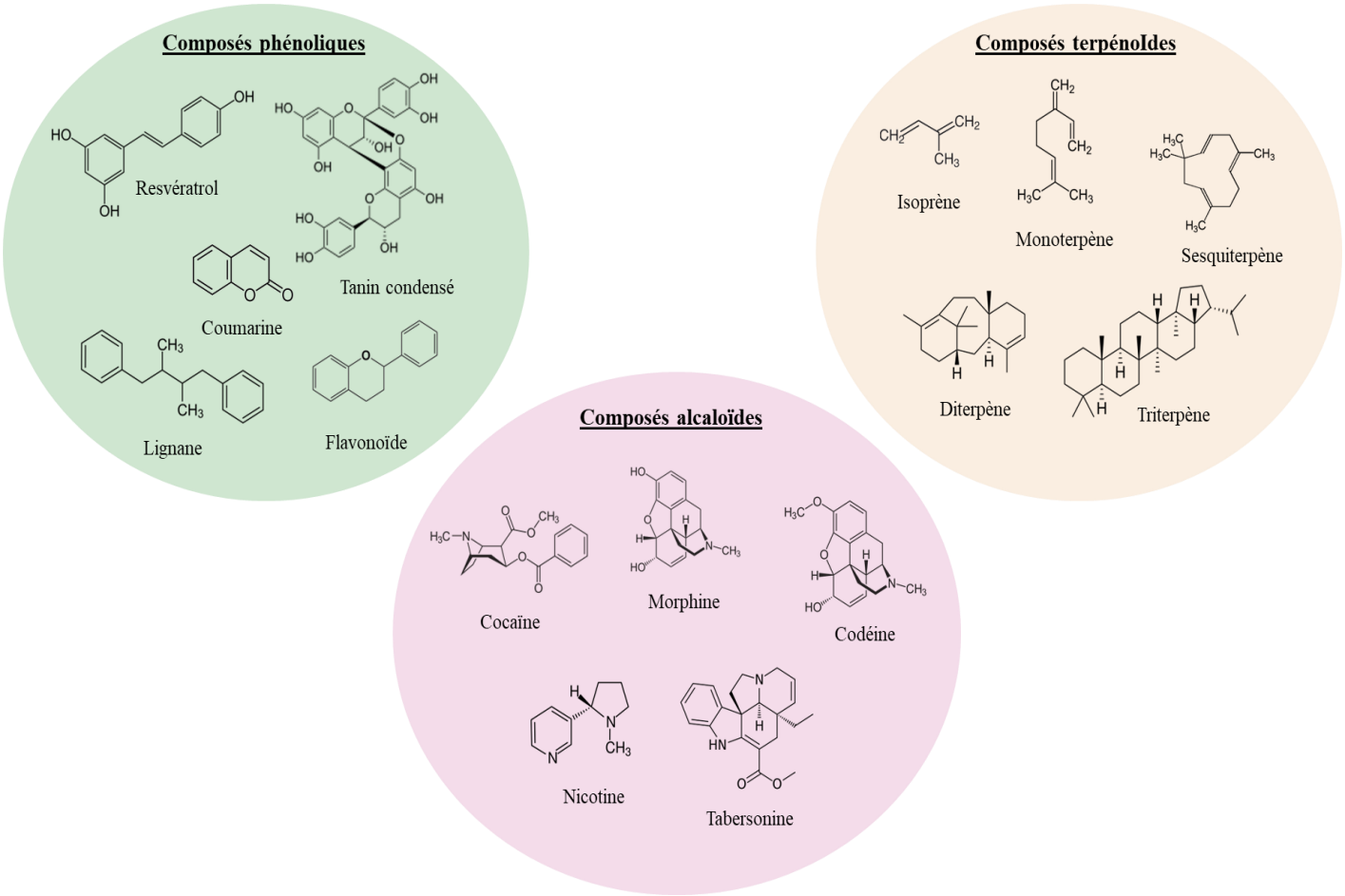
VRSA : Vancomycin-Resistant *S. aureus*

YAT : (-)-Yatéine

YPD : Yeast extract-Peptone-Dextrose

ZFN : Zinc Finger Nucleases

PREAMBULE



<u>Figure 1</u> : Familles de métabolites spécialisés d'intérêt retrouvées dans le règne végétal.

Composés phénoliques : Le resvératrol est un polyphénol disposant d'activités anti-oxydante, anti-inflammatoire et antitumorale, la coumarine est une molécule aromatique utilisée en parfumerie et en tant qu'arôme, les tanins sont utilisés pour le tannage des peaux, la fabrication d'encre ou encore dans l'industrie pharmaceutique (propriétés antioxidante, antivirale, antibactérienne, anticancéreuse, …), les lignanes ont des activités anticancéreuses, antioxydantes, anti-inflammatoire et ostrogéniques, les principales propriétés des flavonoïdes sont antiseptiques, antibactériennes et anti-inflammatoires.
Composés terpénoïdes : Ces composés disposent, en particulier, d'activités antibactérienne et antinéoplasiques, mais sont aussi utilisés en parfumerie et en tant qu'arôme.
Composés alcaloïdes : la cocaïne est un psychotrope et un vasoconstricteur, la morphine disposent de propriétés analgésique et euphorisante, la codéine est utilisée comme analgésique et antitussif narcotique, la nicotine est un agoniste de récepteurs de l'acétylcholine, la tabersonine est un précurseur de synthèse de l'anticancéreux vindoline.

Les plantes assurent la synthèse d'une grande variété de molécules très diverses chimiquement, via leurs métabolismes primaire et spécialisé. Le métabolisme primaire est impliqué dans les fonctions générales et vitales de la plante, telles que la division et la croissance cellulaires, la respiration, la reproduction et le stockage d'acides aminés, de glucides ou de lipides (Bourgaud et al., 2001). Le métabolisme spécialisé quant à lui, est associé à des fonctions plus spécifiques liées à l'interaction des plantes avec leur environnement (Pott et al., 2019). Il a ainsi permis aux végétaux de s'adapter à leur environnement au cours de l'évolution. Les métabolites spécialisés (MS) sont également synthétisés par les végétaux en réponse à des stress biotiques et abiotiques tels que les bioagresseurs, les UV, la gravité, la déshydratation, etc. La présence et la localisation des MS au sein de l'organisme sont corrélés à une fonction spécifique. De façon générale, plus de 100 000 MS ont été répertoriés à ce jour. Ces MS peuvent être organisés en trois superfamilles : les composés phénoliques, les terpénoïdes et les alcaloïdes (**Figure 1**) (Bourgaud et al., 2001). Les composés phénoliques sont caractérisés par la présence d'un noyau benzénique portant au minimum un groupement hydroxyle. Ils représentent les composés les plus abondants du règne végétal et dérivent du métabolisme des phénylpropanoïdes. On retrouve, parmi les composés phénoliques, la lignine, les coumarines, les lignanes, les flavonoïdes ou encore les tanins. Les terpènes sont issus de l'assemblage d'unités isoprènes contenant cinq atomes de carbone. Leur squelette carboné peut être arrangé de façon linéaire ou cyclique. On retrouve notamment, parmi les terpènes, l'isoprène, les monoterpènes, les sesquiterpènes, les diterpènes et les triterpènes composés respectivement d'une, deux, trois, quatre ou six unités isoprènes. Enfin, les alcaloïdes contiennent un ou plusieurs atome(s) d'azote généralement inclu(s) dans un cycle. On retrouve, par exemple, dans cette famille, la cocaïne, la morphine, la codéine, la sérotonine, la vincristine et la nicotine.

Les humains ont toujours utilisé les plantes pour se nourrir et leur jus ou extraits comme remèdes contre diverses pathologies. En effet, les premiers écrits mentionnant des recettes médicinales à base de plantes remontent à 5 000 ans avant Jésus-Christ avec, notamment, la tablette sumérienne de Nippur contenant une douzaine de recettes à base de plantes, dont le pavot à opium (Djurdjevic et al., 2013). Les bienfaits pharmacologiques des végétaux peuvent être attribués à certains métabolites primaires comme les sucres particuliers ou les lipides, mais sont souvent attribués aux MS possédants des activités antimicrobiennes, diurétiques, anti-inflammatoires, antiallergiques, antinéoplasiques, hypertenseurs, hypotenseurs, etc. Les MS sont ainsi devenus indispensables et sont aujourd'hui présents dans de nombreux produits de

santé, cosmétiques et agro-alimentaires. Par conséquent, la valeur et l'importance commerciale des plantes qui les produisent ont considérablement augmenté.

Un approvisionnement en grande quantité de ces MS est nécessaire à leur utilisation dans les différents domaines de l'industrie qui les exploitent et l'extraction à partir de biomasses végétales représente encore la principale étape d'obtention de ces composés. En effet, la synthèse chimique de molécules structurellement simples est faisable et supportable d'un point de vue économique. Néanmoins, la plupart des MS végétaux disposent d'une structure complexe, compliquant ainsi la mise en place d'une synthèse chimique exclusive et rentable de ce type de composés. De nombreuses plantes ont ainsi été exploitées pour leur capacité à fournir ces composés complexes, pouvant conduire à de nombreuses répercussions indésirables comme un fort impact environnemental ou encore la disparition de certaines espèces. C'est le cas, par exemple, de *Podophyllum hexandrum*, plante productrice de podophyllotoxine, précurseur moléculaire naturel indispensable à l'hémisynthèse de l'anticancéreux étoposide, qui a été classée sur la liste des espèces menacées d'extinction. La demande croissante pour l'approvisionnement de MS d'intérêt impose une énorme pression sur les ressources naturelles occasionnant parfois de longues pénuries d'approvisionnement. Des pénuries récurrentes, causées par des contraintes socio-économiques, géopolitiques et climatiques, ont ainsi été observées pour l'anticancéreux vincristine produit par *Catharanthus roseus* (Rabin, 2019). Afin de pallier ces carences régulières, il est donc nécessaire de développer des approches alternatives de production de ces MS d'intérêt. Ces approches sont diverses, on retrouve, par exemple, des approches biotechnologiques de production *in vitro* à partir de suspensions cellulaires végétales, de cals, de racines transformées ou encore de tiges feuillées. Mais ces approches ne sont pas toujours suffisantes pour une production à une échelle industrielle, et certains MS sont difficiles à produire. Une nouvelle alternative d'approvisionnement en pleine expansion repose sur l'ingénierie métabolique de voies de biosynthèse en système hétérologue. Le pré-requis de ce type de développement est une connaissance exhaustive des voies de biosynthèse conduisant aux MS d'intérêt. C'est pourquoi, de nombreux travaux ont été, et sont encore, menés sur l'élucidation et la compréhension des voies métaboliques *in planta*. Les connaissances déjà acquises pour certains métabolismes permettent aujourd'hui d'envisager le transfert des voies métaboliques végétales dans des organismes hétérologues comme la levure afin de les transformer en « cellules usines ». La maîtrise de leur culture à grande échelle pourra alors permettre de sécuriser l'approvisionnement de certains MS d'intérêt, comme cela a déjà été démontré pour l'anti-paludéen artémisinine, dont la bioproduction industrielle à partir de

levure transformée a été réalisée par SANOFI (Paddon et al., 2013). Basées sur le même principe, des souches de levures productrices de vindoline, précurseur de la vincristine, ont été récemment établies (Kulagina et al., 2021c), laissant présager, à court terme, une bioproduction de cet anticancéreux par cette voie.

C'est dans cette optique qu'ont été développés mes travaux de thèse CIFRE, inscrits dans le cadre du projet ETOPOCentre porté par le laboratoire EA2106 Biomolécules et Biotechnologies Végétales de l'Université de Tours et qui réunit le groupe industriel Axyntis et deux partenaires académiques, dont le Laboratoire de Biologie des Ligneux et des Grandes Cultures (LBLGC EA1207) et l'équipe de Méthodologie et hétérocycles de l'Institut de Chimie Organique et Analytique (ICOA UMR7311). Ce projet a été développé dans le cadre du programme Ambition Recherche & Développement de la Région Centre-Val de Loire (ARD CVL) Biomédicaments, financé par la Région Centre-Val de Loire. Le groupe industriel Axyntis a été créé en 2007 et travaille sur la synthèse chimique de molécules d'intérêt pharmaceutique, mais aussi de colorants. Le groupe s'est intéressé à une alternative d'approvisionnement de l'anticancéreux étoposide, originellement extrait du rhizome de *Podophyllum hexandrum*, une espèce menacée d'extinction. Pour cela, ils ont envisagé la création d'un pôle de bioproduction via la culture en bioréacteur d'une levure transformée productrice d'étoposide afin de sécuriser sa production en Région Centre-Val de Loire. Dans le cadre de ma thèse, mon projet principal a donc été de mettre en place une souche de levure stable productrice d'un lignane précurseur d'étoposide industrialisable. Pour cela, une stratégie de production par bioconversion d'un substrat hautement biodisponible a été envisagée, tel qu'il sera décrit dans les sections suivantes. A cette fin, les gènes issus de *P. hexandrum* impliqués dans la synthèse du précurseur d'intérêt ont été clonés dans la souche de levure industrielle via le système CRISPR/Cas9 et l'activité catalytique des enzymes impliquées dans la voie de bioconversion a été optimisée. Ce manuscrit décrira donc principalement la construction de ces souches et les nombreux tests de bioconversion que j'ai pu réaliser pour optimiser le procédé général.

Enfin, il est à noter qu'au début de ma thèse, j'ai également eu l'occasion de m'impliquer dans un projet initié avant mon arrivée au laboratoire BBV visant à produire des tri-hydroxycinnamoyl spermidines (THCSpd) en levure. Ces travaux m'ont permis de développer des compétences et connaissances en ingénierie métabolique en levure, qui ont ensuite été utilisées pour mener à bien mon projet principal. Les THCSpd sont des MS appartenant à la famille des phénolamides aux propriétés pharmacologiques prometteuses mais sous-étudiées

car leur obtention à partir de biomasse végétale est restreinte en termes de quantité et de diversité structurale. Dans ce contexte, l'objectif de ce projet était de réaliser la preuve de concept d'une plateforme de bioproduction à façon en levure de divers THCSpd naturels et non naturels par bioconversion, en fonction des précurseurs ajoutés au milieu de culture. Pour cela, j'ai participé à l'établissement de plusieurs souches de levure exprimant différentes combinaisons de gènes impliqués dans la synthèse de THCSpd issus de plusieurs espèces à l'aide de plasmides à réplication autonome. Un chapitre de cette thèse leur sera donc consacré.

INTRODUCTION

Chapitre I : Ingénierie métabolique

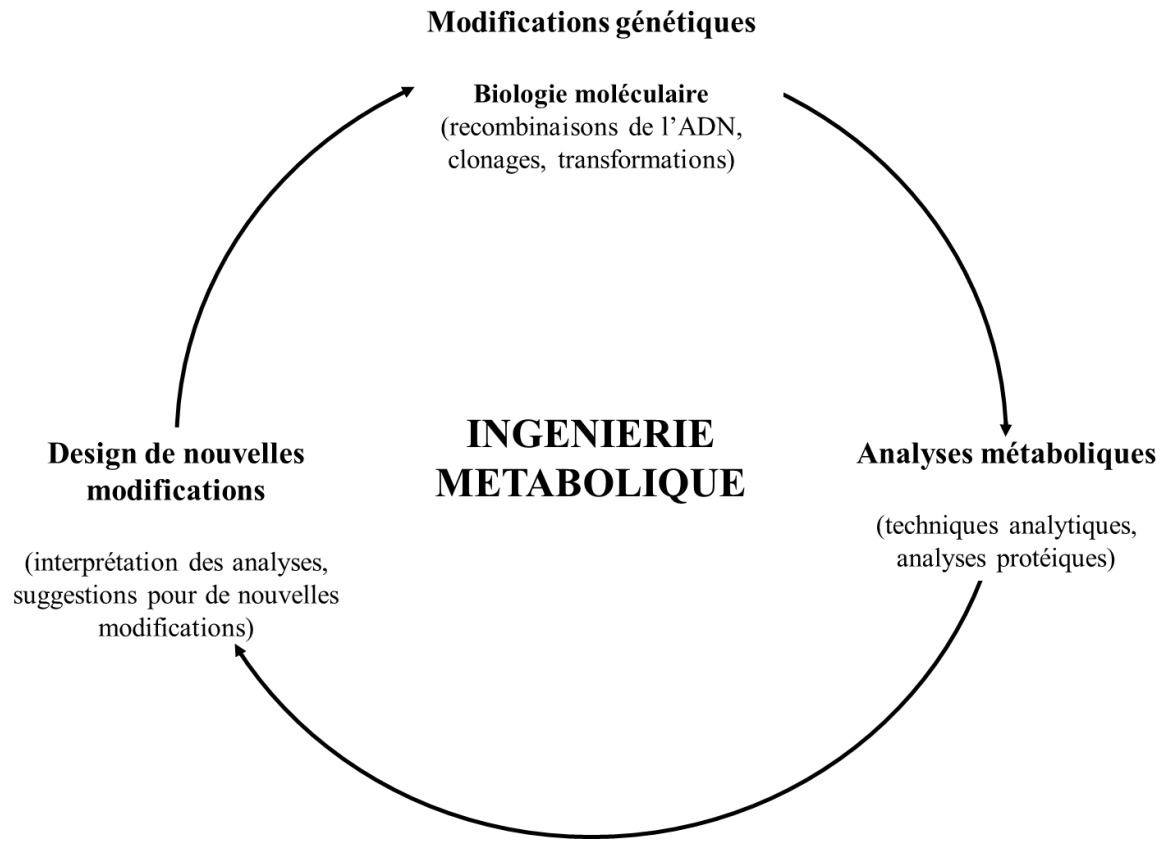
I.1 Généralités sur l'ingénierie métabolique et les biotechnologies

Les biotechnologies sont définies par l'OCDE, Organisation de Coopération et de Développement Économique, comme « l'application des principes scientifiques et de l'ingénierie à la transformation de matériaux par des agents biologiques pour produire des biens et services ». Les biotechnologies touchent un nombre important de domaines et, dans un effort de simplification didactique et de rationalisation, il a été proposé et accepté d'ordonner la grande variété des biotechnologies existantes selon un code de couleurs correspondant aux domaines principaux dans lesquels elles sont appliquées. **Les biotechnologies vertes** sont liées à l'agriculture, l'élevage et l'agro-alimentaire, elles permettent de produire et transformer des produits alimentaires. Elles reposent sur des connaissances anciennes complétées ces dernières décennies par la biologie moléculaire et les connaissances génétiques des organismes végétaux, dont un exemple d'application est la production d'Organismes Génétiquement Modifiés communément appelés OGM. Elles investissent également les champs des biomatériaux et de la production d'énergie. **Les biotechnologies blanches** regroupent les applications industrielles qui emploient les systèmes biologiques, généralement des micro-organismes, comme alternative aux procédés chimiques pour la fabrication de produits à l'échelle industrielle à partir de l'exploitation de biomasses en tant que matière première renouvelable. **Les biotechnologies jaunes** sont directement liées aux problématiques de préservation de l'environnement et à la dépollution par le développement de solutions de traitement ou d'élimination des polluants (bioremédiation). **Les biotechnologies bleues** sont liées à la biodiversité marine et concernent la valorisation du potentiel encore peu connu mais probablement immense des organismes marins. Finalement, **les biotechnologies rouges** concernent le domaine de la santé via la production de protéines thérapeutiques, la mise en place de diagnostics immunologiques, de thérapies géniques ou encore de thérapies anticancéreuses par exemple. Nombre des approches développées dans les différents domaines des biotechnologies nécessitent une édition génétique des organismes utilisés afin de les modifier, de les adapter ou de leur conférer de nouvelles capacités de synthèse - on parle dans ce cas, et de façon générique, d'ingénierie métabolique.

L'ingénierie métabolique est ainsi devenue une approche très importante, notamment pour l'optimisation des processus de fermentation industrielle grâce au développement de

technologies de recombinaison de l'acide désoxyribonucléique (ADN), permettant donc la modification génétique d'un organisme hôte hétérologue. Le développement d'outils très puissants permet désormais l'introduction de modifications génétiques dirigées et d'en analyser méticuleusement les fonctions cellulaires. En effet, les avancées dans le domaine génétique, et notamment les progrès de la génomique et de la transcriptomique, ont permis de mettre en évidence de nouvelles voies métaboliques, mais également les enzymes impliquées dans les différentes étapes de leur biosynthèse. De plus, l'évolution des techniques de visualisation des structures protéiques et l'ingénierie enzymatique associée permettent de concevoir de nouvelles enzymes pour des réactions qui ne sont pas existantes dans le monde du vivant. La maîtrise de nouveaux outils moléculaires, tel que le système CRISPR/Cas9, aide à contrôler de façon précise les voies métaboliques au sein des organismes hétérologues. Toutefois, de nombreux éléments sont à prendre en compte dans l'ingénierie métabolique. Le choix de l'hôte est, à ce titre, une étape importante. En effet, les enzymes issues d'organismes eucaryotes peuvent avoir à subir des modifications post-traductionnelles que les organismes procaryotes sont incapables de réaliser. De même, la conformation des protéines exogènes peut être plus ou moins altérée en fonction de l'organisme hôte, et par conséquent, leur activité ne reflétera pas celle de la protéine native. De plus, les modalités des mécanismes de transport membranaire peuvent différer d'un organisme à l'autre et le choix d'un hôte inadéquat peut ainsi influencer négativement l'adressage et le transport d'une enzyme dans le bon compartiment cellulaire. La toxicité de certains substrats introduits dans le milieu de culture ou produits par les enzymes hétérologues peut aussi être un frein à l'utilisation d'une cellule donnée (Keasling, 2010). En outre, il est important de bien connaître l'organisme hôte sélectionné afin d'adopter la méthode de transformation appropriée et d'adapter la séquence codante de la protéine à cet organisme hôte. Il est également important de pouvoir modifier le métabolisme de la cellule sans en altérer son fonctionnement.

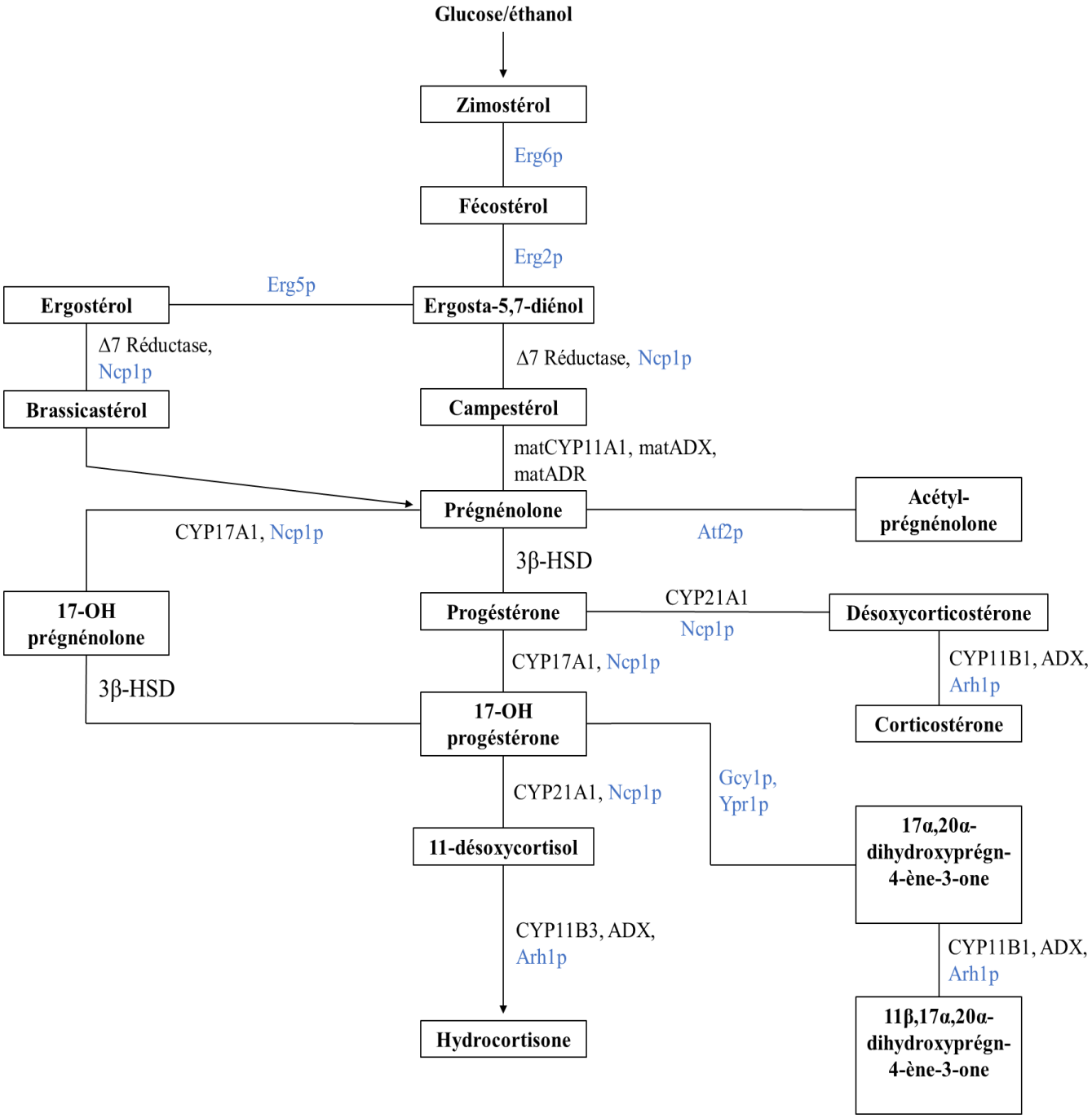
En 1973, Cohen et ses collaborateurs ont réussi à mettre en place la première ingénierie métabolique dans la bactérie *Escherichia coli* en construisant les premiers plasmides de clonage bactériens biologiquement fonctionnels (Cohen et al., 1973). Sur cette première base, de nombreux succès d'ingénierie métabolique de protéines recombinantes ont vu le jour. En particulier, la bioproduction d'insuline humaine par une *E. coli* recombinante a été réalisée via le clonage séparé des gènes codant les chaînes A et B de l'insuline humaine dans des plasmides. Les gènes ont ensuite été fusionnés à l'extrémité C-terminale du gène codant la β-galactosidase bactérienne permettant une transcription et une traduction efficaces de la protéine recombinante



<u>Figure 2</u> : Cycle de la biologie synthétique.

(Goeddel et al., 1979). Cette réussite a constitué une avancée majeure dans le domaine de l'ingénierie métabolique. Les progrès du génie génétique ont ensuite permis l'amélioration des souches via la mutagenèse et le criblage permettant l'introduction de modifications génétiques ciblées dans le génome de l'organisme hôte. En parallèle, le développement des techniques analytiques a permis d'observer au niveau cellulaire les conséquences des modifications génétiques apportées. En particulier, dans les années 90, le développement du couplage de la chromatographie en phase gazeuse et de la spectrométrie de masse (GC-MS de l'anglais « Gas Chromatography-Mass Spectrometry ») a permis d'accélérer l'analyse d'échantillons. La GC-MS est composée d'un chromatographe en phase gazeuse utilisant une colonne capillaire dont les dimensions et les propriétés de la phase dépendent des propriétés chimiques des molécules présentes dans l'échantillon. Les molécules disposant de temps de rétention différents vont être éluées différentiellement de la colonne, permettant ainsi au spectromètre de masse d'ioniser séparément les molécules en sortie de colonne et de détecter les fragments obtenus en fonction de leur rapport masse/charge (m/z). La visualisation et l'analyse des résultats sont permises par le passage des fragments obtenus à travers un détecteur Ultra-Violets (UV) ou un détecteur de masse. La chromatographie en phase liquide couplée à la spectrométrie de masse (LC-MS de l'anglais « Liquid Chromatography-Mass Spectrometry ») utilise le même principe avec une méthode de séparation basée sur l'utilisation d'une phase mobile liquide à travers une phase stationnaire. Le développement de ces techniques analytiques a permis d'améliorer la compréhension des modifications structurales apportées aux molécules. De façon générale, l'ingénierie métabolique implique une amélioration continue des propriétés cellulaires grâce à plusieurs cycles de la biologie synthétique. C'est à ce titre qu'une analyse métabolique des effets induits par les modifications génétiques est nécessaire. Suite à cette analyse, de nouvelles modifications génétiques peuvent être apportées afin d'améliorer les rendements de synthèse et/ou le fonctionnement de la cellule (**Figure 2**). L'ingénierie métabolique permet donc le développement de systèmes biologiques n'existant pas dans la nature via la reconstruction ou la modification de systèmes biologiques existants et permet de diriger la synthèse de molécules complexes avec une grande précision en matière de stéréo- et régio-sélectivité.

En parallèle des nombreuses réussites d'ingénierie métabolique en bactérie, notamment *E. coli*, et afin de s'affranchir des problématiques liées à l'utilisation de cet organisme procaryote, les chercheurs se sont tournés vers l'utilisation d'autres types de cellules hôtes. En particulier, la levure *Saccharomyces cerevisiae* est utilisée comme organisme modèle depuis longtemps et s'est rapidement révélée de haut intérêt dans la mesure où elle dispose d'une



<u>Figure 3</u> : Reconstitution de la voie de biosynthèse de l'hydrocortisone chez *S. cerevisiae*.

Ncp1p : NADPH P450 réductase ; ADX : adrénodoxine ; ADR : adrénodoxine réductase ; Arh1p : homologue lié à l'adrénodoxine réductase ; Atf2p : alcool O-acétyltransférase (acétylprégnénolone acétyltransférase); CYP11A1 : clivage de la chaîne latérale P450 ; CYP17A1 : 17α-stéroïde hydroxylase ; CYP21A1 : 21-stéroïde hydroxylase ; CYP11B1 : 11β-stéroïde hydroxylase ; Erg2p : stérol C8-C7 isomérase ; Erg6p : S-adénosyl méthionine Δ-24-stérol-C-méthyltransférase ; Erg5p : Δ22(23)-stérol désaturase ; 3β-HSD : 3 β-hydroxy stéroïde déshydrogénase ; Gcy1p et Ypr1p : aldo-céto réductases. Les gènes représentés en bleu sont endogènes à *Saccharomyces cerevisiae* et les gènes représentés en noir sont issus de mammifères.

machinerie capable de réaliser des modifications post-traductionnelles et d'un réseau de membranes internes permettant l'ancrage, quand cela est nécessaire, de certaines protéines hétérologues. La levure présente un génome relativement petit et la capacité d'exister sous forme haploïde ou diploïde, ce qui a facilité l'identification et l'édition de ses gènes. Ainsi, en 1996, son génome a été publié comme le premier génome d'un organisme eucaryote entièrement séquencé (Goffeau et al., 1996). Il est composé de 16 chromosomes avec un total de 12,5 Mpbs (paire de bases). La levure dispose de toutes les caractéristiques d'une cellule eucaryote, d'une physiologie bien caractérisée et elle est facilement cultivable et robuste en plus d'être un organisme Generally Recognized As Safe (GRAS). Le concept d'organisme GRAS a été fondé en 1958 par la Food and Drug Administration (FDA) afin de réguler l'utilisation de composés considérés comme sans danger dans les aliments, les médicaments et les cosmétiques. Fort de ces connaissances, de nombreuses molécules d'intérêt ont commencé à être synthétisées par ingénierie métabolique de la levure. En particulier, une des premières réussites majeures de bioproduction par la levure d'une molécule d'intérêt pharmaceutique à partir d'une simple source de carbone a été celle de l'hydrocortisone (Szczebara et al., 2003). L'hydrocortisone est un intermédiaire important dans la synthèse des médicaments stéroïdiens dont la synthèse chimique est difficile et coûteuse. La souche recombinante de *S. cerevisiae* mise en place comporte 13 gènes hétérologues, dont 8 gènes de mammifères (**Figure 3**). Cette première réussite a démontré la faisabilité du transfert génétique et fonctionnel d'une voie complète de biosynthèse complexe dans un micro-organisme conduisant à la bioproduction du métabolite ciblé. Depuis, de nombreux travaux visent à produire des molécules animales ou végétales d'intérêt via l'utilisation de micro-organismes hôtes. Des exemples de bioproduction de molécules végétales par ingénierie métabolique en levure seront détaillés dans le paragraphe 2.b de ce chapitre.

Toutefois, l'utilisation de micro-organismes hôtes moins classiques a aussi émergé ces dernières années, c'est le cas, notamment, de *Pseudomonas putida*. En effet, cette bactérie pathogène pour l'homme produit naturellement des polyhydroxyalcanoates à chaîne moyenne, polymères pouvant être utilisés notamment pour la bioproduction de plastiques biodégradables et de dispositifs médicaux. L'ingénierie métabolique de cet organisme est donc prometteuse pour la bioproduction de polymères à valeur ajoutée (Salvachúa et al., 2019). De plus, *P. putida* dispose d'un métabolisme natif conçu pour générer d'importantes quantités de nicotinamide adénine dinucléotide phosphate (NAD(P)H), cofacteur nécessaire aux enzymes oxydo-réductrices hétérologues notamment présentes dans les voies de biosynthèse végétales.

Finalement, cette bactérie est, non seulement capable d'héberger des voies de biosynthèse hétérologues, mais aussi des voies de synthèse artificielles, ouvrant ainsi la voie à la bioproduction et à l'étude de molécules non-naturelles (Nikel and de Lorenzo, 2018).

Les micro-organismes ne sont pas les seules cellules étudiées pour leur capacité de bioproduction de protéines recombinantes ou de molécules d'intérêt. En effet, des cultures cellulaires initiées artificiellement sont aujourd'hui très couramment utilisées. Citons, par exemple, la culture *in vitro* de cellules végétales dédifférenciées/indifférenciées qui peut, dans certains cas, permettre d'augmenter les rendements de production de MS d'intérêt de la plante dont les cellules sont originaires. Ainsi, la production maximale de crocine, un caroténoïde synthétisé par le safran disposant de propriétés antioxydantes, anticarcinogéniques et antidépressives a été obtenue via une culture en deux étapes *in vitro* de cals de safran (amas de cellules dédifférenciées) (Chen et al., 2003). La culture *in vitro* en deux étapes a permis, dans un premier temps, d'augmenter la biomasse produite, puis, dans un second temps, de produire la crocine à partir de la biomasse formée. Des lignées cellulaires de mammifères sont également largement utilisées. En effet, les anticorps monoclonaux utilisés dans le diagnostic et le traitement de maladies humaines telles que les maladies auto-immunes (Hall, 1989), le cancer (Hale et al., 1988) ou encore les infections virales (Tempest et al., 1991) produits par *E. coli* sont généralement mal assemblés, non glycosylés et mal repliés. L'utilisation de cellules de mammifères CHO (cellules ovariennes de hamster chinois), issues d'une lignée cellulaire couramment utilisée pour la bioproduction de protéines thérapeutiques, ou COS, issues d'une lignée cellulaire de type fibroblaste dérivées de tissu rénal de singe, permet la bioproduction d'anticorps monoclonaux fonctionnels, biologiquement actifs et intacts (Trill et al., 1995). Récemment, les cellules CHO ont été utilisées pour la bioproduction d'anticorps permettant l'assemblage de conjugués anticorps-médicament (ADC) utilisés pour le traitement de certains cancers par le ciblage direct des cellules tumorales ou de leur environnement proche afin de réduire la résistance aux médicaments (Procopio-Melino et al., 2022). L'emploi de ces cellules permet ainsi la bioproduction d'anticorps de très haute qualité selon un processus de fabrication simplifié.

En conclusion, de nombreuses molécules ou protéines d'intérêt peuvent donc être produites par ingénierie métabolique via l'utilisation de diverses cellules hôtes. En particulier, de nombreux produits pharmaceutiques tels que des hormones, des vaccins ou encore des cytokines biosynthétisés par des cellules de levure ont été approuvés par la FDA ces trois dernières décennies (**Figure 4**).

Biopharmaceutical	Expression system	First product	Target
Hormones			
Insulin (rh)	*S. cerevisiae*	1991 (Novolin 1991–2010)	Diabetes mellitus
Glucagon (rh)	*S. cerevisiae*	1999 (Glucagen)	Hypoglycemia
Somatropin (rhGH)	*S. cerevisiae*	2006 (Valtropin 2006–2012)	Growth failure
Insulin glargine (r)	*P. pastoris*	2018 (Semglee)	Diabetes mellitus
Vaccines			
HBsAg (r)	*S. cerevisiae*	1986 (Recombivax)	Hepatitis B
Combination vaccine (r)	*S. cerevisiae*	1996 (Tritanrix-hepB 1996–2014)	Multiple
HPV capsid proteins (r)	*S. cerevisiae*	2006 (Gardasil)	Human papillomavirus (HPX, 4 types)
VLP (partial surface proteins from *P. falciparum* and HBV) (r)	*S. cerevisiae*	2015 (Mosquirix)	*P. falciparum*-caused Malaria; Hepatitis B
Combination vaccine (r)	*H. polymorpha*	2013 (Hexacima)	Multiple
HBsAg (r)	*H. polymorpha*	2017 (HEPLISAV-B)	Hepatitis B
Blood-related			
Anticoagulants			
Hirudin (desirudin) (r)	*S. cerevisiae*	1997 (Revasc 1997–2014)	Venous thrombosis, thrombocytopenia
Clotting factors			
Plasma kallikrein inhibitor (ecallantide) (rh)	*P. pastoris*	2009 (Kalbitor)	Hereditary angioedema
Factor XIII A-subunit (catridecog) (rh)	*S. cerevisiae*	2012 (Novothirteen)	Congenital factor XIII A-subunit deficiency
Proteolytics			
Truncated plasmin (ocriplasmin) (rh)	*P. pastoris*	2012 (Jetrea)	Vitreomacular adhesion/traction
Cytokines			
Colony-stimulating factors GM-CSF (sargamostim) (rh)	*S. cerevisiae*	1991 (Leukine)	Neutropenia
Enzymes			
Urate oxidase (rasburicase) (r)	*S. cerevisiae*	2001 (Fasturtec)	Hyperuricemia

r, recombinant; rh, recombinant human.

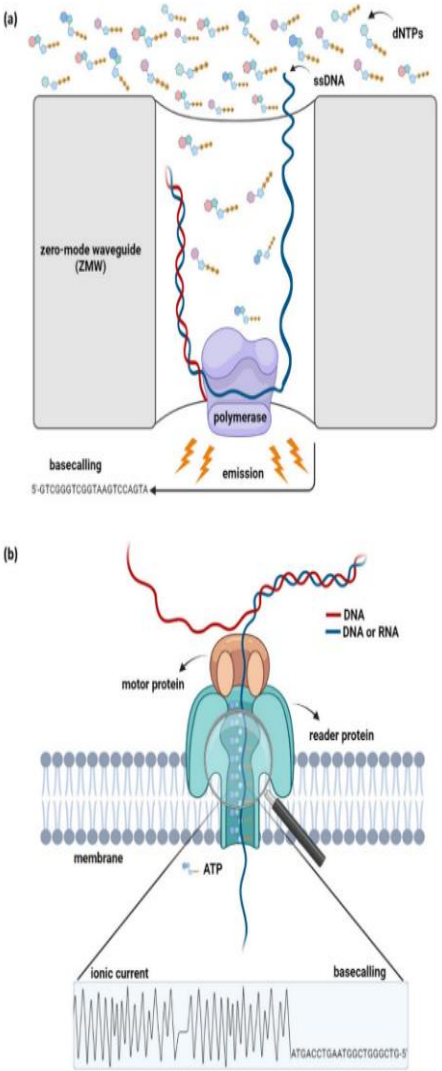
<u>Figure 4</u> : Produits biopharmaceutiques approuvés par la FDA produits dans des cellules de levure. (Kulagina et al., 2021a)

Dans ce mémoire, je me focaliserai sur la bioproduction de molécules végétales d'intérêt via l'utilisation de l'organisme hôte *Saccharomyces cerevisiae*.

I.2 L'ingénierie métabolique pour la production de molécules biosynthétiques végétales

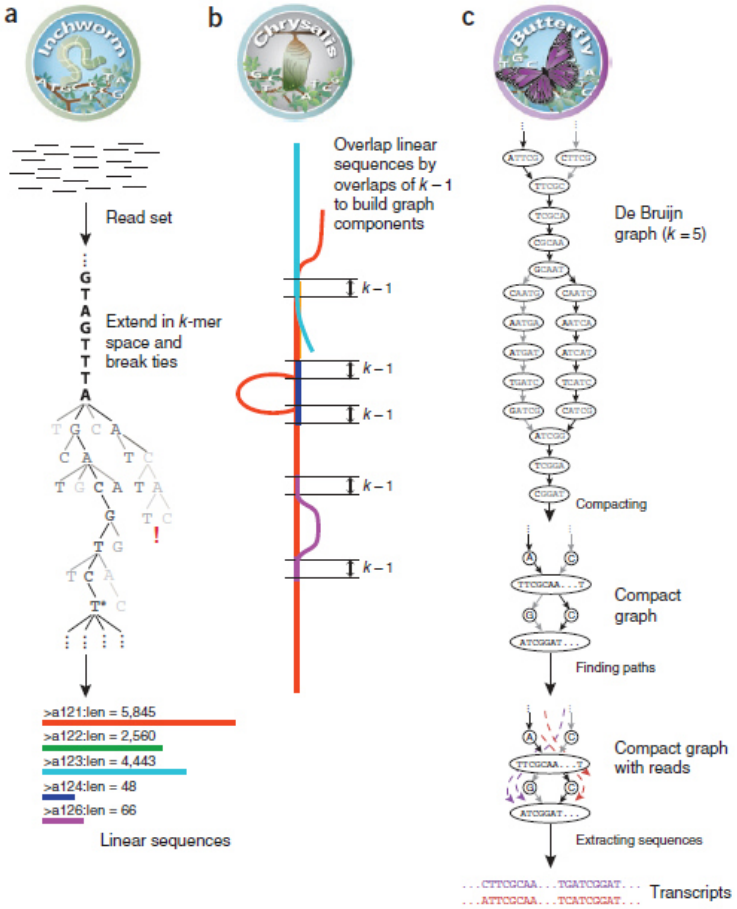
Les métabolismes primaire et spécialisé végétaux sont responsables de l'assemblage d'une vaste collection de molécules chimiquement diverses. Les molécules du métabolisme primaire sont produites par la glycolyse, le cycle de Krebs ou encore la voie du shikimate et sont souvent utilisées comme précurseurs pour la synthèse de nombreux MS. Les MS sont très divers en fonction de l'espèce, de l'organe, du tissu cellulaire et du stade de développement de la plante (Wink, 2010). De plus, leur localisation au sein de la plante est généralement corrélée à une fonction spécifique. Par exemple, les alcaloïdes indolo-monoterpéniques (AIM) synthétisés par *C. roseus* ne sont pas répartis de façon homogène dans la plante. En effet, leur concentration est plus élevée dans les feuilles en développement et diminue lorsque celles-ci sont matures, suggérant leur implication dans le développement de la plante. La fonction peut également être corrélée à une adaptation à des modifications soudaines de l'environnement. Ainsi, lors d'une attaque herbivore, *C. roseus* modifie son profil complet d'AIM et accumule ces molécules dans les cires et l'épiderme, provoquant la mort des chenilles ayant consommé les feuilles de la plante (Dugé de Bernonville et al., 2017 ; Lemos Cruz, 2021a).

Comme mentionné dans le préambule, certains MS produits par les végétaux possèdent des propriétés intéressantes du point de vue pharmaceutique, cosmétique et/ou agro-alimentaire. Les plantes qui les produisent ont donc une importance commerciale et l'extraction de ces composés à partir de matériel végétal peut être coûteuse et compliquée à mettre en place à échelle industrielle. De plus, certains végétaux ne produisent les molécules d'intérêt qu'à un certain stade de développement ou qu'à certaines saisons, ce qui peut avoir de forts impacts sur l'approvisionnement de celles-ci. De même, les plantes ne produisent pas les MS de façon continue mais plutôt en faible quantité. Il est donc nécessaire de disposer d'une biomasse très importante pour extraire des quantités de MS d'intérêt cohérentes avec les demandes du marché, amplifiant ainsi la pression déjà exercée sur les plantes productrices. C'est pourquoi, au cours de ces dernières années, de nombreux travaux ont été menés sur la mise au point et l'utilisation de procédés biotechnologiques, en particulier sur l'ingénierie métabolique des voies de



<u>Figure 5</u> : Représentation schématique du fonctionnement de deux technologies de séquençage de troisième génération. (Athanasopoulou et al., 2022)

a) Séquençage PacBio basée sur l'observation de la lumière émise par l'ADN polymérase
b) Séquençage des nanopores basé sur la détection des altérations du courant électrique lors du passage ADN/ARN



<u>Figure 6</u> : Représentation schématique du fonctionnement de l'outil Trinity. (Grabherr et al., 2011)

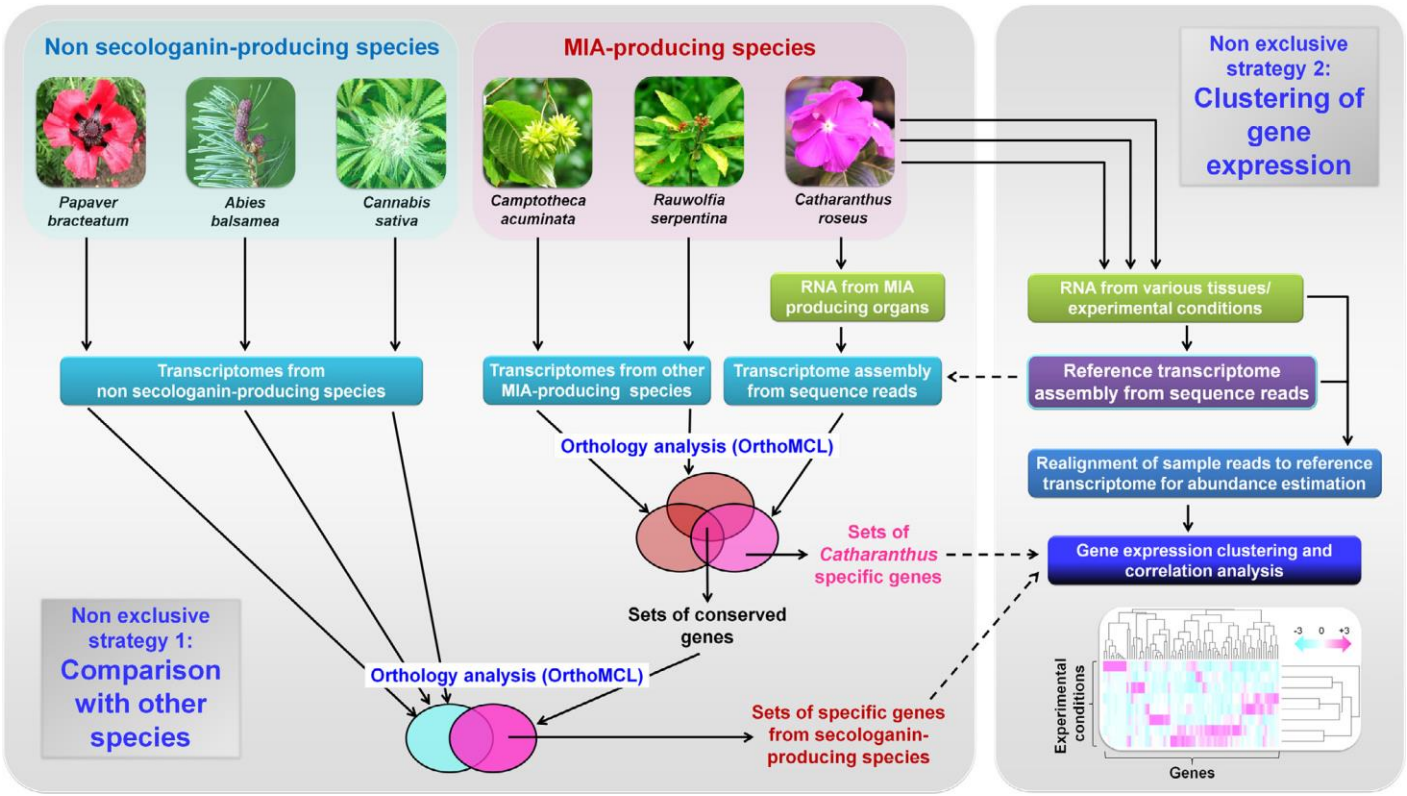
(a) Inchworm assemble l'ensemble de données de lecture (courtes lignes noires, en haut) en recherchant des chemins dans un graphe k-mer (au milieu), résultant en une collection de contigs linéaires (lignes de couleur, en bas), avec chaque k-mer présent une seule fois dans les contigs.
(b) Chrysalis regroupe les contigs (lignes colorées) s'ils partagent au moins un k –1-mer et si les lectures couvrent la jonction entre les contigs, puis elle construit des graphes de Bruijn individuels à partir de chaque pool.
(c) Butterfly prend chaque graphe de Bruijn de Chrysalis (en haut), et coupe les arêtes parasites et compacte les chemins linéaires (au milieu). Il réconcilie ensuite le graphique avec les lectures (flèches colorées en pointillés, en bas) et les paires (non illustrées), et produit une séquence linéaire pour chaque forme d'épissage et/ou transcrit paralogue représenté dans le graphique (en bas, séquences colorées).

voies de biosynthèse végétales en levure, afin d'apporter des alternatives de production des MS et participer à l'approvisionnement des entreprises pharmaceutiques, agro-alimentaires et cosmétiques utilisant ces composés comme matière première. Toutefois, le développement de ces approches d'ingénierie métabolique nécessite comme prérequis une élucidation complète des voies de biosynthèse végétales.

a) Elucidation des voies de biosynthèse végétales

1. Approches bio-informatiques pour l'identification des gènes

Dans l'optique de produire des molécules végétales via une cellule hôte par ingénierie métabolique, la première étape est d'identifier les gènes impliqués dans leur biosynthèse *in planta*. Si la méthode de séquençage de Sanger, nécessitant la construction de plasmides intégrant l'ADN à séquencer, avait fait ses preuves pour le séquençage de l'ADN à la fin des années 70, elle était limitante dans le temps puisqu'il fallait plusieurs années pour réussir à isoler et séquencer chaque gène des voies de biosynthèse. Toutefois, l'évolution des méthodes de séquençage, notamment les NGS (Next Generation Sequencing) ont permis une avancée majeure dans l'identification de gènes puisque de plus en plus de bases de données contenant des collections de séquences d'ARNm (Acide RiboNucléique messager) produites par séquençage d'ARN (RNA-seq) à partir de différents organes ou tissus d'une plante sont mises à disposition. Les technologies Illumina et SOLiD par exemple, permettent le séquençage de l'ARN. La technologie Illumina se base sur le séquençage de fragments de 100 à 500 pbs amplifiés par PCR grâce à un émetteur fluorescent. L'ensemble des ARN issus de l'expression du génome dans une condition ou un organe spécifique constitue le transcriptome. L'émergence des TGS (Third Generation Sequencing) a permis l'ouverture de nouvelles portes grâce à l'obtention de transcrits pleine longueur, simplifiant ainsi la lecture et l'assemblage du transcriptome. Il existe deux technologies de séquençage dites de troisième génération. Le séquençage PacBio est basé sur l'observation de la lumière émise par l'ADN polymérase en fonction du nucléotide synthétisé (**Figure 5.a**), tandis que le séquençage par les nanopores repose sur la détection des altérations du courant électrique lors du passage de l'ADN ou l'ARN dans le nanopore (**Figure 5.b**) (Athanasopoulou et al., 2021). En effet, les nanopores (trous

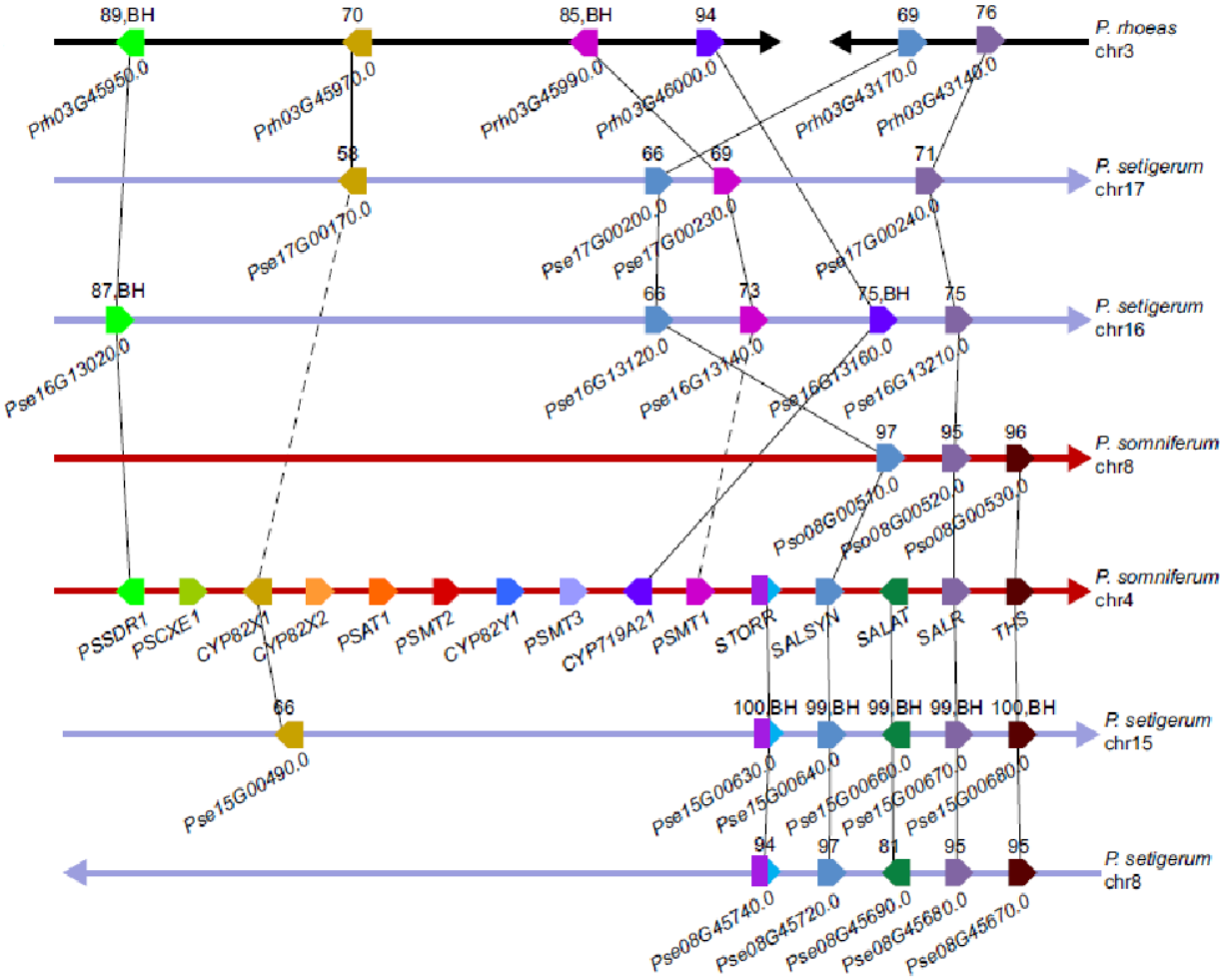


<u>Figure 7</u> : Représentation schématique de l'utilisation de données transcriptomiques à haut débit pour la découverte de gènes biosynthétiques MIA. (Dugé de Bernonville et al., 2015)

Deux stratégies non exclusives peuvent être utilisées : la comparaison d'orthologie avec d'autres espèces (panneau de gauche) et le regroupement de l'expression génique (panneau de droite). La première approche nécessite des données transcriptomiques d'espèces candidates avec des métabolismes alcaloïdes spécifiques (par exemple, des espèces productrices de sécologanine (11) (encart violet) et non productrices de sécologanine (11) (encart bleu)). Ces données sont disponibles dans les projets MPGR et PhytoMetaSyn. Une première comparaison d'orthologie à l'aide de l'algorithme OrthoMCL pourrait permettre d'identifier des gènes conservés codant des enzymes impliquées dans la biosynthèse de précurseurs communs chez les espèces productrices de MIA (*Camptotheca acuminata*, *Rauwolfia serpentina*, *Catharanthus roseus*) mais aussi des enzymes spécifiques dédiées à la synthèse d'alcaloïdes particuliers, tels que les gènes spécifiques de *C. roseus*. Une deuxième analyse d'orthologie peut en outre être effectuée sur l'ensemble de gènes conservés trouvés dans les espèces productrices de MIA pour isoler les candidats de la voie sécoiridoïde en comparant avec les données transcriptomiques d'enzymes productrices de non sécologanine (11). La deuxième stratégie s'appuie sur des analyses de transcriptome à partir de diverses conditions expérimentales pour construire un transcriptome de référence qui sera utilisé pour estimer l'abondance de chaque transcrit dans chaque condition afin de regrouper les gènes en fonction de leurs niveaux d'expression. Plus les conditions expérimentales sont disponibles, plus l'analyse de clustering sera efficace pour regrouper les gènes co-régulés. Notez que ce transcriptome de référence pourrait également servir à l'analyse d'orthologie décrite dans le panneau de gauche, ainsi qu'à vérifier les niveaux d'expression des gènes candidats résultant de ce type d'approche.

d'une dizaine de nanomètres) sont immergés dans un liquide conducteur. Le courant électrique variant en fonction de la taille, de la forme et de la concentration en particules présentes dans le nanopore, l'analyse du flux électrique permet de déterminer la succession de nucléotides passants lors de la translocation de l'ADN dans le nanopore. De plus, de nouvelles approches bioinformatiques permettent désormais de faciliter l'assemblage des transcrits obtenus par RNA-seq (Chang et al., 2015). Si certaines méthodes reposent sur l'alignement des lectures sur un génome de référence pour reconstituer la transcription, de nouveaux outils comme Trinity reposent sur l'assemblage *de novo* de transcrits longs. L'outil Trinity, permet, de plus, de détecter des isoformes et de gérer les lectures appariées sans nécessité d'alignement avec un génome de référence (Grabherr et al., 2011). Dans cet outil, la première étape (Inchworm) consiste en l'assemblage des transcrits présents dans l'échantillon afin de mettre en évidence les fragments chevauchants (**Figure 6.a**). La seconde étape (Chrysalis) consiste en la construction des graphes de De Bruijn (graphiques orientés) à partir des fragments chevauchant permettant de regrouper ensemble les mêmes transcrits épissés alternativement (**Figure 6.b**). La dernière étape (Butterfly), consiste en l'assemblage des graphes de De Bruijn obtenus afin de reconstruire entièrement les transcrits plausibles pour chaque isoforme et/ou paralogue (gène dérivé disposant d'une fonction différente) (**Figure 6.c**).

Une fois le transcriptome assemblé, l'identification de gènes candidats peut être effectuée par plusieurs approches incluant, par exemple, la recherche d'homologie, c'est-à-dire par comparaison avec des gènes préalablement identifiés chez des espèces proches mais également par comparaison des profils d'expression de l'ensemble des gènes avec des gènes préalablement identifiés de ces voies de biosynthèse (**Figure 7**) (Dugé de Bernonville et al., 2015). En effet, les gènes impliqués dans un même processus physiologique ou une même voie de biosynthèse, par exemple, présentent généralement des profils d'expression similaires. Identifier des gènes candidats par la recherche de cette similarité d'expression via l'analyse de réseau de co-expression a ainsi permis de prédire de nombreux gènes candidats et de les caractériser, notamment chez *C. roseus* (Payne et al., 2017 ; Caputi et al., 2018). En outre, la recherche de gènes différentiellement exprimés avec une fonction d'intérêt peut aussi permettre de déterminer si certains gènes sont plus exprimés dans une condition que dans une autre en comparant les transcriptomes d'un organisme placé sous deux conditions différentes. La différence de condition doit être exploitée si une des deux permet une accumulation plus élevée du MS d'intérêt. Cette technique a permis, par exemple, de mettre en évidence une isoforme de la tabersonine/vincadifformine 16-*O*-méthyltransférase Vm16OMT impliquée dans la voie de



<u>Figure 8</u> : Histoire évolutive du groupe de gènes BIA. (Yang et al., 2021)

Les relations synténiques au niveau du groupe de gènes BIA chez trois espèces de Papaver. Les lignes pleines indiquent les relations synténiques détectées et les lignes en pointillés indiquent les meilleurs résultats d'alignement de séquences. Les différentes couleurs indiquent différents gènes. Le nom ou l'ID du gène est étiqueté pour chaque gène. Les nombres au-dessus des gènes représentent les niveaux d'identité de séquence. Les sens des flèches indiquent le chromosome de 5′ à 3′.

biosynthèse des AIM produit par *Vinca minor*, par comparaison entre le transcriptome de plantes exposées ou non à la lumière (Stander et al., 2020). Une analyse du réseau de co-expression entre plusieurs échantillons impliqués dans le même processus biologique peut aussi être intéressante à réaliser puisqu'elle met en évidence les gènes contrôlés par un même programme de régulation transcriptionnel, fonctionnellement liés ou membres du même réseau de régulation génétique (Ruprecht et al., 2017).

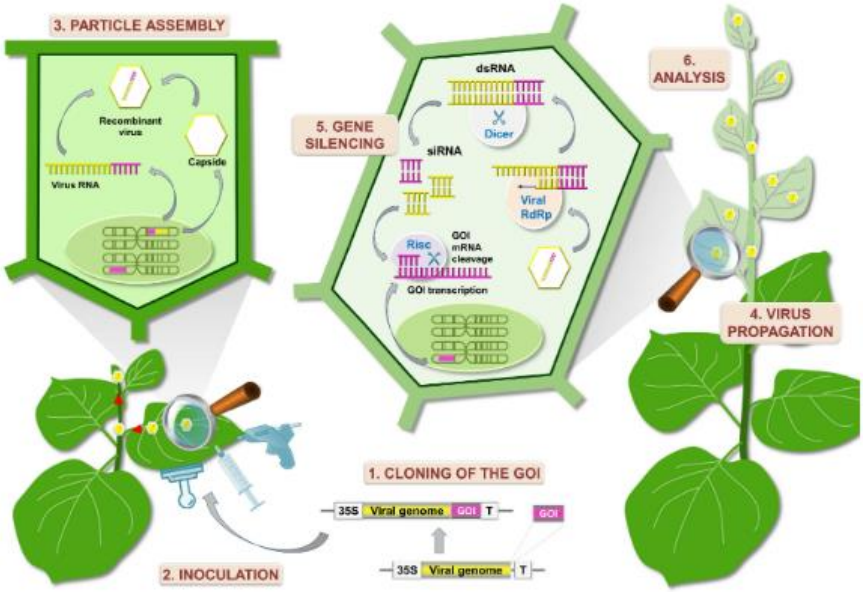
De plus, une analyse génomique, c'est-à-dire de l'organisation des gènes dans le génome, et notamment de conservation de la synténie (conservation de l'ordre des gènes entre espèces apparentées) et de leur fonction, peut également conduire à l'identification de gènes candidats. Dans ce contexte, la mise à disposition de génomes entièrement ou partiellement séquencés a permis une avancée majeure dans l'identification de gènes candidats. Cette identification est facilitée lorsque l'expression des gènes est liée à une fonction précise. En effet, les gènes impliqués dans le métabolisme spécialisé végétal, par exemple, peuvent être organisés en groupes, ou clusters. L'analyse de ces groupes et de leur répétition dans les chromosomes peut être un moyen efficace d'identifier les gènes impliqués dans la synthèse de métabolites spécifiques. Par exemple, la voie de synthèse des alcaloïdes de benzylisoquinoléine (BIA), telles que la morphine et la noscapine, chez *Papaver somniferum* est une des voies de biosynthèse de MS la plus longue caractérisée avec au moins 10 étapes enzymatiques décrites. Une analyse minutieuse de trois espèces de Papaver basée sur la synténie, la phylogénie et des alignements de séquences des gènes impliqués dans cette voie a permis d'identifier, non seulement, des groupes de gènes, mais aussi une importante conservation de ces groupes entre espèces apparentées (**Figure 8**) (Yang et al., 2021). La prédiction des gènes candidats est un processus long et parfois complexe mais qui ne représente que la première étape de l'identification des gènes de voies de biosynthèse. En effet, une fois les gènes candidats prédits, l'activité des enzymes codées doit être déterminée par des approches fonctionnelles afin de s'assurer de la nature et de la spécificité de la réaction catalysée.

2. Validation fonctionnelle des gènes candidats

La validation fonctionnelle des gènes candidats identifiés emploie un éventail de techniques allant des approches biochimiques à celles de génétique inverse. Parmi ces techniques, l'expression hétérologue des gènes candidats dans un organisme hôte tel qu'*E. coli*

est employée afin de purifier les protéines recombinantes correspondantes pour les tester fonctionnellement *in vitro*. Si, historiquement, les essais d'activité enzymatique étaient réalisés sur les protéines recombinantes en condition *in vitro*, ils peuvent également être réalisés *in vivo*, c'est-à-dire directement dans la cellule hôte. Ainsi, Carqueijeiro et al. (2018) ont réalisé, dans leur travaux d'élucidation de la voie de biosynthèse et de décoration de l'AIM tabersonine, précurseur de l'anti-cancéreux vinblastine chez *C. roseus*, la validation fonctionnelle du gène candidat codant la tabersonine 19-*O*-acétyltransférase (TAT) capable d'acétyler les dérivés de 19-hydroxytabersonine des racines par transfert du gène codant dans *S. cerevisiae*. Pour ce faire, le criblage des gènes candidats a été réalisé par expression et tests fonctionnels en levure par apport de 19-hydroxytabersonine dans le milieu de culture. De plus, l'expression hétérologue du gène codant la TAT chez *E. coli* a ensuite permis d'évaluer les propriétés biochimiques de la protéine recombinante via des tests d'activité enzymatique *in vitro* montrant l'intérêt de la combinaison d'approches *in vivo* et *in vitro*.

Les approches de génétique inverse cherchent à comprendre la fonction d'un gène donné en observant les conséquences de la perte de sa fonction sur le métabolisme ou la physiologie de l'organisme initial. Ces techniques permettent de s'affranchir d'éventuels problèmes liés à l'utilisation d'un organisme procaryote et des difficultés inhérentes d'expression des gènes candidats et de mise en place de tests fonctionnels. Ces techniques utilisent l'invalidation génétique par inactivation totale du gène (knock-out) ou par une réduction au silence de l'expression du gène (knock-down). L'invalidation génétique peut être réalisée par modification génomique ou par des approches transitoires telle que la mise sous silence induite par des virus (VIGS : Virus-Induced Gene Silencing). Le knock-out par modification génomique à l'aide du système CRISPR/Cas9 a, notamment, permis de vérifier l'implication dans la régulation de la hauteur du colza des deux gènes codant les monooxygénases MAX1 (Zheng et al., 2020). La technique VIGS, appliquée aux espèces végétales, exploite un mécanisme de défense antiviral médié par l'ARN (Lu et al., 2003). En effet, lorsqu'un virus végétal à génome ARN infecte une cellule, une défense ciblée contre le génome viral est activée suite à la détection d'ARN double brin formé lors de la phase réplicationnelle du virus. Lors de ce mécanisme de défense, l'ARN double brin viral est clivé en ARN interférent (ARNi) double brin par les complexes DICER, des ribonucléases de classe 3, puis ces ARNi sont intégrés au complexe RISC (RNA-Induced Silencing Complex) assurant leur conversion en ARNi simple brin (ARNi sb) via son activité hélicase. Le complexe RISC/ARNi sb va alors scanner les ARN présents dans la cellule végétale et en cas d'identité ARNi sb/ARN, l'hybridation en résultant



<u>Figure 9</u> : Représentation schématique de la technique de répression d'expression génique induite par un virus (VIGS). (Courdavault et al., 2020)

Un fragment du gène d'intérêt (GOI) est cloné dans une construction contenant la séquence du génome viral [1], puis inoculé à des plantes [2], où des particules virales sont produites [3]. Par le mouvement de cellule à cellule à l'aide de plasmodesmes et une progression systémique via le phloème [4], les virus infectent les feuilles nouvellement développées et se répliquent. Cela implique la formation d'ARN double brin qui est clivé par la plante pour libérer l'ARN simple brin. les ARN simple brin sont chargés dans le complexe multiprotéique RISC et leur brin antisens crible l'ARN dans les cellules. L'hybridation de ce complexe entraîne une dégradation spécifique de l'ARN (ou répression de la traduction) de l'ARN viral et des transcrits GOI endogènes provoquant son silençage.

conduira à la dégradation de la cible via l'activité RNAse du complexe RISC. La technique VIGS consiste à intégrer un fragment du gène végétal à réduire au silence (« silencer ») dans le génome viral puis à inoculer ces vecteurs ciblant le gène candidat à la plante. Le mécanisme de défense de la plante ciblera alors les ARN correspondant au gène candidat, résultant ainsi en sa dégradation dans la plante (**Figure 9**). Cette technique à l'aide du virus de la mosaïque du tabac (TMV) et du virus de la mosaïque de la tomate (ToMV) a initialement été développée chez *Nicotiana benthamiana* (Kumagai et al., 1995), puis adaptée à d'autres plantes avec différents virus. Notamment, la technique VIGS a été mise en place à l'aide du virus X de la pomme de terre (PVX) chez *Arabidopsis thaliana*, *Nicotiana tabacum* et la tomate (Angell and Baulcombe, 1999). La technique VIGS s'est progressivement imposée comme une technique de choix chez les plantes de par sa rapidité, sa facilité de mise en œuvre et son efficacité. Elle a notamment été développée pour la validation fonctionnelle des gènes candidats chez *C. roseus* via l'utilisation de dérivés du tobacco rattle virus (TRV) dont l'inoculation chez la plante est réalisée à l'aide d'*Agrobacterium tumefaciens*. Les premiers tests réalisés chez *C. roseus* utilisaient trois procédures différentes pour l'inoculation : la tige pouvait être pincée avec des pinces trempées dans la solution bactérienne (Liscombe et O'Connor, 2011), percée pour y insérer le mélange bactérien (De Luca et al., 2012) ou la plantule entièrement infiltrée (Sung et al., 2014). Un protocole utilisant l'inoculation biolistique, c'est-à-dire, par bombardement de particules contenant l'ADN transformant dans les cellules, a ensuite été mis en place chez *C. roseus* (Carqueijeiro et al., 2015), puis appliqué chez Rauwolfia (Corbin et al., 2017) et enfin étendu à d'autres Apocynaceae (Lemos Cruz et al. 2020). La technique VIGS a ainsi permis de caractériser de nombreux gènes de la voie de biosynthèse des AIM de *C. roseus* dont, par exemple, l'alcool déshydrogénase Geissoschizine Synthase (GS) et le cytochrome P450 CYP71D1V1 ou Geissoschizine Oxidase (GO) dans certains réarrangements structurels de la strictosidine chez *C. roseus*, aboutissant à la formation des AIM de type Strychnos (Tatsis et al., 2017).

Pour compléter la caractérisation des enzymes candidates et la compréhension du flux de synthèse, il peut être intéressant de déterminer la localisation des enzymes *in planta*, plus spécifiquement à l'échelle subcellulaire. En plus de fournir des éléments essentiels à la compréhension de leur implication dans la synthèse des MS, cette connaissance de leur organisation peut fournir des éléments importants pour des approches d'ingénierie métabolique. Pour cela, la protéine d'intérêt peut être fusionnée à une protéine fluorescente de type Green Fluorescent Protein (GFP) permettant de visualiser au microscope sa localisation par

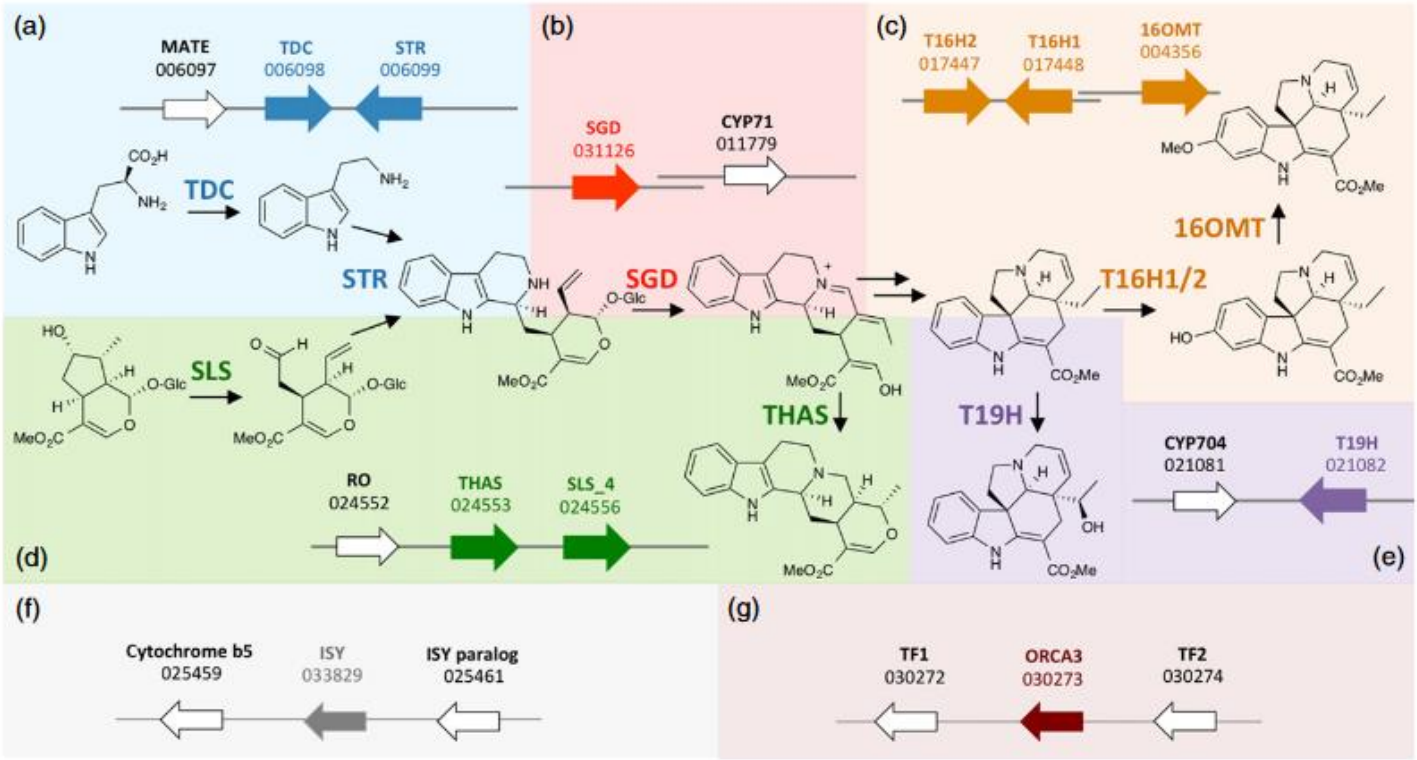
comparaison des profils de fluorescence observés avec ceux de marqueurs de différents compartiments subcellulaires (Foureau et al., 2016). Par exemple, la visualisation par microscopie à épifluorescence des protéines de fusion entre la strictosidine synthase (STR) et la GFP, ou la strictosidine-β-D-glucosidase (SGD) et la GFP respectivement chez *C. roseus* et *Rauvolfia serpentina*, a permis de mettre en évidence la localisation de STR à la vacuole et de la SGD au noyau, suggérant une importance de la séquestration subcellulaire de ces deux enzymes dans le contrôle du flux métabolique (Guirimand et al., 2010). De telles données de localisation des enzymes *in planta* sont donc essentielles à connaître afin de reproduire, optimiser, voire modifier leur organisation dans des cellules hétérologues lors d'approches de reconstitution de voies métaboliques. L'intérêt de la connaissance de l'organisation de ces voies de biosynthèse sera décrit dans la partie I.3.c.5 "Reconfiguration spatiale" de l'Introduction de ce manuscrit.

Une fois les gènes candidats caractérisés, leur clonage en levure peut être initié dans l'optique d'une bioproduction. La connaissance des enzymes impliquées dans les voies de biosynthèse végétales est importante afin d'appréhender les éventuels goulots d'étranglement pouvant réduire les rendements de bioproduction.

3. Réactions enzymatiques des voies de biosynthèse végétales

Les MS sont issus de voies de biosynthèse disposant bien souvent de similarités dans l'enchaînement des réactions enzymatiques qui les composent. En effet, ces voies reposent sur des modifications analogues des intermédiaires biosynthétiques, telles que la réduction, l'oxydation, la méthylation, l'hydroxylation, ou encore l'acylation. Comme, par exemple, la voie de synthèse de la noscapine comprenant au minimum deux cytochromes P450, six méthyltransférases et deux hydroxylases (Li et al., 2018).

A ce titre, des enzymes et des modules réactionnels similaires sont retrouvés dans les différentes voies de biosynthèse comme la succession d'une alcool déshydrogénase, d'un cytochrome P450 et d'une *O*-méthyltransférase ou une acyltransférase dans les voies de biosynthèse végétales. Ceci explique, notamment, que les gènes associés à la synthèse de MS soient parfois regroupés en cluster dans le génome des plantes (Medema and Osbourn, 2016). Par exemple, les gènes codant le cytochrome P450 CYP76 et celui codant la L-DOPA 4,5-



<u>Figure 10</u> : Regroupement physique (cluster) au sein du génome de *C. roseus* des gènes connus de la voie MIA. (Kellner et al., 2015)

Gènes potentiels non caractérisés dans la biosynthèse du MIA (représentés par des flèches blanches) qui se regroupent avec la voie MIA connue et les gènes régulateurs (représentés par des flèches pleines). Des exemples de duplication de gène sont également présentés.

(a) Scaffold WGS 3045674 (29 490 pb) contenant TDC (tryptophane décarboxylase), STR (strictosidine synthase) et un MATE non caractérisé (boîte bleue).
(b) Contigs BAC_SGD c2 et c11 contenant SGD (strictosidine b-glucosidase)et un P450 non caractérisé (boîte rouge).
(c) BAC_T16H contigs c3 et c1 contenant T16H1 (tabersonine 16-hydroxylase 1 (CYP71D12)), T16H2 (tabersonine 16-hydroxylase 2 (CYP71D351)), et 16OMT (16-hydroxytabersonine *O*-méthyltransférase (boîte orange).
(d) Scaffold WGS 3063455 (76 903 pb) contenant le paralogue SLS (secologanine synthase) (SLS_4), THAS (tétrahydroalstonine synthase) et RO (réticuline oxidase-like) non caractérisé (boîte verte).
(e) Scaffold WGS 2964965 (34 685 pb) contenant T19H (tabersonine/lochnéricine 19-hydroxylase (CYP71BJ1)) et un P450 non caractérisé (boîte violette).
(f) Scaffold WGS 3047130 (46 594 pb) contenant ISY (iridoid synthase), le paralogue ISY et un cytochrome b5 non caractérisé (boîte grise).
(g) Scaffold WGS 3061731 (40 593 pb) contenant le facteur de transcription ORCA3 et deux facteurs de transcription non caractérisés similaires (boîte marron).

dioxygénase catalysant les deux premières étapes de la biosynthèse du pigment bétalaïne chez la betterave à sucre sont espacés de moins de 50kb sur le chromosome 2 (Brockington et al., 2015). De même, plusieurs cas de regroupement au même locus génomique de gènes codant des étapes enzymatiques successives différentes ont été mis en évidence dans la voie de synthèse des AIM chez *C. roseus* (**Figure 10**) et dans la voie de synthèse des BIA chez *P. somniferum* (Winzer et al., 2012 ; Kellner et al., 2015). Comme indiqué précédemment, les mêmes familles d'enzymes sont généralement retrouvées dans ces modules réactionnels, incluant notamment, les alcool déshydrogénases, les cytochromes P450, les dioxygénases, les *N*- et *O*-méthyltransférases, les acyltransférases, les UDP-glucuronosyltransférases (UGT) et les estérases. Cette partie décrira les différentes familles d'enzymes généralement retrouvées dans les voies de biosynthèse de MS végétaux.

Réactions d'oxydoréduction et alcool déshydrogénases

L'alcool déshydrogénase (ADH) est une enzyme oxydoréductrice permettant la conversion d'alcools primaires en aldéhydes et d'alcools secondaires en cétones. Les réactions d'oxydation ou, principalement, de réduction catalysées par ce type d'enzyme sont respectivement couplées à une réduction du NAD^+ en NADH ou à une oxydation de NADH en NAD^+. Les ADH végétales ont évolué à partir des formaldéhydes déshydrogénases glutathion-dépendantes (ADH de classe III) (Danielsson et Jörnvall, 1992 ; Shafqat et al., 1996). Il existe différentes familles d'ADH chez les végétaux, notamment, les Short-chain Dehydrogenases/Reductases (SDR) qui catalysent des réactions réversibles et ne disposent pas de zinc dans leur site actif; les Medium-chain Dehydrogenases/Reductases (MDR) qui peuvent avoir, ou non, du zinc dans leur site actif et qui sont généralement actives sous forme de dimère ou de tétramère; et finalement, les alcool cinnamylique déshydrogénases (CAD) spécifiques des plantes qui catalysent l'étape finale de la synthèse des monolignols.

L'activité des ADH chez les végétaux a été observée pour la première fois dans les grains de pollen et les anthères du maïs (Schwartz, 1970). Ces enzymes sont essentielles pour la biosynthèse des lignines chez la plante et participent aux défenses végétales contre les stress biotiques et abiotiques (Umezawa, 2010). En effet, le rôle le plus connu des ADH chez les végétaux est leur implication dans l'adaptation aux conditions d'anaérobie. Précisément, en cas d'hypoxie, les plantes passent d'une glycolyse aérobie à une glycolyse anaérobie, ou

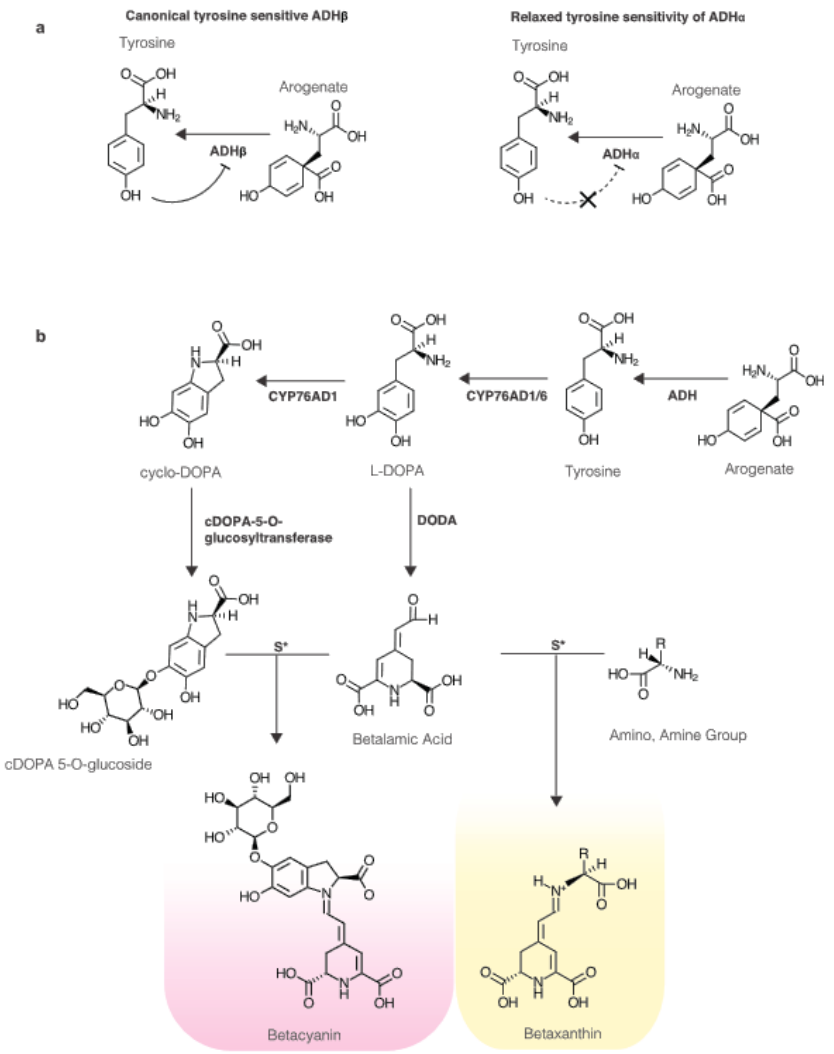


Figure 11 : Les voies de biosynthèse de la tyrosine et de la bétalaïne. (Timoneda et al., 2018)

(a) La boucle de rétroaction négative exercée par la tyrosine sur l'isoforme canonique ADHβ, par rapport à la sensibilité relâchée à la tyrosine de l'isoforme ADHα.

(b) La première étape illustrée est l'étape de décarboxylation de l'arogénate catalysée par l'arogénate déshydrogénase (ADH). Les étapes subséquentes illustrées de bétalaïne comprennent l'hydroxylation de la tyrosine 3' catalysée par les enzymes du cytochrome P450 (CYP76AD1/AD6) pour former la L-DOPA, la formation d'acide bétalamique à partir de la L-DOPA par l'action de la 4,5 DOPA dioxygénase (DODA), la conversion de la L-DOPA en cyclo-DOPA par le cytochrome P450 (CYP76AD1), glucosylation de la cyclo-DOPA via l'action de la cyclo-DOPA 5-O glucosyltransférase et condensation spontanée de l'acide bétalamique avec la cDOPA 5-O-glucoside pour former du violet bétacyanines, ou avec des groupes amino ou amine pour former des bétaxanthines jaunes.

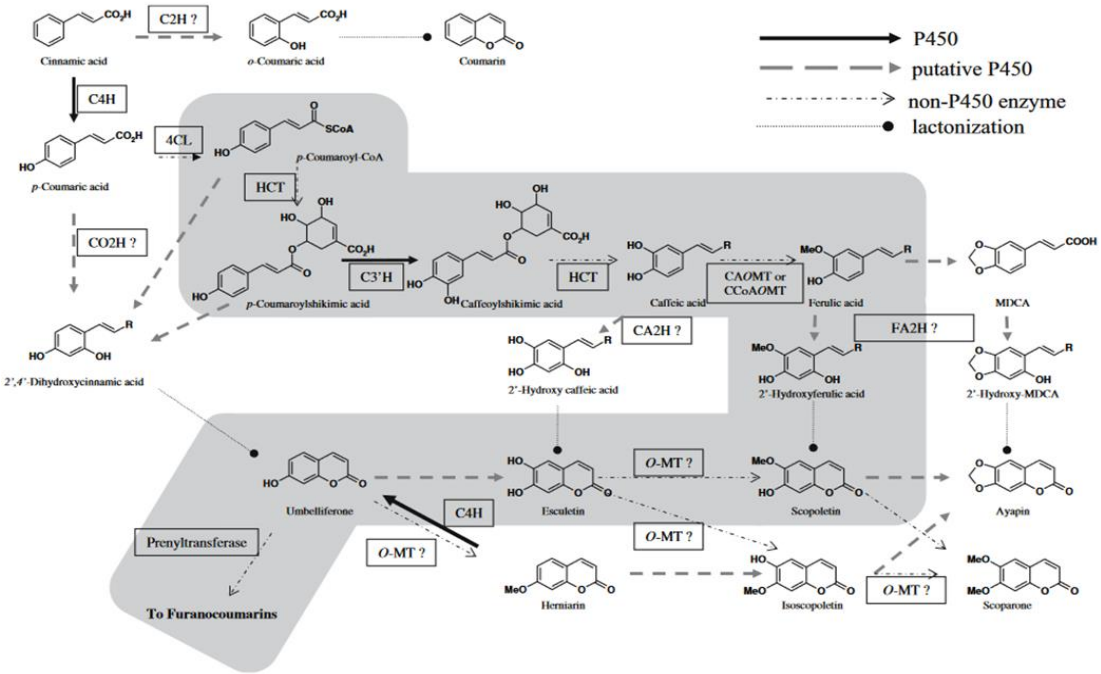


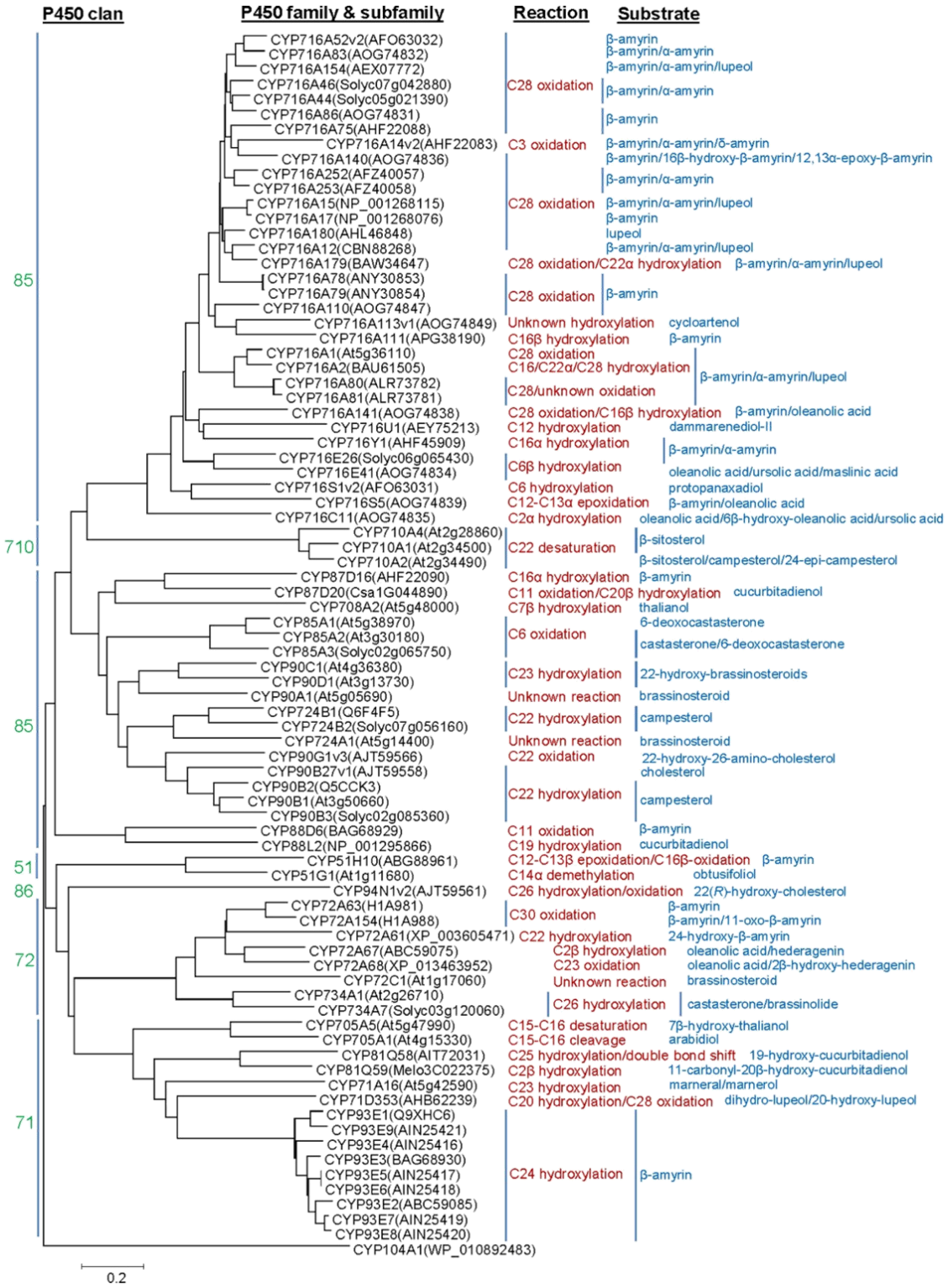
Figure 12 : Voie phénylpropanoïde menant aux coumarines. (Bourgaud et al., 2006)

Les parcours en gris ont été établis sans équivoque. Les composés glycosylés ne sont pas représentés pour plus de clarté. Les enzymes attribuées par un point d'interrogation sont hypothétiques. R = CO2H ou CO-SCoA; C2H : acide cinnamique 2-hydroxylase; C4H : acide cinnamique 4-hydroxylase; 4CL : 4-coumarate-CoA ligase; CO2H : acide 4-coumarique 2-hydroxylase ; HCT : hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl-transférase; C3'H : coumarate 3'-hydroxylase; CAOMT : acide caféique O-méthyltransférase; CCoAOMT : caféoyl CoA O-méthyltransférase, CA2H : acide caféique 2-hydroxylase; FA2H : acide férulique 2-hydroxylase; MDCA2H : méthylène-acide dioxycinnamique 2-hydroxylase; O-MT : O-méthyltransférase.

fermentation, dans laquelle l'ADH1 catalyse l'étape terminale par transformation de l'acétaldéhyde en éthanol (Okimoto et al., 1980 ; Sachs et al., 1980). Le rôle des ADH est notamment crucial dans les grains de pollen pour lesquels l'acétaldéhyde est particulièrement toxique (Strommer, 2011). De plus, les ADH semblent jouer un rôle dans la croissance du tube pollinique (Van Eldik et al., 1997). Ce tube pollinique excrétant de l'éthanol (Bucher et al., 1995), il est probable que les ADH récupèrent l'éthanol produit afin de le convertir en acétaldéhyde pour la production d'énergie et les voies de biosynthèse. En outre, la synthèse par la voie de la lipoxygénase des composés volatils aldéhydes et alcools en C6 et C9 dans les nectaires permettant d'attirer les pollinisateurs dépend d'une ADH pour la conversion aldéhyde/alcool (Hoballah et al., 2005 ; Garabagi et Strommer, 2004 ; Garabagi et al., 2005). De plus, les composés volatils en C6 et C9 déterminent la saveur des fruits, ils jouent ainsi un rôle indirect important dans la dispersion des graines par les animaux (Schwab et al., 2008). Finalement, les volatils produits par la voie de la lipoxygénase sont impliqués dans l'odeur dégagée par les feuilles broyées, constituant une défense contre les insectes prédateurs dans laquelle sont impliquées les ADH (Pichersky et Gershenzon, 2002).

Récemment, des recherches ont été réalisées sur les pigments bétalaïnes utilisés par l'industrie agro-alimentaire comme colorant. La synthèse de ces pigments étant limitée aux Caryophyllales et certains champignons supérieurs (Amanites), les recherches sur leur voie de biosynthèse pourraient permettre une bioproduction en système hétérologue des bétalaïnes (Polturak et Aharoni, 2019). Les bétalaïnes sont synthétisées à partir de l'acide aminé L-tyrosine. Or, la BvADHa, une ADH insensible à l'inhibition de la tyrosine chez *Beta vulgaris* est impliquée dans la synthèse de cet acide aminé, soulignant l'importance de la régulation métabolique primaire pour la diversification du métabolisme spécialisé des plantes (Lopez-Nieves et al., 2018). La co-expression de BvADHa avec les gènes biosynthétiques de bétalaïne a, en effet, permis de multiplier par sept la bioproduction de pigment (**Figure 11**) (Timoneda et al., 2018).

De nombreuses fonctions ont été attribuées aux ADH, notamment dans la synthèse des MS. En effet, le rôle des ADH est primordial dans l'initiation des voies de biosynthèse végétales puisqu'elles permettent la synthèse de molécules composées de groupes électrophiles modifiables par d'autres enzymes. Ainsi, l'étude de différentes MDR issues de *C. roseus* a permis, par exemple, de mieux comprendre le contrôle enzymatique de la stéréosélectivité dans la voie de synthèse des AIM (Stavrinides et al., 2016). Chez *P. hexandrum*, la



<u>Figure 13</u> : Analyse phylogénétique des P450 connues pour catalyser les modifications structurelles sur les squelettes stéroïdiens et triterpéniques. (Ghosh, 2017)

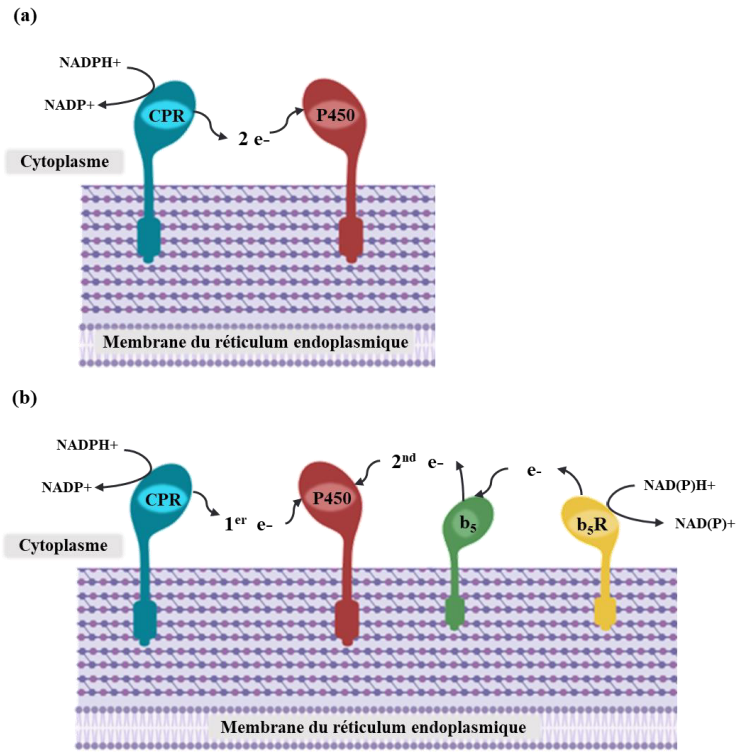
Les séquences complètes d'acides aminés des P450 ont été analysées par MEGA6. L'alignement des séquences a été effectué après l'analyse ClustalW et l'arbre phylogénétique a été construit à l'aide de la méthode Neighbor Joining. Les distances évolutives ont été calculées à l'aide de la méthode de correction de Poisson et sont exprimées en unités du nombre de substitutions d'acides aminés par site. *Agrobacterium tumefaciens* CYP104A1 a été inclus en tant que groupe externe.

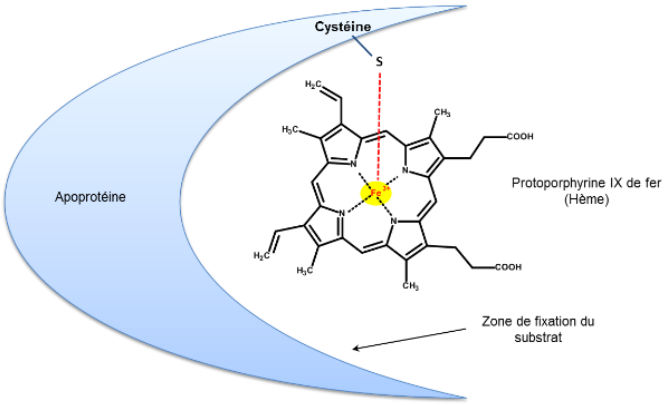
sécoisolaricirésinol déshydrogénase (SDH) est impliquée dans la synthèse du matairésinol, un précurseur de la lignane podophyllotoxine (Xia et al., 2001).

Réactions catalysées par les cytochromes P450

Les cytochromes P450 (P450) forment une superfamille d'hémoprotéines ubiquitaires qui sont retrouvées chez la majorité des êtres vivants. Leur nombre diffère selon les espèces, par exemple, chez *S. cerevisiae,* on ne retrouve que trois gènes codant des P450 (Nelson et al., 2004b) tandis que 55 gènes codent des P450 chez l'Homme (Werck-Reichhart and Feyereisen, 2000) et plus de 244 gènes chez *A. thaliana* (Bak et al., 2011). Cette diversité et cette démultiplication des P450 chez les végétaux s'expliquerait par la nécessité de produire de multiples molécules afin de s'adapter à l'environnement. En effet, cette famille d'enzymes est particulièrement représentée chez les plantes puisque les P450 sont impliquées dans les différents métabolismes spécialisés d'origine végétale et sont considérées comme des enzymes clés de la diversité des MS (Mizutani et Sato, 2011 ; Renault et al., 2014). Par exemple, la voie des phénylpropanoïdes menant aux coumarines implique de nombreuses étapes dépendantes des P450 (**Figure 12**) (Bourgaud et al., 2006). De même, les modifications structurelles médiées par les P450 sont indispensables à la diversification et la fonctionnalisation des triterpènes comme en témoigne le nombre important, environ 80, de P450 dont les fonctions sont liées au métabolisme des triterpènes végétaux (**Figure 13**) (Ghosh, 2017). Les P450 sont aussi impliquées dans le métabolisme des composés alcaloïdiques. Par exemple, CYP82E4v1 est impliquée dans la conversion de la nicotine en nornicotine (Siminszky et al., 2005). En outre, ces enzymes sont également impliquées dans la synthèse de phytohormones et participent à la détoxification d'agents chimiques comme les herbicides.

La réaction principale catalysée par les P450 est la mono-oxygénation permettant l'insertion d'un atome d'oxygène sur un substrat organique. La réaction catalysée par ce type d'enzyme est couplée à une oxydation de NAD(P)H en $NAD(P)^+$. En raison de leur besoin en électrons, les P450 ne sont donc pas des enzymes autonomes et dépendent d'un partenaire donneur d'électrons, principalement la Cytochrome P450 Réductase (CPR). Les CPR appartiennent à la famille des diflavines réductases (Jensen et Møller, 2010) et possèdent deux domaines. Le premier domaine contient une flavine adénine dinucléotide (FAD) ayant une





<u>Figure 15</u> : Représentation schématique d'une cytochrome P450. (Fourreau E., Mémoire de thèse sur l'élucidation de la voie de biosynthèse des alcaloïdes de *Catharanthus roseus* et ingénierie métabolique dans la levure, école doctorale SSBCV)

<u>Figure 14</u> : Schéma des complexes CPR-P450 chez les plantes.

(a) Mécanisme réactionnel d'un complexe CPR-P450. Les deux électrons nécessaires à la réaction sont transférés par la CPR.
(b) Mécanisme réactionnel d'un complexe CPR-P450-b5-b5R. La CPR et le complexe cytochrome b5 (b5) - cytochrome b5 réductase (b5R) transfèrent chacun un électron à la P450.



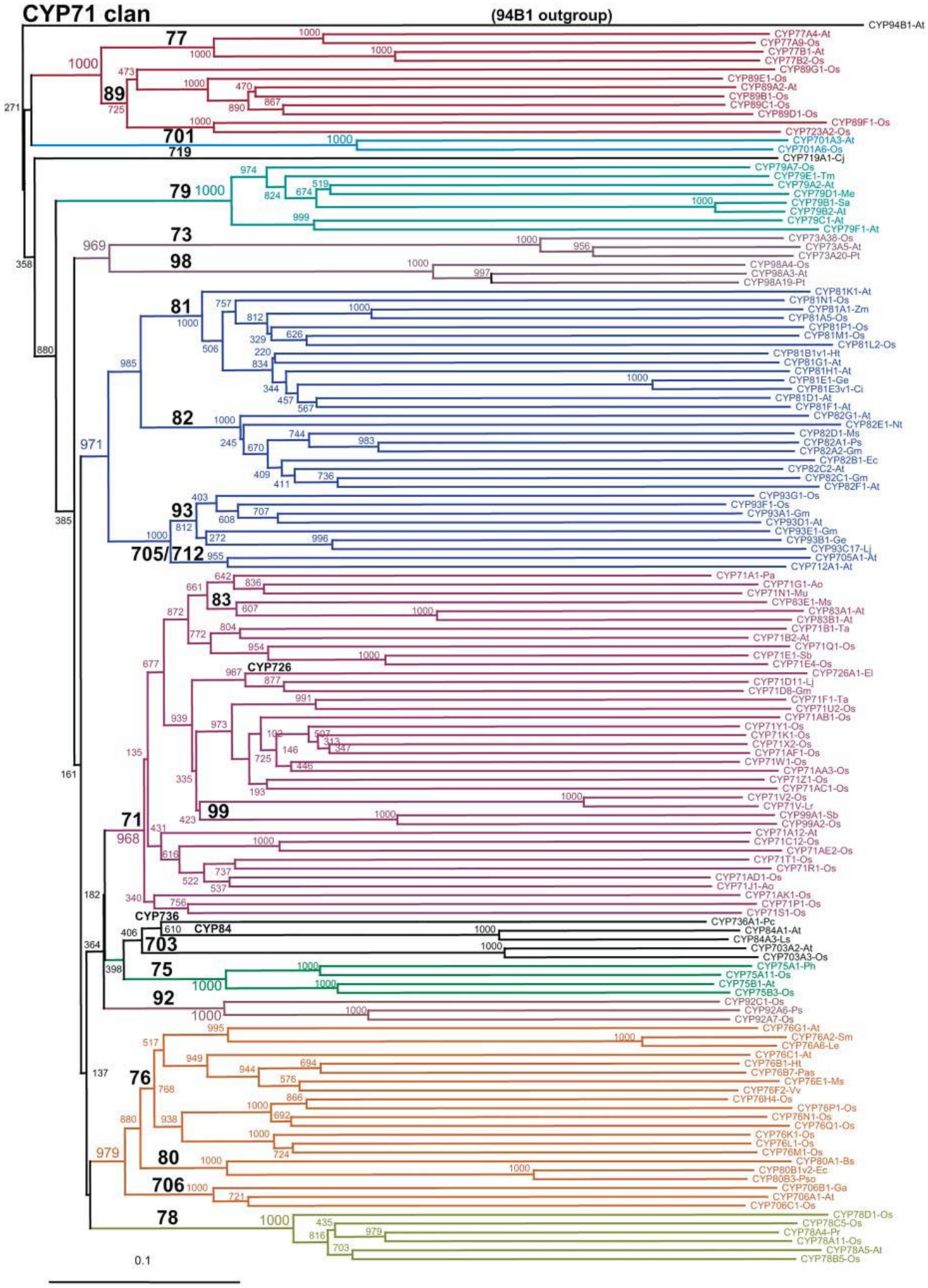
<u>Figure 16</u> : Cladogramme circulaire de 246 P450 pleine longueur d'*A. thaliana*. (Bak et al., 2011)

L'alignement des séquences multiples source a été réalisé dans MUSCLE (http://www.drive5.com/muscle/). Les paramètres étaient : pénalité d'ouverture d'écart : 11 ; pénalité d'extension d'écart : 0,75 ; centre : 0,0 ; Matrice de Gonnet. L'alignement a été affiné une fois (paramètres identiques). L'arbre phylogénétique Neighbor-Join a été créé à l'aide de MEGA4 (http://www.megasoftware.net/) avec le modèle de substitution basé sur la matrice Jones-Taylor-Thornton.

activité oxydoréductrice. Le second domaine contient une flavine mononucléotide (FMN) et sert de domaine transporteur d'électrons de la CPR vers le P450. Il existe deux classes de CPR, les CPR de classe I associées au métabolisme primaire, et les CPR de classe II associées au métabolisme secondaire (Parage et al., 2016). En outre, les cytochromes b_5 peuvent se combiner aux complexes cytochrome P450-CPR afin d'améliorer la réduction du P450 et de stabiliser son association avec le substrat (**Figure 14**) (Renault et al., 2014). En effet, dans ce cas, le complexe cytochrome b_5-cytochrome b_5 réductase (b_5R) assure le transfert d'un électron au P450, permettant ainsi à la CPR de ne délivrer qu'un électron au lieu de deux au P450 (Schenkman and Jansson, 2003). Quoi qu'il en soit, les P450 nécessitent une concentration non-négligeable en cofacteurs capables de fournir les électrons nécessaires à l'activité de celles-ci.

Les P450 sont des hémoprotéines constituées d'une protoporphyrine IX, d'un noyau de fer (hème) et d'une apoprotéine. La protoporphyrine IX et l'apoprotéine sont liées de façon non-covalente par l'intermédiaire d'une cystéine. L'hème est lié aux quatre atomes d'azote de la protoporphyrine IX et au groupement thiolate de l'apoprotéine (**Figure 15**). La grande majorité des P450 eucaryotes possèdent une hélice transmembranaire en N-terminal leur permettant de s'ancrer à la membrane du réticulum endoplasmique (RE) du côté du cytosol et de se regrouper autour des CPR, elles aussi ancrées à la membrane du RE (Schwarz, 1991).

Un système commun de nomenclature des gènes codant les P450 a d'abord été mis en place sur l'identité des séquences protéiques des P450 de 85 eucaryotes et 20 procaryotes (Nelson et al., 1996), puis une nomenclature des P450 végétales a été mise en place sur la base de l'identité des séquences protéiques et de la phylogénie d'*A. thaliana* en 2002 (Werck-Reichhart et al., 2002). Les P450 étaient initialement réparties en deux clades : le clade de type A et le clade de type non-A (Durst and Nelson, 1995 ; Paquette et al., 2000). Avec le temps et la découverte de nouvelles P450, il s'est avéré que les P450 du clade de type A disposaient de séquences très divergentes et une nouvelle nomenclature a été mise en place. Dans cette nomenclature, les P450 d'une même famille partagent au moins 40% d'identité et les P450 d'une même sous-famille partagent au moins 55% d'identité. Le clade de type A est ainsi devenu un des clades les plus importants parmi les nouveaux clades définis. Ce clade initial est ainsi devenu le clade CYP71. La nouvelle nomenclature des P450 végétales propose les clades suivants : CYP51, CYP71, CYP710, CYP711, CYP72, CYP74, CYP81, CYP85, CYP86 et CYP97 (**Figure 16**) (Bak et al., 2011). De plus, certaines familles de P450 végétales existent



<u>Figure 17</u> : Phylogénie du clade CYP71. (Nelson et al., 2004a)

Il s'agit du plus grand ensemble de P450. L'extrême diversification des familles et des sous-familles du clan CYP71 indique une grande utilité pour l'évolution des plantes, mais aucun membre du clan CYP71 n'a été détecté chez *C. reinhardtii*. Cet arbre a 139 séquences dans 26 familles. Il s'agit d'un arbre Neighbor-Joining utilisant une matrice de notation de Gonnett.

sans être représentées chez *A. thaliana* : CYP80, CYP92, CYP99, CYP719, CYP723, CYP725 et CYP726 (Werck-Reichhart et al., 2002).

Les clades CYP51 et CYP710 sont très anciens et déjà présents dans les algues brunes et les champignons. Les P450 appartenant à ces clades sont nécessaires à la synthèse des stérols membranaires, des stéroïdes et des triterpènes. Le clade CYP711 apparaît aussi très tôt dans l'évolution des plantes. Sa fonction semble liée à la synthèse des strigolactones, hormones dérivées des caroténoïdes (Gomez-Roldan et al., 2008). Les P450 appartenant au clade des CYP72 ont des fonctions dans le métabolisme des acides gras, dans la biosynthèse des cytokinines et dans le catabolisme des hormones. Les enzymes du clade CYP74 n'utilisent pas d'oxygène moléculaire ni de partenaire d'oxydo-réduction. En effet, l'oxygène, les réducteurs et les protons nécessaires à la réaction sont apportés par le substrat hydropéroxyde d'acide gras. Les membres du clade CYP85 sont impliqués dans le métabolisme des terpénoïdes, tandis que ceux du clade CYP86 sont impliqués dans le métabolisme des acides gras et des alcanes. Finalement, les P450 du clade CYP97 sont impliquées dans la synthèse des caroténoïdes et localisées dans les chloroplastes. Le clade CYP71 est l'un des plus grands clades des P450 (**Figure 17**) (Nelson et al., 2004a). Les CYP71 participent à la synthèse de métabolites spécialisés impliqués dans la protection des végétaux contre les stress biotiques et abiotiques dans de nombreuses espèces. Cette famille de P450 est retrouvée dans les voies de biosynthèse de toutes les familles de MS. Par exemple, les tabersonine-16-hydroxylases T16H1 (CYP71D12) et T16H2 (CYP71D351) sont impliquées dans la synthèse de l'AIM tabersonine chez *C. roseus* (Besseau et al., 2013); CYP71BE54 et CYP82D61 sont impliquées dans la synthèse de la podophyllotoxine, un lignane retrouvé chez *P. hexandrum* (Lau et Sattely, 2015), et CYP716A14v2 est impliquée dans l'oxydation des triterpènes pentacycliques chez *Artemisia annua* (Moses et al., 2015). La plupart des clades présentés disposent de sous-familles, ce qui prouve l'immensité de la diversité des P450 végétales.

Si les réactions catalysées par les P450 sont souvent associées à des réactions d'hydroxylation, elles peuvent néanmoins catalyser des réactions plus diverses telles que l'époxidation, la déalkylation, la déhydratation, la formation de ponts méthylènedioxy, le couplage de phénols ou encore le clivage de liaisons carbone-carbone (**Figure 18**) (Werck-Reichhart and Feyereisen, 2000 ; Schuler and Werck-Reichhart, 2003 ; Ortiz de Montellano and Voss, 2005 ; Mizutani and Sato, 2011).

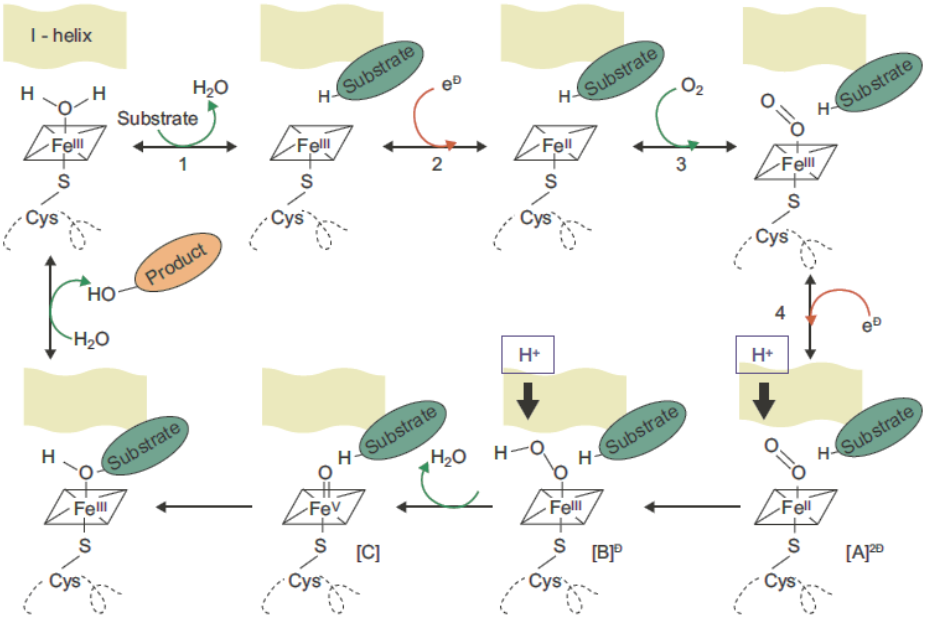


Figure 18 : Mécanisme catalytique des enzymes P450. (Werck-Reichhart and Feyereisen, 2000)

Les P450 sont le plus souvent des mono-oxygénases, catalysant l'insertion d'un des atomes d'oxygène moléculaire dans un substrat, le second atome d'oxygène étant réduit en eau. La réacion la plus fréquemment catalysée est l'hydroxylation (insertion O) en utilisant l'intermédiaire fer-oxo très réactif et électrophile (espèce [C], rangée du bas). La forme hydroperoxo de l'enzyme (espèce [B]-) est également un oxydant électrophile catalysant l'insertion OH+. L'attaque nucléophile peut être catalysée par les espèces [A]2- et [B]- ; la réduction, l'isomérisation ou la déshydratation sont catalysées par les formes sans oxygène de l'enzyme. Ceci, ajouté à la variété des apoprotéines et à la réactivité intrinsèque de tous leurs substrats, explique l'extraordinaire diversité des réactions catalysées par les enzymes P450.

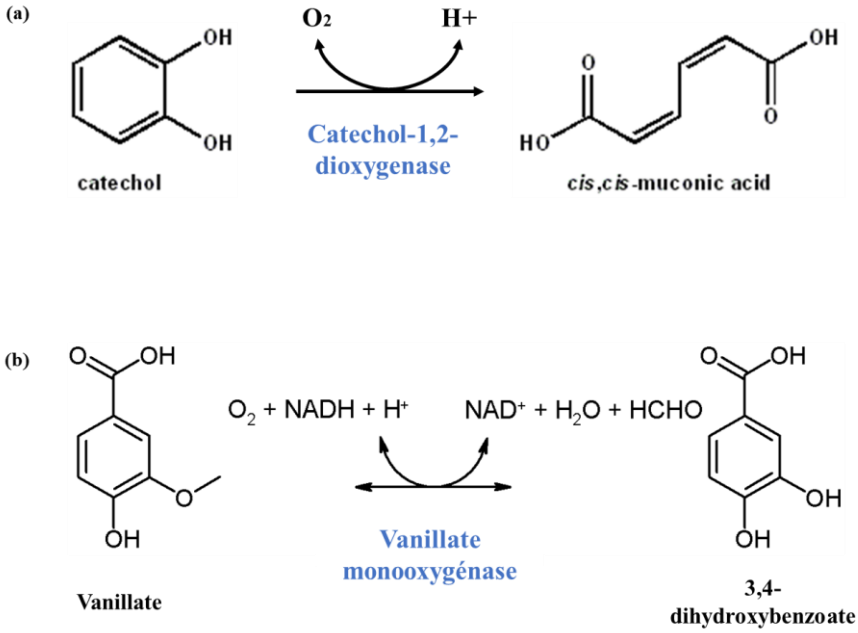


Figure 19 : Exemples de réaction d'une dioxygénase et d'une monooxygénase.

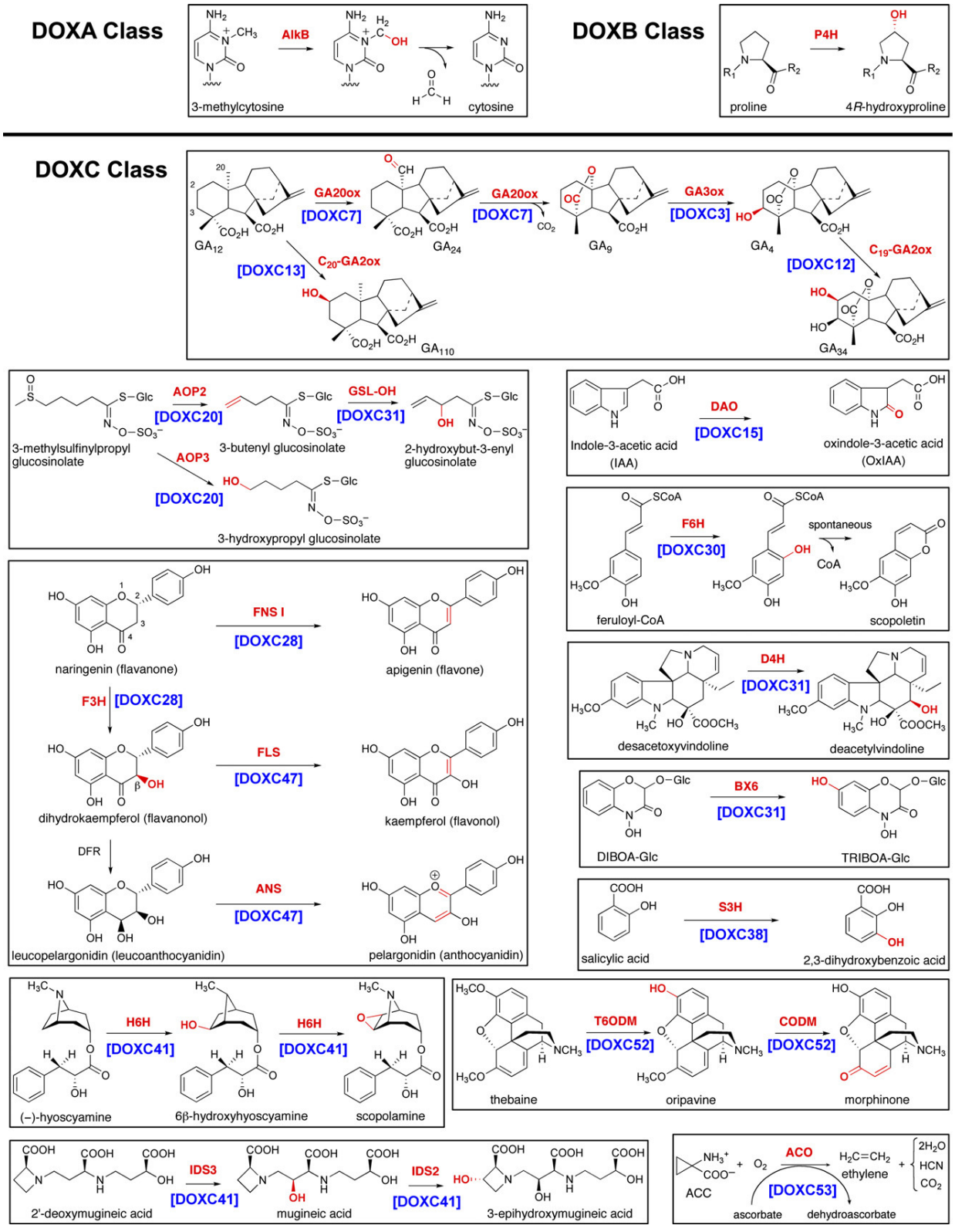
(a) Catechol-1,2-dioxygénase **c**atalysant l'oxydation du catéchol par du dioxygène sans libération de molécule d'eau.
(b) Vanillate monooxygénase catalysant l'oxydation de la vanillate avec libération d'une molécule d'eau.

Réactions d'oxydation et dioxygénases

Les dioxygénases sont des enzymes oxydoréductrices catalysant l'oxydation d'un substrat par du dioxygène sans libération de molécule d'eau, contrairement aux monooxygénases catalysant l'oxydation d'un seul atome d'oxygène avec libération d'une molécule d'eau (**Figure 19**). Les dioxygénases nécessitent un cofacteur métallique (fer(II) ou fer(III)). Il existe plusieurs familles de dioxygénases, notamment, parmi les dioxygénases non-héminiques dépendantes du fer(II), on peut retrouver les dioxygénases 2-oxoglutarate-dépendantes (2OGB), les catéchol dioxygénases, la naphtalène dioxygénase, les lipo-oxygénases, les cyclo-oxygénases, ou encore les β-carotène dioxygénases (Bugg, 2003).

La superfamille des 2OGB représente la deuxième plus grande famille de protéines végétales. En particulier, chez *A. thaliana*, plus de 100 gènes codent des 2OGB. Ces protéines sont localisées sous forme soluble dans le cytosol des cellules végétales. Elles nécessitent du 2-oxoglutarate et de l'oxygène comme co-substrats afin de catalyser l'oxydation d'un substrat (Kawai et al., 2014). Elles sont notamment impliquées, chez les plantes, dans l'hydroxylation et la déméthylation oxydative des acides nucléiques et des histones, dans la 4-hydroxylation de la proline lors de la synthèse des protéines de la paroi cellulaire, dans l'épimérisation, la désaturation et l'halogénation (Wang et al., 2021). Une classification phylogénétique de ces enzymes a été mise en place sur une base d'analyse de séquences (Kawai et al., 2014). Dans cette classification, les 2OGB végétales ont été divisées en trois groupes : la classe DOXA responsable de la réparation de l'ADN (Aravind and Koonin, 2001), la classe DOXB impliquée dans l'hydroxylation de la proline (Gorres and Raines, 2010) et la classe la plus abondante, DOXC, impliquée dans le métabolisme des phytohormones et des MS, elles-mêmes divisées en clades (**Figure 20**) (Kawai et al., 2014).

Les catéchol dioxygénases catalysent le clivage oxydatif du catéchol qui représente une étape importante dans la dégradation des composés aromatiques par les bactéries. Elles sont composées de deux sous-familles d'enzymes : les intradiol catéchol dioxygénases et les extradiol catéchol dioxygénases (Hayaishi, 1966). La naphtalène dioxygénase est composée d'une flavoprotéine réductase dépendante NADPH, d'une ferrédoxine et d'une oxygénase de Rieske formant une chaîne de transfert d'électrons au sein de l'enzyme et est impliquée dans la dégradation des hydrocarbures par les bactéries (Haigler and Gibson, 1990). Les lipo-oxygénases sont des dioxygénases présentes chez les végétaux et impliquées dans de nombreux



<u>Figure 20</u> : Classification des 2OGB végétales. (Kawai et al., 2014)

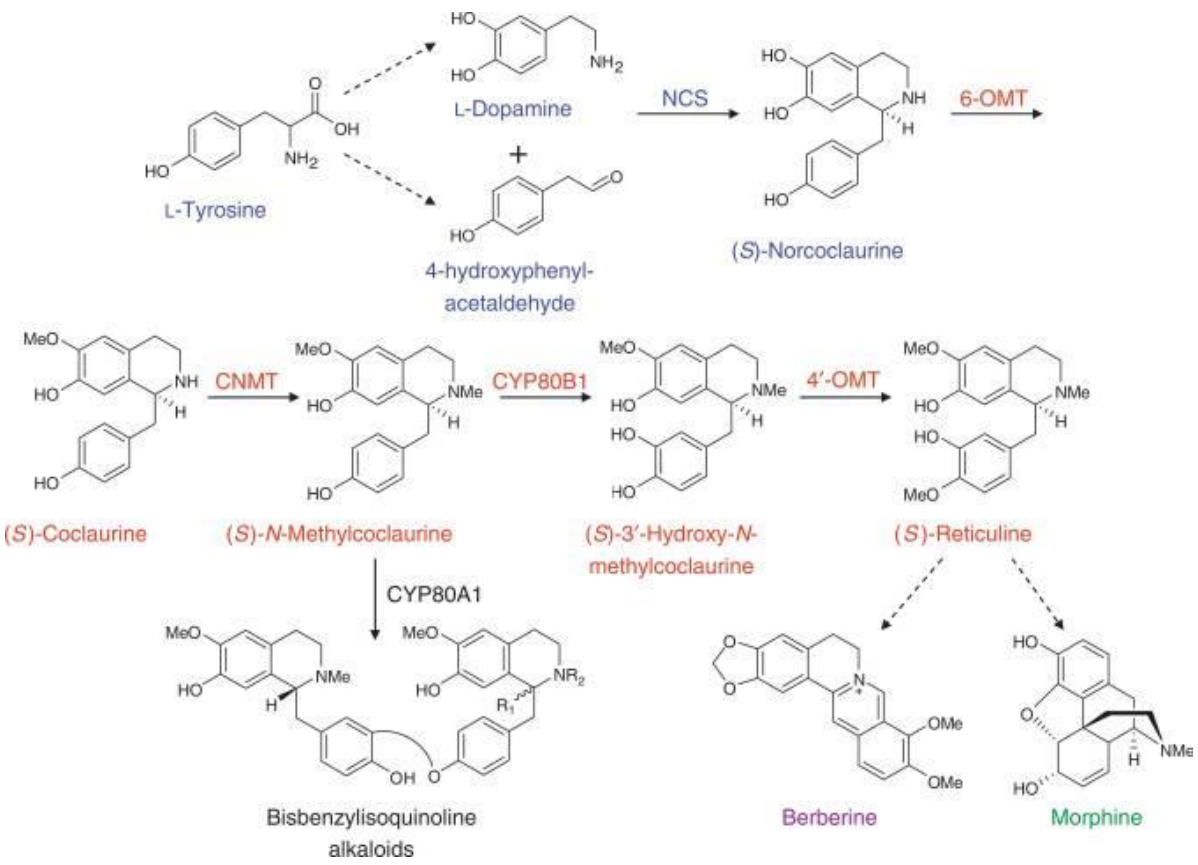
Le clade pour chaque 2OGD est indiqué sous les flèches en bleu.

processus physiologiques mais aussi dans la réponse aux stress environnementaux. Les cyclo-oxygénases, elles, sont impliquées dans la biosynthèse de la prostaglandine, une molécule anti-inflammatoire, chez les mammifères (Bandeira-Melo et al., 2000). La β-carotène dioxygénase catalyse le clivage du β-carotène dans la synthèse de la rhodopsine (Olson and Hayaishi, 1965).

Parmi les dioxygénases caractérisées chez les végétaux, les dioxygénases époxy-caroténoïdes sont connues pour catalyser le clivage oxydatif d'un carotéinoïde afin de synthétiser l'acide abscissique dans de nombreuses espèces végétales. En particulier, l'enzyme AtCCD1 d'*A. thaliana,* une dioxygénase catalysant le clivage des caroténoïdes a été mise en évidence (Schwartz et al., 2001). Sur la base de similarité de séquence, il s'avère que les orthologues de ce gène sont présents dans tout le règne végétal. De même, de nombreuses 2OGB ont été caractérisées. C'est notamment le cas de la désacétoxyvindoline-4-hydroxylase (D4H) catalysant la 4-hydroxylation de la désacétoxyvindoline chez *C. roseus* qui a permis une avancée majeure dans l'élucidation de la voie de synthèse de l'AIM vindoline (Vazquez-Flota et al., 1997). De même, les 2OGB thébaïne-6-*O*-déméthylase (T6ODM) et codéine-*O*-déméthylases (CODM) sont respectivement impliquées dans la synthèse des BIA thébaïne et codéine chez le pavot à opium (Hagel and Facchini, 2010).

Méthylation par les O-méthyltransférases

Les *O*-méthyltransférases (OMT) végétales représentent une grande famille d'enzymes catalysant la méthylation d'un atome d'oxygène de nombreux MS, et en particulier des phénylpropanoïdes, des flavonoïdes (Ibrahim et al., 1987 ; Gauthier et al. 1998) et des alcaloïdes (Lam et al., 2007). Les OMT utilisent la S-adénosyl-L-méthionine (SAM) comme donneur de méthyle et produisent de la S-adénosyl-L-Homocystéine (SAH) (Barton and Nakanishi, 1999). La réaction d'*O*-méthylation joue un rôle important dans la plante puisqu'elle est impliquée dans la synthèse de la lignine (Lewis and Yamamoto, 1990), la tolérance aux stress biotiques et abiotiques, la résistance aux maladies, mais aussi dans la décoration de composés phénoliques impliqués dans la croissance et le développement des plantes (Hahlbrock and Scheel, 1989). Notamment, chez le pois, la synthèse du flavonoïde pisatine, important pour la résistance de la plante au champignon *Nectria hematococca*, implique une 3-*O*-méthylation (Liu et al., 2006). Chez *Vitis vinifera*, l'*O*-méthylation des flavonoïdes anthocyanes dans les



<u>Figure 21</u> : Voie de synthèse de la morphine et de la berbérine chez *P. somniferus*. (Hawkins and Smolke, 2008)

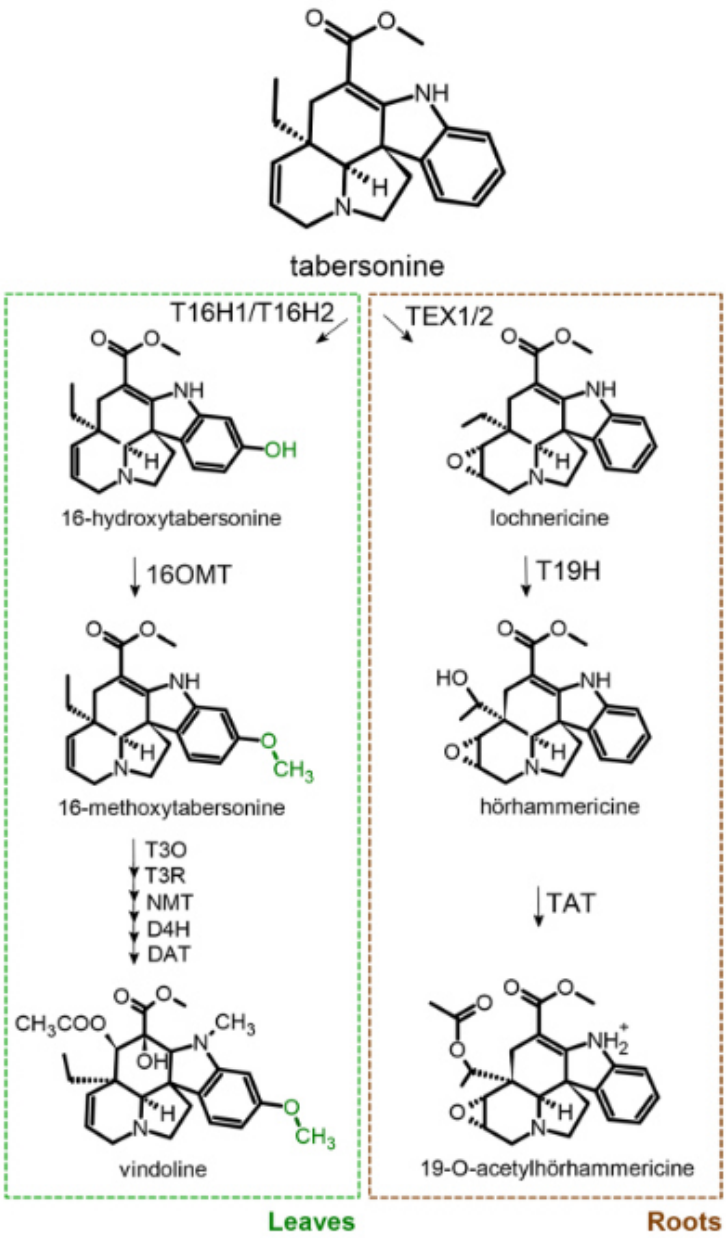
Les schémas de couleurs pour les métabolites et les enzymes sont les suivants : partie en amont de la voie vers la première molécule de squelette norcoclaurine en bleu ; partie médiane de la voie de la norcoclaurine à la réticuline en rouge ; branche de la sanguinarine/berbérine en violet ; branche de la morphine en vert. Les étapes de conversion pour lesquelles l'ensemble complet d'enzymes n'a pas été complètement élucidé ou cloné sont indiquées par des flèches en pointillés.

baies de raisin réalisées par la caféoyl-CoA *O*-méthyltransférase VvCCoAOMT est impliquée dans la résistance des baies lors d'un stress hydrique (Giordano et al., 2016).

Les recherches sur la bioproduction de BIA tels que la morphine issue de *P. somniferum*, tant au niveau de la plante que pour l'ingénierie métabolique, ont montré que la voie de biosynthèse de la morphine implique des enzymes de type OMT. En effet, chez *P. somniferum*, les *O*-méthylations réalisées sur la (R,S)-norlaudanosoline par la 6OMT, et sur la (R,S)-3'-hydroxy-*N*-méthylcoclaurine par la 4'OMT permettent la synthèse de la (R,S)-réticuline, précurseur de la morphine et de la berbérine (**Figure 21**) (Hawkins and Smolke, 2008). De même, la tabersonine-16-*O*-méthyltransférase (T16OMT) de *C. roseus*, a été mise en évidence et est impliquée dans la synthèse de l'AIM vindoline, précurseur des anticancéreux vinblastine et vincristine utilisés couramment dans les protocoles de chimiothérapie, a été mise en évidence (Levac et al., 2008). Plus récemment, l'isoforme T16OMT2 a été caractérisée (Lemos-Cruz et al., 2022). Cette isoforme est à la fois impliquée dans la synthèse de la vindoline et dans la synthèse de certains flavonoïdes, suggérant une relation évolutive entre les métabolismes des alcaloïdes et des flavonoïdes chez *C. roseus*. Finalement, la Matairésinol *O*-méthyltransférase (CtMROMT) est impliquée dans la méthylation des lignanes chez *Carthamus tinctorius*. En effet, cette OMT catalyse la méthylation du matairesinol en 4'-*O*-méthylmatairésinol (Umezawa et al., 2013).

Acylation par les acyltransférases

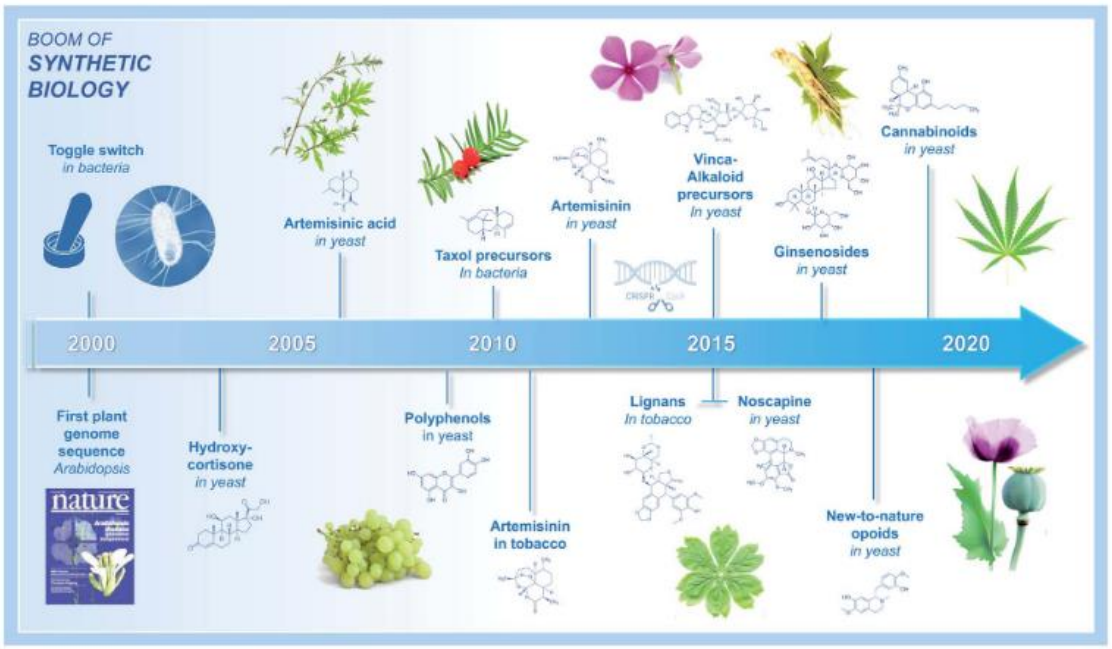
Les acyltransférases sont des enzymes appartenant à la famille des transférases. Elles permettent le transfert d'un groupement acyle et sont retrouvées principalement dans le métabolisme primaire des végétaux. Elles jouent un rôle important dans la formation des chaînes d'acides gras nécessaires à la mise en place des membranes biologiques, au stockage intracellulaire et à la modification des protéines (Röttig and Steinbüchel, 2013). En effet, les acyltransférases permettent de transférer une grande variété de substrats sur les chaînes acyles activées ou leur polymérisation. On retrouve, notamment, parmi cette famille de molécules, les Acyl-CoA-DAG acyltransférases (DGAT), les glycéraldéhyde-3-phosphate acyltransférases (GPAT), les acides lysophosphatidique acyltransférases (LPAAT) et les acyltransférases BAHD. Les DGAT sont impliquées dans la synthèse des cérides qui jouent un rôle dans l'imperméabilisation de la plante et celle des triglycérides qui représentent une forme de réserve



<u>Figure 22</u> : Utilisation de la tabersonine dans les différents organes de *C. roseus.* (Lemos-Cruz et al., 2022)

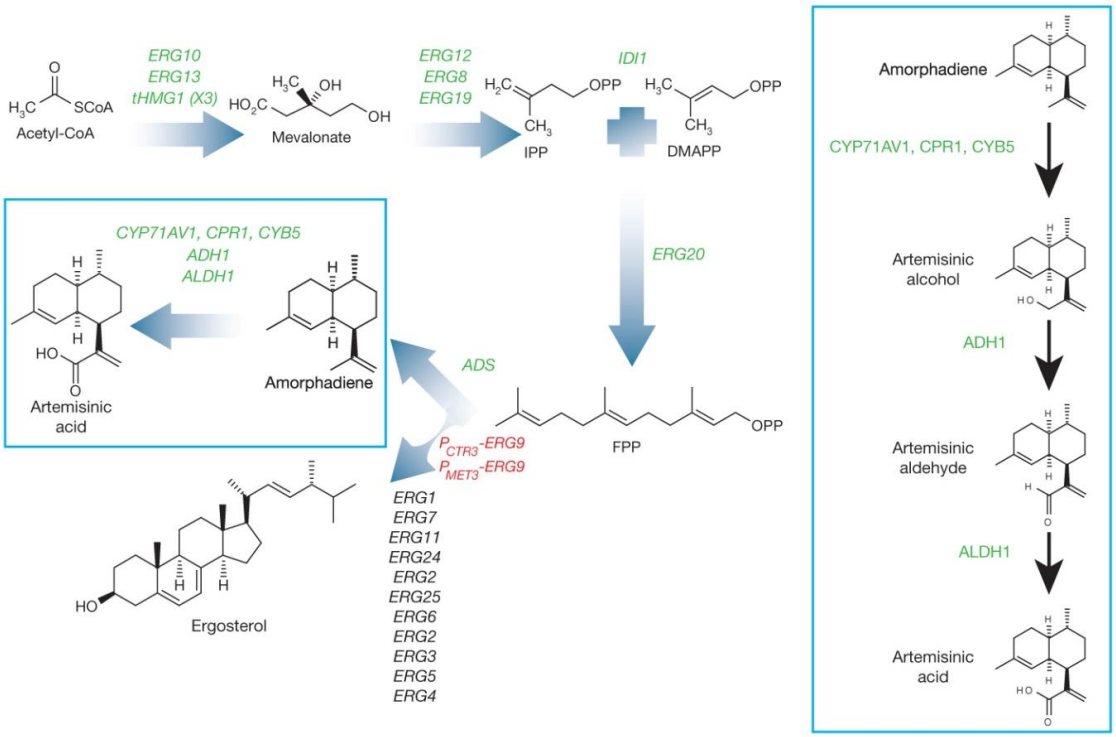
L'hydroxylation en position C-16 est réalisée par la tabersonine 16-hydroxylase (T16H) dans les feuilles et autres organes aériens, tandis que dans les racines, la tabersonine est hydroxylée en position C-19 par la tabersonine 19-hydroxylase (T19H). Le produit T16H est ensuite méthylé par la 16OMT pour donner la 16-méthoxytabersonine, qui après 5 étapes enzymatiques, se transforme en vindoline. Dans les racines, la tabersonine peut être époxydée par les isoformes 1 et 2 de la tabersonine 6,7-époxydase (TEX1/TEX2), avant l'hydroxylation du C-19 par T19H, ce qui donne le produit hörhammericine, qui est acétylé par le dérivé de la tabersonine 19-O-acétyltransférase (TAT), pour produire la 19-O-acétylhörhammericine.

pour les végétaux (Kennedy, 1961). Les GPAT catalysent la synthèse de glycéraldéhyde-3-phosphate, essentiel à la glycolyse (Takeuchi and Reue, 2009). Les LPAAT sont impliquées dans le développement de la plante et dans les mécanismes de défense (Wattelet, 2022). Les BAHD sont, quant à elles, impliquées dans la synthèse de nombreuses molécules végétales, notamment de MS, et rassemblent plusieurs groupes d'acyltransférases. Le nom de cette famille d'enzymes, BAHD, a ainsi été choisi en prenant la première lettre des quatre premiers membres caractérisés : BEAT (Benzylalcohol *O*-acetyltransferase) (Dudareva et al., 1998), AHCT (Anthocyanin *O*-hydroxycinnamoyltrasnsférase) (Fujiwara et al., 1998), HCBT (N-hydroxycinnamoyl/benzoyltransferase) (Yang et al., 1997) et DAT (Deacetylvindoline 4-*O*-acetyltransferase) (St Pierre et al., 1998 ; St Pierre and De Luca, 2000). Plusieurs BAHD ont été caractérisées chez *C. roseus*, dont celle impliquée dans la synthèse de l'AIM vindoline. En effet, l'acétyl-CoA dédacétylvindoline-4-*O*-acétyltransférase (DAT) est une BAHD impliquée dans la dernière étape de la synthèse des AIM menant à la vindoline (St Pierre et al., 1998). La découverte de cette enzyme a d'ailleurs permis de définir la famille des BAHD. Plus récemment, la tabersonine 19-O-acétyltransférase (TAT), a également été caractérisée et est impliquée dans l'acétylation des dérivés 19-hydroxytabersonine dans les racines de *C. roseus* (**Figure 22**) (Carqueijeiro et al., 2018). De même, les acyltransférases BAHD CiSTH1 et CiSHT2 de type spermidine hydroxycinnamoyl transférase ont récemment été mises en évidence dans la chirorée (*Cichorium intybus*) (Bernard et al., 2022). Ces enzymes sont impliquées dans la synthèse de spermines entièrement substituées (tétracoumaroyl spermines) sur la surface des grains de pollen des Astéracées. La suppression de ces gènes via le système CRISPR/Cas9 dans la plante a permis de démontrer que CiSTH2 est impliquée dans les premières étapes de N-acylation conduisant aux di- et tri-coumaroyl spermines, tandis que CiSTH1 est impliquée dans la dernière étape de N-acylation conduisant à la tétracoumaroyl spermine. L'accumulation de spermine tétra-substituée dans les grains de pollen de la chicorée nécessite donc l'expression de ces deux enzymes. Des compléments de description des BAHD seront fournis dans la section III.3 de l'introduction.



<u>Figure 23</u> : Chronologie de la production de métabolites spécialisés végétaux dans des organismes hétérologues. (Courdavault et al., 2021)

Un résumé historique des principales réalisations au sein des plus grandes familles de MS, y compris les polyphénols, les terpènes, les taxanes, les lignanes, les alcaloïdes indoliques, les opioïdes et les cannabinoïdes, entre autres, est décrit ici à partir du début du 21e siècle.



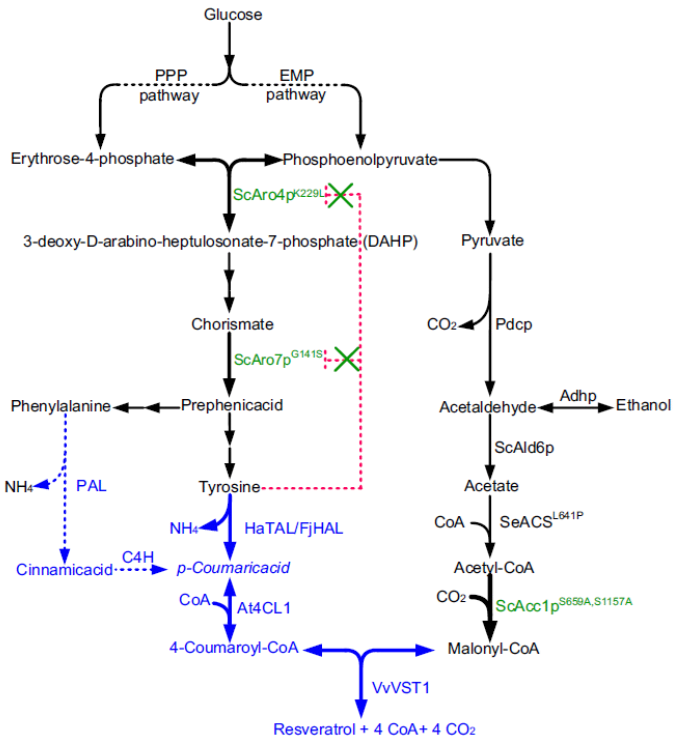
<u>Figure 24</u> : Voie de biosynthèse vers la synthèse de l'acide artémisinique en levure. (Paddon et al., 2013)

Les gènes surexprimés contrôlés par le système d'induction GAL sont représentés en vert. La squalène synthase réprimée par le cuivre ou la méthionine (ERG9) est indiquée en rouge. DMAPP : diphosphate de diméthylallyle ; FPP : diphosphate de farnésyle ; IPP : diphosphate d'isopenténylе. tHMG1 code pour la HMG-CoA réductase tronquée. CYP71AV1, CPR1 et CYB5 oxydent l'amorphadiène en alcool artémisinique; ADH1 oxyde l'alcool artémisinique en aldéhyde artémisinique ; ALDH1 oxyde l'aldéhyde artémisinique en acide artémisinique.

b) Production de métabolites spécialisés végétaux par ingénierie métabolique

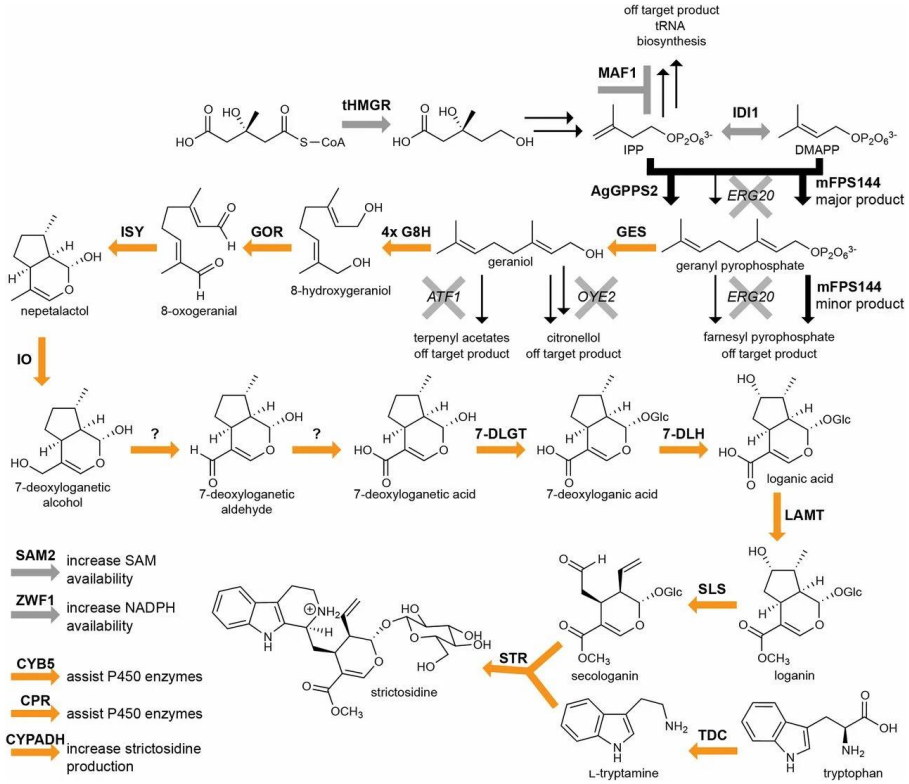
L'ingénierie métabolique pour la bioproduction de MS végétaux est utilisée depuis une vingtaine d'années (Verpoorte et al., 1999) et de plus en plus de molécules végétales d'intérêt pharmaceutique sont produites via l'utilisation d'un organisme hôte. En effet, cette technologie permet de produire des métabolites à la demande, de façon contrôlée grâce au génie des procédés et indépendamment des contraintes inhérentes à l'utilisation de matériel végétal. De plus, si l'hôte microbien sert de plateforme pour synthétiser de nouveaux dérivés de produits naturels, il peut aussi permettre de mieux comprendre les enzymes impliquées dans le métabolisme spécialisé. Au fil des années, de nombreux outils ont été développés pour construire des souches microbiennes productrices de MS (Pickens et al., 2011 ; Hadadi et al., 2016 ; Guerriero et al., 2018 ; Cravens et al., 2019 ; Chen et al., 2020). Ces outils incluent ceux associés à l'identification des gènes candidats comme décrit précédemment mais aussi à l'optimisation de l'ingénierie de voies de biosynthèse hétérologues, comme l'édition du génome, la régulation de l'expression génique ou encore la reconfiguration spatiale. Plusieurs MS ont ainsi pu être synthétisés avec succès via l'utilisation de l'organisme hétérologue *S. cerevisiae*, comme l'acide artémisinique (25 g/L), le resvératrol (531 mg/L), la strictosidine (25,2 mg/L), la thébaïne (0,006 mg/L), le ginsénoside Rh2 (300mg/L), et la noscapine (2,2 mg/L).

Par souci de concision, j'ai choisi de ne décrire que quelques exemples emblématiques de production par ingénierie métabolique tels que ceux de l'acide artéminisique, du resvératrol, de la stricosidine, de la vindoline, de la (S)-réticuline, ou encore de la noscapine. En particulier, la bioproduction d'acide artéminisique issue d'*Artemisia annua* en levure a permis de jeter les bases de l'ingénierie métabolique de voies de biosynthèse végétales en levure (**Figure 23**). L'acide artéminisique est un précurseur de l'artéminisine, un sesquiterpène endopéroxide connu pour son activité antipaludique. La première souche de levure mise en place dans l'optique de produire de l'acide artéminisique en levure est capable de convertir 153 mg/L d'amorphadiène, un précurseur de l'acide artéminisique (Ro et al., 2006). Une souche *S. cerevisiae* productrice d'acide artéminisique a ensuite été mise en place (**Figure 24**) (Paddon et al., 2013). Dans cette souche, les 16 gènes issus de la voie de biosynthèse de l'acide artémisinique d'*A. annua* sont contrôlés par des promoteurs inductibles au galactose (en vert sur la Figure 24). De plus, la réduction d'expression de CPR1, la CPR associée au P450 CYP71AV1 via l'utilisation d'un



<u>Figure 25</u> : Voies de biosynthèses vers la synthèse de resvératrol en levure. (Li et al., 2015)

ScAro4p : DAHP synthase ; ScAro7p : chorismate mutase ; PAL : phenylalanine ammonia-lyase ; C4H: cinnamate-4-hydroxylase ; HaTAL/FjTAL : tyrosine ammonia-lyase d' H. aurantiacus/Flavobacterium johnsoniae; At4CL1 : 4-coumarate-CoA ligase ; VvVST1 : resveratrol synthase de *Vitis vinifera* ; Pdcp : pyruvate decarboxylase ; Adhp : alcohol dehydrogenase ; ScAld6p : aldehydedehydrogenase ; ScDAHPACS1 : acetyl-CoA synthetase ; ScAcc1p : acetyl-CoA carboxylase. Les réactions natives et les intermédiaires sont représentés en bleu. Les flèches en pointillés indiquent les réactions qui n'ont pas été utilisées dans cette étude. Les flèches doubles représentent plusieurs étapes enzymatiques.



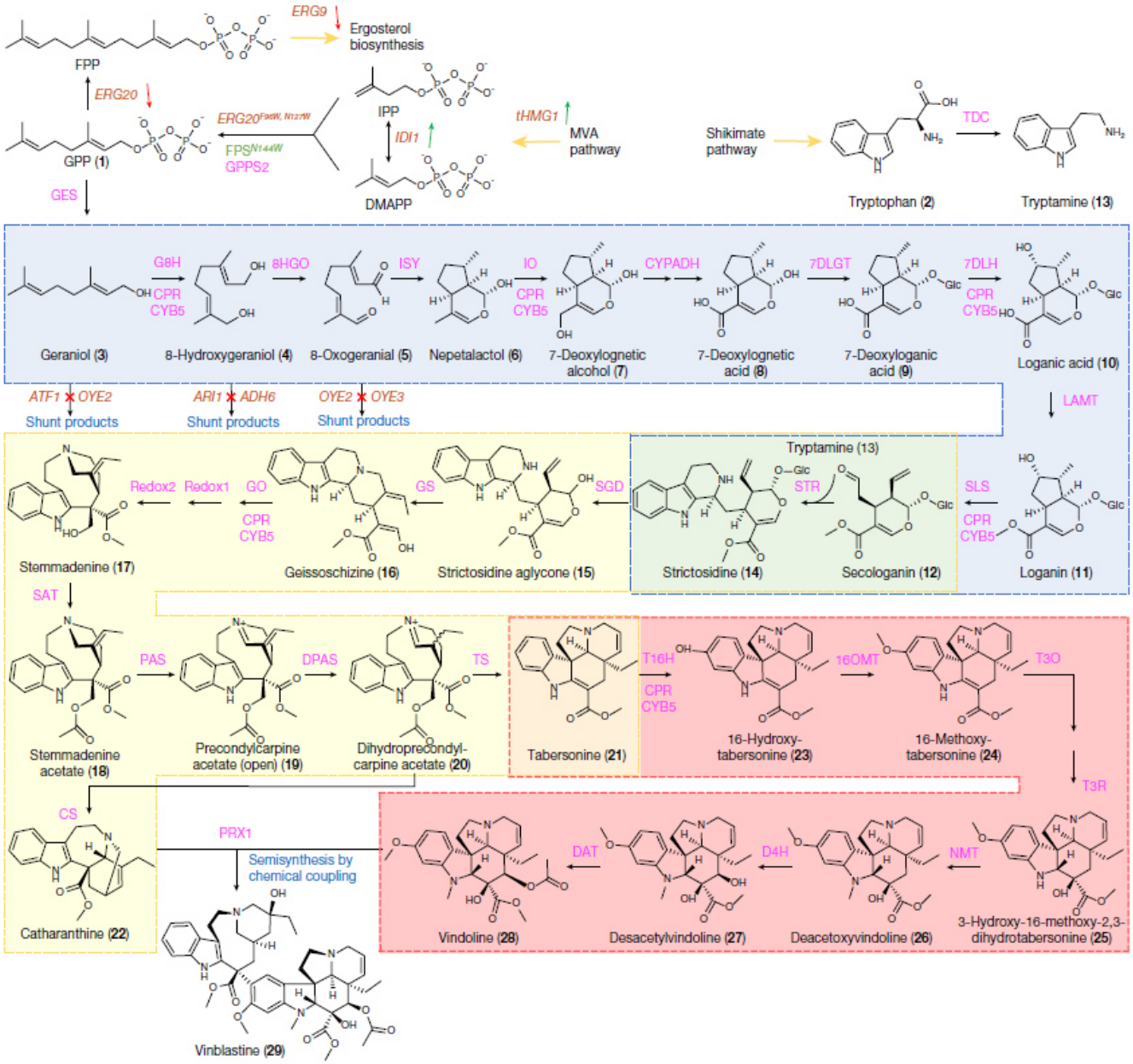
<u>Figure 26</u> : Voie de biosynthèse reconstituée de la stricosidine chez *S. cerevisiae*. (Brown et al., 2015)

Les gènes de *C. roseus* (orange) ont été intégrés dans le génome de *S. cerevisiae* pour permettre la production de strictosidine. Pour augmenter le flux à travers la voie, des secondes copies de gènes natifs de *S. cerevisiae* (gris) ont été intégrées dans le génome. La ligne grise indique que MAF1 est un répresseur de la biosynthèse de l'ARNt. Le X gris représente les gènes qui ont été supprimés pour réduire la consommation d'intermédiaires de voie impliqués dans d'autres voies de biosynthèse. ERG20 a été supprimé et remplacé par AgGPPS2 (*A. grandis*), qui produit exclusivement du pyrophosphate de géranyle, et mFPS144 (*G. gallus*), qui favorise la production de pyrophosphate de géranyle par rapport au pyrophosphate de farnésyle. Les flèches avec des points d'interrogation sont catalysées par IO, mais peuvent également nécessiter des enzymes auxiliaires.
AgGPPS2 : *Abies grandis* géranyl pyrophosphate synthase ; RCR : cytochrome P450 réductase ; CYB5 : cytochrome b5 ; 7-DLGT : glucosyl transférase de l'acide 7-désoxyloganétique; 7-DLH : hydroxylase de l'acide 7-désoxyloganique; DMAPP : pyrophosphate de diméthylallyle; G8H : géraniol 8-hydroxylase; GES : géraniol synthase ; GOR : 8-hydroxygéraniol oxydoréductase; IDI1 : isopentényl pyrophosphate-diméthylallyl pyrophosphate isomérase; IO : acide 7-désoxyloganétique synthase/iridoïde oxydase; IPP : pyrophosphate d'isopentényle; ISY : iridoïde synthase; LAMT : acide loganique *O*-méthyltransférase; mFPS144 : *Gallus gallus* mutant farnésyl pyrophosphate synthase N144W; SAM2 : S-adénosylméthionine synthétase; SLS : sécologanine synthase ; STR : strictosidine synthase; TDC : tryptophane décarboxylase ; tHMGR : 3-hydroxy-3-méthyl-glutaryl-CoA réductase tronquée ; ZWF1 : glucose-6-phosphate déshydrogénase.

promoteur faible, additionnée à l'expression du cytochrome b5 d'*A. annua* CYB5, ainsi que le clonage de l'alcool déshydrogénase ADH1 et de l'aldéhyde déshydrogénase ALDH1 ont permis d'augmenter de façon significative les rendements de bioconversion, atteignant les 25 g/L. Un procédé chimique permettant une conversion de l'acide artémisinique en artémisinine à l'aide d'une source d'oxygène a été développé, permettant ainsi la mise en place d'un procédé de production industriel viable d'artémisinine, assurant ainsi son approvisionnement. Le groupe Sanofi a, en effet, mis en place une chaîne de production par hémisynthèse d'artémisinine via la bioproduction d'acide artémisinique par fermentation en Bulgarie et transformation chimique en Italie. Ce mode de production a été suspendu en 2015 car les parasites *Plasmodium* à l'origine du paludisme présentent une telle résistance à l'artémisinine que ce mode de traitement n'est plus intéressant pour le traitement de cette maladie.

Le resvératrol, polyphénol issu de *Vitis vinifera* ou encore de *Polygonum cuspidatumi,* disposant, notamment, de vertus antioxydantes, est aussi un bon exemple de MS produit en levure (Li et al., 2015). En effet, ce polyphénol utilisé en tant que complément alimentaire et ingrédient cosmétique peut désormais être produit en bioréacteur par *S. cerevisiae* à partir de glucose ou d'éthanol. Le clonage de la tyrosine ammonia-lyase d'*Herpetosiphon aurantiacus* HaTAL, la 4-coumarate-CoA ligase At4CL1 d'*A. thaliana* et la resvératrol synthase VvVST1 de *Vitis vinifera* dans la levure ont d'abord permis d'obtenir 2,73 mg/L de resvératrol à partir du glucose (**Figure 25**). La souche a ensuite été optimisée via la surexpression de la DAHP synthase ScAro4p et de la chorismate mutase ScAro7p et par multiplication du nombre de copies des gènes codant HaTAL, At4CL1 et VvST1. L'approvisionnement en précurseur malonyl-CoA a été augmenté via la surexpression d'une version non régulée de l'acétyl-CoA carboxylase ACC1. Toutes ces stratégies, additionnées au génie des procédés, ont permis d'obtenir une bioproduction en bioréacteur largement supérieure de 415 à 531 mg/L de resvératrol.

La strictosidine est le dernier intermédiaire commun à toutes les voies de biosynthèse des AIM, la reconstitution de la voie de synthèse de cette molécule en système hétérologue peut donc permettre une production rapide et peu coûteuse des AIM. Une souche de levure a ainsi été mise en place pour vérifier la faisabilité de la bioproduction de strictosidine en levure (Brown et al., 2015). Cette souche a été modifiée par introduction de quatorze gènes issus de la voie de synthèse des AIM chez *C. roseus*, de sept gènes supplémentaires issus de *S. cerevisiae*, *Abies grandis* et *Gallus gallus* et de trois délétions de gènes endogènes, permettant la production d'environ 0,5 mg/L de strictosidine (**Figure 26**). Récemment, une souche de levure

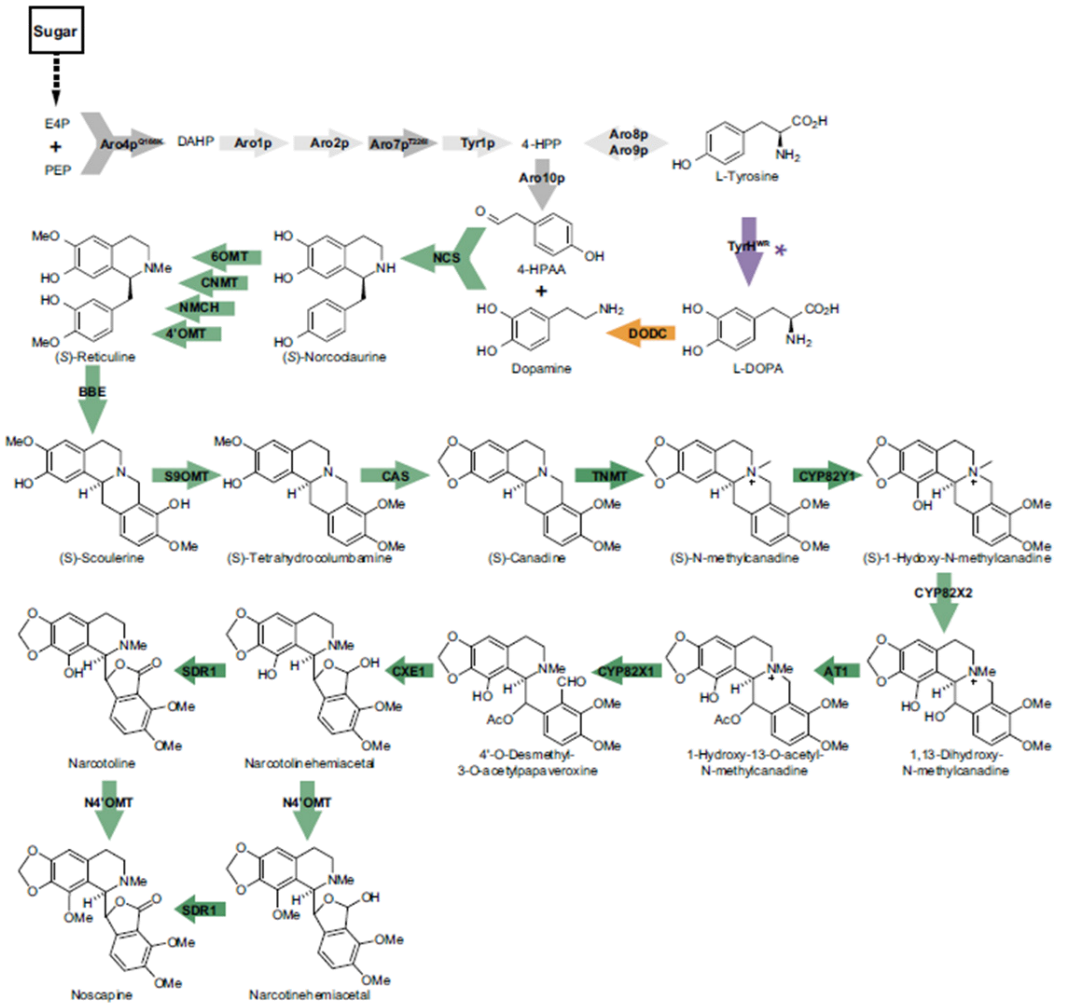


<u>Figure 27</u> : Voie biosynthétique complète pour la production de vinblastine dans la levure. (Zhang et al., 2022)

Les gènes de levure (orange) surexprimés, réprimés ou supprimés sont indiqués par des flèches vertes, des flèches rouges et des croix rouges respectivement. IPP : isopentenylpyrophosphate ; DMAPP : pyrophosphate de diméthylallyle; GPPS : GPP synthase; FPSN144W : variante FPP synthase N144W ; RCR : NADPH-cytochrome P450 réductase ; CYB5 : cytochrome b5 ; GES : géraniol synthase ; G8H : géraniol 8-hydroxylase; 8HGO : 8-hydroxygéraniol oxydoréductase ; ISY : iridoïde synthase; IO : iridoïde oxydase; CYPADH : alcool déshydrogénase 2; 7DLGT : acide 7-désoxyloganétique glucosyl transférase; 7DLH : hydroxylase de l'acide 7-désoxyloganique; LAMT : acide loganique O-méthyltransférase; TDC : tryptophane décarboxylase ; GS : geissoschizine synthase; GO : geissoschizine oxydase; Redox1 : protéine redox 1 ; Redox2 : protéine redox 2 ; SAT : stemmadénine-*O*-acétyltransférase ; CS : catharanthine synthase; TS : tabersonine synthase; T16H : tabersonine 16-hydroxylase; 16OMT : tabersonine 16-*O*-méthyltransférase ; T3O : tabersonine 3-oxygénase; T3R : 16-méthoxy-2,3-dihydro-3-hydroxytabersonine synthase; NMT : 3-hydroxy-16-méthoxy-2,3-dihydrotabersonine-N-méthyltransférase ; D4H : désacétoxyvindoline 4-hydroxylase ; DAT : déacétylvindoline-O-acétyltransférase; PRX1, peroxydase de classe III. Les cases colorées avec des lignes pointillées indiquent modules strictosidine (bleu), tabersonine/catharanthine (jaune) et vindoline (rouge). Les chevauchements entre les modules sont marqués en vert (chevauchement entre les modules strictosidine et tabersonine/catharanthine) et en orange (chevauchement entre les modules tabersonine/catharanthine et vindoline).

capable de produire *de novo* l'AIM ajmalicine a été mise en place via l'intégration de 24 gènes dans le génome de la levure (Liu et al., 2022). Cette souche est capable de produire 119,2 mg/L d'alcaloïdes hétéroyohimbines dont 61,4 mg/L d'ajmalicine. Sachant que la strictosidine est un intermédiaire de la voie de biosynthèse de l'ajmalicine, il est fort probable que les taux de biosynthèse de strictosidine de cette souche soient supérieurs à ceux de la souche de Brown et son équipe. Les AIM anticancéreux vinblastine et vincristine s'accumulent dans les feuilles de *C. roseus*. La forte demande de ces deux molécules et l'impact négatif de la culture massive de *C. roseus* impliquent de récurrentes ruptures de stock en vinblastine et vincristine. La bioproduction de ces molécules par un organisme hôte tel que *S. cerevisiae* a donc été envisagée comme méthode alternative de production. Dans cette optique, une souche de levure capable de produire *de novo* la vindoline et la catharantine, précurseurs de la vinblastine, à partir de pyrophosphate de géranyle et de tryptophane a été mise en place. Dans cette souche, 34 gènes hétérologues végétaux ont été clonés et 10 gènes endogènes ont été inactivés ou surexprimés (**Figure 27**) (Zhang et al., 2022). Une concentration maximale de 25,2 mg/L de strictosidine, précurseur de la vindoline et la catharantine, a été obtenue avec cette souche.

Les BIA constituent une autre grande classe de MS connus pour leurs activités biologiques pharmacologiquement intéressantes. La plupart des BIA, dont la noscapine extraite du pavot à opium, dérivent de la (S)-réticuline. La noscapine et ses dérivés noscapinoïdes sont utilisés en médecine pour leurs propriétés antitussives et anticancéreuses. La culture du pavot étant très sensible aux facteurs environnementaux et la synthèse chimique de la noscapine étant difficilement commercialisable, la mise en place d'une cellule de levure productrice de cette molécule est essentielle pour une future production industrielle de BIA par ingénierie métabolique. La bioproduction de (S)-réticuline a récemment été considérablement améliorée dans une souche de *S. cerevisiae* via le clonage de 18 gènes, dont 8 gènes hétérologues et un gène orthologue de la tyrosine hydroxylase permettant d'orienter le flux métabolique vers la production du précurseur dopamine. De plus, l'ajustement du nombre de copies de chaque gène et la suppression de sept oxydoréductases endogènes impliquées dans la transformation du précurseur 4-hydroxyphénylacétaldéhyde, nécessaire à la synthèse de la (S)-réticuluine, en acide ou en alcool, ont permis la bioproduction de 4,6 g/L de (S)-réticuline (Pyne et al., 2020). D'autre part, Li et son équipe se sont intéressés à la synthèse de la noscapine chez *S. cerevisiae* à partir de sources simples de carbone et d'azote (Li et al., 2018). Leur souche de levure productrice de noscapine dispose de 25 enzymes hétérologues provenant de plantes, de mammifères et de bactéries et 6 enzymes endogènes ont été mutées ou surexprimées (**Figure**



<u>Figure 28</u> : Biosynthèse de novo de la noscapine chez *S. cerevisiae*. (Li et al., 2018)

Les flèches grises clair correspondent aux enzymes de levure non modifiées, les flèches grises foncé aux enzymes de levure surexprimées et mutées, les flèches violettes aux enzymes de mammifères (*Rattus norvegicus*), les flèches orange aux enzymes bactériennes (*Pseudomonas putida*) et les flèches vertes aux enzymes végétales (*Papaver somniferum*, *Coptis japonica*).

Aro10p : phénylpyruvate décarboxylase ; Aro1p : enzyme aromatique pentafonctionnelle ; Aro2p : chorismate synthase bi fonctionnelle et flavine réductase ; Aro4pQ166K : DAHP synthétase ; Aro7pT226I : chorismate mutase ; Aro8p : aminotransférase aromatique I ; Aro9p : aminotransférase aromatique II ; AT1 : 1,13-dihydroxy-N-méthylcandine 13-O-acétyltransférase; BBE : enzyme pont berbérine ; BH4 : 5,6,7,8-tétrahydrobioptérine; CAS : canadine synthase; CNMT : coclaurine N-méthyltransférase; RCR : cytochrome P450 réductase ; CXE1 : 3-O-acétylpapaveroxine carboxylestérase; CYP82X1 : 1-hydroxy-13-O-acétyl-N-méthylcanadine 8-hydroxylase ; CYP82X2 : 1-hydroxy-N-méthylcanadine 13-hydroxylase ; CYP82Y1 : 1-hydroxy N-méthylcanadine synthase ; DAHP : 7-phosphate d'acide 3-désoxy-D-arabino 2-heptulosonique; DODC : L-dopa décarboxylase; E4P : érythrose 4-phosphate; MT2/MT3 : narcotoline-4′-O-méthyltransférase. NCS : (S)-norcoclaurine synthase; NMCH : N-méthylcoclaurine hydroxylase; PEP : phosphoénolpyruvate ; S9OMT : scoulérine 9-O méthyltransférase ; SDR1 : déshydrogénase/réductase à chaîne courte ; Tkl1p : transcétolase ; TNMT : tétrahydroprotoberbérine cis-N-méthyltransférase; Tyr1p : préphénate déshydrogénase; TyrHWR : tyrosine hydroxylase résistante à la rétro-inhibition (R37E, R38E, W166Y); 4-HPAA : 4-hydroxyphénylacétaldéhyde ; 4-HPP : 4-hydroxyphénylpyruvate; 4′OMT : 3′-hydroxy N-méthylcoclaurine 4′-Ométhyltransférase; et 6OMT : norcoclaurine 6-O-méthyltransférase.

28). La souche a, par ailleurs, été optimisée via la régulation du métabolisme endogène de la levure (optimisation de la production du cofacteur NADPH nécessaires aux cytochromes P450), permettant ainsi d'atteindre une bioproduction de 2,2 mg/L de noscapine en bioréacteur lorsque le milieu de culture est optimal pour la souche.

En conclusion, de premiers MS sont produits par ingénierie métabolique en levure. Les souches obtenues représentent une avancée majeure dans la bioproduction de biomédicaments et sont un véritable succès scientifique avec de nombreuses modifications génétiques. Parmi les taux de bioproduction des MS d'intérêt présentés dans cette partie, deux s'avèrent d'ores et déjà compatibles avec une production industrielle, les autres disposent de rendements de biosynthèse encore trop faibles en l'état pour être exploitables à cette échelle (**Tableau 1**). En effet, les bioproductions peuvent commencer à devenir commercialement viables lorsque les concentrations atteignent le gramme par litre. Il est donc essentiel de mettre en place de nouveaux outils afin d'améliorer la production de métabolites spécialisés végétaux par les cellules de levure « usines ».

I.3 La levure, un châssis cellulaire

a) Modalités de bioproduction en levure

La bioproduction de MS par transfert des voies de biosynthèse végétales en levure peut être mise en place selon deux stratégies. La première consiste en l'intégration dans la levure de la voie de biosynthèse complète initiée par la prise en charge d'un substrat endogène, en général issu du métabolisme primaire de la levure. Cette stratégie dite de synthèse *de novo* permet de produire le MS d'intérêt sans fournir de précurseurs spécifiques, la fermentation de la levure en présence de sucre et d'acide aminés suffisant à produire la molécule d'intérêt. Une telle reconstruction métabolique est menée étape par étape, mais les interactions entre le métabolisme exogène et celui de la levure sont peu prévisibles, et il est souvent nécessaire de de revenir sur la construction de la souche pour optimiser chaque étape de la voie. Cette technique est donc longue à mettre en place et la souche doit toujours être retravaillée afin d'atteindre des taux de synthèse compatibles avec une production industrielle. De plus, les voies de synthèse *de novo* impliquent généralement un nombre non négligeable d'étapes

<u>Tableau 1</u> : Taux de production en levure des métabolites spécialisés présentés.

Les taux de production représentées en vert sont compatibles avec une production industrielle**.**

Molécule	Famille	Production (mg/L)	Référence
Acide artéminisique	Terpène	25000	Paddon et al., 2013
Resvératrol	Polyphénol	415 à 531	Li et al., 2015
Strictosidine	Alcaloïde	25,2	Brown et al., 2015
Vindoline	Alcaloïde	266	Kulagina et al., 2021
(S)-réticuline	Alcaloïde	4600	Pyne et al., 2020
Noscapine	Alcaloïde	2,2	Li et al., 2018

enzymatiques dont le fonctionnement de concert peut être difficile à atteindre. Il en résulte ainsi la nécessité de transférer un nombre élevé de gènes hétérologues et donc de disposer de mécanismes de transfert compatibles avec une intégration multiple de gènes exogènes. Finalement, la synthèse étant réalisée à partir d'un précurseur endogène, il est parfois requis de modifier le métabolisme de la levure afin d'améliorer le flux métabolique vers la production de ce précurseur tout en évitant d'altérer la croissance ou d'autres fonctions physiologiques de la cellule. Comme présenté précédemment, par exemple, le clonage d'un gène orthologue de la tyrosine hydroxylase et la suppression de sept oxydoréductases endogènes ont été nécessaires chez la levure afin d'améliorer la biosynthèse, respectivement, de la dopamine et de la 4-hydroxyphénylacétaldéhyde, précurseurs de la (S)-réticuline (Pyne et al., 2020).

La seconde stratégie de bioproduction de MS repose sur une approche dite de bioconversion. Elle consiste en l'intégration d'une partie de la voie de biosynthèse, en général les étapes finales conduisant au métabolite ciblé, permettant ainsi de réduire le nombre d'étapes enzymatiques dans la voie hétérologue. Si les levures ainsi générées ne disposent pas d'une capacité autonome de bioproduction des MS, ceci est compensé par l'apport, aux levures en culture, d'un intermédiaire biosynthétique des MS ciblés, généralement hautement biodisponible et qui sera converti en la molécule d'intérêt grâce aux gènes hétérologues transférés. Cette stratégie est plus simple à mettre en place pour les voies de biosynthèse comportant de nombreuses étapes car elle permet de s'affranchir d'éventuels problèmes liés au fonctionnement concerté d'un très grand nombre d'enzymes souvent difficile à obtenir, aux interactions entre le métabolisme exogène et celui de la levure ainsi que d'une modification des flux métaboliques endogènes. Bien que cette stratégie puisse s'avérer initialement plus coûteuse du point de vue de l'industriel en raison de la nécessité de produire ou purifier le précurseur, les temps de développement et d'optimisation des souches sont largement réduits, les rendements de synthèse sont souvent plus élevés et les souches, plus robustes en raison d'un degré moindre d'édition génétique. Une souche de levure productrice de vindoline a ainsi été récemment mise en place selon le principe de conversion de la tabersonine (Kulagina et al., 2021c), dont l'approvisionnement est facile est déjà réalisé à l'échelle industrielle par extraction à partir de graines de *Voacanga africana* accumulatrices de très fortes teneurs en tabersonine. La voie de bioconversion clonée dans la levure via le système CRISPR/Cas9 nécessite 7 étapes enzymatiques. Cette souche a été optimisée en régulant finement la balance du nombre de copies des gènes impliqués dans la voie de biosynthèse de la vindoline chez *C. roseus*, permettant ainsi d'éviter la fuite métabolique vers la production de vindorosine, un produit sans

intérêt particulier issu d'une voie parallèle à la voie de la vindoline avec un précurseur commun. De plus, les rendements de bioconversion ont été améliorés grâce à une optimisation du pH et de la composition du milieu de culture. Ces stratégies d'optimisation ont permis d'obtenir une souche capable de convertir 88% de tabersonine en vindoline, correspondant à une concentration de 266 mg/L via une culture en bioréacteur. Ce taux de bioproduction en levure est le plus fort rapporté jusqu'à présent pour la synthèse d'un AIM. C'est pour cette raison que les travaux que j'ai conduit durant ma thèse ont exclusivement reposé sur des approches de bioconversion.

b) Mécanismes de transfert de gènes en levure

Quelles que soient les stratégies de bioproduction employées, des transferts de gènes multiples sont nécessaires afin de générer les cellules usines. Ceux-ci impliquent, en premier lieu, un transfert d'ADN transformant dans les levures qui peut impliquer l'électroporation, la conjugaison, la transformation de sphéroplastes ou la compétence naturelle des levures. Dans une optique d'ingénierie métabolique, il est important que l'ADN transféré dans la cellule soit maintenu au cours des divisions cellulaires. Ce maintien peut être réalisé par réplication autonome ou suite à une intégration chromosomique (Riley and Guss, 2021).

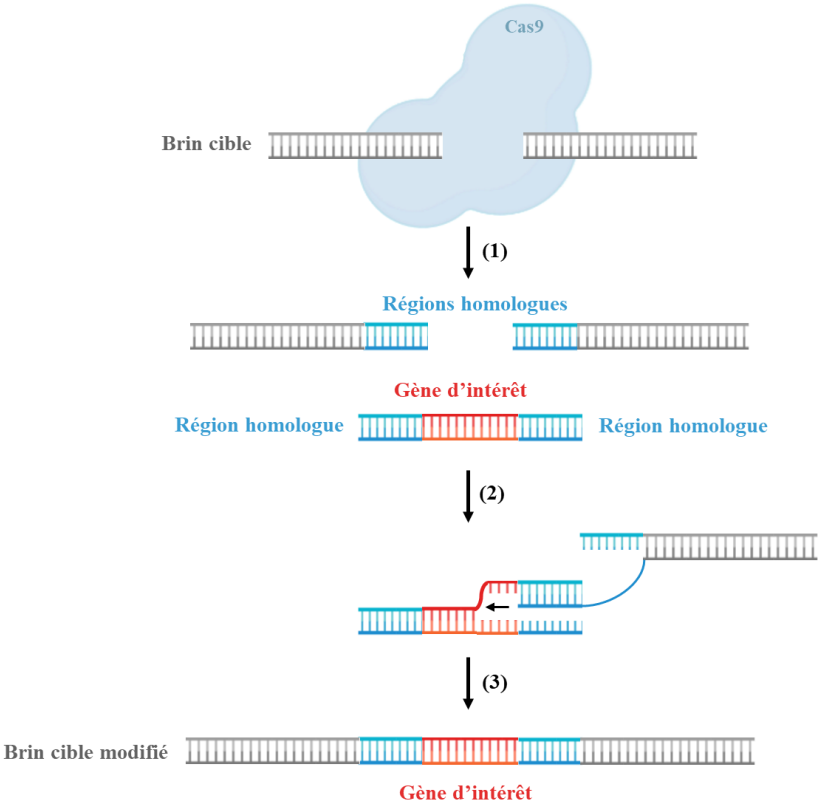
Le maintien de l'ADN transféré par réplication autonome consiste en l'apport dans l'organisme hôte d'un vecteur, très souvent de nature plasmidique, capable de se répliquer de façon indépendante des chromosomes grâce à une séquence appelée origine de réplication. C'est une méthode efficace pour obtenir un taux de transformation élevé et rapide et peut servir de preuve de concept lors de la validation de gènes candidats car elle garantit d'obtenir un nombre élevé de copies du gènes d'intérêt. Toutefois, elle reste peu compatible, voire incompatible pour le développement de cellules « usines » productrices de molécules à l'échelle industrielle. En effet, en plus d'une très grande variabilité entre souches, cette technique implique d'exercer une pression de sélection continue pour assurer le maintien des plasmides lors des divisions cellulaires et donc de cultiver obligatoirement les levures sur milieu sélectif contenant un antibiotique ou permettant une sélection basée sur l'auxotrophie, plus largement employée. Un organisme auxotrophe n'est pas capable de synthétiser un composé organique nécessaire à son développement. La plupart du temps, les auxotrophies employées en ingénierie métabolique sont liées à des déficiences en acides aminés ou de bases causées par l'invalidation

de certains de leurs gènes de biosynthèse. Une souche auxotrophe à un acide aminé ne pourra ainsi se développer que si le milieu de culture est supplémenté avec cet acide aminé ou si un gène de synthèse fonctionnel est apporté pour compenser la mutation du gène endogène. Dans la stratégie de réplication autonome, le plasmide transformé dans les levures porte ce gène ramenant donc l'autotrophie à l'acide aminé ou à la base. C'est le cas, par exemple, de la série de plasmides pESC d'Agilent Technologies conçus pour l'expression de deux gènes par plasmide chez *S. cerevisiae*. Chacun de ces plasmides dispose d'un gène marqueur de sélection ramenant l'autotrophie (HIS3 codant l'imidazole glycérol-phosphate déshydratase impliquée dans la biosynthèse de l'histidine, TRP1 codant la phosphoribosyl anthranilate isomérase impliquée dans la biosynthèse du tryptophane, LEU2 codant la bêta-isopropylmalate déshydrogénase impliquée dans la biosynthèse de la leucine et URA3 codant l'orotidine-5'-phosphate décarboxylase impliquée dans la biosynthèse de l'uracile). De ce fait, il est donc nécessaire de (i) retirer cet acide aminé/base du milieu de culture afin de garantir la pression de sélection et le maintien du plasmide mais aussi (ii) de disposer de souches présentant une à plusieurs auxotrophies pour l'apport de plusieurs plasmides. Si ces souches sont facilement générées par génie génétique ou directement disponibles, leur emploi est limité à l'échelle industrielle puisque les souches cultivées sont désormais dépourvues d'éléments de sélection. Elle s'avère en outre coûteuse car nécessitant d'utiliser des milieux supplémentés en produits chers et peut altérer la croissance de la souche. Une illustration de l'utilisation de ce type d'approche sera fournie dans la partie 1 de cette thèse relative au développement d'une souche productrice de phénolamides.

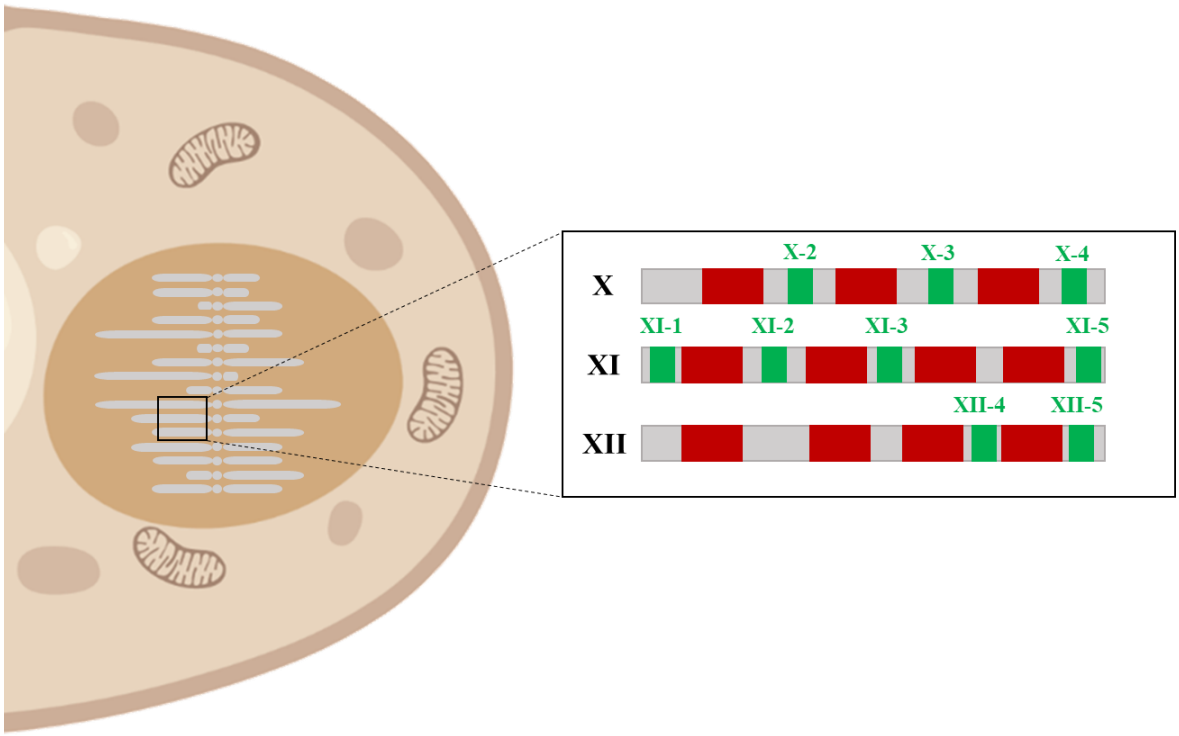
Le maintien de l'ADN transféré dans les levures peut également être assuré par son intégration dans le génome des levures qui peut être réalisée selon plusieurs approches. La recombinaison homologue à un locus ciblé du génome est restée la technique d'intégration chromosomique la plus utilisée depuis 1987 (Ma et al., 1987). Le principe de cette technique est d'exploiter la capacité des levures à intégrer avec une grande efficacité des fragments d'ADN exogènes dans leur génome sur la base d'une simple homologie de séquence et de réactions de crossing-over (échange de séquences génétiques entre chromosomes homologues lors de la division cellulaire). L'insertion d'un transgène est permise par simple transformation d'un fragment d'ADN transformant, souvent sous la forme d'un plasmide linéarisé ou d'un produit PCR, contenant le transgène flanqué par les séquences homologues chevauchantes du locus génomique d'insertion, orientées de façon inversée. La sélection des cellules transformées ayant procédé à la recombinaison nécessite, là encore, d'employer un gène marqueur de

sélection (antibiotiques, antifongiques ou d'auxotrophie) et, potentiellement, des souches auxotrophes. Différents loci peuvent ainsi être ciblés selon les séquences homologues employées qui délimitent une zone à intégrer comprenant le(s) transgène(s) et le marqueur de sélection. Fréquemment, cette intégration est réalisée au niveau des gènes conférant l'auxotrophie des souches et la recombinaison homologue permet alors, à la fois l'insertion d'un transgène, et l'insertion d'une copie fonctionnelle du marqueur métabolique qui restaure l'autotrophie. Toutefois, ces stratégies présentent des limites parfois similaires à celles décrites pour la réplication autonome, incluant, par exemple, le faible nombre de gènes marqueurs qui peut limiter le transfert des transgènes, la nécessité d'invalider préalablement les marqueurs métaboliques dans la souche de levure qui peut altérer la croissance, les limites liées à l'emploi des agents de sélection en cas de cultures industrielles de la souche, ou encore la localisation de certains marqueurs métaboliques à des loci à faible activité transcriptionnelle.

Plus récemment, des techniques d'édition du génome plus efficaces que la recombinaison homologue classique ont vu le jour. En théorie, l'édition du génome permet de modifier n'importe quel locus génomique par insertion, délétion ou remplacement d'un ou plusieurs nucléotides sans avoir recours à des marqueurs de sélection. Le principe de l'édition du génome repose sur la capacité des cellules à réparer les coupures double brin de l'ADN génomique (ADNg) pour assurer leur survie. Deux mécanismes de réparation existent : le NHEJ (Non-Homologous End-Joining) qui raccorde les fragments clivés mais génère souvent des mutations ponctuelles incontrolées, et la recombinaison homologue qui utilise une séquence homologue comme modèle pour recopier la région à remplacer, permettant une réparation fidèle de la cassure double brin. Ainsi, pour intégrer un transgène à un locus ciblé, il suffit de créer un clivage double brin de l'ADNg au niveau de ce locus, et de fournir à la cellule, pour réparation, un fragment d'ADN dont les extrémités sont homologues à la zone clivée mais contenant le gène à transférer (Caldecott, 2008). L'une des clefs de l'édition du génome est donc de pouvoir cliver l'ADNg de façon séquence-spécifique. Différentes techniques permettant ce ciblage, comme les enzymes ZFN (zinc finger nucleases) introduites en 2005 (Urnov, 2010), les enzymes TALENS (Transcription activator-like effector nucleases) (Hockemeyer et al., 2011) mises au point en 2010 ou encore l'enzyme Cas9 (CRISPR-associated protein 9) du système CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats) découverte en 2012. Les systèmes ZFN et TALENS sont des enzymes de restriction artificielles générées en fusionnant un domaine de liaison à l'ADN en doigts de zinc ou TALE (effecteurs de type activateur de transcription) respectivement à un domaine de



<u>Figure 29</u> : Mécanisme de fonctionnement du système CRISPR/Cas9.



<u>Figure 30</u> : Loci d'intégration dans le génome de la levure. (Mikkelsen et al., 2012)

Loci permettant l'intégration de cassettes d'expression dans le génome de la levure. Vert : Loci d'intégration; Rouge: éléments génétiques essentiels.
Ces loci sont situés dans des zones du génome à forte activité transcriptionnelle et entre des éléments génétiques essentiels à la vie de la levure, assurant une stabilité génétique des éléments intégrés.

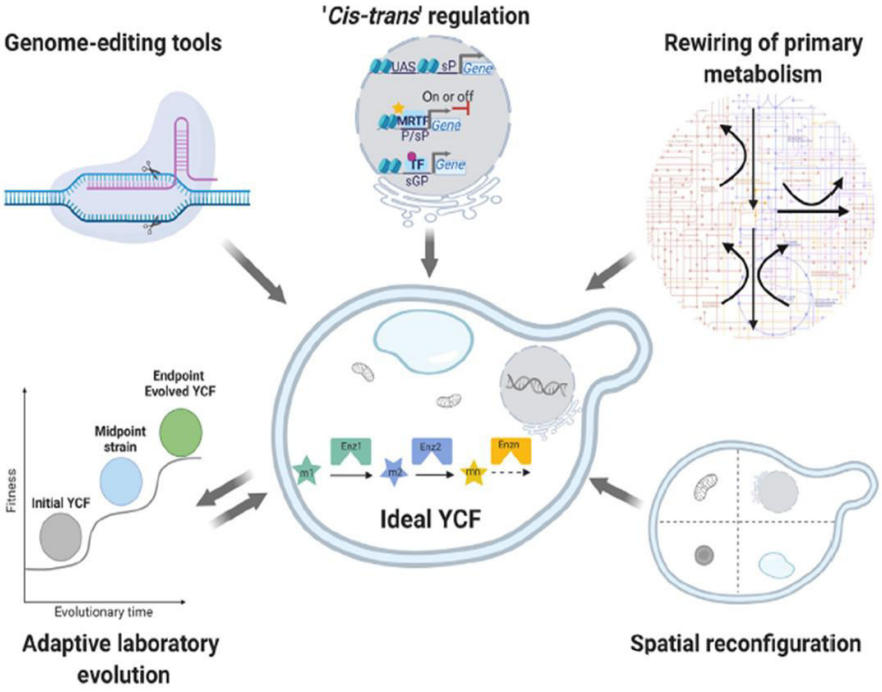
clivage de l'ADN. Le domaine TALE peut être conçu de façon à permettre une fixation sur la séquence d'ADN souhaitée. Le système CRISPR/Cas est un mécanisme de défense naturel des bactéries et des archées contre les éléments génétiques mobiles, en particulier les bactériophages (Barrangou, 2015). Cas9 est associée au système immunitaire adaptatif des bactéries, le système CRISPR (Jinek et al., 2012). Cette protéine est une endonucléase d'ADN guidée par un ARN guide (ARNg) capable de couper l'ADN double brin. L'utilisation de Cas9, contrairement aux systèmes ZFN et TALENS, ne nécessite que la conception d'un ARNg complémentaire à la séquence cible à cliver et non l'ingénierie chronophage d'un domaine TALE spécifique. CRISPR/Cas9 joue désormais un rôle crucial d'édition du génome dans l'ingénierie métabolique (Zhang et al., 2020). En effet, il permet de muter ou désactiver des gènes précis rapidement, mais aussi d'insérer des cassettes de gènes à un endroit défini du génome sans nécessiter l'intégration d'un marqueur de sélection. Le système CRISPR/Cas9 a ainsi été adapté pour l'édition du génome de la levure (**Figure 29**). Le principe d'édition du génome via le système CRISPR/Cas9 repose sur le clonage de deux vecteurs, le premier permettant l'expression de Cas9, le second contenant l'ARNg permettant de cibler une partie précise du génome. Le découplage de ces vecteurs permet de disposer de plasmides de plus petites tailles et de faciliter ainsi le clonage des ARNg. En plus des vecteurs Cas9 et ARNg, l'insertion dans la levure d'une séquence linéaire composée de bras homologues à la région génomique permet une insertion ciblée de la séquence d'ADN située entre les deux bras lors de la réparation de la zone clivée. La séquence linéaire peut contenir une séquence codante ou une série de codon stop permettant le knock-out du gène visé. Toutefois, certains éléments sont à prendre en compte pour l'utilisation de cette technique. Par exemple, le plasmide Cas9 ne doit pas contenir de promoteurs trop puissants, sous peine d'entraîner une toxicité chez la levure (Jakočiūnas et al., 2016). De plus, l'ARNg doit pouvoir se replier correctement afin d'assurer son interaction avec le Cas9. Désormais, la plupart des transferts de voies hétérologues en levure sont réalisés via le système CRISPR/Cas9.

Le choix du site d'intégration génomique est important dans le contrôle de l'expression des transgènes. Contrairement à la recombinaison homologue classique où l'intégration n'était pas aisément contrôlable, les intégrations ciblées de type CRISPR/Cas9 peuvent être facilement réalisées à un locus sélectionné du génome. Ainsi, l'utilisation du système CRISPR/Cas9 pour le transfert de gènes a donc permis l'exploitation de sites préférentiels du génome de la levure pour l'intégration de gènes hétérologues et définis comme "hot-spots". Onze premières zones génomiques pour l'insertion de gènes d'intérêt dans la levure ont été identifiées (**Figure 30**)

(Mikkelsen et al., 2012). Chaque site d'intégration est séparé des autres par des éléments génétiques essentiels au maintien de la croissance des cellules afin de conserver une certaine stabilité génétique. En effet, l'intégration d'éléments répétitifs tels que des promoteurs et des terminateurs identiques peut générer des recombinaisons incontrôlées lors de la transformation des souches. La séparation spatiale des hot-spots permet l'expression de copies multiples du même gène ou de promoteurs et de terminateurs identiques avec un risque moindre de recombinaison incontrôlée. Ces sites d'intégration permettent ainsi d'intégrer les gènes d'intérêt en dehors des locis associés aux marqueurs de sélection, de s'affranchir des effets du contexte génomique sur l'expression des enzymes hétérologues telle que la répression mais aussi, et surtout, de bénéficier d'un contexte local positif d'expression. Enfin, une fois la stratégie de bioproduction et le mécanisme de transfert en levure déterminés, plusieurs stratégies peuvent être employées pour optimiser les rendements de bioproduction des cellules usines de levure.

c) Outils et stratégies d'optimisation des cellules « usines » de levure

Nous avons vu que de nombreuses souches de levure modifiées sont en mesure de produire des MS mais très peu de ces souches sont capables de produire des MS à des niveaux compatibles avec une exploitation à l'échelle industrielle (**Tableau 1**). Des stratégies classiques existent dans le but d'améliorer les rendements de biosynthèse en levure, notamment pour optimiser les étapes limitantes des voies transférées. Ces approches incluent, par exemple, les modifications de composition des milieux de culture, l'ingénierie enzymatique tel qu'illustré par les travaux de Rabeharindranto et al. (2019) sur la voie de biosynthèse des caroténoïdes en levure, ou encore la multiplication des copies des gènes comme pratiqué par Kulagina et al. (2021c) pour la biosynthèse de l'AIM vindoline. Chaque approche peut ainsi lever des verrous d'activité enzymatique, mais également présenter des effets négatifs dont il est nécessaire de connaître ou prévoir les effets. C'est, en effet, le cas de l'augmentation du nombre de copies de gènes car l'accumulation d'un même transcrit peut avoir des effets répressifs sur l'expression génique, comme le RNA silencing qui inhibe l'expression des gènes via des interactions séquence-spécifiques. Ce phénomène, impliqué dans le maintien de l'intégrité des génomes pourrait, en effet, inhiber la transcription en cas d'accumulation trop importante d'un même



<u>Figure 31</u> : Illustration des différents outils de biologie de synthèse utilisés pour la construction et l'optimisation des cellules de levure usines. (Guirimand et al., 2021)

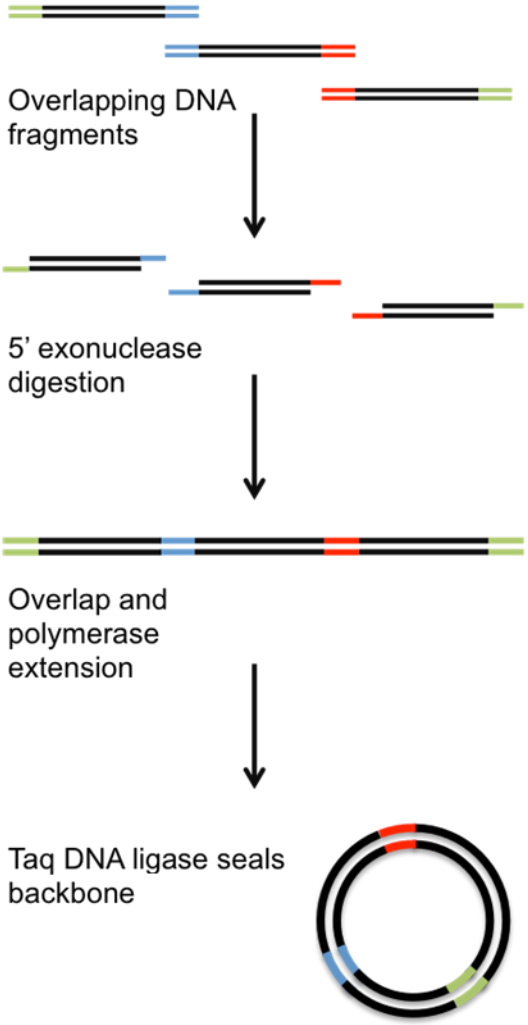
Les différents outils de biologie synthétique utilisés pour la construction et l'optimisation de cellules « usine » de levure sont représentés par les flèches en gris, comprenant les outils avancés d'édition du génome, les régulation transcriptionnelle (cis) et post-transcriptionnelle (trans), le recâblage du métabolisme primaire de la levure et l'évolution adaptative en laboratoire (ALE).
YCF : Yeast Cell Factory.

transcrit, et donc inhiber l'efficacité de l'étape enzymatique associée. Pour pallier à cette éventualité, le clonage de gènes codant des isoformes des enzymes de fonctions identiques mais disposant d'une séquence partiellement identique peut permettre d'augmenter le nombre de copies de gènes en s'affranchissant d'éventuels problèmes liés à l'accumulation d'un même transcrit.

De surcroît, des évolutions des approches existantes et de nouvelles stratégies ne cessent de se développer afin d'améliorer les rendements de synthèse dans les souches de levure « usines » (**Figure 31**). Parmi celles-ci, le développement d'outils d'édition du génome, la possibilité de réguler la transcription et la traduction, la régulation du métabolisme primaire ainsi que la capacité de recyclage des cofacteurs nécessaires aux enzymes recombinantes, la reconfiguration spatiale, le contournement des fuites métaboliques et, finalement, les stratégies adaptatives d'évolution sont autant de pistes permettant l'amélioration des rendements de bioconversion qui seront présentées dans cette partie.

1. Développement d'outils d'édition du génome et d'assemblage de fragments d'ADN multiples

Comme décrit précédemment, les outils d'édition du génome nouvellement développés facilitent l'intégration de gènes exogènes dans la levure sans nécessité de marqueurs (Zhang et al., 2020). Malgré cette évolution, ils ne permettent l'intégration que d'un ou d'un très faible nombre de fragments d'ADN par transformation. Or, le transfert en levure d'une nouvelle voie de biosynthèse nécessite l'insertion d'un nombre parfois conséquent de gènes: étapes enzymatiques plus ou moins nombreuses, augmentation du nombre de copies de gènes afin d'améliorer le flux métabolique, surexpression de gènes endogènes pour améliorer la production de cofacteurs. Comme présenté antérieurement, par exemple, l'augmentation du nombre de copies du P450 T16H et de l'*O*-méthyltransférase 16OMT spécifique de la voie de biosynthèse de la vindoline a permis de réduire le flux vers la synthèse de vindorosine indésirable dans la levure (Kulagina et al., 2021c ; Lemos-Cruz et al., 2021). En effet, les enzymes hétérologues T16H et Tabersonine-3-oxygénase (T3O) impliquées dans la voie de bioconversion de la vindoline sont en compétition pour la tabersonine, précurseur de la voie. Il est donc essentiel d'améliorer les outils d'édition du génome afin de permettre ces intégrations multiples et de travailler plus rapidement et de façon robuste sur l'optimisation des souches de



<u>Figure 32</u> : Fonctionnement de l'assemblage de Gibson pour la construction de plasmides multigéniques.

Les fragments d'ADN sont amplifiés par PCR de façon à disposer, à leur extrêmité, de séquences homologues avec les séquences voisines (vert, rouge ou bleu sur la figure). Les différents fragments obtenus sont ensuite digérés par une exonucléase afin d'obtenir des coupures simple brin franches permettant aux différents fragments chevauchants de se lier à l'aide d'une ligase.

levures usines. Pour cela, de nouvelles approches offrent désormais la possibilité d'intégrer de multiples gènes dans un même locus du génome de la levure via la construction de plasmides intégratifs comportant plusieurs gènes. La génération de ces constructions comportant plusieurs cassettes d'expression est facilitée, notamment, par la technique de fusion USER et l'assemblage de Gibson, permettant l'hybridation de fragments d'ADN par homologie des séquences situées aux extrémités des fragments d'ADN à assembler (**Figure 32**) (Geu-Flores et al. 2007 ; Gibson et al., 2009), ou encore la possibilité de synthétiser des gènes de synthèse. Il est, de plus, désormais possible d'intégrer plusieurs fragments d'ADN simultanément dans des chromosomes différents via le système CRISPR/Cas9 en utilisant conjointement plusieurs ARNg (Degreif et al., 2018). Ces approches sont composées d'un plasmide d'expression constitutive contenant différents ARNg et de l'endonucléase Cas9 pouvant être exprimée de façon constitutive ou inductible. Notamment, le système « CasEMBLR » mis en place par Jakočiunas et son équipe a permis l'assemblage et l'intégration en une unique étape de la voie des caroténoïdes contenant 15 gènes dans trois loci ciblés du génome de la levure (Jakočiūnas et al., 2018). Cette technique a également permis l'intégration de 10 gènes dans deux loci ciblés et la délétion de deux des gènes endogènes simultanément afin de construire une souche productrice de tyrosine (Jakočiūnas et al., 2018). De même, la méthode alternative CRISPR/Cas12a a permis l'insertion simultanée de trois gènes de la voie des β-carotènes dans le génome de la levure (Verwaal et al., 2018). Cas12a est une endonucléase d'ADN guidée par un ARNg de type II découverte dans plusieurs génomes bactériens et un génome archéen (Makarova et al., 2015).

2. Outils de régulation transcriptionnelle de l'expression génique

La régulation de l'expression génique au niveau transcriptionnel par l'utilisation de promoteurs et terminateurs améliorés est un levier puissant pour l'ingénierie de cellules « usines » de levure. Effectivement, la sélection de promoteurs appropriés est importante pour l'optimisation des rendements de synthèse. En effet, en plus de déterminer le niveau de transcription, le promoteur peut être régulé différemment en fonction des conditions de croissance de la levure (Sun et al., 2012). De nombreux promoteurs constitutifs de S. cerevisiae sont classiquement utilisés en ingénierie métabolique et s'avèrent être efficaces pour

l'expression de gènes hétérologues, notamment les promoteurs TEF1, TEF2 (codant des facteurs d'allongement translationnel EF-1 alpha), TPI1 (codant la triose phosphate isomérase), PGK1 (codant la 3-phosphoglycérate kinase), ACT1 (codant l'actine) et TDH3 (codant la glycéraldéhyde-3-phosphate déshydrogénase) (Partow et al., 2010). Les promoteurs constitutifs ne sont pas régulés et permettent une transcription continue du gène qui leur est associé, tandis que les promoteurs inductibles activent la transcription du gène qui leur est associé en réponse à la présence d'un composé ou d'une condition environnementale particulière. Il est, par exemple, possible de réguler la transcription via l'utilisation de promoteurs indépendants du glucose. En effet, de nombreux promoteurs natifs sont associés à la glycolyse et, de fait, moins efficaces en cas de carence en glucose (Peng et al., 2015). C'est pourquoi de nouveaux promoteurs ont été mis en place en modifiant le site de liaison de l'activateur transcriptionnel du gène Gcr1 lié à la glycolyse (Rajkumar et al., 2019), permettant une forte transcription génique dans un milieu pauvre en glucose. Ces promoteurs ont été validés dans une souche de levure productrice de vanilline-β-glucoside capable de produire ce MS pendant la phase éthanol durant laquelle le milieu est appauvri en glucose. Ces promoteurs peuvent ainsi être utilisés dans les souches de levure où la bioproduction a lieu en phase éthanol, lorsque le bioproduit est toxique pour la levure et que le découplage des phases de croissance et de production est nécessaire, ou lorsque plusieurs sources de carbones sont utilisées. De même, des promoteurs synthétiques forts ont été développés sur la base de promoteurs natifs via une modification de leurs sites de liaison, permettant une induction de la transcription dans des conditions acides (Rajkumar et al., 2016). Par exemple, le promoteur YGP1v7, construit par ajout d'un site de liaison aux sites de régulation transcriprionnelle Msn2/4p, Rap1p, Swi5p, Swi4p et Rlm1p du promoteur natif YGP1 codant une glycoprotéine, a permis une transcription 5 fois plus élevée en pH acide par rapport au promoteur natif. Plusieurs bibliothèques de promoteurs ont ainsi été mises en place en utilisant des séquences d'activation (Decoene et al. 2019) ou encore en modifiant la région centrale du promoteur située entre la boîte TATA et le site de démarrage de la transcription (Liu et al., 2020b) afin d'améliorer l'efficacité de transcription.

La force des promoteurs et des terminateurs conditionne l'efficacité de l'initiation et de la terminaison de la transcription. Le choix des promoteurs et des terminateurs est donc une étape importante afin de réguler l'activité des protéines hétérologues en levure. En outre, l'adéquation du couple promoteur-terminateur et de la force de chacun est déterminante pour l'optimisation de la transcription d'un gène hétérologue (D'Ambrosio et al., 2022). En effet, les promoteurs et les terminateurs interagissent au cours de la transcription, formant une boucle

Première lettre	Deuxième lettre U		Deuxième lettre C		Deuxième lettre A		Deuxième lettre G		Troisième lettre
U	UUU	Phénil-	UCU	sérine	UAU	tyrosine	UGU	cystéine	U
	UUC	alanine	UCC		UAC		UGC		C
	UUA	leucine	UCA		UAA	codons	UGA	codon stop	A
	UUG		UCG		UAG	stop	UGG	tryptophane	G
C	CUU	leucine	CCU	proline	CAU	histidine	CGU	arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	isoleucine	ACU	thréonine	AAU	asparagine	AGU	sérine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	lysine	AGA	arginine	A
	AUG	méthionine	ACG		AAG		AGG		G
G	GUU	valine	GCU	alanine	GAU	acide aspartique	GGU	glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	acide glutamique	GGA		A
	GUG		GCG		GAG		GGG		G

Figure 33 : Codons utilisés préférentiellement chez les plantes (Murray et al., 1989) **et la levure.** (Sharp et al., 1986)**.**

Les codons utilisés préférentiellement par les plantes sont indiqués en vert et ceux par *Saccharomyces cerevisiae* en orange.

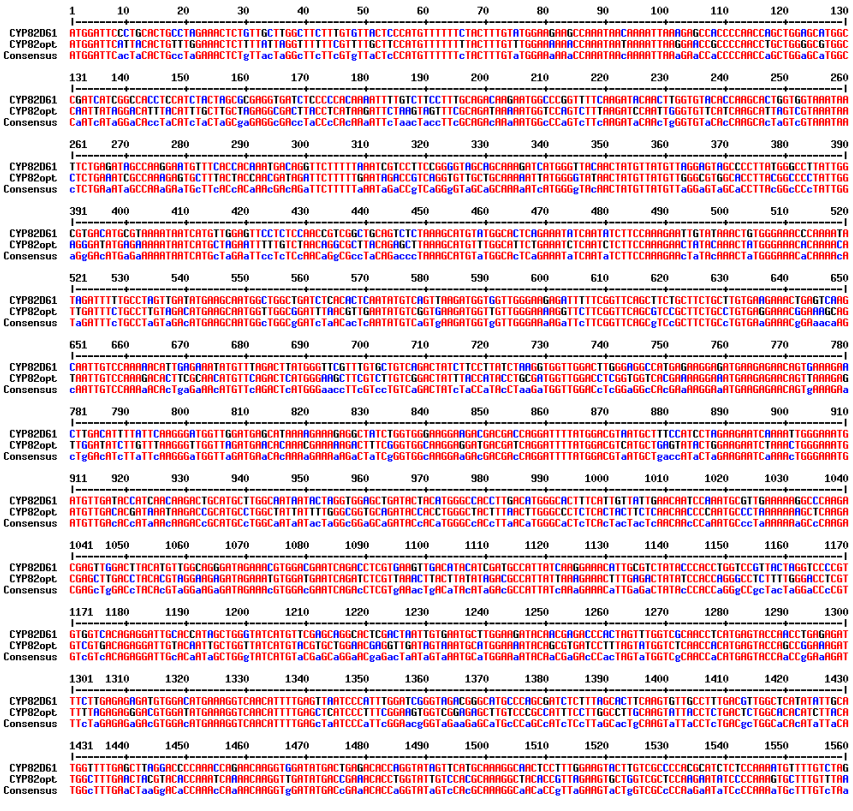


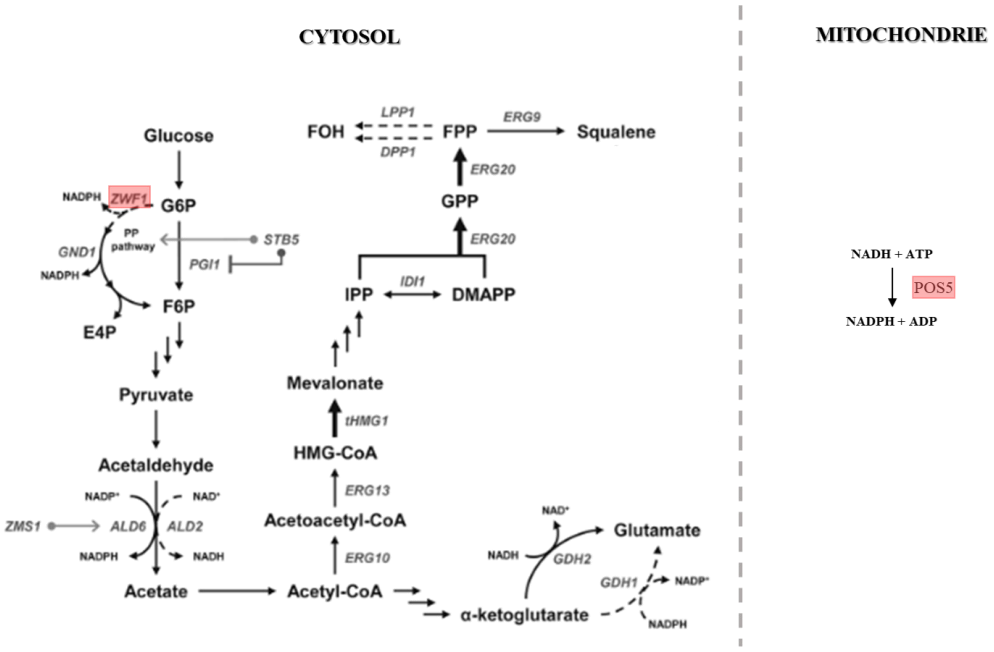
Figure 34 : Exemple de comparaison des séquences nucléotidiques du cytochrome P450 CYP82D61 natif de *Podophyllum hexandrum* et de sa version optimisée pour la levure.

Les nucléotides affichés en bleu correspondent aux différences entre les séquences. Les deux séquences disposent de 76% d'identité. La version optimisée a été mise en place via la plateforme web d'optimisation COOL (Codon Optimization OnLine).

génique capable d'accélérer la transcription (Al-Husini et al., 2020). Par exemple, l'amplitude d'expression en levure du gène optimisé pour la levure codant la citrine (yeCitrine) contrôlé par un même promoteur (TEFmut3) et différents terminateurs a été mesurée par fluorescence (Curran et al., 2013). Ces mesures ont mis en évidence une différence de fluorescence en fonction du terminateur utilisé, allant d'une fluorescence relative d'environ 0,6 pour le terminateur CYC1 (codant l'isoforme 1 du cytochrome C) à plus de 3 pour le terminateur PRM9 (codant la protéine membranaire régulée par les phéromones 9).

3. Outils d'optimisation de la traduction

L'optimisation de la capacité des ARNm issus des transgènes à être traduits est aussi un puissant levier en ingénierie métabolique. La dégénérescence du code génétique offre la possibilité de coder un même acide aminé à partir de plusieurs codons différents. Chaque cellule utilise préférentiellement un codon parmi ceux possibles pour la synthèse de chaque acide aminé, on parle de biais d'usage du code (Fox and Erill, 2010). Cette préférence dépend de l'organisme, mais aussi de l'origine de l'ARN (mitochondrie, noyau, chloroplaste). Ce biais d'usage du code influence l'efficacité de traduction et peut ainsi parfois altérer l'expression de protéines hétérologues chez la levure (Tuller et al., 2010). En effet, les ARN de transfert (ARNt), permettant la reconnaissance des codons par les ribosomes, sont présents en proportion variable dans les cellules, notamment, en de très faibles quantités lorsque les codons sont rares chez un organisme, causant alors un biais d'expression de ces codons rares. Ainsi, les codons contenus dans les séquences d'ADN provenant de plantes peuvent n'être que rarement utilisés dans un organisme hôte tel que la levure (**Figure 33**) (Gustafsson et al., 2004). C'est pourquoi l'optimisation des codons est couramment utilisée afin d'améliorer l'expression de gènes hétérologues dans le contexte d'ingénierie métabolique. Cette optimisation consiste à remplacer les codons considérés comme rares dans l'organisme hôte par des codons plus fréquents dans la séquence d'ADN codant la protéine hétérologue (**Figure 34**). Pour faciliter l'optimisation des codons dans la levure, des algorithmes basés sur les fréquences d'utilisation des codons ont été mis en place afin de générer plusieurs versions d'une séquence codante (Lanza et al., 2014). Plusieurs séquences de gènes hétérologues ont ainsi été optimisées pour la levure afin d'en améliorer la traduction, et ainsi les rendements de biosynthèse de molécules d'intérêt. Par exemple, le clonage dans *S. cerevisiae* d'une version optimisée du gène codant la catéchol 1,2-



<u>Figure 35</u> : Voie de biosynthèse du squalène chez *S. cerevisiae*. (Kim et al., 2018)

Les enzymes en rouge correspondent aux enzymes endogènes surexprimées (ZWF1: glucose-6-phosphate déshydrogénase; POS5: NADH kinase).

dioxygénase d'*Acinetobacter baylyi* disposant de 77% d'identité avec la séquence native a permis d'augmenter par 2,9 les rendements de synthèse de l'acide muconique, un précurseur de polymère (Lanza et al., 2014).

Parmi les autres éléments employés pour optimiser la traduction de protéines hétérologues en levure, l'optimisation du site d'initiation de la traduction représente un outil puissant. En effet, chez les eucaryotes, la traduction est initiée lorsque le codon initiateur AUG situé à l'extrémité 5' de l'ARNm est reconnu par le complexe composé de l'ARNt initiateur porteur de l'acide aminé méthionine et de la petite sous-unité ribosomique. Ce complexe scanne la partie 5' non traduite (5'-UTR) de l'ARNm à la recherche d'un codon AUG, et certaines séquences situées en amont peuvent affecter la traduction, y compris les séquences d'initiation de la traduction. Une étude a permis de caractériser et classer différents sites d'initiation de la traduction en fonction de leur impact sur l'expression des protéines dans les cellules de levure (Petersen et al. 2018). En particulier, le site d'initiation TCGGTC situé en amont du codon AUG a permis d'augmenter de 50% l'expression de la GFP par rapport au site d'initiation consensus.

4. Régulation du métabolisme primaire et recyclage des co-facteurs

La plupart des voies de biosynthèse de MS intègrent des enzymes nécessitant des cofacteurs, comme, par exemple, les cytochromes P450 dont l'activité nécessite le NADPH, les *O*-méthyltransférases dépendantes de la SAM, ou encore certaines acyltransférases dépendantes de l'acétyl-coenzyme A (CoA). La production native de ces cofacteurs par la levure peut représenter un facteur limitant dans l'ingénierie métabolique, notamment lorsque de nombreuses enzymes hétérologues sont surexprimées dans des cellules hôtes. Il est alors important d'améliorer la synthèse ou la régénération de ces cofacteurs dans la levure en optimisant le métabolisme primaire de la cellule hôte et, par conséquent, en modifiant les flux métaboliques natifs (Li et al., 2018). En particulier, plusieurs méthodes ont été mises en place pour la régénération du NADPH. Parmi celles-ci, la surexpression des gènes endogènes ZWF1, codant la glucose-6-phosphate déshydrogénase et POS5, codant la NADH kinase, impliqués dans la régénération du NADPH chez *S. cerevisiae* a permis une augmentation de 27,5 fois la synthèse de squalène, un intermédiaire de la voie de l'ergostérol, synthétisé par la HMG-CoA

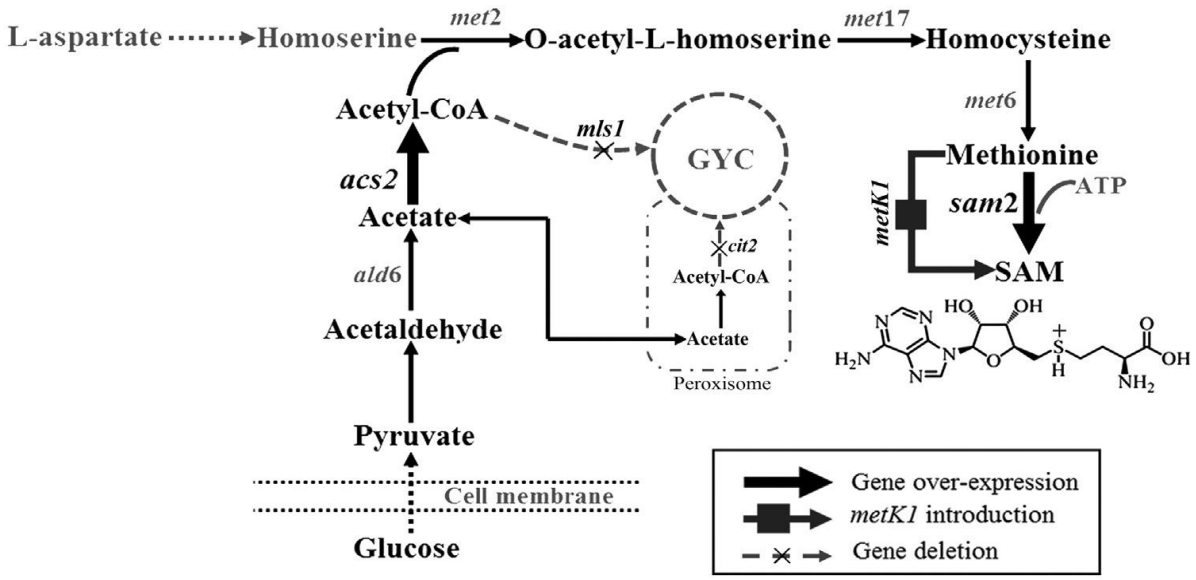


Figure 36 : Biosynthèse de S-Adénosyl-L-méthionine (SAM) dans une souche *S. cerevisiae* recombinante. (Chen et al., 2016)

Abréviations de gènes et de protéines : ald6 : aldéhyde déshydrogénase NADP-dépendante ; acs2 : acétyl-CoA synthase; met2 : homosérine O-acétyltransférase ; met17 : O-acétyl-homosérine amino carboxypropyltransférase ; met6 : méthionine synthase ; sam2 : méthionine adénosyltransférase ; metK1 : méthionine adénosyltransférase de *L. infantum* ; mls1 : malate synthase cytosolique et cit2 : citrate synthase peroxysomale.

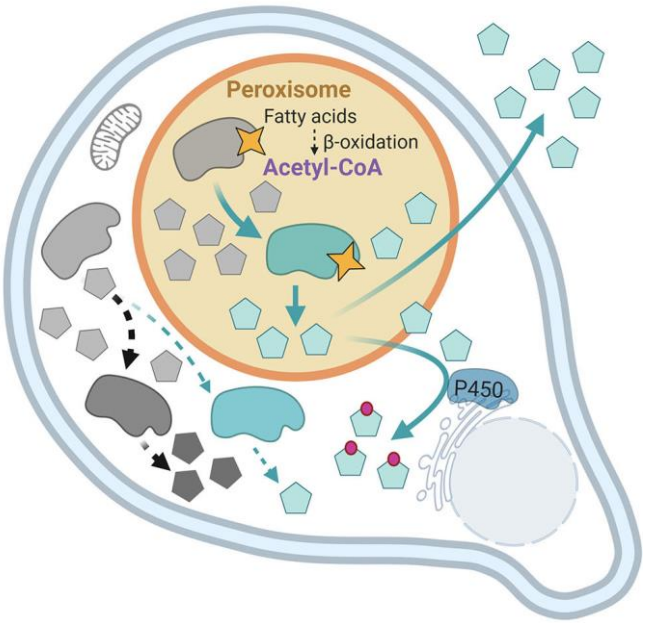
réductase HMG1 et la squalène synthase ERG9, toutes deux consommatrices de NADPH (Paramasivan and Mutturi, 2017). En effet, ZWF1 catalyse la première étape de la voie des pentoses phosphates connue pour produire du NADPH et POS5 est impliquée dans la synthèse du NADPH mitochondrial (**Figure 35**). D'autre part, la substitution de l'aldéhyde déshydrogénase ALD2 endogène génératrice de NADH par l'aldéhyde déshydrogénase ALD6 génératrice de NADPH a permis d'obtenir une bioproduction 11 fois plus élevée de protopanaxadiol, un ginsénoside médicinal, dans une cellule « usine » de levure (Kim et al., 2018).

A l'instar de la stratégie de bioproduction du resvératrol précédemment évoquée, et afin d'améliorer la production de SAM chez *S. cerevisiae*, une stratégie combinant l'optimisation du métabolisme de l'acétyl-CoA et le clonage de la méthionine adénosyl transférase (METK1) de *Leishmania infantum* permettant d'éliminer l'inhibition de la SAM synthase a été mise en place (Chen et al., 2016). Le clonage de cette enzyme, couplé à la surexpression de l'acétyl-CoA synthase ACS2 et la suppression de la malate synthase MLS1 impliquées dans le métabolisme de l'acétyl-CoA ont permis une production 9,18 fois plus élevée de SAM par la souche de levure recombinante cultivée en bioréacteur (**Figure 36**).

Des approches combinées de modification du flux métabolique primaire peuvent ainsi permettre d'améliorer la synthèse de cofacteurs chez *S. cerevisiae*. Récemment, des souches de *S. cerevisiae* ont été mises en place afin de recycler simultanément *in vivo* les cofacteurs $FADH_{2,}$, utilisé principalement par les déshydrogénases, SAM et NADPH (Chen et al., 2022). Ces travaux ont fait l'objet de la rédaction d'un commentaire publié dans la revue Frontiers in Bioengineering and Biotechnology et présenté en **Annexe II** (Perrin et al., 2022).

5. Reconfiguration spatiale

Les voies de biosynthèse de MS d'intérêt présentent généralement une compartimentation complexe *in planta*, à la fois à l'échelle cellulaire et subcellulaire. C'est le cas, par exemple, de la voie de biosynthèse entièrement caractérisée des AIM vincristine et vinblastine chez *C. roseus* qui présente une compartimentation unique avec l'intervention de trois types cellulaires et l'implication de cinq compartiments subcellulaires distincts (Courdavault et al., 2014 ; Guirimand et al., 2020). Lors du transfert de ces voies en levure, il



<u>Figure 37</u> : Utilisation du peroxysome pour limiter la concurrence entre les substrats. (Kulagina et al., 2021b)

L'adressage d'enzymes natives/hétérologues aux peroxysomes de *Saccharomyces cerevisiae* empêche la compétition pour les substrats communs entre les voies métaboliques hétérologues et natives, ce qui se traduit par un meilleur rendement. Les formes grises représentent les enzymes natives (gris foncé : enzyme concurrente, gris clair : enzyme commune fournissant le substrat) et les flèches en pointillés noires montrent les voies métaboliques natives. Les formes vertes représentent les enzymes hétérologues et les flèches vertes indiquent les voies de biosynthèse hétérologues. Les étoiles jaunes représentent le ciblage peroxysomal via l'ajout du signal PTS1/ePTS1. Les pentagones bleu clair représentent le produit/intermédiaire ; les pentagones gris représentent les composés synthétisés nativement. Dans certains cas, l'approche de double localisation est bénéfique pour la production. De plus, la proximité du RE avec les peroxysomes est potentiellement bénéfique pour les enzymes P450.

est possible de préserver l'organisation/l'adressage des enzymes à l'identique ou de procéder à des reconfigurations de ces voies. Bien souvent, le maintien de l'organisation initiale des protéines hétérologues est suffisant pour obtenir des rendements de biosynthèse élevés. Toutefois, une relocalisation cellulaire des enzymes hétérologues peut être avantageuse dans les souches à faibles rendements de conversion (Huttanus and Feng, 2017). En effet, ces quantités infimes de bioproduction sont parfois liées à des conditions physiochimiques non optimales pour les enzymes hétérologues ou à une accumulation d'intermédiaires biosynthétiques toxiques. Dans ces conditions, la reconfiguration de la compartimentation intracellulaire des voies de biosynthèse est une approche pouvant être utilisée (Kulagina et al., 2021b). De plus, cette approche permet de favoriser les flux métaboliques soit en regroupant toutes les enzymes dans un seul organite, soit en se rapprochant des conditions *in planta* par isolation des enzymes dans des organites spécifiques.

En particulier, le peroxysome est peu à peu devenu un organite attractif pour cette approche puisqu'il dispose naturellement d'une forte concentration protéique, suggérant que cet organite est capable de maintenir active une grande quantité de protéines (Sibirny, 2016). Le peroxysome est un organite dépourvu d'ADN impliqué dans la β-oxydation des acides gras. Il dispose d'une taille particulièrement modelable et d'une machinerie d'import protéique très efficace. La perméabilité de sa membrane permet, en outre, la diffusion de métabolites d'une large gamme de tailles et de structures. L'identification de séquences peptidiques ciblant le peroxysome permet désormais d'orienter très facilement les enzymes hétérologues vers cet organite. En effet, l'ajout de la séquence d'adressage au peroxysome PTS1 à l'extrémité C-terminale d'une protéine permet une relocalisation efficace des enzymes hétérologues vers le peroxysome des levures (Lee et al., 2009). L'adressage des enzymes hétérologues cytoplasmiques vers un organite permet de créer une proximité plus élevée entre les protéines et les cofacteurs qui leur sont nécessaires, en particulier l'acétyl-CoA spécialement concentrée dans le peroxysome. Récemment, il a été démontré que l'adressage des enzymes de la voie de synthèse du squalène au peroxysome permettait à la levure de stocker plus facilement le squalène produit (**Figure 37**) (Liu et al., 2020a). En effet, la compartimentation des enzymes hétérologues couplée à une fermentation en bioréacteur de la souche a permis d'obtenir 11 g/L de squalène, concentration la plus élevée jamais rapportée. En outre, le transfert au peroxysome de certaines enzymes hétérologues peut permettre de séquestrer les intermédiaires toxiques pour la levure, limitant ainsi leur effet néfaste sur la croissance cellulaire.

Comme décrit précédemment, la plupart des voies de biosynthèses végétales comportent des cytochromes P450 ancrées à la membrane du RE. Le RE est un organite lié à la membrane du noyau qui permet la synthèse des protéines et des lipides, mais qui est aussi impliqué dans le transfert de vésicules vers l'appareil de Golgi où « mûrissent » les protéines. Une augmentation de la taille de cet organe pourrait permettre, non seulement à un nombre plus élevé de cytochromes P450 de s'y ancrer, mais aussi d'améliorer la synthèse des protéines. La mutation du gène codant la phosphatase d'acide phosphatidique PAH1 entraîne un agrandissement de la taille du RE, augmentant ainsi la production de cytochromes P450 hétérologues telles que CYP716A12, CYP72A68, CYP72A67 et la CPR MtCPR1 de *Medicago truncatula* impliquées dans la synthèse de saponines triterpéniques (Arends et al., 2017). De façon remarquable, l'expansion du RE causée par la mutation du gène codant PAH1 a permis d'augmenter jusqu'à 16 fois les rendements de bioconversion de différents saponines triterpéniques. De même, la surexpression du gène codant le facteur de transcription pour la biosynthèse de lipides INO2 a permis d'accroître suffisamment la taille du RE pour augmenter de 8 fois la bioproduction de protopanaxadiol via l'amélioration de l'activité du cytochrome P450 protopanaxadiol synthase dans une souche de *S. cerevisiae* (Kim et al., 2019).

6. Contournement des fuites métaboliques

Lors du transfert de voies de biosynthèse, la formation de sous-produits indésirables générés soit par des voies endogènes concurrentes, soit par les enzymes hétérologues elles-mêmes peut réduire drastiquement les rendements de bioconversion des cellules « usine » de levure vers le produit cible. Afin de pallier ces éventuels problèmes, plusieurs stratégies peuvent être mises en place. La suppression des voies endogènes secondaires concurrentes est une stratégie couramment employée pour limiter les fuites métaboliques (Cravens et al., 2019). Par exemple, *S. cerevisiae* métabolise le géraniol, un intermédiaire de la voie de biosynthèse hétérologue de l'AIM strictosidine (Brown et al., 2015). La délétion des gènes endogènes ATF1 codant l'alcool acétyltransférase et l'old yellow enzyme OYE2 codant la NADPH-oxydoréductase a réduit la métabolisation de cet intermédiaire par la levure, permettant ainsi de multiplier par 6 la bioproduction de strictosidine. En effet, les gènes codant ATF1 et OYE2 sont impliqués dans l'acétylation des terpénols et dans la réduction du géraniol en citronellol respectivement (Steyer et al., 2013). De même, la bioproduction du terpénoïde népétalactone

par *S. cerevisiae* a été augmentée considérablement après délétion de deux homologues endogènes des old yellow enzymes OYE2 et OYE3 (Davies et al., 2021). En effet, le géraniol étant un intermédiaire de la voie de biosynthèse du népétalactone qui conduit à la strictosidine notamment, l'inhibition de la métabolisation de cette molécule par la levure, en plus de la multiplication du nombre de copies des gènes hétérologues, a permis de produire 3,10 mg/L de népétalactone. Finalement, la réorientation de la voie endogène d'Ehrlich, impliquée dans la transformation des aldéhydes toxiques pour la levure en acides ou en alcools via l'inactivation de l'aldéhyde réductase ARI1, l'alcool déshydrogénase ADH6, la réductase putative YPR1, l'aldéhyde réductase YDR541C, l'aryl-alcool déshydrogénase putative AAD3, la 3-méthylbutanal réductase GRE2 et la déshydrogénase HFD1 toutes impliquées dans la détoxification des aldéhydes a permis d'augmenter les rendements de synthèse de divers fragments structurels de tétrahydroisoquinoline (THIQ) en levure (Pyne et al., 2020). En effet, la voie d'Ehrlich permet une conversion rapide du 4-hydroxyphenyl acétaldéhyde nécessaire à la synthèse de THIQ.

Une reconfiguration spatiale de la voie de biosynthèse peut aussi permettre de contourner la formation de sous-produits indésirables causée par les enzymes hétérologues elles-mêmes. En effet, il arrive que deux enzymes hétérologues soient en compétition pour un même substrat. Dans ce cas, les taux de bioconversion pour la bioproduction du MS d'intérêt seront plus élevés si le substrat et l'enzyme qui le catalyse sont spatialement proches, notamment dans un même organite. De plus, le métabolisme de la levure peut varier en fonction de son environnement, notamment le contrôle des enzymes, le taux d'absorption du substrat, l'excrétion du produit et la croissance des cellules varient en fonction du mode de culture des levures en bioréacteur (continue ou discontinue) (Wang and Hatzimanikatis, 2006), des fuites métaboliques pourraient donc intervenir en cas de stress des cellules.

7. Stratégies adaptatives d'évolution

Finalement, il arrive que certains intermédiaires ou produits des voies de biosynthèse végétales soient toxiques pour la levure et non séquestrables dans un organite, ou que certaines enzymes dépensent des éléments nécessaires à la croissance des levures comme c'est le cas des méthyltransférases consommatrices de cystéine. Dans ces conditions, il peut être intéressant de coupler les stratégies d'ingénierie métabolique présentées précédemment à des stratégies

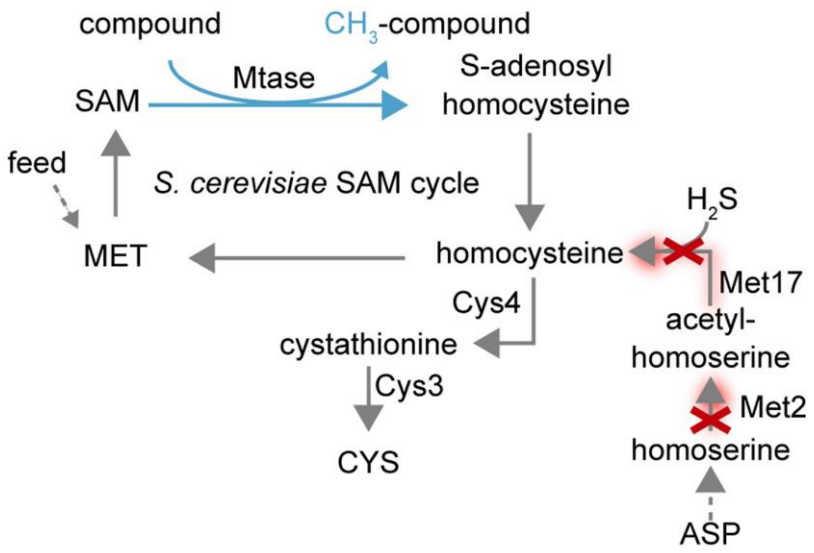


Figure 38 : lllustration de la conception de sélection couplée à la croissance chez *S. cerevisiae*. (Luo et al., 2019)

Les réactions natives sont colorées en gris et les réactions hétérologues sont colorées en bleu. Les lignes pointillées représentent de multiples réactions enzymatiques. Les croix rouges indiquent la suppression du gène. Les lignes ombrées rouges indiquent des réactions inactives après la suppression du gène (croix rouge). Les cofacteurs, les sous-produits et les co-substrats sont omis.
ASP : acide aspartique ; CYS : cystéine; Cys3 : cystathionine-γ-lyase; Cys4 : cystathionine-β-synthase; MET : méthionine ; Mtase : méthyltransférase ; SAM : S-adénosy-L-méthionine.

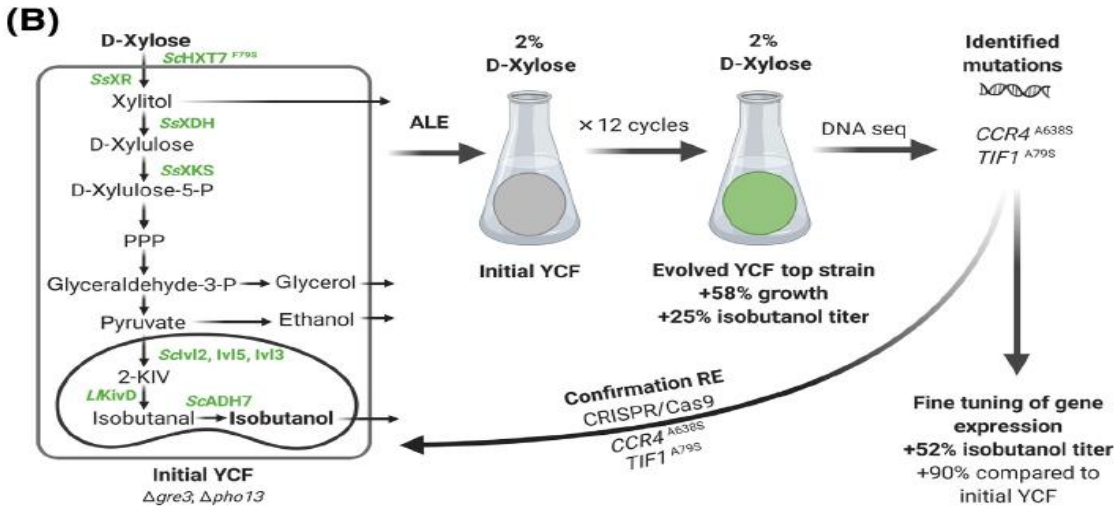


Figure 39 : Couplage de l'ingénierie métabolique classique et d'un stratégie adaptative d'évolution pour la synthèse d'isobutanol chez *S. cerevisiae*. (Promdonkoy et al., 2020; Guirimand et al., 2021)

La stratégie d'évolution a été menée en présence de 2 % de D-xylose comme seule source de carbone sur 12 cycles de culture cellulaire. La voie a été simplifiée, et les cofacteurs et certaines étapes ont été omis pour plus de simplicité.
Les enzymes en vert ont été surexprimées.
YCF : cellule usine de levure ; RE : rétro-ingénierie ; HXT7 : Transporteur de glucose à haute affinité avec une mutation ponctuelle F79S ; XR : xylose réductase ; XDH : xylitol déshydrogénase ; XKS : xylulokinase ; Ilv2 acétolactate synthase ; Ilv5 : acétohydroxyacide réductoisomérase ; Ilv3 : dihydroxyacide déshydratase ; kivD : cétoacide décarboxylase mitochondriale ; ADH7 : alcool déshydrogénase mitochondriale.

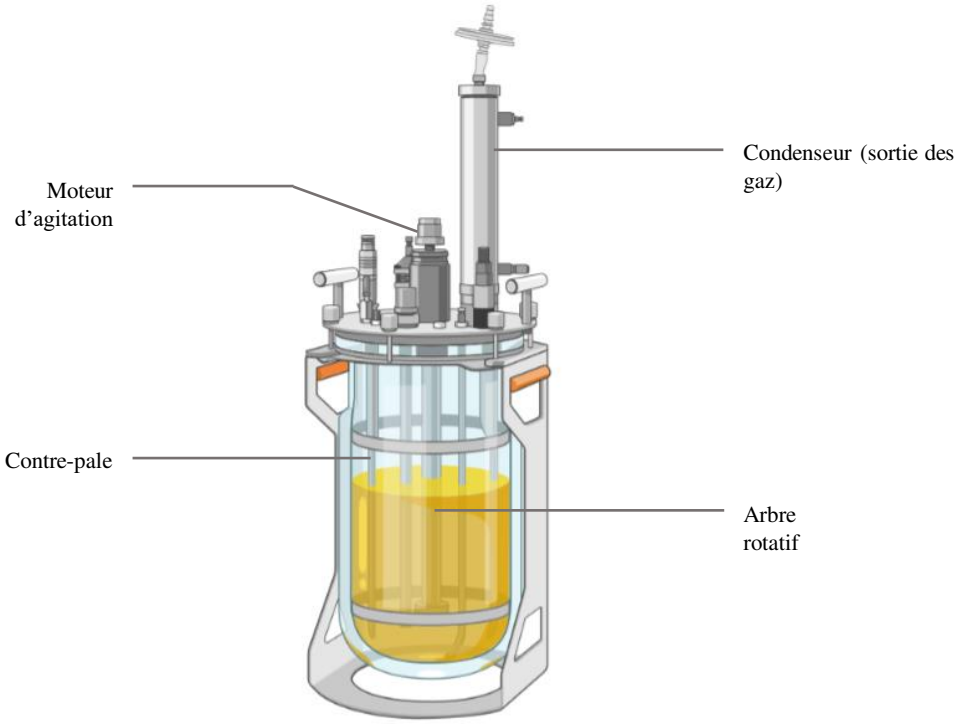
adaptatives d'évolution. En effet, ces stratégies reposent sur l'amélioration génétique de la levure et son adaptation aux contraintes générées par l'ingénierie métabolique. Lors de l'évolution adaptative, la levure, soumise à une pression de sélection environnementale, va augmenter la fréquence des allèles bénéfiques à sa survie dans ce nouvel environnement et diminuer la fréquence des allèles non nécessaires ou délétères.

Une stratégie adaptative d'évolution en laboratoire a permis, par exemple, la sélection d'une souche de levure mutante hyperproductrice de l'acide aminé cystéine via blocage de la biosynthèse d'homocystéine par inactivation des gènes codant l'homosérine *O*-acétyltransférase (MET2) et l'*O*-acétylhomosérine sulfhydrylase (MET17) (Luo et al., 2019). Afin d'éliminer les méthyltransférases endogènes SAM-dépendantes concurrentes, les gènes codant la phosphatidyléthanolamine méthyltransférase (CHO2) et la phospholipide méthyltransférase (OPI3) ont été inactivés. Les problèmes de croissance liés à l'activité des méthyltransférases de type *N* et *O* sont donc limités dans cette souche (**Figure 38**) (Luo et al., 2019). La sélection de cette souche a permis de doubler l'activité des méthyltransférases hétérologues. De même, l'évolution adaptative a été mise en place sur une souche de *S. cerevisiae* afin d'améliorer ses capacités à utiliser le D-xylose (Promdonkoy et al., 2020). Le séquençage du génome des souches évoluées a permis d'identifier une mutation ponctuelle dans le gène codant le récepteur de chimiokine de type 4 CCR4 et dans le gène codant l'intermédiaire de transcription TIF1. Ces mutations ont permis d'améliorer la croissance des souches de 23% et 14% respectivement. Cette stratégie couplée à la surexpression d'enzymes endogènes a permis d'augmenter les rendements de biosynthèse du biocarburant avancé isobutanol de 110% par rapport à la souche non-évoluée (**Figure 39**). La stratégie adaptative d'évolution peut donc permettre d'augmenter les rendements de biosynthèse en optimisant génétiquement la levure face aux contraintes générées par l'ingénierie métabolique, mais cet outil ne peut pas être mis en place dans la même échelle de temps que les actions proposées dans les parties précédentes.

8. Perspectives d'optimisation complémentaires des cellules « usines » de levure

L'ingénierie métabolique conduisant à la mise en place de cellules « usines » de levure productrices de MS d'intérêt, difficilement synthétisables par voie chimique, est désormais une approche couramment utilisée. Cependant, beaucoup de défis restent à relever, autant dans la

compréhension des mécanismes de régulation du métabolisme de *S. cerevisiae*, que dans l'anticipation du comportement des enzymes hétérologues. Bien que certains verrous soient limitants pour la synthèse de divers MS, de plus en plus de stratégies permettent de pallier aux goulots d'étranglement. Ces stratégies, en constante évolution, peuvent être combinées pour améliorer les rendements de biosynthèse et la mise en place de cellules « usines » de levure capables de produire des MS d'intérêt à échelle industrielle. La levure représente un châssis cellulaire très robuste mais est un organisme limitant pour la synthèse de certaines molécules comme les dérivés halogénés. Or, l'halogénation de certains composés, tels que la chloration ou la fluoration, représente un intérêt majeur dans l'industrie pharmaceutique puisque ces composés présentent des caractéristiques pharmaceutiques améliorées (Runguphan et al., 2010). Par exemple, la dexaméthasone, un dérivé fluoré de l'hydrocortisone, est 40 fois plus puissante que l'hydrocortisone et beaucoup plus stable *in vivo* (Guirimand et al., 2021). Toutefois, l'halogénation chimique des composés est un processus complexe, de faible efficacité (Vigalok, 2015) et le développement de toute solution alternative s'avérerait d'intérêt. Dans ce contexte, l'utilisation de cellules usines pour réaliser ces modifications est une option de choix et la modification du métabolisme primaire de la levure permettant de faciliter la synthèse de ce type de composés est donc un défi majeur à relever. De premiers exemples confirment la faisabilité de ce type d'approche. En particulier, la souche de levure productrice de noscapine est capable de produire des dérivés noscapinoïdes halogénés lorsque des substrats halogénés tels que la 3-fluoro-tyrosine, la 3-chloro-tyrosine, ou encore la 3-iodo-tyrosine sont ajoutés au milieu de culture (Li et al., 2018). Dans cette souche, la voie endogène de synthèse de la tyrosine a été modifiée pour orienter le flux métabolique vers la production de cette molécule essentielle à la synthèse de noscapine. La biosynthèse de composés halogénés pourrait, en outre, être réalisée via l'expression et l'ingénierie d'enzymes telles que des halogénases. En effet, de telles enzymes ont déjà été caractérisées, notamment, la protéine 2 liée à la tryptophane aminotransférase TAR2 impliquée dans la voie de synthèse de l'auxine halogénée acide 4-chloroindole-3-acétique chez le pois (Ross et al., 2012), ou encore la 7-L-tryptophane halogénase dépendante de la flavine RebH issue d'une espèce bactérienne d'actinomycète capable de chlorer la tryptamine chez *C. roseus* (Glenn et al., 2011). Ces halogénases pourraient être exprimées en levure afin de produire des MS halogénés. Ce type d'approche pourra alors être généralisé via l'identification ou la modification de nouvelles halogénases capables de travailler sur des précurseurs de MS d'intérêt.



<u>Figure 40</u> : Schéma d'un bioréacteur pour l'expérimentation en laboratoire.

Le moteur d'agitation permet d'homogénéiser et d'oxygéner le milieu de culture grâce à la rotation de l'arbre souvent composé de turbines. Le condenseur permet une sortie des gaz, notamment le CO_2 produit au cours de la fermentation. Les contre-pales permettent de « casser » les bulles d'air formées par les turbines et donc participent à l'aération du milieu de culture. Il arrive que les bioréacteurs soient composés d'un diffuseur d'air pour augmenter les capacités d'aération.

d) Optimisation de la conduite de culture de levures par le génie des procédés

Afin de produire des MS d'intérêt à une échelle industrielle, les cellules « usines » de levure sont cultivées en bioréacteur de gros volumes (jusqu'à 100 000 L). Ces bioréacteurs permettent aux levures de se multiplier et de synthétiser la molécule d'intérêt sous des conditions de culture entièrement contrôlées. Pour cela, une montée progressive en échelle est essentielle afin de contrôler les conditions de culture et de vérifier le comportement de la souche modifiée en volume croissant. En effet, les tests permettant la mise en place de cellules « usines » de levure en laboratoire sont généralement réalisés dans de petits volumes de culture (de 100µL pour des tests en tube à 100 mL pour des tests en Erlenmeyer). La montée en échelle sur des bioréacteurs de petite contenance (entre 1,5 L et 10 L) permet de vérifier le comportement de la souche et d'adapter la conduite de la culture afin d'optimiser les rendements de biosynthèse. Les volumes sont ensuite implémentés à une échelle pilote (bioréacteurs de 10 à 100 L), puis à l'échelle industrielle (10 000 à 100 000 L). Les conditions de culture en bioréacteurs sont moins exigeantes avec les micro-organismes tels que les levures, les bactéries ou les virus qu'avec les cellules de mammifères.

Les bioréacteurs sont composés d'un réservoir (en verre pour l'expérimentation en laboratoire, en acier inoxydable pour l'industrie ou pour ceux de très grande contenance) qui est le contenant de la culture, et de systèmes de contrôle des conditions de fonctionnement du bioréacteur, comme le thermomètre, le débitmètre, le sonde $\underline{O_2}$, le pH mètre, …, d'échangeurs de chaleur (réchauffement/refroidissement), d'aération et d'agitation (homogénéisation et oxygénation) (**Figure 40**). Les bioréacteurs peuvent fonctionner selon trois modes: le mode discontinu (ou batch), le mode semi-continu (ou fed-batch) et le mode continu (ou chémostat). Le mode batch consiste en l'intégration de la culture et de tous les éléments nécessaires à la culture au début du bioprocessus sans ajout ni prélèvement jusqu'à la fin. Le mode fed-batch, souvent employé pour la bioproduction de MS d'intérêt par les levures, consiste en une intégration initiale de la culture dans la cuve, puis en une intégration continue des substrats nécessaires à la culture (sucre, substrat, etc…) comme dans le cas des bioconversions. Dans ce mode, aucun prélèvement n'est réalisé avant la fin du bioprocessus. Finalement, le mode chémostat consiste en l'intégration continue de culture et des substrats nécessaires et en un prélèvement continue de façon à garder un volume constant dans le réservoir.

A titre d'exemple, citons le cas de la synthèse de l'AIM vindoline. Initialement de 26,8 mg/L, avec la souche de levure « usine » optimisée, elle a été augmentée à 34,6 mg/L via l'optimisation de l'apport en nutriments et du pH dans le milieu de culture à petite échelle (Erlenmeyer). Les rendements de bioconversion obtenus après une première culture en bioréacteur de 2 L de la souche ont été deux fois inférieurs à ceux obtenus dans les expériences à petite échelle. Ainsi, la conduite de la culture en bioréacteur a dû être optimisée notamment par une alimentation continue des substrats de croissance et une alimentation contrôlée et continue du précurseur tabersonine. Ces premières modifications ont conduit à une synthèse de 266 mg/L de vindoline après 96 heures de culture (Kulagina et al., 2021c). Cet exemple illustre l'importance de l'étude et du contrôle de la conduite de la culture dans l'augmentation des taux de biosynthèse de MS d'intérêt.

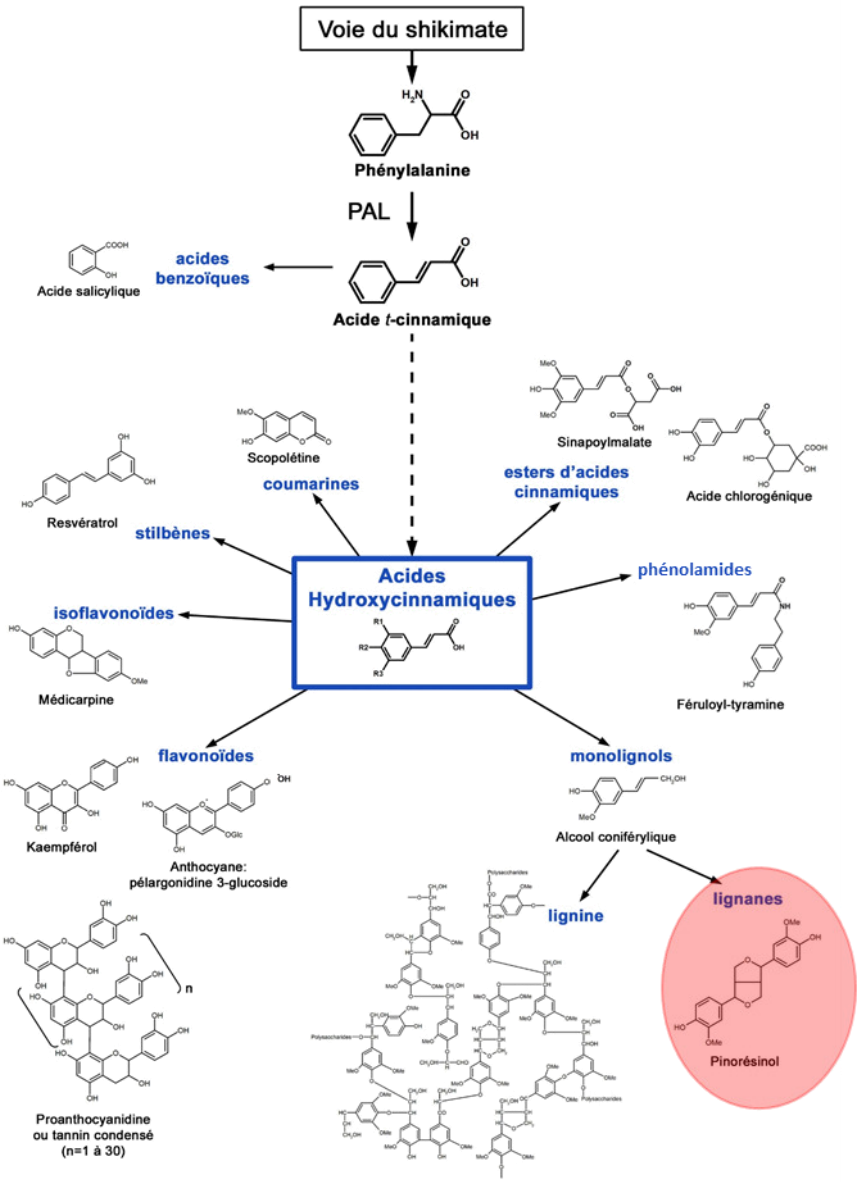


Figure 41 : Principales familles de composés dérivant du métabolisme des phénylpropanoïdes. (Besseau, non publié)

PAL : phénylalanine ammoniac lyase

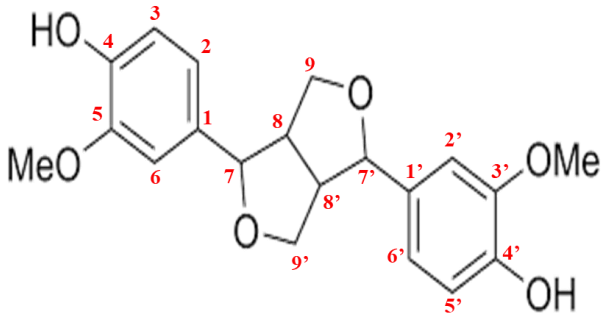


Figure 42 : Base de la nomenclature des lignanes issues du couplage de deux alcool coniféryliques par une liaison β-β.

Couplage de deux alcool coniféryliquespermettant la synthèse du pinorésinol. La numérotation des carbones est représentée en rouge. La première unité phénylpropane est numérotée de 1 à 6, le groupe propyle est numéroté de 7 à 9 à partir du cycle benzénique. Les nombres sont suivis d'un prime dans la deuxième unité. Les deux unités sont liées entre les positions 8 et 8'.<

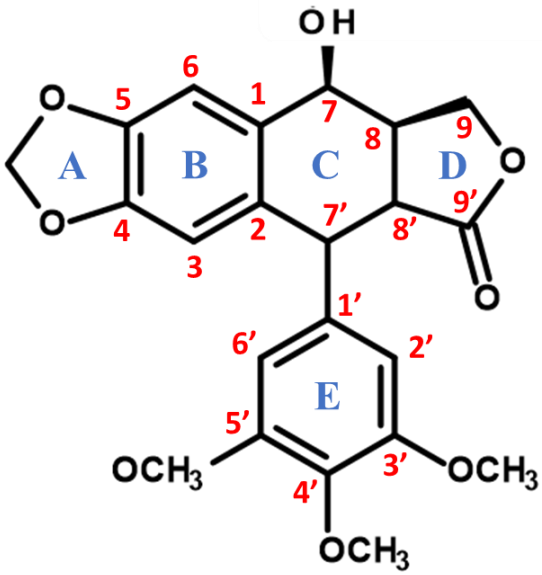
Chapitre II : Les vertus pharmaceutiques de la podophyllotoxine

La podophyllotoxine (PPT) est un lignane extrait du rhizome de la plante himalayenne *Podophyllum hexandrum* disposant d'activités biologiques intéressantes, en particulier, des activités anti-mitotique et inhibitrice de la topoisomérase II lui conférant des propriétés anticancéreuses. Les effets secondaires liés, en partie, à l'activité anti-mitotique de cette molécule ont amené les chercheurs à étudier de nombreux dérivés de PPT moins toxiques pour la santé humaine. Notamment, l'étoposide, un dérivé de la PPT, dispose d'une activité inhibitrice de la topoisomérase II, mais pas d'activité anti-mitotique. Cette molécule est ainsi devenue l'une des plus utilisées dans les protocoles de chimiothérapie.

II.1 Généralités sur la podophyllotoxine

Les lignanes sont des métabolites spécialisés dérivant du métabolisme des phénylpropanoïdes (**Figure 41**). Ces composés polyphénoliques sont formés par dimérisation de monolignols (Fuss, 2007). Différentes liaisons peuvent intervenir dans le couplage des monolignols, créant ainsi une grande variété de squelettes carbonés. Par exemple, les dimères de monolignols peuvent être liés par les carbones en β-β, β-5 ou par une liaison éther β-O-4 (Moss, 2000 ; Boerjan et al., 2003). Le lignane le plus répandu est le pinorésinol, issu de la polymérisation de deux alcools coniféryliques en β-β (**Figure 42**). La diversité des lignanes est d'autant plus grande que ces squelettes carbonés peuvent subir des décorations et modifications plus ou moins importantes (Xu et al., 2018). Les lignanes sont présents chez un très grand nombre de plantes, mais leurs structures sont variables d'une espèce à l'autre.

Les lignanes peuvent démontrer des activités antivirales, fongicides, antibactériennes et cytotoxiques, il est donc fort probable que ceux-ci soient impliqués dans la défense des plantes contre des agents pathogènes, mais aussi dans les interactions avec d'autres organismes comme les champignons ou les insectes (Fuss, 2007 ; Hano et al., 2021). De plus, les lignanes semblent avoir des activités inhibitrices d'enzymes, de transport de nucléotides, de synthèse d'ADN ou même de protection contre le stress oxydatif dans la plante (Macrae et Towers, 1984 ; Gaross et al., 2018). Leur activité, localisation et concentration sont très variables d'une espèce végétale à l'autre. De nombreux lignanes issus de plantes utilisées en médecine traditionnelle ont été étudiés afin de mettre en évidence de nouveaux agents thérapeutiques (Gordaliza et al.,



<u>Figure 43</u> : Structure de la (-)-podophyllotoxine.

La podophyllotoxine est composée de 5 cycles lettrés en bleu. La numérotation des carbones est représentée en rouge.

Class	Fungal Endophyte	Host Plant
Sordariomycetes	*Fusarium oxysporum* *Fusarium solani* *Fusarium* sp. *Pseudallescheria* sp.	*Juniperus recurva, Podophyllum hexandrum, Dysosma versipellis, Sinopodophyllum hexandrum*
Agaricomycetes	*Trametes hirsuta*	*Podophyllum peltatum*
Leotiomycetes	*Phialocephala fortinii*	*Podophyllum peltatum*
Dothideomycetes	*Alternaria tenuissima*	*Sinopodophyllum emodi*
Eurotiomycetes	*Aspergillus fumigatus, Penicillium implicatum*	*Juniperous communis, Dysoma veitchii, Diphylleia sinensis*
Ascomycetes	*Monilia* sp.	*Dysoma veitchi*
Zygomycetes	*Mucor fragilis*	*Sinopodophyllum hexandrum*

<u>Figure 44</u> : Sources fongiques signalées comme source de podophyllotoxine. (Shah et al., 2021)

2004). En particulier, l'activité antitumorale de la podophyllotoxine (PPT) issue du rhizome de *Podophyllum hexandrum* a été démontrée (Macrae et Towers, 1984).

La (7α,7'α,8α,8'β)-7-hydroxy-3',4',5'-triméthoxy-4,5-méthylénédioxy-2,7'-cyclo-lignano-9',9-lactone, plus couramment appelée podophyllotoxine (PPT), est un lignane majeur des racines et des rhizomes chez *Podophyllum hexandrum* et *Podophyllum peltatum* dont elle tire son nom (Giri et al., 2000). La structure de cette molécule a été mise en évidence en 1958 (Hartwell and Schrecker, 1958), elle est composée de 5 cycles (A à E) (**Figure 43**). Le double cycle B et C forme une structure de type tétraline prolongée de part et d'autre par un hétérocycle de type lactone à 5 carbones en D et dioxolane en A. Le cycle C possède une fonction hydroxyle en C-7 et un cycle aromatique E en C-7'. Le cycle E n'appartenant pas au plan formé par les cycles A à D, il confère un encombrement stérique jouant un rôle important dans l'activité de la molécule, mais aussi dans sa solubilité dans l'eau (Bromberg, 1998). Enfin, le cycle E est décoré par trois groupements méthoxyle. La PPT a été isolée pour la première fois en 1980 (Podwyssotzki, 1980) et est restée l'objet de nombreuses recherches du fait de son activité antimitotique particulièrement intéressante.

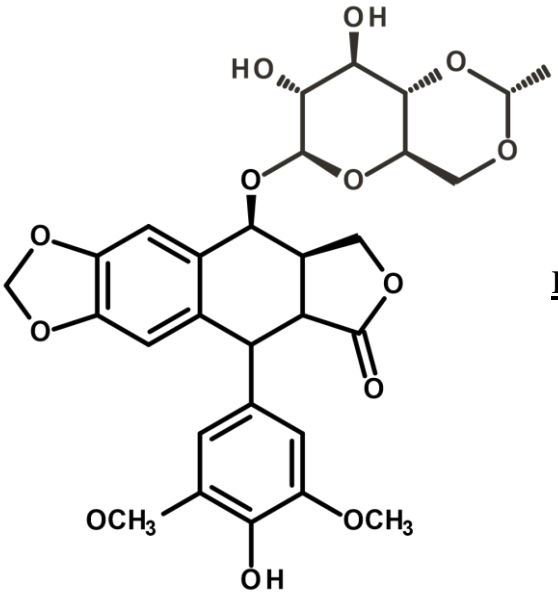
La PPT est produite dans un nombre restreint d'espèces et de genres végétaux. En dehors du genre *Podophyllum*, elle a été décrite chez *Diphylleia* (Berberidaceae), *Haplophyllum* (Rutaceae), *Anthriscus* (Apiaceae), *Thuja*, *Callitris, Juniperus* et *Thujopsis* (Cupresceae), *Hernandia* (Hernandiaceae), *Teucrium* et *Hyptis* (Lamiaceae), *Commiphora* et *Bursera* (Burseraceae), *Polygala* (Polygalaceae), *Linun* (Linaceae) (Yu et al., 2017, Ardalani et al., 2017). Il a également été mis en évidence que certains organismes fongiques produisent de la PPT en symbiose avec des plantes (**Figure 44**) comme *Trametes hirsuta* (Puri et al., 2006), *Sebacina vermifera*, *Phialocephala fortinii* (Eyberger et al., 2006), *Fungus Alternaria* (Cao et al., 2007), *Fusarium oxysporum* (Kour et al., 2008), *Aspergillus fumigatus* (Kusari et al., 2009), *Fusarium solani* (Nadeem et al., 2012), ou encore *Mucor fragilis* (Huang et al., 2014). Cependant, les rhizomes de *P. hexandrum* restent la source la plus importante de PPT (Kumar et al., 2015).

II.2 Utilisation de la podophyllotoxine en santé humaine

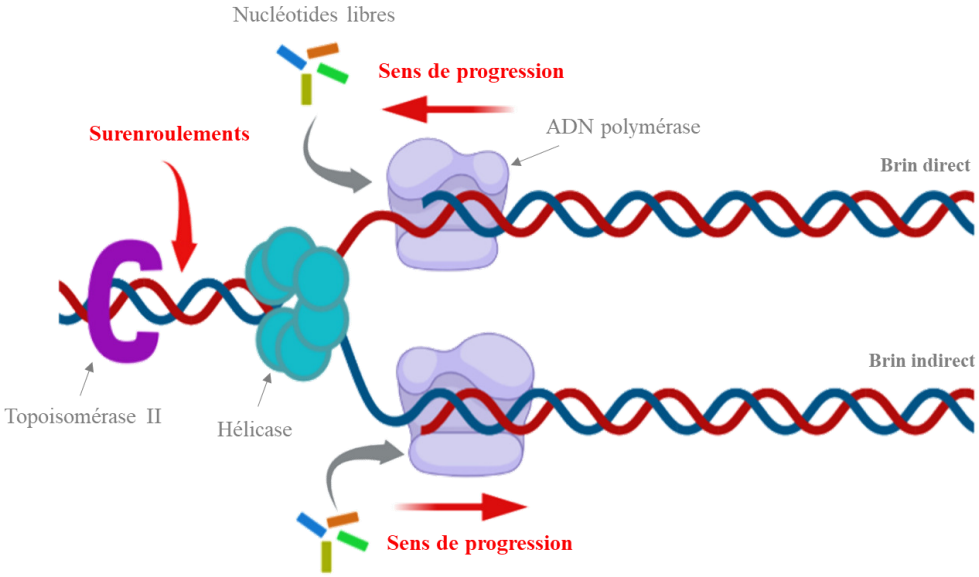
a) Propriétés pharmacologiques de la podophyllotoxine

La podophylline est une résine obtenue par un procédé de précipitation à partir d'un extrait de *Podophyllum hexandrum.* Elle possède de nombreuses vertus médicinales dues à sa richesse en lignanes. En effet, ces composés, incluant l'épipodophyllotoxine, la podophyllotoxone, la déméthylpodophyllotoxine mais surtout la PPT qui en est le composé majoritaire, lui confèrent ses propriétés pharmacologiques (Anand et al., 2022). La PPT dispose de nombreuses activités intéressantes du point de vue pharmaceutique. Ainsi, il a été démontré que la PPT présente des activités anti-virales contre le virus de l'Herpès (Bedows and Hatfield, 1998), le PapillomaVirus Humain (HPV) (Syed et al., 1995) ou encore le Virus de l'Immunodéficience Humaine (HIV) (Chen et al., 2007), auxquelles s'ajoute une activité radioprotectrice (Goel et al., 1998) et une activité anti-oxydante (Li et al., 2012). Elle semble également efficace pour calmer les crises dues au psoriasis (Leander and Rosen, 1988). Mais la PPT est particulièrement étudiée car ses activités anti-mitotique et inhibitrice de la topoisomérase II lui confèrent des propriétés anti-cancéreuses importantes. En effet, la PPT peut se lier au site colchicine de la tubuline, ce qui empêche l'assemblage et la polymérisation des microtubules et entraîne l'arrêt du cycle cellulaire en métaphase (Buss et al., 1995). L'action sur la topoisomérase II entraîne la mort cellulaire suite à l'absence de réparation d'éventuelles cassures de l'ADNg (Long and Minocha, 1983 ; Ross et al., 1984).

Malgré cette propriété intéressante, l'utilisation de la PPT en tant qu'anti-cancéreux a été interdite, du fait des nombreux effets secondaires liés à sa toxicité gastro-intestinale, sa neurotoxicité, et son action sur l'altération de la moelle osseuse. Cependant, la PPT est utilisée en tant que base structurelle pour la production de molécules semi-synthétiques moins cytotoxiques comme l'étoposide et le teniposide, toutes deux utilisées en santé humaine pour le traitement de plusieurs types de cancer. De plus, certains dérivés de PPT présentent un intérêt pharmaceutique différent, par exemple dans le traitement de l'arthrite rhumatoïde (dérivé CPH-82) (Lerndal and Svensson, 2000). La PPT est aussi intéressante d'un point de vue agronomique puisqu'elle dispose d'une activité insecticide contre les larves de *Brontispa longissima*, insecte ravageur du cocotier et d'autres espèces de palmiers (Zhang et al., 2010).



<u>Figure 45</u> : Structure de l'étoposide.



<u>Figure 46</u> : Schéma d'une fourche de réplication les contraintes topologiques qui y sont associées.

Les surenroulements causés par l'ouverture de la double hélice d'ADN par l'hélicase sont relâchés grâce à une coupure transitoire de l'ADN double brin par la topoisomérase II.

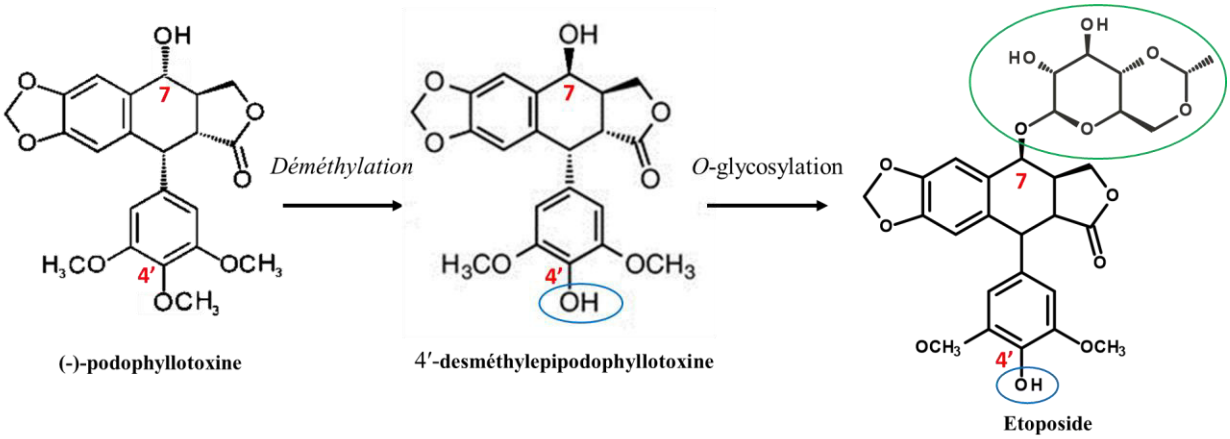
<u>Figure 47</u> : Etoposide sous forme liquide utilisée dans les protocoles de chimiothérapie. (Mandiamart)



b) L'étoposide

Le 4'-déméthyl-épipodophyllotoxine-9-4,6-O-éthylidène-p-Dglucopyranoside, plus couramment appelé étoposide, est une forme dérivée de la PPT disposant d'une activité inhibitrice de la topoisomérase II (Hande, 1998) moins puissante que la PPT elle-même (**Figure 45**) (Greenspan et al., 1950). Cette molécule a été conçue lors d'un projet de synthèse d'une série de dérivés de PPT, testés par rapport à une activité antinéoplasique moins toxique que celle de PPT et compatible avec une utilisation comme molécule anti-cancéreuse (Stähelin and Von Wartburg, 1991). Les topoisomérases sont une famille d'enzymes nucléaires capables de couper transitoirement l'ADN, permettant ainsi à la cellule d'adapter la topologie de son ADN et d'éviter les surenroulements (**Figure 46**) (Chen, 1994 ; Froelich-Ammon and Osheroff, 1995). Ces topoisomérases sont essentielles à la vie de la cellule, notamment lors des étapes de réplication, de transcription, de recombinaison et de réparation de l'ADN. Il existe deux formes de topoisomérases : les topoisomérases de type I capables de couper l'ADN simple brin mises en évidence en 1977 (Wang, 1977) et les topoisomérases de type II capables de couper l'ADN double brin mises en évidence en 1976 (Gellert et al., 1976). L'étoposide forme un complexe stable entre l'ADN et la topoisomérase II, empêchant ainsi la possibilité de coupure de l'ADN par l'enzyme. Lors de la division cellulaire, la formation de fourches de réplication implique de nombreuses contraintes topologiques dans l'ADN telles que des surenroulements qui doivent être relâchés via la coupure transitoire de double brin d'ADN. La topoisomérase étant dans l'incapacité de réaliser ces coupures, l'ADN subit de nombreuses cassures double brin conduisant à la mort de la cellule (Long and Minocha, 1983 ; Ross et al., 1984). L'activité de l'étoposide est particulièrement importante dans les cellules ayant un fort taux de multiplication telles que les cellules cancéreuses. C'est pourquoi des essais cliniques ont été réalisés sur cette molécule dès le début des années soixante-dix et sa commercialisation en tant que molécule antinéoplasique a été approuvée en 1983 aux Etats-Unis (Hande, 1998).

L'étoposide est aujourd'hui un des médicaments anti-cancéreux les plus utilisés en chimiothérapie (Meresse et al., 2004) puisqu'elle est efficace contre de nombreux cancers, tels que les cancers du poumon à petites cellules, les carcinomes testiculaires, les leucémies aigües lymphoblastiques et myéloblastiques, les lymphomes malins et les choriocarcinomes placentaires (**Figure 47**) (Sinkule, 1984 ; Gordaliza et al., 2004). Bien qu'elle soit largement utilisée dans les protocoles de chimiothérapie, l'étoposide présente un risque de développement de leucémie secondaire chez les patients traités (2 à 12%). Des analogues d'étoposide et de PPT



<u>Figure 48</u> : Synthèse chimique de l'étoposide à partir de la podophyllotoxine.

La première étape de déméthylation permet de remplacer le groupement méthoxyle par un groupement hydroxyle en C-4' (cercle bleu), ce qui a pour conséquence de changer le plan de la fonction hydroxyle en C-7. La seconde étape est une *O*-glycosylation permettant l'ajout d'un groupement éthylidène-β-D-glucopyranoside en C-7 (cercle vert).

ont donc été synthétisés afin de comprendre le mécanisme impliqué dans cet effet secondaire (Murphy et al. 2020). Enfin, il a été récemment montré que l'étoposide pourrait être utilisée dans le traitement du choc cytokinique associé à certains patients atteints de la COVID-19 (Zhao et al., 2021).

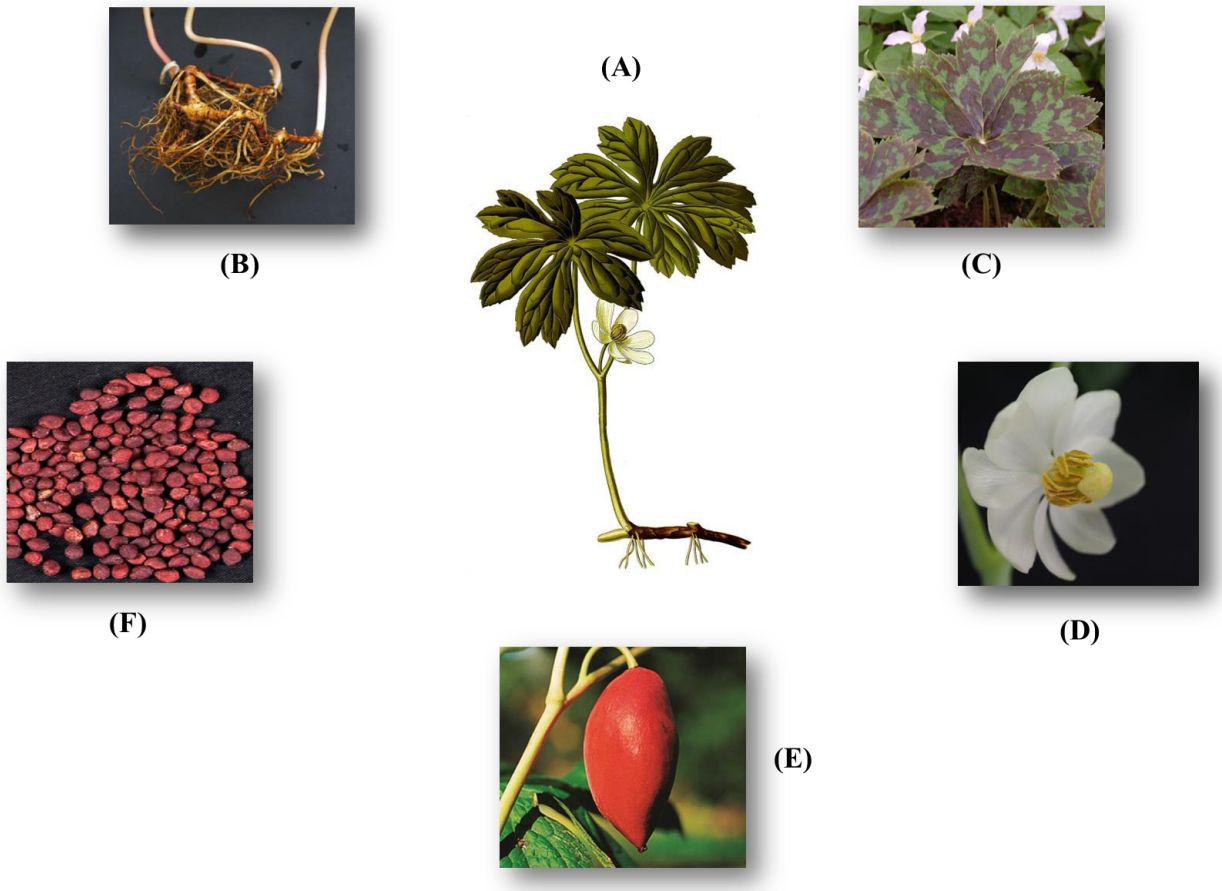
La première synthèse d'étoposide et les premiers tests biologiques ont été réalisés en 1966 (Stähelin and Von Wartburg, 1991). Elle est actuellement obtenue par synthèse chimique à partir de PPT via la substitution du groupement méthoxyle en groupement hydroxyle en position C-4', le passage à un épimère de la fonction hydroxyle en C-7 et la formation stéréosélective d'une liaison *O*-glycosidique permettent l'ajout d'un groupement éthylidène-β-D-glucopyranoside en C-7 (**Figure 48**) (Sinkule, 1984). Cette transformation chimique est réalisée en deux étapes. La première étape consiste en la conversion de PPT en 4′-deméthylépipodophyllotoxine (4'dEPT) par une réaction chimique couplée de déméthylation et changement de la stéréochimie (Stähelin et Von Wartburg, 1991). Une *O*-glycosylation permet ensuite d'ajouter le groupement éthylidène-β-D-glucopyranoside en C-7 (Liu et al., 2016). La formation de la liaison *O*-glycosidique peut être obtenue grâce à un couplage de type Koenigs-Knorr d'un dérivé de glucose ou grâce à une attaque stéréocontrolée du groupe hydroxyle d'un dérivé de glycopyranose (Allevi et al., 1993).

II.3 Approvisionnement en podophyllotoxine

a) Podophyllum hexandrum, une espèce menacée d'extinction

La production des médicaments anticancéreux comme l'étoposide décrite précédemment, mais aussi le téniposide et l'étopophose (Who, 2000) est réalisée par hémi-synthèse à partir de PPT extraite de biomasse végétale. Si la PPT a été décrite dans plusieurs genres végétaux, le genre Podophyllum est celui qui en accumule le plus et parmi elles, *P. hexandrum* reste la plus intéressante. Elle accumule, par exemple, 17 fois plus de PPT dans ses rhizomes que *P. peltatum* (Jackson and Dewick, 1984).

P. hexandrum est une plante himalayenne qui a longuement été utilisée par les indigènes et les Indiens d'Amérique comme remède purgatif pour les intestins (Nadeem et al., 1999). Mais elle est aussi utilisée dans d'autres régions comme cicatrisant, antiseptique, anti-



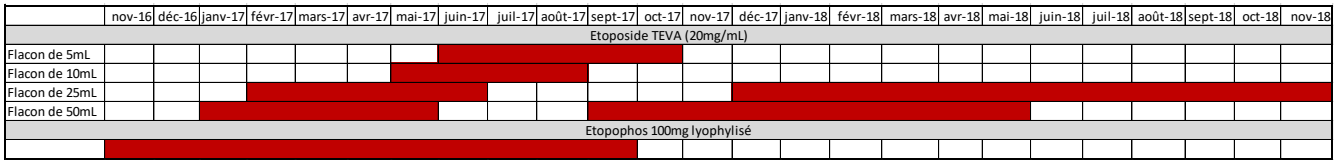
<u>Figure 49</u> : *Podophyllum hexandrum* et ses différents organes.

(A) Dessin de la plante entière (Köhler, Medicinal-Pflanzen, 1887)
(B) Rhizomes poussant près de la surface du sol.
(C) Feuilles pouvant mesurer jusqu'à 25cm couramment tachetées.
(D) Fleur blanche.
(E) Fruit de couleur orangée de la taille d'un citron.
(F) Graines plates, ovales et de couleur brun foncé.

diarrhéique, et pour bien d'autres raisons encore (Anand et al., 2022). La famille des Berberidaceae est composée de plantes angiospermes dicotylédones réparties en 18 genres, dont le genre herbacé *Podophyllum*. Ce genre est composé de six espèces, notamment *P. peltatum* (Amérique du nord) et *P. hexandrum* (Inde). Il est parfois difficile de rechercher les espèces de ce genre car elles ont été déplacées d'un genre à l'autre et ont parfois été renommées. Par exemple, *P. hexandrum* peut être nommée *Podophyllum emodi* ou *Sinopodophyllum hexandrum* dans certains manuscrits. *Podophyllum* dispose de plusieurs noms vernaculaires comme la pomme de mai, la mandragore, le citron sauvage, la baie du raton laveur, la pomme du diable ou encore la plante parapluie (Maqbool, 2011).

Podophyllum dispose de rhizomes pérennes poussant près de la surface du sol, sa partie aérienne est annuelle, émergeant généralement vers la mi-avril. La tige, pouvant atteindre 15 à 30 centimètres, dispose de paires de feuilles pouvant mesurer jusqu'à 25cm couramment tachetées (Holm, 1899). Les bourgeons éclosent en mai et laissent apparaître une fleur blanche rosée permettant à un fruit de couleur orangée et riche en eau de mûrir jusqu'à atteindre la taille d'un petit citron (Pearce and Thieret, 1993). *Podophyllum* est auto-incompatible, c'est-à-dire qu'il ne peut pas y avoir de fécondation dans une même fleur, ni entre les fleurs d'une même plante, nécessitant ainsi la présence de pollinisateurs extérieurs comme les insectes (Laverty and Plowright, 1988) pour faciliter les fécondations croisées entre deux plantes distinctes. Les graines présentes dans le fruit (quelques dizaines/fruits), sont plates, ovales et de couleur brun foncé. Les plantes ne produisent qu'un unique rhizome (**Figure 49**) (Chatterjee, 1952).

P. hexandrum est une espèce endémique à la région Himalayenne, elle pousse donc à des altitudes comprises entre 3500 et 4000 mètres (Chattopadhyay et al., 2002a). De plus, cette plante se reproduit difficilement puisqu'elle préfère la multiplication végétative à la reproduction sexuée grâce à son rhizome. Le rhizome et les racines sont les organes accumulant le plus de PPT, ce qui implique d'arracher la plante entière pour extraire la molécule. Or, il s'agit d'une plante à croissance lente puisqu'il faut attendre 6 ans pour que celle-ci soit suffisamment mature pour accumuler de la PPT (Chatterjee, 1952). Les contraintes de culture de cette plante, couplées au fort intérêt pharmaceutique qu'elle représente, ont entraîné une récolte sauvage intensive de celle-ci conduisant à une réelle menace de disparition. *P. hexandrum* a ainsi été inscrite sur la liste de la Convention sur le commerce des espèces de faune et de flore sauvages menacées d'extinction en 1992 et, depuis 1999, le gouvernement Indien a interdit l'exportation de tout spécimen prélevé dans la nature. Sachant que l'éventail de plantes connues produisant de la PPT est insuffisant pour répondre à la demande croissante



<u>Figure 50</u> : Durée des ruptures de stock des médicaments contenant de l'étoposide de novembre 2016 à novembre 2018.

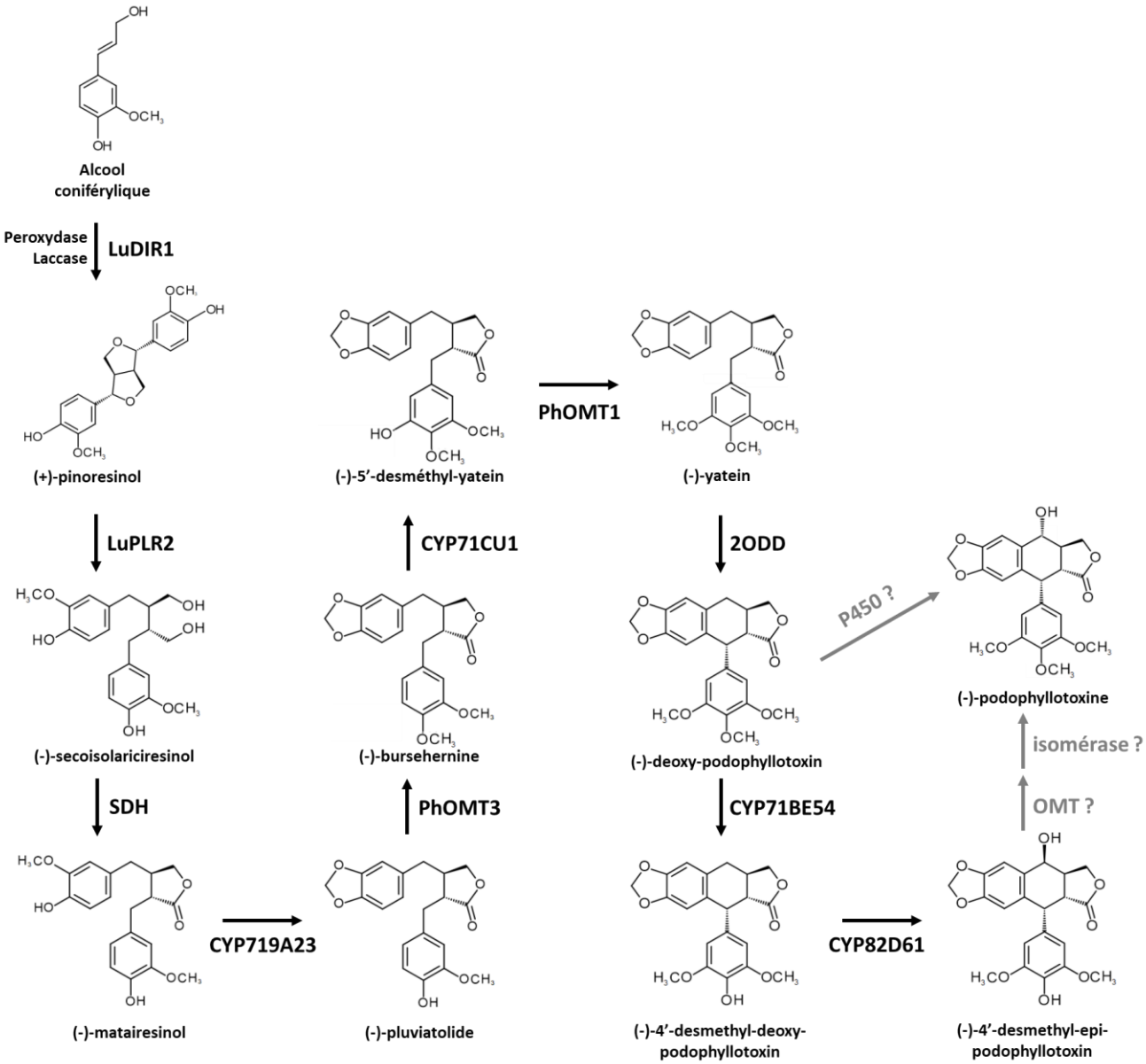
Les périodes de pénuries sont représentées en rouge.

en étoposide (Gordaliza et al., 2004), des ruptures totales et mondiales en étoposide ont lieu de façon récurrente (**Figure 50**) (Vidal, 2017) et ne cessent de s'amplifier.

b) Productions alternatives de podophyllotoxine

Les contraintes liées à cette espèce, telles que son faible taux de germination ou encore sa croissance lente (Badhwar and Sharma, 1963) ont amené beaucoup d'équipes à chercher des alternatives afin de produire de la PPT. Une synthèse chimique de la PPT à l'échelle industrielle est difficilement envisageable. En effet, la présence de quatre centres chiraux et d'un fragment trans-lactone sensible aux bases complique sa synthèse qui nécessite beaucoup d'étapes de protection et diminue ainsi les rendements (Ward, 1992 ; Bush and Jones, 1995 ; Berkowitz et al., 2000 ; Canel et al., 2000). La synthèse chimique de l'épipodophylloxotine à partir du 6-bromopipéronal est la plus courte connue avec cinq étapes (Chi and Maimone, 2014).

C'est donc vers des procédés liés aux biotechnologies végétales que se sont orientées les études. Par exemple, Van Uden et son équipe ont initié, en 1989, des cultures cellulaires *in vitro* à partir de racines de *P. hexandrum* et montré une production de PPT dans des cultures de cals et de suspensions cellulaires indifférenciées. L'ajout dans le milieu de culture de précurseurs de la voie de synthèse de la PPT permet également d'augmenter son accumulation mais à un niveau relativement faible (Van Uden et al., 1989). En 1990, il a été montré que des cultures *in vitro* d'embryons somatiques pouvaient être une bonne alternative de production (Arumugam and Bhojwani, 1990). En 1997, Sadowska et son équipe ont mis en évidence la possibilité d'obtenir des plantules et des cals de *P. peltatum* via des cultures d'embryons et d'embryons divisés (Sadowska et al., 1997). En 1998, un protocole de micropropagation de *P. peltatum in vitro* a été développé afin de produire plus rapidement des plantes à haut rendement (Moraes-Cerdeira et al., 1998). Chattopadhyay et son équipe ont travaillé, de 2001 à 2003 sur la production de PPT via des cultures de suspensions cellulaires de *P. hexandrum* (Chattopadhyay et al., 2001 ; Chattopadhyay et al., 2002a ; Chattopadhyay et al., 2002b ; Chattopadhyay et al., 2003). En 2006, Bedir et son équipe ont mis en évidence que les feuilles blessées par écrasement et séchées produisent plus de PPT que les feuilles saines (Bedir et al., 2006). En 2009, un mode de production de PPT par culture racinaire de *P. hexandrum* a été établi via l'ajout d'hormones exogènes (Li et al., 2009). En 2000, une forte accumulation de PPT chez *Linum album* a été mise en évidence (Empt et al., 2000) et Garden a tenté, en 2003,



<u>Figure 51</u> : Voie de biosynthèse de la podophyllotoxine.

LuDIR1 : protéine dirigeante de *Linum usitatissimum* ; LuPLR2 : pinorésinollaricirésinol réductase de *L. usitatissimum* ; SDH : sécoisolaricirésinol déshydrogénase ; CYP719A23 : pluviatolide synthase de *P. hexandrum* ; PhOMT3 : pluviatolide 4-*O*-méthyltransférase de *P. hexandrum* ; CYP71CU1 : bursehernine 5′-hydroxylase de *P. hexandrum* ; PhOMT1 : 5'-déméthylyateine *O*-transférase de *P. hexandrum* ; 2ODD : désoxypodophyllotoxine synthase de *P. hexandrum*; CYP71BE54 : cytochrome P450 de *P. hexandrum* ; CYP82D61 : cytochrome P450 de *P. hexandrum*. La dernière étape de la voie conduisant à la podophyllotoxine n'a pas encore été découverte.

de produire de la PPT grâce à la culture en bioréacteur de suspensions cellulaires de cette espèce (Garden, 2003). Malgré toutes ces tentatives, l'accumulation de PPT au sein de ces différents systèmes biotechnologiques reste insuffisante pour être économiquement viable et/ou difficilement transposable à l'échelle industrielle.

II.4 Découverte de la voie de biosynthèse de la podophyllotoxine *in planta*

La caractérisation des enzymes impliquées dans la voie de synthèse de la PPT a permis d'envisager une alternative d'approvisionnement via l'ingénierie métabolique de cette voie dans un organisme hôte hétérologue.

Les lignanes sont synthétisés à partir de la voie des phénylpropanoïdes et en particulier, à partir de l'alcool conifé rylique pour former la PPT (**Figure 51**) (Lewis et al., 1998). Les premières étapes de la voie de biosynthèse ont été caractérisées pour la première fois chez *Linum usitatissimum* et *Forsythia intermedia*. Ainsi trois étapes enzymatiques sont nécessaires pour former le matairésinol. Dans un premier temps, des peroxidases catalysent l'association de deux alcools coniféryliques. L'intervention d'une protéine dirigeante (LuDIR1) située dans l'apoplaste des cellules favorise un couplage β-β pour former le (+)-pinorésinol (Davin et al., 1997 ; Dasilay et al., 2015). Ensuite, une pinorésinol/laricirésinol réductase (LuPLR2 ; EC:1.23.1.3) exprimée dans les racines catalyse les réductions successives du (+)-pinorésinol en (+)-laricirésinol puis en (-)-sécoisolaricirésinol (Von Heimendahl et al., 2005). Finalement une sécoisolaricirésinol déshydrogénase (FiSDH ; EC:1.1.1.331) également identifiée chez *P. peltatum* est capable de convertir la (-)-sécoisolaricirésinol en (-)-matairésinol (Xia et al., 2001).

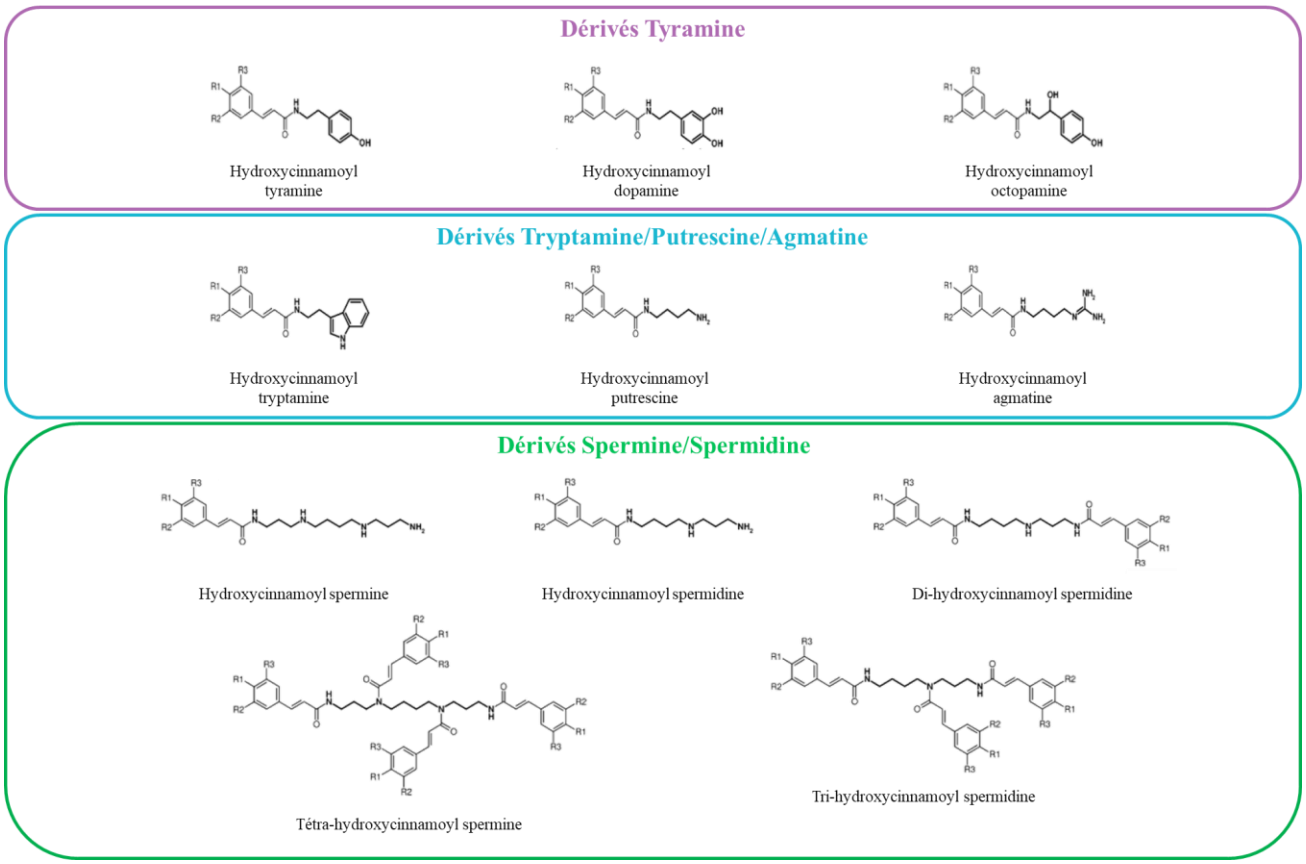
En 2013, Marques et son équipe ont mis en évidence l'activité pluviatolide synthase de CYP719A23 (EC:1.14.19.72) de *P. hexandrum* à partir du (-)-matairésinol (Marques et al., 2013). Les quatre enzymes suivantes ont également été caractérisées chez *P. hexandrum* par Lau and Sattely pour la synthèse de déoxypodophyllotoxine (DPT) à partir du (-)-pluviatolide (**Figure 51**) (Lau and Sattely, 2015). Le (-)-pluviatolide est d'abord converti en bursehernine via une méthylation en C-4' par la pluviatolide 4-*O*-méthyltransférase (PhOMT3 ; EC:2.1.1.323). La bursehernine 5′-hydroxylase CYP71CU1 (EC:1.14.14.131) ajoute ensuite un oxygène moléculaire en C-5′ de la bursehernine, pour obtenir la (-)-5-desméthyl-yateine. Cette dernière est ensuite transformée en (-)-yateine (YAT) via une méthylation en C-5'OH par la 5'-

déméthylyateine *O*-transférase (PhOMT1 ; EC:2.1.1.330). La dioxygénase désoxypodophyllotoxine synthase (2ODD ; EC:1.14.20.8) cyclise la YAT précédemment formée entre les carbones C-2 et C-7' pour obtenir la DPT. La DPT est le premier squelette de type aryltetraline lactone, précurseur des lignanes de type PPT. La 2ODD est une dioxygénase dépendante du 2-oxoglutarate et du Fe(II). Chez *P. hexandrum*, la 2ODD est principalement exprimée dans les feuilles et les tiges. Lorsqu'elle est exprimée dans la levure, elle est située dans le cytosol. Elle est composée de 311 acides aminés pour un poids moléculaire de 34,5 kDa et son site actif se situe entre les acides aminés 159 et 258.

La DPT doit être hydroxylé pour aboutir à la PPT, mais aucune enzyme capable de réaliser cette réaction n'a été identifiée à ce jour. Cependant, Lau et Sattely ont identifié deux cytochromes P450 de *P. hexandrum* capables d'utiliser la DPT comme substrat. En effet, une première P450, CYP71BE54 (EC:1.14.14-), catalyse la synthèse de (-)-4′-demethyldésoxypodophyllotoxine (4'dDPT) par déméthylation du groupement C-4'OH de la DPT. La seconde P450, CYP82D61 (EC:1.14.14.132), catalyse l'hydroxylation du groupement C-7 de la 4'dDPT pour former le 4'dEPT. Ces P450 sont dépendantes du NADPH et appartiennent au clade des CYP71. Elles sont ancrées à la membrane du RE et exprimées spécifiquement dans le rhizome de *P. hexandrum*. CYP71BE54 est composée de 508 acides aminés pour un poids moléculaire de 56,4 kDa et son domaine transmembranaire d'ancrage à la membrane du RE est situé entre les acides aminés 1 et 21. CYP82D61 est composée de 520 acides aminés pour un poids moléculaire de 57,7 kDa. Etonnamment, les modèles prédisent que CYP82D61 dispose de deux domaines transmembranaires d'ancrage à la membrane du RE, un premier situé de façon classique entre les acides aminés 6 et 26 et un second, avec une confiance de prédiction plus faible, entre les acides aminés 108 et 130. Pour obtenir la PPT à partir de 4'dEPT, deux réactions seraient nécessaires, une isomérisation et une méthylation. Ainsi deux voies de biosynthèse de PPT à partir de la DPT sont envisageables *in planta* mais aucune n'a été validée pour l'instant.

La synthèse de PPT à partir de l'alcool conifférylique est donc une voie complexe *in planta* comprenant au moins dix étapes enzymatiques. La voie de synthèse du pinorésinol à partir de l'alcool conifférylique est retrouvée dans de nombreuses plantes telle qu'*A. thaliana*. Quelques espèces suivent la voie des lignanes vers la PPT jusqu'au matairésinol ou jusqu'à la YAT, alors que les lignanes en aval sont retrouvés dans des plantes bien plus spécialisées, ce qui limite sa production dans la nature.

La création d'une souche de levure capable de bioconvertir la YAT hautement biodisponible en 4'dEPT directement utilisable pour une synthèse chimique d'étoposide est la stratégie alternative d'approvisionnement en étoposide qui sera présentée dans ces travaux de thèse.



<u>Figure 52</u> : Exemples de la diversité des phénolamides.

R1 et R2 = H; OH ou OMe.

Chapitre III : Exploration du potentiel pharmaceutique des tri-hydroxycinnamoyl spermidines

Les tri-hydroxycinnamoyl spermidines (THCSpd) sont des phénolamides accumulés exclusivement dans le manteau pollinique des dicotylédones. Même si les connaissances sur la fonction biologique des THCSpd dans la qualité du pollen demeurent aujourd'hui assez floues, ces molécules disposent d'activités pharmaceutiques très intéressantes. Cependant, la connaissance des applications prometteuses des différentes THCSpd est limitée par l'approvisionnement de ces molécules.

III.1 Généralités sur les tri-hydroxycinnamoyl spermidines

a) Les phénolamides

Les phénolamides, parfois appelés phénylamides ou acides aminés hydroxycinnamiques (HCAA), sont des MS retrouvés essentiellement chez les angiospermes. Ces composés sont formés d'une amine N-acylée par un groupement phénolique (**Figure 52**). Les acides phénoliques peuvent être conjugués à des monoamines aromatiques telles que la tyramine, la tryptamine et la dopamine ou de polyamines aliphatiques telles que la putrescine, la spermidine, la spermine et l'agmatine (Elejalde-Palmett et al., 2015 ; Roumani et al., 2021). Les groupements phénoliques principalement retrouvés chez les phénolamides sont les hydroxycinnamates, des dérivés hydroxylés de l'acide cinnamique tels que le *p*-coumarate, le caféate, le férulate, le 5-OHférulate ou encore le sinapate, issus du métabolisme des phénylpropanoïdes (**Figure 41**) (Strack et al., 1990). Les polyamines peuvent être mono- ou poly-substituées sur l'ensemble de leurs fonctions amines, procurant plusieurs degrés possibles de saturation, ce qui conduit à une grande diversité de ces composés (Bassard et al., 2010). Théoriquement, il pourrait exister plus de 5 000 phénolamides dans le monde végétal. La féruloylputrescine a été le premier phénolamide décrit au milieu du XXème siècle (Ryabinin and Ilina, 1949).

Bien que le rôle des phénolamides ne soit pas complètement élucidé, ils ont été corrélés ou impliqués dans les réponses de défense végétale contre les stress biotiques tels que les pathogènes et les insectes. Ainsi, la synthèse des coumaroyl et féruloyl tyramines par plusieurs Solanacées est induite après traitement aux éliciteurs ou lors d'infections fongiques et bactériennes (Negrel and Martin, 1984 ; Keller et al., 1996 ; Schmidt et al., 1998 ; Newman et al., 2001 ; Zacarés et al., 2007) tout comme la synthèse des coumaroyl et féruloyl agmatine chez Arabidopsis (Muroi et al., 2009). De plus, des activités antimicrobiennes contre Pyrenophora et l'oïdium chez *Hordeum vulgare* ont été observées pour la dicoumaroyl-spermidine (Walters et al., 2001), ainsi qu'une activité antivirale de la féruloyl tyramine et de certains dérivés de putrescine dans la défense du tabac contre le virus de la mosaïque du tabac (Martin-Tanguy et al., 1973, Martin-Tanguy et al., 1976 ; Negrel and Jeandet, 1987). Ces observations argumentent l'implication des phénolamides dans la protection des végétaux contre les agents pathogènes. Toutes les classes de phénolamides semblent être impliquées dans la défense des plantes contre les pathogènes. Ce rôle protecteur implique des activités antibactériennes ou antifongiques directes des phénolamides mais aussi un effet indirect via le renforcement de la paroi cellulaire secondaire afin de réduire la digestion de celle-ci par les enzymes hydrolytiques des pathogènes (Roumani et al., 2021). L'induction de la synthèse de phénolamides a également été rapportée en réponse aux attaques par des insectes dans de nombreuses plantes. Par exemple, *Nicotiana attenuata* synthétise de la caféoyl putrescine et de la dicaféoyl spermidine en réponse à une attaque d'insectes herbivores (Kaur et al., 2010). La synthèse de phénolamides semble également pouvoir agir comme facteur de diminution de la ponte des insectes. En effet, l'accumulation de la caféoyl putrescine et l'ajout de *p*-coumaroylputrescine synthétique dans les cotylédons de plants de poivron ont inhibé l'oviposition de *Liriomyza trifolii* sur les feuilles les produisant (Tebayashi et al., 2007).

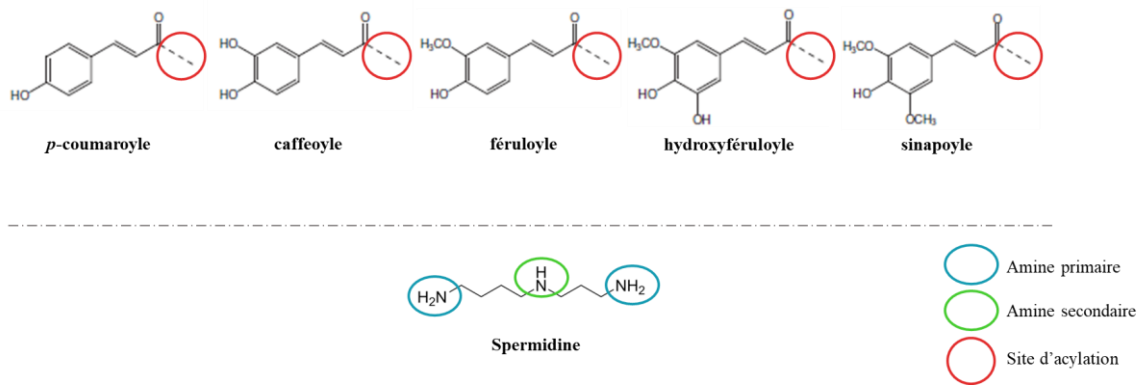
Même si les études sur la synthèse de phénolamides en réponse à un stress abiotique ont été moins nombreuses, il semblerait que la synthèse de dérivés de putrescine et de spermidine soit initiée en réponse à de nombreux stress abiotiques tels que les carences en minéraux, les températures élevées, l'excès d'eau et l'exposition aux UV-B ou à un niveau élevé d'ozone (O_3) chez le blé et le tabac (Klapheck, 1983 ; Edreva et al., 1995 ; Edreva et al., 1998 ; Königshofer and Lechner, 2002 ; Demkura et al., 2010). A l'inverse, il a été montré que certains facteurs abiotiques inhibent la synthèse de phénolamides, tels que le stress salin et la limitation en azote (Shevyakova et al., 2006 ; Larbat et al., 2014). Plusieurs hypothèses restent à confirmer sur le rôle des phénolamides dans le piégeage de radicaux libres, dans l'activité antioxydante et dans

la protection contre les UV (Edreva et al., 1998 ; Choi et al., 2007 ; Georgiev et al., 2012 ; Abe et al., 2017).

Enfin, des phénolamides s'accumulent constitutivement dans les organes reproducteurs, les graines ou bien les racines chez un certain nombre d'espèces végétales (Elejalde et al., 2015). On retrouve une accumulation de dérivés de putrescine et de spermidine mono- ou di-substitués dans les graines de *Pentaclethra macrophylla*, du maïs, et d'*A. thaliana* (Mbadiwe, 1973 ; Moreau et al., 2001 ; Luo et al., 2009), les racines de *Capsicum annuum*, *Aphelandra tetragona* et *Capparis spinosa* (Yoshihara et al., 1981 ; Hedberg et al., 1996 ; Fu et al., 2008), et les bourgeons floraux ainsi que dans les organes reproducteurs de *N. tabacum*, des Fagales et de diverses espèces du genre Hamamelididae (Buta and Izac, 1972 ; Martin-Tanguy et al., 1973 ; Cabanne et al., 1981 ; Meurer et al., 1988). De façon originale, il a été montré que les spermidines tri- et tétra-substituées s'accumulent à la surface des grains de pollen chez les dicotylédones. Les phénolamides auraient donc un possible rôle dans la croissance et le développement floral, comme dans l'initiation florale et la fertilité des fleurs. Notamment, les mono- et di-putrescines, la spermidine et les phénolamides de spermidine se comportent comme des régulateurs de croissance des plantes (Tiburcio et al., 1993 ; Martin-Tanguy, 1997 ; Edreva et al., 2007). En effet, leur synthèse semble liée aux processus biologiques de développement tels que la prolifération cellulaire, la formation de cals, la différenciation cellulaire, l'embryogénèse ou encore l'organogénèse (Edreva et al., 2007). Néanmoins, aucune preuve directe de ces hypothèses n'a été fournie, même si plusieurs indices semblent les confirmer tels que le fait que l'accumulation de phénolamides au niveau des feuilles supérieures et des apex semble précéder l'initiation florale chez les Solanaceae, les Brassicaceae, les Asteraceae et les Araceae (Cabanne et al., 1981 ; Martin-Tanguy, 1985 ; Martin-Tanguy, 1997) ou que les plantes mutantes qui ne parviennent pas à fleurir n'accumulent pas de phénolamides (Martin-Tanguy, 1985).

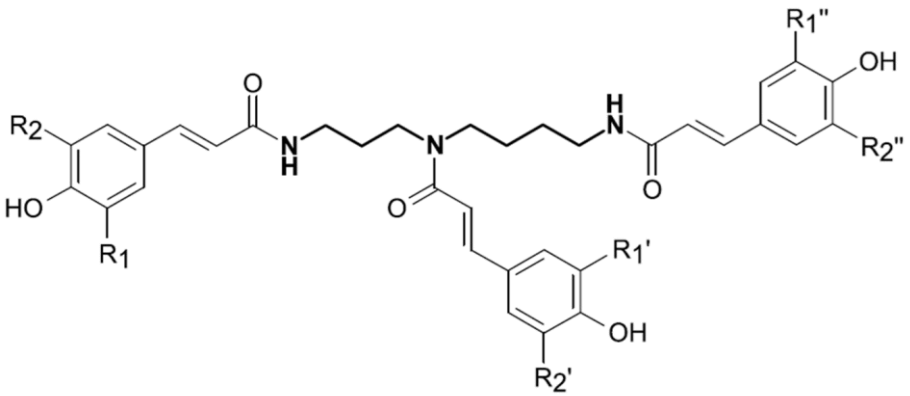
b) Diversité des THCSpd

Les phénolamides sont des molécules mono-substituées sur des amines primaires. Les polyamines, elles, peuvent être mono- ou di-substituées sur leur amine primaire terminale. Enfin, il existe des cas particuliers de polyamines disposant de fonctions amines secondaires comme la spermidine et la spermine qui peuvent être substituées sur l'ensemble de leurs



<u>Figure 53</u> : Structures des différents acides hydroxycinnamiques et de la spermidine.

Tous les groupements du squelette de spermidine peuvent être substitués par les groupements hydroxycinnamiques représentés.



<u>Figure 54</u> : Structure de la tri-hydroxycinnamoyl spermidine (THCSpd).

R1 et R2 = H, OH ou OMe.

fonctions azotées. En particulier, les tri-hydroxycinnamoyl spermidines (THCSpd) représentent une famille particulière de polyamines substituées sur l'ensemble de leur fonction aminées. En effet, ils sont composés d'un squelette de spermidine, ou moins couramment, de spermine, dont tous les groupements amines ont été substitués par des groupements hydroxycinnamiques (**Figure 53**).

Les THCSpd ont été découvertes pour la première fois en 1990 dans les fleurs de Rosaceae (Strack et al., 1990). Depuis, des études ont mis en évidence la présence spécifique de ces composés dans le manteau pollinique, et plus particulièrement dans celui des dicotylédones (Elejalde-Palmett et al., 2015). En effet, des THCSpd ont été mises en évidence dans le pollen de chêne (Bokern et al., 1995), du tournesol (Lin and Mullin, 1999), de l'arachide (Sobolev et al., 2008), du sureau (Kite et al., 2013), du câpre (Wiese et al., 2013), du noisetier (Meurer et al., 1986), d'Aphelandra (Werner et al., 1995), d'Hippeastrum x hortorum (Youhnovski et al., 2001) et d'Arabidopsis (Grienenberger et al., 2009). Ainsi, les THCSpd peuvent être considérées comme un marqueur universel de la paroi du pollen des Eudicotylédones. La THCSpd la plus courante est la tri-*p*-coumaroyl spermidine (**Figure 54**). Cependant dans la plante, cinq hydroxycinnamates différents peuvent substituer les divers groupements amines du squelette spermidine, ce qui permet d'envisager jusqu'à 125 structures de THCSpd probables. Certaines structures de THCSpd ont été décrites à ce jour (Li et al., 2018) avec une analyse exhaustive du modèle d'accumulation de ces molécules dans le pollen d'*A. thaliana* (Handrick et al., 2010). Cette étude a permis, grâce à une approche de spectrométrie de masse en tandem, d'identifier 14 THCSpd différents dans le pollen d'*A. thaliana*. Récemment, il a été démontré dans le pollen de *Cichorium intybus* que les dérivés de spermidine entièrement substituées ont laissé place à des dérivés de spermine entièrement substitués telle que la N',N'',N''',N''''-tétracoumaroylspermine. Ce glissement des THCSpd d'une base spermidine vers la spermine représentent un marqueur de la famille des Asteraceae (Delporte et al., 2018). Mais l'accumulation de phénolamides entièrement substitués dans le manteau pollinique des Eudicotylédones est un trait évolutif conservé (Elejalde et al., 2015 ; Bernard et al., 2022), suggérant que ces molécules joueraient un rôle dans le développement du pollen (Fellenberg and Vogt, 2015).

Les connaissances sur la fonction biologique des THCSpd dans la qualité du pollen sont peu développées. Cependant, il est probable que ces composés soient impliqués dans la formation de l'exine du grain de pollen et dans la protection contre les UV du matériel génétique qui assurera la formation du gamète pour la fécondation (Fellenberg and Vogt, 2015). Les

THCSpd peuvent aussi être impliqués dans la résistance structurelle du manteau pollinique, dans la compatibilité pollen-stigmate ou encore dans le pouvoir germinatif du grain de pollen. De plus, sachant que les THCSpd sont spécifiques des angiospermes et que seuls les angiospermes réalisent la pollinisation via les insectes, il est possible que ces composés puissent être impliqués dans l'interaction entre la plante et les insectes pollinisateurs (Austel et al., 2021). Mais ces fonctions ne sont que des hypothèses et des études doivent être menées pour les vérifier. Par exemple, l'étude du mutant *Atsht* d'*A. thaliana*, qui ne produit plus de phénolamides dans le pollen (Grienenberger et al., 2009), présente des modifications morphologiques du grain de pollen et de la structure de la sporopollénine du manteau pollinique. Malgré ces modifications, celui-ci est toujours viable et fertile.

III.2 Les tri-hydroxycinnamoyl spermidines et leurs bienfaits

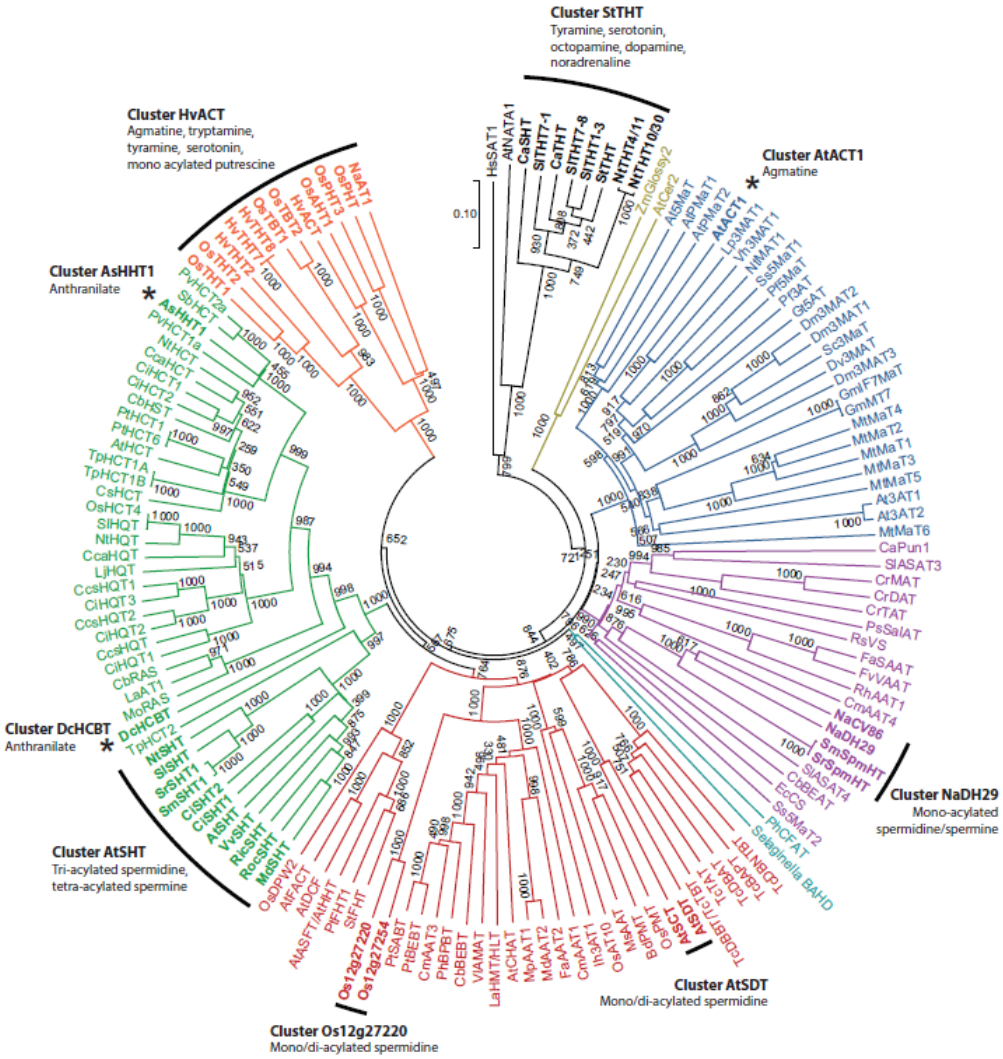
Les THCSpd disposent d'activités biologiques très intéressantes du point de vue pharmaceutique. En effet, la tri-*p*-coumaroyl spermidine est capable d'inhiber *in vitro* 70% de l'activité de la protéase du VIH-1, enzyme nécessaire à la synthèse des protéines virales essentielles, et donc à la virulence du virus (Ma et al., 2001). Elle a, de plus, été identifiée comme antagoniste du transporteur de la sérotonine (SERT) et pourrait donc avoir un effet antidépresseur (Zhao et al., 2010). Par ailleurs, la tri-*p*-coumaroyl spermidine est responsable de la protection du pollen de *Helianthus annuus* contre la détérioration fongique (Kyselka et al., 2018), suggérant une activité antifongique potentiellement intéressante. Plus particulièrement, elle inhibe la croissance d'*Aspergillus niger*, *Penicillium verucosum* et *Fusarium culmorum*. D'autre part, le conjugué tri-dihydrocaféoyl polyamine présente une activité antibactérienne importante contre les souches de *Staphylococcus aureus* résistantes à la vancomycine (VRSA) (Yingyongnarongkul et al., 2008). En outre, certains isomères de tétra-*p*-coumaroyl spermine et tétra-*p*-coumaroyl thermospermine extraits du pollen de *Matricaria chamomilla* sont des antagonistes des récepteurs à la neurokinine-1 (NK1), ils disposent donc d'actions antiémétique, antidépresseur et anxiolytique, ainsi qu'une activité antitumorale contre le cancer du sein (Yamamoto et al., 2002 ; Park et al., 2017). Finalement, la N1,N5,N10-tricaféoylspermidine dispose d'une activité inhibitrice de la catéchol-*O*-méthyltransférase humaine impliquée dans la maladie de Parkinson et la dépression (Miyata et al., 2022). Le pollen est utilisé en tant que complément alimentaire depuis longtemps, il est réputé notamment

pour ses bienfaits sur le métabolisme énergétique, le fonctionnement du système nerveux et contre la fatigue. Les bienfaits sur la santé attribués au pollen pourraient être liés en partie à l'accumulation de THCSpd dans les grains de pollen.

III.3 Voie de biosynthèse

Les phénolamides sont à la jonction entre le métabolisme primaire des amines, et le métabolisme spécialisé des phénylpropanoïdes (Bassard et al., 2010). Leur biosynthèse ne nécessite qu'une seule enzyme spécifique à activité *N*-hydrocycinnamoyltransférase à partir de précurseurs préexistants issus d'autres métabolismes. A ce jour, 43 *N*-hydroxycinnamoyltransférases ont été caractérisées (Roumani et al., 2021). Elles appartiennent à deux familles d'acyltransférases distinctes: principalement à la superfamille des acyltransférases BAHD (Roumani et al., 2021), mais quelques *N*-hydroxycinnamoyltransférases sont des homologues de la superfamille des acyltransférases GNAT (Gen-5-related N-acyltransferases).

La superfamille des GNAT regroupe des *N*-acyltransférases retrouvées dans tous les organismes vivants. Ces enzymes acylent le groupement amine de diverses molécules et sont divisées en 17 sous-familles, dont les SSAT (diamine *N*-acyltransférase ou spermidine/spermine *N*1-acyltransférase) impliquées dans le catabolisme des polyamines (Lou et al, 2016) à partir desquelles un cluster de *N*-hydroxycinnamoyltransférases spécifiques de l'agmatine, la tryptamine, la tyramine, la serotonine et la putrescine a émergé. La superfamille des BAHD regroupe des acyltransférases capables de *O*-acyler ou *N*-acyler une grande variété de substrats dépendamment d'un acyl-CoA. Elles sont impliquées dans la synthèse de nombreux MS tels que la lignine, les composés phénoliques, les alcaloïdes, les phytoalexines ou encore les anthocyanes. Plus de 50 BAHD ont été caractérisées et un classement en huit clades et sous-clades a été mis en place sur la base du type de substrat utilisé (D'auria, 2006 ; Tuominen et al., 2011). Parmi les BAHD, les *N*-hydroxycinnamoyltransférases ne forment pas un cluster unique homogène mais huit clusters indépendants qui ont émergé séparément dans plusieurs des clades de BAHD. Par exemple, les agmatine coumaroyl transférases (ACT) qui *N*-acylent l'agmatine chez *Hordeum vulgare* et *A. thaliana* (Burhenne et al., 2003 ; Muroi et al., 2009), la spermidine disinapoyl transférase (SDT) et la spermidine dicoumaroyltransférase (SCT) qui *N*-acylent les amines primaires de la spermidine chez *A. thaliana* (Luo et al., 2009),



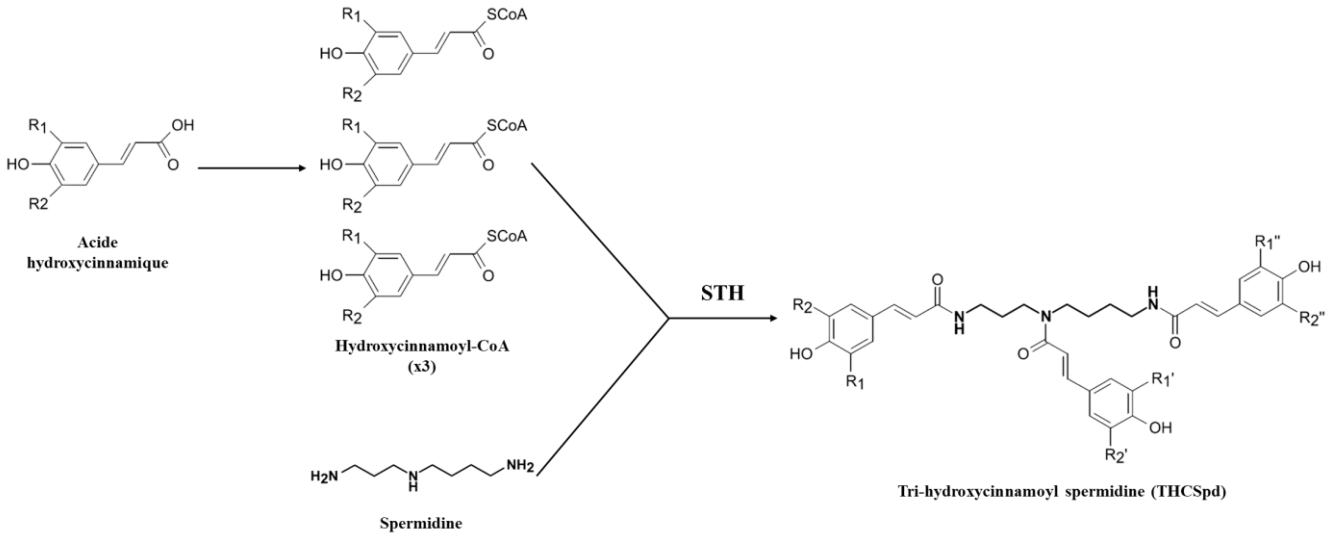
<u>Figure 55</u> : Phylogénie des N-hydroxycinnamoyltransférases caractérisées. (Roumani et al., 2021)

Les séquences de protéines ont été alignéees à l'aide de ClustalX et l'arbre phylogénique a été construit à l'aide de Neighbor-Joining et d'une analyse bootstrap. Les noms en noir correspondent à la famille DAT des acyltransférases GNAT pour laquelle seuls deux membres représentatifs caractérisés chez les mammifères et les plantes sont mentionnés (HsSAT1 et AtNATA1). Les noms colorés représentent les clades d'acyltransférases BAHD pour lesquels chaque membre caractérisé est mentionné : clade Ia (bleu), clade II (jaune), clade IIIa (magenta), clade IIIb (cyan), clade IV (orange), clade Va (rouge), et clade Vb (vert). Les N-hydroxycinnamoyltransférases sont indiquées en gras et les clusters sont mis en évidence par des lignes sombres ou marquées d'un astérisque. Les accepteurs d'amine utilisés par les membres de chaque groupe de N-hydroxycinnamoyltransférase sont mentionnés.

ou encore l'hydroxycinnamoyl-CoA-spermidine acyltransférase (DH29) et l'hydroxycinnamoyl-CoA-putrescine (AT1) *N*-acylant respectivement la spermidine et la putrescine chez *Nicotania attenuata* (Onkokesung et al., 2012). Les *N*-hydroxycinnamoyltransférases ont ainsi été classées phylogénétiquement selon neuf clusters : StTHT, HvACT, AtACT1, DcHCBT, AsHHT1, NaDH29, AtSDT, Os12g27220 et AtSHT (**Figure 55**) (Roumani et al., 2021). Ces neufs clusters se singularisent par les amines *N*-acylées. La spécificité de substrat, la catégorie et le nombre de groupes aminés acylés sur les accepteurs, la distribution au sein des plantes et entre espèces sont autant de facteurs qui différencient chaque cluster de *N*-hydroxycinnamoyltransférases et qui explique la diversité des formes et des fonctions des phénolamides.

L'enzyme *N*-hydroxycinnamoyltransférase responsable de la synthèse des THCSpd a été caractérisée pour la première fois chez *A. thaliana* (Grienenberger et al., 2009). Cette enzyme particulière est la Spermidine *N*-HydroxycinnamoylTransférase (SHT). C'est une *N*-hydroxycinnamoyltransférase appartenant au clade Vb des BAHD (Tuominen et al., 2011) et forment un cluster unique de *N*-hydroxycinnamoyltransférases issus de la néofonctionnalisation, c'est à dire, de la divergence fonctionnelle, dans l'ancêtre commun des dicotylédones, de la shikimate *O*-hydrocycinnamoyl transférase impliquée dans la synthèse de la lignine. Elle est capable de catalyser la *N*-substitution successivement sur les groupements amine primaires et secondaires de la spermidine, permettant une saturation complète de celle-ci (**Figure 56**) (Grienenberger et al., 2009). La première SHT a été identifiée chez *A. thaliana*, mais d'autres SHT ont, depuis, été caractérisées chez *Malus domestica* (Elejalde-Palmett et al., 2015), *Solanum melongena* et son parent sauvage *Solanum richardii* (Peng et al., 2016) et chez *Cichorium intybus* (Delporte et al., 2018). Ces différents exemples montrent que les SHT disposent d'une grande promiscuité de substrat pour les donneurs d'acyle, ce qui explique la diversité des substitutions des THCSpd disponibles chez les végétaux. Au niveau des accepteurs d'acyle, les SHT caractérisées sont spécifiques de la spermidine sauf chez la chicorée où une modification de spécificité vers la spermine est à l'origine du changement de la nature des THCSpd observés chez les Astéracées comme indiqué précédemment (Bernard et al., 2022).

Les enzymes du cluster SHT sont exclusivement exprimées dans les cellules du tapetum des étamines (Grienenberger et al., 2009 ; Elejalde-Palmett et al., 2015). Le tapetum est une couche de cellules nutritives dans les anthères des fleurs en développement, qui assure le développement et la maturation des grains de pollen (Vogt, 2018). La synthèse des THCSpd



<u>Figure 56</u> : Synthèse de la tri-hydroxycinnamoyl spermidine (THCSpd).

Les STH sont capables de catalyser la *N*-substitution successive des groupements amine primaires et secondaires de la spermidine, permettant une saturation complète de celle-ci.
R1 et R2 = H, OH ou OMe.

serait réalisée au niveau du tapetum, qui, en se résorbant, permettrait le dépôt des THCSpd synthétisées sur le pollen, expliquant ainsi l'accumulation de ces composés sur le manteau du pollen.

Les *N*-hydroxycinnamoyltransférases sont généralement les uniques enzymes impliquées dans la biosynthèse des phénolamides, la nature des substitutions ajoutées aux groupements amine étant dépendante de chaque enzyme. La synthèse des THCSpd chez les Brassicaceae semble pourtant montrer que des exceptions existent. Un cytochrome P450, CYP398A8 appartenant au clade des CYP98A, et une *O*-méthyltransférase d'*A. thaliana* (Tapetum-Specific Methyltransferase AtTSM1) sont impliqués dans la modification des groupements associés aux fonctions amines après la formation des THCSpd (Matsuno et al., 2009 ; Vogt, 2010). La multiplicité de structures des THCSpd est d'autant plus grande que d'autres enzymes peuvent être impliquées dans leur diversification. En effet, la diversité chimique impliquée dans les voies de biosynthèse des dérivés de phénylpropanoïdes est dictée par les cytochromes P450 et les *O*-méthyltransférases (Vogt, 2010), il est donc probable que ces familles d'enzymes soient aussi impliquées dans la diversification des phénolamides.

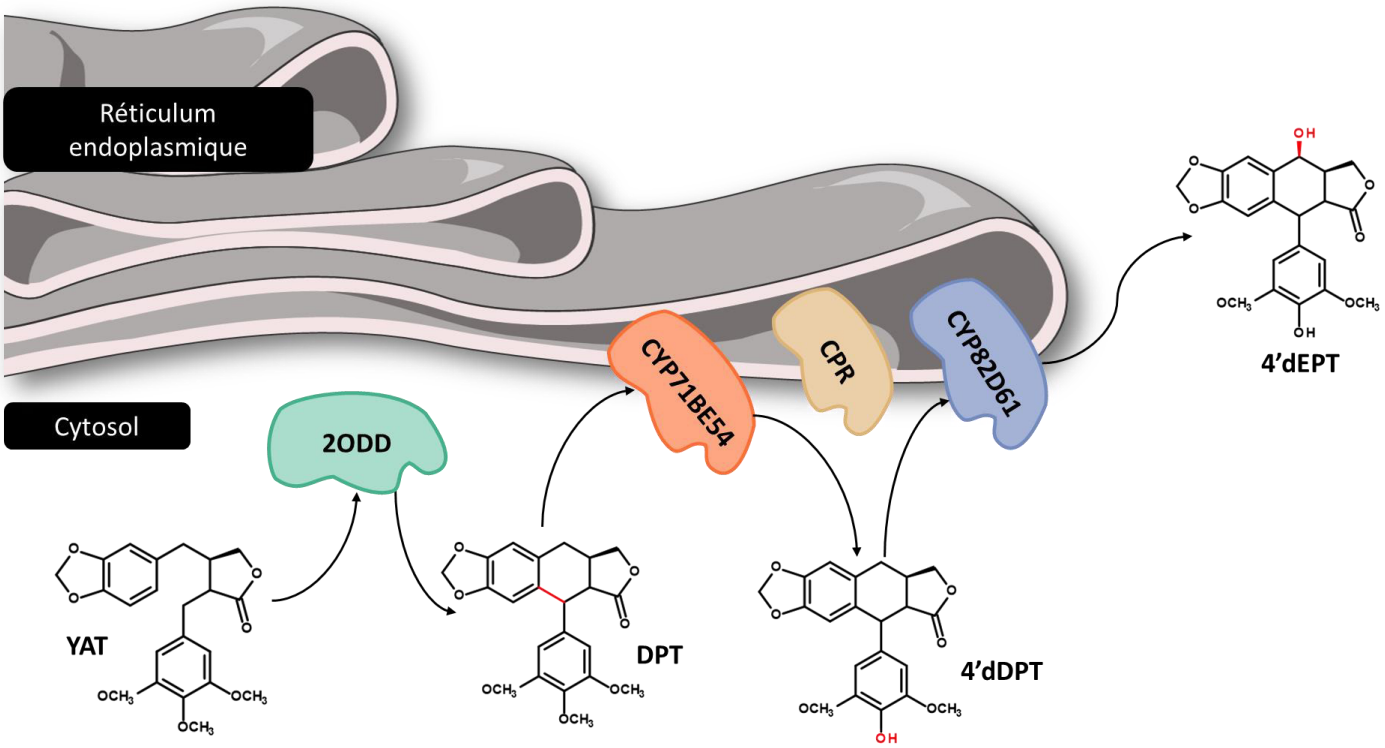
III.4 Limites à l'étude des THCSpd

Les THCSpd disposent d'activités très intéressantes du point de vue pharmaceutique, mais la connaissance de leurs applications prometteuses est impactée par l'approvisionnement de ces molécules. En effet, les connaissances actuelles ont pu être mises en évidence grâce à la purification de THCSpd à partir de pollen ou bien une synthèse chimique. Cependant, des études plus importantes, et plus abouties, de ces molécules restent limitées par leur accumulation stricte dans l'enveloppe des grains de pollen. De plus, les pollens sont produits de façon saisonnière et pendant une courte période, limitant ainsi les possibilités d'accès aux THCSpd. Aussi, l'absence de produits commerciaux correspondant à ces molécules en limite considérablement les études pharmacologiques, voire les applications futures potentielles. Ces barrières constituent un frein dans la recherche sur les propriétés de ces métabolites et sur leur diversité. Aujourd'hui, la plupart des études réalisées sur ces molécules portent sur la tri-*p*-coumaroyl spermidine, la THCSpd la plus courante. Mais jusqu'à 125 THCSpd peuvent être envisagées compte tenu des cinq acides hydroxycinnamiques présents chez les végétaux et des différentes combinaisons possibles avec les groupements hydroxycinnamiques. De plus,

l'implication d'autres familles d'enzymes dans la synthèse de ces composés contribue à complexifier et à diversifier leur structure chimique. Cette diversité probable de structures pourrait sans doute être corrélée à une diversité d'effets pharmaceutiques intéressants. De plus, l'accès à ces structures pourraient permettre de mieux comprendre le rôle des THCSpd chez la plante.

Il est possible de synthétiser chimiquement certains THCSpd, mais les substitutions des groupements amine obtenues sont identiques, ce qui limite la diversité des molécules pouvant être produites. De plus, il serait nécessaire d'ajouter des groupements protecteurs des amines puis de les supprimer successivement afin d'ajouter les hydroxicinnamates (Hu and Hesse, 1996), ce qui diminuerait drastiquement les rendements de synthèse. Ces travaux de thèse présenteront la mise en place de souches de levures productrices de THCSpd permettant d'accéder à la structure et la diversité de ces molécules.

OBJECTIF ET ORGANISATION DE LA THESE



<u>Figure 57</u> : Schéma de la stratégie de bioconversion de 4'dEPT à partir de YAT en levure.

Les modifications apportées par les enzymes sur la structure des molécules sont représentées en rouge.
YAT : (-)-yateine ; DPT : déoxypodophyllotoxine; 4'dDPT : (−)-4′-demethyldésoxypodophyllotoxine; 4'dEPT : 4′-desméthylepipodophyllotoxine.
2ODD : désoxypodophyllotoxine synthase de *P. hexandrum*; CYP71BE54 : cytochrome P450 de *P. hexandrum* ; CYP82D61 : cytochrome P450 de *P. Hexandrum ;* CPR : Cytocrhome P450 réductase.

L'étoposide est un lignane dérivé de podophyllotoxine présentant une forte activité inhibitrice des topoisomérases de type II et qui s'avère être, aujourd'hui, un des médicaments anticancéreux les plus utilisés en chimiothérapie. Malgré la valeur économique accrue de cette molécule, les rhizomes de la plante himalayenne *Podophyllum hexandrum* restent la principale source de podophyllotoxine. La forte demande en étoposide imposant une énorme pression sur cette ressource naturelle, *P. hexandrum* a été inscrite sur la liste de la Convention sur le commerce des espèces de faune et de flore sauvages menacées d'extinction, occasionnant de longues pénuries d'approvisionnement en étoposide. Les progrès des biotechnologies et des outils génétiques ont permis la production alternative de nombreux métabolites spécialisés végétaux dans des organismes hôtes hétérologues telle que la levure. Récemment, la voie de biosynthèse de la (-)-4'-desméthyl-épipodophyllotoxine (4'dEPT) à partir de (-)-yatéine (YAT) a été élucidée chez *P. hexandrum* (**Figure 51**). Néanmoins, la synthèse de la PPT à partir de l'alcool conifférylique est, non seulement longue avec au moins dix étapes enzymatiques, mais elle est, de plus, incomplète puisque la/les dernière(s) étape(s) ne sont pas connues. Cependant, comme présenté au paragraphe II.2.b de l'Introduction, la synthèse chimique de l'étoposide à partir de la PPT passe par la synthèse d'un intermédiaire: la 4'dEPT dont la voie de biosynthèse est connue. La création d'une souche de levure capable de bioconvertir la YAT hautement biodisponible en 4'dEPT directement utilisable pour une synthèse chimique d'étoposide semble donc être une possible stratégie alternative d'approvisionnement en étoposide. En effet, la YAT étant hautement biodisponible notamment chez certaines espèces de la famille des Cupressacea comme les Juniperus, elle pourra être facilement extraite et purifiée en grande quantité afin de produire de l'étoposide via une stratégie de bioconversion en levure. La stratégie de bioconversion employée dans ces travaux permettra ainsi de produire de la 4'dEPT à partir de YAT par clonage des trois enzymes 2ODD, CYP71BE54 et CYP82D61 (**Figure 57**). Ces deux dernières enzymes sont des cytochromes P450 qui doivent être réduites par une CPR afin d'être actives (Paragraphe I.2.a.3 de l'Introduction). La CPR endogène de la levure ne disposant pas d'une activité suffisante pour réduire les P450 hétérologues, le clonage d'AtATR1 ou AtATR2 d'*A. thaliana* dans les souches de levures contenant des P450 végétales hétérologues est utilisé de façon routinière. Une de ces deux CPR sera donc clonée dans la souche de levure « usine » productrice de 4'dEPT.

Dans ces travaux, des approches d'ingénierie métabolique pour la production de MS en levure, et plus spécifiquement de lignanes, ont donc été développées. Pour ce faire, plusieurs stratégies d'expression des transgènes ont été explorées. La réplication autonome via la

transformation de plasmides en levure a été la première stratégie mise en place. Pour ce faire, j'ai d'abord choisi de tester cette approche sur les phénolamides. En effet, les THCSpd sont des phénolamides composés d'un squelette de spermidine entièrement substitué par des acides hydroxycinnamiques. Ces MS disposent également d'activités pharmaceutiques intéressantes, telles que des activités antivirales, antidépressives, antifongiques, antibactériennes, antiémétiques, anxiolytiques et antitumorales. Cependant, l'étude des applications de ces molécules est altérée par l'approvisionnement de ces composés principalement accumulés dans les grains de pollen des Eudicotylédones. Un système de production alternatif de ces composés par ingénierie métabolique de la voie de synthèse des THCSpd en levure permettra d'étudier plus profondément les fonctions et les applications potentielles de la diversité des THCSpd.

La stratégie de bioconversion à partir des souches de levure productrices de THCSpd employée dans la **première partie** de ces travaux a permis de vérifier la possibilité de bioproduction de différentes structures de THCSpd à partir de divers acides hydroxycinnamiques et de la spermidine par clonage des gènes codant la 4-coumarate-CoA ligase (4CL) d'*A. thaliana* et la spermidine N-hydroxycinnamoyltransférase (SHT). En effet, la 4CL permet, dans un premier temps, de produire les hydroxycinnamoyl-CoA à partir des acides hydroxycinnamiques nécessaires à la synthèse des THCSpd. Dans un second temps, la SHT utilise les hydroxycinnamoyl-CoA produites en tant que donneurs d'acyle et la spermidine ajoutée au milieu de culture en tant qu'accepteur d'acyle pour la synthèse de THCSpd. Le clonage de ces enzymes a été réalisé par réplication autonome. En effet, cette stratégie de clonage est plus efficace et rapide à mettre en œuvre pour une preuve de concept, le but n'étant pas de produire de façon quantitative, mais d'exploiter la diversité des THCSpd et de vérifier la possibilité de mise en place d'une telle souche. L'ajout d'acides hydroxicinnamiques naturels et/ou plus exotiques a permis de produire différentes structures de THCSpd. L'utilisation de la souche produite comme plateforme de bioproduction de ces molécules à échelle laboratoire permettra une avancée majeure dans l'étude et la compréhension du rôle des THCSpd naturelles, mais permettra aussi de produire des THCSpd non naturelles qui pourraient avoir des activités biologiques intéressantes. Ces travaux ont fait l'objet d'un article publié en 2021 dans « ACS Synthetic Biology ».

La technique d'expression de transgènes par réplication autonome a ensuite été appliquée pour la production du lignane 4'dEPT. Ainsi, dans la **deuxième partie** de mes

travaux, différentes souches de levures à réplication autonome ont été mises en place afin de vérifier l'activité catalytique en levure de chacune des enzymes de la voie de bioconversion. L'étape d'hydroxylation terminale catalysée par un cytochrome P450 représentant un goulot d'étranglement pour la synthèse de 4'dEPT en levure, les rendements d'activité catalytique de cette dernière étape ont été optimisés notamment via le suivi de l'accumulation de l'enzyme, le criblage d'isoformes ou d'orthologues de l'enzyme, et l'optimisation de la réduction du P450 associé à cette étape.

Suite à l'optimisation de l'étape d'hydroxylation terminale de la voie de synthèse, la mise en place de la souche de levure « usine » productrice de 4'dEPT via intégration par le système CRISPR/Cas9 des gènes d'intérêt dans le génome de la souche industrielle *Saccharomyces cerevisiae* CEN-PK a été initiée. En effet, le système d'expression par réplication autonome présente de nombreux inconvénients pour une bioproduction à échelle industrielle, notamment, une altération de la croissance des levures et l'utilisation de marqueurs de sélection. Cette **troisième partie** de mes travaux présente donc la mise en place de la souche finale industrialisable et l'optimisation des rendements de bioconversion via l'optimisation de la réduction et de la traduction du cytochrome P450 catalysant l'hydroxylation finale de la voie de bioconversion, ainsi qu'un travail sur la balance des copies de gènes transférés. Cette partie des résultats fera bientôt l'objet d'un article.

Finalement, la **quatrième, et dernière partie** des résultats présente l'initiation de la montée en échelle de la bioproduction de 4'dEPT à partir de la souche la plus optimale. Pour ce faire, un protocole d'extraction et de purification de la YAT à partir de biomasse végétale a été mis en place, permettant de réaliser une première montée en échelle de la culture en bioréacteur.

Ces travaux de thèse présentent donc une approche globale de production de métabolites spécialisés végétaux en levure via l'exploitation de plusieurs stratégies d'expression des transgènes. Ces productions hétérologues ont une visée industrielle pour la production de 4'dEPT, et une visée plus fondamentale pour l'étude de la fonction des THCSpd.

RESULTATS

Partie 1 : Production de phénolamides par ingénierie métabolique

Les tri-hydroxycinnamoyl spermidines (THCSpd) sont des phénolamides composés d'un squelette de spermidine N-acylée sur l'ensemble de ses fonctions aminées par des groupements dérivés des acides hydroxycinnamiques. Les THCSpd sont des métabolites retrouvés spécifiquement dans l'enveloppe de pollen des Eudicotylédones. Bien que leur fonction biologique dans la plante ne soit pas encore élucidée, ces composés disposent d'activités pharmaceutiques intéressantes. Des activités antivirales contre le VIH-1 (Ma et al., 2001), antidépressives (Zhao et al., 2010), antifongiques (Kyselka et al., 2018), antibactériennes (Yingyongnarongkul et al., 2008), antiémétiques, anxiolytiques et antitumorales (Yamamoto et al., 2002 ; Park et al., 2017) ont été rapportées pour certaines THCSpd. Cependant, l'étude des applications de ces molécules est altérée par les difficultés d'approvisionnement de ces composés. Le pollen, qui est l'unique source de THCSpd, est produit de façon saisonnière, en petite quantité chez les dicotylédones et il est difficile à collecter. En outre, si plus de 125 structures de THCSpd peuvent être envisagées sur la base des cinq hydroxycinnamates majeurs produits par les plantes capables d'acyler la spermidine, assez peu ont été décrites dans la nature pour l'instant. Des systèmes de production alternatifs doivent donc être envisagés pour faciliter l'accès à ces molécules d'intérêts. Cette première partie des résultats présente un article décrivant la production de THCSpd par ingénierie métabolique en levure. Ces travaux collaboratifs que j'ai réalisés au début de ma thèse m'ont permis de me familiariser avec les techniques de clonage, de cultures et d'analyses métaboliques utilisées au laboratoire.

L'objectif de ces travaux n'est pas de produire une souche de levure industrialisable capable de produire de grandes quantités de THCSpd, mais de développer une souche de laboratoire pour produire une grande diversité de structures de THCSpd. Le transfert en levure de la voie de biosynthèse de THCSpd a donc été réalisée à l'aide de vecteurs d'expression à réplication autonome, afin de mettre en place rapidement différentes souches et de vérifier la faisabilité d'une production de ce type de molécules en levure. Une stratégie de bioconversion des THCSpd à partir d'acides hydroxycinnamiques ajoutés au milieu de culture a été choisie pour contrôler la nature des THCSpd produits en fonction des précurseurs ajoutés au milieu de culture. Les souches de levures transformantes disposent de deux enzymes hétérologues : la N-hydroxycinnamoyltransférase (SHT), enzyme clef permettant l'acylation d'esters-CoA sur la

spermidine, et la 4-coumarate-CoA ligase 4CL permettant la synthèse des esters-CoA donneurs d'acyle à partir des acides hydroxycinnamiques ajoutés au milieu de culture.

Dans un premier temps, la tri-*p*-coumaroyl spermidine, la THCSpd la plus étudiée, a pu être synthétisée en cultivant une levure contenant les gènes hétérogènes codant 4CL5 et SHT d'*A. thaliana* en présence d'acide *p*-coumarique et de spermidine. Ces résultats ont mis en lumière le détournement par le métabolisme de la levure de certains intermédiaires dans la synthèse de THCSpd. Le niveau de production de la souche a donc été amélioré en adaptant le châssis de la levure par mutation/substitution des gènes endogènes PAD1 et TSC13. Dans un second temps, afin d'explorer la diversité des THCSpd pouvant être produites, des levures recombinantes exprimant des combinaisons d'isoformes de 4CL d'*A. thaliana* et d'orthologues de SHT issus de 5 espèces différentes ont été cultivées en présence de divers acides hydroxycinnamiques seuls ou en mélanges. Ainsi, huit THCSpd naturelles ont été obtenues et nous ont permis d'identifier les meilleures combinaisons d'enzymes pour l'acylation de la spermidine en fonction du (des) acide(s) hydroxycinnamique(s) ajouté(s) dans le milieu de culture. Finalement, la promiscuité de substrats de 4CL et SHT ont été exploitées pour produire onze THCSpd non naturelles en cultivant les levures recombinantes en présence d'acides hydroxycinnamiques exotiques, d'acides dihydroxycinnamiques ou d'acides hydroxycinnamiques halogénés. Au total, dix-neuf THCSpd comportant 12 substitutions d'acyle différentes ont été produites dans le cadre de ces travaux prouvant ainsi la possibilité de synthèse à façon d'une grande diversité de THCSpd à l'aide d'une plateforme de production en levure. Cet outil permettra d'étudier la diversité des THCSpd naturelles mais aussi de synthétiser des analogues dont les applications biologiques peuvent être prometteuses.

ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

Exploiting Spermidine *N*-Hydroxycinnamoyltransferase Diversity and Substrate Promiscuity to Produce Various Trihydroxycinnamoyl Spermidines and Analogues in Engineered Yeast

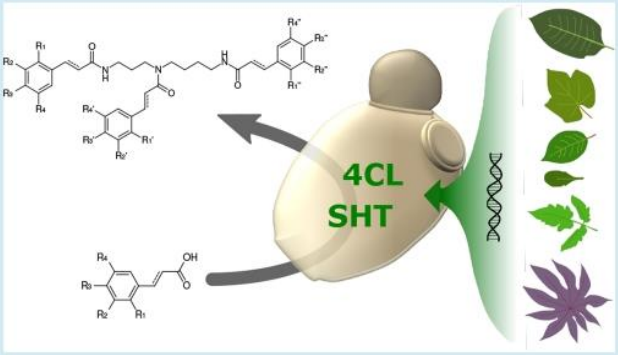
Jennifer Perrin,[§] Natalja Kulagina,[§] Marianne Unlubayir, Thibaut Munsch, Inês Carqueijeiro, Thomas Dugé de Bernonville, Johan-Owen De Craene, Marc Clastre, Benoit St-Pierre, Nathalie Giglioli-Guivarc'h, David Gagneul, Arnaud Lanoue, Vincent Courdavault, and Sébastien Besseau*

Cite This: *ACS Synth. Biol.* 2021, 10, 286−296

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Trihydroxycinnamoyl spermidines (THCSpd) are plant specialized metabolites with promising pharmacological activities as antifungals, antibacterial, antiviral, and antidepressant drugs. However, their characterization and potential pharmaceutical exploitation are greatly impaired by the sourcing of these compounds, restricted to the pollen of core Eudicot plant species. In this work, we developed a precursor-directed biosynthesis of THCSpd in yeast using a dual enzymatic system based on 4-coumarate-CoA ligases (4CL) and spermidine *N*-hydroxycinnamoyltransferases (SHT). The system relies on the yeast endogenous spermidine pool and only requires hydroxycinnamic acids as exogenous precursors. By exploring 4CL isoforms and SHT diversity among plants, we have driven the production of 8 natural THCSpd, using single or mixed hydroxycinnamic acid precursors. Substrate promiscuities of 4CL and SHT were genuinely exploited to produce 8 new-to-nature THCSpd from exotic hydroxycinnamic and dihydrohydroxycinnamic acids, together with 3 new-to-nature THCSpd containing halogenated hydroxycinnamoyl moieties. In this work, we established a versatile and modular biotechnological production platform allowing the tailor-made THCSpd synthesis, constituting pioneer metabolic engineering for access to these valuable natural products.



KEYWORDS: *trihydroxycinnamoyl spermidines, phenolamides, yeast engineering, N-hydroxycinnamoyltransferase, precursor-directed biosynthesis, 4-coumarate-CoA ligase*

Phenolamides are specialized metabolites related to alkaloids, restricted to vascular plants, and combining a phenolic moiety with an amine through an amide bond.[1] Trihydroxycinnamoyl spermidines (THCSpd) are particular phenolamides composed of a spermidine backbone, fully substituted on amino groups with hydroxycinnamoyl moieties (*i.e.*, *p*-coumaroyl, caffeoyl, feruloyl, 5-hydroxyferuloyl, or sinapoyl). Since the discovery of tri-*p*-coumaroyl spermidine 30 years ago,[2] THCSpd have been demonstrated as universal components of the pollen coat in core eudicots,[1] the main occurrence of these compounds reported to date in nature with only scarce exceptions, such as in roots of *Microdesmis keayana*.[3] Although their biological function in pollen fitness remains enigmatic,[4] interesting biological and pharmacological activities have been attributed to these singular compounds. On the basis of bioassay-guided fractionation of a pollen extract, tri-*p*-coumaroyl spermidine was identified as the compound responsible for sunflower pollen protection against fungal spoilage through growth inhibition of mycotoxigenic fungi like *Aspergillus niger*, *Penicillium verucosum*, and *Fusarium culmorum*.[5] Besides, tri-*p*-coumaroyl spermidine was reported as antagonist of serotonin transporter (SERT), one of the favorite targets of common antidepressant drugs,[6] and has been found to inhibit *in vitro* 70% of the HIV-1 protease activity, which is essential for virus virulence.[7] An antibacterial activity has been demonstrated for tridihydrocaffeoyl polyamines analogues against vancomycin-resistant *Staphylococcus aureus* strains.[8] Finally, the THCSpd-related tetra-*p*-coumaroyl

Received: July 23, 2020
Published: January 15, 2021



ACS Publications

© 2021 American Chemical Society

spermine/thermospermine are neurokinin-1 receptor (NK1) antagonists with potential application in treatment of mood disorders and anxiety, as well as in cancers through antitumor activity.[9,10]

However, knowledge on promising THCSpd applications is greatly impaired by the sourcing of these compounds. Although THCSpd are widely distributed in core eudicots, their accumulation, strictly restricted to pollen grain, is a major constraint:[1] THCSpd are produced seasonally during a short period, in a very low quantity within the entire plant, and pollens are intricate to collect in substantial amounts. Consequently, high scale purification of THCSpd from biological material is not attractive for suppliers, and the research in the field is usually done from homemade purifications of THCSpd.[6,7] In addition to the quantitative limits, access to THCSpd diversity is a second constraint. Most studies have focused on tri-*p*-coumaroyl spermidine, the most current THCSpd. However, considering the five hydroxycinnamic acids produced in plants and the combinations of various hydroxycinnamoyl moieties among spermidine amino groups, up to 125 diversely substituted THCSpd can be envisaged. To date, only a few of these have been reported,[11] and an exhaustive study of THCSpd accumulation pattern in pollen was only performed in *Arabidopsis*,[12] limiting knowledge and *de facto* access to THCSpd diversity from biological materials. An alternative could be chemical synthesis, which allows introducing various hydroxycinnamoyl moieties on spermidine, but the three amino group substitutions should be identical.[5,8,13]

To improve functional studies and increase the discovery rate of THCSpd applications, alternative production systems have to be developed. Biotechnological production in cell factories can be considered to increase availability and range of THCSpd. For decades, yeast has turned out to be one of the most efficient and valuable heterologous production systems for complex plant specialized metabolites.[14,15] Yeast has been engineered for precursor-directed biosynthesis or *de novo* production of polyphenol derivatives from hydroxycinnamic acids, including flavonoids, dihydrochalcones, stilbenes, or anthranilates.[16−21] However, such bioproduction systems rely on the deciphering of plant biosynthesis pathways prior transposition into microorganisms.

The metabolism of THCSpd involves a single enzyme with *N*-hydroxycinnamoyltransferase activity, able to use precursors from the specialized metabolism of phenylpropanoids (hydroxycinnamoyl-CoA) as acyl donors, and polyamine primary metabolisms (spermidine) as acyl acceptor.[22] Such enzymes are called spermidine *N*-hydroxycinnamoyltransferase (SHT) and belong to the BAHD acyltransferase superfamily.[23] SHTs define a subgroup in the clade Vb of BAHD,[24] which displays high sequence homology, together with a restricted expression pattern in tapetum cells of etamines, which is responsible for universal THCSpd accumulation on the pollen coat.[1,22] In contrast to other *N*-hydroxycinnamoyltransferases, SHTs have the unique feature to catalyze N-substitution on both primary and secondary amino groups, allowing full saturation of spermidine with a unique enzyme.[22] Identified for the first time in *Arabidopsis*, SHT has been further characterized in apple trees, eggplants, and chicory.[1,25,26] In this last example, neofunctionalization occurred with a shift in acyl acceptor specificity from spermidine to spermine, as generally observed in Asteraceae. A large substrate promiscuity among acyl donors seems to be a common trait of characterized SHT, while affinities differ and seem to be one of the main factors responsible for THCSpd diversity among species, in addition to the pool of hydroxycinnamic acids available in tissues.[1,25,26]

In this study, 4-coumarate-CoA ligase (4CL) and SHT enzymes were coexpressed in engineered *S. cerevisiae* strains to perform precursor-directed biosynthesis of THCSpd from hydroxycinnamic acids and spermidine. In this dual enzymatic system (Figure 1), 4CL converted the stable and commercially

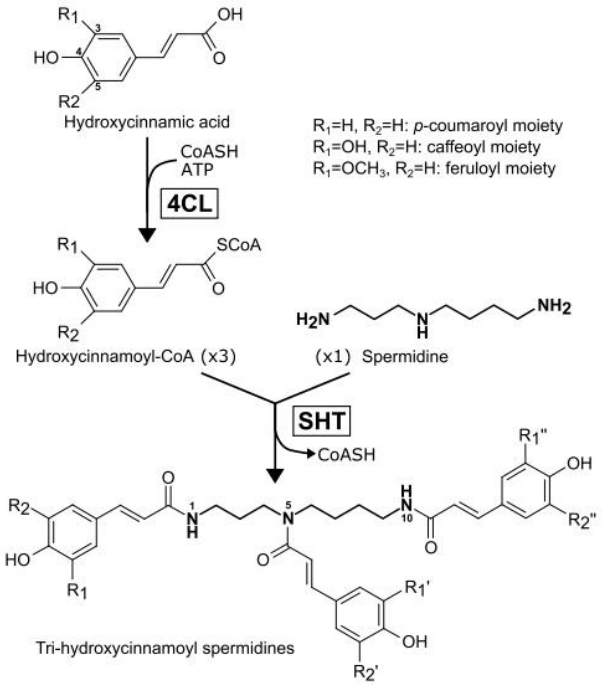


Figure 1. Enzymatic reactions catalyzed by 4-coumarate-CoA ligases (4CL) and spermidine *N*-hydroxycinnamoyltransferases (SHT).

available hydroxycinnamic acids into corresponding CoA esters,[27,28] which are in turn used by SHT as the acyl donors, as previously reported for microorganisms engineered with BAHD enzymes.[18,19,29] In a first attempt, we produced the common tri-*p*-coumaroyl spermidine using the *Arabidopsis* 4CL5 and SHT. By exploiting the large substrate promiscuity of 4CL and SHT, as well as the distinct substrate affinities within sequences available among plants, we further produced a large diversity of THCSpd from various hydroxycinnamic acids and analogues.

RESULTS AND DISCUSSION

Production of THCSpd in Engineered Yeast. To evaluate the ability of producing THCSpd in *S. cerevisiae*, we first assessed precursor-directed biosynthesis of tri-*p*-coumaroyl spermidine using a strain harboring At4CL5 and AtSHT. Both enzymes were coexpressed in yeast using the pESC series of episomal vectors, under galactose inducible promoters. Culture of the transformed yeast strain in the presence of *p*-coumaric acid and spermidine as exogenous precursors (Figure 2) led to the accumulation in cell culture supernatant of a compound with the same retention time (13.28 min), mass to charge ratio (ESI+ 584.1), and mass fragmentation profile (Supporting Information; Figure S1) as N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

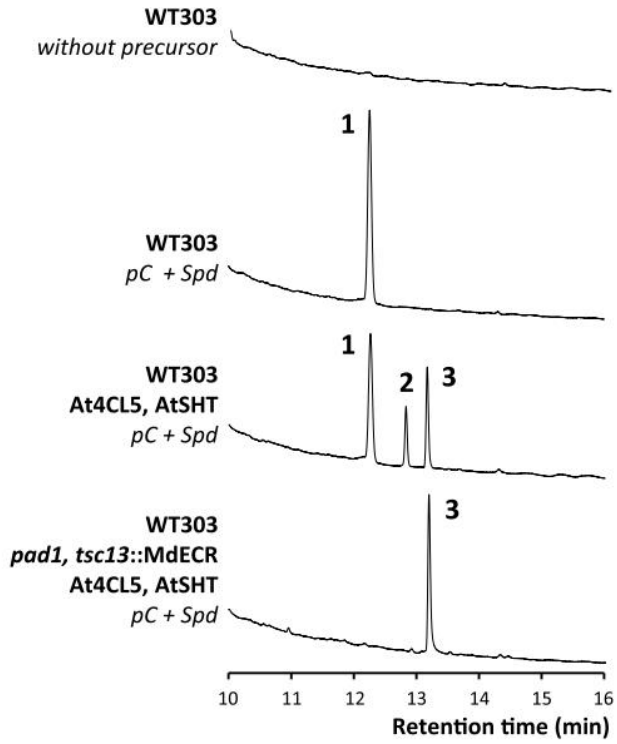


Figure 2. Production of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine in yeast. Supernatants of yeast growth media were analyzed by LC-MS/MS. UV chromatograms (Maxplot) of wildtype yeast (WT303) grown without exogenous precursor or fed with *p*-coumaric acid (pC) and spermidine (Spd) are compared to wild type yeast (WT303) or the *pad1*, *tsc13*::MdECR mutant both expressing At4CL5 and AtSHT, and fed with *p*-coumaric acid (pC) and spermidine (Spd). MS and MS/MS profile of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (3) obtained is presented in Supporting Information (Figure S1). Styrene-derivative (1) and *p*-dihydrocoumaric-derivative (2) are byproducts and correspond to 4-hydroxystyrene (Supporting Information; Figure S2) and *N*-*p*-dihydrocoumaroyl, N',N''-di-*p*-coumaroyl spermidine (Supporting Information; Figure S1) respectively.

In addition to the product of interest, two main byproducts appeared (Figure 2). The first one was also observed in untransformed yeast fed with *p*-coumaric acid and corresponds to 4-hydroxystyrene (Supporting Information; Figure S2). As previously reported, the yeast phenyl lactate decarboxylase enzymes Pad1 and Fdc1 metabolize hydroxycinnamic acids (mainly cinnamic, *p*-coumaric and ferulic acids) into styrene derivatives.[30] The second byproduct was not observed in untransformed yeast and corresponds to a *p*-dihydrocoumaroyl spermidine derivative (*i.e.*, *N*-*p*-dihydrocoumaroyl, N',N''-di-*p*-coumaroyl spermidine; Supporting Information; Figure S1). In engineered yeast dedicated to flavonoid production, the yeast enoyl reductase Tsc13, involved in long fatty acids synthesis, has been reported to reduce *p*-coumaroyl-CoA into *p*-dihydrocoumaroyl-CoA leading to further phloretic acid (so-called *p*-dihydrocoumaric acid) accumulation.[31] In our case, the *p*-dihydrocoumaroyl-CoA was used as a substrate by SHT to acylate spermidine and competed with *p*-coumaroyl-CoA. To circumvent these artifactual activities and substantially increase accumulation of the product of interest (Figure 2), a *pad1* mutant was first produced using a KanMX4 cassette,[32] which is sufficient to lose artifactual decarboxylase activity.[30] The essential gene *TSC13* ORF was then complemented by the MdECR ORF from apple tree, which displays the same essential function, but lacks its side activity on hydroxycinnamoyl-CoA.[31]

Starting from the yeast mutant *pad1*, *tsc13*::MdECR chassis and taking advantage of At4CL5 and AtSHT large substrate affinity among hydroxycinnamoyl moieties,[1,22,27,28] we investigated the production of other THCSpd by growing recombinant yeasts with various hydroxycinnamic acids as precursors (*i.e.*, cinnamic, caffeic, or ferulic acids). N^1,N^5,N^{10}-Triferuloyl spermidine was successfully produced in a very high quantity compared to N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine; however, caffeic and cinnamic acids were not metabolized at all (Figure 3). The synthesis of THCSpd from *p*-coumaric and

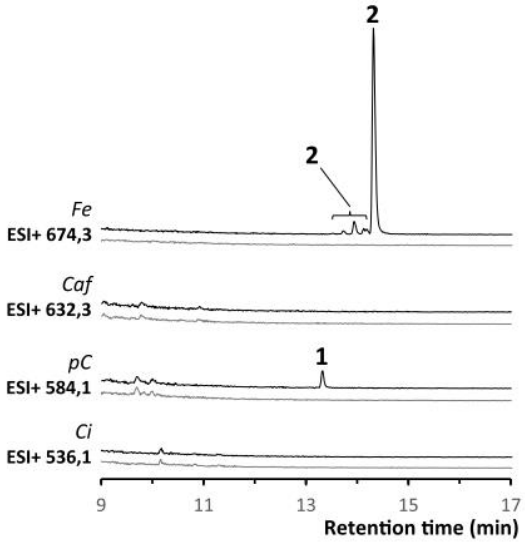


Figure 3. Screen for hydroxycinnamic acids as acyl donor precursors for THCSpd production in yeast. LC-MS mass chromatograms of the culture medium of yeast mutant *pad1*, *tsc13*::MdECR expressing At4CL5 and AtSHT, fed with spermidine in addition to cinnamic acid (Ci), *p*-coumaric acid (pC), caffeic acid (Caf), or ferulic acid (Fe) (upper dark line). Lower gray lines correspond to control yeast grown without precursors. Selected mass to charge ratios in positive ion mode are used depending on attempted products. MS and MS/MS profile of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (1) and N^1,N^5,N^{10}-triferuloyl spermidine (2) are presented in Supporting Information (Figure S1). Additional lower peaks prior N^1,N^5,N^{10}-triferuloyl spermidine (2) correspond to photoisomers.[45]

ferulic acids were finally assessed without the addition of spermidine in the culture medium, since the polyamine metabolism of yeast includes spermidine and putrescine as main compounds. No significant differences were found in the production level of both N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine and N^1,N^5,N^{10}-triferuloyl spermidine, with or without exogenous applications of spermidine (Figure 4). The pool of endogenous spermidine in yeast was thus not limiting in our production conditions, and further investigations were performed without spermidine in the culture medium. However, for N^1,N^5,N^{10}-triferuloyl spermidine, the most efficiently produced THCSpd in yeast, data distribution and *p*-value of production suggest that the spermidine pool is close to being a limiting factor (Figure 4). Engineering of yeast polyamine biosynthesis should thus be considered in a future attempt to optimize the production of THCSpd.[33,34]

Exploring 4CL and SHT Diversity in Plants to Improve THCSpd Synthesis. To improve THCSpd production, we

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

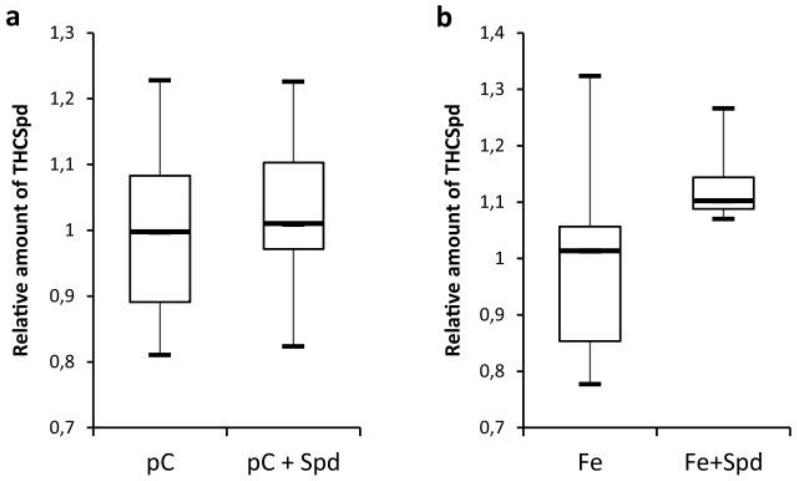


Figure 4. Evaluation of the ability of yeast to exploit endogenous spermidine pool for the production of THCSpd. The production level of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (a) and N^1,N^5,N^{10}-triferuloyl spermidine (b) in yeast mutant *pad1*, *tsc13*::MdECR expressing At4CL5 and AtSHT were compared with or without exogenous spermidine (Spd) in the culture medium in addition to *p*-coumaric acid (pC) or ferulic acid (Fe). Relative amounts were measured by LC-MS. There is no statistical difference between treatments (Wilcoxon rank sum test, $p = 0.9372$ for pC and $p = 0.06494$ for Fe; $n = 6$).

explored the diversity of 4CL and SHT enzymes and screened for combination couples. SHT was first discovered in *Arabidopsis* and an orthologue was further characterized in apple tree.[1] Their distinct acyl donor specificities, in accordance with the divergence of the THCSpd accumulated in the pollen coat of these two species, prompted us to isolate additional SHTs to expand the diversity of the generated products. In a previous work, phenolamide contents were investigated in the pollen grains among core eudicots species.[1] On the basis of these metabolic data, five species were selected with divergent THCSpd accumulation profiles, considering they could contain SHT with different acyl donor affinity. Putative SHTs from tobacco, tomato, castor bean, grapevine, and dog rose were thus cloned through genomic and transcriptomic data analysis. The optimized yeast strain WT303 *pad1*, *tsc13*::MdECR was transformed with At4CL5 and one of the seven SHTs, and further fed with various hydroxycinnamic acids (Figure 5a,c,e). First of all, the newly cloned NtSHT, SlSHT, RicSHT, VvSHT, and RocSHT effectively catalyzed the synthesis of THCSpd, thus confirming they correspond to *bona fide* SHTs (GenBank MN787045, MN787044, MN787043, MN787047, and MN787046). Second, in our experimental conditions, a significantly higher production of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine and N^1,N^5,N^{10}-triferuloyl spermidine was obtained with NtSHT and MdSHT, respectively, compared to AtSHT. Finally, RicSHT was shown to produce an additional THSCSpd N^1,N^5,N^{10}-tricaffeoyl spermidine. Besides, 4CLs are well characterized in *Arabidopsis* and among the four isoforms, At4CL1, 2, and 5 have marked substrates preferences.[27,28] Each 4CL was associated with the best SHT producers of THCSpd (Figure 5b,d,f). The highest synthesis rates were obtained for isoforms 2 and 5 with *p*-coumaric acid, isoform 5 with caffeic acid, and isoforms 1 and 5 with ferulic acid. These results are significantly different from comparative 4CL isoforms and SHT orthologues catalytic efficiency reported through *in vitro* studies.[22,27,28,35] Experiments were conducted in rich and nonselective medium to ensure the highest yeast biomass and the most efficient interest product synthesis. A control experiment, conducted in minimum selective medium and shown in Supplemental Figure S3, indicates that plasmid stability seems not to be a significant factor in the differences observed between 4CLs and SHTs. However, several parameters can reduce intrinsic enzyme efficiency and influence the conversion rate in a specific biological system, such as enzyme expression level (transcript stability, protein stability/folding in yeast) and enzyme activity in the yeast cell environment. Therefore, these results specifically indicate the most interesting sequences to use in yeast and to focus on for further sequence optimizations. To summarize, the couples At4CL2/NtSHT, At4CL5/RicSHT, and At4CL5/MdSHT are among the best THCSpd producers in yeast depending on hydroxycinnamic acid used (Table 1). However, a marked difference in THCSpd amounts is still observed as shown in Supplemental Figure S4: N^1,N^5,N^{10}-triferuloyl spermidine ≫ N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine ≫ N^1,N^5,N^{10}-tricaffeoyl spermidine.

Production of a Wide Diversity of Natural THCSpd. Feeding recombinant yeasts with a single hydroxycinnamic acid ensured identical decoration of spermidine amino groups. However, the substrate diversity of both 4CL and SHT could be exploited to create THCSpd with different substitution on the three spermidine amino groups. To this aim, yeast harboring the most versatile At4CL and SHT was fed with a combination of hydroxycinnamic acids. The production of THCSpd, containing both *p*-coumaroyl and feruloyl moieties, was first attempted using At4CL5 and NtSHT; however, only N^1,N^5,N^{10}-triferuloyl spermidine was produced. Although At4CL5 and NtSHT are among the most efficient to use *p*-coumaric and ferulic acids/thioesters (Figure 5), the gap previously observed in the production level of N^1,N^5,N^{10}-triferuloyl spermidine over N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine is largely in favor of ferulic acid incorporation. To mitigate this effect, a couple of enzymes efficiently using *p*-coumaric acid and displaying a weak affinity for ferulic acid was tested. Thereby, the yeast strain expressing At4CL2 and NtSHT grown in the presence of the same amount of *p*-coumaric and ferulic acid produced a mix of *N*,*N*′-di-*p*-coumaroyl *N*″-feruloyl spermidine and *N*-*p*-coumaroyl *N*′,*N*″-diferuloyl spermidine in addition to N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine and

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

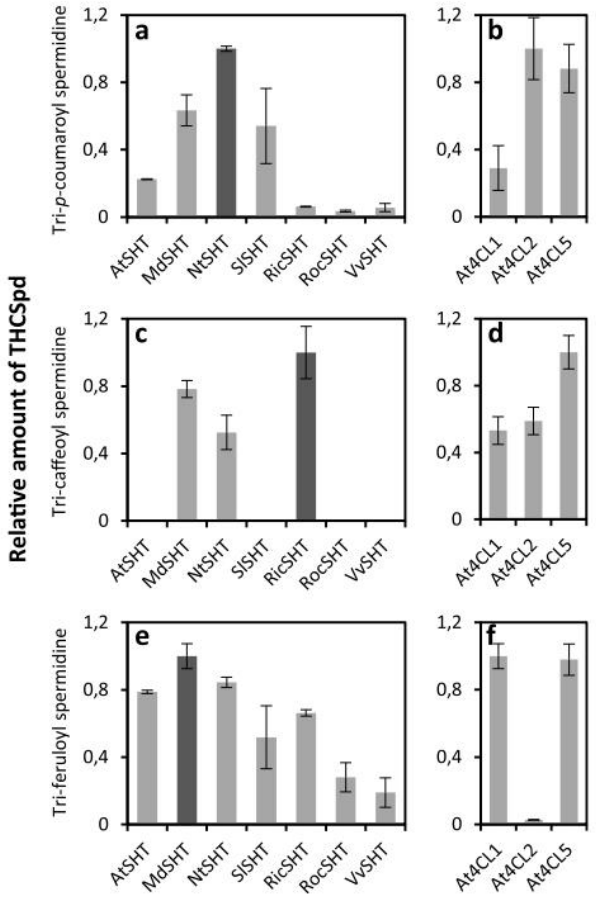


Figure 5. Comparative THCSpd production level using yeast expressing various At4CL isoforms and SHT orthologues. The production levels of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (a,b), N^1,N^5,N^{10}-tricaffeoyl spermidine (c,d), and N^1,N^5,N^{10}-triferuloyl spermidine (e,f) in yeast fed with respective hydroxycinnamic acids were compared among strains expressing At4CL5 and an SHT orthologue from *Arabidopsis thaliana*, *Malus x domestica*, *Nicotiana tabacum*, *Solanum lycopercecum*, *Ricinus communis*, *Rosa canina*, or *Vitis vinifera* (a,c,e) and strains expressing the most efficient SHT (in dark gray: NtSHT for *p*-coumaric acid, RicSHT for caffeic acid and MdSHT for ferulic acid) coupled with At4CL isoforms 1, 2, or 5 (b,d,f). To calculate relative amounts, the conditions with the highest production levels were fixed to 1 for each THCSpd. Error bars correspond to standard deviations of the data from biological triplicates (n = 3). A comparative production rate between the three THCSpd and control experiments are presented as Supporting Information (Figure S3 and Figure S4).

N^1,N^5,N^{10}-triferuloyl spermidine (Figure 6, Table 1). A similar approach was used to produce THCSpd containing caffeoyl and feruloyl moieties; however, the huge preference for ferulic acid over caffeic acid in our enzymatic system prompted us to use a large excess of caffeic acid in the culture medium. By feeding the yeast strain expressing At4CL5 and RiSHT with 10-fold more caffeic acid than ferulic acid, a significant amount of N-caffeoyl N',N''-diferuloyl spermidine and traces of N,N'-dicaffeoyl N''-feruloyl spermidine were obtained (Table 1).

The same assumptions did not result in the synthesis of spermidine acylated with both *p*-coumaroyl and caffeoyl moieties, and only N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine accumulates. Compared to the combination of ferulic acid and caffeic acid, the difference in substrate preferences are too important for *p*-coumaric acid over caffeic acid. To circumvent this limit, we took advantage of the unique and singular THCSpd metabolism of Brassicaceae. SHTs cannot use the bulky 5OH-feruloyl and sinapoyl moieties to produce THCSpd,[1,22] while a specific occurrence of these moieties has been reported in THCSpd of Brassicaceae pollen coat.[12] In fact, two cytochromes P450 (CYP98A8 and CYP98A9) and an *O*-methyltransferase (TSM1) able to use THCSpd as substrate have emerged specifically in this plant family.[22,36] These enzymes catalyze the C3 hydroxylation of *p*-coumaroyl moiety and the C5 hydroxylation and further methylation of feruloyl moiety already associated with spermidine.[22,36] Thus, the expression of CYP98A8 and the cytochrome P450 reductase ATR2 in the best recombinant yeast producer of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine fed with *p*-coumaric acid led to the synthesis of N,N'-di-*p*-coumaroyl N''-caffeoyl spermidine (Table 1). Similarly, yeast harboring CYP98A8 and ATR2, together with At4CL5 and MdSHT, produced N,N'-diferuloyl N''-5-hydroxyferuloyl spermidine from ferulic acid (Table 1). However, in our yeast system, CYP98A8 hydroxylation activity was weak and limited to a single moiety instead of the three. The poor catalytic conversion by CYP98A8 could be the consequence of an intrinsic low enzyme activity in yeast or of an efficient THCSpd export after synthesis, which prevents further enzymatic modifications.

Substrate Screening for the Production of New-to-Nature THCSpd Analogues. Phloretic acid is reported to accumulate in yeast expressing an At4CL isoform and fed with *p*-coumaric acid, because of Tsc13 endogenous activity on corresponding CoA ester.[31] However, as previously mentioned, spermidine acylated with a *p*-dihydrocoumaroyl moiety was detected instead of free phloretic acid in recombinant yeast coexpressing AtSHT and At4CL5 (Figure 1). Considered as a potential limiting step for the optimization of natural THCSpd synthesis, Tsc13 was removed from the genetic background of the yeast. However, this result highlighted the ability of SHTs to use reduced hydroxycinnamoyl-CoA as acyl donors. To exploit this new-to-nature THCSpd production, yeast mutant *pad1* harboring an At4CL and a SHT, transformed with an additional copy of the *TSC13* gene, were fed with hydroxycinnamic acids. Large amounts of *N*-*p*-dihydrocoumaroyl N',N''-di-*p*-coumaroyl spermidine and *N*-dihydroferuloyl N',N''-diferuloyl spermidine were obtained by feeding *p*-coumaric acid to the strain expressing the couple At4CL2/NtSHT, and ferulic acid to the strain expressing At4CL5/MdSHT, respectively (Table 2). In these conditions, only a single spermidine amino group was substituted by dihydrohydroxycinnamoyl moiety, the others were substituted by conventional hydroxycinnamoyl moieties. This could result from either a weak reduction of the hydroxycinnamoyl-CoA pool since it depends on a side activity of Tsc13,[31] or the poor SHT affinity for dihydrohydroxycinnamoyl-CoA compared to hydroxycinnamoyl-CoA.

Hydroxycinnamoyl moiety reduction *in planta* occurs on CoA esters,[37] just after the 4CL enzymatic step. However, the very large substrate promiscuity of 4CLs offers the ability to use dihydrohydroxycinnamic acid as substrate.[38] Therefore, as an alternative, dihydrohydroxycinnamic acids were directly fed as a precursor in the yeast culture medium. The use of dihydroferulic acid with the recombinant yeast coexpressing At4CL5 and MdSHT led to the accumulation of N^1,N^5,N^{10}-tridihydroferuloyl spermidine (Table 2), while neither the use of *p*-dihydrocoumaric acid nor dihydrocaffeic acid led to the production of the corresponding THCSpd. Interestingly,

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

Table 1. Natural THCSpd Produced in Yeast[a]

compound	t_R (min)	$[M + H]^+$	fragmentation	yeast strain	vector	precursor
N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine	13.28	ESI+ 584	438; 420; 318; 292; 275; 204; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT	*p*-coumaric acid
N^1,N^5,N^{10}-triferuloyl spermidine	14.27	ESI+ 674	498; 480; 348; 322; 305; 234; 177	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT	Ferulic acid
N^1,N^5,N^{10}-tricaffeoyl spermidine	10.87	ESI+ 632	470; 452; 333; 308; 291; 220; 163	WT303	pESC-LEU-At4CL5, pESC-URA-RicSHT	Caffeic acid
N,*N*′-di-*p*-coumaroyl *N*″-feruloyl spermidine	13.60	ESI+ 614	468; 450; 421; 348; 322; 305; 292; 275; 234; 204; 177; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT	*p*-coumaric acid, Ferulic acid (1:1)
N-*p*-coumaroyl *N*′,*N*″-diferuloyl spermidine	13.92	ESI+ 644	480; 468; 450; 348; 322; 305; 275; 234; 204; 177; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT	*p*-coumaric acid, Ferulic acid (1:1)
N-caffeoyl *N*′,*N*″-diferuloyl spermidine	13.20	ESI+ 660	498; 484; 465; 322; 305; 291; 234; 220; 177; 163	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-RicSHT	Caffeic acid, Ferulic acid (10:1)
N,*N*′-di-*p*-coumaroyl *N*″-caffeoyl spermidine	12.40	ESI+ 600	454; 436; 292; 275; 220; 204; 163; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT, pESC-TRP-CYP98A8-ATR2	*p*-coumaric acid
N,*N*′-diferuloyl *N*″-5-hydroxyferuloyl spermidine	13.15	ESI+ 690	514; 496; 481; 338; 322; 305; 250; 193; 177	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT, pESC-TRP-CYP98A8-ATR2	Ferulic acid

[a]Detailed fragmentation pattern is available in the Supporting Information (Figure S1). The most efficient condition for the production of each THCSpd is indicated.

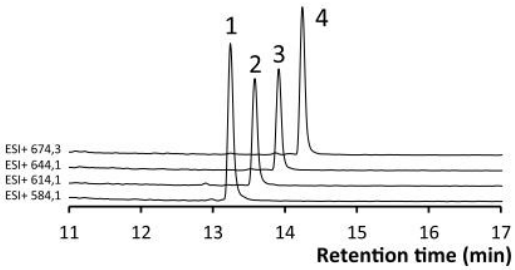


Figure 6. Production of THCSpd heterogeneously substituted with *p*-coumaroyl and feruloyl moieties. The culture medium of yeast mutant *pad1*, *tsc13*::MdECR expressing At4CL2/NtSHT and fed with a ratio of 1:10 of ferulic acid:*p*-coumaric acid was analyzed by LC-MS. Mass chromatograms at ESI+ 584.1, ESI+ 614.1, ESI+ 644.1, and ESI+ 674.1 are overlaid in the same scale range. N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (1); *N*,*N*′-di-*p*-coumaroyl *N*″-feruloyl spermidine (2); *N*-*p*-coumaroyl *N*′,*N*″-diferuloyl spermidine (3); N^1,N^5,N^{10}-triferuloyl spermidine (4). MS and MS/MS profile of compounds is presented in the Supporting Information (Figure S1).

mono- and di-*p*-dihydrocoumaroyl spermidine intermediates were detected in the culture medium. SHT seems thus to be the limiting step for dihydrohydroxycinnamoyl moiety incorporation, reminding the SHT preferences for ferulic moiety on these reduced acid forms. Finally, reduced and conventional hydroxycinnamic acids in a 5:1 ratio were combined in yeast culture medium to obtain *N*,*N*′-di-*p*-dihydrocoumaroyl *N*″-*p*-coumaroyl spermidine and *N*,*N*′-didihydroferuloyl *N*″-feruloyl spermidine (Table 2).

In addition to dihydrohydroxycinnamic acids, the 4CLs were also demonstrated to catalyze the synthesis of CoA esters form various analogues and unnatural hydroxycinnamic acids.[18,39,40] Therefore, less common hydroxycinnamic acids were assessed to feed 4CL/SHT recombinant yeasts. Among acyl donors tested, *o*-coumaric and *m*-coumaric acids led to N^1,N^5,N^{10}-tri-*o*-coumaroyl spermidine and N^1,N^5,N^{10}-tri-*m*-coumaroyl spermidine synthesis respectively in yeast harboring At4CL2 and NtSHT, and 4-methylcinnamic acid led to N^1,N^5,N^{10}-tri-4-methylcinnamoyl spermidine synthesis in yeast harboring At4CL5 and MdSHT (Table 2), whereas 4-nitrocinnamic acid was not incorporated in THCSpd. Although benzoic acid has been reported as a substrate of 4CLs and benzoyl-CoA as a substrate of few BAHD,[19] spermidine acylation was not observed using benzoic acid or 4-hydroxybenzoic acid. Finally, halogenated hydroxycinnamic acids (*i.e.*, 4-fluorocinnamic, 4-bromocinnamic and 4-chlorocinnamic acids), which are also CoA esterified by 4CL,[39] were used in feeding assays with the recombinant yeasts harboring At4CL5 and MdSHT. These combinations resulted in the synthesis of three new-to-nature THCSpd as N^1,N^5,N^{10}-tri-4-fluorocinnamoyl spermidine N^1,N^5,N^{10}-tri-4-bromocinnamoyl spermidine and N^1,N^5,N^{10}-tri-4-chlorocinnamoyl spermidine (Table 2).

***De Novo* Synthesis of N^1,N^5,N^{10}-Tri-*p*-coumaroyl Spermidine in Yeast.** The expression in yeast of a tyrosine ammonia-lyase (TAL), which catalyzes the conversion of tyrosine into *p*-coumaric acid, enabled the synthesis of various polyphenols and hydroxycinnamic esters directly from the endogenous tyrosine pool, instead of feeding hydroxycinnamic acids.[17,38] Therefore, to produce THCSpd without the addition of precursors to the culture medium, we expressed the highly efficient FjTAL from *Flavobacterium johanensis*,[41] in yeast previously engineered to produce N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (Figure 7). The corresponding recombinant yeast strain produced N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine without supplying any external precursors in the culture medium. By using a standard of N^1,N^{10}-di-*p*-coumaroyl spermidine, the production of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine was estimated at 10.2 mg/L ± 1.5 of culture in 48 h in the yeast mutant *pad1*, *tsc13*::MdECR expressing FjTAL, At4CL2, and NtSHT, without further culture condition optimization. This thus established that the synthesis of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (the less complex THCSpd) can be easily achieved from sugar. However, the extension to other natural and new-to-nature THCSpd, previously obtained in this work, remains critical since the wide substrate promiscuities of 4CLs and SHT impair any metabolic flux control toward the desired final compound.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

Table 2. New-to-Nature THCSpd Analogues Produced in Yeast[a]

compound	t_R (min)	$[M + H]^+$	fragmentation	yeast strain	vector	precursor
N-*p*-dihydrocoumaroyl *N*′,*N*″-di-*p*-coumaroyl spermidine	12.93	ESI+ 586	440; 422; 320; 294; 277; 206; 147	WT303 *pad1*	pESC-LEU-At4CL2, pESC-URA-NtSHT, pESC-HIS-TSC13	*p*-coumaric acid
N-dihydroferuloyl *N*′,*N*″-diferuloyl spermidine	13.40	ESI+ 676	500; 482; 324; 307; 248; 236; 177	WT303 *pad1*	pESC-LEU-At4CL5, pESC-URA-MdSHT, pESC-HIS-TSC13	Ferulic acid
N^1,N^5,N^{10}-tridihydroferuloyl spermidine	13.57	ESI+ 680	502; 484; 307; 236; 179	WT303 *pad1*	pESC-LEU-At4CL5, pESC-URA-MdSHT	Dihydroferulic acid
N,*N*′-di-*p*-dihydrocoumaroyl *N*″-*p*-coumaroyl spermidine	12.68	ESI+ 588	442; 425; 320; 277; 206; 147	WT303 *pad1*	pESC-LEU-At4CL2, pESC-URA-NtSHT	*p*-coumaric acid, *p*-dihydrocoumaric acid (1:5)
N,*N*′-didihydroferuloyl *N*″-feruloyl spermidine	13.51	ESI+ 678	502; 484; 307; 250; 236; 177	WT303 *pad1*	pESC-LEU-At4CL5, pESC-URA-MdSHT	Ferulic acid, Dihydroferulic acid (1:5)
N^1,N^5,N^{10}-tri-*o*-coumaroyl spermidine	16.34	ESI+ 584	438; 420; 292; 275; 204; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT	*o*-coumaric acid
N^1,N^5,N^{10}-tri-*m*-coumaroyl spermidine	14.74	ESI+ 584	420; 275; 204; 147	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL2, pESC-URA-NtSHT	*m*-coumaric acid
N^1,N^5,N^{10}-tri-4-methylcinnamoyl spermidine	26.65	ESI+ 578	416; 273; 202; 145	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT	4-methylcinnamic acid
N^1,N^5,N^{10}-tri-4-fluorocinnamoyl spermidine	22.62	ESI+ 590	442; 424; 277; 206; 149	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT	4-fluorocinnamic acid
N^1,N^5,N^{10}-tri-4-bromocinnamoyl spermidine	27.19	ESI+ 771−775	545−549; 338−339; 265−269; 209−211	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT	4-bromocinnamic acid
N^1,N^5,N^{10}-tri-4-chlorocinnamoyl spermidine	26.25	ESI+ 637−643	456−459; 292−295; 222−224; 165−167	WT303 *pad1*, *tsc13*::MdECR	pESC-LEU-At4CL5, pESC-URA-MdSHT	4-chlorocinnamic acid

[a]Detailed fragmentation pattern is available in the Supporting Information (Figure S1).

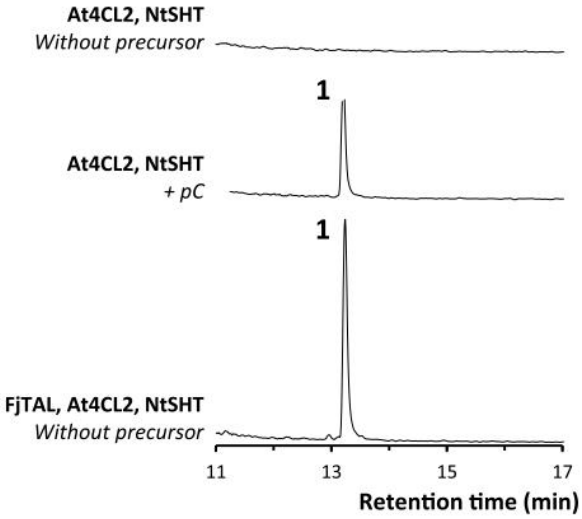


Figure 7. *De novo* synthesis of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine (1) in engineered yeast. LC-MS chromatograms (ESI+ 584.1) of the culture medium of yeast mutant strain *pad1*, *tsc13*::MdECR expressing At4CL2 and NtSHT without exogenous precursor or with *p*-coumaric acid, compared to the strain expressing the additional FjTAL without any precursor.

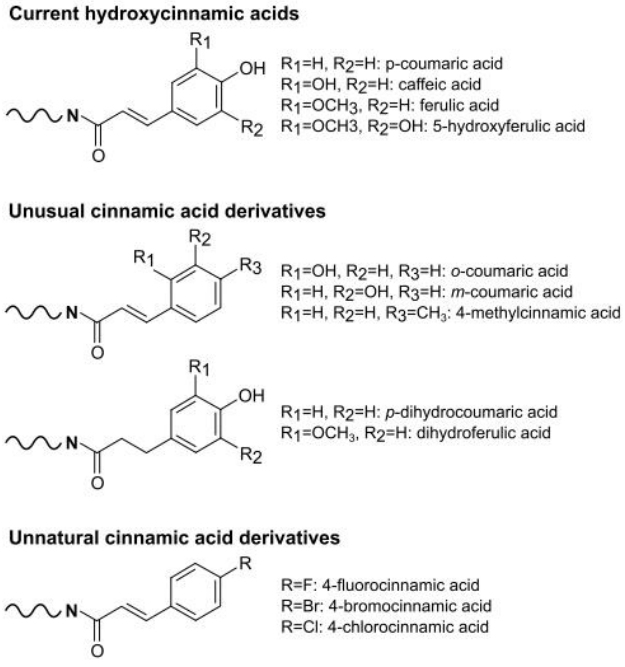


Figure 8. Diversity of acyl groups transferred on spermidine by SHT to produce THCSpd and analogues.

■ CONCLUSIONS

This work establishes a modular and optimized precursor-directed biosynthesis system in yeast to produce THCSpd. Up to 12 different acyl substitutions were obtained on spermidine (Figure 8), leading to 19 THCSpd, including both identical and differential substitutions among the three spermidine amino groups. Such a diversity is generated by combining specific 4CL isoforms with SHT orthologues, together with feeding single or double precursor(s) in various ratios, and the use of THCSpd modifier enzymes such as CYP98A8. This biotechnological production platform is a pioneer work for tailor-made THCSpd synthesis for academic research or pharmaceutical applications, and could solve the restrictive amount and diversity of THCSpd available in plants. Finally, beyond the synthesis of various natural THCSpd, the production of new-to-nature analogues with promising biological applications is established, including the potentially bioactive halogenated THCSpd.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

METHODS

Chemicals. Cinnamic acid, *p*-coumaric acid, caffeic acid, ferulic acid, and spermidine were purchased from Sigma-Aldrich. Dihydroferulic acid, *p*-dihydrocoumaric acid, dihydrocaffeic acid, *o*-coumaric acid, *m*-coumaric acid, and 4-methylcinnamic acid were purchased from Extrasynthese. 4-Fluorocinnamic acid, 4-bromocinnamic acid, and 4-chlorocinnamic acid were purchased from Alfa Aesar.

Construction of the Yeast Expression Plasmids. The expression of recombinant proteins in yeast was performed with the pESC series of episomic vectors. These vectors include a bidirectional galactose inducible promoter *GAL1/GAL10* to control expression of two distinct ORFs cloned in SpeI and *Bam*HI restriction sites. Four different plasmids based on metabolic markers *HIS3*, *URA3*, *LEU2*, or *TRP1* were used and could be cotransformed. To clone gene of interest, ORF sequences were amplified form cDNA or gDNA by PCR using a high fidelity polymerase and specific primers including restriction site for SpeI, XbaI, or *Bam*HI depending on the sequence cloned (Supporting Information; Table S1). PCR products were digested with appropriate restriction enzyme, purified with a commercial kit (PCR-product purification kit, Macherey Nagel) and ligated into the SpeI or *Bam*HI site of a linearized pESC vector depending on sequence cloned. Bacterial clones were selected on LB-ampicillin plates (100 μg/mL) and screened for insertion and correct orientation using colony PCR. The final clones were sequenced with GAL1 Fwd and Rev primers for cloning in *Bam*HI or GAL10 Fwd and Rev primers for cloning in SpeI to control insertion sites and the absence of mutations.

At4CL1, At4CL2, and At4CL5 were amplified from *Arabidopsis thaliana* seedlings cDNA, AtSHT, and CYP98A8 from *Arabidopsis thaliana* flowers cDNA, ATR2 from *Arabidopsis thaliana* stems cDNA, MdSHT from *Malus x domestica L.* flower buds cDNA, NtSHT, SlSHT, RcoSHT, VvSHT, and RcaSHT from *Nicotiana tabacum*, *Solanum lycopersicum*, *Ricinus communis*, *Vitis vinifera*, and *Rosa canina L.* gDNA, respectively, *TSC13* from *Saccharomyces cerevisiae* gDNA and FjTAL from *Flavobacterium johnsoniae* gDNA. To prepare cDNA, total RNA were purified from plant samples as previously described[42] and 1 μg were reverse transcripted using oligo (dT)18 and RevertAid Reverse Transcriptase (Fermentas) according to the manufacturer's instructions. To prepare plant gDNA, plant leaves were frozen in liquid nitrogen, ground using mortar and pestle, and solubilized in a CTAB buffer (100 mM Tris-HCl pH8, 20 mM EDTA, 10 mM β-mercaptoethanol, 2% CTAB, 1% PVP40 and 1 M NaCl). Nucleic acids were recovered with two successive phenol/chloroform/isoamylic acid (1/1/1, v/v/v) extractions, precipitated with sodium acetate/ethanol, and finally solubilized in water. The preparation of yeast gDNA was performed as previously described.[32] Finally, gDNA of *F. johnsoniae* were kindly provided by Dr. K.D. Kim (Korea University, Seoul, South Korea). The final constructs are presented in Supporting Information (Table S2). *S. cerevisiae* were transformed by electroporation and selected on CSM medium plates (0.8% yeast nitrogen base w/o amino acids, 2% glucose and amino acids dropout mix) lacking combination of URA, LEU, HIS, or TRP depending on plasmids used.[43]

Construction of Yeast Strains. The *Saccharomyces cerevisiae* strain WT303 (*MAT* a; *ade2*-1; *his3*-11,15; *leu2*-3,112; *ura3*-1; *trp1*-1) was used as a starting material. The *PAD1* coding sequence was removed by inserting a KanMX4 cassette through conventional homologous recombination. To this aim, a PCR product was amplified from gDNA of the yeast mutant ydr538W obtained from the Saccharomyces Genome Deletion Project, with specific primers of *PAD1* promoter and terminator (Supporting Information; Table S1). PCR products were transformed into WT303 strain, selected on YPD medium plates supplemented with 200 mg/L Geneticin (G418). The recombinations were further confirmed on several clones using PCR on gDNA to obtain the WT303 *pad1* strain (*MAT* a; *ade2*-1; *his3*-11,15; *leu2*-3,112; *ura3*-1; *trp1*-1; *pad1*Δ0::KanMX4). The *TSC13* coding sequence was further replaced by insertion of the MdECR (very-long-chain enoyl-CoA reductase from apple tree) coding sequence using CRISPR-Cas9 technology and the pWS vectors set (Addgene). To prepare the template DNA, three DNA fragments corresponding to the MdECR ORF and 500 bp of *TSC13* gene in upstream of the ATG codon and downstream of the stop codon, were preamplified using a high fidelity polymerase, specific primers (Supporting Information; Table S1) and apple tree leaf cDNA or yeast gDNA. An overlap PCR was performed to combine the three DNA fragments and the PCR product was further cloned in the pCR2.1 vector using A-tailing strategy to generate pCR2.1-TSC13/MdECR. Besides, a gRNA specific to the *TSC13* ORF sequence was designed on E-CRISP[44] and inserted in the gRNA scaffold of pWS082 vector, instead of the GFP sequence. Briefly, the vector was PCR amplified using a high fidelity polymerase and specific primers containing gRNA sequence (Supporting Information; Table S1), after PCR product phosphorylation, self-ligation and transformation, bacteria were selected on LB-ampicilin plates. A clone without green fluorescence under UV light was selected as pWS082-TSC13. Yeast WT303 *pad1* mutant was transformed with 1 μg of DNA template produced by PCR on the pCR2.1-TSC13/MdECR vector, 200 ng of the pWS082-TSC13 vector digested by EcoRV and 100 ng of the CAS9 containing vector pWS158, digested by BsmBI and purified with the phenol/chloroform method. Yeasts were selected on CSM-URA medium plates and genomic modifications were confirmed on several clones using PCR on gDNA. A positive yeast strain was curated of the recombinant vector pWS158 by overnight growth on YPD to obtain the final WT303 *pad1 tsc13*::MdECR strain (*MAT* a; *ade2*-1; *his3*-11,15; *leu2*-3,112; *ura3*-1; *trp1*-1; *pad1*Δ0::KanMX4; *tsc13*Δ::MdECR).

Precursor Feeding Experiments. Seed cultures were prepared by inoculating single colony from CSM agar selective plate into a Falcon 50 mL containing 10 mL YPD (2% bacto peptone, 1% yeast extract, and 2% glucose). Yeasts were grown for 16 h at 30 °C to approximately 6−7 of OD600. After a short centrifuge step at 3000*g*, cells were washed with YPGal medium (2% bacto peptone, 1% yeast extract, and 2% galactose) to remove glucose traces and finally suspended in 10 mL of YPGal for routine experiments, otherwise culture OD600 was adjusted to 1. After 5 h of incubation to induce *GAL1/GAL10* driven expression, seed cultures were divided into several 2 mL subcultures in Falcon tubes or 150 μL subcultures in microplates, depending on the feeding experiments. Substrates were added at this point if needed: 500 μM of spermidine and 200 μM of hydroxycinnamic acids or analogues. Yeasts were incubated 24 or 48 h at 30 °C under shaking. Yeast growth were performed in the dark to avoid photoisomerization of the THCSpd produced.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

LC-MS/MS Analysis of Phenolamides. Yeast cultures were centrifuged 5 min at 13 000*g*, a fraction of the supernatant was mixed with 9 volumes of methanol and centrifuged once again. Culture media were analyzed on a Waters ACQUITY UPLC system coupled to a DAD detector and a XEVO-TQD mass spectrometer equipped with an electrospray ionization (ESI) source controlled by Masslynx 4.1 software (Waters). Samples were separated on a Waters Acquity HSS T3 C18 column (150 × 2.1 mm, i.d. 1.8 μm) with a flow rate of 0.4 mL min^{-1} at 55 °C, using an acetonitrile-formic acid water linear gradient from 5:95 to 50:50 over 18 min. The injection volume was 5 μL. The capillary and sample cone voltages were 3000 V and 30−50 V, respectively. The cone and desolvation gas flow rates were 60 and 800 Lh^{1-}. Both UV and mass chromatograms were used to detect compounds. MS detection was carried out in positive mode both in full scan and in the selected ion monitoring mode targeting $[M + H]^+$ of products attempted. Compounds were identified according to their retention time, m/z ratio and fragmentation profile by comparison with the literature.[1,12] MS/MS data were acquired in daughter scan mode in positive ionization mode electrospray. The fragmentation was performed on molecular ion $[M + H]^+$ using cone voltage of 30 V and two different collision energies of 50 eV and a ramp from 5 to 60 eV. Argon stream was used as collision gas. Quantification of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine was performed on UV chromatograms in the Maxplot detection mode from a calibration curve of N^1,N^{10}-di-*p*-coumaroyl spermidine.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00391.

Table S1: Primer used in this study; Table S2: Plasmid constructs produced and used in this study; Figure S1: LC-MS/MS characterization of THCSpd produced in this study; Figure S2: MS spectrum (ESI+) of the styrene-derivative (1) detected in the Figure 2 and identified as 4-hydroxystyrene; Figure S3: Comparison of the stability of plasmids used in this study, during yeast growth in rich medium; Figure S4: Comparative production rate of N^1,N^5,N^{10}-tri-*p*-coumaroyl spermidine, N^1,N^5,N^{10}-tricaffeoyl spermidine, and N^1,N^5,N^{10}-triferuloyl spermidine in yeast expressing 4CL/SHT couples with the highest conversion efficiency as shown in Figure 5 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Sébastien Besseau − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France;* orcid.org/0000-0002-5518-2580; Email: sebastien.besseau@univ-tours.fr

Authors

Jennifer Perrin − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Natalja Kulagina − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Marianne Unlubayir − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Thibaut Munsch − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Inês Carqueijeiro − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Thomas Dugé de Bernonville − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Johan-Owen De Craene − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Marc Clastre − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Benoit St-Pierre − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Nathalie Giglioli-Guivarc'h − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

David Gagneul − *UMR Transfrontalière BioEcoAgro No. 1158, Univ. Lille, INRAE, Univ. Liège, UPJV, ISA, Univ. Artois, Univ. Littoral Côte d'Opale, ICV − Institut Charles Viollette, F-59000 Lille, France*

Arnaud Lanoue − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France;* orcid.org/0000-0002-7377-8177

Vincent Courdavault − *EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, F-37200 Tours, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00391

Author Contributions

[§]JP and NK contributed equally. JP, NK, MU, TM, IC, JDC, TDB, and SB conducted experiments and analyzed data. JP, NK, and SB performed cloning. JP, NK, IC, and SB performed yeast transformation and precursor-directed biosynthesis experiments. TM and MU conducted LC-MS analyses. JODC and SB performed genetic manipulation of yeast. TDB exploited transcriptomic/genomic data. MC, BStP, NGG, DG, VC, and SB conceived and designed experiments. SB supervised the research and wrote the manuscript. All the authors edited the document and approved the final manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge Dr. K.D. Kim (Korea University, Seoul, South Korea) for providing gDNA of *Flavobacterium johnsoniae* and Dr. Sylvie Friant (Université de Strasbourg − UMR7156, Strasbourg, France) for providing the yeast *pad1* mutant strain. We thank the botanical garden of Tours (France) for plant biological materials. We also acknowledge financial support from ARD2020 Biopharmaceutical program of the Région Centre Val de Loire (BioPROPHARM, CatharSIS and ETOPOCentre projects).

■ REFERENCES

(1) Elejalde-Palmett, C., De Bernonville, T. D., Glevarec, G., Pichon, O., Papon, N., Courdavault, V., St-Pierre, B., Giglioli-Guivarc'H, N.,

Lanoue, A., and Besseau, S. (2015) Characterization of a spermidine hydroxycinnamoyltransferase in Malus domestica highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *J. Exp. Bot. 66*, 7271−7285.
(2) Strack, D., Eilert, U., Wray, V., Wolff, J., and Jaggy, H. (1990) Tricoumaroylspermidine in flowers of rosaceae. *Phytochemistry 29*, 2893−2896.
(3) Zamble, A., Sahpaz, S., Hennebelle, T., Carato, P., and Bailleul, F. (2006) N1,N5,N10-Tris (4-hydroxycinnamoyl) spermidines from Microdesmis keayana Roots. *Chem. Biodiversity 3*, 982−989.
(4) Vogt, T. (2018) Unusual spermine-conjugated hydroxycinnamic acids on pollen: Function and evolutionary advantage. *J. Exp. Bot. 69*, 5311−5318.
(5) Kyselka, J., Bleha, R., Dragoun, M., Bialasová, K., Horáčková, Š., Schätz, M., Sluková, M., Filip, V., and Synytsya, A. (2018) Antifungal Polyamides of Hydroxycinnamic Acids from Sunflower Bee Pollen. *J. Agric. Food Chem. 66*, 11018−11026.
(6) Zhao, G., Qin, G. W., Gai, Y., and Guo, L. H. (2010) Structural identification of a new tri-p-coumaroylspermidine with serotonin transporter inhibition from safflower. *Chem. Pharm. Bull. 58*, 950−952.
(7) Ma, C. M., Nakamura, N., and Hattori, M. (2001) Inhibitory effects on HIV-1 protease of tri-p-coumaroylspermidine from Artemisia caruifolia and related amides. *Chem. Pharm. Bull. 49*, 915−917.
(8) Yingyongnarongkul, B. E., Apiratikul, N., Aroonrerk, N., and Suksamrarn, A. (2008) Synthesis of bis, tris and tetra-(dihydrocaffeoyl)polyamine conjugates as antibacterial agents against VRSA. *Arch. Pharmacal Res. 31*, 698−704.
(9) Park, S. B., Song, K., and Kim, Y. S. (2017) Tetra-cis/trans-Coumaroyl Polyamines as NK 1 Receptor Antagonists from Matricaria chamomilla. *Planta Medica Int. Open 4*, 43−51.
(10) Yamamoto, A., Nakamura, K., Furukawa, K., Konishi, Y., Ogino, T., Higashiura, K., Yago, H., Okamoto, K., and Otsuka, M. (2002) A new nonpeptide tachykinin NK1 receptor antagonist isolated from the plants of Compositae. *Chem. Pharm. Bull. 50*, 47−52.
(11) Li, Z., Zhao, C., Zhao, X., Xia, Y., Sun, X., Xie, W., Ye, Y., Lu, X., and Xu, G. (2018) Deep Annotation of Hydroxycinnamic Acid Amides in Plants Based on Ultra-High-Performance Liquid Chromatography-High-Resolution Mass Spectrometry and Its in Silico Database. *Anal. Chem. 90*, 14321−14330.
(12) Handrick, V., Vogt, T., and Frolov, A. (2010) Profiling of hydroxycinnamic acid amides in Arabidopsis thaliana pollen by tandem mass spectrometry. *Anal. Bioanal. Chem. 398*, 2789−2801.
(13) Hu, W., and Hesse, M. (1996) Synthese Der P-Cumaroyl-spermidine. *Helv. Chim. Acta 79*, 548−559.
(14) Paddon, C. J., and Keasling, J. D. (2014) Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol. 12*, 355−367.
(15) Courdavault, V., O'Connor, S. E., Oudin, A., Besseau, S., and Papon, N. (2020) Towards the Microbial Production of Plant-Derived Anticancer Drugs. *Trends in Cancer 6*, 444−448.
(16) Rodriguez, A., Strucko, T., Stahlhut, S. G., Kristensen, M., Svenssen, D. K., Forster, J., Nielsen, J., and Borodina, I. (2017) Metabolic engineering of yeast for fermentative production of flavonoids. *Bioresour. Technol. 245*, 1645−1654.
(17) Borja, G. M., Rodriguez, A., Campbell, K., Borodina, I., Chen, Y., and Nielsen, J. (2019) Metabolic engineering and transcriptomic analysis of Saccharomyces cerevisiae producing p-coumaric acid from xylose. *Microb. Cell Fact. 18*, 1−14.
(18) Eudes, A., Baidoo, E. E. K., Yang, F., Burd, H., Hadi, M. Z., Collins, F. W., Keasling, J. D., and Loqué, D. (2011) Production of tranilast [N-(3′,4′-dimethoxycinnamoyl)- anthranilic acid] and its analogs in yeast Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol. 89*, 989−1000.
(19) Eudes, A., Mouille, M., Robinson, D. S., Benites, V. T., Wang, G., Roux, L., Tsai, Y. L., Baidoo, E. E. K., Chiu, T. Y., Heazlewood, J. L., Scheller, H. V., Mukhopadhyay, A., Keasling, J. D., Deutsch, S., and Loqué, D. (2016) Exploiting members of the BAHD acyltransferase family to synthesize multiple hydroxycinnamate and benzoate conjugates in yeast. *Microb. Cell Fact. 15*, 1−16.
(20) Eichenberger, M., Lehka, B. J., Folly, C., Fischer, D., Martens, S., Simón, E., and Naesby, M. (2017) Metabolic engineering of Saccharomyces cerevisiae for de novo production of dihydrochalcones with known antioxidant, antidiabetic, and sweet tasting properties. *Metab. Eng. 39*, 80−89.
(21) Li, M., Schneider, K., Kristensen, M., Borodina, I., and Nielsen, J. (2016) Engineering yeast for high-level production of stilbenoid antioxidants. *Sci. Rep. 6*, 1−8.
(22) Grienenberger, E., Besseau, S., Geoffroy, P., Debayle, D., Heintz, D., Lapierre, C., Pollet, B., Heitz, T., and Legrand, M. (2009) A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J. 58*, 246−259.
(23) D'Auria, J. C. (2006) Acyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol. 9*, 331−340.
(24) Tuominen, L. K., Johnson, V. E., and Tsai, C. J. (2011) Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among Populus paralogues. *BMC Genomics 12*, 12.
(25) Delporte, M., Bernard, G., Legrand, G., Hielscher, B., Lanoue, A., Molinié, R., Rambaud, C., Mathiron, D., Besseau, S., Linka, N., Hilbert, J., and Gagneul, D. (2018) A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in the pollen coat of the Asteraceae family. *J. Exp. Bot. 69*, 5355−5371.
(26) Peng, H., Yang, T., Whitaker, B. D., Trouth, F., Shangguan, L., Dong, W., and Jurick, W. M. (2016) Characterization of spermidine hydroxycinnamoyl transferases from eggplant (Solanum melongena L.) and its wild relative Solanum richardii Dunal. *Hortic. Res. 3*, 1−9.
(27) Ehlting, J., Büttner, D., Wang, Q., Douglas, C. J., Somssich, I. E., and Kombrink, E. (1999) Three 4-coumarate:coenzyme A ligases in arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. *Plant J. 19*, 9−20.
(28) Hamberger, B., and Hahlbrock, K. (2004) The 4-coumarate:CoA ligase gene family in Arabidopsis thaliana comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc. Natl. Acad. Sci. U. S. A. 101*, 2209−2214.
(29) Moglia, A., Comino, C., Lanteri, S., de Vos, R., de Waard, P., Teris, A., van, B., Luca, G., Retta, S. F., and Beekwilder, J. (2010) Production of novel antioxidative phenolic amides through heterologous expression of the plant's chlorogenic acid biosynthesis genes in yeast. *Metab. Eng. 12*, 223−232.
(30) Mukai, N., Masaki, K., Fujii, T., Kawamukai, M., and Iefuji, H. (2010) PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in Saccharomyces cerevisiae. *J. Biosci. Bioeng. 109*, 564−569.
(31) Lehka, B. J., Eichenberger, M., Bjørn-Yoshimoto, W. E., Vanegas, K. G., Buijs, N., Jensen, N. B., Dyekjçr, J. D., Jenssen, H., Simon, E., and Naesby, M. (2017) Improving heterologous production of phenylpropanoids in Saccharomyces cerevisiae by tackling an unwanted side reaction of Tsc13, an endogenous double-bond reductase. *FEMS Yeast Res.*, 17.
(32) Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature 418*, 387−391.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

(33) Valdés-Santiago, L., and Ruiz-Herrera, J. (2014) Stress and polyamine metabolism in fungi. *Front. Chem. 1*, 1−10.
(34) Kim, S. K., Jo, J. H., Park, Y. C., Jin, Y. S., and Seo, J. H. (2017) Metabolic engineering of Saccharomyces cerevisiae for production of spermidine under optimal culture conditions. *Enzyme Microb. Technol. 101*, 30−35.
(35) Elejalde-Palmett, C., Dugé de Bernonville, T., Glevarec, G., Pichon, O., Papon, N., Courdavault, V., St-Pierre, B., Giglioli-Guivarc'H, N., Lanoue, A., and Besseau, S. (2015) Characterization of a spermidine hydroxycinnamoyltransferase in Malus domestica highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *J. Exp. Bot. 66*, 7271−7285.
(36) Matsuno, M., Compagnon, V., Schoch, G. A., Schmitt, M., Debayle, D., Bassard, J.-E., Pollet, B., Hehn, A., Heintz, D., Ullmann, P., Lapierre, C., Bernier, F., Ehlting, J., and Werck-Reichhart, D. (2009) Evolution of a novel phenolic pathway for pollen development. *Science 325*, 1688−1692.
(37) Ibdah, M., Berim, A., Martens, S., Valderrama, A. L. H., Palmieri, L., Lewinsohn, E., and Gang, D. R. (2014) Identification and cloning of an NADPH-dependent hydroxycinnamoyl-CoA double bond reductase involved in dihydrochalcone formation in Malus × domestica Borkh. *Phytochemistry 107*, 24−31.
(38) Eudes, A., Juminaga, D., Baidoo, E. E. K., Collins, F. W., Keasling, J. D., and Loqué, D. (2013) Production of hydroxycinnamoyl anthranilates from glucose in Escherichia coli. *Microb. Cell Fact. 12*, 1−10.
(39) Werner, S. R., Chen, H., Jiang, H., and Morgan, J. A. (2010) Synthesis of non-natural flavanones and dihydrochalcones in metabolically engineered yeast. *J. Mol. Catal. B: Enzym. 66*, 257−263.
(40) Kufs, J. E., Hoefgen, S., Rautschek, J., Bissell, A. U., Graf, C., Fiedler, J., Braga, D., Regestein, L., Rosenbaum, M. A., Thiele, J., and Valiante, V. (2020) Rational Design of Flavonoid Production Routes Using Combinatorial and Precursor-Directed Biosynthesis. *ACS Synth. Biol. 9*, 1823.
(41) Jendresen, C. B., Stahlhut, S. G., Li, M., Gaspar, P., Siedler, S., Förster, J., Maury, J., Borodina, I., and Nielsen, A. T. (2015) Highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and Saccharomyces cerevisiae. *Appl. Environ. Microbiol. 81*, 4458−4476.
(42) Daudu, D., Allion, E., Liesecke, F., Papon, N., Courdavault, V., Dugé de Bernonville, T., Mélin, C., Oudin, A., Clastre, M., Lanoue, A., Courtois, M., Pichon, O., Giron, D., Carpin, S., Giglioli-Guivarc'h, N., Crèche, J., Besseau, S., and Glévarec, G. (2017) CHASE-Containing Histidine Kinase Receptors in Apple Tree: From a Common Receptor Structure to Divergent Cytokinin Binding Properties and Specific Functions. *Front. Plant Sci. 8*, 1−15.
(43) Navarro Gallón, S. M., Elejalde-Palmett, C., Daudu, D., Liesecke, F., Jullien, F., Papon, N., Dugé de Bernonville, T., Courdavault, V., Lanoue, A., Oudin, A., Glévarec, G., Pichon, O., Clastre, M., St-Pierre, B., Atehortùa, L., Yoshikawa, N., Giglioli-Guivarc'h, N., and Besseau, S. (2017) Virus-induced gene silencing of the two squalene synthase isoforms of apple tree (Malus × domestica L.) negatively impacts phytosterol biosynthesis, plastid pigmentation and leaf growth. *Planta 246*, 45−60.
(44) Heigwer, F., Kerr, G., and Boutros, M. (2014) E-CRISP: Fast CRISPR target site identification. *Nat. Methods 11*, 122−123.
(45) Sobolev, V. S., Sy, A. A., and Gloer, J. B. (2008) Spermidine and flavonoid conjugates from peanut (Arachis hypogaea) flowers. *J. Agric. Food Chem. 56*, 2960−2969.

https://dx.doi.org/10.1021/acssynbio.0c00391
ACS Synth. Biol. 2021, 10, 286−296

Supporting Information

Exploiting Spermidine N-HydroxycinnamoylTransferase diversity and substrate promiscuity to produce various tri-hydroxycinnamoyl spermidines and analogs in engineered yeast.

Jennifer Perrin*, Natalja Kulagina*, Marianne Unlubayir, Thibaut Munsch, Inês Carqueijeiro, Thomas Dugé de Bernonville, Johan-Owen De Craene, Marc Clastre, Benoit St-Pierre, Nathalie Giglioli-Guivarc'h, David Gagneul, Arnaud Lanoue, Vincent Courdavault, and Sébastien Besseau

Table S1. Primer used in this study.

Primer name	Sequence (5'-3')	Use
GAL1 Fwd	ATTTTCGGTTTGTATTACTTC	Sequencing / colony PCR
GAL1 Rev	GTTCTTAATACTAACATAACT	Sequencing / colony PCR
GAL10 Fwd	GGTGGTAATGCCATGTAATATG	Sequencing / colony PCR
GAL10 Rev	GGCAAGGTAGACAAGCCGACAAC	Sequencing / colony PCR
At4CL1 BamHI Fwd	AGTGGATCCATGGCGCCACAAGAACAAGCAG	ORF cloning
At4CL1 BamHI Rev	AGTGGATCCTCACAATCCATTTGCTAGTTTTGCCC	ORF cloning
At4CL2 BamHI Fwd	AGTGGATCCATGACGACACAAGATGTGATAGT	ORF cloning
At4CL2 BamHI Rev	AGTGGATCCTAGTTCATTAATCCATTTGCTAGTC	ORF cloning
At4CL5 SpeI Fwd	GACACTAGTATGGTGCTCCAACAACAAACGC	ORF cloning
At4CL5 SpeI Rev	GACACTAGTCTATTTAGAGCACATGGTTTCCAA	ORF cloning
AtSHT BamHI Fwd	AGTGGATCCATGGCTCCCATAACTTTTAGAAAATC	ORF cloning
AtSHT BamHI Rev	AGTGGATCCTAGATATCTTCATAAAAGTGTTTCTTG	ORF cloning
MdSHT BamHI Fwd	AGCGGATCCATGGTGGTGAAATTCAAAGAC	ORF cloning
MdSHT BamHI Rev	AGCGGATCCTCATATAATATCTTCATAGAAG	ORF cloning
NtSHT SpeI Fwd	CTGAGAACTAGTATGAAGGTTAGCTTGAAGCAC	ORF cloning
NtSHT SpeI Rev	CTGAGAACTAGTTCAACATTCAATTTCTTCATAGAAGAAG	ORF cloning
SlSHT SpeI Fwd	CTGAGAACTAGTATGAAAGTTATCTTAAAGAATCAATGGG	ORF cloning
SlSHT SpeI Rev	CTGAGAACTAGTTCAACATTCAATATCTTCATAGAAAAATTTC	ORF cloning
RcoSHT SpeI Fwd	CTGAGAACTAGTATGATTGTCATCATCAAGGGCTC	ORF cloning
RcoSHT SpeI Rev	CTGAGAACTAGTTCAGATATCTTCATAGAAGTACTTCTTAA	ORF cloning
VvSHT SpeI Fwd	CTGAGAACTAGTATGGTTGTAGTGAGCTTGAAAG	ORF cloning
VvSHT SpeI Rev	CTGAGAACTAGTTCAGTCTATTGAAGATTCTGCC	ORF cloning
RcaSHT SpeI Fwd	CTGAGAACTAGTATGGAGGTGAAGTTCAAAGGCTG	ORF cloning
RcaSHT SpeI Rev	CTGAGAACTAGTTTAAATATGAGCCACTTGAAGGCA	ORF cloning
FjTALSpeI Fwd	GACACTAGTATGAATACAATAAATGAATATCTAAG	ORF cloning
FjTALSpeI Rev	GACACTAGTTAACTGTTGATCAAATAATCTT	ORF cloning
CYP98A8 XbaI Fwd	GACTCTAGATGATTATATATCTAATTTCATTAC	ORF cloning
CYP98A8 XbaI Rev	GACTCTAGATTAATCTAAAGGTAAAGGTATTG	ORF cloning
ATR2 BamHI Fwd	GACGGATCCATGAAAAACATGATGAATTATAAATTAAAAC	ORF cloning
ATR2 BamHI Rev	GACGGATCCTTACCATACATCTCTAAGATATCTTCC	ORF cloning
ScTSC13 SpeI Fwd	GACACTAGTATGCCTATCACCATAAAAAGCCG	ORF cloning
ScTSC13 SpeI Rev	GACACTAGTCAAAATACAAATGGAATCAAGAATGC	ORF cloning
PAD1 Fwd	TAGCCCGGCAGAAGCCATAC	Locus amplification
PAD1 Rev	TCTAAAGCTGGATTTAGCTTCCTC	Locus amplification
TSC13 500pb up Fwd	CACGGGAACCACCTCTGCTAC	Template arms cloning
TSC13 500pb up Ref	GCGAAACTACGGTCACCTTCATTTTCAAATTGAATTCAGAATATGTATC	Template arms cloning
TSC13 500pb down Fwd	TGGTCCGATCATTATTGGTGCTTCC	Template arms cloning
TSC13 500pb down Rev	GGTGATCCTACCTCCATTCCTGTGATCCTCCTGTTTGCAACATG	Template arms cloning
MdECR Fwd	ATGAAGGTGACCGTAGTTTCGC	ORF amplification and cloning
MdECR Rev	TCACAGGAATGGAGGTAGGATCACC	ORF amplification and cloning
TSC13 gRNA Fwd	AAAGTCCCATTCGCCACCCGAAG	gRNA cloning
TSC13 gRNA Rev	GCATAGTAGCTAAAGAGAATGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning
TSC13 Fwd	ATGCCTATCACCATAAAAAGCCG	ORF amplification
TSC13 Rev	TCAAAATACAAATGGAATCAAGAATGC	ORF amplification

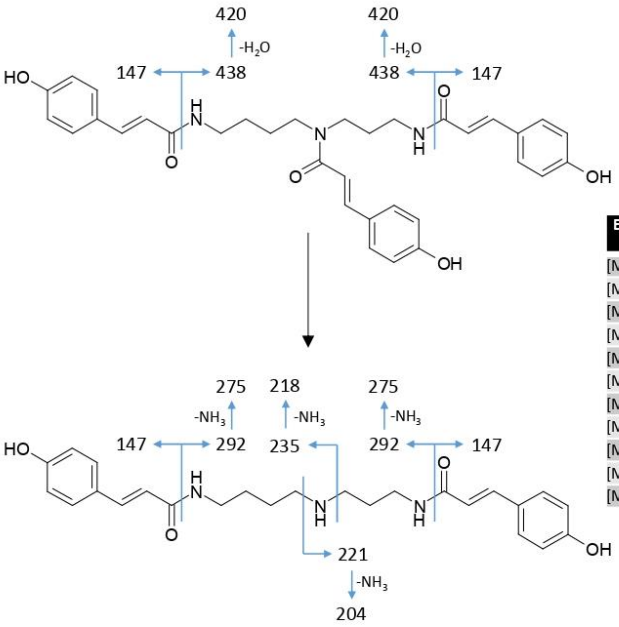
Table S2. Plasmid constructs produced and used in this study.

Plasmid name	pGAL1/tCYC1 cassette (BamHI)	pGAL10/tADH1 cassette (SpeI)
pESC-LEU-At4CL1	At4CL1	
pESC-LEU-At4CL2	At4CL2	
pESC-LEU-At4CL5		At4CL5
pESC-URAAtSHT-	AtSHT	
pESC-URA-MdSHT	MdSHT	
pESC-URA-NtSHT		NtSHT
pESC-URA-SlSHT		SlSHT
pESC-URA-RcoSHT		RcoSHT
pESC-URA-VvSHT		VvSHT
pESC-URA-RcaSHT		RcaSHT
pESC-TRP-CYP98A8-ATR2	ATR2	CYP98A8
pESC-HIS-ScTSC13		ScTSC13
pESC-LEU-FjTAL-At4CL2	At4CL2	FjTAL

Figure S1. LC-MS/MS characterization of THCSpd produced in this study.

For each THCSpd expected, the pseudo-molecular ions *m/z* ratio were estimated based on the calculated molecular weight of the compounds. Pseudo-molecular ion fragmentations of THCSpd were predicted based on the scheme proposed by Handrick *et al.* 2010.

Example of N^1, N^5, N^{10}-tri-*p*-coumaroyl spermidine putative fragmentation scheme and deducted *m/z* values:

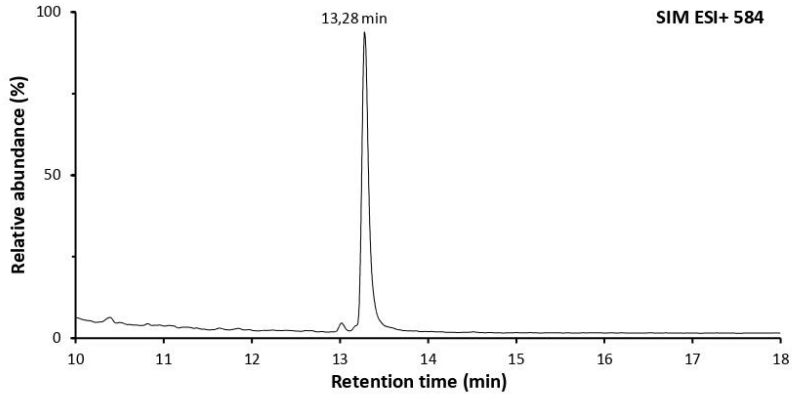


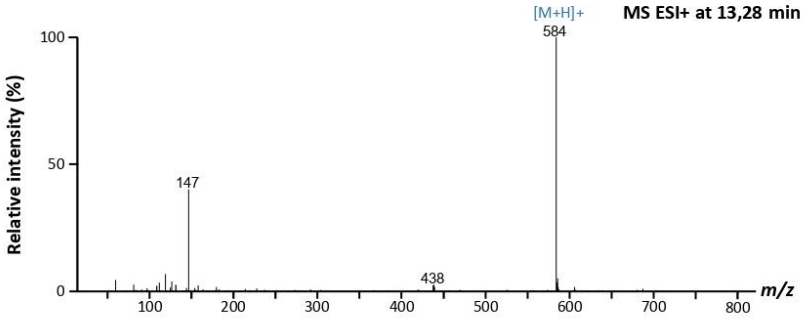
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	584
[M+Na]+	606
[M+H-pC]+	438
[M+H-pC-H_2O]+	420
[M+H-pC-pC]+	292
[M+H-pC-pC-NH_3]+	275
[M+H-pC-pC-C3chain]+	235
[M+H-pC-pC-C3chain-NH_3]+	218
[M+H-pC-pC-C4chain]+	221
[M+H-pC-pC-C4chain-NH_3]+	204
[M+H-pC-pC-Spd]+	147

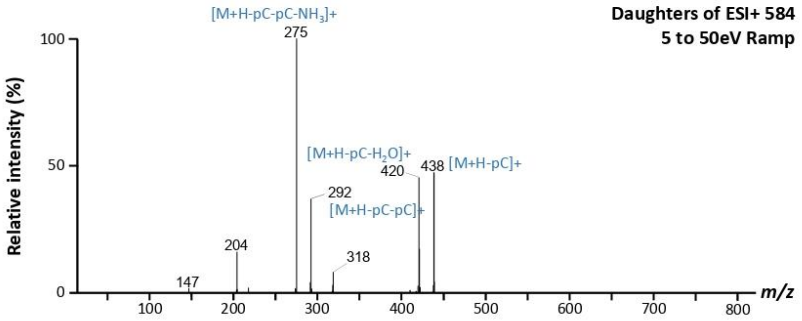
pC, *p*-coumaroyl; oC, *o*-coumaroyl; mC, *m*-coumaroyl; MetCi, 4-methylcinnamoyl; Caf, caffeoyl; Fe, feruloyl; dpC, *p*-dihydrocoumaroyl; dFe, dihydroferuloyl; OHFe, 5-hydroxyferuloyl; ClCi, 4-chlorocinamoyl; BrCi, 4-bromocinnamoyl; Fci, 4-fluorocinamoyl.
For some THCSpd, several peaks appear on mass chromatograms which correspond to photoisomers.

N^1, N^5, N^{10}-tri-*p*-coumaroyl spermidine (MW 583 g/mol)

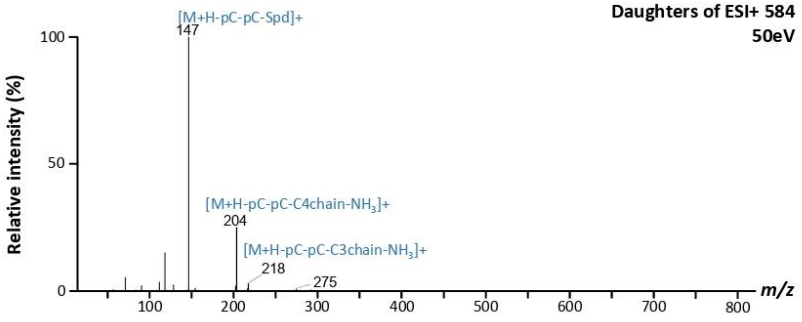
Expected pseudo-molecular ion [M+H]+: *m/z*=584





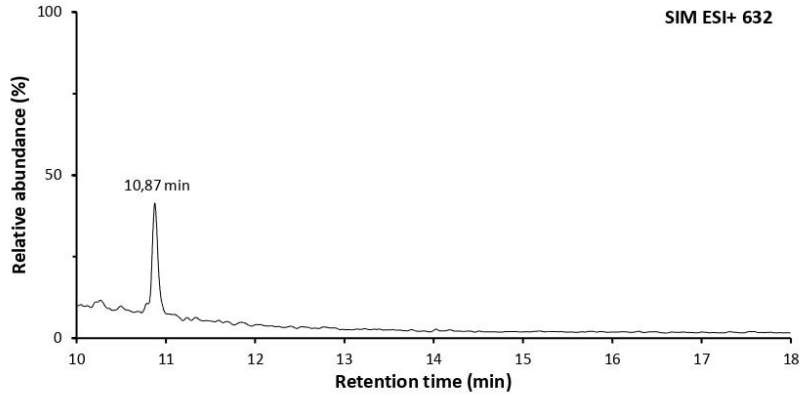


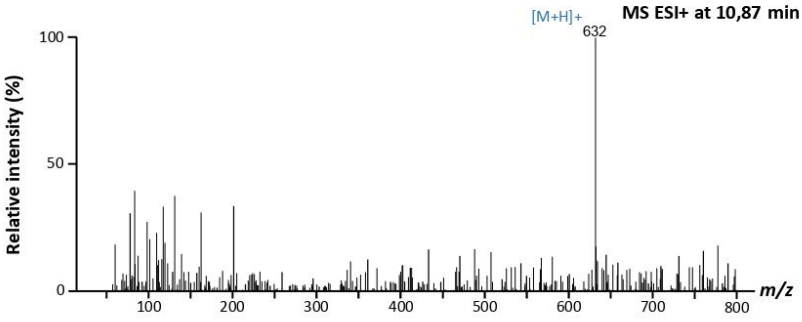
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	584
[M+Na]+	606
[M+H-pC]+	438
[M+H-pC-H_2O]+	420
[M+H-pC-pC]+	292
[M+H-pC-pC-NH_3]+	275
[M+H-pC-pC-C3chain]+	235
[M+H-pC-pC-C3chain-NH_3]+	218
[M+H-pC-pC-C4chain]+	221
[M+H-pC-pC-C4chain-NH_3]+	204
[M+H-pC-pC-Spd]+	147

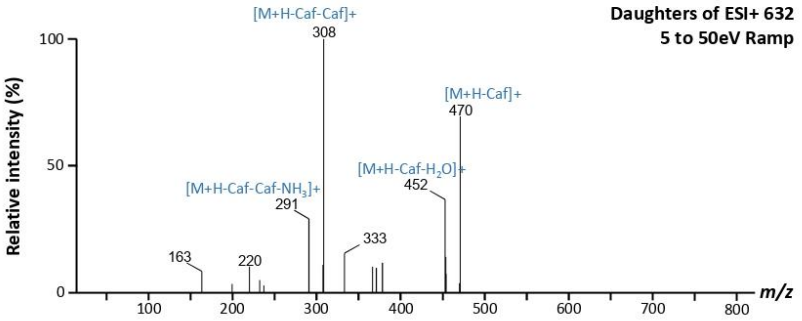


N^1, N^5, N^{10}-tri-caffeoyl spermidine (MW 631 g/mol)

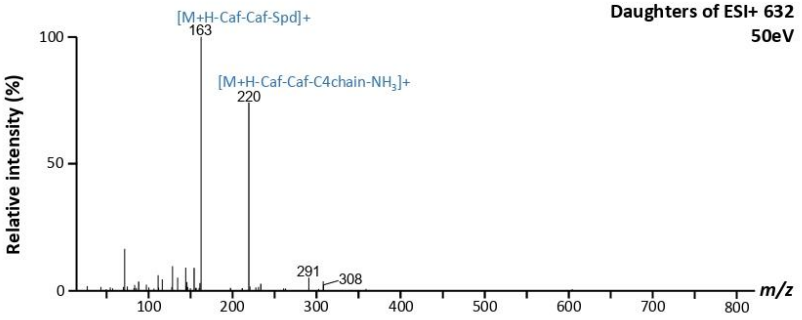
Expected pseudo-molecular ion [M+H]+: *m/z*=632





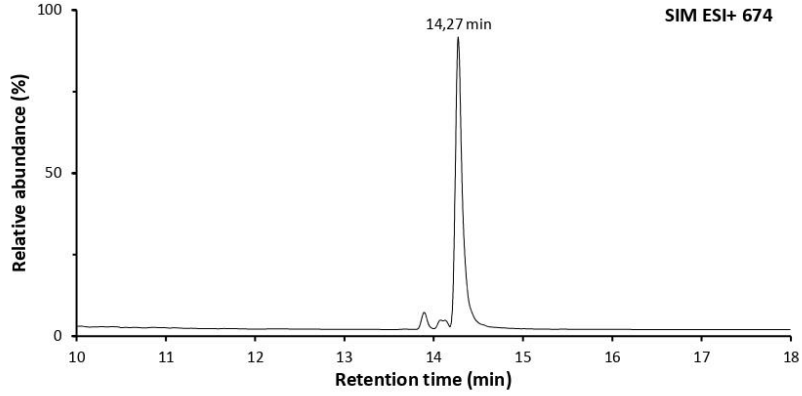


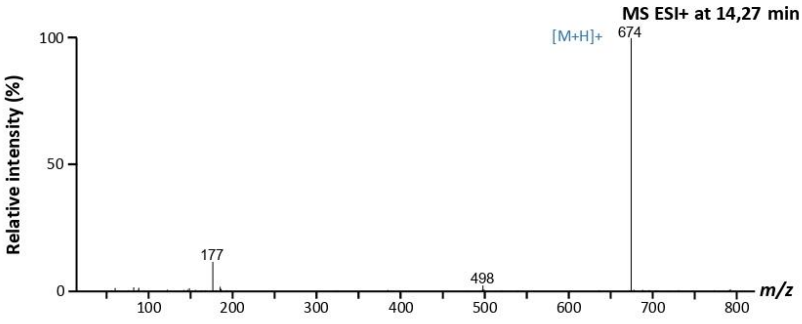
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	632
[M+Na]+	654
[M+H-Caf]+	470
[M+H-Caf-H_2O]+	452
[M+H-Caf-Caf]+	308
[M+H-Caf-Caf-NH_3]+	291
[M+H-Caf-Caf-C3chain]+	251
[M+H-Caf-Caf-C3chain-NH_3]+	234
[M+H-Caf-Caf-C4chain]+	237
[M+H-Caf-Caf-C4chain-NH_3]+	220
[M+H-Caf-Caf-Spd]+	163

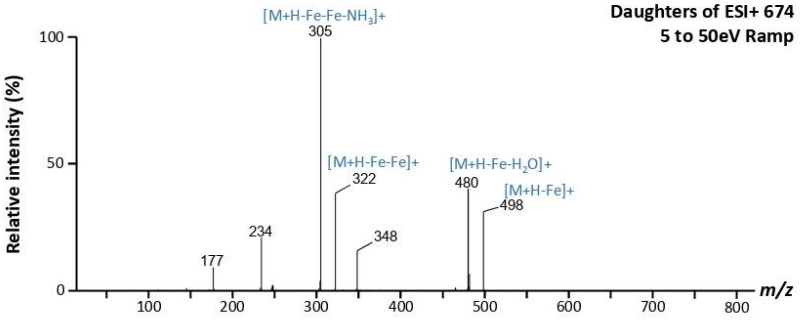


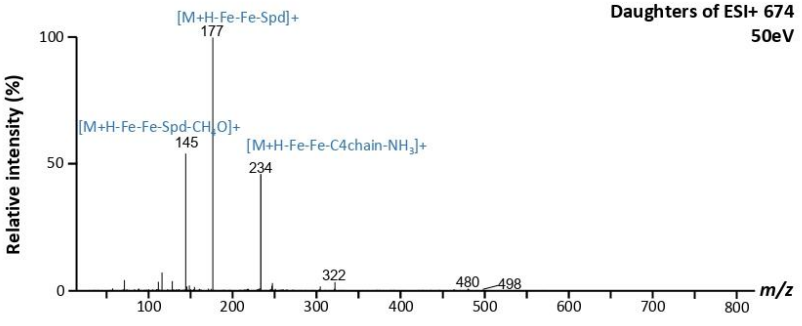
N^1, N^5, N^{10}-tri-feruloyl spermidine (MW 673 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=674





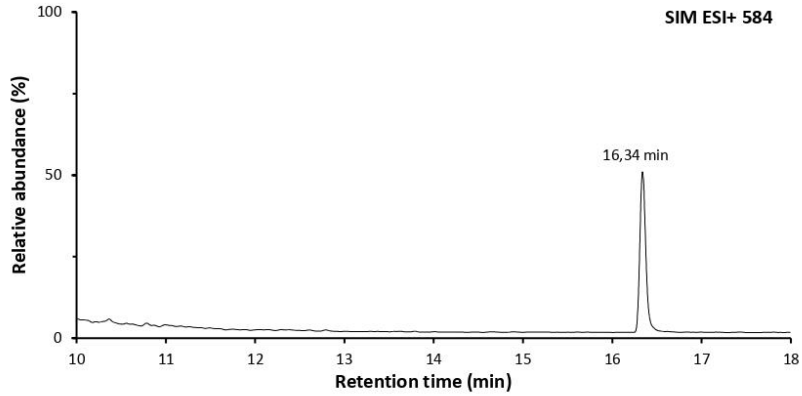


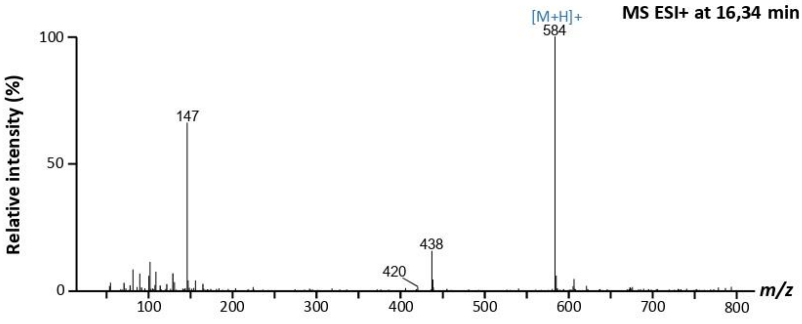


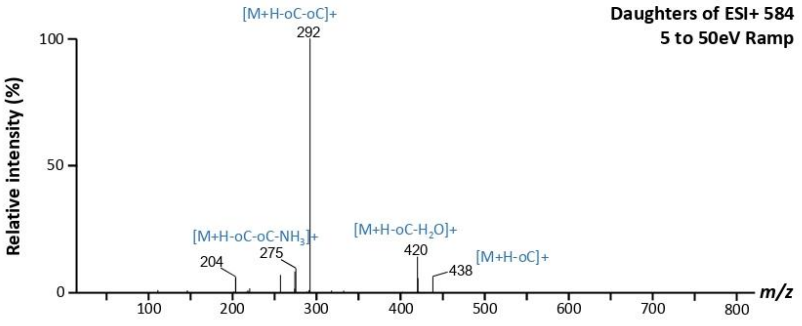
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	674
[M+Na]+	696
[M+H-Fe]+	498
[M+H-Fe-H_2O]+	480
[M+H-Fe-Fe]+	322
[M+H-Fe-Fe-NH_3]+	305
[M+H-Fe-Fe-C3chain]+	265
[M+H-Fe-Fe-C3chain-NH_3]+	248
[M+H-Fe-Fe-C4chain]+	251
[M+H-Fe-Fe-C4chain-NH_3]+	234
[M+H-Fe-Fe-Spd]+	177
[M+H-Fe-Fe-Spd-CH_4O]+	145

N^1, N^5, N^{10}-tri-*o*-coumaroyl spermidine (MW 583 g/mol)

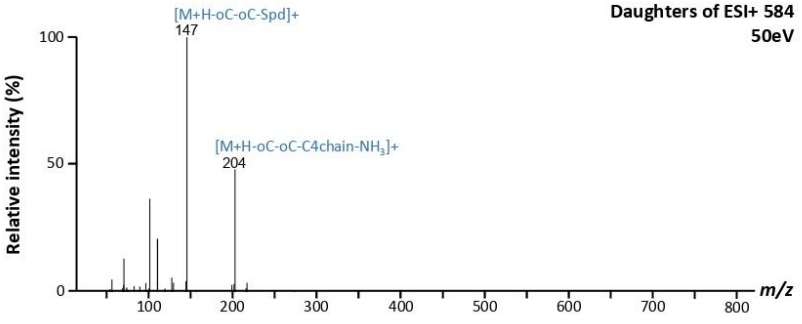
Expected pseudo-molecular ion [M+H]+: *m/z*=584





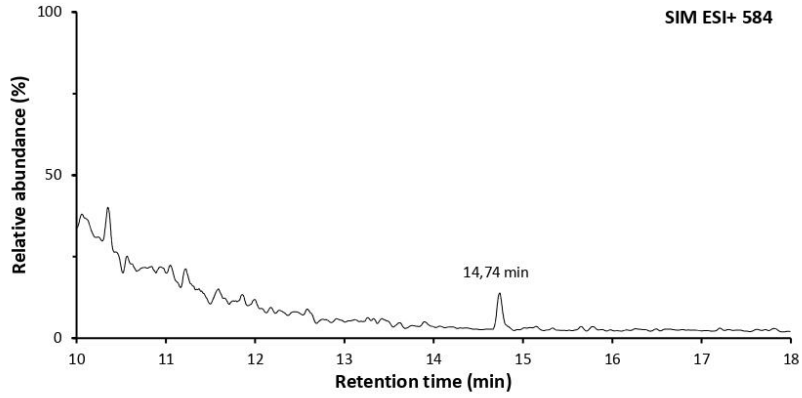


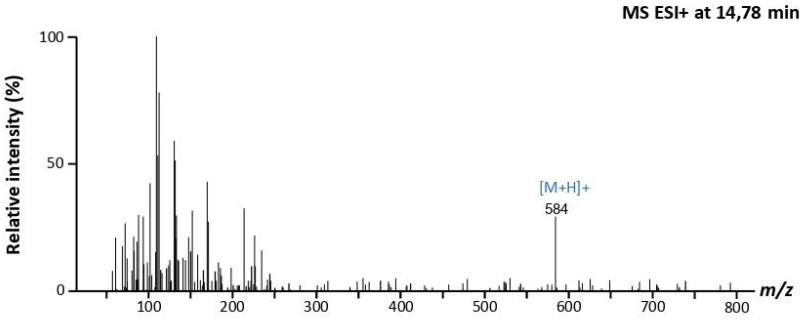
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	584
[M+Na]+	606
[M+H-oC]+	438
[M+H-oC-H_2O]+	420
[M+H-oC-oC]+	292
[M+H-oC-oC-NH_3]+	275
[M+H-oC-oC-C3chain]+	235
[M+H-oC-oC-C3chain-NH_3]+	218
[M+H-oC-oC-C4chain]+	221
[M+H-oC-oC-C4chain-NH_3]+	204
[M+H-oC-oC-Spd]+	147

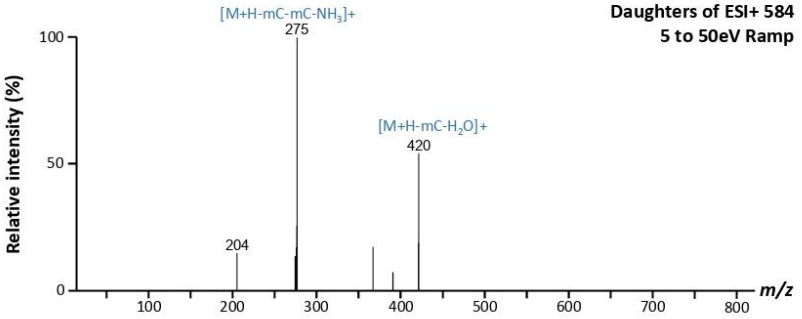


N^1, N^5, N^{10}-tri-*m*-coumaroyl spermidine (MW 583 g/mol)

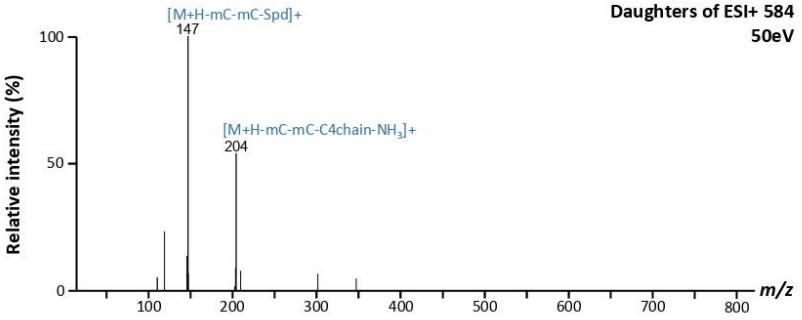
Expected pseudo-molecular ion [M+H]+: *m/z*=584





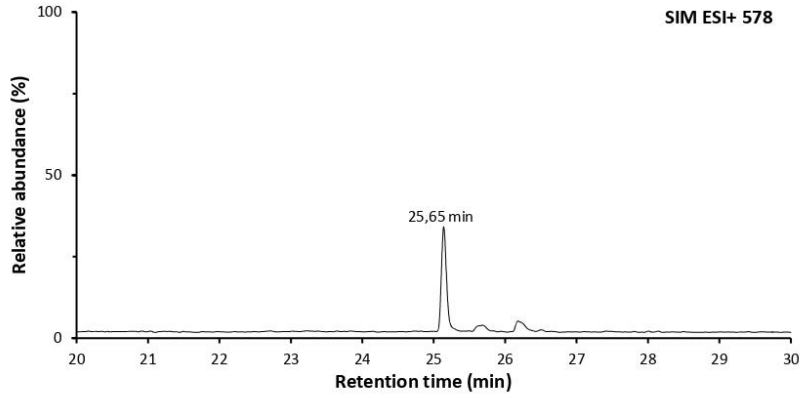


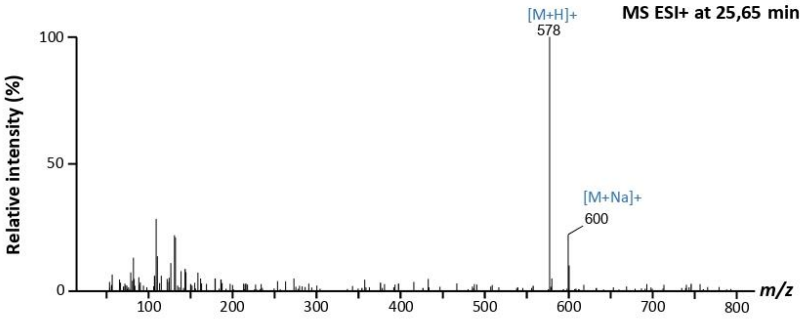
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	584
[M+Na]+	606
[M+H-mC]+	438
[M+H-mC-H_2O]+	420
[M+H-mC-mC]+	292
[M+H-mC-mC-NH_3]+	275
[M+H-mC-mC-C3chain]+	235
[M+H-mC-mC-C3chain-NH_3]+	218
[M+H-mC-mC-C4chain]+	221
[M+H-mC-mC-C4chain-NH_3]+	204
[M+H-mC-mC-Spd]+	147

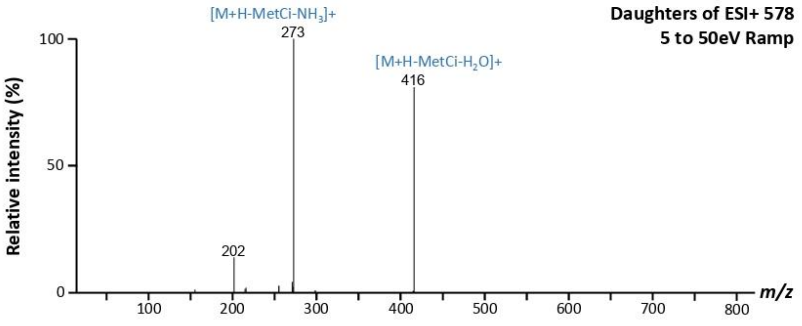


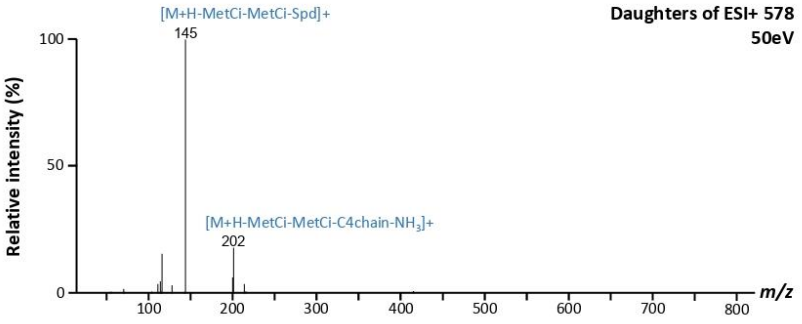
N^1, N^5, N^{10}-tri-4-methylcinnamoyl spermidine (MW 577 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=578





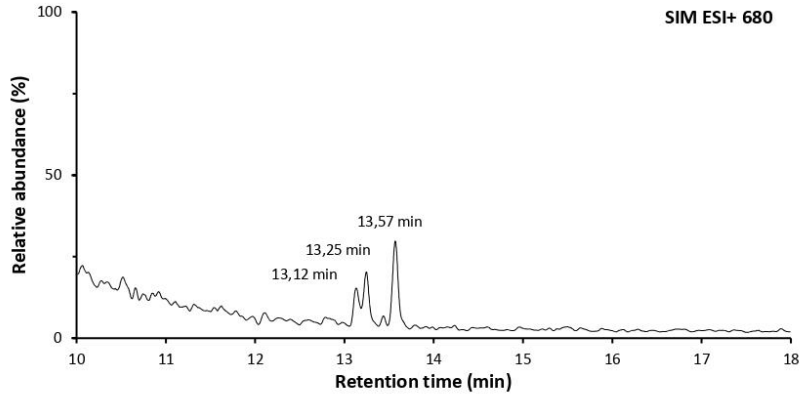


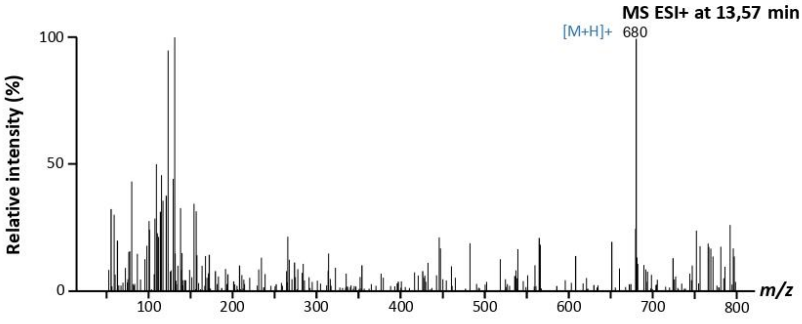


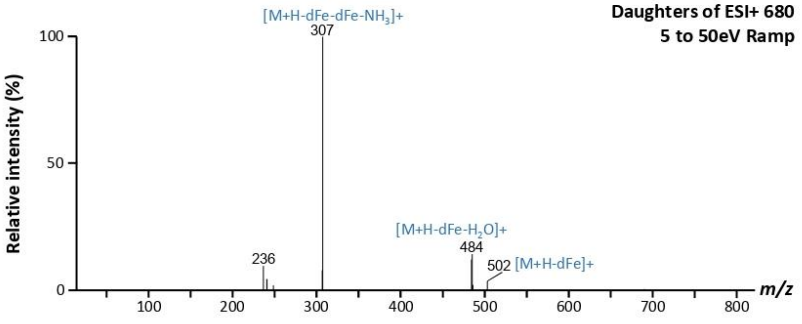
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	578
[M+Na]+	600
[M+H-MetCi]+	434
[M+H-MetCi-H_2O]+	416
[M+H-MetCi-MetCi]+	290
[M+H-MetCi-MetCi-NH_3]+	273
[M+H-MetCi-MetCi-C3chain]+	233
[M+H-MetCi-MetCi-C3chain-NH_3]+	216
[M+H-MetCi-MetCi-C4chain]+	219
[M+H-MetCi-MetCi-C4chain-NH_3]+	202
[M+H-MetCi-MetCi-Spd]+	145

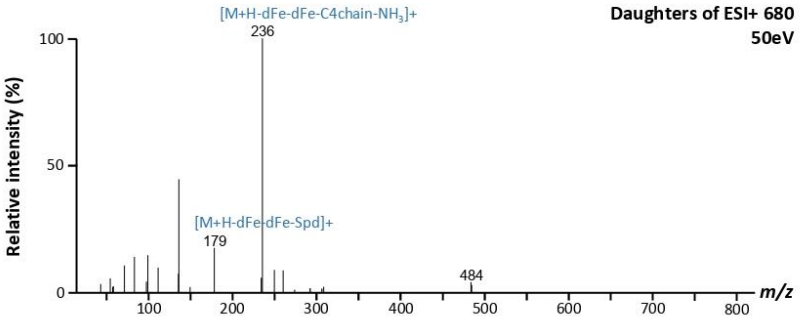
N^1, N^5, N^{10}-tri-dihydroferuloyl spermidine (MW 679 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=680





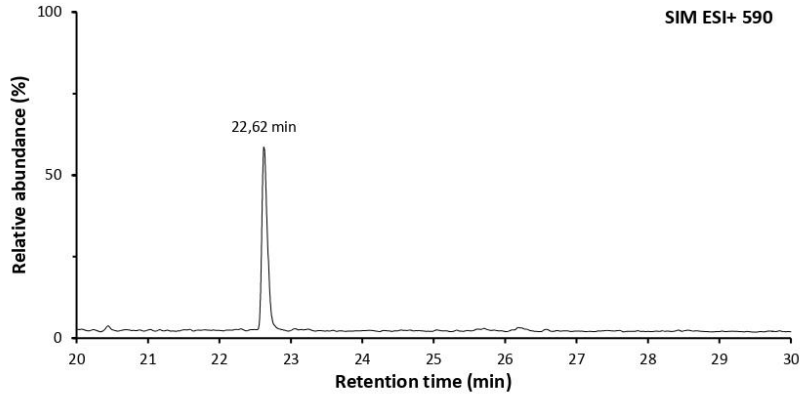


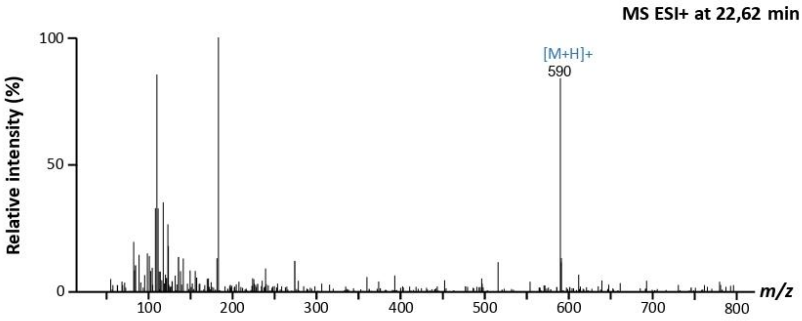


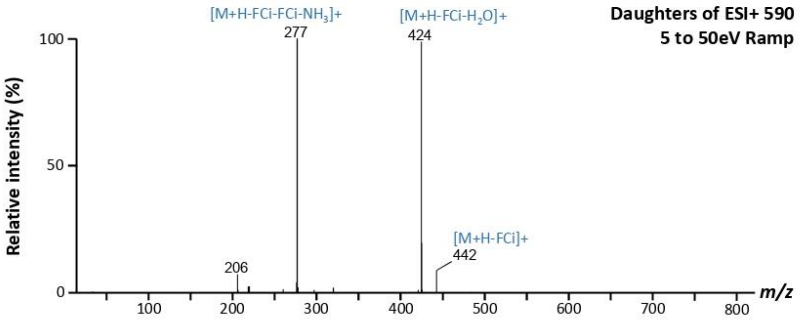
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	680
[M+Na]+	702
[M+H-dFe]+	502
[M+H-dFe-H_2O]+	484
[M+H-dFe-dFe]+	324
[M+H-dFe-dFe-NH_3]+	307
[M+H-dFe-dFe-C3chain]+	267
[M+H-dFe-dFe-C3chain-NH_3]+	250
[M+H-dFe-dFe-C4chain]+	253
[M+H-dFe-dFe-C4chain-NH_3]+	236
[M+H-dFe-dFe-Spd]+	179
[M+H-dFe-dFe-Spd-CH_4O]+	147

N^1, N^5, N^{10}-tri-4-fluorocinnamoyl spermidine (MW 589 g/mol)

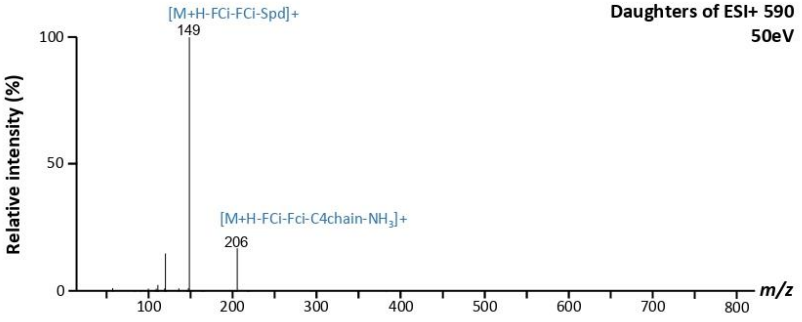
Expected pseudo-molecular ion [M+H]+: *m/z*=590





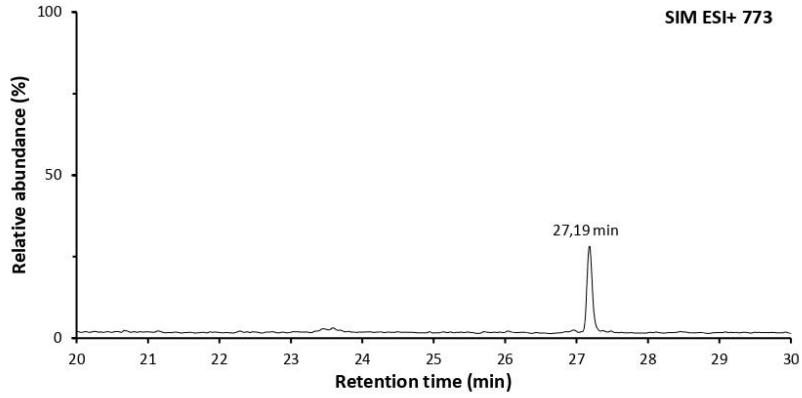


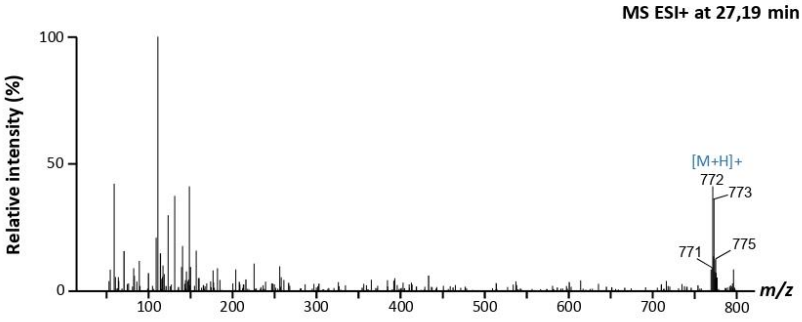
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	590
[M+Na]+	612
[M+H-FCi]+	442
[M+H-FCi-H_2O]+	424
[M+H-FCi-FCi]+	294
[M+H-FCi-FCi-NH_3]+	277
[M+H-FCi-FCi-C3chain]+	237
[M+H-FCi-FCi-C3chain-NH_3]+	220
[M+H-FCi-FCi-C4chain]+	223
[M+H-FCi-FCi-C4chain-NH_3]+	206
[M+H-FCi-FCi-Spd]+	149

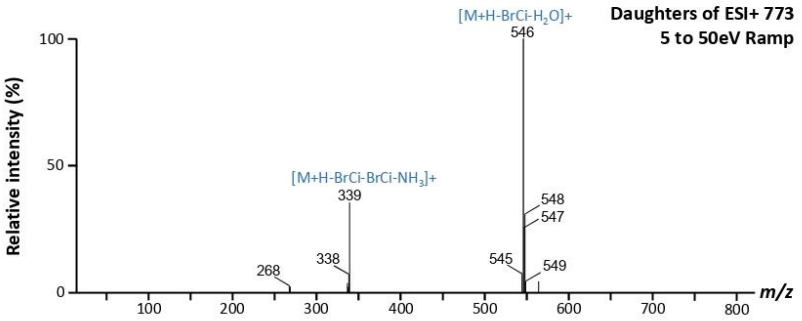


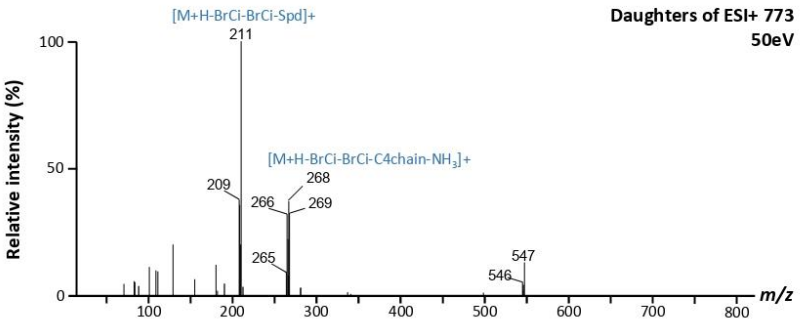
N^1, N^5, N^{10}-tri-4-bromocinnamoyl spermidine (MW 772 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=773





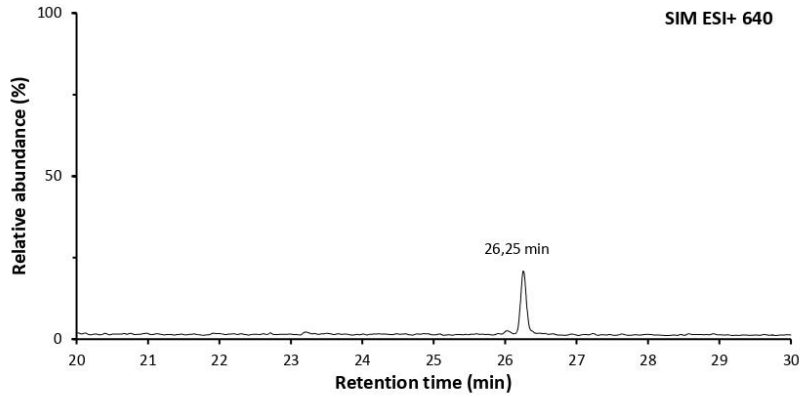


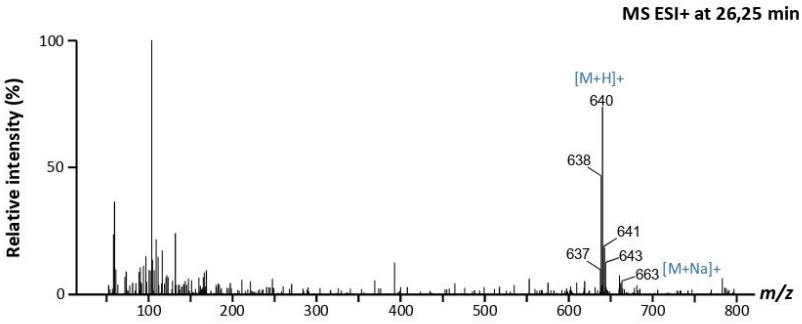


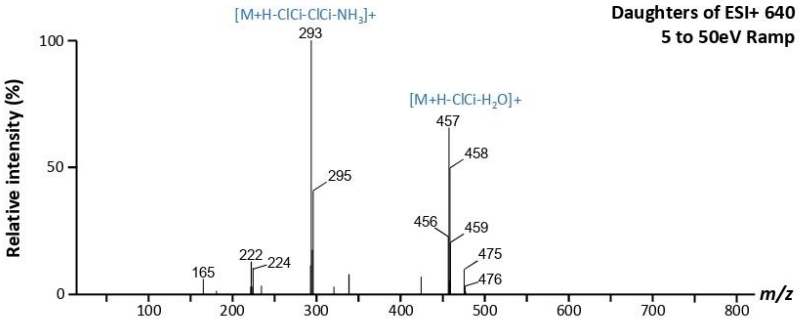
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	773
[M+Na]+	795
[M+H-BrCi]+	564
[M+H-BrCi-H_2O]+	546
[M+H-BrCi-BrCi]+	355
[M+H-BrCi-BrCi-NH_3]+	338
[M+H-BrCi-BrCi-C3chain]+	298
[M+H-BrCi-BrCi-C3chain-NH_3]+	281
[M+H-BrCi-BrCi-C4chain]+	284
[M+H-BrCi-BrCi-C4chain-NH_3]+	267
[M+H-BrCi-BrCi-Spd]+	211

N^1, N^5, N^{10}-tri-4-chlorocinnamoyl spermidine (MW 639 g/mol)

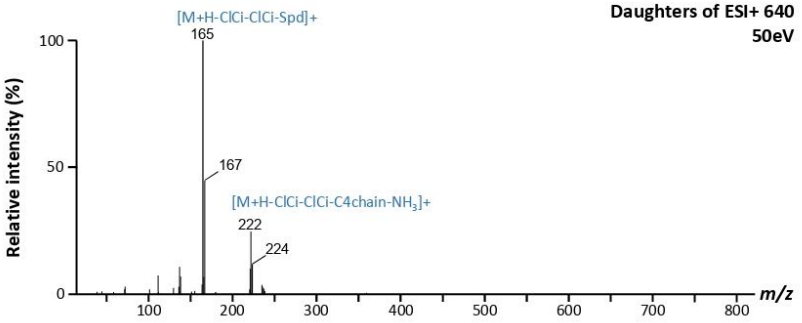
Expected pseudo-molecular ion [M+H]+: *m/z*=640





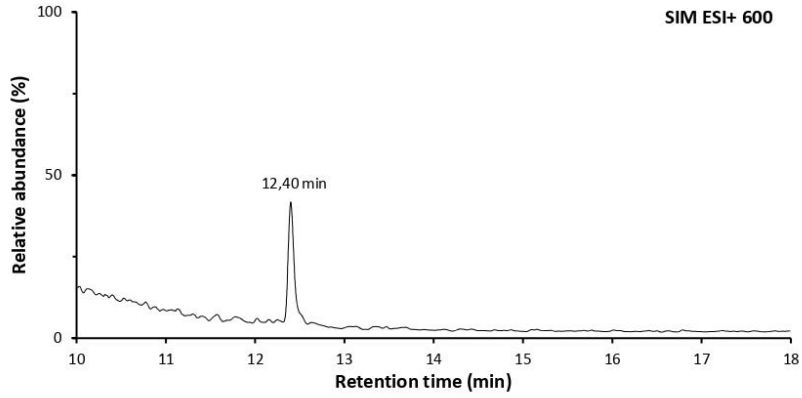


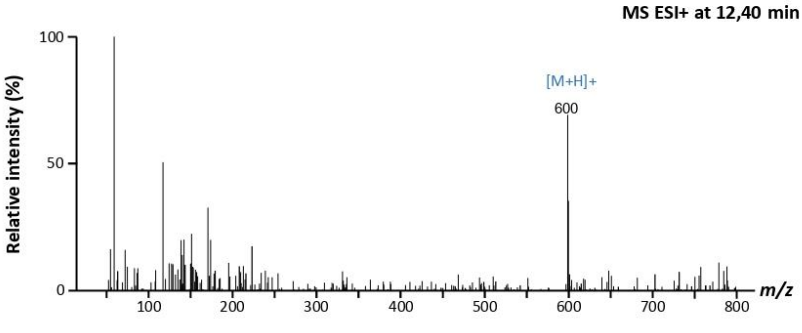
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	640
[M+Na]+	662
[M+H-ClCi]+	475
[M+H-ClCi-H_2O]+	457
[M+H-ClCi-ClCi]+	310
[M+H-ClCi-ClCi-NH_3]+	293
[M+H-ClCi-ClCi-C3chain]+	253
[M+H-ClCi-ClCi-C3chain-NH_3]+	236
[M+H-ClCi-ClCi-C4chain]+	239
[M+H-ClCi-ClCi-C4chain-NH_3]+	222
[M+H-ClCi-ClCi-Spd]+	165

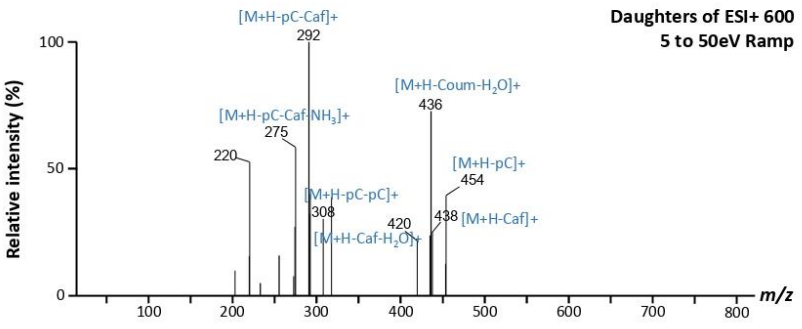


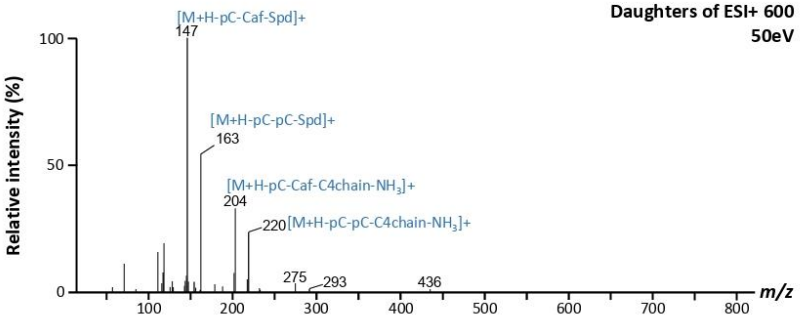
N, N'-di-*p*-coumaroyl N''-caffeoyl spermidine (MW 599 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=600





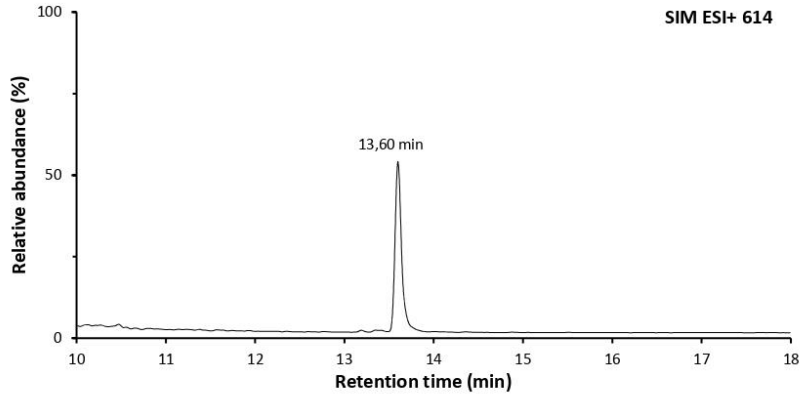


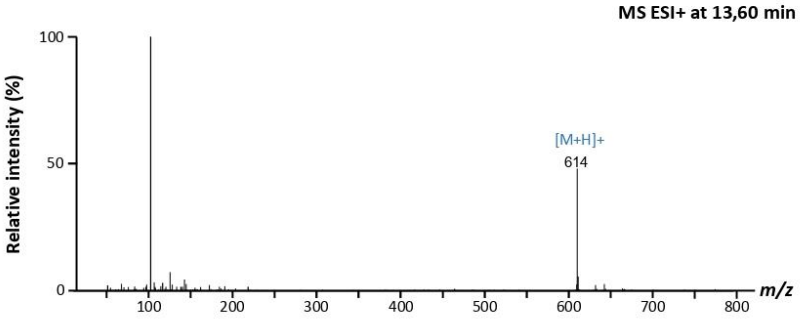


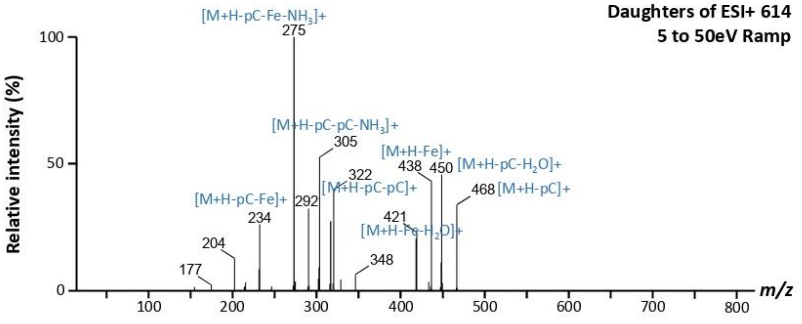
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	600
[M+Na]+	622
[M+H-pC]+	454
[M+H-pC-H_2O]+	436
[M+H-Caf]+	438
[M+H-Caf-H_2O]+	420
[M+H-pC-pC]+	308
[M+H-pC-pC-NH_3]+	291
[M+H-pC-Caf]+	292
[M+H-pC-Caf-NH_3]+	275
[M+H-pC-pC-C3chain]+	251
[M+H-pC-pC-C3chain-NH_3]+	234
[M+H-pC-Caf-C3chain]+	235
[M+H-pC-Caf-C3chain-NH_3]+	218
[M+H-pC-pC-C4chain]+	237
[M+H-pC-pC-C4chain-NH_3]+	220
[M+H-pC-Caf-C4chain]+	221
[M+H-pC-Caf-C4chain-NH_3]+	204
[M+H-pC-pC-Spd]+	163
[M+H-pC-Caf-Spd]+	147

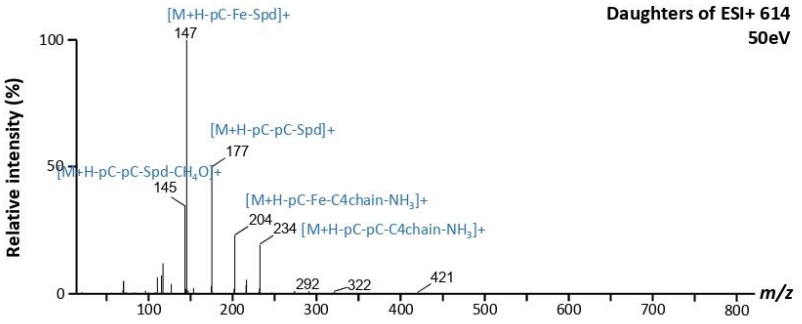
N, N'-di-*p*-coumaroyl N''-feruloyl spermidine (MW 613 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=614





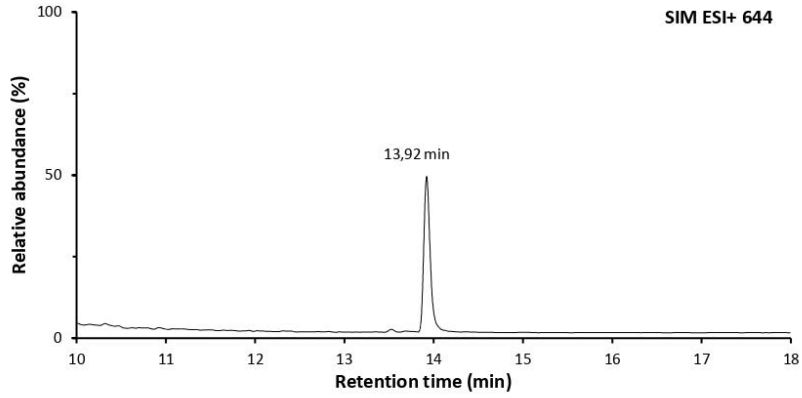


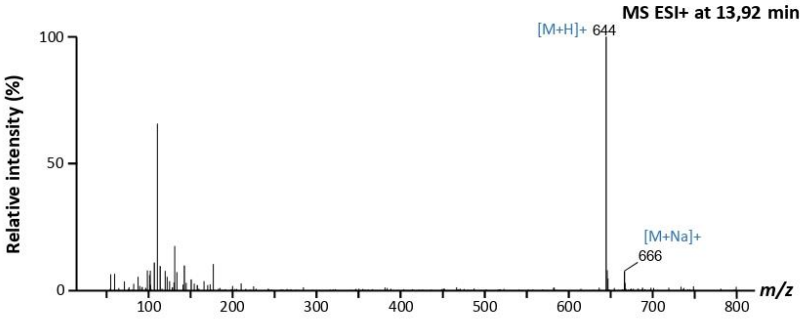


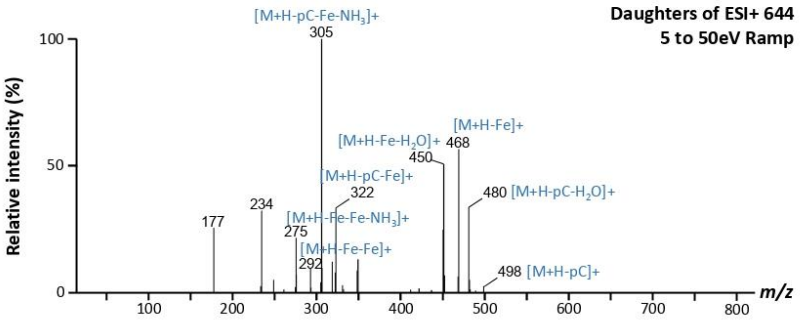
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	614
[M+Na]+	636
[M+H-pC]+	468
[M+H-pC-H_2O]+	450
[M+H-Fe]+	438
[M+H-Fe-H_2O]+	420
[M+H-pC-pC]+	322
[M+H-pC-pC-NH_3]+	305
[M+H-pC-Fe]+	292
[M+H-pC-Fe-NH_3]+	275
[M+H-pC-pC-C3chain]+	265
[M+H-pC-pC-C3chain-NH_3]+	248
[M+H-pC-Fe-C3chain]+	235
[M+H-pC-Fe-C3chain-NH_3]+	218
[M+H-pC-pC-C4chain]+	251
[M+H-pC-pC-C4chain-NH_3]+	234
[M+H-pC-Fe-C4chain]+	221
[M+H-pC-Fe-C4chain-NH_3]+	204
[M+H-pC-pC-Spd]+	177
[M+H-pC-pC-Spd-CH_4O]+	145
[M+H-pC-Fe-Spd]+	147

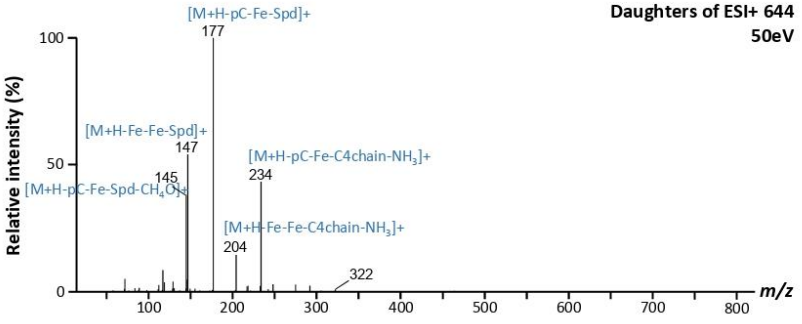
N-*p*-coumaroyl N', N''-di-feruloyl spermidine (MW 643 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=644





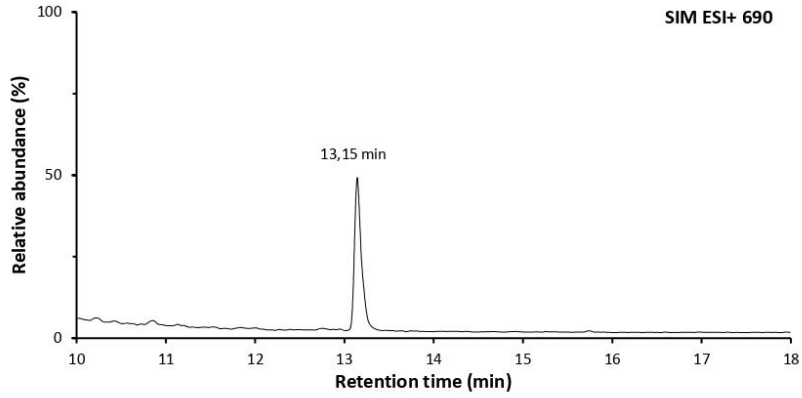


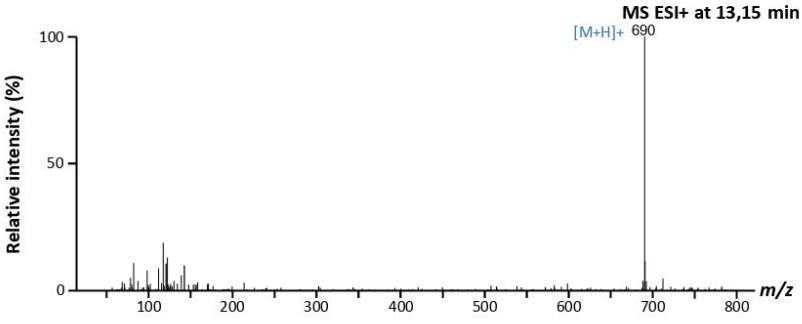


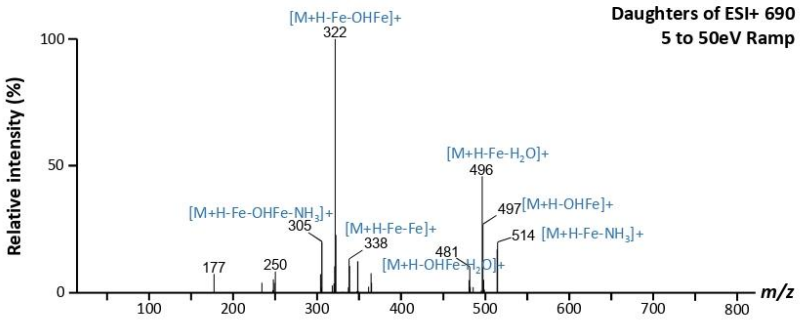
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	644
[M+Na]+	666
[M+H-pC]+	498
[M+H-pC-H_2O]+	480
[M+H-Fe]+	468
[M+H-Fe-H_2O]+	450
[M+H-pC-Fe]+	322
[M+H-pC-Fe-NH_3]+	305
[M+H-Fe-Fe]+	292
[M+H-Fe-Fe-NH_3]+	275
[M+H-pC-Fe-C3chain]+	265
[M+H-pC-Fe-C3chain-NH_3]+	248
[M+H-Fe-Fe-C3chain]+	235
[M+H-Fe-Fe-C3chain-NH_3]+	218
[M+H-pC-Fe-C4chain]+	251
[M+H-pC-Fe-C4chain-NH_3]+	234
[M+H-Fe-Fe-C4chain]+	221
[M+H-Fe-Fe-C4chain-NH_3]+	204
[M+H-pC-Fe-Spd]+	177
[M+H-pC-Fe-Spd-CH_4O]+	145
[M+H-Fe-Fe-Spd]+	147

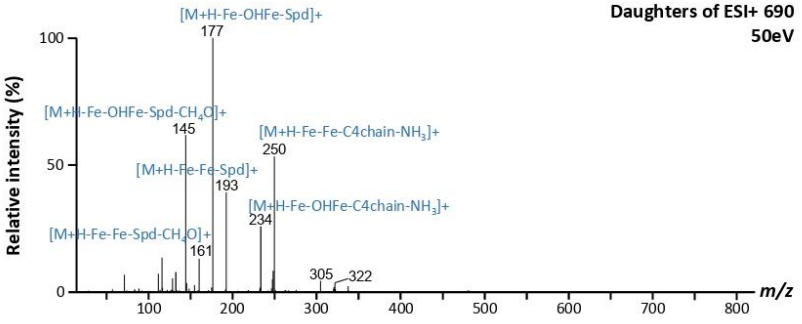
N, N'-di-feruloyl N''-5-hydroxyferuloyl spermidine (MW 689 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=690





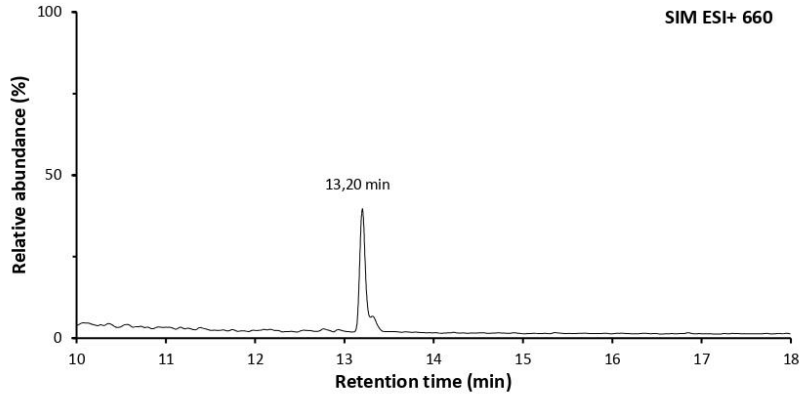


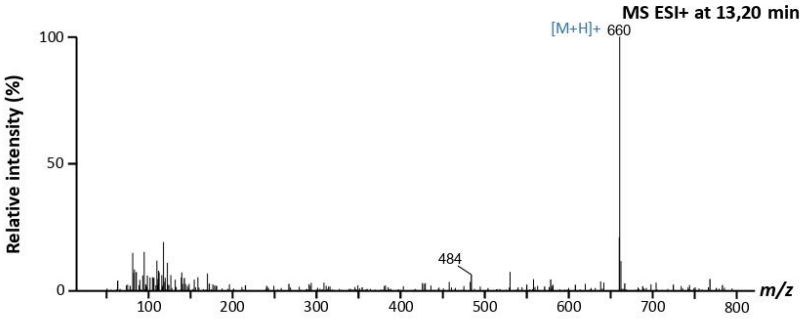


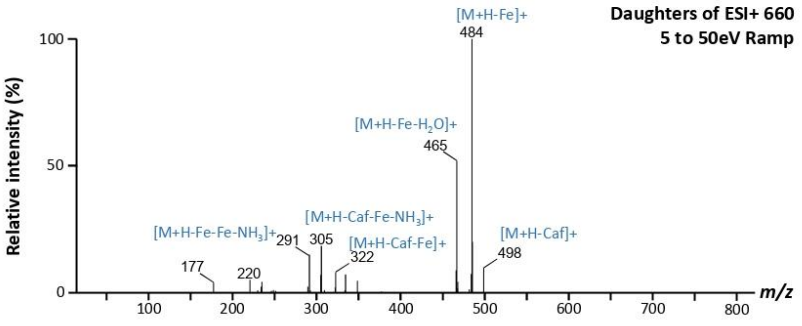
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	690
[M+Na]+	712
[M+H-Fe]+	514
[M+H-Fe-H_2O]+	496
[M+H-OHFe]+	498
[M+H-OHFe-H_2O]+	480
[M+H-Fe-Fe]+	338
[M+H-Fe-Fe-NH_3]+	321
[M+H-Fe-OHFe]+	322
[M+H-Fe-OHFe-NH_3]+	305
[M+H-Fe-Fe-C3chain]+	281
[M+H-Fe-Fe-C3chain-NH_3]+	264
[M+H-Fe-OHFe-C3chain]+	265
[M+H-Fe-OHFe-C3chain-NH_3]+	248
[M+H-Fe-Fe-C4chain]+	267
[M+H-Fe-Fe-C4chain-NH_3]+	250
[M+H-Fe-OHFe-C4chain]+	251
[M+H-Fe-OHFe-C4chain-NH_3]+	234
[M+H-Fe-Fe-Spd]+	193
[M+H-Fe-Fe-Spd-CH_4O]+	161
[M+H-Fe-OHFe-Spd]+	177
[M+H-Fe-OHFe-Spd-CH_4O]+	145

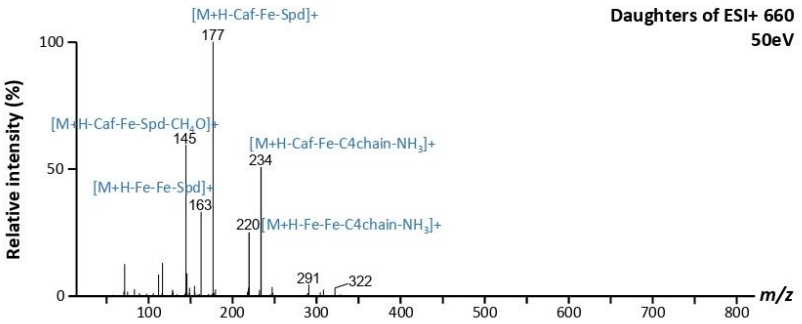
N-caffeoyl N', N''-di-feruloyl spermidine (MW 659 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=660





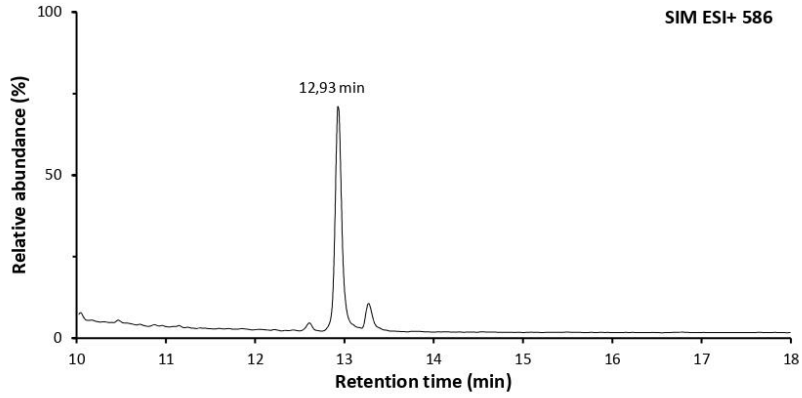


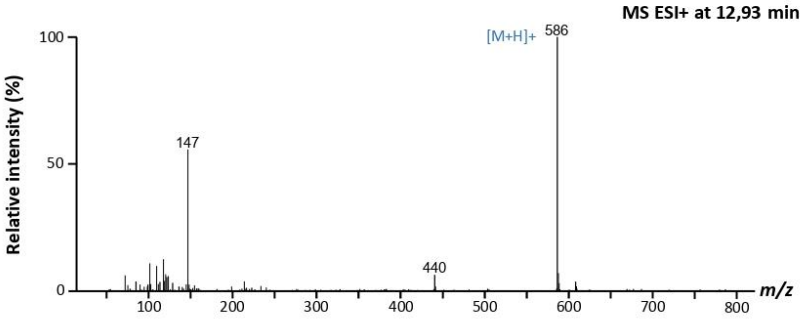


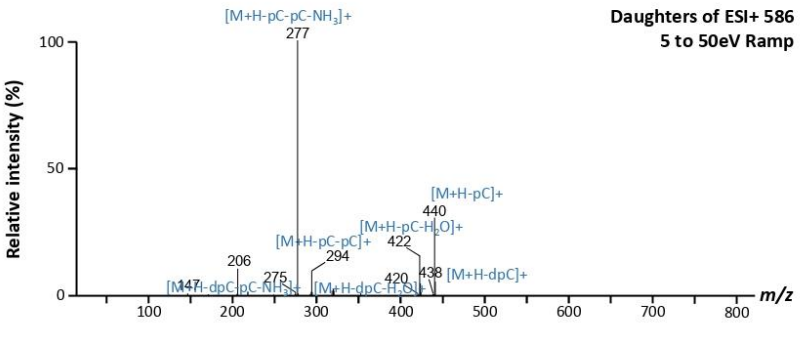
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	660
[M+Na]+	682
[M+H-Caf]+	498
[M+H-Caf-H_2O]+	480
[M+H-Fe]+	484
[M+H-Fe-H_2O]+	466
[M+H-Caf-Fe]+	322
[M+H-Caf-Fe-NH_3]+	305
[M+H-Fe-Fe]+	308
[M+H-Fe-Fe-NH_3]+	291
[M+H-Caf-Fe-C3chain]+	265
[M+H-Caf-Fe-C3chain-NH_3]+	248
[M+H-Fe-Fe-C3chain]+	251
[M+H-Fe-Fe-C3chain-NH_3]+	234
[M+H-Caf-Fe-C4chain]+	251
[M+H-Caf-Fe-C4chain-NH_3]+	234
[M+H-Fe-Fe-C4chain]+	237
[M+H-Fe-Fe-C4chain-NH_3]+	220
[M+H-Caf-Fe-Spd]+	177
[M+H-Caf-Fe-Spd-CH_4O]+	145
[M+H-Fe-Fe-Spd]+	163

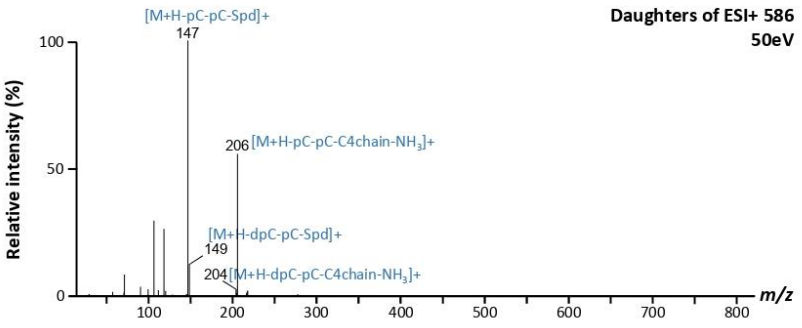
N-*p*-dihydrocoumaroyl N', N''-di-*p*-coumaroyl spermidine (MW 585 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=586





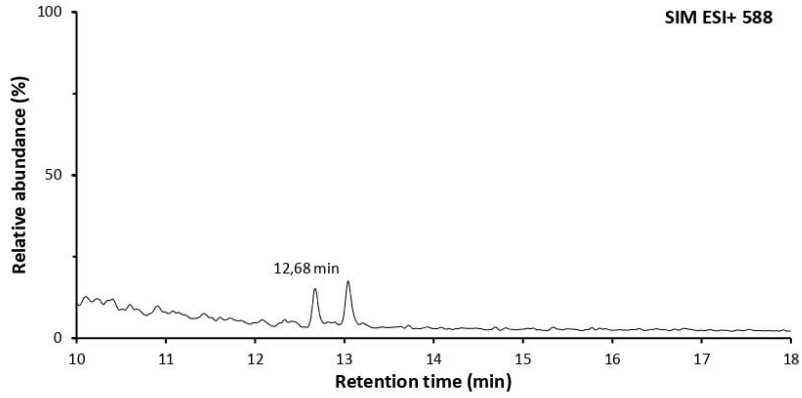


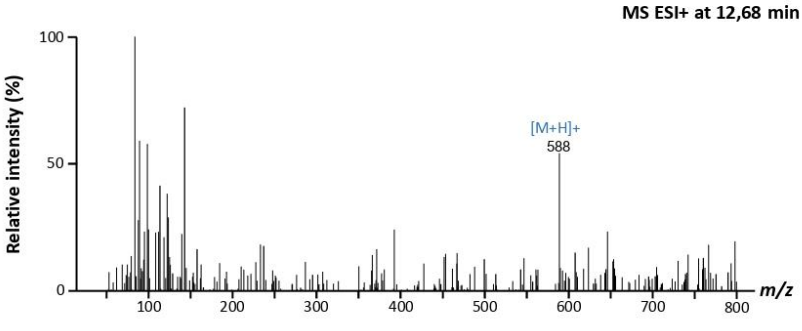


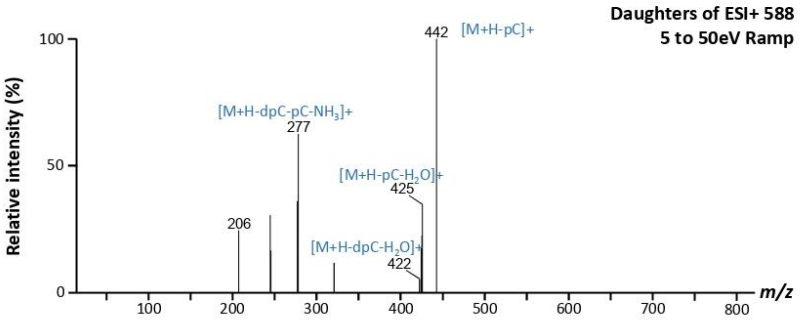
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	586
[M+Na]+	602
[M+H-dpC]+	438
[M+H-dpC-H_2O]+	420
[M+H-pC]+	440
[M+H-pC-H_2O]+	422
[M+H-dpC-pC]+	292
[M+H-dpC-pC-NH_3]+	275
[M+H-pC-pC]+	294
[M+H-pC-pC-NH_3]+	277
[M+H-dpC-pC-C3chain]+	235
[M+H-dpC-pC-C3chain-NH_3]+	218
[M+H-pC-pC-C3chain]+	237
[M+H-pC-pC-C3chain-NH_3]+	220
[M+H-dpC-pC-C4chain]+	221
[M+H-dpC-pC-C4chain-NH_3]+	204
[M+H-pC-pC-C4chain]+	223
[M+H-pC-pC-C4chain-NH_3]+	206
[M+H-dpC-pC-Spd]+	149
[M+H-pC-pC-Spd]+	147

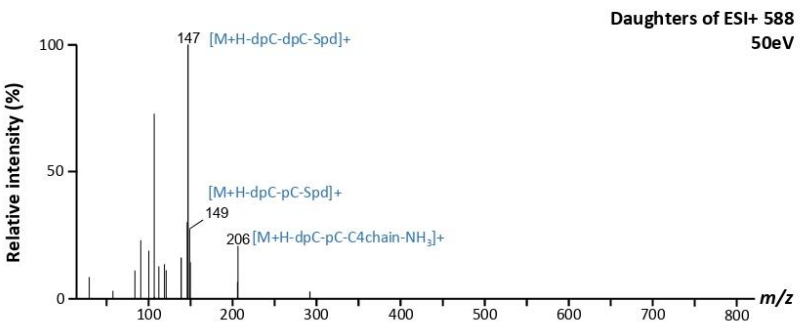
N, N'-di-*p*-dihydrocoumaroyl N''-*p*-coumaroyl spermidine (MW 587 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=588





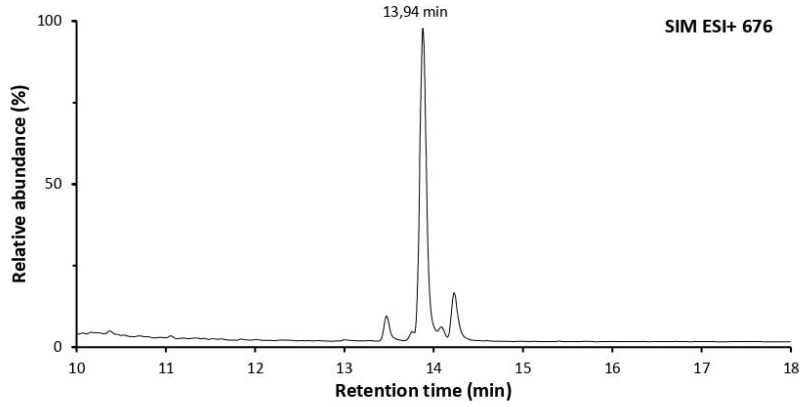


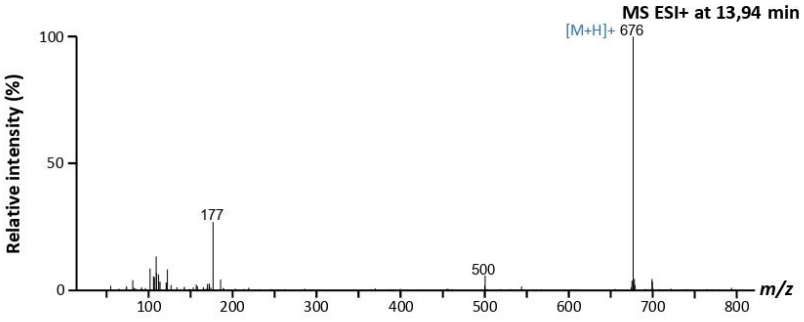


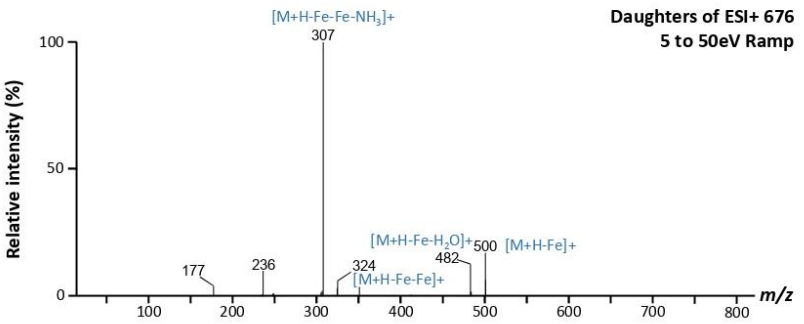
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	588
[M+Na]+	610
[M+H-dpC]+	440
[M+H-dpC-H_2O]+	422
[M+H-pC]+	442
[M+H-pC-H_2O]+	424
[M+H-dpC-pC]+	294
[M+H-dpC-pC-NH_3]+	277
[M+H-dpC-dpC]+	292
[M+H-dpC-dpC-NH_3]+	275
[M+H-dpC-pC-C3chain]+	237
[M+H-dpC-pC-C3chain-NH_3]+	220
[M+H-dpC-dpC-C3chain]+	235
[M+H-dpC-dpC-C3chain-NH_3]+	218
[M+H-dpC-pC-C4chain]+	223
[M+H-dpC-pC-C4chain-NH_3]+	206
[M+H-dpC-dpC-C4chain]+	221
[M+H-dpC-dpC-C4chain-NH_3]+	204
[M+H-dpC-pC-Spd]+	149
[M+H-dpC-dpC-Spd]+	147

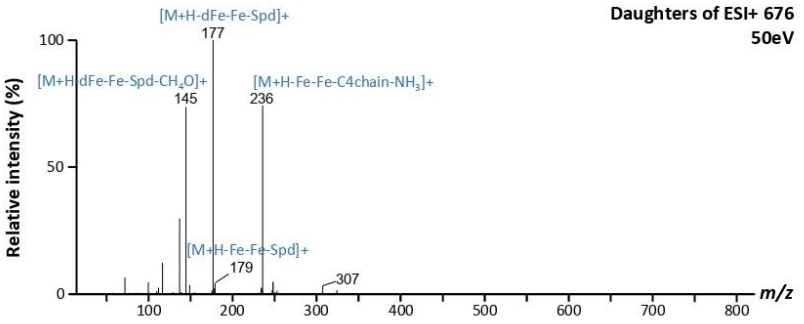
N-dihydroferuloyl N', N''-di-feruloyl spermidine (MW 675 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=676





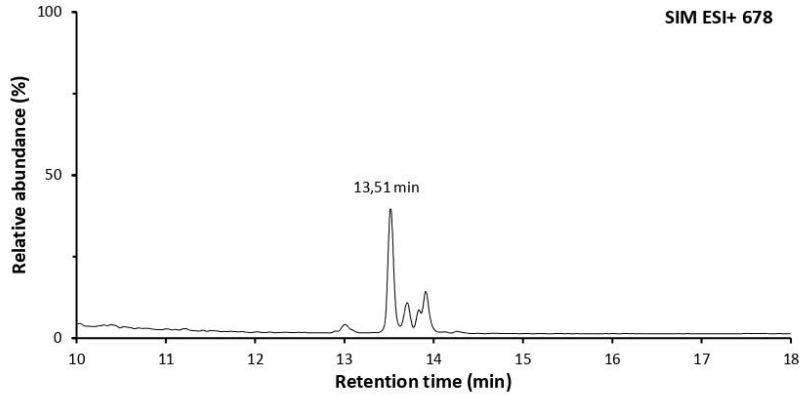


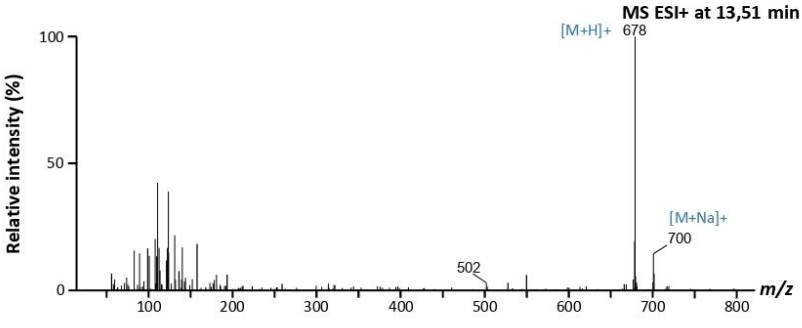


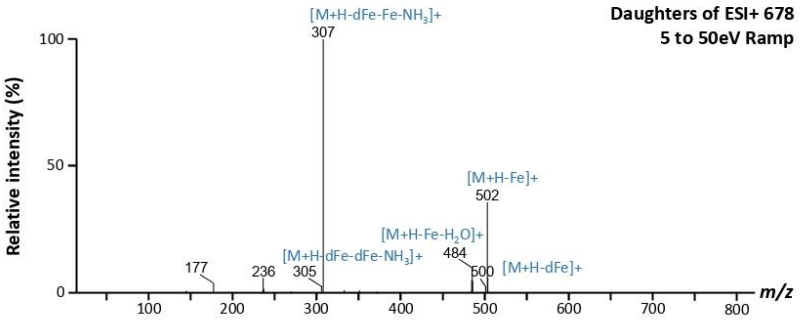
Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	676
[M+Na]+	698
[M+H-Fe]+	500
[M+H-Fe-H_2O]+	482
[M+H-dFe]+	498
[M+H-dFe-H_2O]+	480
[M+H-Fe-Fe]+	324
[M+H-Fe-Fe-NH_3]+	307
[M+H-dFe-Fe]+	322
[M+H-dFe-Fe-NH_3]+	305
[M+H-Fe-Fe-C3chain]+	267
[M+H-Fe-Fe-C3chain-NH_3]+	250
[M+H-dFe-Fe-C3chain]+	265
[M+H-dFe-Fe-C3chain-NH_3]+	248
[M+H-Fe-Fe-C4chain]+	253
[M+H-Fe-Fe-C4chain-NH_3]+	236
[M+H-dFe-C4chain]+	251
[M+H-dFe-C4chain-NH_3]+	234
[M+H-Fe-Fe-Spd]+	179
[M+H-dFe-Fe-Spd]+	177
[M+H-dFe-Fe-Spd-CH_4O]+	145

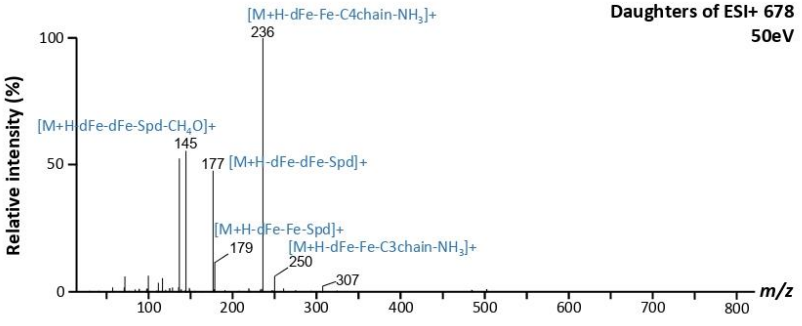
N, N'-di-dihydroferuloyl N''-feruloyl spermidine (MW 677 g/mol)

Expected pseudo-molecular ion [M+H]+: *m/z*=678









Expected pseudo-molecular ions and fragmentation	*m/z*
[M+H]+	678
[M+Na]+	700
[M+H-dFe]+	500
[M+H-dFe-H_2O]+	482
[M+H-Fe]+	502
[M+H-Fe-H_2O]+	484
[M+H-dFe-Fe]+	324
[M+H-dFe-Fe-NH_3]+	307
[M+H-dFe-dFe]+	322
[M+H-dFe-dFe-NH_3]+	305
[M+H-dFe-Fe-C3chain]+	267
[M+H-dFe-Fe-C3chain-NH_3]+	250
[M+H-dFe-dFe-C3chain]+	265
[M+H-dFe-dFe-C3chain-NH_3]+	248
[M+H-dFe-Fe-C4chain]+	253
[M+H-dFe-Fe-C4chain-NH_3]+	236
[M+H-dFe-dFe-C4chain]+	251
[M+H-dFe-dFe-C4chain-NH_3]+	234
[M+H-dFe-Fe-Spd]+	179
[M+H-dFe-dFe-Spd]+	177
[M+H-dFe-dFe-Spd-CH_4O]+	145

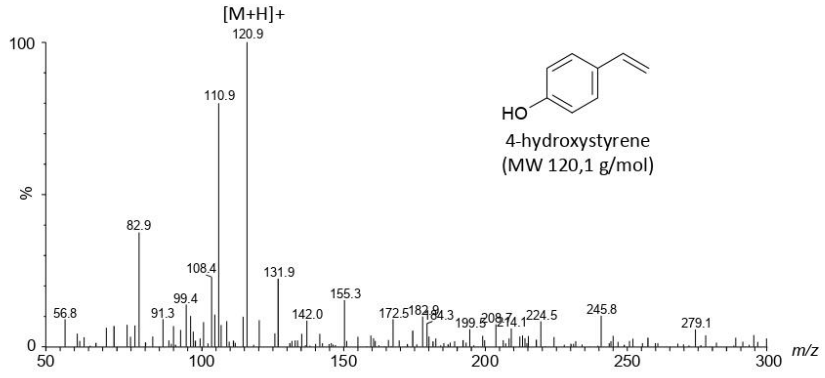


Figure S2. MS spectrum (ESI+) of the styrene-derivative (1) detected in the figure 2 and identified as 4-hydroxystyrene.

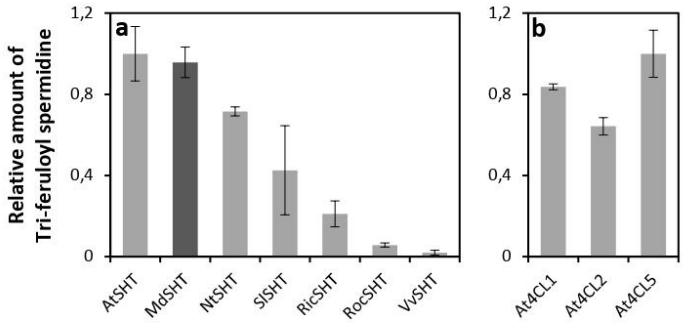


Figure S3. Comparison of the stability of plasmids used in this study, during yeast growth in rich medium. All along the study, yeast transformed with episomal expression systems were grown in rich and non-selective medium to ensure the highest biomass and the most efficient interest product synthesis. Therefore, the stability of the different plasmids used in yeast grown without selection pressure was investigated and compared. To this aim, the experiments showed in Figure 5e and 5f were repeated in equivalent conditions and using minimum selective medium (CSM-Ura-Leu with 2% glucose or 2% galactose) instead of YPD/YPGal medium: yeast fed with ferulic acid were compared among strains expressing At4CL5 and an SHT orthologue from *Arabidopsis thaliana*, *Malus* x *domestica*, *Nicotiana tabacum*, *Solanum lycopercecum*, *Ricinus communis*, *Rosa canina* or *Vitis vinifera* (a) and strains expressing MdSHT coupled with At4CL isoforms 1, 2 or 5 (b). To calculate relative amounts, the conditions with the highest production levels were fixed to 1. Error bars correspond to standard deviations of the data from biological triplicates (n=3).

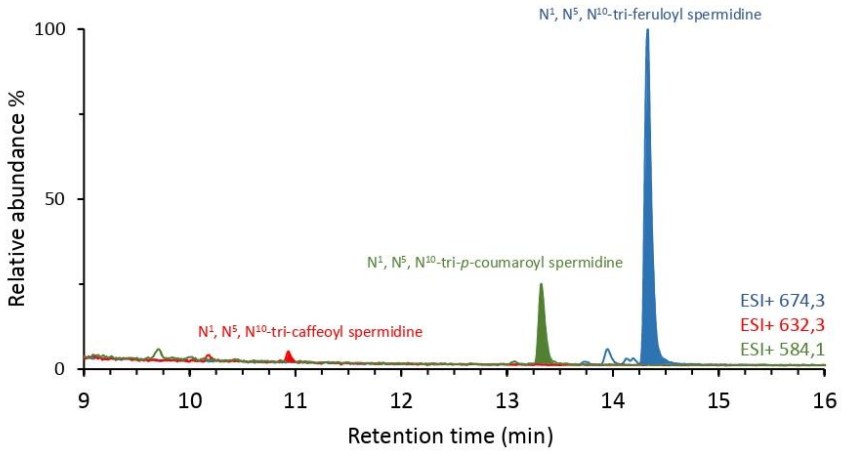
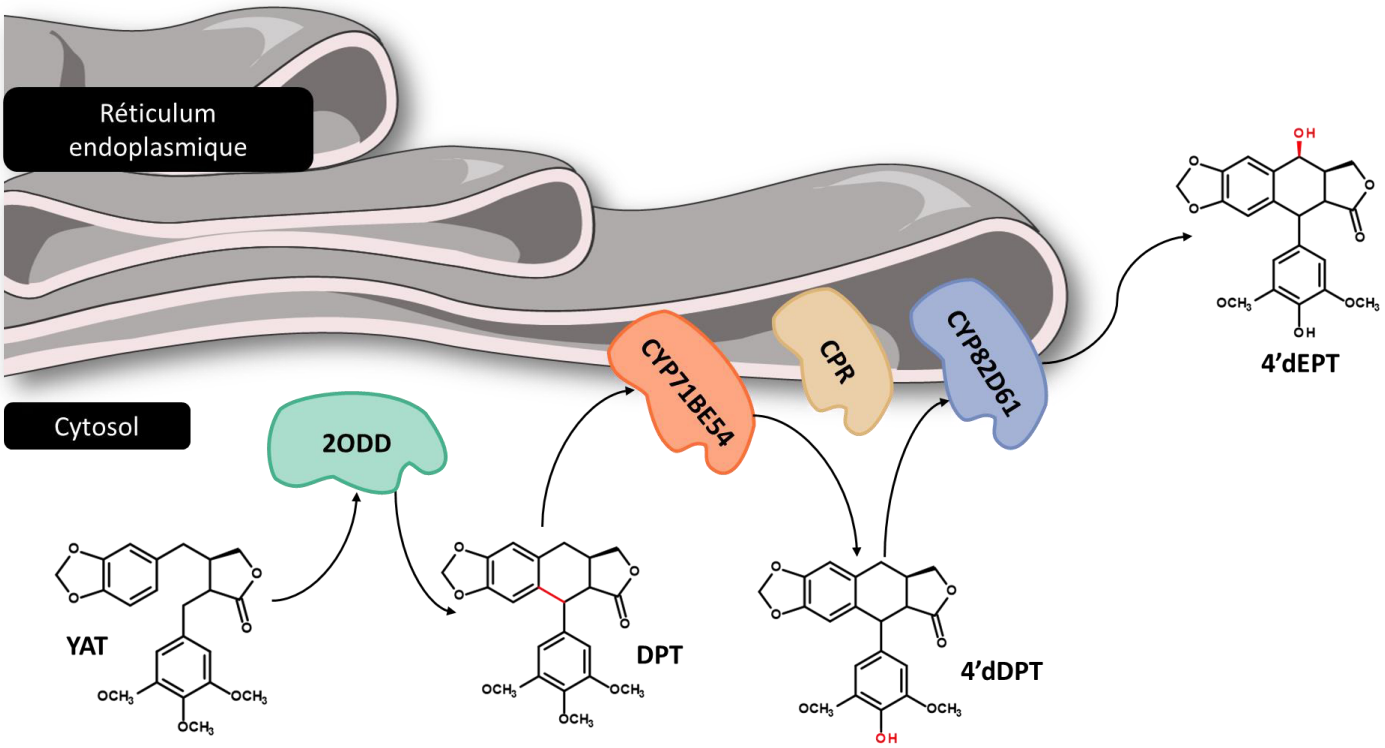


Figure S4. Comparative production rate of N^1, N^5, N^{10}-tri-*p*-coumaroyl spermidine, N^1, N^5, N^{10}-tri-caffeoyl spermidine and N^1, N^5, N^{10}-tri-feruloyl spermidine in yeast expressing 4CL/SHT couples with the highest conversion rate as shown in figure 5. LC-MS mass chromatograms superimposition of the culture medium of yeast mutant *pad1*, *tsc13*::MdECR expressing At4CL2/NtSHT (green), At4CL5/RicSHT (red) or At4CL5/MdSHT (blue) fed with *p*-coumaric acid (green), caffeic acid (red) or ferulic acid (blue). Selected mass to charge ratios in positive ion mode are used depending on attempted products: ESI+ 584,1 for N^1, N^5, N^{10}-tri-*p*-coumaroyl spermidine, ESI+ 632,3 for N^1, N^5, N^{10}-tri-caffeoyl spermidine and ESI+ 674,3 for N^1, N^5, N^{10}-tri-feruloyl spermidine.



<u>Figure 57</u> : Schéma de la stratégie de bioconversion de 4'dEPT à partir de YAT en levure.

Les modifications apportées par les enzymes sur la structure des molécules sont représentées en rouge.
YAT : (-)-yateine ; DPT : déoxypodophyllotoxine; 4'dDPT : (−)-4′-demethyldésoxypodophyllotoxine; 4'dEPT : 4′-desméthylepipodophyllotoxine.
2ODD : désoxypodophyllotoxine synthase de *P. hexandrum*; CYP71BE54 : cytochrome P450 de *P. hexandrum* ; CYP82D61 : cytochrome P450 de *P. Hexandrum ;* CPR : Cytocrhome P450 réductase.

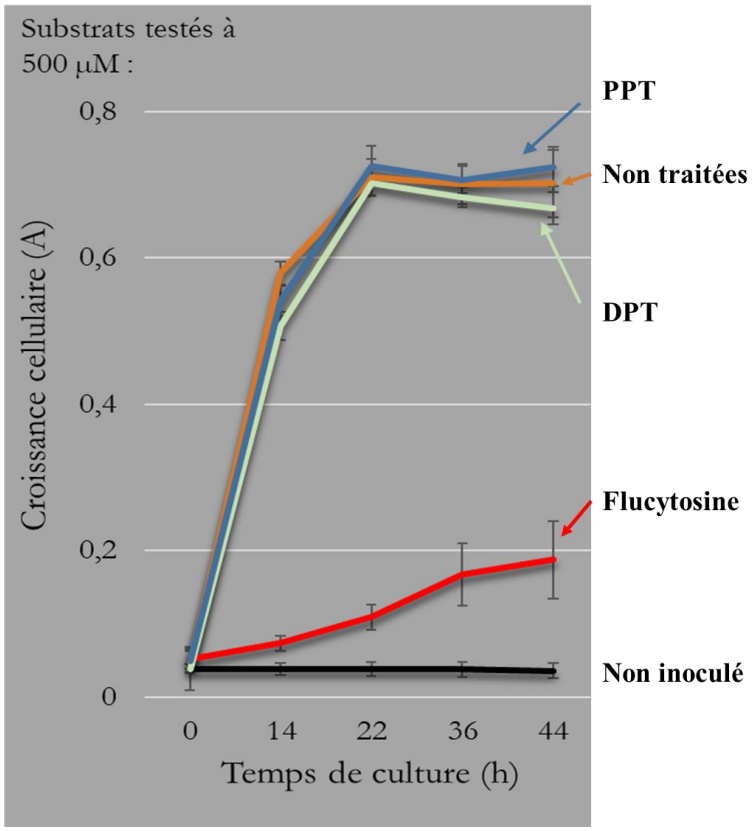
Partie 2 : Test de faisabilité d'un système de production de 4'dEPT en levure: levée de verrous techniques

L'objectif majeur de la thèse est la construction d'une souche de levure capable de produire de la 4'dEPT à une échelle industrielle par bioconversion de YAT. En effet, un système de bioconversion à partir de YAT, précurseur hautement biodisponible, permet de limiter la voie de biosynthèse transférée dans la levure à trois enzymes : la dioxygénase désoxypodophyllotoxine synthase 2ODD et les deux cytochromes P450 CYP71BE54 et CYP82D61. Avant de commencer à travailler sur une telle souche, plusieurs essais de faisabilité ont été réalisés sur la levure de laboratoire compatible avec l'utilisation de vecteurs d'expression à réplication autonome déjà utilisés dans la partie 1 pour la production de THCSpd. Si cette méthode n'est pas adaptée pour une exploitation à l'échelle industrielle étant donné l'instabilité génétique des levures dont les transgènes sont portés par des plasmides, elle offre cependant l'avantage d'être éprouvée pour l'expression de protéines recombinantes végétales et les vecteurs d'expression à réplication autonome assurent un haut niveau d'expression des gènes transférés. Les résultats ainsi obtenus et présentés ci-dessous ont tout d'abord permis de s'assurer de la non-toxicité pour la levure des lignanes dérivés de la PPT, prérequis à la poursuite de l'utilisation de cet organisme pour la bioconversion de la YAT. Dans un second temps, la fonctionnalité et l'efficacité catalytique des enzymes 2ODD, CYP71BE54 et CYP82D61 ont été évaluées en système levure. A partir de ces premiers essais enzymatiques, l'identification d'un défaut d'activité majeur dans la bioconversion de 4'dEPT à partir de 4'dDPT catalysée par CYP82D61 nous a conduit à rechercher des solutions permettant d'améliorer les rendements de bioconversion de 4'dDPT en 4'dEPT en levure.

2.1 Résultats

2.1.1 Evaluation de la toxicité des dérivés de PPT sur les levures

La PPT et certains de ses dérivés disposent d'activités biologiques sur des éléments clés de la machinerie cellulaire conférant une défense aux plantes contre des agents pathogènes, telles que les activités anti-mitotique et inhibitrice de la topoisomérase II de la PPT (Buss et al., 1995). Les plantes productrices de tels composés se sont prémunies de différentes façons contre



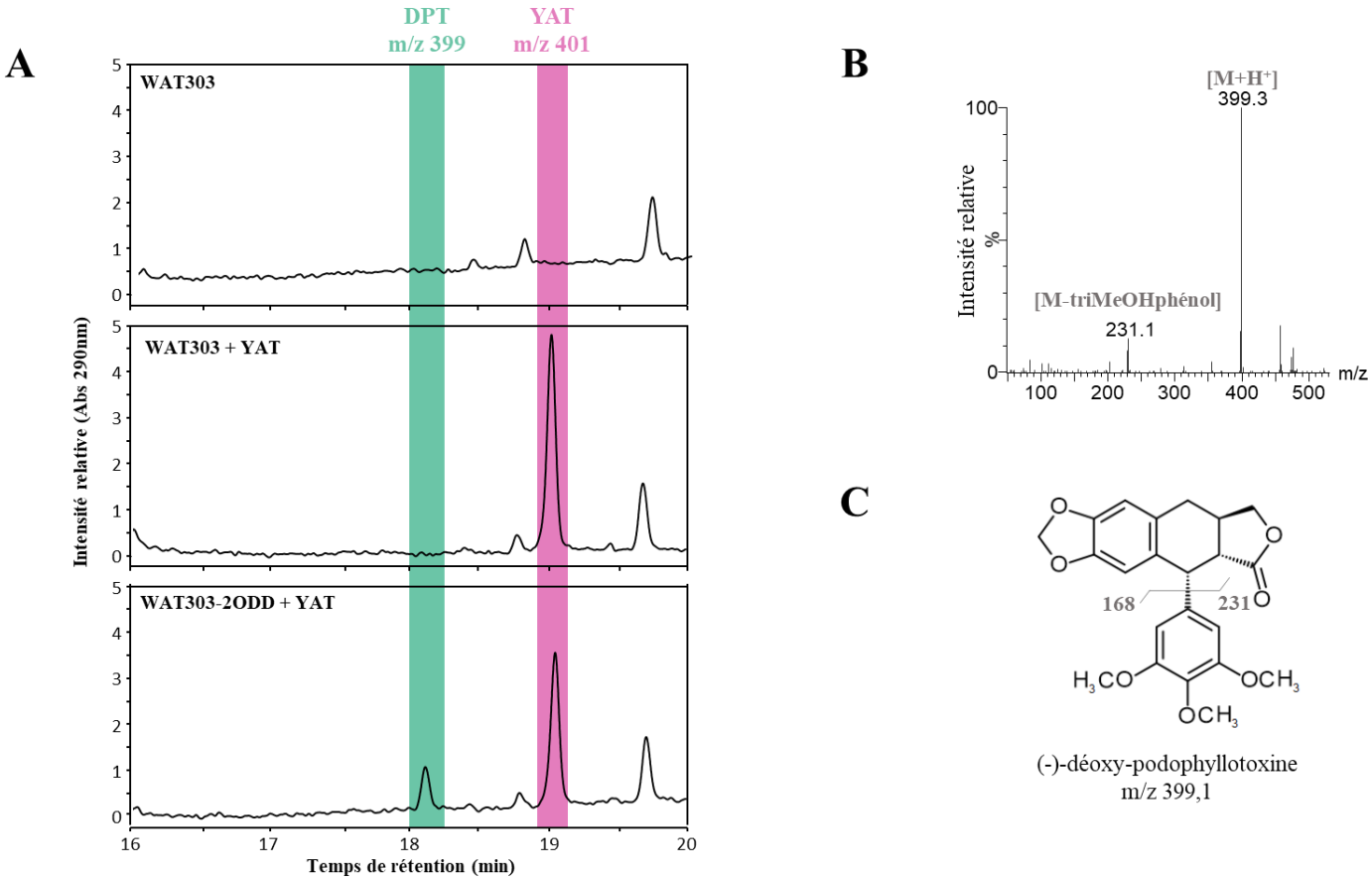
<u>Figure 58</u> : Test de toxicité de dérivés de lignanes.

Test de toxicité sur la souche WAT303 vis-à-vis de la PPT et la DPT (Le Peillet, 2018, mémoire de stage).
La flucytosine sert de témoin de toxicité, le milieu de culture non inoculé est un témoin de croissance des levures.

ces effets toxiques au fur et à mesure de l'émergence évolutive de leurs voies de biosynthèse. Ainsi, la synthèse artificielle de ces composés dans la levure, non prémunie contre ces effets, pourrait être délétère pour les cellules. Il est donc essentiel de vérifier que les lignanes de type PPT ne sont pas toxiques pour la levure afin de s'assurer de la faisabilité de la bioconversion de YAT en 4'dEPT dans ce système. Ne disposant pas de YAT, de 4'dDPT, ni de 4'dEPT au laboratoire au début du projet de thèse, les tests de toxicité ont été réalisés avec la DPT et la PPT, lignanes de type aryltétralines proches des substrats produits lors de la bioconversion, à partir d'une pré-culture de levures WAT303 préalablement diluée à une absorbance de 0,1 à 590nm. La croissance des levures cultivées en présence de 500µM de DPT ou 500µM de PPT a été suivie en mesurant la turbidité du milieu à 590nm (**Figure 58**). Ces cinétiques de croissance sur 44 heures ont été comparées à celles de levures non traitées ou traitées avec un fongicide (flucytosine). En premier lieu, les levures non traitées suivent une phase de croissance exponentielle classique avant d'atteindre un plateau après 22 heures de culture. En présence de flucytosine, la croissance des levures est beaucoup plus lente du fait de son activité antifongique. Les courbes de croissance des levures en présence de DPT et de PPT sont similaires à celles des levures non traitées, ce qui suggère que ces molécules n'ont pas d'effet toxique détectable sur les levures jusqu'à 200 mg/L. La stratégie de bioproduction de 4'dEPT en levure peut donc être envisagée.

2.1.2 Evaluation de l'efficacité catalytique des enzymes de la voie de biosynthèse du 4'dEPT en levure

Sachant que les dérivés de PPT ne sont pas toxiques pour les levures, l'efficacité catalytique des enzymes permettant la bioconversion de YAT en 4'dEPT a pu être vérifiée. Ainsi, les activités de la 2ODD, CYP71BE54 et CYP82D61 ont été testées individuellement en levure par ajout des substrats dans le milieu de culture. Cette étape nous permet de vérifier la stabilité et la fonctionnalité de chaque enzyme exprimée chez la levure et d'identifier de potentielles étapes limitantes dans la stratégie de bioconversion envisagée.

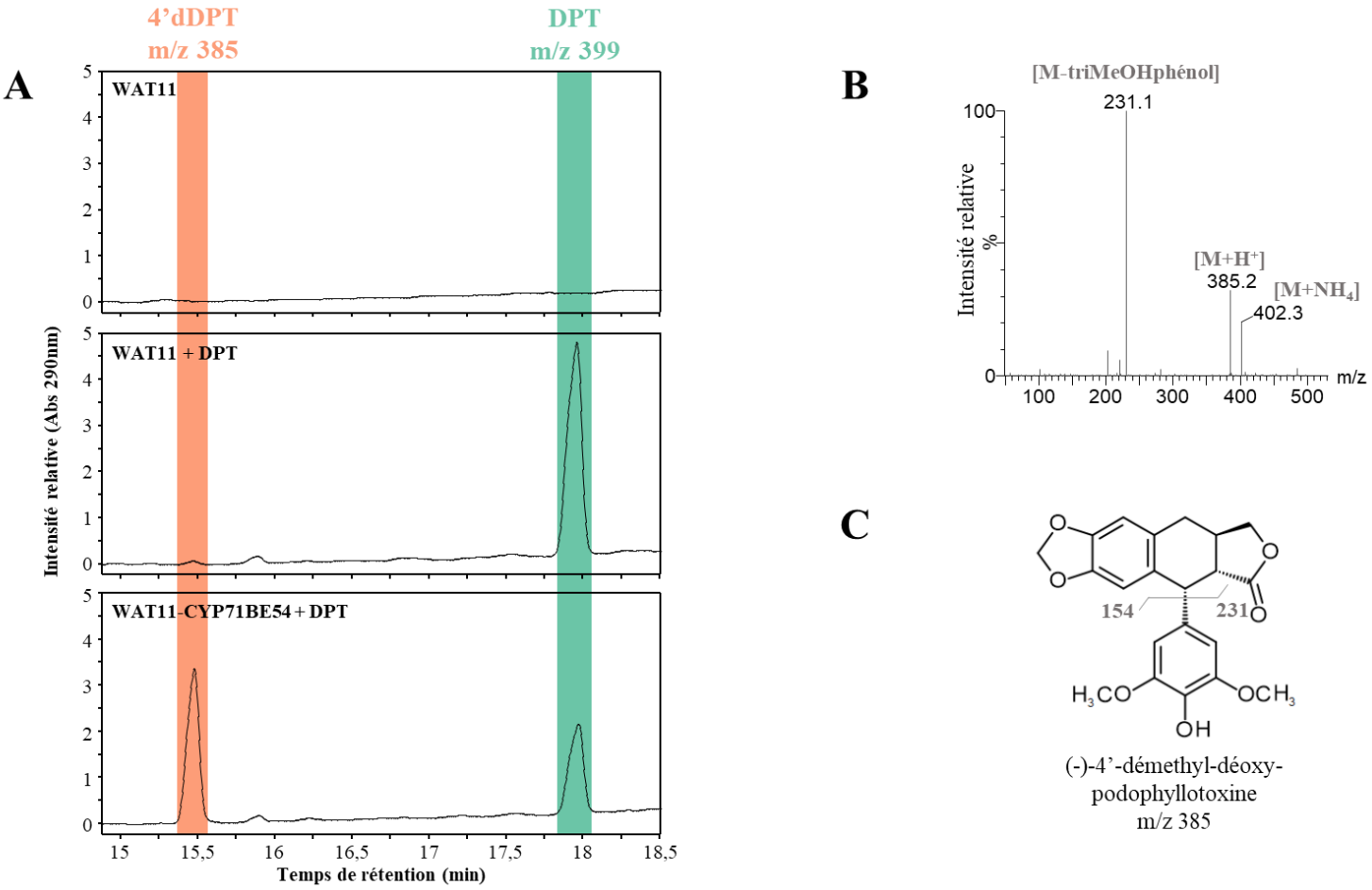


<u>Figure 59</u> : Activité catalytique de la 2ODD en levure.

(A) Chromatogrammes UV (290 nm) du surnageant de cultures de levures WAT303 non transformées sans ajout de précurseur (levure contrôle), de levures WAT303 non transformées (YAT) ou transformées (2ODD) en présence de YAT après 24H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.
(B) Spectre de masse du pic obtenu pour la DPT. M = Molécule
(C) Structure chimique de la DPT. Les numéros en gris indiquent les masses molaires de chaque partie fragmentable de la molécule.

a) Activité de la 2ODD en levure

Afin de vérifier l'activité de la 2ODD dans notre système levure, la séquence codante de l'enzyme issue *P. hexandrum* a été clonée dans le vecteur d'expression pESC-Trp et transformée dans la souche de levure WAT303. Un test de bioconversion a été réalisé sur la souche obtenue en présence de 200µM de YAT ajoutée au milieu de culture. Les chromatogrammes UV de la **Figure 59.A** présentent les analyses en UPLC des surnageants de cultures après 24 heures d'incubation de la souche exprimant la 2ODD et d'une levure sauvage. Les profils métaboliques de levures sauvages cultivées avec et sans YAT diffèrent par la présence d'un seul pic visible à 19,1 min de temps de rétention. Il s'agit de la YAT, comme l'a confirmé son spectre de masse avec un ion majoritaire au rapport masse sur charge (m/z) de 401 en ionisation positive. On peut noter que la quantité de YAT détectée après 24 heures de culture est la même qu'au départ (résultats non présentés), indiquant une absence de métabolisation par la levure sauvage. Par contre, pour la souche transformée avec le gène codant la 2ODD, le pic de la YAT diminue légèrement et sa consommation est couplée à l'apparition d'un produit à 18,1 min de rétention. Ce pic, absent des profils métaboliques de la levure sauvage, a un spectre de masse correspondant à la DPT. En effet, sur le spectre de masse en ionisation positive du produit obtenu (**Figure 59.B**), le m/z majoritaire à 399 est celui de la DPT protonée et le m/z de 231 correspond à un ion fils issue de la perte du groupement aromatique en C-7' de la DPT (**Figure 59.C**). Le rendement de conversion de YAT en DPT mesuré en 24 heures varie de 30 à 100% en fonction des clones de levures recombinantes utilisées. Cette variabilité observée de l'efficacité de conversion s'explique potentiellement par l'utilisation d'un système d'expression basé sur des plasmides à réplication autonome dont le nombre de copies diffère parmi les clones de levures obtenus après transformation. Ces résultats démontrent que la 2ODD s'exprime convenablement dans la levure et qu'elle y est fonctionnelle, ce qui n'est pas toujours le cas avec les dioxygénases. En effet, les dioxygénases nécessitent un cofacteur métallique (Fe^{2+}) et du 2-oxoglutarate pour être fonctionnelles, cofacteurs limitants dans la levure. Ces résultats confirment également que la YAT est internalisée par les levures, ce qui est indispensable pour le développement d'une stratégie de bioconversion d'un précurseur exogène. Enfin, l'absence de consommation de la YAT par les levures sauvages démontre qu'elles ne sont pas capables de le métaboliser naturellement, évitant ainsi des pertes de flux métabolique au cours de la bioconversion.

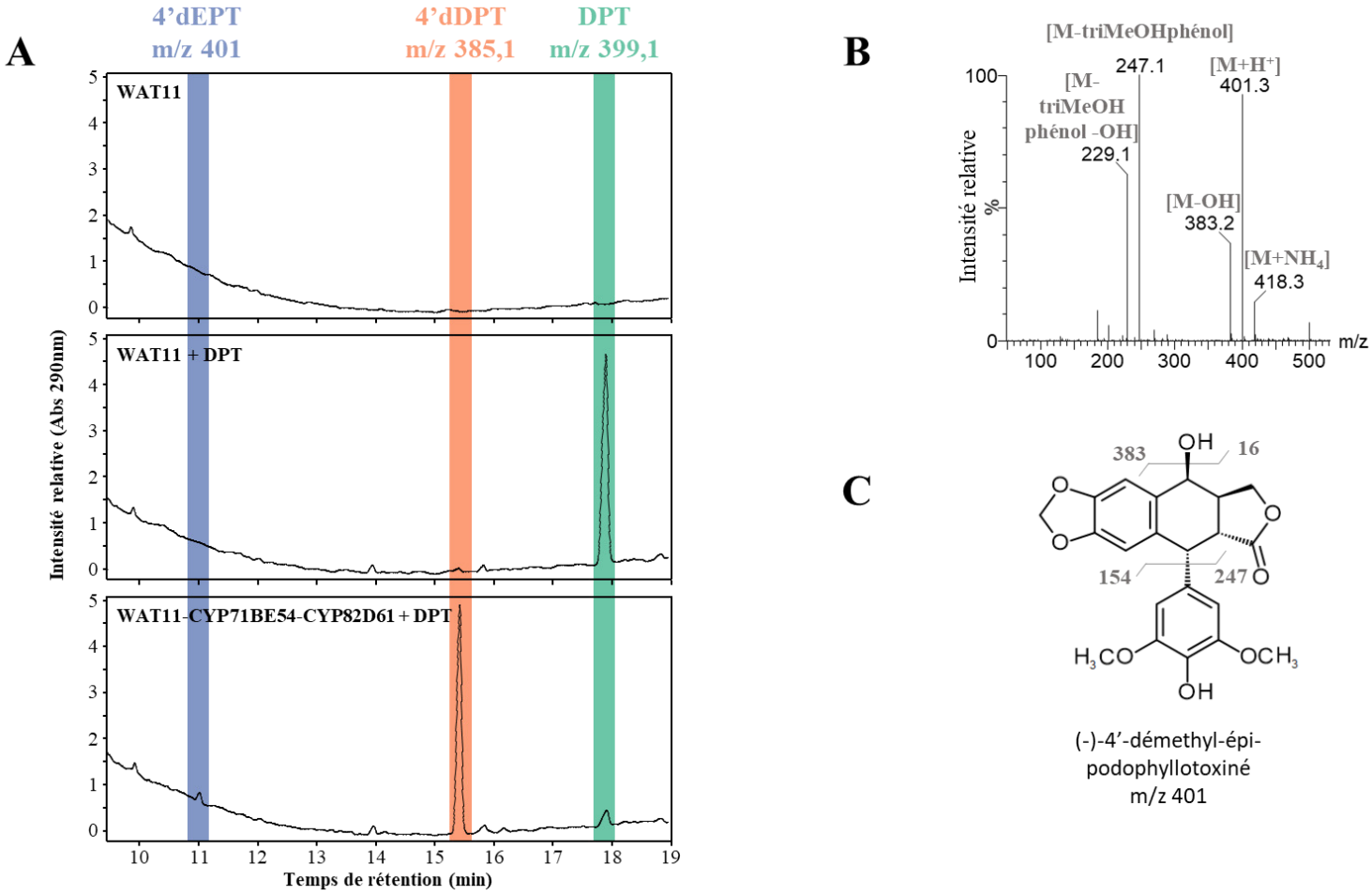


<u>Figure 60</u> : Activité catalytique de CYP71BE54 en levure.

(A) Chromatogrammes UV (290 nm) du surnageant de cultures de levures WAT11 non transformées sans ajout de précurseur (levure contrôle), de levures WAT11 non transformées (DPT) ou transformées (CYP71BE54) en présence de DPT après 24H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.
(B) Spectre de masse du pic obtenu pour la 4'dDPT. M = Molécule
(C) Structure chimique de la 4'dDPT. Les numéros en gris indiquent les masses molaires de chaque partie fragmentable de la molécule.

b) Activité de CYP71BE54 en levure

L'activité catalytique en levure de la seconde enzyme de la voie de bioconversion a ensuite été vérifiée via le clonage de la séquence codante du gène CYP71BE54 de *P. hexandrum* dans le vecteur d'expression pESC-Trp et transformation dans la souche de levure WAT11, qui correspond à la souche WAT303 exprimant la CPR d'*A. thaliana* AtATR1. En effet, les cytochromes P450 nécessitent d'être réduites par une CPR afin d'être actives, comme décrit dans l'introduction de ce manuscrit. L'activité de la CPR endogène de *S. cerevisiae* n'étant pas suffisante pour assurer la réduction de cytochromes P450 hétérologues (Parage et al., 2016), les CPR AtATR1 ou AtATR2 d'*A. thaliana* sont utilisées de façon routinière en levure pour réduire les P450 végétales. Un test de bioconversion a été réalisé sur la souche obtenue, cultivée en présence de 200µM de DPT. Sur les chromatogrammes UV présentés **Figure 60.A** de l'analyse en UPLC du surnageant de culture de la souche sauvage et de la souche transformée en présence de DPT, celle-ci est visible à 17,9 min de temps de rétention. Ainsi, les profils métaboliques de la levure sauvage et de la levure transformée cultivées avec du DPT ne diffèrent que par la présence d'un seul pic visible à 15,4 min de temps de rétention. Ce pic, absent des profils métaboliques de la levure sauvage, a un spectre de masse correspondant à la 4'dDPT. Comme le montre la **Figure 60.B**, le produit obtenu forme deux adduits en ionisation positive, aux m/z de 385 et 402 compatibles avec les formes protonée et ammonium de la 4'dDPT dont la masse molaire est de 384,4 g/mol. De plus, comme pour la DPT, la perte du groupement aromatique en C-7' génère un ion fils au m/z de 231 (**Figure 60.C**). Par contre, en comparant les spectres de la DPT et de la 4'dDPT des **Figures 59.B** et **60.B**, on remarque que les adduits de l'ion mère de 4'dDPT ne sont pas les ions majoritaires, ce qui indique une forte fragilité de la 4'dDPT par rapport à la DPT. Cette différence notable ne permet pas d'envisager une quantification des lignanes de type PPT sur la base des spectres de masse. Ainsi, les quantifications réalisées dans la suite de ce mémoire ont été faites à partir des chromatogrammes UV. Enfin, le chromatogramme correspondant à la levure sauvage en présence de DPT ne montre pas de consommation de cet intermédiaire par les levures puisque le pic ne diminue pas, ce qui indique que les levures ne sont pas capables de métaboliser la DPT (**Figure 59.A**). L'ensemble de ces résultats atteste de l'efficacité catalytique de CYP71BE54 en levure pour la conversion de la DPT en 4'dDPT. Comme pour la 2ODD, l'utilisation de plasmides à réplication autonome entraîne une hétérogénéité clonale et les taux de bioconversion de 200 µM de DPT en 24 heures varient de 60 à 99% en fonction du clone de levure utilisé (résultats non représentés).



<u>Figure 61</u> : Activité catalytique de CYP82D61 en levure.

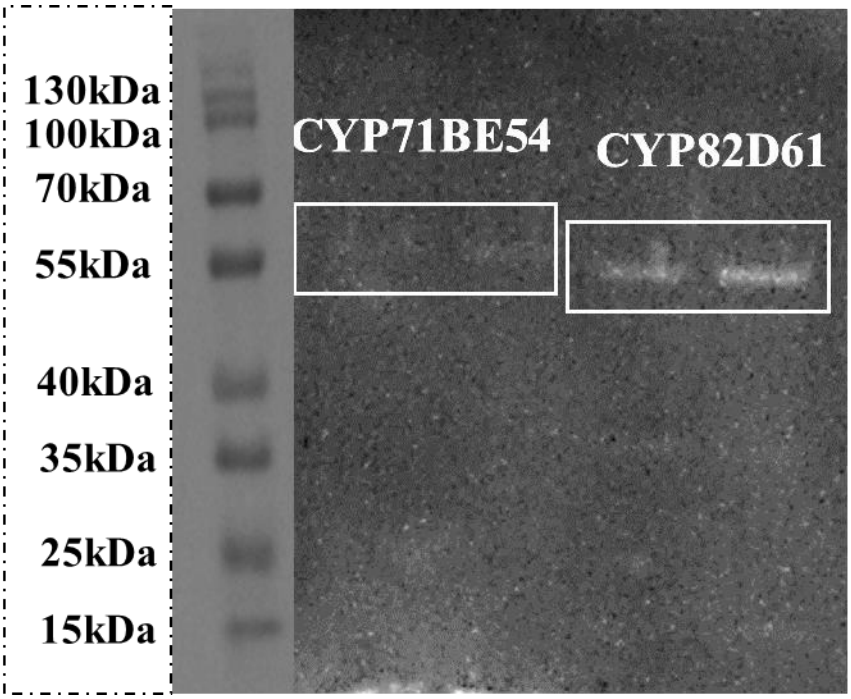
(A) Chromatogrammes UV (290 nm) du surnageant de cultures de levures WAT11 non transformées sans ajout de précurseur (levure contrôle), de levures WAT11 non transformées (DPT) ou transformées (CYP71BE54 et CYP882D61) en présence de DPT après 24H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.
(B) Spectre de masse du pic obtenu pour la 4'dEPT. M = Molécule
(C) Structure chimique de la 4'dEPT. Les numéros en gris indiquent les masses molaires de chaque partie fragmentable de la molécule.

L'efficacité catalytique de CYP71BE54 dans le système levure est ainsi très encourageante pour notre stratégie de bioconversion.

c) Activité de CYP82D61 en levure

Enfin, dans le but de vérifier l'activité catalytique en levure de l'enzyme catalysant la dernière étape de conversion de 4'dDPT en 4'dEPT, la séquence codante du gène CYP82D61 de *P. hexandrum* a été clonée dans le vecteur d'expression pESC-His et co-transformé avec le plasmide pESC-Trp-CYP71BE54 dans la souche de levure WAT11. Le test de bioconversion a été réalisé sur la souche obtenue en présence de 200µM de DPT dans le milieu de culture. En effet, ne disposant pas de 4'dDPT au laboratoire, CYP82D61 a été co-exprimée avec CYP71BE54 de manière à produire *in vivo* de la 4'dDPT directement accessible à CYP82D61. Sur le chromatogramme UV du surnageant de culture de la levure sauvage présenté **Figure 61.A**, on retrouve la DPT ajoutée au milieu de culture à 17,9 min de temps de rétention. Dans la souche co-transformée, on observe une forte diminution de la DPT et l'apparition du 4'dDPT à 15,4 min de rétention, liée à l'activité de CYP71BE54 comme décrit précédemment. Un second pic mineur est également détecté à 11,1 min de rétention (**Figure 61.A**). Ce produit réactionnel, absent des analyses réalisées précédemment avec CYP71BE54 (**Figure 60.B**) et dépendant de la présence de CYP82D61, correspond à la 4'dEPT (**Figure 61.B**). Sur le spectre de masse, les adduits protoné et ammonium de la 4'dEPT sont détectés aux m/z de 401 et 418 respectivement. A l'image de la DPT et de la 4'dDPT, la 4'dEPT se fragmente à deux endroits, sur la liaison du groupement aromatique en C-7', mais aussi au niveau du groupement hydroxyle nouvellement ajouté en C-7 (**Figure 61.C**). Ces deux fragmentations individuelles ou cumulées sont à l'origine des ions fils aux m/z 247, 383 et 229. Ces résultats indiquent que CYP82D61 est capable de convertir la 4'dDPT en 4'dEPT mais avec une activité trop faible en l'état pour pouvoir être exploitable dans notre système de production en levure.

Afin de s'assurer du niveau d'expression de CYP82D61 dans la levure, une étiquette FLAG a été ajoutée à la protéine pour la détection par Western Blot. Une version de CYP71BE54 fusionnée à un FLAG a également été utilisée comme contrôle positif puisque cette enzyme n'a pas de problème d'expression ou de stabilité au vu de son activité en levure. L'étiquette FLAG est placée en C-terminale des deux protéines pour ne pas interférer avec le



<u>Figure 62</u> : Comparaison de l'accumulation de CYP71BE54 et CYP82D61 en levure par Western Blot.

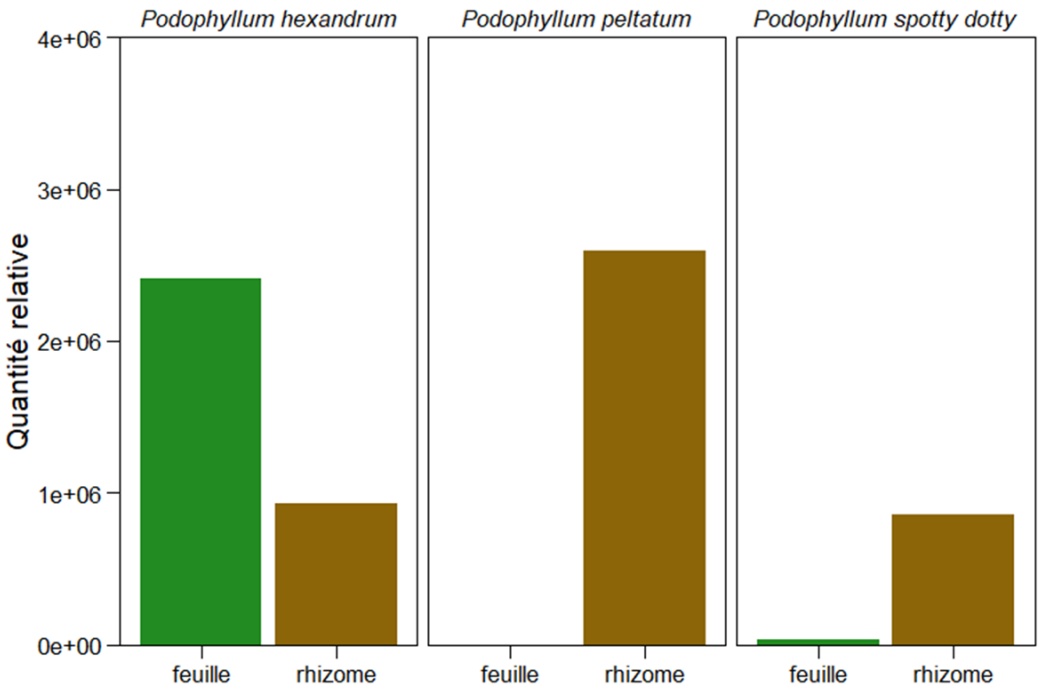
Western Blot réalisé sur les fractions totales des cultures de WAT303/CYP71BE54 et WAT303/CYP82D61. Taille attendue pour CYP71BE54 et CYP82D61 : 55kDa.

domaine transmembranaire de ces P450. Après transformation des plasmides dans la souche WAT303 (pESC-His et pESC-Ura respectivement), un Western Blot à l'aide d'un anticorps anti-FLAG a été réalisé sur les culots cellulaires issus des cultures des souches. Le résultat présenté **Figure 62** montre la détection spécifique d'une protéine dans chaque souche de levure, dont le poids moléculaire, autour de 55kDa, correspond à celui des deux cytochromes P450, avec une taille prédite légèrement supérieure (56,4 kDa pour CYP71BE54 et 57,7 kDa pour CYP82D61). L'intensité des bandes de la version taguée de CYP82D61 est légèrement plus intense que celles de la version taguée de CYP71BE54, suggérant que CYP82D61 ne présente pas de défaut d'accumulation dans la levure. Plusieurs autres hypothèses peuvent donc être formulées pour expliquer la faible activité de CYP82D61 observée dans nos levures, incluant un manque de spécificité de l'enzyme pour la 4'dDPT ou encore un défaut de réduction de l'enzyme par la CPR exprimée dans la souche de levure employée.

2.1.3 Alternatives à l'utilisation de CYP82D61 pour catalyser l'étape finale de bioproduction de 4'dEPT

L'étape de conversion de 4'dDPT en 4'EPT étant une étape limitante pour envisager la bioproduction de 4'dEPT en quantité, nous avons orienté nos recherches vers une alternative à l'utilisation de CYP82D61 issue de *P. hexandrum*. Pour identifier d'autres enzymes capables de réaliser cette bioconversion en levure avec de meilleurs rendements, nous avons exploré la diversité génétique de CYP82D61, en recherchant des isoformes et des orthologues de l'enzyme dans le genre Podophyllum, mais aussi en utilisant le P450 humain CYP3A4 rapporté pour avoir une activité de synthèse de 4'dEPT (Vasilev et al., 2006).

La présence du substrat 4'dDPT dans le milieu des levures est nécessaire pour valider l'activité des différents candidats testés dans cette partie. Sachant que nous ne disposons pas de ce substrat au laboratoire, à la différence de la YAT, les gènes codant CYP71BE54 et la 2ODD de *P. hexandrum* ont été clonés dans le vecteur d'expression pESC-Trp et insérés dans la souche WAT11 de *S. cerevisiae* (contenant une CPR d'*A. thaliana*). La souche de levure obtenue convertit 70% de la YAT ajoutée au milieu de culture en 4'dDPT qui peut alors être disponible pour la dernière étape de conversion. Cette souche, nommée WAT11-2ODD-CYP71BE54, a donc été transformée avec les constructions de gènes candidats qui ont été testés dans cette partie.



<u>Figure 63</u> : Comparaison de l'accumulation des dérivés de 4'dDPT hydroxylés en C-7 dans les feuilles et le rhizome de *Podophyllum hexandrum*, *Podophyllum peltatum* et *Podophyllum spotty dotty.*

Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

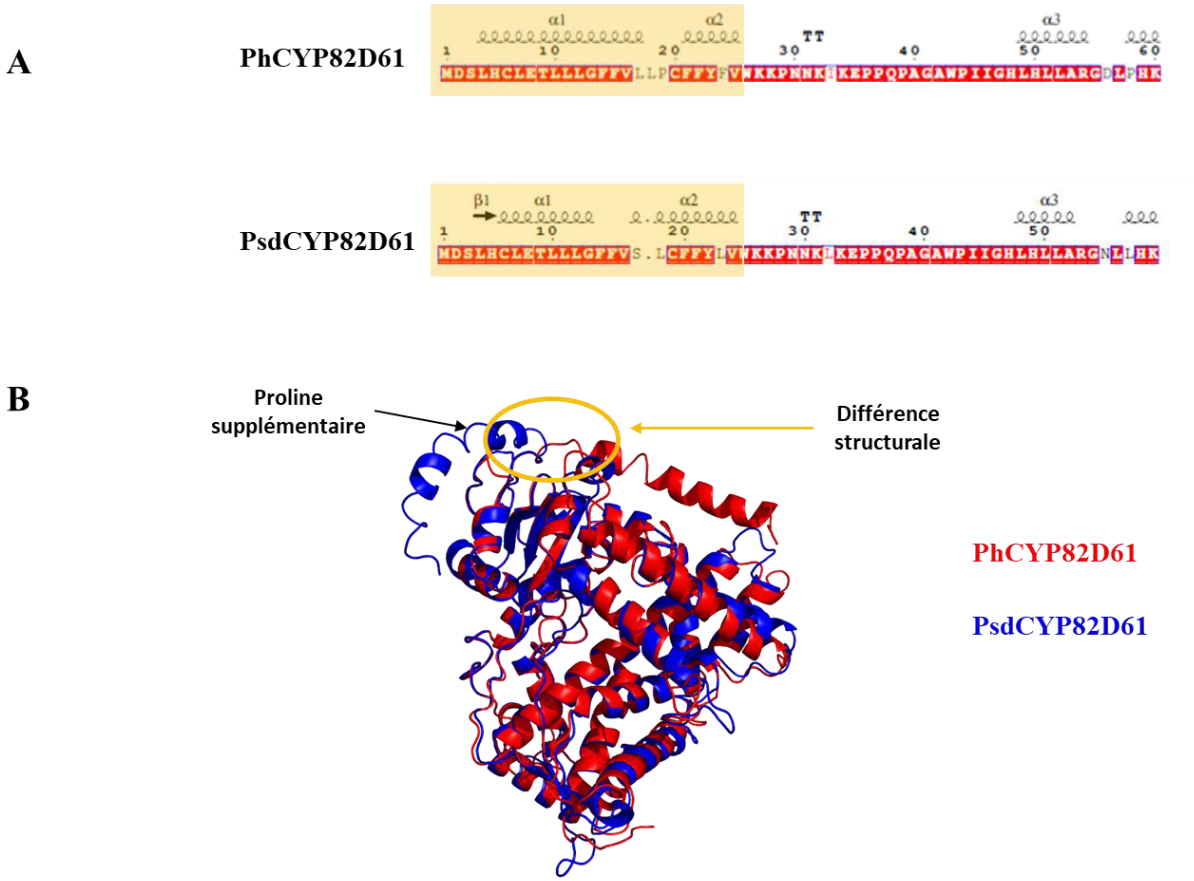
a) Recherche d'isoformes de CYP82D61 plus efficaces pour la synthèse de 4'dEPT

La faible activité de CYP82D61 peut être due à une faible spécificité de l'enzyme pour la 4'dDPT. On peut imaginer qu'un autre P450, proche de CYP82D61, soit plus efficace pour catalyser la conversion de 4'dDPT en 4'dEPT *in planta*. Une isoforme potentielle de CYP82D61 a donc été recherchée via des analyses transcriptomiques sur différents organes de *P. hexandrum*. Une séquence d'intérêt, annotée DN717, a été identifiée sur la base de son homologie de séquence avec CYP82D61 (93% d'identité nucléotidique, 89% d'identité protéique) et de son profil d'expression corrélé avec ceux des gènes en amont dans la voie de biosynthèse des lignanes.

La séquence de DN717 a été amplifiée par RT-PCR à partir d'ARN de feuilles de *P. hexandrum* puis clonée dans le vecteur d'expression pESC-His. Le vecteur pESC-His-DN717 a ensuite été transformé dans la souche WAT11-2ODD-CYP71BE54 et l'activité de l'enzyme a été comparée à celle de CYP82D61 en cultivant la souche en présence de YAT. Toutefois, les résultats obtenus n'ont montré aucune activité détectable de DN717 sur la 4'dDPT (données non représentées). L'absence de synthèse de 4'dEPT par DN717 laisse à penser que cette activité, même faible, est peut-être spécifique à CYP82D61. Nous avons donc décidé de comparer l'activité catalytique en levure d'orthologues de CYP82D61 issues d'autres espèces de Podophyllum accumulatrices de 4'dEPT.

b) Comparaison des activités catalytiques de CYP82D61 entre espèces du genre Podophyllum

L'accumulation de lignanes tels que la YAT et la DPT est un trait conservé chez les espèces du genre Podophyllum, en revanche, la PPT ou ses formes dérivées (hydroxylées, méthoxylées, glycosylées…) s'y accumulent en quantités variables. Ces différences entre espèces peuvent s'expliquer par une dérive génétique des enzymes de fin de voie de biosynthèse. Nous avons exploré l'effet de cette variabilité sur l'efficacité catalytique de la CYP82D61 en comparant l'activité catalytique d'orthologues issues de différentes espèces de Podophyllum. En plus de *P. hexandrum*, deux autres espèces ont été utilisées : *Podophyllum peltatum* et *Podophyllum spotty dotty*. En premier lieu, les profils d'accumulation de lignanes



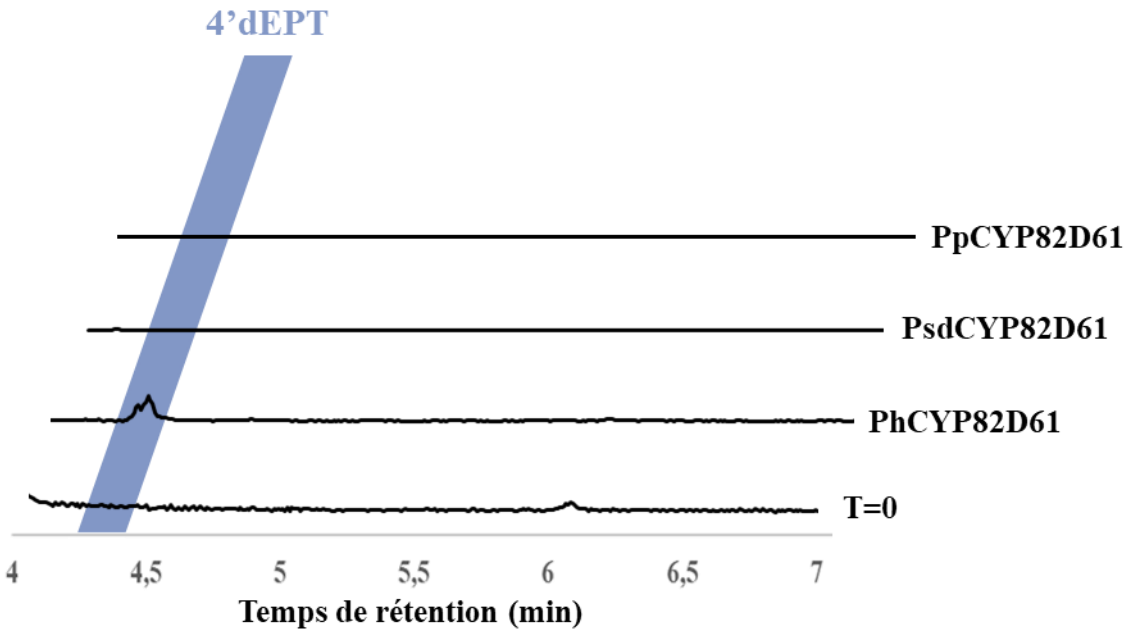
<u>Figure 64</u> : Comparaison des structures protéiques de PhCYP82D61 et PsdCYP82D61.

(A) Prédiction linéaire des feuillets β et des hélices α au niveau du DTM des protéines.
(B) Prédiction de la structure tridimensionnelle des protéines.

ont été établis par l'analyse en UPLC-MS/MS d'extraits métaboliques réalisés sur les feuilles et les rhizomes des trois espèces. Comme nous l'avons observé précédemment, les lignanes de type 4'dEPT ou PPT se fractionnent préférentiellement en MS/MS sur le C-7' (perte du groupement aromatique) et le C-7 (perte du groupement hydroxyle). Ainsi, le suivi sur les chromatogrammes de masse de l'ions fils au m/z 229 est un indicateur des lignanes post-DPT hydroxylés en C-7 (le groupement hydroxyle en C-7 n'étant pas clivé). La **Figure 63** représente les quantités cumulées de ces composés dans les extraits des trois plantes. Seuls trois dérivés de DPT hydroxylés en C-7 ont été détectés et identifiés comme étant la 4'dEPT, la PPT, et leur versions glycosylées (4'dEPT-Gluc et PPT-Gluc). Les résultats obtenus montrent que les trois espèces accumulent ces dérivés de façon abondante mais avec des différences notables sur le type de dérivés hydroxylés en C-7 et leur localisation dans les organes de la plante. Ces résultats indiquent que ces espèces disposent toutes d'une activité de conversion de 4'dDPT en 4'dEPT efficace. Les gènes codant PsdCYP82D61 et PpCYP82D61 ont donc été recherchés. Ne disposant pas de données transcriptomiques ou génomiques de *P. spotty dotty* et *P. peltatum*, les séquences de ces orthologues ont été amplifiées à l'aide des amorces permettant l'amplification du gène codant PhCYP82D61. Les séquences codantes ainsi obtenues ont été clonées dans le vecteur d'expression pESC-His et séquencées via la méthode Sanger.

Les séquences codantes de PsdCYP82D61 et PpCYP82D61 disposent de 96% d'identité nucléotidique avec celle de PhCYP82D61 et les trois protéines ont 94% d'identité entre elles. Parmi les différences notables, PsdCYP82D61 et PpCYP82D61 ont une proline délétée au niveau de l'acide aminé 19 par rapport à PhCYP82D61. Ce résidu est situé au centre de l'hélice α du domaine transmembranaire prédit pour l'ancrage à la membrane du RE (**Figure 64.A**). Cette différence peut entraîner des divergences structurales du domaine transmembranaire entre les orthologues, comme prédit par les modélisations de la **Figure 64.B**, et impacter la stabilité ou la localisation des enzymes. De plus, PpCYP82D61 dispose d'une proline supplémentaire par rapport aux séquences protéiques des CYP82D61 de *P. hexandrum* et *P. spotty sotty* localisée au niveau d'un feuillet β situé dans le site actif de la protéine. Ces différences laissent à penser que PhCYP82D61 et ses orthologues chez *P. spotty sotty* et *P. peltatum* peuvent avoir une stabilité et/ou une efficacité catalytique différente, justifiant, de fait, l'intérêt de comparer leurs activités.

Les vecteurs d'expression pESC-His contenant les séquences codantes des gènes PhCYP82D61, PsdCYP82D61 ou PpCYP82D61 ont été transformés dans la souche de levure WAT11-2ODD-CYP71BE54. Des tests de bioconversion ont été réalisés sur les souches



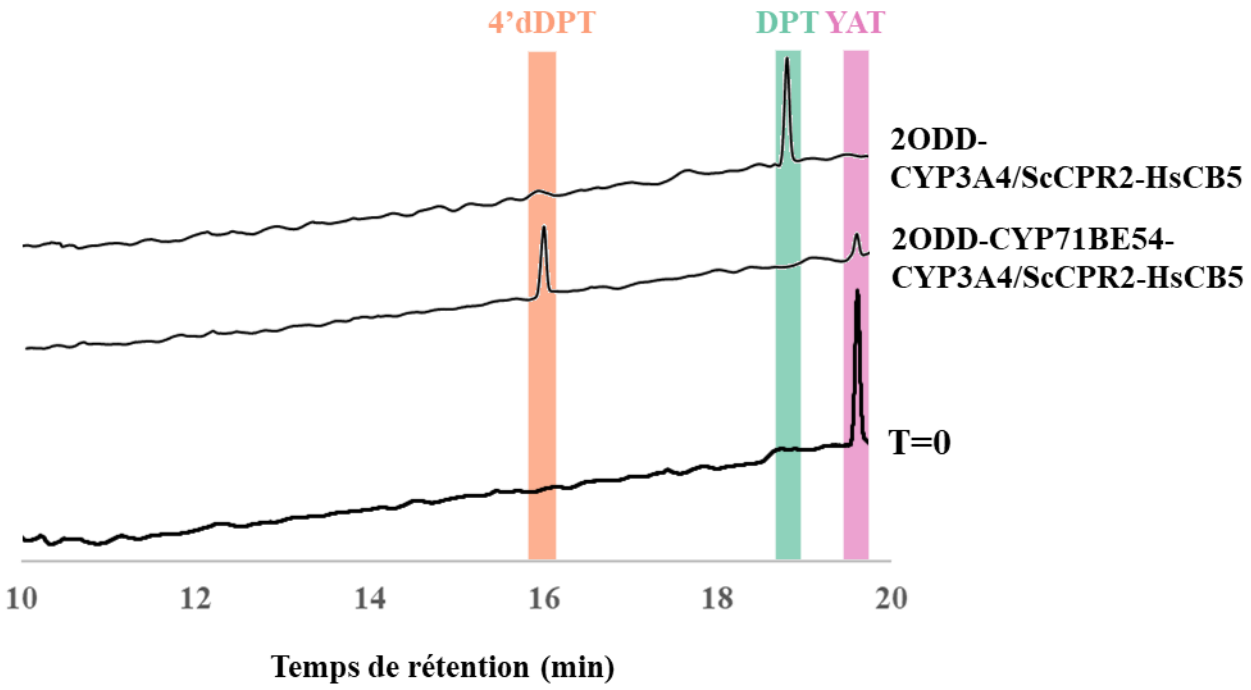
<u>Figure 65</u> : Comparaison des activités catalytiques des orthologues de CYP82D61 en levure.

Single Ion Monitoring (SIM) en ionisation positive m/z 401 issus du surnageant de culture de levures WAT11 contenant AtATR1 transformées (2ODD, CYP71BE54, PhCYP82D61/PsdCYP82D61 ou PpCYP82D61) en présence de YAT à T=0 ou après 72H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

obtenues en présence de 200µM de YAT. Les chromatogrammes de masse au m/z 401 en ionisation positive (ion moléculaire protoné de la 4'dEPT) obtenus à partir des extraits de culture des différentes souches présentés **Figure 65** montrent que seul le chromatogramme issu de la souche contenant la CYP82D61 de *P. hexandrum* dispose d'un pic à 11,1 min correspondant à la 4'dEPT, suggérant qu'aucune 4'dEPT n'a été produite dans le milieu de culture des autres souches. Ces résultats indiquent que les orthologues de CYP82D61 issues de *P. spotty dotty* et *P. peltatum* ne s'accumulent pas dans la cellule ou ne sont pas capables de catalyser la bioconversion de 4'dDPT en 4'dEPT en levure, même après 72 heures de culture.

c) Exploitation de la CYP3A4 humaine

L'absence d'amélioration en levure de la bioconversion de la 4'dDPT en 4'dEPT par l'isoforme DN717 de CYP82D61 ou ses orthologues issues d'autres espèces de Podophyllum nous a conduit à orienter nos recherches vers des cytochromes P450 alternatifs à PhCYP82D61. Une activité d'hydroxylation en C-7 de la DPT a été rapportée pour le cytochrome P450 3A4 humaine (CYP3A4) (Vasilev et al., 2006). En effet, CYP3A4 est la principale enzyme hépatique humaine qui assure à elle seule près de 30% du catabolisme des xénobiotiques par le foie. Sa grande poche catalytique lui permet d'oxyder de façon peu spécifique une grande diversité de substrats, tels que la cyclosporine, les statines, les taxanes, les antibiotiques macrolides, mais aussi les lignanes (Julsing et al., 2008). Ainsi, il était envisageable de remplacer CYP82D61 par CYP3A4 dans notre stratégie de bioconversion, en amont ou en aval de CYP71BE54. Il a été démontré, par ailleurs, qu'une version active de CYP3A4 pouvait être produite chez *S. cerevisiae* (Hayashi et al., 2000 ; Liao et al., 2006). Cependant, dans le système levure, des problèmes de stabilité de l'enzyme liés à une dégradation protéique importante nécessitent de fusionner CYP3A4 à la CPR de la levure ScCPR2 sous forme d'une protéine soluble (Sakaki et al., 1994). De plus, la co-expression de la fusion entre CYP3A4 et ScCPR2 avec le cytochrome b5 humain (HsCB5) a montré une augmentation significative de l'activité de CYP3A4 en levure (Hayashi et al., 2000). Fort de ces connaissances, nous avons décidé de construire deux vecteurs d'expression, l'un dans pESC-His, contenant la séquence codante de CYP3A4 fusionnée à la séquence codante de ScCPR2, l'autre dans pESC-Leu contenant la séquence codante de HsCB5. Ne connaissant pas les spécificités de substrat de CYP3A4 pour la DPT et la 4'dDPT, les plasmides obtenus ont été co-transformés avec le plasmide pESC-Trp-



<u>Figure 66</u> : Alternative d'hydroxylation de la DPT et de la 4'dDPT par CYP3A4.

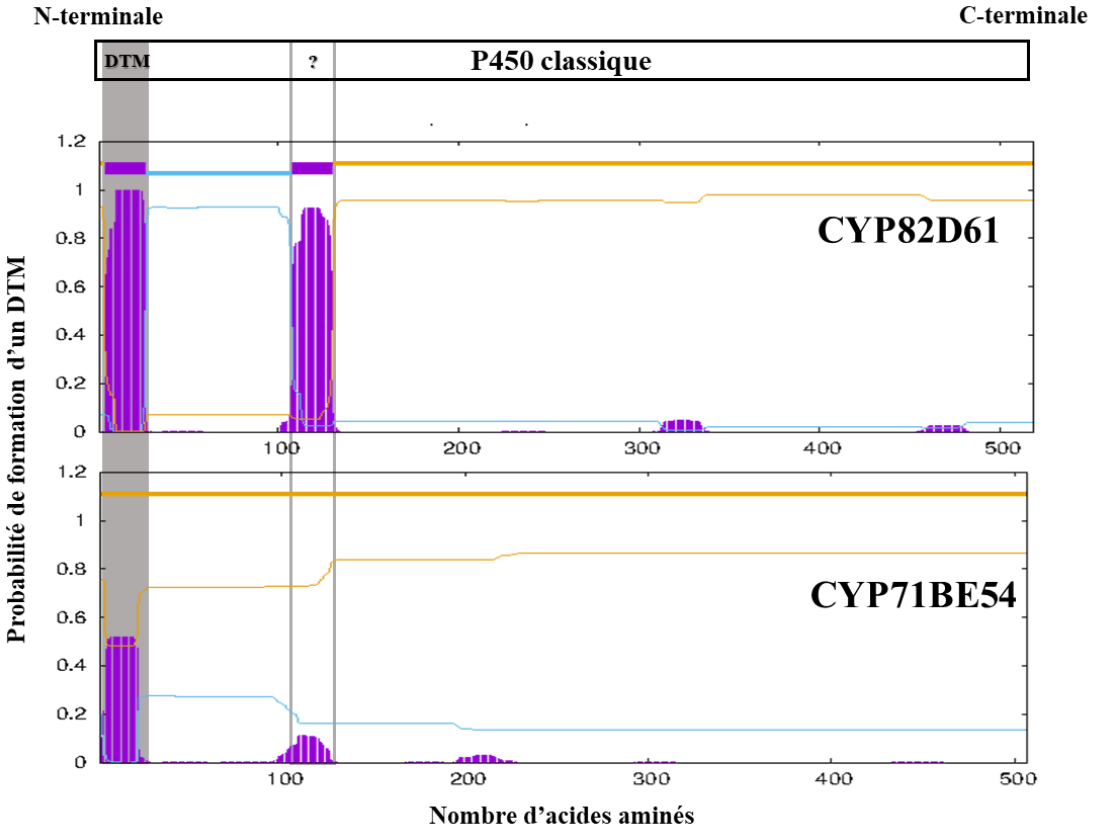
Chromatogramme UV (290 nm) du surnageant de cultures de levures WAT11-2ODD-CYP3A4/ScCPR2-HsCB5 et WAT11-2ODD-CYP71BE54-CYP3A4/ScCPR2-HsCB5 en présence de YAT à T=0 ou après 48H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

2ODD ou pESC-Trp-2ODD-CYP71BE54 dans la souche de levure WAT11, puis les levures ont été cultivées en présence de 200µM de YAT. Le chromatogramme UV obtenu à partir de l'analyse du surnageant de culture de la souche WAT11-2ODD-CYP3A4/ScCPR2-HsCB5 montre que le pic correspondant à la YAT a totalement disparu et que cette disparition est couplée à l'apparition d'un pic correspondant à la DPT, suggérant que la 2ODD a entièrement converti la YAT disponible (**Figure 66**). En revanche, aucun autre pic n'apparaît, CYP3A4 ne semble donc pas capable d'hydroxyler la DPT en levure et dans nos conditions de tests. De plus, sur le chromatogramme UV obtenu à partir de la culture de la souche WAT11-2ODD-CYP71BE54-CYP3A4/ScCPR2-HsCB5, la diminution de taille du pic correspondant à la YAT est corrélée à l'apparition d'un pic correspondant à la 4'dDPT (le pic correspondant à la DPT n'est pas visible). Ces résultats suggèrent que la 2ODD et CYP71BE54 ont converti la quasi-totalité de la YAT en 4'dDPT. Comme précédemment, aucun autre pic n'apparaît sur le chromatogramme UV, confirmant l'incapacité de CYP3A4 à métaboliser la 4'dDPT en levure. Ainsi, même si CYP3A4 est active sur une grande diversité de substrats et qu'elle a été rapportée pour hydroxyler la DPT, son incapacité à fonctionner efficacement dans nos conditions/systèmes n'en fait pas une alternative intéressante pour améliorer les rendements de bioproduction de 4'dEPT.

Sur la base de ces résultats, il apparaît donc que CYP82D61 est la seule enzyme ayant permis une bioconversion de 4'dDPT en 4'dEPT en levure parmi celles testées. De ce fait, nous avons décidé, pour la suite de nos travaux, de nous focaliser sur celle-ci en essayant d'améliorer son efficacité.

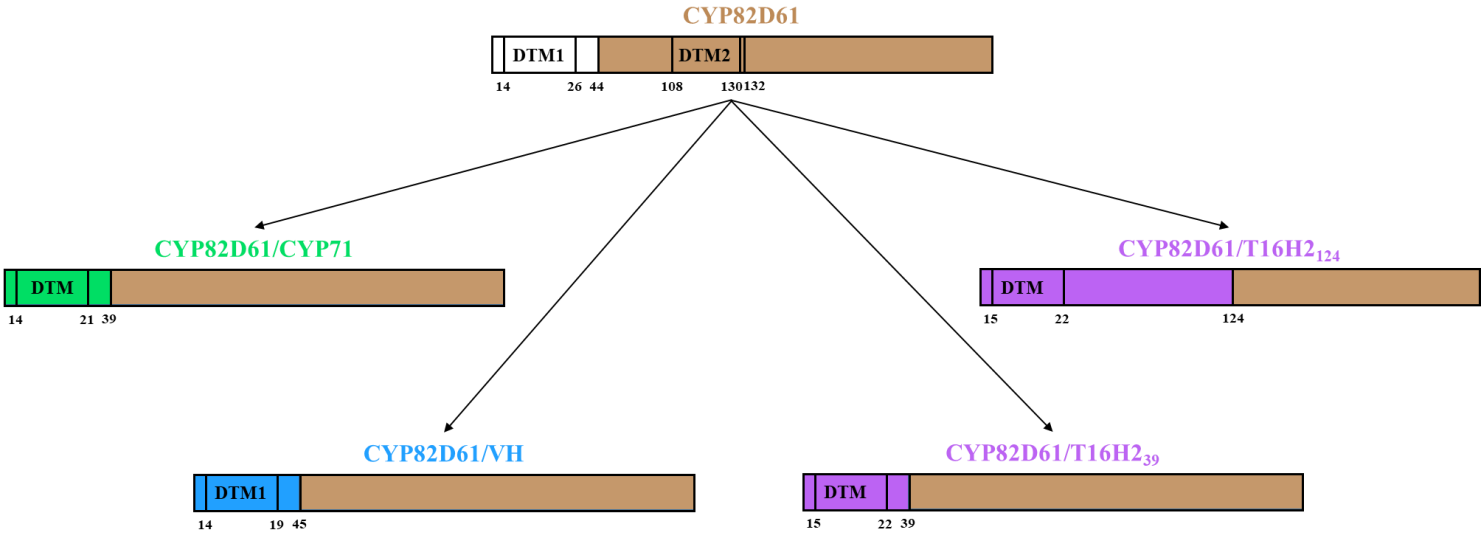
2.1.4 Optimisation de la réduction de CYP82D61 dans la levure

N'ayant pas trouvé d'alternative enzymatique efficace pour remplacer CYP82D61, nous avons orienté nos recherches vers une amélioration de son efficacité catalytique dans la levure. Les cytochromes P450 sont des protéines ancrées à la surface du RE par un domaine transmembranaire (DTM). Cette localisation est importante car l'activité catalytique des P450 dépend de leur réduction par des CPR membranaires. L'expression d'un P450 dans un système hétérologue peut générer un défaut d'ancrage au RE et/ou un défaut d'interaction avec les CPR (Cabello-Hurtado et al., 1999). A partir de ce postulat, trois stratégies ont été mises en œuvre pour s'affranchir d'un potentiel problème de réduction de CYP82D61 dans la levure : la



<u>Figure 67</u> : Prévision des DTM de CYP82D61 et CYP71BE54 par le logiciel TMHMM.

Le logiciel prédictif *TransMembrane based Hidden Markov Model* (TMHMM) est capable d'identifier les zones pouvant correspondre à des DTM avec un degré élevé de précision (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0).



<u>Figure 68</u> : Représentation schématique des protéines chimériques construites à partir de CYP82D61.

L'extrémité N-terminale des protéines est représentée à gauche et la C-terminale à droite.

substitution du domaine transmembranaire de CYP82D61 avec celui de cytochromes P450 actifs en levure, la fusion de CYP82D61 à une CPR d'*A. thaliana* ou de *S. cerevisiae* et le criblage de CPR végétales d'espèces différentes pour réduire efficacement CYP82D61 dans la levure.

a) Substitution du domaine transmembranaire d'ancrage à la membrane du réticulum endoplasmique

La grande majorité des P450 s'ancrent à la membrane du RE via leur DTM, une hélice α située en N-terminale de la protéine. Il est possible que le défaut d'activité de CYP82D61 observé dans les levures soit lié à une singularité du DTM de l'enzyme qui perturbe son ancrage à la membrane du RE de la levure. De plus, la prédiction du DTM de CYP82D61 fait ressortir en fait deux séquences DTM proches du côté N-terminale au lieu d'une seule habituellement. Le premier se situe, de façon classique, entre les résidus 14 et 26 de la protéine, mais un second se situerait entre les résidus 108 à 130 (**Figure 67**). Afin de résoudre un potentiel défaut d'ancrage à la membrane du RE de CYP82D61 et, par conséquent, un défaut d'interaction avec les CPR, son premier DTM a été substitué par celui de trois cytochromes P450 dont l'ancrage au RE de la levure a été confirmé via leur fort niveau d'activité en levure : CYP71BE54 de *P. hexandrum*, la tabersonine 16 Hydroxylase 2 (T16H2) de *C. roseus* (Besseau et al., 2013) et la Vinorine Hydroxylase (VH) de *R. serpentina* (Dang et al., 2017). De plus, en raison de la présence de deux domaines transmembranaires potentiels dans la structure de CYP82D61, les deux DTM de la protéine ont été substitués par la séquence correspondante de T16H2 (**Figure 68**).

La fusion des protéines a été réalisée par hybridation des fragments amplifiés par PCR. Les séquences codantes des gènes obtenues ont chacune été clonée dans le vecteur d'expression pESC-Leu et les vecteurs obtenus ont été transformés dans la souche de levure WAT11-2ODD-CYP71BE54. Des tests de bioconversion ont été réalisés sur les souches obtenues cultivées 48 heures en présence de 200µM de YAT. Le chromatogramme de masse au m/z 401 en ionisation positive obtenu avec le surnageant de la souche de levure contenant la version sauvage de CYP82D61 est utilisé comme contrôle. On y retrouve bien l'apparition d'une faible quantité de 4'dEPT à 11,1 min (**Figure 69**). En comparaison, aucune trace de 4'dEPT n'est détectée avec

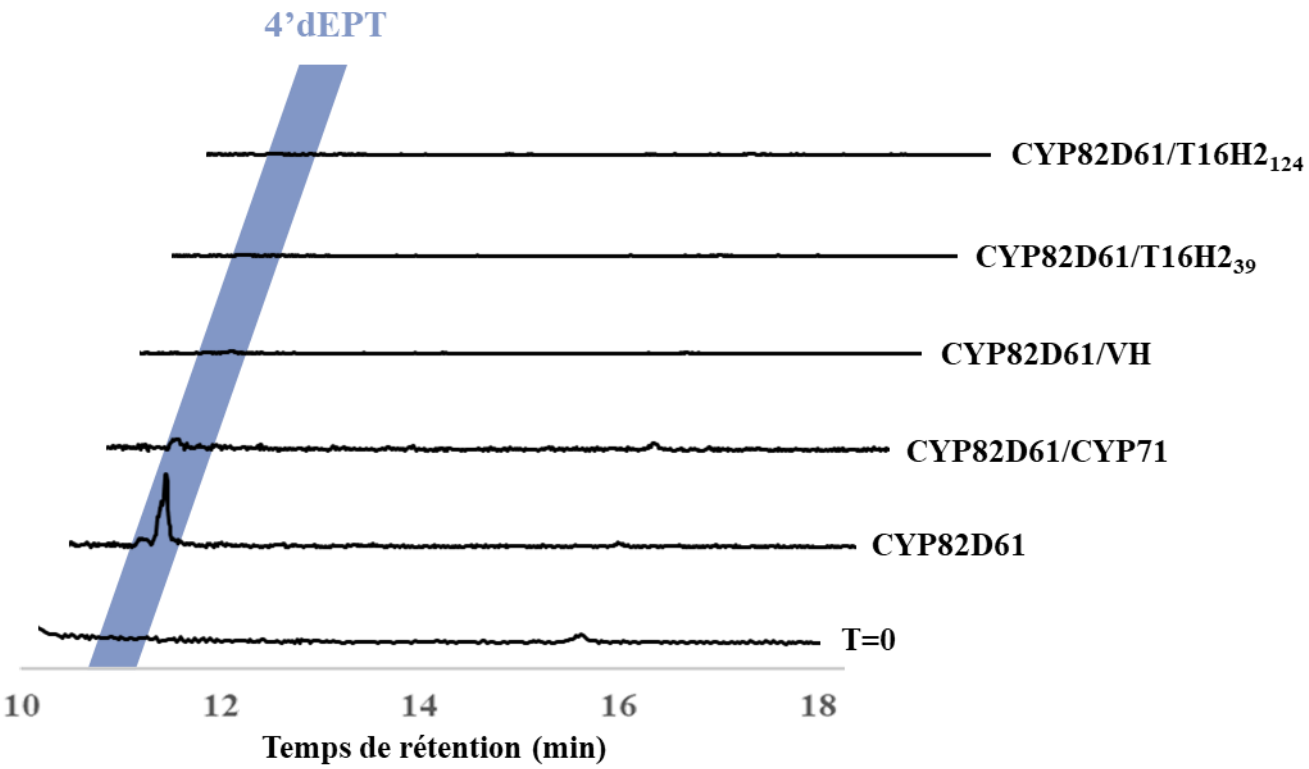


Figure 69 : Effet de la substitution du DTM de CYP82D61 sur son activité catalytique.

Single Ion Monitoring (SIM) en ionisation positive m/z 401 issus du surnageant de culture de levures WAT11 contenant AtATR1 transformées (2ODD, CYP71BE54, CYP82D61 chimériques) en présence de YAT à T=0 ou après 48H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

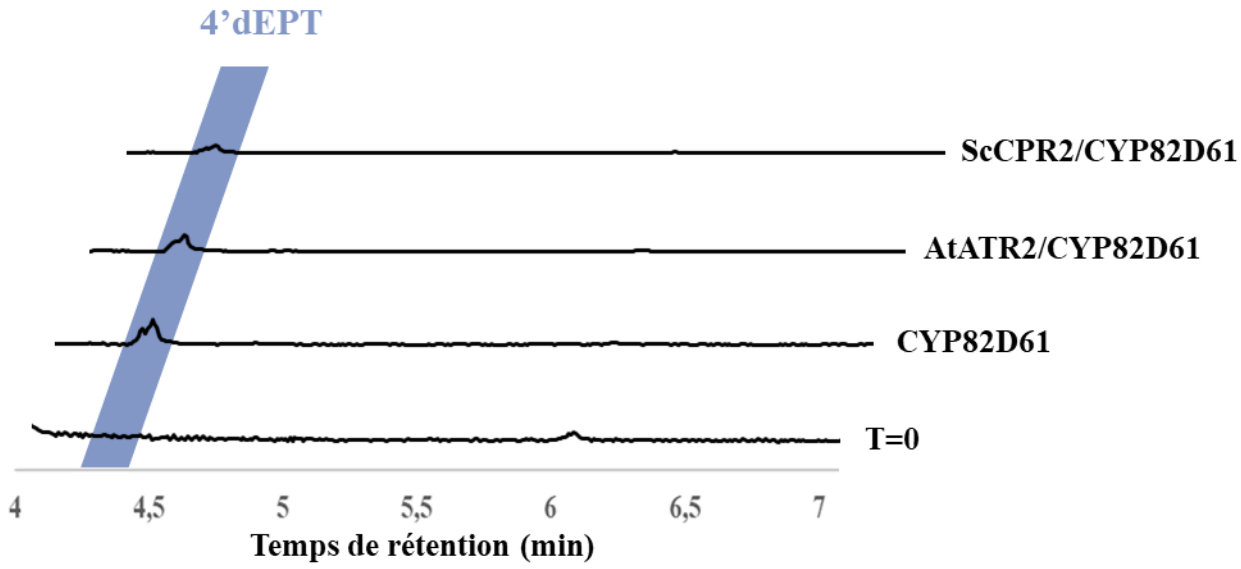


Figure 70 : Activité enzymatique des CYP82D61 fusionnées à la CPR.

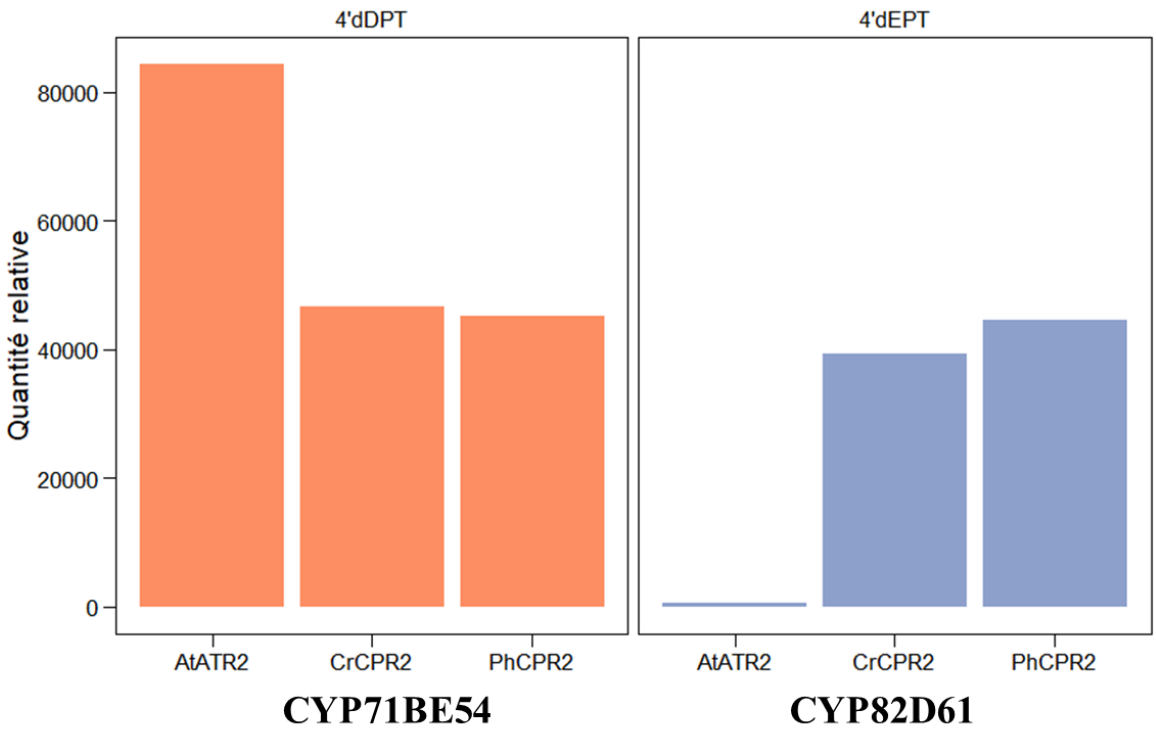
Single Ion Monitoring (SIM) en ionisation positive m/z 401 issus du surnageant de culture de levures WAT11 transformées (2ODD, CYP71BE54, CYP82D61, AtATR2/CYP82D61 ou ScCPR2/CYP82D61) en présence de YAT à T=0 ou après 48H de culture. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

les constructions de CYP82D61 chimériques. Non seulement, les changements de DTM n'ont pas permis d'améliorer l'efficacité catalytique de CYP82D61, mais ils ont, de plus, entraîné la perte totale d'activité de l'enzyme. A ce stade, il est difficile d'expliquer ce résultat, mais la modification de l'extrémité N-terminale a pu perturber le repliement de la protéine ou même altérer son expression. Cette approche non concluante a donc été mise de côté au profit de la réalisation d'une protéine de fusion CYP82D61/CPR.

b) Fusion de CYP82D61 à une CPR

Ne réussissant pas à améliorer la réduction de CYP82D61 par la CPR en optimisant son ancrage à la membrane du RE, nous avons réalisé des protéines de fusion à partir des deux partenaires. Ce type de construction est couramment utilisée en ingénierie métabolique chez les bactéries lorsqu'un cytochrome P450 doit y être transféré. En effet, l'absence de système membranaire interne chez les bactéries nécessite l'utilisation de versions solubles, dépourvues du DTM des cytochromes P450 et des CPR. Cependant, l'absence d'ancrage à une membrane diminue les probabilités de rencontre des deux protéines, d'où la nécessité de les rapprocher par l'intermédiaire d'une protéine de fusion (Sakaki et al., 1994). Cette stratégie a, par ailleurs, déjà été utilisée en ingénierie métabolique chez la levure pour augmenter l'efficacité catalytique de CYP3A4 (Liao et al., 2006).

PhCYP8261 a donc été fusionnée aux CPR d'*A. thaliana* AtATR2 et de *S. cerevisiae* ScCPR2. La séquence complète de CYP82D61 a été conservée pour éviter les pertes d'activité observées précédemment lorsque le DTM de la protéine est modifié et les CPR ont été ajoutées en C-terminale, sans leur DTM. Les séquences codantes obtenues ont été clonées dans le plasmide d'expression pESC-His, puis transformées dans la souche de levure WAT11-2ODD-CYP71BE54. Les essais de bioconversion ont été réalisés avec 200µM de YAT ajoutés au milieu de culture des levures. Les chromatogrammes de masse au m/z 401 en ionisation positive obtenus à partir des milieux de culture des différentes souches cultivées 48 heures sont représentés **Figure 70**. Un pic à 4,3 min de rétention correspondant à la 4'dEPT est bien présent sur le profil métabolique du milieu de culture des levures exprimant la version native de CYP82D61 à 48 heures en comparaison du T0. Ce pic de 4'dEPT est retrouvé de façon équivalente pour les souches exprimant les protéines de fusion. La fusion de CYP82D61 aux CPR n'a donc pas altéré son activité, mais ne l'a pas pour autant améliorée. Ces résultats laissent



<u>Figure 71</u> : Impact de l'origine des CPR co-exprimées sur l'efficacité catalytique de CYP71BE54 et CYP82D61 en levure.

Quantités relatives obtenues par analyses UPLC-PDA-MS des surnageants de cultures de levures WAT303 transformées (CYP71BE54, CYP82D61, AtATR2 ou PhCPR2 ou CrCPR2) en présence de DPT après 48H de culture. Gradient linéaire de 20 minutes.

donc à penser qu'il n'y a pas de défaut d'ancrage de CYP82D61 à la membrane du RE de la levure, ou encore d'une incapacité à interagir avec les CPR. Cela suggère, en retour, que les CPR utilisées jusqu'à présent dans nos souches de levures ne parviennent peut-être pas à réduire correctement CYP82D61.

c) Criblage de CPR végétales pour leur efficacité de réduction de CYP82D61

Les résultats précédents semblent indiquer qu'il n'y a sans doute pas de défaut d'interaction entre CYP82D61 et les CPR dans la levure, mais plutôt une faible capacité de réduction de CYP82D61 par AtATR2. En effet, il existe parfois des spécificités de réduction de P450 par certaines CPR, la faible efficacité de réduction des P450 végétaux par ScCPR2 par rapport à AtATR1 ou AtATR2 en levure en est une preuve (Parage et al., 2016). Il est donc probable que CYP82D61 ne soit pas efficacement réduite par AtATR1 ou AtATR2, et qu'elle dispose d'une spécificité avec d'autres CPR végétales. Afin de vérifier cette hypothèse, les séquences codantes de trois CPR issues de trois organismes différents ont été clonées dans le vecteur d'expression pESC-Trp : AtATR2 d'*A. thaliana*, PhCPR2 de *P. hexandrum* et CrCPR2 de *C. roseus*. Le choix de *P. hexandrum* s'explique aisément par la correspondance de l'espèce entre CYP82D61 et PhCPR2, laissant supposer une correspondance plus grande entre les deux protéines. En parallèle, la CrCPR2 de *C. roseus* a également été testée en raison des succès obtenus avec elle pour la conversion de tabersonine en vindoline en levure (Kulagina et al., 2021c). Chacun des vecteurs obtenus a été co-transformé avec les plasmides pESC-Ura-CYP71BE54 et pESC-His-CYP82D61, contenant les séquences codant CYP71BE54 et CYP82D61 respectivement, dans la souche de levure WAT303. Après 48 heures de culture des souches en présence de 200µM de DPT, les quantités relatives de 4'dDPT et 4'dEPT du milieu ont été mesurées à partir des données d'analyses en UPLC (**Figure 71**). Ces résultats montrent, dans un premier temps, que de la 4'dDPT est formée à partir de la DPT dans les trois souches utilisées, suggérant que l'activité catalytique de CYP71BE54 n'est pas altérée si celle-ci est réduite par AtATR2, PhCPR2 ou CrCPR2. Ainsi, de la 4'dDPT est bien disponible pour mesurer l'activité de CYP82D61. Les quantités relatives de 4'dEPT dans les surnageants de cultures des souches transformées avec les gènes codant CrCPR2 et PhCPR2 sont nettement plus élevées que dans le surnageant de culture de la souche transformée avec le gène codant

AtATR2 (62 et 71 fois plus respectivement), indiquant que CYP82D61 est beaucoup plus active sur la 4'dDPT lorsqu'elle est réduite par PhCPR2 ou CrCPR2.

Le verrou principal de notre stratégie de bioconversion en levure a ainsi pu être levé. La faible activité catalytique de CYP82D61 en levure était due à un défaut de réduction de l'enzyme par les CPR d'*A. thaliana* dans nos conditions de test. Bien que les niveaux d'expression relative des différentes CPR n'aient pas été analysés par Western blot, il apparaît que l'activité catalytique de CYP82D61 est meilleure lorsque celle-ci est co-exprimée et donc potentiellement réduite par PhCPR2 ou CrCPR2. En outre, l'activité de CYP71BE1 n'étant pas altérée par l'utilisation de ces CPR, le clonage de PhCPR2 ou CrCPR2 à la place d'AtATR1 ou d'AtATR2 dans la levure permet d'améliorer nettement les rendements de bioconversion de 4'dDPT en 4'dEPT.

2.2 Conclusion

Dans un premier temps, les essais réalisés via l'emploi de plasmides auto-réplicatifs dans les souches de levure WAT303 ou WAT11 nous ont permis de vérifier, la bénignité, dans les conditions testées, des dérivés de lignanes pour la levure, assurant ainsi un flux des substrats et des produits dans les cellules où sont situées les enzymes hétérologues impliquées dans la voie de conversion de la 4'dEPT à partir de la YAT. Les premiers essais enzymatiques ont établi une première preuve de concept de notre stratégie de bioconversion pour la production d'un précurseur d'étoposide. En effet, la mesure de l'activité des enzymes en levure, étape par étape, a démontré que la 2ODD et CYP71BE54 de *P. hexandrum* sont suffisamment actives pour envisager une bioconversion à plus grande échelle. Par contre, l'activité de CYP82D61 étant très insuffisante, plusieurs stratégies ont dû être mises en place pour optimiser les rendements de conversion de la 4'dDPT en 4'dEPT. En effet, une première analyse par Western Blot a permis de vérifier que la protéine s'accumulait correctement dans les cellules. De ce fait, la faible activité de CYP82D61 sur la 4'dDPT en levure, comparée à sa forte accumulation dans les cellules nous a laissé penser, dans un premier temps, que la conversion de 4'dDPT en 4'dEPT pouvait ne pas être l'activité principale de CYP82D61 *in planta* et qu'une autre enzyme pourrait être impliquée dans cette réaction chez *P. hexandrum*. Nous avons, de ce fait, vérifié l'activité catalytique de DN717, l'isoforme la plus proche de CYP82D61 ayant été mise en évidence par analyses transcriptomiques. Toutefois, aucune activité n'a pu être mesurée pour

DN717, laissant supposer qu'elle ne catalyse pas cette réaction *in planta*. Bien que d'autres enzymes, moins identiques à CYP82D61, puissent catalyser cette réaction, nous avons choisi de nous orienter vers la recherche d'orthologues de CYP82D61 issues d'autres espèces de Podophyllum accumulatrices de 4'dEPT. Celles-ci ont ainsi été clonées pour tester leur activité catalytique et leur utilisation potentielle pour augmenter les rendements de bioconversion. Toutefois, malgré les différences structurales des orthologues par rapport à CYP82D61 de *P. hexandrum*, notamment au niveau de leurs sites catalytiques, la conversion de 4'dDPT en 4'dEPT n'a pas été améliorée dans les souches de levures construites. En outre, sachant que le cytochrome P450 3A4 humain (CYP3A4) est actif sur un grand nombre de substrats, et notamment sur la DPT, nous avons décidé de cloner ce P450 avec les gènes codant 2ODD seule ou avec CYP71BE54 en levure pour vérifier sa capacité à métaboliser la DPT et la 4'dDPT. Les problèmes liés à la dégradation de cette enzyme en levure nous ont orienté vers une fusion de CYP3A4 à la CPR de *S. cerevisiae* ScCPR2 (Liao et al., 2006). Cependant, les résultats obtenus dans nos conditions de travail montrent que CYP3A4 n'est capable de catalyser, ni la DPT, ni la 4'dDPT. Ainsi, la recherche d'alternatives à l'utilisation de CYP82D61 pour la conversion de 4'dDPT en 4'dEPT en levure n'ayant pas été fructueuse, nous nous sommes focalisés sur l'optimisation des rendements de bioconversion avec cette enzyme. Pour cela, nous avons, dans un premier temps, vérifié qu'il n'y avait pas de défaut d'ancrage de CYP82D61 à la membrane du RE de la levure en remplaçant le domaine transmembranaire d'ancrage par celui de P450 dont l'ancrage au RE des levures a été établi. De plus, pour pallier un éventuel défaut d'ancrage de CYP82D61 au RE empêchant sa réduction par les CPR, nous avons facilité le rapprochement CYP82D61/CPR en créant des protéines de fusion avec la CPR ScCPR2 de *S. cervevisiae* ou AtATR2 d'*A. thaliana*. Globalement, les résultats obtenus ont confirmé que le défaut d'activité de CYP82D61 n'était pas causé par un défaut d'ancrage à la membrane du RE de la levure. Notre dernière hypothèse portait sur une défaillance de réduction de CYP82D61 par les CPR d'*A. thaliana*. Pour l'étayer, les gènes codant la 2ODD, CYP71BE54 et CYP82D61 de *P. hexandrum* ont été exprimés dans une souche de levure avec les gènes codant PhCPR2 de *P. hexandrum* ou CrCPR2 de *C. roseus*. Les tests réalisés dans ces conditions ont révélé une augmentation drastique des quantités de 4'dEPT dans le surnageant de culture de ces souches, nous permettant de lever le verrou à notre stratégie de bioconversion reposant sur la réduction de CYP82D61.

Au final, ces travaux, basés sur le transfert de la voie de conversion de la YAT en 4'dEPT par utilisation de plasmides auto-réplicatifs, ont permis d'établir la preuve de concept

Tableau 2 : Liste des amorces utilisées pour l'amplification des gènes d'intérêt.

Nom	Séquence	Site de restriction	Fonction
Ph2ODD_fw	CTGAGAACTAGTTCGGTCATGGGTTCTACA	SpeI	Amplification 2ODD de *P. hexandrum*
Ph2ODD_rev	CTGAGAACTAGTTCATGCACCTGTGTACGC	SpeI	Amplification 2ODD de *P. hexandrum*
AtATR2_fw	CTGAGAACTAGTTCGGTCATGAAAAACATGATGAATTATAAATTAAAAC	SpeI	Amplification AtATR2 d'*Arabidopsis thaliana*
AtATR2_rev	CTGAGAACTAGTTTAGTGGTGATGGTGATGATGCCATACATCTCTAAGATATCTTCC	SpeI	Amplification AtATR2 d'*Arabidopsis thaliana*
CYP71BE54_fw	CTGAGAACTAGTTCGGTCATGGAGTTCCTT	SpeI	Amplification CYP71BE54 de *P. hexandrum*
CYP71BE54_rev	CTGAGAACTAGTCTATTCTTTTGTAGTCGA	SpeI	Amplification CYP71BE54 de *P. hexandrum*
CYP71BE54_rev_His	CTGAGAACTAGTCTAGTGGTGATGGTGATGATGTTCTTTTGTAGTCGATCTTTCCAGC	SpeI	Amplification CYP71BE54 de *P. hexandrum* avec tag Histidine
CYP82D61_fw	CTGAGATCTAGATCGGTCATGGATTCCCTG	XbaI	Amplification CYP82D61 de *P. hexandrum, P. spotty dotty* et *P. peltatum*
CYP82D61_rev	CTGAGATCTAGACTAGACAAAACATTTTGG	XbaI	Amplification CYP82D61 de *P. hexandrum, P. spotty dotty* et *P. peltatum*
CYP82D61_rev_His	CTGAGATCTAGATTAGTGGTGATGGTGATGATGGACAAAACATTTTGGAGAGATGCG	XbaI	Amplification CYP82D61 de *P. hexandrum* avec tag Histidine
DN717_fw	CTGAGATCTAGAATGTATTCACTGCCTTTCCTAGAAAC	XbaI	Amplification DN717 de *P. hexandrum*
DN717_rev	CTGAGATCTAGACCGACAAAACATTTTGGAGAGATGC	XbaI	Amplification DN717 de *P. hexandrum*
CYP82D61_DTM_fw	CCGATCATCGGCCACCTCCAT	XbaI	Amplification de CYP82D61 sans domaine transmembranaire de *P. hexandrum*
CYP71BE54_DTM_rev	ATGGAGGTGGCCGATGATCGGTAGCTTCCGTGGCCCTGGGG		Amplification du domaine transmembranaire de CYP71BE54 de *P. hexandrum*
VH_DTM_fw	CTGAGATCTAGATCGGTCATGGATCTCCTCCAGATTCTGCT	XbaI	Amplification du domaine transmembranaire de VH de *Rauvolfia serpentina*
VH_DTM_rev	ATGGAGGTGGCCGATGATCGGCCAGGCACCTGCTGGTTCTG		Amplification du domaine transmembranaire de VH de *Rauvolfia serpentina*
T16H2_DTM_fw	CTGAGATCTAGATCGGTCATGGAGTTGTATTATTTTTCCACC	XbaI	Amplification du domaine transmembranaire de T16H2 de *Catharanthus roseus*
T16H2_DTM_rev	ATGGAGGTGGCCGATGATCGGAATTGGGGGAGGCCCCAGAG		Amplification du domaine transmembranaire de T16H2 de *Catharanthus roseus*
AtATR2_Fu_fw	CTGAGAACTAGTCTCGAGCTGTTACCATCTTTTTCGGTACAC	SpeI/XhoI	Amplification AtATR2 d'*Arabidopsis thaliana* pour fusion CPR/P450
AtATR2_Fu_rev	CTGAGAACTAGTTTACCATACATCTCTAAGATATCTTCC	SpeI	Amplification AtATR2 d'*Arabidopsis thaliana* pour fusion CPR/P450
ScCPR2_Fu_fw	CTGAGAACTAGTCTCGAGCTGCTGTCAGCTCGGGCAACA	SpeI/XhoI	Amplification ScCPR de *Saccharomyces cerevisiae* pour fusion CPR/P450
ScCPR2_Fu_rev	CTGAGAACTAGTTTACCAGACATCTTCTTGGTATCTAC	SpeI	Amplification ScCPR de *Saccharomyces cerevisiae* pour fusion CPR/P450
CYP82D61_Fu_fw	CTGAGACTCGAGTCGGTCATGGATTCCCTGCACTGCCT	XhoI	Amplification CYP82D61 de *P. hexandrum* pour fusion CPR/P450
CYP82D61_Fu_rev	CTGAGACTCGAGCGACAAAACATTTTGGAGAGATGC	XhoI	Amplification CYP82D61 de *P. hexandrum* pour fusion CPR/P450
PhCPR2_fw	CTGAGAACTAGTTCGGTCATGCAATCAAGCTCAGCTTTC	SpeI	Amplification PhCPR2 de *P. hexandrum*
PhCPR2_rev	CTGAGAACTAGTTCACCATACGTCACGGAGATATC	SpeI	Amplification PhCPR2 de *P. hexandrum*
CrCPR2_fw	CTGAGAGCTAGCTCGGTCATGGATTCTAGCTCGGAGAAGTTGTC	NheI	Amplification CrCPR2 de *Catharanthus roseus*
CrCPR2_rev	CTGAGAGCTAGCTCAGTGGTGATGGTGATGATGCCAGACATCTCGGAGATACC	NheI	Amplification CrCPR2 de *Catharanthus roseus*
GAL1_fw	ATTTTCGGTTTGTATTACTTC		PCR sur colonie des gènes clonés en BamHI vecteur pESC
CYC1_rev	GTTCTTAATACTAACATAACT		PCR sur colonie des gènes clonés en BamHI vecteur pESC
GAL10_fw	GGTGGTAATGCCATGTAATATG		PCR sur colonie des gènes clonés en SpeI vecteur pESC
ADH1_rev	GGCAAGGTAGACAAGCCGACAAC		PCR sur colonie des gènes clonés en SpeI vecteur pESC

de la production d'un précurseur d'étoposide par bioconversion en levure. Toutefois, l'emploi de ce type de plasmides n'étant pas optimal pour une montée en échelle de la bioproduction de par l'hétérogénéité de bioconversion entre les clones, l'utilisation de marqueurs d'auxotrophie, et l'altération potentielle de croissance des cellules causée par ces marqueurs, les gènes codant la 2ODD, CYP71BE54, CYP82D61 et CrCPR2 ou PhCPR2 doivent désormais être intégrés dans le génome de la levure pour envisager une production à plus grande échelle.

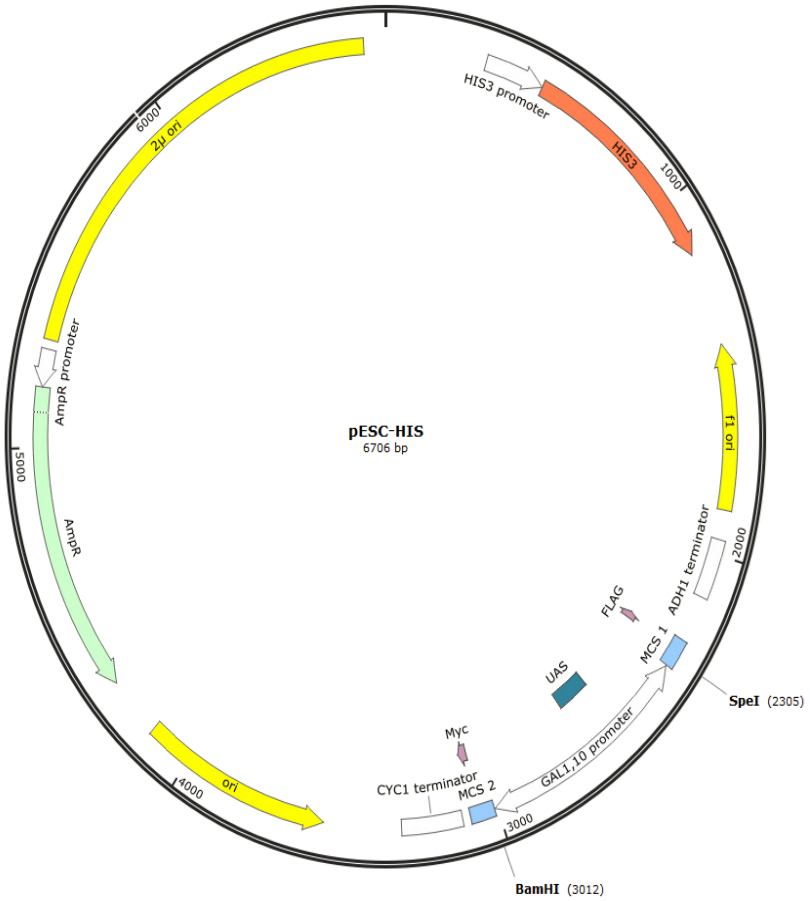
2.3 Matériel et méthodes

2.3.1 Test de toxicité

Afin de vérifier la toxicité du substrat et des produits de la voie de synthèse de la 4'dEPT, des tests ont été réalisés sur le souche de levure WAT303. Pour cela, les cellules ont été mises en culture liquide pendant 16 heures sous agitation (160 tr/min) à 28°C. Une dilution a été réalisée sur les cultures afin d'obtenir une absorbance de 0,1 à 590nm. Les cultures diluées ont ensuite été réparties sur une plaque 96 puits (*Microwell #055260, Dutscher*) (98µL par puits) en présence de 500µM DPT, 500µM PPT ou 500µM de flucytosine, un puissant fongicide. La plaque a ensuite été incubée à 28°C et la croissance des levures a été mesurée à différents temps (0H, 14H, 22H, 36H et 44H) par mesure spectrométrique à 590nm.

2.3.2 Gènes d'intérêt

Les ARN totaux du rhizome de *Podophyllum hexandrum*, du rhizome de *Podophyllum peltatum*, du rhizome de *Podophyllum spotty dotty*, des feuilles d'*Arabidopsis thaliana,* des feuilles de *Catharanthus roseus* et des feuilles de *Rauwolfia serpentina* ont été extraits à l'aide du kit NucleoSpin® RNA Plant and Fungi de Macherey-Nagel. L'ADN complémentaire (ADNc) a été obtenu à partir d'1µ de ces ARN en suivant les instructions de la RevertAid Reverse Transcriptase de Thermofisher. Les gènes d'intérêt ont été amplifiés par PCR (Polymerase Chain Reaction) grâce à l'ADN polymérase Phusion (Phusion™ High-Fidelity, ThermoFisher) à partir de l'ADNc. Les amorces utilisées pour les différentes PCR permettent l'ajout du site d'initiation TCGGTC en amont du codon ATG comme présenté au paragraphe



<u>Figure 72</u> : Exemple de carte d'un plasmides pESC utilisé pour les clonages.

2µ ori: origine de réplication du plasmide chez *S. cerevisiae ;* AmpR et AmpR promoter: cassette de résistance à l'ampiciline ; Ori: origine de réplication haute copies pUC chez *E. coli ;* HIS3 et HIS3 promoter: cassette ; CYC1 terminator: terminateur de transcription pour le gène cloné en BamHI ; GAL1,10 promoter : promoteur de transcription bidirectionnel ; ADH1 terminator: terminateur de transcription pour le gène cloné en SpeI ; f1 ori: origine de réplication du bactériophage f1 ; HIS3 : marqueur d'auxotrophie pour l'histidine ; FLAG et MYC: épitope pour marquage en N-terminal de la protéine.

I.3.c.3 de l'Introduction et contiennent des sites de restriction SpeI, XbaI ou NheI (**Tableau 2**) permettant la digestion par l'enzyme de restriction de l'ORF (Open Reading Frame) (données supplémentaires) et la ligation dans le plasmide intégratif sélectionné grâce à la T4 DNA Ligase de ThermoFisher.

2.3.3 Construction des plasmides

Les plasmides pESC-His, pESC-Leu, pESC-Ura et pESC-Trp permettant une réplication autonome dans la levure des gènes clonés ont été commandés chez Sigma-Aldrich **(Figure 72)**. Ce sont des vecteurs navette, c'est-à-dire qu'ils sont capables de se maintenir aussi bien dans les bactéries que dans les levures. En effet, l'origine de réplication du plasmide pUC (ori) et la cassette de résistance à l'ampicilline permettent la réplication du plasmide et la sélection chez *Escherichia coli* et l'origine de réplication 2µ et le marqueur métabolique permettent la réplication autonome et la sélection du plasmide chez *Saccharomyces cerevisiae*. Chacun de ces plasmides possède un marqueur métabolique différent (histidine, leucine, uracile et tryptophane respectivement) permettant l'expression de plusieurs plasmides différents dans la levure et une sélection simplifiée. Ils contiennent les couples promoteur/terminateur de transcription pGAL10/tADH1 et pGAL1/tCYC1 de la levure qui permettent le clonage d'un à deux gène(s) par plasmide. Les promoteurs pGAL10 et pGAL1 sont inductibles au galactose. Enfin, les épitopes FLAG et MYC ont été ajoutés en amont et en aval des sites d'insertion afin de pouvoir ajouter un tag en C-terminale des protéines clonées. En effet, si les gènes clonés ne disposent pas de codon STOP, les protéines exprimées seront sous forme taguée, s'ils disposent d'un codon STOP, les protéines exprimées seront sous forme native.

La souche bactérienne *E. coli* DH5α (Invitrogen) a été utilisée comme intermédiaire de clonage. Elle dispose du génotype F- Φ80*lac*ZΔM15 Δ (*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(rk-,mk+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ-. Les mutations situées dans le gène de la recombinase A1 (*recA1*) et celui de l'endonucléase A1 (*endA1*) diminuent les phénomènes de recombinaison et de dégradation de l'ADN respectivement, permettant une meilleure stabilité des plasmides transformés. Ces cellules ont été transformées par choc thermique avec le produit de ligation et cultivées dans du milieu Luria-Bertani (LB) additionné de 100µg/mL d'ampicilline à 37°C pour sélectionner les transformants. Une PCR sur colonie a été réalisée sur les clones ayant poussé à l'aide de l'ADN polymérase DreamTaq (ThermoFisher). Les plasmides ont été extraits des cellules positives à la PCR sur colonies avec le kit NucleoSpin®

Tableau 3 : Liste des différents plasmides construits pour réplication autonome en levure.

Nom du plasmide	Description
pESC-Trp-2ODD	Ph2ODD clonée en SpeI pESC-Trp
pESC-Trp-2ODD-CYP71BE54	Ph2ODD clonée en SpeI et CYP71BE54 en BamHI pESC-Trp
pESC-His-CYP82D61	CYP82D61 clonée en SpeI pESC-His
pESC-Ura-CYP71BE54_tag	CYP71BE54 tagguée clonée en SpeI pESC-Ura pour Western Blot
pESC-His-CYP82D61_tag	CYP82D61 tagguée clonée en SpeI pESC-His pour Western Blot
pESC-His-DN717	DN717 clonée en SpeI pESC-His
pESC-His-PsdCYP82D61	PsdCYP82D61 clonée en SpeI pESC-His
pESC-His-PpCYP82D61	PpCYP82D61 clonée en SpeI pESC-His
pESC-His-ScCPR2/CYP3A4	Fusion ScCPR2/CYP3A4 clonée en SpeI pESC-His
pESC-Leu-HsCB5	HsCB5 clonée en SpeI pESC-Leu
pESC-Leu-CYP82D61	CYP82B61 clonée en SpeI pESC-Leu
pESC-Leu-CYP82D61/CYP71BE54	CYP82B61 avec le DTM de CYP71BE54 clonée en SpeI pESC-Leu
pESC-Leu-CYP82D61/VH	CYP82B61 avec le DTM de VH clonée en SpeI pESC-Leu
pESC-Leu-CYP82D61/T16H2	CYP82B61 avec le DTM de T16H2 clonée en SpeI pESC-Leu
pESC-His-AtATR2/CYP82D61	Fusion AtATR2/CYP82D61 clonée en SpeI pESC-His
pESC-His-ScCPR2/CYP82D61	Fusion ScCPR2/CYP82D61 clonée en SpeI pESC-His
pESC-Ura-CYP71BE54	CYP71BE54 clonée en SpeI pESC-Ura
pESC-Trp-AtATR2	AtATR2 clonée en SpeI pESC-Trp
pESC-Trp-PhCPR2	PhCPR2 clonée en SpeI pESC-Trp
pESC-Trp-CrCPR2	CrCPR2 clonée en SpeI pESC-Trp

Plasmid de Machery-Nagel et séquencés via la méthode Sanger. Les différents plasmides obtenus sont répertoriés dans le **Tableau 3**.

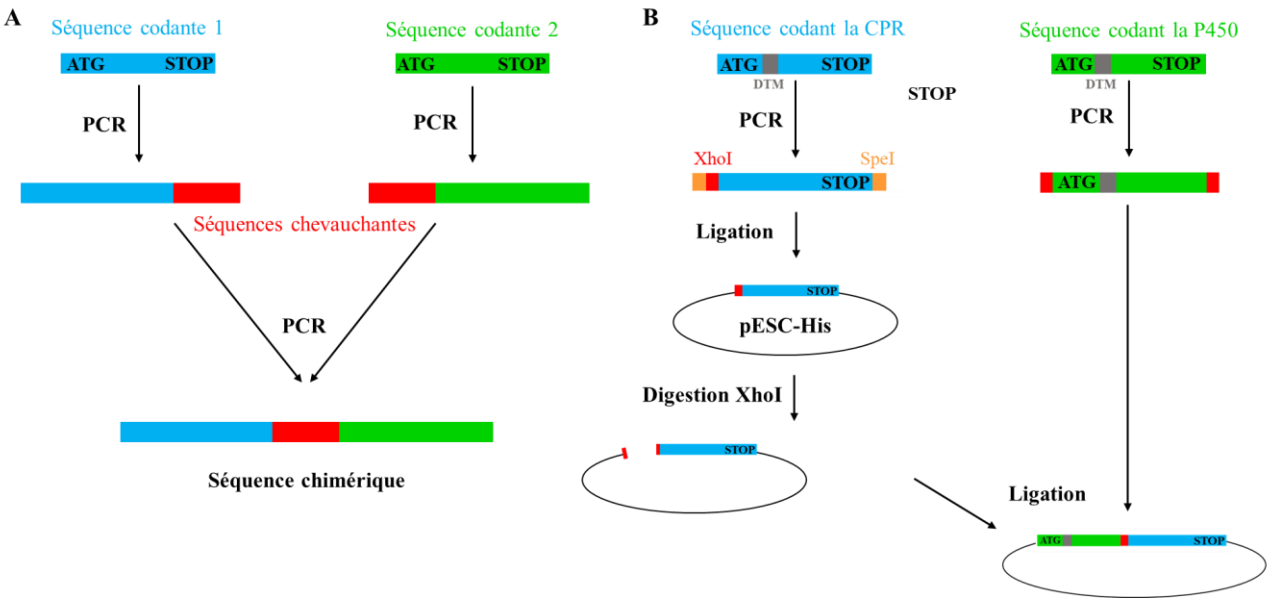
2.3.4 Fusions de fragments d'ADN

Afin de réaliser la substitution du domaine transmembranaire (DTM) de CYP82D61, des amorces ont été élaborées afin d'amplifier par PCR CYP82D61 sans son DTM d'un côté et le DTM souhaité d'un autre. Ces amorces ont été pensées afin qu'une séquence chevauchante soit présente sur chacun des fragments pour permettre une hybridation entre ceux-ci (**Figure 73.A**). Ainsi, deux PCR ont été réalisées avec l'ADN polymérase Phusion (Phusion™ High-Fidelity, ThermoFisher). La première a été réalisée à partir de plasmides disponibles au laboratoire afin d'amplifier séparément les deux séquences d'ADN à assembler. Après purification des deux séquences d'ADN à l'aide du kit NucleoSpin® Gel and PCR Clean-up de Macherey-Nagel, une seconde PCR a été réalisée en utilisant les deux fragments d'ADN comme matrice. Les deux fragments ont ainsi été « fusionnés » par hybridation. Le gène obtenu a pu être digéré par l'enzyme de restriction SpeI et une ligation a permis son insertion dans le vecteur d'intérêt.

Dans un second temps, des cytochromes P450 ont été fusionnées à des CPR. Pour cela, le plasmide pESC-His a d'abord été digéré par l'enzyme XhoI, puis une incubation de 5 minutes à température ambiante en présence de la T4 DNA polymérase (ThermoFisher) a permis l'obtention de bouts francs permettant une re-ligation du vecteur délété du site de restriction XhoI grâce à la T4 DNA ligase (ThermoFisher). Des amorces ont été élaborées afin d'obtenir les CPR et les P450 amplifiées par PCR avec de part et d'autre les sites de restriction SpeI (pour ligation avec le vecteur) et XhoI (pour la fusion P450/CPR) (**Figure 73.B**). La fusion des P450 aux CPR n'a donc pas été réalisée comme précédemment par hybridation mais par une succession de digestions/ligations.

2.3.5 Souches de levure

Les souches utilisées dans cette partie du travail sont les souches haploïdes *Saccharomyces cerevisiae* WAT303 et WAT11. La souche WAT303 dispose du génotype



<u>Figure 73</u> : Schémas des méthodes utilisées dans ce projet pour la fusion entre deux fragments d'ADN.

(A) Fusion par PCR combinatoire à l'aide de séquences chevanchantes.
(B) Fusion par succession de digestions/ligations.

MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) et est couramment utilisée pour l'expression et l'étude de protéines recombinantes. La souche WAT11 dispose du même génotype que WAT303 mais exprime de plus la CPR d'*A. thaliana* AtATR1 (Urban et al., 1997).

2.3.6 Transformation des levures

La préparation des levures électro-compétentes permet aux cellules d'être transformées par électroporation. Pour cela, les souches ont été inoculées dans 10mL de milieu liquide Yeast extract-Peptone-Dextrose (YPD) (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) pendant 16 heures sous agitation (200 tr/min) à 28°C. Puis 5 mL de culture ont été ensemencées dans 50 mL de YPD liquide pendant 5 heures sous agitation modérée (160 tr/min) à 28°C. Après une incubation de 15 minutes dans la glace, la culture a été centrifugée 1 minute à 3000g, 4°C. Le culot a ensuite été délicatement repris dans 20 mL d'AcLiTE-DTT froid (10mM Tris-HCl pH 7,5, 1mM EDTA, 150mM Acétate de lithium, 10mM DTT). Après incubation d'une heure à 28°C, la suspension cellulaire a été centrifugée 1 minute à 3000g, 4°C et le culot a été lavé deux fois avec 20 mL de sorbitol 1M glacé. Finalement, après une dernière centrifugation de 1 minute, 3000g, 4°C, le culot de levure a été remis en suspension dans 3 mL de sorbitol 1M glacé. Les levures électro-compétentes ont été aliquotées et stockées à -80°C.

Pour réaliser les transformations, 200µL de cellules électro-compétentes ont été mélangés à 1µg de plasmide dans la glace. Le mélange a ensuite été transféré dans une cuve à électroporation à espace inter-électrode de 2mm (Molecular BioProducts). Pour réaliser l'électroporation, un courant électrique d'1,5 kV a été appliqué pendant 5 ms sur la cuve à l'aide d'un électroporateur (BioRad MicroPulser Electroporator). Finalement, les cellules ont été étalées sur boîte de Pétri contenant du milieu CSM (7g/L YNB, 0,75 g/L acides aminés, 20 g/L glucose) ne contenant pas l'acide aminé correspondant à l'auxotrophie apportée par le plasmide.

2.3.7 Test de conversion enzymatique

Les cellules transformées ont été inoculées dans 5 mL de YPD liquide pendant 16 heures sous agitation (200 tr/min) à 28°C. Les cultures ont ensuite été centrifugées 2 minutes à 3800

rpm (rotation par minute) puis lavées dans 30 mL d'eau stérile. Après une seconde centrifugation, les culots cellulaires ont été remis en suspension dans 5 mL de YPG (10 g/L yeast extract, 20 g/L peptone, 20 g/L galactose) puis cultivés pendant 5 heures sous agitation (200 tr/min) à 28°C afin d'induire l'expression des gènes clonés. Des fractions de 200µL ont été récupérées et mises en culture 24, 48 ou 72 heures sous agitation (200 tr/min) à 28°C en présence de 200µM de YAT ou DPT.

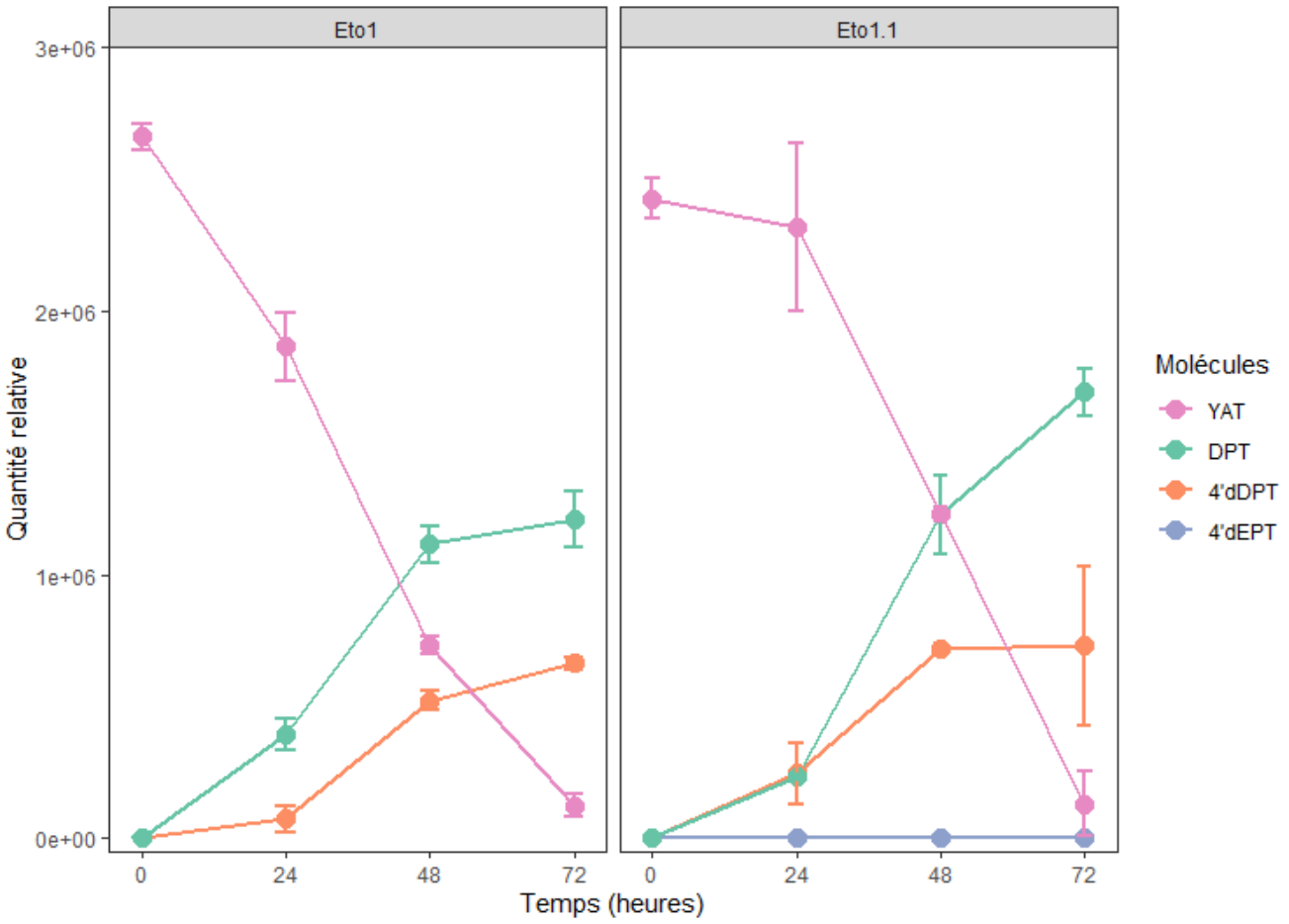
2.3.8 Analyses métaboliques

Le milieu extracellulaire de culture de levures a été obtenu par centrifugation (17 000g, 10mins) et dilué dix fois dans du méthanol 100%. Les échantillons dilués ont été à nouveau centrifugés (17 000g, 10mins) et le surnageant a été analysé par UPLC-MS sur le système Aquity™ Ultra-High Performance Liquid Chromatography (UPLC), couplé à un détecteur à barrette de diode (PDA) et à un spectromètre de masse (MS) contrôlé par le logiciel Masslynx 4.1 (Waters). La séparation des molécules a été effectuée sur une colonne C18 (Waters Acquity HSS T3C18, 150 x 2,1 mm, 1,8 µm) avec un débit de 0,4 mL/min à 55°C sur un volume d'injection de 5µL. La phase mobile est constituée d'un solvant A contenant 0,1% d'acide formique et de l'eau et d'un solvant B contenant 0,1% d'acide formique et de l'acétonitrile. La séparation chromatographique a d'abord été réalisée selon un gradient linéaire de 20 minutes avec une variation de 5 à 55 % de solvant B, puis elle a été optimisée selon un gradient linéaire de 8 minutes avec une variation de 20 à 75% de solvant B. La détection massique a été réalisée en électrospray avec une ionisation positive (ES+), le voltage du capillaire était de 3000 V et celui du cône de la source à 30 et 50 V. Le débit du cône et de la désolvatation des gaz était respectivement de 60 et de 800 L/h.

Les lignanes ont été identifiés selon leur temps de rétention (19 min (YAT), 18,1 min (DPT), 15,5 min (4'dDPT) et 11,1 min (4'dEPT) avec le premier gradient puis 7,2 min (YAT), 6,8 min (DPT), 5,8 min (4'dDPT) et 3,9 min (4'dEPT) avec le second gradient), leur valeur m/z (YAT et 4'dEPT : 401, DPT : 399,1, 4'dDPT : 385,1) et leur spectre de masse.

2.3.9 Western Blot

Les souches ont été inoculées dans 5 mL de YPD liquide pendant 16 heures sous agitation (200 tr/min) à 28°C. Les cultures ont été centrifugées 2 minutes à 3800 rpm puis lavées dans 30 mL d'eau stérile. Après une deuxième centrifugation, les culots cellulaires ont été remis en suspension dans 2,5 mL de Tris-HCl 50mM, pH 7,5. Des billes de verre ont été ajoutées à l'affleurement et les cellules ont été broyées 5 fois à fréquence maximale (30 secousses par seconde) pendant 30sec avec 1 minute dans la glace entre chaque broyage. Après avoir ajouté 2,5 mL de Tris-HCl 50mM, pH 7,5, le surnageant a été chargé sur un gel d'électrophorèse SDS-PAGE (Sodium Dodecyl Sulfate PolyacrylAmide Gel Electrophoresis) contenant du tétraméthyléthylènediamine (TEMED) afin de séparer les protéines présentes dans l'échantillon. Après transfert membranaire, la membrane contenant les protéines a été trempée dans un anticorps primaire anti-FLAG monoclonal de lapin, puis, après rinçage, dans un anticorps secondaire de chèvre anti-anticorps de lapin. Les protéines ont finalement été révélées par chimioluminescence.



<u>Figure 74</u> : Activités catalytiques de la 2ODD, CYP71BE54 et CYP82D61 dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT, DPT, 4'dDPT et 4'dEPT dans les extraits issus du surnageant de cultures des souches Eto1 et Eto1.1 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

Partie 3 : Ingénierie métabolique de la levure pour la bioconversion de YAT en 4'dEPT : construction d'une souche industrialisable

Les souches construites dans la partie précédente nous ont permis de lever le verrou concernant l'activité catalytique de CYP82D61 en levure et de vérifier la faisabilité de la bioconversion de YAT en 4'dEPT en levure. Mais ces souches ne sont pas utilisables à l'échelle industrielle puisque les gènes transférés sont portés par des plasmides à réplication autonome, il est donc désormais nécessaire d'intégrer les gènes d'intérêt dans le génome de la levure pour en garantir la stabilité génétique, et d'utiliser la souche de levure CEN.PK113 adaptée pour la bioproduction à grande échelle. Cette partie du mémoire présente la mise en place de cette souche industrialisable via l'insertion des gènes hétérologues, sous contrôle de promoteurs constitutifs forts, dans des « hots spot » du génome de la levure à l'aide de l'outil d'édition du génome CRISPR/Cas9, ainsi que les stratégies adoptées pour optimiser le niveau de production et les rendements de bioconversion de la souche.

3.1 Résultats

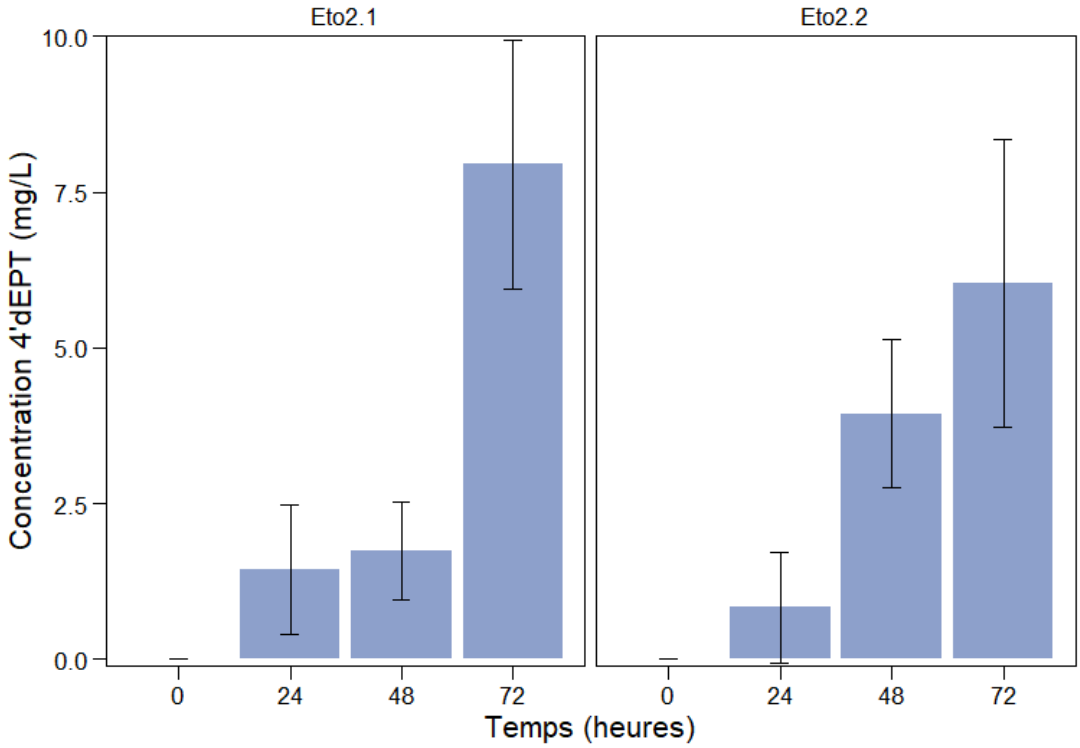
3.1.1 Première souche stable complète pour la biosynthèse de 4'dEPT

Une première souche stable dans le fond génétique de la levure haploïde CEN.PK113-7D a été construite par intégrations successives, via l'utilisation du système CRISPR/Cas9, d'une cassette d'expression contenant les séquences codantes de la 2ODD et CYP71BE54 au locus XII-4 et d'une cassette d'expression contenant la séquence codante d'AtATR2 (CPR *d'A. thaliana*) au locus X-3. A partir de cette souche, notée Eto1, une souche Eto1.1 a été construite par l'ajout d'une cassette d'expression de la séquence codant PhCYP82D61 au locus X-4. L'ensemble des cassettes d'expression utilisées sont constituées de couples promoteurs/terminateurs constitutifs forts (**Tableau 9**). Un test de bioconversion de 80 mg/L de YAT a été réalisé sur les souches Eto1 et Eto1.1, les variations au cours du temps des quantités relatives en YAT, DPT, 4'dDPT et 4'dEPT ont été suivies en analysant le surnageant de culture en UPLC (**Figure 74**). La quantité de YAT ajoutée au milieu de culture de la souche Eto1 diminue au cours du temps, pour être presque entièrement consommée à 72 heures de

culture. Cette consommation du précurseur est couplée à une augmentation des quantités de DPT et de 4'dDPT, représentant respectivement environ ⅔ et ⅓ des produits obtenus à 72 heures. Le profil de bioconversion obtenu avec la souche Eto1.1 est similaire à celui de la souche Eto1, et aucune production de 4'dEPT n'est détectée. Ces résultats montrent que la conversion de YAT en DPT catalysée par la 2ODD est quasi-totale après 72 heures de culture. En revanche, le taux de conversion de la DPT en 4'dDPT catalysée par CYP71BE54 n'est que de 27,3 à 30,1% en fonction de la souche et celui de la dernière étape catalysée par CYP82D61 est nulle. Ces taux de conversion sont équivalents lorsque le test est réalisé sur plusieurs clones différents issus de la transformation (données non représentées), indiquant une homogénéité entre les clones, contrairement à la partie précédente lorsque les levures étaient transformées par des plasmides à réplication autonome, où les niveaux de conversion variaient entre les clones. Des différences d'efficacité de conversion sont également observées entre les souches stables et les souches à réplication autonome. Ces deux différences sont dépendantes du nombre de copies des gènes d'intérêts présents dans la levure: un nombre variable de copies est présent dans les cellules ayant été transformées par un plasmide à réplication autonome, contre une seule copie dans les souches stables. En conclusion, l'amélioration du flux métabolique vers la synthèse de 4'dEPT dans cette première levure nécessite plusieurs ajustements afin d'augmenter la conversion de DPT en 4'dDPT et permettre celle de 4'dDPT en 4'dEPT.

3.1.2 Optimisation de la réduction de CYP82D61 dans la souche stable

Nous avons démontré dans la partie précédente que la réduction de CYP82D61 par PhCPR2 de *P. hexandrum* et CrCPR2 de *C. roseus* permettait d'obtenir de meilleurs rendements de conversion de 4'dDPT en 4'dEPT qu'avec AtATR1 ou AtATR2, sans altérer l'efficacité de l'activité de CYP71BE54. Nous avons donc évalué l'effet du remplacement d'AtATR2 utilisée dans la souche Eto1.1 par PhCPR2 ou CrCPR2. Pour cela, une souche Eto2 a été mise en place par intégration des cassettes d'expression de 2ODD, CYP71BE54 et CYP82D61 utilisées dans la souche Eto1.1. Deux souches ont ensuite été construites à partir d'Eto2 en intégrant une cassette d'expression de CrCPR2 (Eto2.1) ou PhCPR2 (Eto2.2) au locus XII-5. Les capacités de chaque souche à produire de la 4'dEPT à partir de 80 mg/L de YAT ont été évaluées au cours de la culture et sont représentées **Figure 75**. Les concentrations ont été calculées à partir d'une



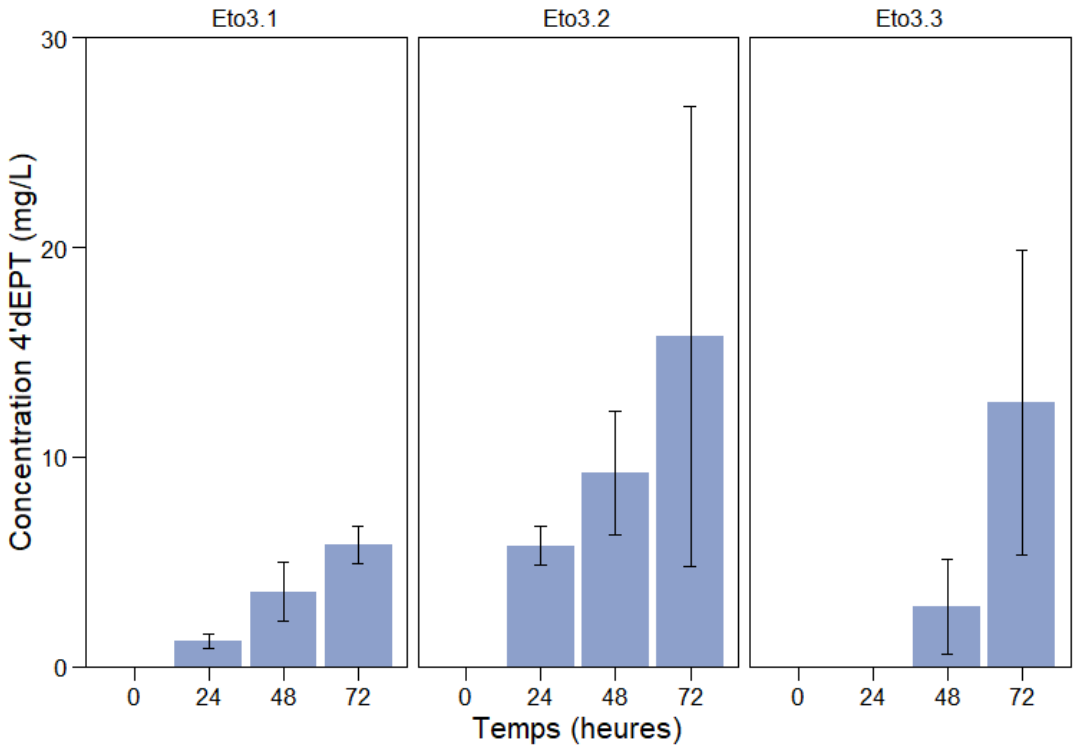
<u>Figure 75</u> : Impact de l'expression de CrCPR2 (Eto2.1) ou PhCPR2 (Eto2.2) sur l'efficacité catalytique de CYP82D61 dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) des quantités de 4'dEPT dans le surnageant de culture des souches Eto2.1 et Eto2.2 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

courbe de calibration de 4'dEPT purifiée. La figure montre que la concentration en 4'dEPT dans le surnageant de culture augmente au cours du temps pour les deux souches. Lorsque CYP71BE54 et CYP82D61 sont réduites par CrCPR2, la production moyenne de 4'dEPT est de 7,9±2,0 mg/L en 72 heures, contre 6,02±2,3 mg/L avec PhCPR2, ce qui représente un taux de conversion d'environ 10%. Cependant, une analyse statistique calculée à l'aide d'un test de Wilcoxon-Mann-Whitney montre que la valeur de probabilité p est supérieure à 0,05, indiquant que la différence de conversion entre les deux souches n'est pas significative. Ces résultats indiquent que PhCPR2 et CrCPR2 réduisent CYP71BE54 et CYP82D61 aussi efficacement dans la souche stable. Nous avons décidé de sélectionner PhCPR2 pour réduire les P450 dans la souche finale productrice de 4'dEPT afin de conserver une certaine stabilité puisque toutes les séquences utilisées proviennent de *P. hexandrum*.

3.1.3 Optimisation de la traduction de CYP82D61

Si les niveaux de production de 4'dEPT ont largement été améliorés par l'utilisation de PhCPR2 pour réduire CYP71BE54 et CYP82D61, le taux de conversion de 10% de la YAT présente dans le milieu de culture reste beaucoup trop faible pour une montée en échelle de la bioproduction. Pour améliorer la conversion de 4'dDPT en 4'dEPT, des versions codons optimisés de la séquence codante de CYP82D61 ont été employées à la place de la séquence native. En effet, les codons utilisés chez *P. hexandrum* n'ont pas la même fréquence d'utilisation par *S. cerevisiae*. Ainsi, l'adaptation de la séquence de CYP82D61 aux codons les plus fréquents de la levure devrait permettre une meilleure traduction des ARNm et une augmentation de la production de l'enzyme. Deux séquences codantes de PhCYP82D61 ont été synthétisées en optimisant les codons pour la levure à partir de deux algorithmes différents. Le premier est la plateforme web d'optimisation COOL (Codon Optimization OnLine) (Chin et al., 2014). Cette plateforme intègre les indices d'utilisation individuelle des codons, le contexte de codon, l'indice d'adaptation des codons de l'hôte et la fréquence de nucléotides G et C pour générer des séquences codantes optimisées. Le second algorithme utilisé est celui du logiciel GENEius mis en place par Eurofins Genomics (Köhler et al., 2006). Ce logiciel permet d'éviter les codons rares de l'organisme hôte, d'utiliser des codons à haute fréquence dans l'organisme sélectionné, de supprimer les motifs indésirables tels que les structures répétées ou en épingle à cheveux et d'uniformiser la distribution en bases G et C. Les séquences optimisées de

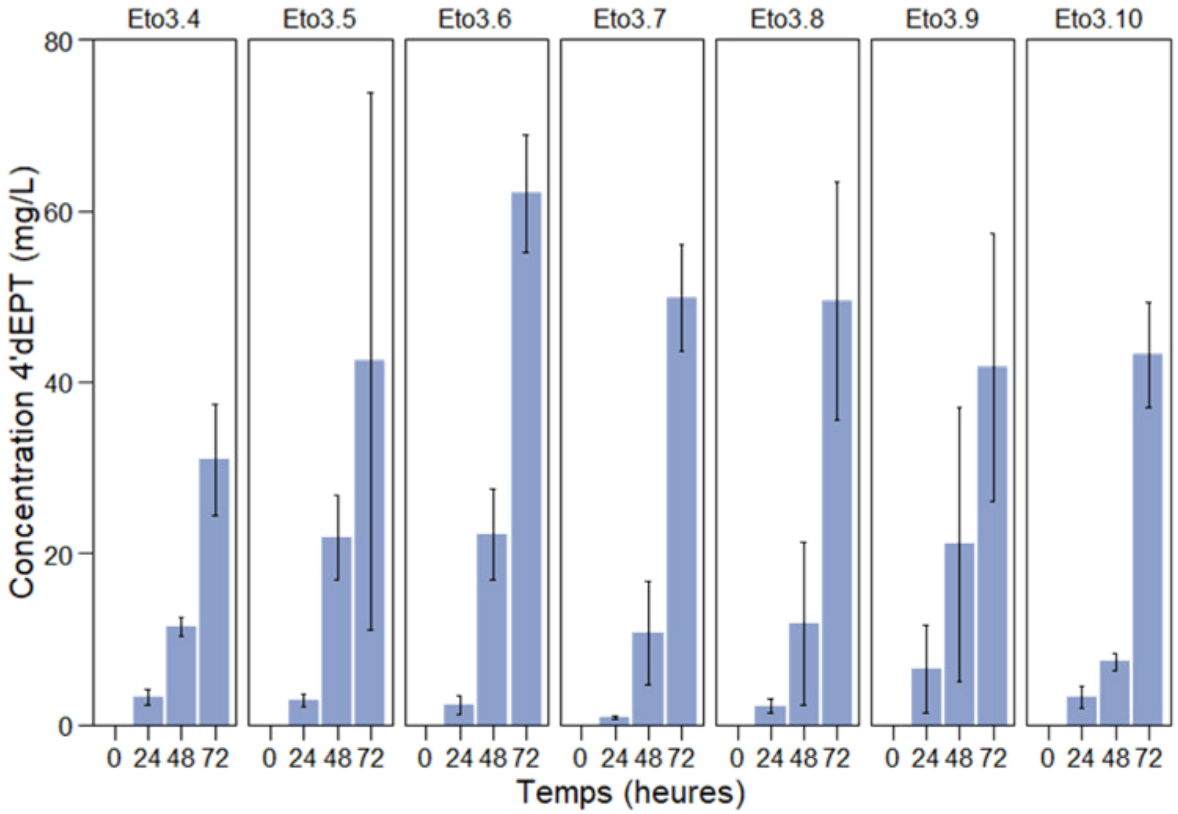


<u>Figure 76</u> : Comparaison des activités catalytiques des versions optimisées de CYP82D61 dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) des quantités de 4'dEPT dans le surnageant de culture des souches Eto3.1, Eto3.2 et Eto3.3 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

G et C. Les séquences optimisées de CYP82D61 obtenues par COOL et GENEius ont été respectivement nommées CYP82opt et CYP82opt_bis. Ces deux séquences disposent de 76% d'identité avec la séquence nucléotidique native codant CYP82D61.

Pour pouvoir tester l'efficacité de ces séquences CYP82D61 optimisées, une souche, notée Eto3, a été construite par intégration via le système CRISPR/Cas9 des cassettes d'expression contenant PhCPR2, 2ODD et CYP71BE54 aux loci XII-5 et XII-4, à l'image de la souche Eto2.2. Trois souches ont ensuite été générées par intégration de cassettes d'expression contenant PhCYP82D61 (Eto3.1), CYP82opt (Eto3.2) ou CYP82opt_bis (Eto3.3) au locus X-4 de la souche Eto3. La bioproduction de 4'dEPT à partir de YAT des trois souches a ensuite été évaluée dans les mêmes conditions que précédemment (**Figure 76**). La concentration en 4'dEPT dans le surnageant de culture augmente progressivement dès 24 heures de culture pour les souches Eto3.1 et Eto3.2, jusqu'à atteindre une concentration finale à 72 heures de culture de 5,8±0,9 mg/L et 15,8±11 mg/L respectivement. La concentration en 4'dEPT dans le surnageant de culture de la souche Eto3.3 n'augmente qu'à partir de 48 heures de culture, mais de façon plus drastique puisqu'entre 48 et 72 heures de culture, elle passe de 2,9±2,3 mg/L à 12,6±7,3 mg/L. Les valeurs p obtenues à l'aide de tests de Wilcoxon-Mann-Whitney sur les concentrations en 4'dEPT à 72 heures de culture sont inférieures à 0,05 entre les souches Eto3.1 et Eto3.2 et entre les souches Eto3.1 et Eto3.3, montrant que les différences sont significatives entre la souche contenant la séquence native de CYP82D61 et celles contenant les séquences optimisées. Par contre, le test de Wilcoxon-Mann-Whitney n'a montré aucune différence significative entre les concentrations en 4'dEPT obtenues à 72 heures de culture des souches Eto3.2 et Eto3.3. Ces résultats démontrent que l'optimisation pour la levure de la séquence codante de CYP82D61 permet d'améliorer l'efficacité catalytique de la conversion de 4'dDPT en 4'dEPT via une production sans doute plus élevée de la protéine, et ainsi de doubler la production finale de 4'dEPT. Bien que significative, cette amélioration n'est pas encore suffisante et la dernière étape de la voie de biosynthèse est toujours limitante. Pour augmenter encore plus la quantité d'enzymes CYP82D61 dans la levure, nous avons travaillé à multiplier le nombre de copies de CYP82D61 dans le génome de la levure, en combinant versions native et optimisées.



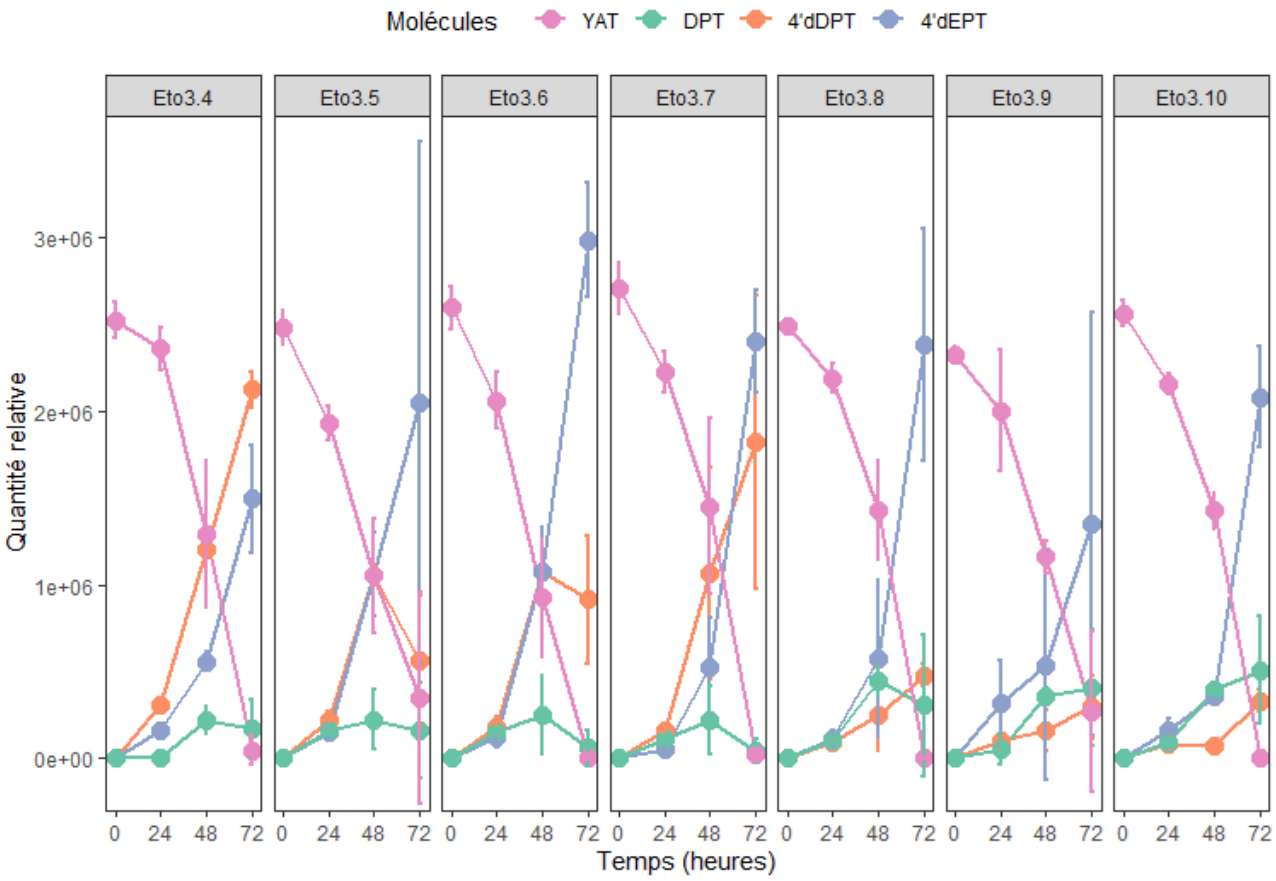
<u>Figure 77</u> : Production de 4'dEPT en fonction du nombre de copies de CYP82D61 et ses versions optimisées dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) des quantités de 4'dEPT dans le surnageant de culture des souches Eto3.4 à Eto3.10 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

3.1.4 Multiplication du nombre de copies de CYP82D61

L'optimisation de la réduction de CYP82D61 par PhCPR2 et de sa traduction via l'utilisation de séquences codantes optimisées pour la levure a permis une augmentation drastique des rendements de conversion de YAT en 4'dEPT. Grâce à ces stratégies, une bioconversion maximale de 15,8±11 mg/L de 4'dEPT a été obtenue dans la souche Eto3.2. Ces résultats très encourageants, demeurent toutefois encore insuffisants. En effet, dans les souches Eto3.1, Eto3.2 et Eto3.3, la 4'dDPT est l'intermédiaire de synthèse qui s'accumule le plus (données non représentées), suggérant que la synthèse de 4'dEPT à partir de la 4'dDPT représente toujours un goulot d'étranglement. C'est pourquoi, nous avons entrepris d'augmenter le nombre de copies de CYP82D61 pour accroître encore l'efficacité de conversion de 4'dDPT en 4'dEPT. En effet, il est courant d'utiliser cette stratégie dans les travaux d'ingénierie de voies métaboliques hétérologues en levure pour augmenter la bioproduction des composés d'intérêts (Li et al., 2015 ; Kulagina et al., 2021c). Dans le but de multiplier le nombre de copies des gènes codant la CYP82D61 native et ses versions optimisées, plusieurs souches ont été mises en place à partir de la souche Eto3 contenant les gènes codants PhCPR2, 2ODD et CYP71BE54. Pour vérifier la balance optimale des copies de gènes permettant une maximisation des rendements de conversion, plusieurs configurations ont été testées (**Tableau 9**), à commencer par le doublement du nombre de copies de CYP82D61 native ou optimisées. Des cassettes d'expression à deux copies de PhCYP82D61, de CYP82opt ou de CYP82opt_bis ont été construites puis intégrées dans le génome de la souche Eto3 via le système CRISPR/Cas9 respectivement au locus XI-3, X-4 ou XI-2. Les souches Eto3.4, Eto3.5 et Eto 3.6 ainsi générées ont été mises en culture 72 heures en présence de 80 mg/L de YAT, la production de 4'dEPT mesurée dans le surnageant de culture est présentée **Figures 77**. Pour chacune des trois souches, la concentration en 4'dEPT augmente au cours du temps. Dans l'ensemble, ces données montrent des concentrations en 4'dEPT supérieures à celles obtenues dans la **Figure 76** avec une concentration finale minimale de 31±6,5 mg/L pour la souche Eto3.4 et maximum de 62,1±6,9 mg/L pour la souche Eto3.6, multipliant jusqu'à quatre les concentrations obtenues avec une seule copie de CYP82D61 et démontrant ainsi l'intérêt de multiplier le nombre de copies du gène.

Dans un second temps, le nombre de copie de CYP82D61 a encore été augmenté mais en combinant les versions native et optimisées pour éviter une accumulation trop importante de

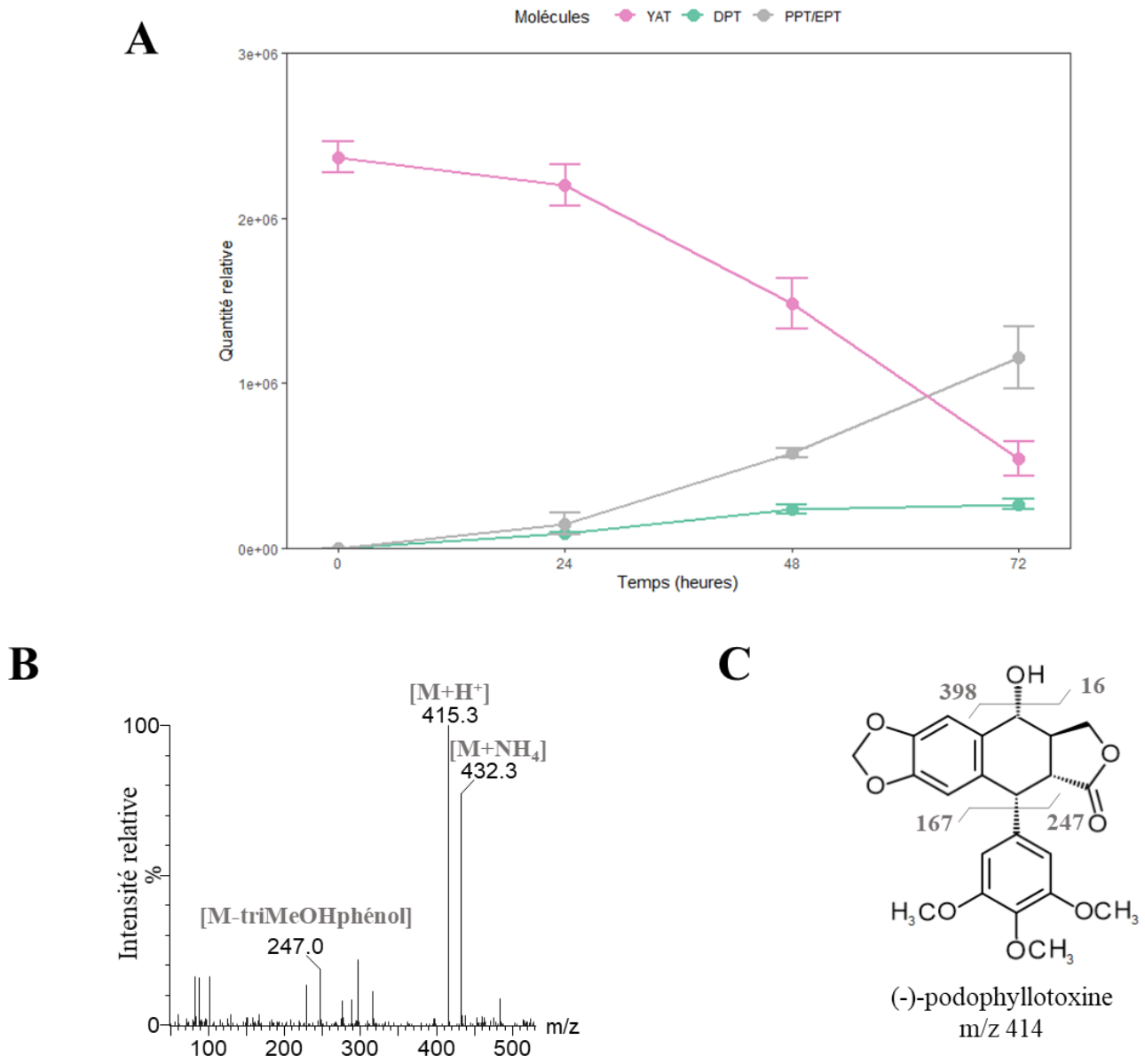


<u>Figure 78</u> : Evolution au cours du temps de la quantité relative de chaque intermédiaire en fonction du nombre de copies de CYP82D61 et ses versions optimisées dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT, DPT, 4'dDPT et 4'dEPT dans les extraits issus des surnageants de cultures des souches Eto3.4 à Eto3.10 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

transcrits identiques qui pourrait entraîner un phénomène de régulation épigénétique par interférence d'ARN (ARNi) (Goto and Nakayama, 2012). Parmi les souches construites à partir d'Eto3, la souche Eto3.7 possède une copie de PhCYP82D61 native et une de CYP82opt au locus X-4, la souche Eto3.8 possède deux copies de PhCYP82D61 natives et deux de CYP82opt aux loci X-4 et XI-3, la souche Eto3.9 possède deux copies de CYP82opt et deux de CYP82opt_bis aux loci X-4 et XI-2 et enfin, la souche Eto3.10 possède deux copies des trois variants de CYP82D61 aux loci X-4, XI-2 et XI-3. Les concentrations finales de 4'dEPT (**Figure 77**) mesurées à 72 heures dans les cultures des quatre souches alimentées par 80 mg/L de YAT ne montrent pas de différences significatives (tests de Wilcoxon-Mann-Whitney) par rapport aux souches contenant deux copies de CYP82D61 optimisées (Eto3.5 et Eto3.6). Au regard de la concentration finale en 4'dEPT, toutes les souches, hormis Eto3.6, semblent convertir la YAT en 4'dEPT de façon équivalente avec une concentration moyenne en 4'dEPT de 48,1±13,3 mg/L après 72 heures ce qui équivaut à environ 60% de la YAT de départ.

Les quantités de chaque intermédiaire de la voie de biosynthèse ont été reportées dans la **Figure 78** pour les souches Eto3.4 à Eto3.10. Cependant, ne disposant pas de toutes les molécules pour faire des quantifications absolues, la **Figure 78** présente des quantifications relatives, basées sur l'aire des pics de chaque composé mesurée à partir des chromatogrammes en UV de l'analyse en UPLC. Cette méthode de quantification est soumise à des approximations entre composés, car des divergences importantes ont été observées dans les courbes de calibration de lignanes purs dont nous disposons au laboratoire (notamment la YAT et la 4'dEPT). Elle reste pour autant informative, puisqu'au-delà de ces approximations, la **Figure 78** permet de visualiser le flux métabolique au fil de la bioconversion par les levures. Ainsi, pour toutes les souches, on peut observer une diminution progressive de la YAT du milieu de culture, jusqu'à une consommation totale à 72 heures de culture et l'augmentation progressive de la quantité de 4'dEPT qui semble s'accélérer après les 24 premières heures de culture. La quantité de DPT semble relativement faible et reste constante au cours de la culture, suggérant que CYP71BE54 consomme la DPT au fur et à mesure de sa biosynthèse. Par contre, la quantité de 4'dDPT à 72 heures de culture diffère entre les souches. En effet, une quantité non négligeable de 4'dDPT est observable dans le surnageant de culture des souches Eto3.4, Eto3.5, Eto3.6 et Eto3.7, mais est absente des souches Eto3.8, Eto3.9 et Eto3.10 alors que les quantités de 4'dEPT produites dans ces souches ne sont pas significativement différentes. Cette observation suggère que l'augmentation trop importante du nombre de copies de CYP82D61 puisse être responsable d'une perte du flux métabolique au niveau de la DPT ou de la 4'dDPT.



<u>Figure 79</u> : Production de PPT/EPT par la levure en absence de CYP71BE54.

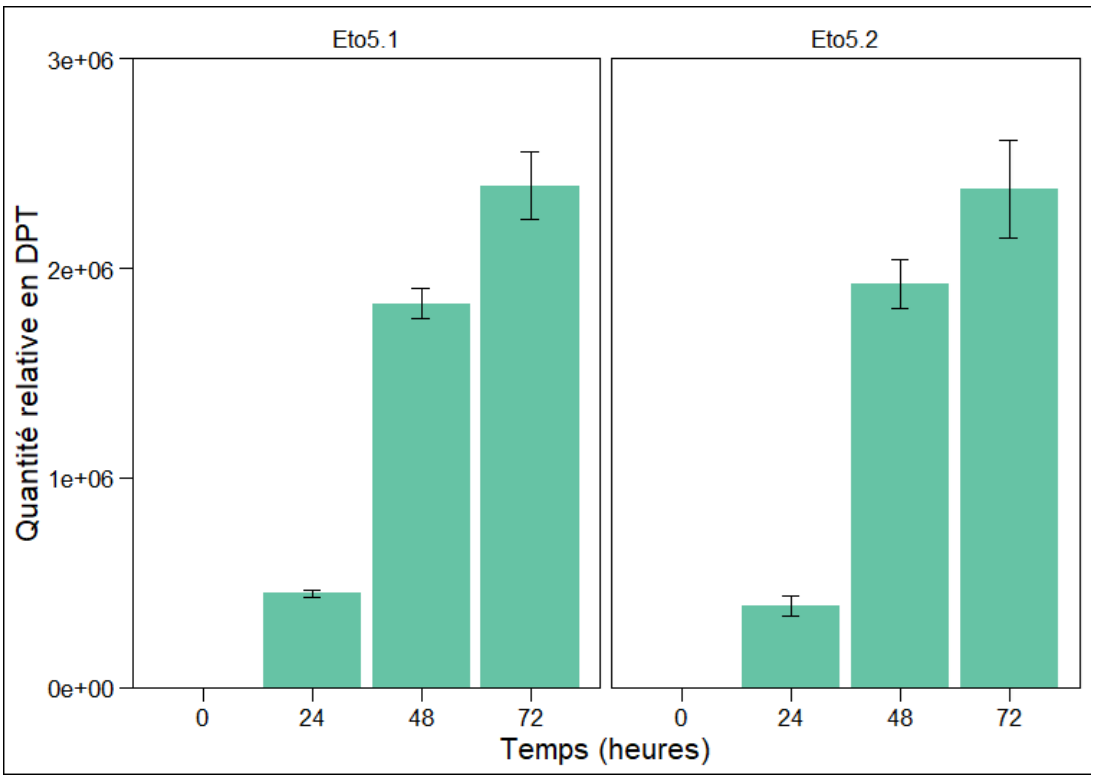
(A) Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT, DPT, et PPT/EPT dans les extraits issus des surnageants de cultures de la souche Eto4 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.
(B) Spectre de masse du pic obtenu pour la PPT/EPT. M = Molécule
(C) Structure chimique de la PPT. Les numéros en gris indiquent les masses molaires de chaque partie fragmentable de la molécule.

Dans nos conditions de bioconversion à partir de 80 mg/L de YAT, il serait intéressant de conserver une souche à deux copies de CYP82D61 et d'envisager une cinétique de culture plus longue. Cependant, lorsque la première souche productrice de 4'dEPT sera finalisée, elle devra ensuite être testée et éprouvée en bioréacteur, ce qui implique d'augmenter les quantités de YAT ajoutées au milieu de culture. Dans cette optique, nous avons préféré conserver les souches Eto3.9 et Eto3.10 plus appropriées pour une montée en échelle puisqu'elles disposent de quatre et six copies de CYP82D61 respectivement, et de rechercher les causes possibles de détournement du flux métabolique vers la synthèse de 4'dEPT afin de trouver une solution à ces pertes de rendement de production.

3.1.5 Recherche de fuites métaboliques

Avant de continuer l'optimisation de nos souches de levures productrices de 4'dEPT, nous avons voulu identifier l'origine de la perte de flux métabolique observée au niveau de la DPT ou de la 4'dDPT dans les souches précédentes. Dans un premier temps, la question d'une métabolisation de la DPT par la levure s'est posée. La souche sauvage *S. cerevisiae* CEN.PK113-7D a été mise en culture pendant une heure en présence de 80 mg/L de DPT et le surnageant de culture a été analysé en UPLC après extraction au méthanol. La quantité de DPT du milieu de culture ne varie pas significativement avec le temps (données non représentées) et la recherche non ciblée de lignanes dérivés de la DPT potentiellement produits par la levure sauvage n'a rien donné. La levure ne semble donc pas métaboliser la DPT pour produire d'autres molécules structurellement proches.

Les fuites métaboliques observées, notamment dans les souches Eto3.9 et Eto3.10, semblent liées à l'augmentation de l'activité de CYP82D61 dans les levures. Ainsi, la capacité de cette dernière à utiliser la DPT comme substrat a été envisagée puisque la DPT est structurellement proche de la 4'dDPT. La souche de levure Eto4 a été construite par intégration via le système CRISPR/Cas9 des gènes codants PhCPR2 (locus XI-2), 2ODD (locus XII-4), PhCYP82D61 (locus XI-3) et CYP82opt (loci XI-3 et X-4) dans le génome de *S. cerevisiae* CEN.PK113-7D. La souche Eto4 a été mise en culture pendant 72 heures en présence de 80 mg/L de YAT. La **Figure 79** présente les quantités relatives de YAT et ses dérivés présents dans le surnageant de culture obtenues par UPLC/MS. La quantité relative de YAT diminue au cours du temps dans le surnageant de culture de la souche Eto4, mais les quantités de DPT



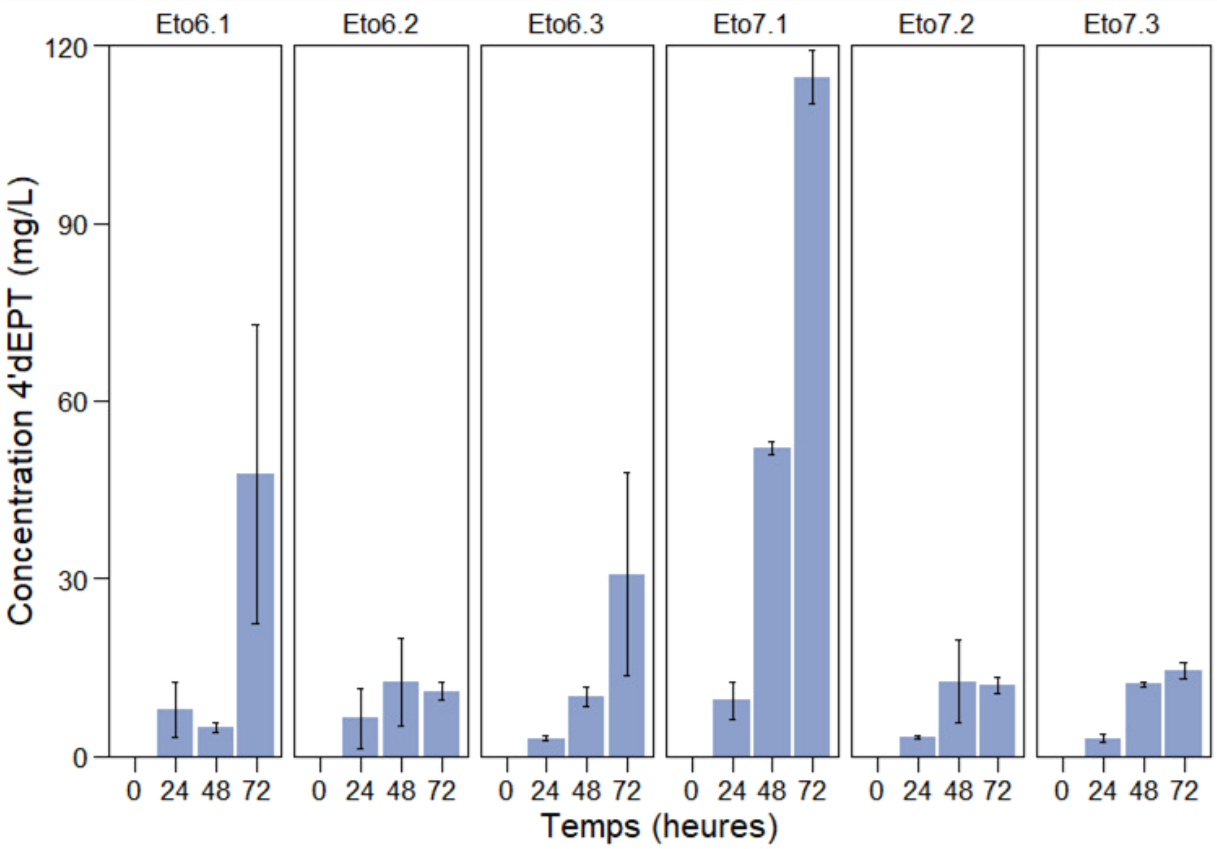
<u>Figure 80</u> : Production de DPT par les souches comportant la 2ODD native (Eto5.1) et la 2ODD orientée au peroxysome (Eto5.2).

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de DPT dans les extraits issus des surnageants de cultures des souche Eto5.1 et Eto5.2 après ajout de 80mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

produites sont relativement faibles par rapport à la consommation de la YAT. Une recherche non ciblée de lignanes dérivés de la DPT a permis de détecter un composé dont l'accumulation au cours du temps suit la disparition de la YAT. Le temps de rétention (5,15min) et le spectre de masse (**Figure 79.B**) de ce composé correspondent à ceux de la PPT ou de son isomère, l'épipodophyllotoxine (EPT) (**Figure 79.C**). Ces résultats suggèrent que CYP82D61 serait capable de catalyser la conversion de DPT en PPT et/ou EPT, montrant la présence d'une fuite métabolique au sein même de la voie de biosynthèse hétérologue. Une analyse approfondie des chromatogrammes obtenus en UPLC des surnageants de culture de toutes les souches testées dans nos travaux semble cependant indiquer qu'il y a peu de PPT/EPT accumulé. Si CYP71BE54 et CYP82D61 sont en compétition pour utiliser la DPT dans nos souches de levures capables de produire de la 4'dEPT à partir de YAT, la faible spécificité de CYP82D61 pour la DPT ne semble pas être à l'origine d'un détournement majeur du flux métabolique. Ces résultats ne permettent donc pas de mettre en évidence l'origine des pertes de flux observées pour les souches Eto3.9 et Eto3.10. Une analyse non ciblée de lignanes dérivés de la DPT plus poussée serait nécessaire à réaliser sur ces souches afin de comprendre l'origine de ces pertes d'intermédiaires. Par ailleurs, la capacité de synthèse de PPT/EPT par CYP82D61 directement à partir de la DPT indique qu'il existerait peut-être une voie alternative et plus directe pour la biosynthèse de PPT *in planta* chez Podophyllum.

3.1.6 Relocalisation de la 2ODD au peroxysome

Ne pouvant pas expliquer l'origine exacte de la perte d'intermédiaires métaboliques des souches Eto3.9 et Eto3.10, nous nous sommes attachés à renforcer les efficacités catalytiques des premières enzymes de la voie de biosynthèse de 4'dEPT afin d'augmenter le flux métabolique et réduire l'accumulation des intermédiaires. La 2ODD est une dioxygénase oxoglutarate dépendante de co-facteurs nécessaires à son activité qui sont produits dans le peroxysome. De fait, et de façon classique, les dioxygénases hétérologues clonées en levure sont délocalisées au peroxysome (Kulagina et al., 2021b). Nous avons donc vérifié si la relocalisation de la 2ODD au peroxysome pouvait être un moyen d'améliorer son efficacité catalytique dans la levure. Pour cela, les souches Eto5.1 et Eto5.2 ont été construites par intégration via le système CRISPR/Cas9 du gène codant la 2ODD native ou la 2ODD supplémentée de la séquence d'adressage au peroxysome PTS1 (Lee et al., 2009) au locus XII-



<u>Figure 81</u> : Production de 4'dEPT en fonction du nombre de copies de 2ODD et CYP71BE54 dans la souche stable.

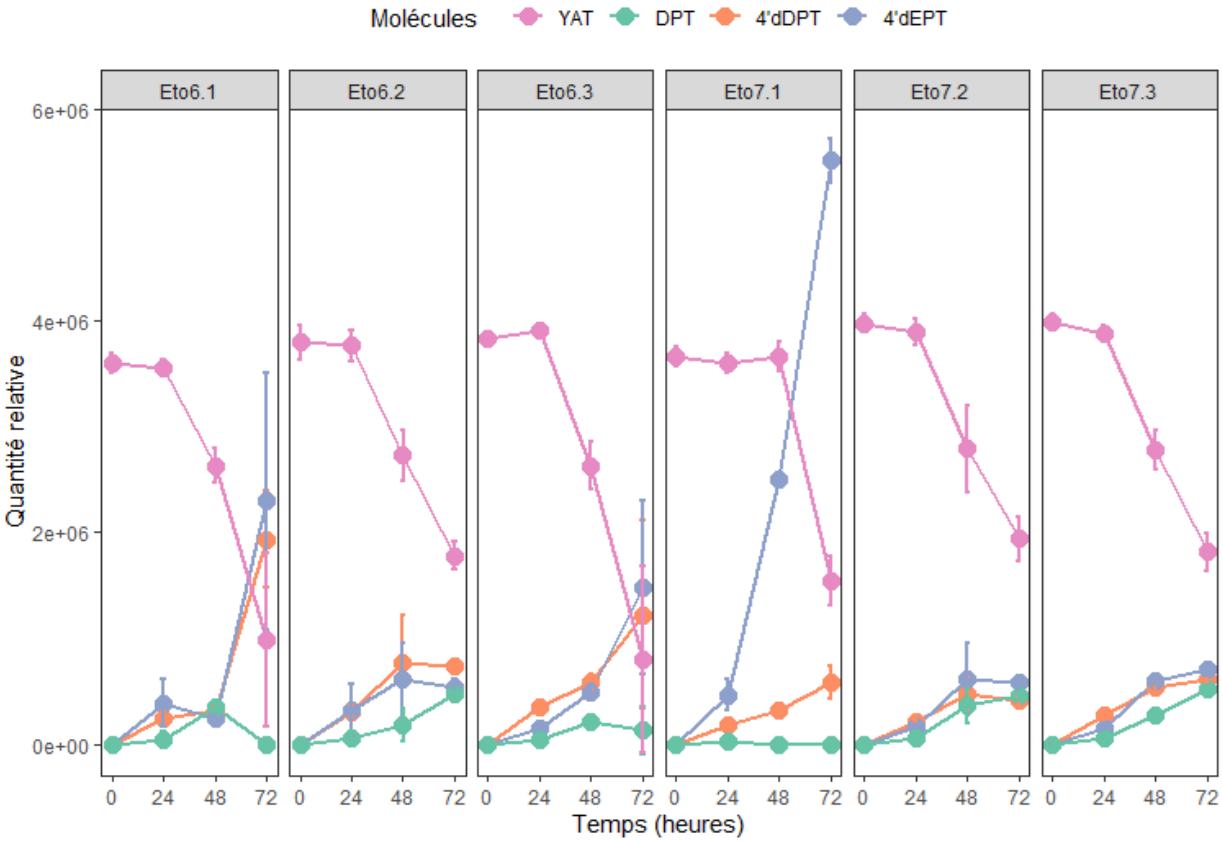
Mesures à partir du chromatogramme UV (290nm) des quantités de 4'dEPT dans le surnageant de culture des souches Eto6.1 à 6.3 à Eto7.1 à 7.3 après ajout de 180mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

4 du génome de la souche CEN.PK113-7D. Les deux souches obtenues ont été mises en culture pendant 72 heures en présence de 80 mg/L de YAT. L'analyse des surnageants de culture présentée dans la **Figure 80** montre une augmentation équivalente au cours du temps des quantités relatives de DPT entre les souches Eto5.1 et Eto5.2. Ce résultat indique que l'efficacité de bioconversion de la YAT par la 2ODD est similaire, que celle-ci soit dans le cytosol ou dans le peroxysome de la levure. La relocalisation de la 2ODD au peroxysome a donc été mise de côté et une stratégie d'augmentation du nombre de copies de 2ODD et CYP71BE54 a été développée.

3.1.7 Multiplication du nombre de copies de 2ODD et CYP71BE54

La multiplication du nombre de copies des gènes codant CYP82D61 couplée aux stratégies d'amélioration de la réduction de CYP82D61 par PhCPR2 et d'optimisation de la traduction par optimisation de la séquence nous ont permis d'augmenter significativement la production de 4'dEPT. Les souches Eto3.9 et Eto3.10 qui synthétisent respectivement 41,7±15,6 mg/L et 43,2±6,0 mg/L de 4'dEPT après 72 heures de culture à partir de 80 mg/L de YAT, ont été retenues pour la suite. Cependant, ces souches ne convertissent que 50% de la YAT en 4'dEPT et semblent présenter un détournement du flux métabolique. Pour augmenter la vitesse de bioconversion et atténuer les pertes d'intermédiaires, le nombre de copies de 2ODD et CYP71BE54 a été augmenté.

Six nouvelles souches ont été construites par ajout successif dans les souches Eto3.9 et Eto3.10 de trois cassettes d'expression contenant chacune une copie de 2ODD et une de CYP71BE54 aux loci XI-5, X-2 et XI-1. Les souches Eto3.9 et Eto3.10 ont donc permis de construire les souches Eto6.1/Eto6.2/Eto6.3 et Eto7.1/Eto7.2/Eto7.3 respectivement contenant deux, trois ou quatre copies de 2ODD et de CYP71BE54 (**Tableau 9**). Le test de bioconversion habituel a été réalisé sur les souches obtenues mais, cette fois-ci, en présence de 180 mg/L de YAT pour s'assurer de l'efficacité des souches lorsque la concentration initiale en YAT est un peu plus élevée. Les concentrations en 4'dEPT mesurées dans les surnageants de culture des différentes souches sont représentées **Figure 81**, et les niveaux d'accumulation de la YAT et des intermédiaires métaboliques au cours de la culture dans la **Figure 82**. Dans les souches Eto6.1 et Eto7.1 pour lesquelles le nombre de copies de 2ODD et de CYP71BE54 a été doublé par rapport aux souches parentes (Eto3.9 et Eto3.10), les concentrations finales en 4'dEPT



<u>Figure 82</u> : Evolution au cours du temps de la quantité relative de chaque intermédiaire en fonction du nombre de copies de 2ODD et CYP71BE54 dans la souche stable.

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT, DPT, 4'dDPT et 4'dEPT dans les extraits issus des surnageants de cultures des souches Eto6.1 à 6.3 à Eto7.1 à 7.3 après ajout de 180mg/L de YAT. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

obtenues sont respectivement de 47,6±25,3 mg/L et 114,7±4,4 mg/L. Le gain de production de 4'dEPT par la souche Eto6.1 n'est qu'1,2 fois plus élevé que la souche parente Eto3.9 et le rendement de bioconversion de la YAT de 52% dans la souche Eto3.9 chute ici à 26,4%. Cependant, toute la YAT n'est pas consommée en 72 heures et la très forte accumulation de 4'dDPT, occasionnée par l'augmentation des copies de 2ODD et CYP71BE54, indique que CYP82D61 est redevenue l'étape limitante (4 copies dans Eto6.1). En revanche, la production de 4'dEPT de la souche Eto7.1 est multipliée respectivement par 2,6 et 2,4 par rapport aux souches Eto3.10 et Eto6.1. L'augmentation de concert du nombre de copies de 2ODD, de CYP71BE54 et de CYP82D61 permet donc une très forte amélioration de la production de 4'dEPT, avec un taux de conversion de la YAT initiale de 63,7%. La YAT n'étant pas complètement consommée à 72 heures, l'efficacité de conversion réelle de YAT en 4'dEPT de la souche Eto7.1 est donc proche des 100% et une culture plus longue permettrait sans doute de convertir l'intégralité de la YAT de départ. Ce gain d'efficacité de bioconversion est sans doute permis par l'amélioration du flux métabolique dans la souche Eto7.1, puisque la DPT ne s'y accumule pas et qu'il y a peu de 4'dDPT. Ceci est sans doute à l'origine d'une réduction des pertes d'intermédiaires observées précédemment. Pour finir, les souches contenant trois ou quatre copies de 2ODD et de CYP71BE54 (Eto6.2, Eto6.3, Eto7.2 et Eto7.3) présentent une chute considérable dans la production de 4'dEPT sans pour autant accumuler de grandes quantités de DPT ou de 4'dDPT. La démultiplication des copies de 2ODD et de CYP71BE54, en utilisant toujours les mêmes séquences codantes, semble responsable de l'instabilité observée dans la conversion de la YAT en 4'dEPT.

L'ensemble de ces résultats met en évidence l'importance de l'optimisation de la balance du nombre de copies de chaque gène dans la voie de biosynthèse hétérologue. En effet, une augmentation trop importante de certains gènes peut influencer de façon répressive le flux métabolique et inhiber la synthèse du produit final. Dans ces travaux, il semblerait que la voie de bioconversion optimale pour la synthèse de 4'dEPT à partir de YAT ne nécessite au maximum que deux copies des gènes codant les deux premières enzymes de la voie de biosynthèse (2ODD et CYP71BE54) et à minima quatre copies du gène codant la dernière enzyme (CYP82D61). La souche Eto7.1 semble donc être la meilleure candidate pour une montée en échelle de la synthèse de 4'dEPT et une augmentation des niveaux de production de 4'dEPT via une optimisation de la conduite de la culture en bioréacteur.

Souche	Taux de croissance μ (h-1)	Biomasse (g/L)
CEN.PK113-7D	0,39	5,62
Souche 1	0,41	4,66
Souche 2	0,39	5,31
Souche 3	0,38	5,16
Souche 4	0,31	3,89
Souche 5	0,35	4,37
Souche 6	0,29	3,82

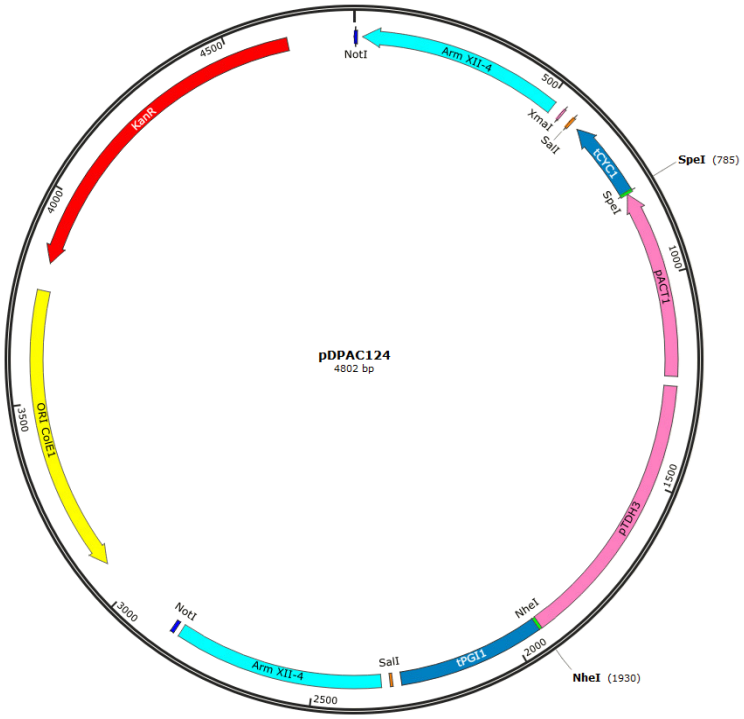
Tableau 4 : Tests de croissance réalisés sur des souches CEN.PK113-7D contenant un nombre variable de copies de gènes codant des P450.

Récapitulatif du taux de croissance et de la biomasse finale de chaque souche.

La souche représentée en gris est la souche sauvage, la souche en noir contient quatre copies de gènes codant des P450, les souches représentées en bleu et vert contiennent six copies de gènes codant des P450, en rouge et rose huit copies de gènes codant des P450 et la souche représentée en orange contient dix copies de gènes codants des P450.

3.1.8 Altération de la croissance cellulaire

Les travaux menés jusqu'à présent ont permis de mettre en place la souche Eto7.1 via le clonage dans le génome de la souche *S. cerevisiae* CEN.PK113-7D d'une copie du gène codant PhCPR2, deux copies des gènes codant 2ODD et CYP71BE54 et six copies de gènes codants CYP82D61 et ses versions optimisées. Cette souche est capable de produire 114,7±4,4 mg/L de 4'dEPT à partir de 180 mg/L de YAT. Une augmentation supplémentaire du nombre de copies des gènes de biosynthèse impacte la stabilité de la souche qui perd ses capacités de production de 4'dEPT. De plus, avec huit cytochromes P450 sur les onze gènes transférés, la consommation importante des cofacteurs tels que le NADPH et la suraccumulation de cytochromes P450 recombinants, ancrés à la membrane du RE, risque d'impacter le métabolisme primaire de la levure. Un test de croissance a donc été réalisé afin de vérifier l'impact de l'expression de multiples cytochromes P450 sur les levures. Pour cela, plusieurs souches contenant un nombre variable de copies de gènes codants CYP71BE54 et CYP82D61 native ou optimisée ont été mises en culture dans 30 mL de YPD en Erlenmeyer à baffles de 250mL pendant 8 heures à 30°C sous agitation de 120rpm. Le taux de croissance μ a été calculé à partir de la mesure de la densité optique de la culture à 590nm. La souche 1 contient quatre copies de gènes codant des P450, les souches 2 et 3 en contiennent six, les souches 4 et 5 en contiennent huit et la souche 6 contient 10 copies de gènes codant des P450. Le **Tableau 4** présente le taux de croissance et la biomasse à 8 heures de culture de chaque souche. Le taux de croissance des souches 1, 2 et 3 est similaire à celui de la souche sauvage CEN.PK113-7D, tandis que le taux de croissance des souches 4, 5 et 6 est inférieur à celui de la souche sauvage indiquant une vitesse de croissance plus faible. De façon corrélée, les souches 1, 2 et 3 ont une biomasse en fin de culture proche de celle de la souche sauvage, tandis que les souches 4, 5 et 6 ont une biomasse plus faible que la souche sauvage. Ces résultats montrent que la capacité de croissance et de division des levures est altérée à partir du moment où la souche contient huit copies de gènes codant des cytochromes P450 recombinants. La mesure du taux de croissance de la souche Eto7.1 n'a pas encore pu être réalisée mais, disposant de huit gènes codants des cytochromes P450, il est probable que sa croissance soit légèrement altérée. L'utilisation d'une souche de levure à croissance lente n'est évidemment pas souhaitable, mais peut être contrebalancée par le gain de productivité obtenu par la souche. Ce paramètre devra être pris en compte lors de la montée en échelle de la culture de la souche Eto7.1.



<u>Figure 83</u> : Exemple de carte d'un plasmide intégratif utilisé pour les clonages.

KanR: cassette de résistance à la kanamycine ; OR1 ColE1: origine de réplication en bactérie ; Arm: bras permettant la recombinaison homologue ; pTDH3//pACT1: promoteurs de transcription ; tPGI1/tCYC1 : terminateurs de transcription ; NotI: sites de restriction pour la linéarisation de la cassette d'intégration ; SpeI/NheI: sites de restriction correspondant aux sites de clonage.

<u>Tableau 5</u> : Liste des amorces utilisées pour l'amplification des gènes d'intérêt.

Nom	Séquence	Site de restriction	Fonction
AtATR2_fw	CTGAGAACTAGTTCGGTCATGAAAAACATGATGAATTATAAATTAAAAC	SpeI	Amplification AtATR2 d'*Arabidopsis thaliana*
AtATR2_rev	CTGAGAACTAGTTTAGTGGTGATGGTGATGATGCCATACATCTCTAAGATATCTTCC	SpeI	Amplification AtATR2 d'*Arabidopsis thaliana*
Ph2ODD_fw	CTGAGAACTAGTTCGGTCATGGGTTCTACA	SpeI	Amplification 2ODD de *P. hexandrum*
Ph2ODD_rev	CTGAGAACTAGTTCATGCACCTGTGTACGC	SpeI	Amplification 2ODD de *P. hexandrum*
PhCYP71_fw	CTGAGAACTAGTTCGGTCATGGAGTTCCTT	SpeI	Amplification CYP71DBE54 de *P. hexandrum*
PhCYP71_rev	CTGAGAACTAGTCTATTCTTTTGTAGTCGA	SpeI	Amplification CYP71BE54 de *P. hexandrum*
PhCYP82_fw	CTGAGATCTAGATCGGTCATGGATTCCCTG	XbaI	Amplification CYP82D61 de *P. hexandrum*
PhCYP82_rev	CTGAGATCTAGACTAGACAAAACATTTTGG	XbaI	Amplification CYP82D61 de *P. hexandrum*
CrCPR2_fw	CTGAGAGCTAGCTCGGTCATGGATTCTAGCTCGGAGAAGTTGTC	NheI	Amplification CrCPR2 de *Catharanthus roseus*
CrCPR2_tag_rev	CTGAGAGCTAGCTCAGTGGTGATGGTGATGATGCCAGACATCTCGGAGATACC	NheI	Amplification CrCPR2 de *Catharanthus roseus*
PhCPR2_fw	CTGAGAACTAGTTCGGTCATGCAATCAAGCTCAGCTTTC	SpeI	Amplification PhCPR2 de *P. hexandrum*
PhCPR2_rev	CTGAGAACTAGTTCACCATACGTCACGGAGATATC	SpeI	Amplification PhCPR2 de *P. hexandrum*
CYP82opt_fw	CTGAGAACTAGTTCGGTCATGGATTCATTACACTGTTTG	SpeI	Amplification CYP82D61 codons optimisés
CYP82opt_rev	CTGAGAACTAGTTTAAACAAAGCACTTTGG	SpeI	Amplification CYP82D61 codons optimisés
Ph2ODDperox_rev	CTGAGAACTAGTTCATAACTTAGATGCACCTGTGTACGCCTTGT	SpeI	Amplification 2ODD de *P. hexandrum* orientée au peroxysome

3.2 Conclusion

En définitive, le clonage dans le génome de la levure via le système CRISPR/Cas9 des gènes codant les enzymes hétérologues 2ODD, CYP71BE54 et CYP82D61 de la voie de conversion de la YAT en 4'dEPT permet de stabiliser la souche, puisque les rendements sont homogènes entre les différents clones, ainsi que de s'affranchir de l'utilisation de marqueurs de sélection auxotrophes ou d'antibiotiques. L'initialisation de la mise en place de la souche via le clonage d'une seule copie de chaque gène a permis de mettre en évidence les taux de conversion de chaque étape de la voie de biosynthèse et de sélectionner la CPR PhCPR2 pour la réduction des cytochromes P450 CYP71BE54 et CYP82D61. En effet, même si les souches codant les CrCPR2 de *C. roseus* et PhCPR2 de *P. hexandrum* convertissent la YAT en 4'dEPT avec des rendements similaires, le clonage de PhCPR2 permettra de stabiliser la souche et de n'avoir que des gènes issus de *P. hexandrum* dans la souche industrielle finale. Le clonage d'une copie de chaque gène (PhCPR2, 2ODD, CYP71BE54, CYP82D61) a permis d'obtenir une conversion de 6,0±,2,3 mg/L de 4'dEPT à partir de 80 mg/L de YAT. Ces résultats sont bien plus encourageants que les taux de conversion quasi nuls obtenus au début du projet, mais restent encore insuffisants pour une industrialisation de la souche. Afin d'optimiser les rendements finaux, le clonage de deux versions optimisées pour la levure de la séquence codant CYP82D61 a permis de mettre en évidence la présence d'un biais de codon sur la séquence native de CYP82D61. En effet, certains codons semblent être rarement utilisés par la levure, impactant la traduction de ce P450 dans cet organisme. Le clonage des deux versions optimisées de CYP82D61 à la place de la séquence native ont permis de doubler les rendements de conversion de 4'dEPT. Finalement, dans le but de pousser davantage le flux métabolique vers la synthèse de 4'dEPT, plusieurs configurations de multiplication de copies des gènes codant la CYP82D61 native et ses versions optimisées ont été testées, montrant qu'il est nécessaire d'intégrer au moins quatre copies des gènes codant CYP82D61 et ses versions optimisées dans la souche pour optimiser la bioproduction de 4'dEPT. Une fois cette dernière étape de la voie de synthèse optimale, la multiplication des copies de gènes codant 2ODD et CYP71BE54 a été initiée pour éviter l'accumulation de DPT et de 4'dDPT dans le surnageant de culture et optimiser le flux métabolique vers une biosynthèse maximale de 4'dEPT. Contre toute attente, une multiplication trop importante des copies de ces deux gènes semble inhiber la synthèse de 4'dEPT, suggérant qu'il est très important de prêter attention à la balance du nombre de copies de chaque gène et qu'une multiplication excessive des gènes hétérologues n'est pas forcément optimale pour

Tableau 6 : Liste des différents plasmides construits et des loci d'intégration ciblés associés.

Nom du plasmide	Cassette d'expression	Locus d'intégration ciblé
pPDHFC103-AtATR2	pTEF2-AtATR2-tCYC1	X-3
pDPAC124-Ph2ODD	pTDH3-Ph2ODD-tPGI1	XII-4
pDPAC124-Ph2ODDperox	pTDH3-Ph2ODDperox-tPGI1	XII-4
pDPAC124-Ph2ODD-PhCYP71BE54	pACT1-PhCYP71BE54-tCYC1 - pTDH3-Ph2ODD-tPGI1	XII-4
pPETA102-Ph2ODD	pPGK1-Ph2ODD-tENO2	X-2
pPETA102-Ph2ODD-PhCYP71BE54	pTEF1-PhCYP71BE54-tADH1 - pPGK1-Ph2ODD-tENO2	X-2
pDPTA111-PhCYP71BE54	ptDH3-PhCYP71BE54-tPGI1	XI-1
pDPTA111-Ph2ODD-PhCYP71BE54	pTEF1-Ph2ODD-tADH1 - ptDH3-PhCYP71BE54-tPGI1	XI-1
pPEAC115-Ph2ODD	pACT1-Ph2ODD-tCYC1	XI-5
pPEAC115-Ph2ODD-PhCYP71BE54	pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2	XI-5
pPETA104-PhCYP82D61	pTEF1-PhCYP82D61-tADH1	X-4
pPETA104-CYP82D61opt	pTEF1-CYP82D61opt-tADH1	X-4
pPETA104-CYP82D61opt_CYP82D61opt	pTEF1-CYP82D61opt-tADH1 - pPGK1-CYP82D61opt-tENO2	X-4
pDPTA125-CrCPR2	pTEF1-CrCPR2-tADH1	XII-5
pDPTA125-PhCPR2	pTEF1-PhCPR2-tADH1	XII-5
pPETA102-PhCYP71BE54	pTEF1-PhCYP71BE54-tADH1	X-2
pPETA104-PhCYP82D61-CYP82D61opt	pTEF1-PhCYP82D61-tADH1 - pPGK1-CYP82D61op-tENO2	X-4
pDPAC113-PhCYP82D61	pACT1-PhCYP82D61-tCYC1	XI-3
pDPAC113-PhCYP82D61-CYP82D61opt	pACT1-PhCYP82D61-tCYC1 - pTDH3-CYP82D61opt-tPGI1	XI-3
pDPAC113-PhCYP82D61-PhCYP82D61	pACT1-PhCYP82D61-tCYC1 - pTDH3-PhCYP82D61-tPGI1	XI-3
pDPTA125-PhCYP82D61	pTEF1-PhCYP82D61-tADH1	XII-5
pDPTA125-PhCYP82D61-CYP82D61opt	pTEF1-PhCYP82D61-tADH1 - pTDH3-CYP82D61op-tPGI1	XII-5
pDPAC112-CYP82D61opt_bis-CYP82D61opt_bis	pACT1-CYP82D61opt_bis-tCYC1 - pTDH3-CYP82D61opt_bis-tPGI1	XI-2

augmenter les rendements de conversion. Tous les résultats précédents ont permis la sélection de la souche Eto7.1 contenant une copie du gène codant PhCPR2, deux copies des gènes codant 2ODD et CYP71BE54, deux copies du gène codant la CYP82D61 native et deux copies de chaque version optimisée de CYP82D61. Cette souche finale convertit 64% de la YAT initialement ajoutée au milieu et a permis la synthèse de 114,7±4,4 mg/L finale de 4'dEPT. Par contre, la présence d'une fuite métabolique vers la production de PPT/EPT est à surveiller dans les prochaines étapes de la montée en échelle. Finalement, la croissance des levures étant altérée par la multiplication du nombre de copies de gènes hétérologues codant des cytochromes P450, il sera important de veiller à ce que la croissance de la souche Eto7.1 soit suffisante pour envisager une culture en bioréacteur, et que celle-ci n'ait pas d'impact sur les rendements de bioconversion.

La souche ayant été optimisée au niveau génétique, une montée en échelle de la culture a pu être initiée et sera présentée dans la partie suivante. L'optimisation de la conduite de la culture via le génie des procédés permettra d'augmenter encore les rendements de bioconversion de cette souche.

3.3 Matériel et méthodes

Dans cette partie, les méthodes d'amplification des gènes d'intérêt à partir de matériel végétal et d'analyses métaboliques ont été réalisées comme indiqué dans les paragraphes 2.3.2 et 2.3.8 respectivement de la partie précédente. En revanche, la construction des plasmides, la souche de levure, les transformations en levure et les tests de conversion enzymatique ont été réalisés différemment dans cette partie.

3.3.1 Construction des plasmides

pPEAC115, pPETA104, pDPTA125, pDPAC113 et pDPAC112 ont été préalablement construits au laboratoire (Kulagina et al., 2021c) (**Figure 83**). Ils ont été agencés à partir du squelette pDONR221. La cassette de clonage de ce plasmide d'origine a été remplacée par deux sites de restriction NotI situés en amont et en aval de deux bras correspondants aux différents loci d'intégration dans le génome de la levure (**Figure 30**) (Mikkelsen et al., 2012). Un site de

Tableau 7 : Liste des ARNg correspondant à chaque locus d'intégration.

Nom du vecteur	Locus d'intégration ciblé	Plasmide ARNg correspondant
pPETA102	X-2	pCfB3020
pDPTA111	XI-1	pCfB3043
pPEAC115	XI-5	pCfB3046
pDPAC124	XII-4	pCfB3049
pPETA104	X-4	pCfB3042
pDPTA125	XII-5	pCfB3050
pFCPDH103	X-3	pCfB3041
pDPAC113	XI-3	pCfB3045
pDPAC112	XI-2	pCfB3044

Tableau 8 : Amorces utilisées pour les PCR de vérification.

Nom	Séquence	Locus cible
Univ_DOWN_F	CAGGACTAGTGCTGAGGCATTAAT	
X-2_DOWN_R	GAGAACGAGAGGACCCAACAT	X-2
X-3_DOWN_R	CCGTGCAATACCAAAATCG	X-3
X-4_DOWN_R	GACGGTACGTTGACCAGAG	X-4
XI-1_DOWN_R	GAAGACCCATGGTTCCAAGGA	XI-1
XI-3_DOWN_R	CACATTGAGCGAATGAAACG	XI-3
XI-5_DOWN_R	GCATGGTCACCGCTATCAGC	XI-5
XII-2_DOWN_R	GGCCCTGATAAGGTTGTTG	XII-2
XII-4_DOWN_R	CGTGAAATCTCTTTGCGGTAG	XII-4
XII-5_DOWN_R	GTGGGAGTAAGGGATCCTGT	XII-5
Univ_UP_F	CGAGCTCGCTGAGGACTTAAT	
X-2_UP_R	TGCGACAGAAGAAAGGGAAG	X-2
X-3_UP_R	TGACGAATCGTTAGGCACAG	X-3
X-4_UP_R	CTCACAAAGGGACGAATCCT	X-4
XI-1_UP_R	CTTAATGGGTAGTGCTTGACACG	XI-1
XI-3_UP_R	GTGCTTGATTTGCGTCATTC	XI-3
XI-5_UP_R	CTCAATGATCAAAATCCTGAATGCA	XI-5
XII-2_UP_R	CGAAGAAGGCCTGCAATTC	XII-2
XII-4_UP_R	GAACTGACGTCGAAGGCTCT	XII-4
XII-5_UP_R	CCACCGAAGTTGATTTGCTT	XII-5

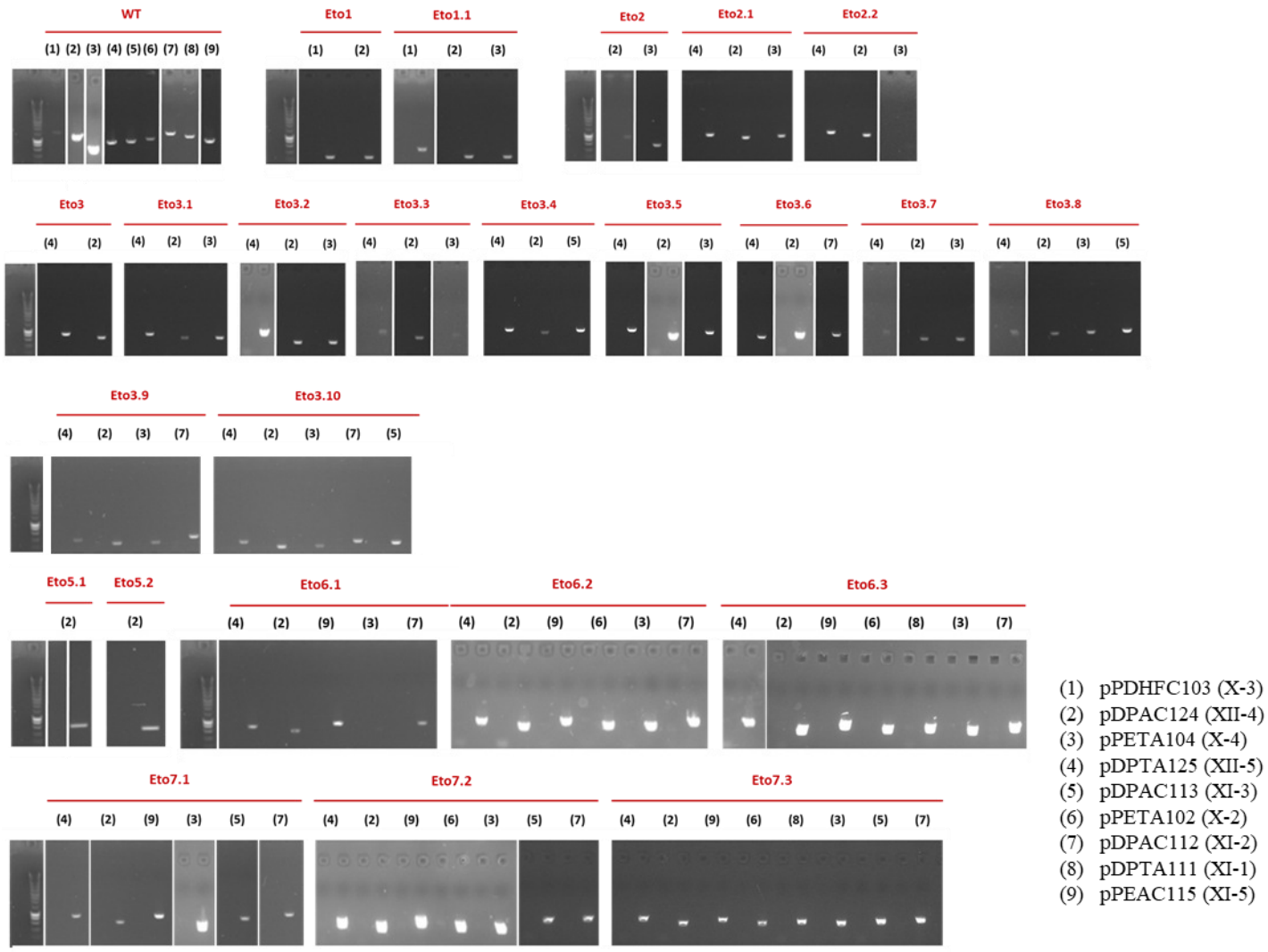
restriction SalI a été ajouté au milieu des deux bras. Les promoteurs et les terminateurs ont été amplifiés par PCR à l'aide de l'ADN polymérase Phusion à partir de l'ADN génomique (ADNg) de la souche *S. cerevisiae* CEN.PK113-7D. Les promoteurs et les terminateurs ont été purifiés et combinés par PCR combinatoire. Les amplicons obtenus ont été digérés par SalI et intégrés par ligation au niveau du site de restriction SalI situé entre les deux bras des plasmides squelettes. Les cassettes d'expression sont composées de deux couples promoteur/terminateur inversés pour synthétiser des promoteurs bidirectionnels artificiels. Ces séquences régulatrices varient entre les différents plasmides. Finalement, des sites de restriction SpeI et NheI ont été introduits entre chaque couple promoteur/terminateur afin de cloner les différentes ORF. Les amorces utilisées pour les PCR permettant d'amplifier les différentes ORF sont répertoriées dans le **Tableau 5**. Les bras entourant la cassette d'expression varient en fonction du locus ciblé et permettent l'intégration de la cassette dans le génome de la levure au niveau du locus d'intégration ciblé via une recombinaison homologue. Les plasmides construits disposent d'une cassette de résistance à la kanamycine permettant leur sélection en bactérie.

Les cellules *E. coli* transformées avec les plasmides d'intérêt, comme présenté paragraphe 2.3.3 de la partie précédente, ont été sélectionnées sur milieu LB supplémenté de 50µg/mL de kanamycine. Les plasmides obtenus dans cette partie du travail sont répertoriés dans le **Tableau 6**.

Les plasmides contenant les ARNg (**Tableau 7**) (Jessop-Fabre et al., 2016) et le plasmide d'expression Cas9 (pCfB2312) ont été commandés chez Addgene (EasyClone - MarkerFree Vector Set). Ils disposent respectivement d'une cassette de résistance à la nourséothricine et au sulfate de généticine G-418, permettant la sélection des levures transformées.

3.3.2 Souche de levure

La souche de départ utilisée pour effectuer les premières intégrations chromosomiques des cassettes d'intérêt est la souche haploïde *Saccharomyces cerevisiae* CEN.PK113-7D de génotype *MATα MAL2-8C SUC2* commandée chez EUROSCARF (Van Dijken et al., 2000). Cette souche dispose d'une forte vitesse de croissance et est adaptée à la production de molécules d'intérêts en bioréacteur, elle est donc optimisée pour une montée à l'échelle



<u>Figure 84</u>: Gels d'électrophorèses obtenus après amplification par PCR sur l'ADNg de chaque souche.

Les PCR ont été réalisées avec les amorces présentées Tableau 4 permettant, pour chaque souche et chaque locus, d'avoir un contrôle négatif (à gauche) et une amplification si la cassette a été intégrée au locus ciblé (à droite).

industrielle des cultures. Le plasmide d'expression Cas9 a été transformé dans cette souche afin de permettre les transformations avec les plasmides ARNg.

3.3.3 Transformation des levures

Différentes intégrations via le système CRISPR/Cas9 ont été réalisées pour introduire les cassettes dans les différents « hot-spots » du génome de la levure (Shao et al., 2009 ; Mikkelsen et al., 2012 ; Kulagina et al., 2021c). Les plasmides intégratifs contenant les gènes d'intérêt ont été digérés avec l'enzyme de restriction NotI. La souche a été transformée avec 1µg du plasmide intégratif linéarisé et 500ng d'ARNg selon la méthode LiAc/PEG (Gietz et al., 2007). Les transformants ont été sélectionnés sur un milieu solide YPD contenant 200mg/L de sulfate de généticine G-418 et 100mg/L de nourséothricine. L'intégration du gène dans la souche stable a été confirmé par PCR (DreamTaq, ThermoFisher) sur colonie en utilisant les amorces décrites dans le **Tableau 8** et confirmée par PCR sur l'ADNg extrait des différentes souches (**Figure 84**). L'ADNg des levures a été extrait par suspension d'une colonie dans 100µL de 200mM de lithium d'acétate et 1% de SDS, puis incubée 5 minutes à 70°C. 300µL d'éthanol 100% ont été ajouté à la suspension cellulaire et agités par vortex. Le mélange a été centrifugé 3 minutes à 15 000g puis l'éthanol a été retiré et le culot a été lavé avec 600µL d'éthanol 70%. Le tube a été centrifugé 3 minutes à 15 000g et le culot a été séché avant d'être mis en suspension dans 100µL d'eau MilliQ puis de nouveau centrifugé 1 minute à 15 000g. Les PCR ont été réalisées sur 1µL de surnageant contenant l'ADNg.

Les vecteurs contenant les ARNg ont été retirés des souches via la culture sur milieu solide YPD ne contenant que le sulfate de généticine G-418. En effet, le plasmide contenant le Cas9 disposant d'une cassette de résistance à la généticine G-418, cet antibiotique est toujours ajouté dans le milieu de culture des souches dans lesquelles de nouvelles intégrations sont prévues. Les différentes souches obtenues sont répertoriées dans le **Tableau 9**.

3.3.4 Test de conversion enzymatique

Les cellules transformées ont été inoculées dans 5mL de milieu liquide YPD pendant 16 heures sous agitation (200 tr/min) à 28°C. Les cultures ont ensuite été centrifugées 2 minutes

Tableau 9 : Liste des souches construites et le génotype associé.

Souche	Génotype	Source
WT	CEN.PK113-7D (*MATa MAL2-8C, SUC2)*	H0
Eto1	H0; X-3::pTEF2-AtATR2-tCYC1; XII-4::pACT1-PhCYP71BE54-tCYC1 - pTDH3-Ph2ODD-tPGI1	H1
Eto1.1	H1; X-4::pTEF1-PhCYP82D61-tADH1	
Eto2	H0; XII-4::pACT1-PhCYP71BE54-tCYC1 - pTDH3-Ph2ODD-tPGI1; X-4::pTEF1-PhCYP82D61-tADH1	H2
Eto2.1	H2; XII-5::pTEF1-CrCPR2-tADH1	
Eto2.2	H2; XII-5::pTEF1-PhCPR2-tADH1	
Eto3	H0; XII-5::pTEF1-PhCPR2-tADH1; XII-4::pACT1-PhCYP71BE54-tCYC1 - pTDH3-Ph2ODD-tPGI1	H3
Eto3.1	H3; X-4::pTEF1-PhCYP82D61-tADH1	
Eto3.2	H3; X-4::pTEF1-CYP82opt-tADH1	
Eto3.3	H3; X-4::pTEF1-CYP82opt_bis-tADH1	
Eto3.4	H3; XI-3::pACT1-PhCYP82D61-tCYC1 - pTDH3-PhCYP82D61-tPGI1	
Eto3.5	H3; X-4::pTEF1-CYP82opt-tADH1 - pPGK1-CYP82opt-tENO2	
Eto3.6	H3; XI-2::pACT1-CYP82opt_bis-tCYC1 - pTDH3-CYP82opt_bis-tPGI1	
Eto3.7	H3; X-4::pTEF1-PhCYP82D61-tADH1 - pPGK1-CYP82opt-tENO2	H4
Eto3.8	H4; XI-3::pACT1-PhCYP82D61-tCYC1 - pTDH3-CYP82opt-tPGI1	
Eto3.9	H3; X-4::pTEF1-CYP82opt-tADH1 - pPGK1-CYP82opt-tENO2; XI-2::pACT1-CYP82opt_bis-tCYC1 - pTDH3-CYP82opt_bis-tPGI1	H5
Eto3.10	H3; X-4::pTEF1-PhCYP82D61-tADH1 - pPGK1-CYP82op-tENO2; XI-3::pACT1-PhCYP82D61-tCYC1 - pTDH3-CYP82opt-tPGI1; XI-2::pACT1-CYP82opt_bis-tCYC1 - pTDH3-CYP82opt_bis-tPGI1	H6
Eto4	H0; XI-2::pACT1-CYP82opt_bis-tCYC1; XII-4::pTDH3-Ph2ODD-tPGI1; XI-3::pACT1-PhCYP82D61-tCYC1 - pTDH3-CYP82opt-tPGI1; X-4::pTEF1-CYP82opt-tADH1	
Eto5.1	H0; XII-4::pTDH3-Ph2ODD-tPGI1	
Eto5.2	H0; XII-4::pTDH3-Ph2ODDperox-tPGI1	
Eto6.1	H5; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2	
Eto6.2	H5; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2; X-2::pTEF1-PhCYP71BE54-tADH1 - pPGK1-Ph2ODD-tENO2	
Eto6.3	H5; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2; X-2::pTEF1-PhCYP71BE54-tADH1 - pPGK1-Ph2ODD-tENO2; XI-1::pTEF1-Ph2ODD-tADH1 - pTDH3-PhCYP71BE54-tPGI1	
Eto7.1	H6; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2	
Eto7.2	H6; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2; X-2::pTEF1-PhCYP71BE54-tADH1 - pPGK1-Ph2ODD-tENO2	
Eto7.3	H6; XI-5::pACT1-Ph2ODD-tCYC1 - pPGK1-PhCYP71BE54-tENO2; X-2::pTEF1-PhCYP71BE54-tADH1 - pPGK1-Ph2ODD-tENO2; XI-1::pTEF1-Ph2ODD-tADH1 - pTDH3-PhCYP71BE54-tPGI1	

à 3800 rpm puis lavées dans 25 mL d'eau stérile. Après une seconde centrifugation, les culots cellulaires ont été remis en suspension dans 5 mL de YPD frais puis des fractions de 200µL ont été récupérées et mises en culture 24, 48 ou 72 heures sous agitation (200 tr/min) à 28°C en présence de 80 mg/L ou 180 mg/L de YAT.

3.3.5 Courbes de calibration

Les courbes de calibration de la YAT et la 4'dEPT ont été obtenues en réalisant différentes dilutions de stock de YAT et 4'dEPT dans du DMSO (Diméthylsulfoxyde). Le stock de YAT a été réalisé à partir de YAT commandée chez PhytoLab. Le stock de 4'dEPT a été obtenu à partir d'une purification de milieu de culture d'une souche productrice de 4'dEPT. Les différentes dilutions ont été réalisées dans la phase mobile correspondant au gradient utilisé pour les analyses en UPLC (95% H_2O + 0,1% Acide Formique ; 5% Acétonitrile + 0,1% Acide Formique). Une analyse UPLC/MS de solutions contenant 50mg/L, 25mg/L, 10mg/L, 5mg/L, 2mg/L, 1mg/L, 0,2mg/L ou 0,1mg/L de YAT ou de 4'dEPT a permis d'obtenir les courbes de calibration de ces molécules grâce à l'équation de la droite obtenue.

Les équations des droites obtenues ont permis de déterminer les résultats suivants :

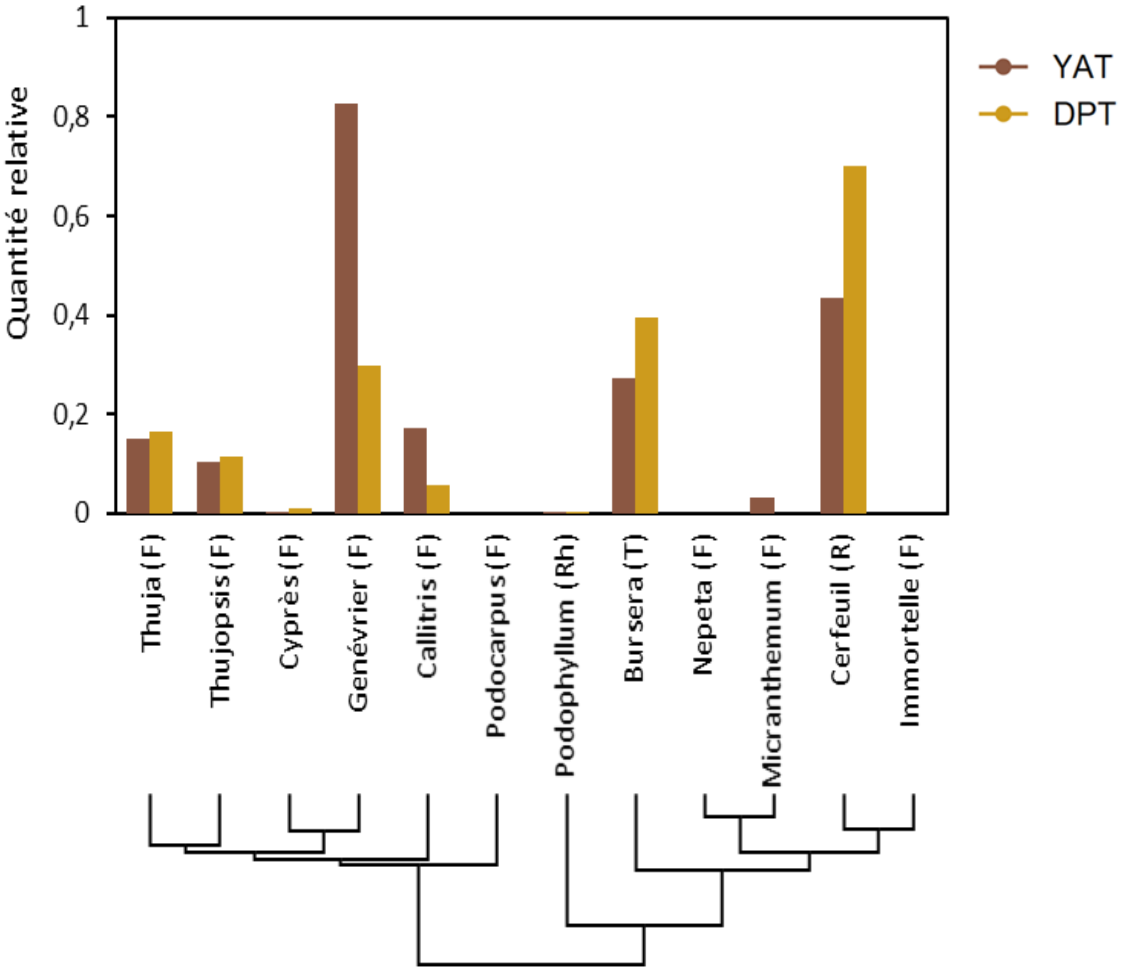
- Concentration YAT (mg/L) = (Aire + 15918) / 19596
- Concentration 4'dEPT (mg/L) = (Aire - 8276) / 47961

L'aire correspond à l'aire sous le pic du spectre UV obtenu par UPLC/MS.

La DPT et la 4'dDPT n'existant pas en standard analytique et étant difficiles à purifier à partir du surnageant de culture, les courbes de calibration de ces intermédiaires réactionnels n'ont pas pu être réalisées pour le moment. La DPT et la 4'dDPT n'étant pas quantifiables, les résultats comprenant ces intermédiaires ont été exposés de façon relative en mesurant l'aire sous le pic correspondant du spectre UV obtenu en UPLC/MS.

3.3.6 Analyses statistiques

Pour chaque condition (temps, souche), cinq répétitions ont été réalisées et analysées. Une analyse en composantes principales sur ces répétitions a permis de sélectionner les trois réplicats les plus représentatifs. La moyenne et l'écart-type ont été calculés à partir de ces trois échantillons sélectionnés. Enfin, la valeur de probabilité p a été calculée à l'aide d'un test de Wilcoxon-Mann-Whitney sur les cinq répétitions. Si la valeur p est supérieure à 0,05, la différence entre les échantillons comparés n'est pas significative.



<u>Figure 85</u> : Recherche d'espèces végétales productrices de YAT.

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT et de DPT dans les extraits réalisés.

F : Feuilles ; Rh : Rhizome ; T : Tige ; R : Racine. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

Partie 4 : Biosourcing et purification de YAT pour une montée de la production de 4'dEPT à l'échelle d'un bioréacteur

Suite à nos travaux d'ingénierie moléculaire, nous avons généré une souche de levure (Eto7.1), potentiellement compatible avec une utilisation industrielle, et capable de produire 114,7±4,4 mg/L de 4'dEPT, précurseur d'étoposide, par conversion en 72 heures d'environ 65% de la YAT ajoutée au milieu de culture. La production de 4'dEPT par cette souche de levure se base sur une stratégie de bioconversion de YAT, nécessitant de fait de grandes quantités initiales de cette molécule pour envisager, à terme, une production industrielle d'étoposide. Celle-ci est disponible sur le marché avec une pureté avoisinant les 98% mais reste particulièrement coûteuse, de l'ordre de 15 000 € du gramme. L'utilisation de cette molécule issue du commerce n'est donc pas envisageable à grande échelle par Axyntis, imposant de fait le développement d'une approche d'approvisionnement autonome. La YAT étant présente chez de nombreux végétaux, un de mes objectifs de thèse a été, *in fine*, de jeter les bases de la mise en place d'une filière d'approvisionnement de YAT en Région Centre-Val de Loire en sélectionnant des espèces naturellement riches en YAT, cultivables en France, puis en adaptant à plus grande échelle un protocole d'extraction et de purification de cette molécule à partir de grandes quantités du matériel végétal collecté. Au-delà d'une future possible utilisation industrielle, l'obtention de grandes quantités de YAT par ces approches s'avérait être un élément essentiel dans le cadre de mon travail, en me permettant d'initier des montées en échelle des cultures de levures et de la production de 4'dEPT en petit bioréacteur de 2L.

4.1 Résultats

4.1.1 Identification d'espèces végétales productrices de YAT

L'extraction et la purification de grandes quantités de YAT à partir de matériel végétal est nécessaire pour assurer à terme la pérennité de la production d'étoposide par bioconversion. Une recherche a donc été effectuée afin de déterminer la ou les espèces végétales les plus appropriées pour la mise en place d'un protocole d'extraction et de purification de YAT à grande échelle. Une priorité certaine a été mise sur les plantes cultivables sous nos latitudes mais d'autres plantes connues pour accumuler des dérivés de PPT ont également été analysées



J. virginiana



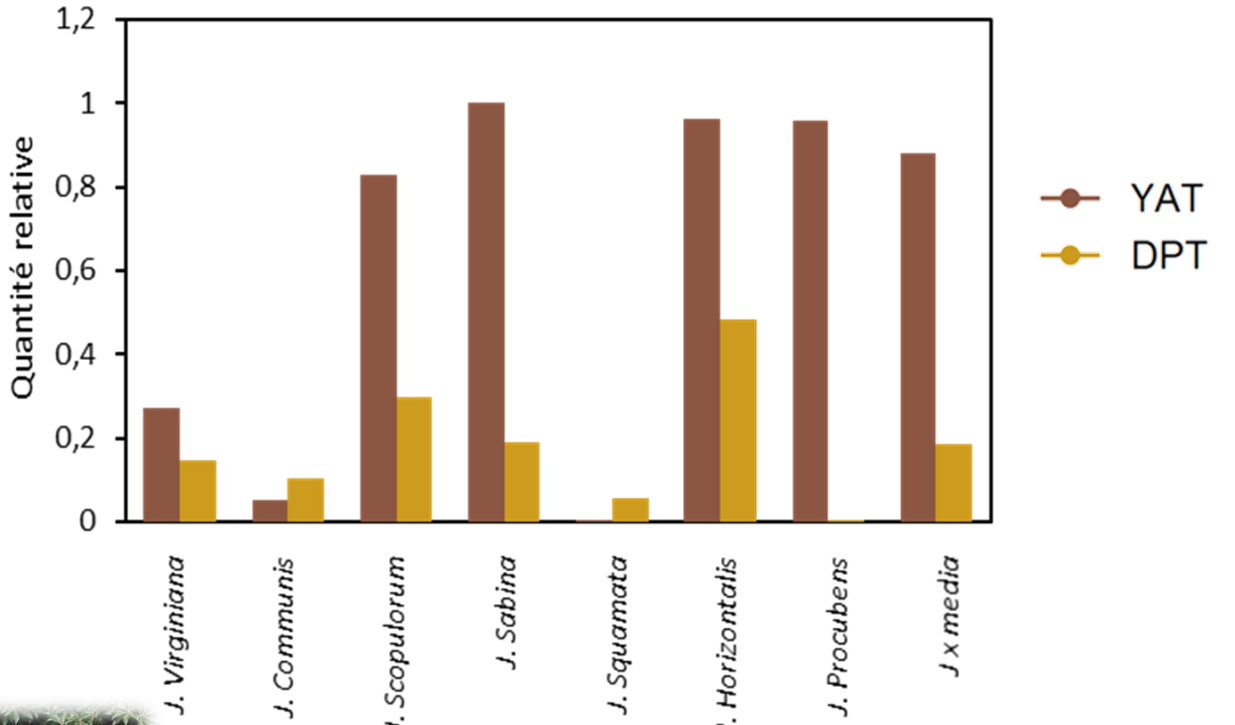
J. communis



J. scopulorum



J. sabina





J. squamata



J. horizontalis



J. procubens



J. x media

<u>Figure 86</u> : Recherche de l'espèce de Juniperus accumulant le plus de YAT.

Mesures à partir du chromatogramme UV (290nm) de la quantité relative de YAT et de DPT dans les extraits réalisés. Mesures UPLC-PDA-MS, gradient linéaire de 8 minutes.

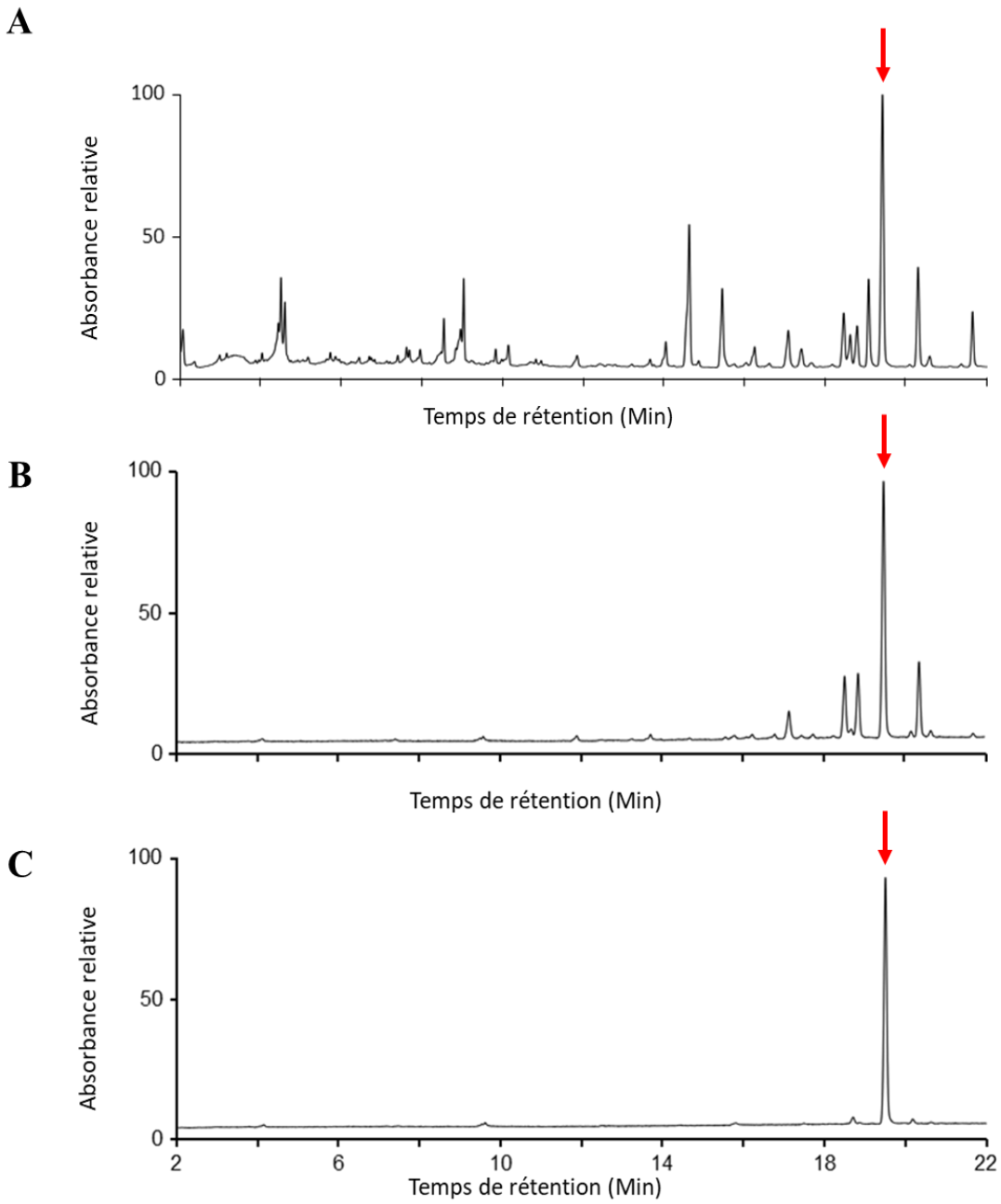
en contrôle. Des extractions au méthanol ont ainsi été réalisées sur différents organes de dicotylédones d'intérêt : Nepeta, Micranthemum, immortelle, Bursera, cerfeuil et Podophyllum. Sachant que la YAT est accumulée chez les plantes de la famille des Cupressaceae (Hwang et al., 2004), des extractions au méthanol ont également été réalisées sur les feuilles de plantes provenant de différents genres issus de cette famille : Thuja, Thujopsis, Cyprès, Genevrier, Callitris et Podocarpus. Les extractions ont d'abord été réalisées au méthanol après broyage au mortier pilon des tissus végétaux congelés dans l'azote liquide. Les extraits obtenus ont été analysés en UPLC pour déterminer la quantité relative de YAT dans chaque extrait à partir des chromatogrammes UV (**Figure 85**). La DPT a également été dosée, car il est également envisageable de réaliser les bioconversions en levure à partir d'un mélange de YAT et de DPT puisque la souche construite est capable de convertir ces deux substrats.

Les résultats présentés **Figure 85** ne montrent aucune accumulation de YAT, ni de DPT dans les extractions réalisées à partir de Podocarpus, Podophyllum, Nepeta et Immortelle. Dans les extraits de Cyprès et de Micranthemum, les quantités de YAT et de DPT sont trop faibles pour être exploitées (inférieur à 1 mg/gMF (gramme de matière fraîche)). Ces six espèces ne sont donc pas appropriées pour une extraction à grande échelle de YAT ou de DPT. Ces résultats peuvent s'expliquer par l'utilisation de ces intermédiaires métaboliques *in planta* pour la synthèse de différents produits finaux (ex: PPT chez Podophyllum), mais aussi par une moindre accumulation engendrée par les conditions de culture de ces plantes. A l'inverse, les extraits réalisés à partir des feuilles de Thuja, Thujopsis et Callitris présentent une quantité notable de YAT, avec environ 6 mg/gMF, et des quantités relatives de DPT du même ordre de grandeur, tandis que les feuilles de Genévrier, les tiges de Bursera et les racines du Cerfeuil présentent des quantités bien plus importantes de ces composés. En particulier, l'extrait de feuille de genévrier (Juniperus) dispose de la plus grande quantité de YAT accumulée parmi les espèces testées, avoisinant les 30,4 mg/gMF. En outre, les quantités relatives en DPT sont plus importantes dans les extraits de tige de Bursera et de racine de cerfeuil. Sur la base de ces résultats, il apparaît donc que les feuilles de genévrier représentent la matière végétale la plus intéressante pour la mise en place d'un protocole d'extraction et de purification de YAT à grande échelle. Afin d'affiner ces résultats, des extractions au méthanol ont été réalisées de façon similaire sur plusieurs espèces du genre Juniperus pour déterminer les espèces les plus intéressantes. Des extractions ont donc été réalisées à partir de feuilles de *Juniperus virginiana*, *Juniperus communis*, *Juniperus scopulorum*, *Juniperus sabina*, *Juniperus squamata*, *Juniperus horizontalis*, *Juniperus procubens* et *Juniperus x* media. Les résultats obtenus présentés **Figure**

86 montrent que les quantités relatives en YAT et DPT dans les feuilles de genévrier sont très variables en fonction des espèces. Ainsi, *J. virginiana*, *J. communis* et *J. squamata* accumulent nettement moins de YAT que *J. scopulorum*, *J. sabina*, *J. horizontalis*, *J. procubens* et *J. x media* (33.8 mg de YAT/gMF en moyenne) suggérant que ces dernières soient les espèces les plus appropriées pour la mise en place d'un protocole d'extraction de YAT à partir de matériel végétal. Toutefois il est important de préciser à ce stade que la sélection des espèces végétales a uniquement été faite sur un critère d'accumulation maximale de YAT sans prendre en compte les modalités de mise en culture, la vitesse de croissance ou le port de la plante qui peuvent drastiquement influencer la quantité finale de YAT/DPT purifiée. A titre d'exemple, l'espèce *J. horizontalis* présente, comme son nom l'indique, un port étalé ne facilitant pas la mise en place de grandes zones de culture de cette espèce pour la production finale d'étoposide. A l'inverse, une espèce comme *J. scopulorum* de port élancé et de grande taille s'avère plus adaptée pour ce type de culture. A terme, il pourrait également être envisageable, en plus de mettre en place des zones de culture, de valoriser les déchets issus de l'élagage chez les particuliers ou les entreprises disposant d'espaces verts contenant l'espèce d'intérêt telle que *J. scopulorum* souvent cultivée dans les espaces verts. Toutefois, des problèmes de traçabilité devront être pris en compte car ils peuvent limiter l'utilisation de ce type de source dans une perspective industrielle.

4.1.2 Mise en place d'un protocole d'extraction et de purification de la YAT

Une fois les espèces végétales les plus intéressantes sélectionnées, des premiers essais de purification de YAT ont été réalisés sur le site d'Axyntis. Pour ce faire, les parties aériennes non ligneuses de quatre plants de *J. scopulorum* Blue Arrow ont été broyées, permettant d'obtenir 1,85 kg de matrice végétale. En vue d'une purification de YAT à partir du matériel végétal, deux extractions au méthanol successives ont été réalisées afin d'éviter une perte de YAT due à la saturation du méthanol. La **Figure 87.A** présente le chromatogramme UV obtenu en UPLC après l'extraction. De nombreux pics apparaissent sur ce chromatogramme, en accord avec le caractère non spécifique de l'extraction. Le pic principal de cet extrait a un temps de rétention (19min), et un spectre de masse correspondant à la YAT, indiquant que la YAT est bien le lignane majoritaire extrait à partir de la partie verte de *J. scopulorum* Blue Arrow.



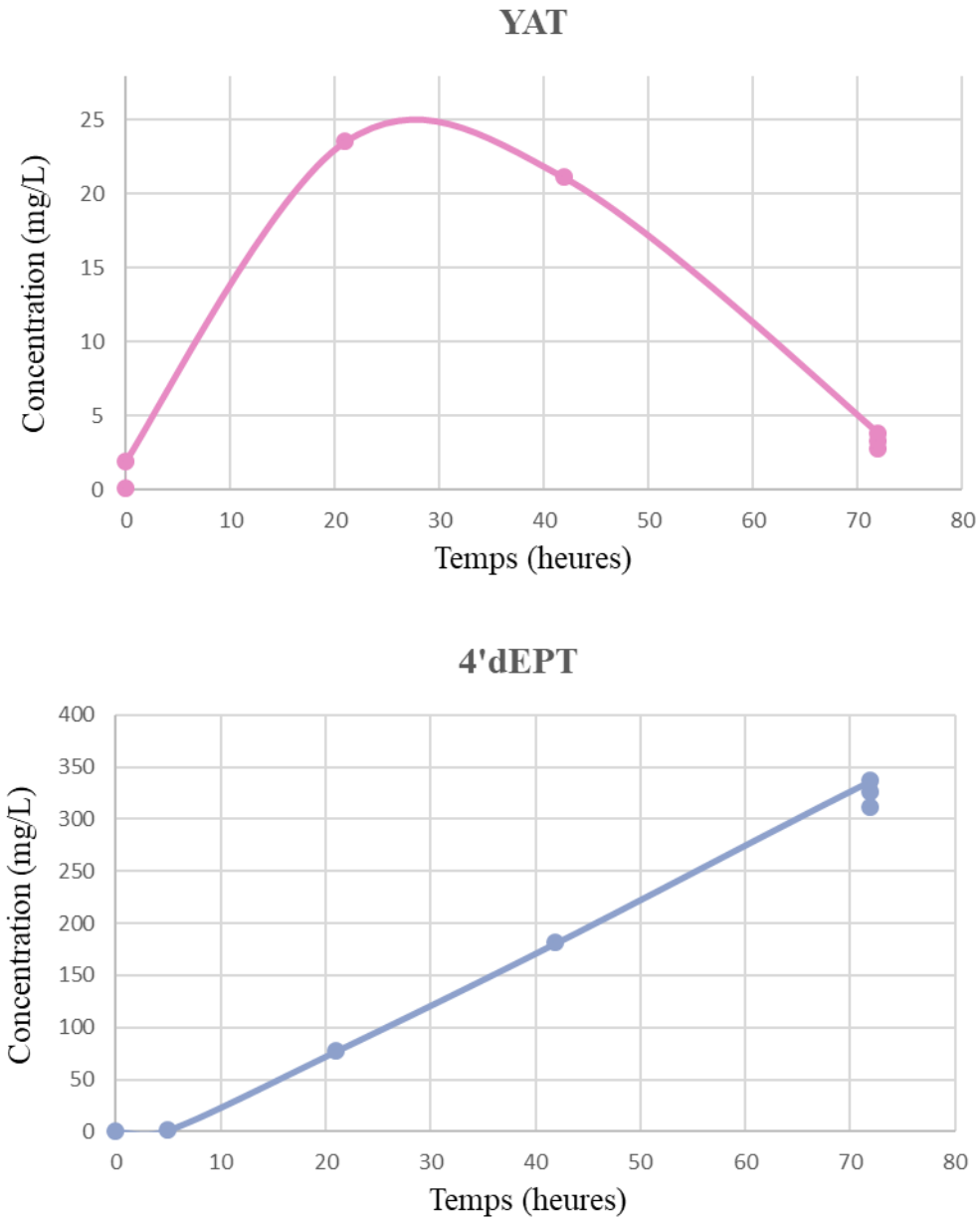
<u>Figure 87</u> : Evolution de la purification de YAT à partir d'extrait de *J. scopulorum*.

Chromatogrammes UV (290nm) obtenus par UPLC-PDA-MS après extraction (A), première purification (B) et seconde purification (C) de YAT à partir de *J. scopulorum* Blue Arrow. Gradient linéaire de 20 minutes. Les flèches en rouge indiquent le pic correspondant à la YAT sur le spectre UV.

Une première étape de purification a ensuite été réalisée à partir de l'extrait obtenu sur colonne de silice avec une phase mobile acétate d'éthyle/heptane. Les fractions d'intérêt contenant de la YAT ont été assemblées et distillées à sec, permettant l'obtention d'un premier résidu sec de 14g. La **Figure 87.B** présente le chromatogramme UV obtenu en UPLC après la première étape de purification. Le pic précédemment déterminé comme étant celui de la YAT est toujours le pic majoritaire. De plus, sur la trentaine de pics présents dans le chromatogramme précédent, il n'en reste désormais plus qu'une dizaine. Cette première étape de purification a donc permis d'éliminer les ⅔ des molécules présentes dans le premier extrait, tout en gardant une concentration majoritaire de YAT. A ce stade, le niveau de pureté de la YAT sur la base du chromatogramme UV est estimé à 51%.

Une seconde étape de purification a été réalisée à partir de 2g du résidu sec obtenu lors de la première étape, sur colonne de silice avec une phase mobile H_2O/acétonitrile. Un extrait sec de 760 mg a alors été obtenu, dont le chromatogramme UV est représenté **Figure 87.C.** Le pic correspondant à la YAT y est très largement majoritaire, la pureté estimée de la YAT dans cet extrait est de 95%.

Finalement, les résultats issus de ce premier essai d'extraction et de purification de la YAT à partir des parties aériennes non ligneuses de *J. scopulorum* sont très encourageants pour une montée en échelle du procédé. En effet, le protocole d'extraction et les deux étapes de purification réalisées ont permis d'obtenir une fraction contenant de la YAT pure à 95% pouvant être utilisée en tant que précurseur à apporter aux levures productrices de 4'dEPT. Par la suite, le protocole de la seconde étape de purification a été appliqué aux 12g restants de résidu sec issus de la première étape de la purification pour avoir une quantité de YAT suffisante en vue de réaliser une première montée en échelle de la bioproduction en bioréacteur de 2L. Le protocole de purification de YAT mis au point est donc efficace, mais certaines étapes restent à optimiser pour une montée à une échelle industrielle, notamment l'étape de broyage de la partie verte nécessitera d'être mécanisée. Enfin, une montée des étapes de purification, à une échelle pilote, sur des colonnes contenant 5 kg de silice permettra d'avoir un aperçu des rendements de purification à échelle industrielle.



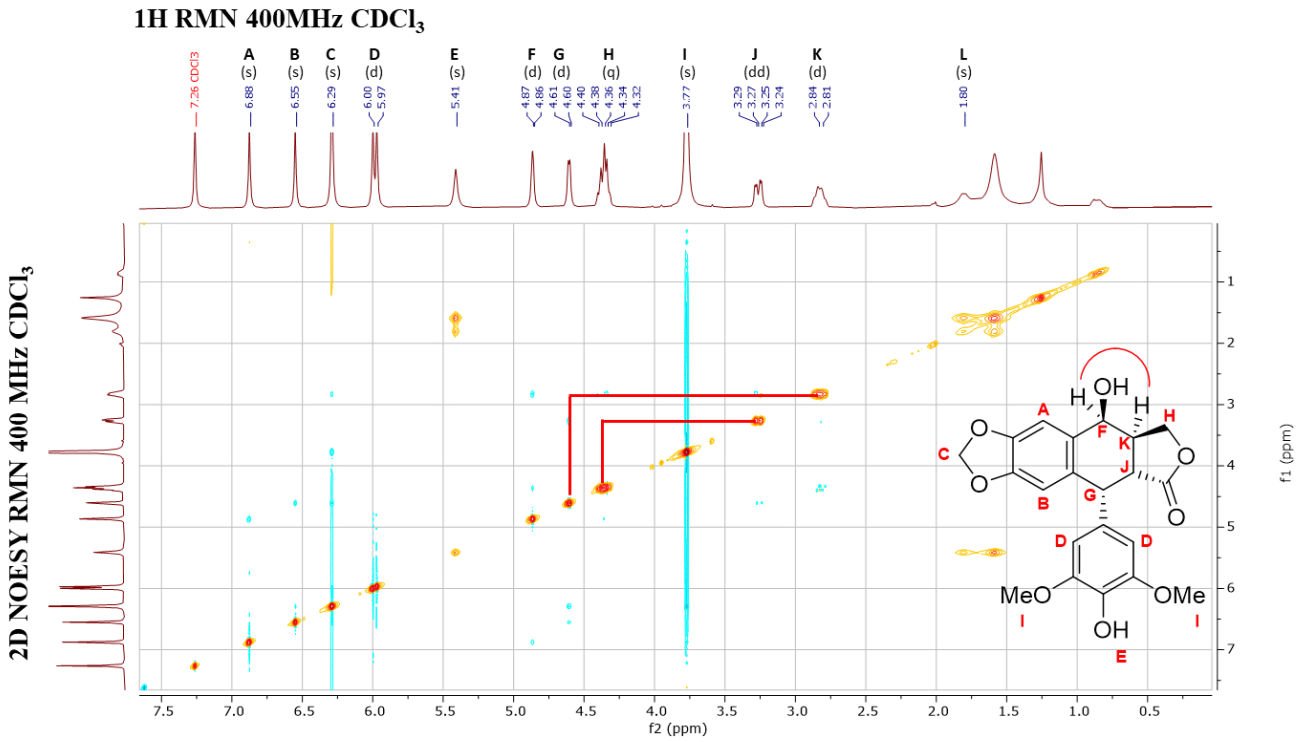
<u>Figure 88</u> : Première montée en échelle de la culture de la souche EtoA.

Mesures à partir du chromatogramme UV (290nm) des quantités de YAT et 4'dEPT dans le surnageant de culture en bioréacteur de la souches EtoA. Mesures UPLC-PDA-MS, gradient linéaire de 20 minutes.

4.1.3 Première montée en échelle de la bioproduction

Une première montée en échelle de la bioproduction de 4'dEPT a été effectuée en bioréacteur de 2L par les collaborateurs en charge des bioprocédés au laboratoire EA2106 BBV (Dr. Jillian Marc et Céline Melin) et à laquelle j'ai pu assister. Cette montée en échelle a été réalisée sur une souche stable intermédiaire nommée EtoA, générée précocement durant mon travail mais non présentée dans ce manuscrit. Dans cette souche, une cassette d'expression contenant PhCPR2 a été intégrée au locus XI-2, quatre cassettes d'expression contenant la 2ODD et la CYP71BE54 ont été intégrées aux loci XII-4, X-2, XI-5 et XI-1, et trois cassettes d'expression contenant PhCYP82D61 et CYP82opt ont été intégrées aux loci X-4, XI-3 et XII-5 du génome de la souche CEN.PK113 via le système CRISPR/Cas9. Bien qu'il s'agisse d'une souche non optimale, cette étape était primordiale car elle avait pour but, non seulement, de vérifier la robustesse en bioréacteur d'une souche productrice de 4'dEPT contenant de nombreuses copies des gènes de la voie de biosynthèse, mais surtout, de produire des quantités de 4'dEPT suffisantes afin de réaliser une analyse RMN de sa structure.

Cette culture en bioréacteur a été réalisée pendant 72 heures selon un mode de culture fed-batch avec ajout continue de milieu de culture YPD à pH 6, de glucose, et de YAT. Au total, 2,22g de YAT ont été ajoutés de façon continue durant les 21 premières heures au milieu de culture. Les variations au cours du temps des quantités de YAT et 4'dEPT ont été suivies en analysant le surnageant de la culture en UPLC (**Figure 88**). La quantité de YAT dans le milieu de culture augmente progressivement jusqu'à environ 21 heures de culture, correspondant à l'ajout de YAT dans le milieu, puis diminue au cours du temps jusqu'à être presque entièrement consommée à 72 heures de culture, suggérant que la YAT est bien métabolisée par les 2ODD. En parallèle, la quantité de 4'dEPT augmente progressivement au cours du temps, jusqu'à atteindre les 324±13 mg/L, correspondant à une masse totale de 454±13 mg de 4'dEPT produite en 72 heures pour les 1,5 litre de culture. Bien qu'il ne s'agisse que d'une faible conversion de YAT (environ 20%), largement explicable par le caractère préliminaire de la souche EtoA, ces travaux ont ainsi permis d'obtenir une quantité suffisante de 4'dEPT pour en réaliser une analyse structurale. Le milieu de culture contenant la 4'dEPT a été envoyé chez Axyntis pour réaliser une purification semblable à celle réalisée pour la YAT. Un échantillon contenant la 4'dEPT purifiée à partir du milieu de culture a ensuite été analysé par spectrométrie de Résonance Magnétique Nucléaire (RMN) à l'ICOA (UMR7311) partenaire du projet ETOPOCentre. Cette analyse a ainsi confirmé la stéréochimie de la 4'dEPT produite



<u>Figure 89</u> : Analyses par RMN de la structure de la 4'dEPT produite dans le milieu de culture.

(configuration β), validant définitivement l'approche de bioconversion en levure pour la production de précurseurs d'étoposide (**Figure 89**).

4.2 Conclusion

La recherche de plantes accumulatrices de YAT, réalisée à partir de l'analyse des extraits méthanoliques de plusieurs espèces végétales, a permis de mettre en évidence que le genre Juniperus est celui qui accumule le plus de YAT. Au sein de ce genre, l'espèce *J. scopulorum* a été sélectionnée car elle présente le meilleur taux d'accumulation. Cette espèce présente également une croissance verticale et est retrouvée dans de nombreux espaces verts, ce qui la rend très accessible en vue de récoltes à grande échelle. De même, elle présente une facilité de culture en parcelle ainsi que la possibilité de mise en valeur des déchets d'élagage. L'ensemble de ces propriétés a orienté le choix de cette espèce pour la mise en place d'un protocole d'extraction et de purification de YAT.

Un premier protocole a été mis en place afin d'extraire et purifier de la YAT à partir de 1,85kg de biomasse issue de plants de *J. scopulorum* Blue Arrow. Ce protocole, basé sur une extraction à chaud et aux ultrasons, a permis d'obtenir, après deux étapes de purification successives sur colonne de silice, un extrait de 760 mg contenant de la YAT avec un taux de pureté d'environ 95%. Outre la nécessité de mécaniser l'étape de broyage de la matière végétale, le protocole d'extraction/purification de la YAT mis en place dans ces travaux peut aisément être transféré à une échelle pilote, avec des purifications sur des colonnes contenant 5 kg de silice.

Ainsi, l'obtention d'une quantité importante de YAT a permis de réaliser une première montée en échelle en bioréacteur de la culture de la souche EtoA productrice de 4'dEPT à partir de YAT. Cette souche n'est pas la souche optimale sélectionnée pour une utilisation à l'échelle industrielle, mais elle a permis de vérifier la robustesse en bioréacteur d'une souche transformée avec un nombre important de gènes codant les enzymes de la voie de biosynthèse de la 4'dEPT, et d'initier l'étape chimique d'obtention d'étoposide via la validation de sa structure par analyse RMN.

4.3 Matériel et méthodes

4.3.1 Extraction de la YAT à partir de matière végétale

Afin d'extraire la YAT à partir de matière végétale, la partie verte des plants de *J. scopulorum* Blue Arrow a été broyée au mortier et au pilon après congélation à l'azote liquide. Deux extractions successives ont ensuite été réalisées sur la matrice végétale obtenue par ajout de 3L de méthanol 100% pour 1kg de matrice végétale. Pour chaque extraction, le mélange méthanol/matrice végétale a été chauffé à 50°C pendant 4 heures dans un bain à ultrasons. Les extraits obtenus ont ensuite été distillés à sec.

4.3.2 Purification de la YAT

Deux étapes de purification ont été mises en place à partir de l'extrait obtenu après extraction au méthanol. La première étape a été réalisée sur colonne de 300 g de silice Chromatorex Si GS60 20/45µm avec une phase mobile comprenant 20% d'acétate d'éthyle et 80% d'heptane. Les différentes fractions issues de cette première étape ont été récupérées et analysées par HPLC (High Performance Liquid Chromatography) et CCM (Chromatographie sur couche mince) afin de vérifier la présence de la YAT. Les fractions d'intérêts ont été distillées à sec afin de réaliser une seconde étape de purification sur une colonne de 300 g de silice Chromatorex C18 SPS100-10THE avec une phase mobile comprenant 60% d'H_2O + 0.1% d'acide formique et 40% d'acétonitrile. Comme décrit pour la première étape, les différentes fractions ont été récupérées et analysées par HPLC et CCM puis les fractions d'intérêt contenant la YAT ont été regroupées et distillées à sec.

CONCLUSIONS GENERALES

Les métabolites spécialisés (MS) issus de plantes médicinales représentent encore une source précieuse pour le traitement de nombreuses maladies (Favarin et al., 2013 ; Erhirhie et al., 2021). Nombre d'entre eux sont accumulés en de très faibles quantités dans ces plantes alors qu'ils en sont toujours extraits directement pour la synthèse de médicaments. De ce fait, la pression imposée sur les ressources végétales par une demande croissante des industries occasionne des pénuries d'approvisionnement en MS d'intérêt. De plus, l'épuisement des ressources altère notre environnement déjà fragilisé. Il est donc nécessaire de trouver des alternatives de production de ces molécules précieuses. La synthèse chimique des MS est l'une de ces possibilités, bien qu'elle s'avère souvent complexe, coûteuse et limitée en raison de leur complexité structurale. L'ingénierie métabolique des voies de synthèses végétales en système hétérologue propose une autre alternative pour la production de MS végétaux par d'autres organismes. Cette approche est plus appropriée à la complexité structurale des MS et plus en phase avec des stratégies de développement durable. En effet, les cellules « usines » présentent aujourd'hui des rendements de production qui deviennent compatibles avec une utilisation industrielle tout en conservant des processus d'extraction relativement simples (Guirimand et al., 2021). La levure *S. cerevisiae* est ainsi utilisée pour le développement d'approches de production de MS végétaux. En effet, cet organisme reconnu comme GRAS, est robuste et facile à cultiver dans de gros volumes. De plus, son génome est largement connu et *S. cerevisiae* est un organisme eucaryote, facilitant ainsi le clonage de gènes d'origine végétale nécessitant d'être épissés et dont les produits d'expression nécessitent parfois d'être localisés à un organite défini pour être fonctionnels (Goffeau et al., 1996). Plusieurs outils permettent désormais d'assurer la construction de souches de levure « usines » productrices de molécules d'intérêt. Parmi les procédures amonts nécessitant l'élucidation des voies de biosynthèse des métabolites à transférer en levure, le développement d'outils d'identification des gènes et de validation fonctionnelle des gènes végétaux candidats, tel que la technique VIGS (Lu et al., 2003), permet d'initier sereinement l'ingénierie métabolique en fournissant les gènes à intégrer aux levures. De plus, le développement croissant de mécanismes de transfert efficaces de gènes en levure permet désormais l'insertion de multiples gènes hétérologues à des « hot-spots » du génome de la levure (Geu-Flores et al. 2007 ; Gibson et al., 2009 ; Mikkelsen et al., 2012 ; Zhang et al., 2020). En outre, des stratégies d'optimisation des cellules usines ne cessent de se développer, permettant de réguler de plus en plus finement l'expression génique au niveau transcriptionnel (Al-Husini et al., 2020), d'optimiser la traduction (Lanza et al., 2014), de réguler le métabolisme primaire de la levure et le recyclage des cofacteurs essentiels aux enzymes hétérologues (Chen et al., 2022), mais aussi de contrôler la localisation de ces enzymes et de contourner les fuites

métaboliques induisant des réductions de rendement. Enfin, les stratégies adaptatives d'évolution permettent dorénavant de sélectionner des souches plus robustes et plus adaptées à la production de certains MS (Luo et al., 2019). Fort de ces différentes connaissances, il est désormais possible de proposer le développement de stratégies de bioproduction, basées sur la création de cellules usines de levure pour la production de MS en tension d'approvisionnement tel que l'étoposide, et selon une échelle de temps compatible avec un développement industriel.

L'étoposide est un lignane utilisé en chimiothérapie depuis de nombreuses années (Meresse et al., 2004). Cette molécule pharmaceutique est obtenue par synthèse chimique à partir de PPT (Sinkule, 1984). Le sourcing de PPT ne dépend actuellement que de l'extraction à partir du rhizome de *P. hexandrum*, plante menacée d'extinction, ce qui explique les ruptures d'approvisionnement récurrentes qui ont été observées pour l'étoposide. De ce fait, nous avons initié, au cours de cette thèse, le développement d'une approche alternative de synthèse d'un précurseur d'étoposide par ingénierie métabolique en levure afin de sécuriser l'accès à ce composé via la relocalisation de sa production en Région Centre-Val de Loire. Nous avons ainsi opté pour une bioproduction de 4'dEPT, molécule précédant la PPT dans la voie de biosynthèse car celle-ci, déjà déméthylée en position C-4', ne nécessite qu'une étape d'hémisynthèse pour la synthèse d'étoposide. Cette bioproduction de 4'dEPT s'effectue par bioconversion de YAT, un précurseur hautement biodisponible dans des espèces végétales cultivables en Région Centre (ex: Juniperus). Elle ne nécessite qu'un transfert limité de gènes de la voie de biosynthèse car la conversion de la YAT en 4'dEPT n'implique que trois étapes catalysée par une dioxygénase 2ODD et deux cytochromes P450 CYP71BE54 et CYP82D61 tel que décrit chez *P. hexandrum* (**Figure 57**) (Lau and Sattely, 2015). Ce transfert restreint de gènes a également été choisi de façon à maximiser les rendements de bioconversion en limitant les risques de détournements d'intermédiaires métaboliques, mais aussi pour pouvoir être réalisable sur les trois années de travail de cette thèse. Néanmoins, avant d'initier ces travaux relatifs à l'étoposide, j'ai choisi de m'initier aux approches d'ingénierie métabolique en levure via le métabolisme des THCSpd alors en cours à mon arrivée au laboratoire. En effet, les THCSpd sont des phénolamides disposant de propriétés biologiques très intéressantes du point de vue pharmaceutique (Ma et al., 2001 ; Yingyongnarongkul et al., 2008 ; Zhao et al., 2010 ; Park et al., 2017 ; Kyselka et al., 2018 ; Miyata et al., 2022). Toutefois, ces composés ne s'accumulent que dans le pollen des Dicotylédones, réduisant ainsi l'accès aux connaissances sur ces molécules précieuses. De ce fait, le laboratoire souhaitait mettre en place des souches de levures productrices de différents THCSpd à partir de divers acides hydroxycinnamiques via l'expression d'orthologues de la

SHT et de la 4CL en levure. Une telle approche permettrait, en effet, de vérifier la possibilité de production et d'accès à la diversité de ces molécules en levure et de fournir, à terme, des quantités compatibles avec une étude plus approfondie de leur propriétés pharmacologiques et modes d'action.

Pour mener à bien mes travaux, deux systèmes différents ont été utilisés afin d'exprimer les gènes végétaux chez *S. cerevisiae* : les plasmides auto-réplicatifs et l'intégration chromosomique. Dans un premier temps, l'utilisation des plasmides auto-réplicatifs pour l'expression des gènes hétérologues végétaux a permis de développer des souches de levure productrices de THCSpd naturels et nouveaux. Ces souches, dont les taux de biosynthèse restent relativement bas (au maximum 20mg/L), pourront néanmoins servir de plateformes pour la production et le screening de THCSpd disposant potentiellement d'un intérêt pharmaceutique. De façon notable, l'emploi du système des plasmides auto-réplicatifs a permis de contrôler les décorations ajoutées à la spermidine car une construction de la souche par intégration stable des gènes hétérologues aurait compliqué la synthèse de THCSpd « exotiques ». En effet, plusieurs combinaisons d'isoformes de 4CL d'*A. thaliana* et d'orthologues de SHT peuvent être employées pour l'acylation de la spermidine en fonction de/des acide(s) hydroxycinnamique(s) ajouté(s) au milieu de culture. De plus, l'ajout des acides hydroxycinnamiques dans le milieu de culture permet de facilement contrôler la nature des décorations ajoutées à la spermidine. Ces travaux, valorisés par la publication d'un article dans la revue ACS Synthetic Biology, m'ont ainsi fourni les bases théoriques et pratiques nécessaires à l'application d'une telle stratégie pour la bioproduction de lignanes, but principal de mes travaux de thèse.

La stratégie d'expression de gènes hétérologues basée sur l'utilisation de plasmides auto-réplicatifs a été appliquée à la production de 4'dEPT. Ce système a ainsi permis de vérifier l'activité des enzymes de la voie de synthèse de la 4'dEPT en levure et de trouver une solution pour augmenter l'activité catalytique de la dernière enzyme de la voie, CYP82D61. En effet, l'activité catalytique de cette enzyme étant initialement très faible en levure, plusieurs stratégies ont été testées afin d'améliorer le rendement de cette dernière étape enzymatique. En particulier, les activités catalytiques d'une isoforme et de plusieurs orthologues de CYP82D61, ainsi que de la CYP3A4 humaine ont été vérifiées sur la 4'dDPT en levure. De plus, le système de plasmides auto-réplicatifs nous a permis de moduler la structure de CYP82D61, en particulier, son domaine transmembranaire d'ancrage à la membrane du réticulum endoplasmique qui a été substitué par celui de cytochromes P450 connues pour être actives en levure, et CYP82D61 a été fusionnée à la cytochrome P450 réductase (CPR) d'*A. thaliana* (AtATR2) ou de *S.*

cerevisiae (ScCPR2). Ces modulations structurales ont été mises en place afin de pallier un potentiel problème d'ancrage de la protéine à la membrane du RE de la levure pouvant altérer l'interaction entre le P450 et la CPR qui résulterait en un défaut de réduction du P450. Alors qu'aucune de ces approches n'a permis d'augmenter significativement l'activité de CYP82D61, c'est, finalement, l'expression des CPR de *P. hexandrum* (PhCPR2) et de *C. roseus* (CrCPR2) à la place d'AtATR2 qui a permis d'établir que cette faible activité était vraisemblablement liée à un défaut de réduction par AtATR2. Cette question de spécificité d'interaction/réduction des P450 par les CPR a d'ailleurs été reprise dans de nombreux travaux. Ainsi, en sélectionnant les CPR permettant de réduire efficacement les P450 CYP450 et CYP72A63 issus de la réglisse, les titres les plus élevés de biosynthèse des triterpénoïdes acide glycyrrhétinique (18,9 ± 2,0 mg/L) et 11-oxo-β-amyrine (108,1 ± 4,6 mg/L) ont été obtenus en levure (Zhu et al., 2018). De même, la réduction de la P450 CYP720B1 par la CPR issue de *Taxus cuspidate* a permis d'augmenter les rendements de biosynthèse d'acide lévopimarique (23,13 mg/L) par rapport à sa réduction par AtATR1 ou AtATR2 (Liu et al., 2018).

Au-delà d'apporter une réponse biologique à la faible activité de CYP82D61 en système hétérologue, l'expression d'une CPR adaptée a été le premier élément permettant une véritable augmentation de la conversion de YAT en 4'dEPT. Il est ainsi fort possible que la faible activité de CYP82D61 rapportée par Lau et Sattely (2015) par expression hétérologue chez *N. benthamiana* découle également d'une faible capacité de la CPR de *N. benthamiana* à réduire ce P450. De façon plus remarquable, c'est la conduite de ces travaux et l'emploi des plasmides auto-réplicatifs qui ont établi la preuve de concept de la production de 4'dEPT en cellules usines de levure, et c'est, à mon sens, l'un de ses avantages principaux. Il est à noter que cette approche s'avère particulièrement rapide à mettre en œuvre tout en obtenant des taux de transformation élevés à chaque étape de la procédure. Toutefois, cette approche n'est pas exempte de certaines limites. Ainsi, le nombre de copies de plasmides peut varier d'une cellule à l'autre et ce phénomène peut être d'autant plus amplifié que plusieurs plasmides portant des marqueurs distincts sont transformés dans une même levure (ex. cas des plasmides de la série pESC employés dans le premier et second chapitre de cette thèse). Cette variation peut, de ce fait, entraîner des combinaisons différentes de profil d'expression des enzymes hétérologues dans les cellules transformées avec plusieurs plasmides (Albertsen et al., 2011). Par exemple, 95% des cellules de levure ont des niveaux d'expression similaires des protéines Yellow Fluorescent Protein (YFP), Red Fluorescent Protein (RFP) et Cyan Fluorescent Protein (CFP) lorsque les gènes codant sont intégrés dans le génome, tandis que seulement 6% des cellules possèdent des

niveaux identiques d'expression d'YFP, RFP et CFP lorsque les gènes codant son co-exprimés dans différents plasmides (Jensen et al., 2014). En outre, l'emploi de plasmides auto-réplicatifs peut entraîner des problèmes de stabilité génétique dans les souches ainsi créées dans la mesure où les levures peuvent perdre les plasmides transférés au cours de la culture. A titre d'exemple, seuls 28% des cellules de levure ont maintenu les deux plasmides contenant les gènes codant l'amorphadiène synthase (ADS) et l'amorphadiène oxydase (AMO) impliquées dans la synthèse d'artéminisine initialement clonés après 120 heures de culture (Ro et al., 2008). Enfin, l'utilisation de plasmides entraîne une faible reproductibilité et une grande variabilité des taux de production entre les cellules, contrairement à l'intégration génomique qui présente une forte reproductibilité (Isogai et al., 2021). La synthèse de benzylglucosinolate dérivé de la phénylalanine est, par exemple, 8,4 fois plus élevée lorsque les gènes codants sont intégrés dans le génome de la levure que lorsque les gènes sont clonés dans des plasmides malgré un niveau d'expression plus faible des protéines (Wang et al., 2022).

Ainsi, dans un second temps, et fort des premiers résultats positifs obtenus avec le système des plasmides auto-réplicatifs, nous avons choisi de transférer les gènes assurant la conversion de YAT en 4'dEPT par intégration dans le génome des levures. Cette intégration chromosomique a, plus spécifiquement, été réalisée dans le génome de la souche industrielle *Saccharomyces cerevisiae* CEN.PK113-7D via le système CRISPR/Cas9. En plus de garantir une stabilité génétique plus grande des cellules ainsi générées, cette approche permet de facilement ajuster le ratio des copies de gènes transférés pour l'adapter le plus finement possible au flux métabolique recherché. C'est d'ailleurs cette stratégie qui a permis d'optimiser les rendements de bioconversion de YAT en 4'dEPT grâce à l'insertion de versions optimisées pour la levure de CYP82D61 et la multiplication du nombre de copies des différentes enzymes de la voie. La souche Eto7.1 finale sélectionnée produit 114,7±4,4 mg/L de 4'dEPT à partir de 180 mg/L de YAT, soit un rapport de bioconversion de l'ordre de 65%. Ce taux de bioconversion est légèrement inférieur à celui obtenu par la souche de levure de Kulagina et al. (2021c) capable de bioconvertir 88% de la tabersonine ajoutée au milieu de culture en vindoline. Toutefois, cette souche a été optimisée, non seulement via une régulation de la balance du nombre de copies des gènes, mais aussi via une optimisation de la composition du milieu de culture. Ces résultats laissent donc à penser que les rendements de bioconversion de YAT en 4'dEPT par la souche Eto7.1 pourraient largement atteindre les 90% lorsque la conduite de la culture en plus gros volume sera optimisée. En plus d'une optimisation des rendements de conversion, cette approche, basée dans notre cas sur l'emploi de la technique

CRISPR/Cas s'avère particulièrement avantageuse dans l'optique d'une production industrielle car elle ne nécessite pas de sélection, contrairement aux autres approches d'intégration ou de l'emploi des plasmides auto-réplicatifs, comme j'ai pu le décrire dans la section Introduction. Cette absence d'éléments de sélection (auxotrophie ou antifongique) est souvent garante d'une absence d'altération des taux de croissance des levures malgré une intégration d'un grand nombre de copies des gènes exogènes, et s'avère, en outre, être un prérequis à une utilisation industrielle à terme. Ces avantages s'accompagnent d'une meilleure reproductibilité de résultats et d'une plus grande stabilité génétique tel que décrit précédemment.

Néanmoins, la procédure de transfert de gènes par intégration CRISPR/Cas9 peut s'avérer chronophage dans sa mise en place, en particulier si elle nécessite de démultiplier le nombre de copies des gènes hétérologues. C'est, en particulier, le cas avec le système plasmidique que j'ai utilisé (Partie 3 des Résultats) car ceux-ci ne permettent l'intégration que de deux gènes par loci. Ceci s'avère être une problématique à prendre en compte car cela limite considérablement le nombre de gènes total à transférer, compte tenu du nombre fini de hot-spots utilisables, et augmente, en retour, le temps de préparation des constructions moléculaires et d'obtention des souches transformées. Néanmoins, dans ce cas de figure, l'emploi de méthodes de clonage alternatives telles que les assemblages Gibson ou USER décrits dans la section Introduction peuvent permettre de pallier cette limite en augmentant le nombre de gènes intégrés à chaque locus. D'autres approches ont également été développées afin de faciliter des insertions multiples. C'est notamment le cas d'une méthode basée sur l'utilisation des éléments transposables (transposons) comme loci d'intégration permettant le clonage de gènes multicopies dans la levure (Wang et al., 2018). En effet, les transposons se déplacent de manière autonome dans le génome de la levure (transposition) grâce à une transposase capable de réaliser des coupures double brin dans l'ADN. L'utilisation du système CRISPR/Cas9 ciblant ces éléments transposables permet alors l'insertion de cassettes de gènes aux loci du génome où se déplacent les transposons. Cette méthode a permis l'intégration de plus de 10 copies du même gène en seule une étape (Wang et al., 2018). De façon similaire, le système « CasEMBLR », également présenté en Introduction, permet l'assemblage de plusieurs gènes hétérologues dans le génome de la levure en une seule étape de transformation grâce à l'utilisation de plasmides « tout-en-un » contenant plusieurs ARN guides (Jakočiunas et al., 2018). Il sera ainsi indispensable de considérer ces approches plus efficientes en termes d'assemblage et de clonage pour assurer un développement de souches performantes en terme de synthèse des MS mais aussi en un laps de temps acceptable.

Bien que chacune des deux approches employées pour l'expression des gènes hétérologues présente ses avantages et limites respectifs, leur combinaison peut également être une alternative à considérer pour la production à une échelle industrielle de MS d'intérêt par ingénierie métabolique en levure. En effet, le nombre de copies des gènes hétérologues apporté via les plasmides auto-réplicatifs peut être directement plus élevé que le nombre obtenu avec des intégrations chromosomiques. Il peut alors être considéré d'exprimer l'entièreté d'une voie de biosynthèse de MS via des intégrations chromosomiques et d'apporter des copies complémentaires, via des plasmides auto-réplicatifs, notamment des gènes codant des enzymes constituant des goulots d'étranglement de la voie ou celles en compétition avec d'autres enzymes de détournement des intermédiaires biosynthétiques. Cette association de méthodes permet d'optimiser les rendements de conversion, en s'affranchissant d'étapes de transformation chronophages tout en limitant le nombre de marqueurs de sélection ayant des impacts négatifs sur la croissance des cellules et le coût des milieux de culture. La production récente *de novo* de l'AIM strictosidine, précurseur des anticancéreux vinblastine et vincristine, en est un bon exemple (Zhang et al., 2022). Cette souche, contenant 34 gènes hétérologues végétaux, est capable de produire 25,2 mg/L de strictosidine grâce au clonage par réplication autonome des gènes codant les enzymes limitantes dans la voie de synthèse.

L'apport de gènes de synthèse optimisés dans les voies de biosynthèse hétérologues en levure présente, en outre, un fort intérêt de par leur facilité d'obtention et l'avantage qu'ils apportent afin d'optimiser la traduction des gènes codés dans les organismes hétérologues. En effet, comme présenté en Introduction, le biais d'usage du codon peut influencer l'efficacité de traduction des protéines hétérologues. Ainsi, le clonage de gènes de synthèse disposant de codons optimisés pour l'organisme hôte peut permettre d'augmenter les rendements de biosynthèse. Nous en avons fait le constat, notamment, lorsque les gènes de synthèse codant des versions codons-optimisés de CYP82D61 ont été clonés dans la souche Eto3, nous permettant de doubler les rendements de biosynthèse de 4'dEPT par rapport à la souche contenant la version native du gène codant CYP82D61. Ces résultats sont en adéquation avec plusieurs exemples de la littérature. Notamment, la première production de resvératrol en levure a été permise par l'optimisation pour la levure de la tyrosine ammonia lyase (TAL) issue de *Rhodobacter* permettant la synthèse de l'acide *p*-coumarique à partir de la tyrosine dans la voie de synthèse du resvératrol (**Figure 25**) (Wang et al., 2011). De même, les rendements de biosynthèse en levure de l'acide muconique ont été multipliés par 2,9 lorsque la séquence

codant la catéchol 1,2-dioxygénase d'*Acinetobacter baylyi* a été optimisée pour la levure (Lanza et al., 2014).

Comme indiqué précédemment, la stratégie retenue dans mon travail de thèse s'appuie sur une approche de bioconversion de 4'dEPT à partir de YAT hautement biodisponible. Cette approche permet, en effet, de diminuer le nombre d'étapes enzymatiques, et donc de mettre en place plus rapidement la souche productrice, mais elle permet aussi d'augmenter les rendements de biosynthèse. En effet, une synthèse *de novo* impliquant un nombre important d'étapes enzymatiques hétérologues, si chacune de ces étapes ne convertit pas la totalité du substrat, le rendement final de biosynthèse sera, de fait, nettement impacté. Le rendement de biosynthèse de la vindoline en levure a, par exemple, été multiplié par 10 lorsque la levure convertit la tabersonine ajoutée au milieu de culture par rapport à la souche productrice de vindoline *de novo*. Précisément, la souche convertissant la tabersonine via 7 étapes enzymatique est capable de produire 266 mg/L de vindoline (Kulagina et al., 2021c), tandis que seulement 25,2 mg/L de strictosidine, précurseur de la vindoline, ont été produits par la souche *de novo*, avec le clonage de 34 gènes hétérologues (Zhang et al., 2022). De même, la production *de novo* de vanilline en levure est de seulement 700 µg/L après intégration des 8 gènes exogènes impliqués dans la synthèse de vanilline, de la 4-coumarate 3-hydroxylase issue de *Saccharothrix espanaensis* afin d'augmenter la production d'acide caféique, et la mutation de deux gènes endogènes pour réduire la rétroaction négative de synthèse du précurseur tyrosine (Qiu et al., 2022). En plus de diminuer le temps de mise en place des souches, et d'augmenter les rendements de biosynthèse, la stratégie de bioconversion permet d'envisager la biosynthèse de molécules nouvelles via l'ajout de substrats exotiques, tels que les substrats halogénés comme présenté au paragraphe I.3.c.8 de l'Introduction (Li et al., 2018).

Néanmoins, cette approche nécessite l'apport d'un précurseur d'origine végétal, ainsi, la biosynthèse est donc toujours dépendante des plantes, et il n'est pas toujours facile d'extraire et de purifier des quantités suffisantes pour une production industrielle. C'est donc pour cette raison qu'un effort important a été consacré à la recherche des espèces végétales pouvant être les meilleures sources de YAT. Ainsi, la recherche d'espèces végétales accumulatrices de YAT a été effectuée sur la base de l'analyse métabolique d'extraits méthanoliques réalisée chez différentes plantes de la famille des Cupressaceae connues pour accumuler la YAT (Hwang et al., 2004). Cette approche a permis de mettre en évidence une accumulation accrue de YAT chez les espèces du genre Juniperus, et en particulier chez *J. scopulorum*, *J. sabina*, *J. horizontalis*, *J. procubens* et *J. x media*. Un protocole d'extraction au méthanol des lignanes

présentes chez *J. scopulorum* a donc été mis en place. Ce protocole s'inspire des travaux réalisés sur le lin afin d'extraire de la PPT et son dérivé, le 6-méthoxypodophyllotoxine (Kartal et al., 2004). Dans ces travaux, les deux molécules d'intérêt ont été extraites selon trois protocole différents : soit par extraction aqueuse (extraction liquide-liquide à l'aide d'acétate d'éthyle); soit par extraction au méthanol par chauffage et bain ultrasons; soit par extraction à l'acétone, puis extraction liquide-liquide à l'aide d'acétate d'éthyle. Même si ces trois méthodes sont utilisables pour l'extraction, elles peuvent être utilisées différentiellement en fonction des objectifs. Ainsi, la deuxième méthode permet d'extraire de façon plus efficace les lignanes de type PPT. C'est pourquoi nous avons décidé de nous focaliser sur celle-ci pour nos travaux. Deux étapes de purification successives à l'acétate d'éthyle et à l'heptane, puis à l'acétonitrile ont ainsi permis de purifier la YAT de l'extrait sec obtenu à 95%. La YAT obtenue permettra de réaliser une montée en échelle de la culture de la souche Eto7.1 avec des équipements présents en région Centre-Val de Loire, permettant ainsi une production entièrement locale de l'étoposide. Ce second versant de ma thèse m'a ainsi permis de me familiariser avec le travail en entreprise, avec les enjeux liés à une montée en échelle de la bioproduction et s'avère donc parfaitement complémentaire du travail de laboratoire que j'ai également conduit. J'ai pu, en effet, appréhender les contraintes liées à l'industrie, telles que les restrictions d'utilisation de certains solvants, ou encore la limitation du coût énergétique des différents procédés. Ceci m'a permis, en outre, d'observer, à travers la montée en échelle de l'extraction et la purification de YAT, les différences de procédés entre l'échelle laboratoire et l'échelle pilote/industrielle. C'est donc sur cette double formation, complémentaire sur de nombreux points, que je bâtirai la suite de ma carrière professionnelle.

En conclusion, ma thèse s'inscrivant dans le projet ETOPOCentre réunissant le groupe industriel Axyntis et les laboratoires EA2106 BBV de l'Université de Tours, EA1207 LBLGC et UMR7311 ICOA de l'Université d'Orléans, nous avons, en fin de thèse, été en mesure de transmettre des premiers lots de 4'dEPT issue de cultures en bioréacteur de la souche EtoA (Partie 4 des Résultats) à l'équipe de Méthodologie et hétérocycles de l'ICOA, spécialiste de la chimie de synthèse dirigée par le Pr. Isabelle Gillaizeau. L'équipe a pu adapter l'étape de glycosylation, normalement développée pour la PPT, à la 4'dEPT, permettant, de fait, la production d'un premier lot d'étoposide entièrement issue de la région Centre-Val de Loire à l'automne 2022. Il s'agit là d'un élément fort du projet attestant de la pertinence et du potentiel de l'approche retenue au cours de ma thèse. Ce résultat met donc un point d'honneur au travail que j'ai développé depuis plus de trois ans.

PERSPECTIVES

L'ensemble des travaux que j'ai mené durant ma thèse ont, notamment, permis de jeter les bases solides d'une production d'un précurseur d'étoposide par des approches de bioproduction de type bioconversion. En plus d'avoir généré une souche hautement efficiente dans la conversion de la yatéine en 4'dEPT, j'ai également contribué à la mise en place d'une procédure d'obtention de la yatéine via la sélection des plantes accumulatrices et le développement d'un protocole de purification efficace de ce composé. Néanmoins, ces premiers succès ne constituent qu'une première étape qui nécessitera le développement de travaux complémentaires. Ainsi, même si un premier essai de culture en bioréacteur (2 litres) de la souche Eto7.1 a confirmé sa robustesse dans la conversion de la yatéine (données non présentées), une optimisation de ce procédé de culture devra être réalisée à l'échelle du laboratoire. A l'instar des travaux menés par l'équipe pour la bioconversion de tabersonine en vindoline qui a été presque triplée grâce à une optimisation de la conduite de la culture, notamment via la détermination du pH optimal et une alimentation continue de la tabersonine et des substrats de croissance (Kulagina et al., 2021c).

Après ces optimisations à petite échelle, la souche Eto7.1 devra être cultivée en bioréacteur de grande capacité pour la production à grande échelle de 4'dEPT. La première étape de ce scale-up de production passera par une culture sur le réacteur de 50 litres disponible sur le plateau technique du Bio^3 Institute de Tours (https://bio3institute.fr/) que le laboratoire BBV emploi déjà dans ses projets de bioproduction d'alcaloïdes anticancéreux. Pour cela, il sera nécessaire, dans un premier temps, de vérifier le comportement de la souche dans de plus gros volumes de culture puis d'adapter/optimiser encore la conduite de culture à ces grands volumes afin de conserver des rendements de bioconversion. Si ces rendements venaient à être impactés, que ce soit à l'échelle du réacteur de 2 litres ou de plus grands volumes, le squelette moléculaire de la souche pourra être retravaillé afin d'optimiser les différentes étapes enzymatiques de la voie de bioconversion. Par exemple, il serait envisageable d'augmenter le nombre de copies des gènes hétérologues. Le clonage de versions optimisées pour la levure et/ou d'orthologues de la 2ODD et de CYP71BE54 pourrait permettre d'optimiser le flux métabolique en évitant l'accumulation de transcrits identiques pouvant inhiber la synthèse des protéines. De telles isoformes ont d'ores et déjà été identifiées chez d'autres espèces végétales (*Anthriscus sylvestris* ou *Juniperus* sp.) dans le cadre du projet ETOPOCentre et sont en cours de caractérisation.

Notre stratégie de bioconversion de 4'dEPT en levure nécessite de cultiver les levures en présence de YAT relativement pure, ajoutant un coût non négligeable au processus de

production de 4'dEPT. Le protocole d'extraction de YAT à partir de *J. scopulorum*, efficace à échelle laboratoire, devra donc être optimisé à échelle industrielle. Toutefois, l'accès à ces plantes de croissance lente, tout comme l'étape d'extraction actuellement chronophage et coûteuse du point de vue énergétique pourrait limiter le déploiement de notre approche. De ce fait, il sera peut-être opportun de trouver, à terme, une alternative à ce protocole d'extraction afin de minimiser les coûts et le temps. Ainsi, l'une d'entre elles pourrait reposer sur la mise en place d'une souche capable de bioconvertir de la 4'dEPT à partir d'une substrat plus facile d'accès que la YAT. A ce titre, les graines de lin représentent une source riche en lignanes, et des extractions réalisées sur cet organe ont mis en évidence l'accumulation de lignanes impliquées dans la voie de synthèse de la 4'dEPT (Sicilia et al., 2003) tels que le sécoisolaricirésinol, le matairesinol et le pinorésinol, lignanes retrouvés dans la voie de synthèse de la 4'dEPT à partir d'alcool conifférylique (Davin et al., 1997 ; Dasilay et al., 2015 ; Von Heimendahl et al., 2005 ; Xia et al., 2001 ; Marques et al., 2013 ; Lau and Sattely, 2015) (**Figure 51**). La mise en place d'une souche de levure capable de bioconvertir de la 4'dEPT à partir de ces trois substrats, composée de 9 étapes enzymatiques (catalysées par la pinorésinol/laricirésinol réductase PLR2, la sécoisolaricirésinol déshydrogénase SDH, l'activité pluviatolide synthase de CYP719A23, la pluviatolide 4-*O*-méthyltransférase OMT3, la bursehernine 5′-hydroxylase CYP71CU1, la 5'-déméthylyateine *O*-transférase OMT1, la dioxygénase désoxypodophyllotoxine synthase 2ODD et les P450 CYP71BE54 et CYP82D61), permettrait de s'affranchir des problèmes liés aux coûts d'extraction et de purification du substrat, mais aussi à la culture de plants de Juniperus. Le tourteau de lin est un coproduit de l'industrie alimentaire obtenu après extraction de l'huile des graines de lin dans lequel les lignanes d'intérêt sont accumulés. En 2014, la France a produit 23 319 tonnes de lin, ce qui représente une quantité importante de lignanes non valorisés. Utiliser ce coproduit comme substrat pour la bioconversion de 4'dEPT en levure permettrait de revaloriser les déchets de l'industrie alimentaire pour produire de l'étoposide en s'affranchissant des étapes de purification nécessaires actuellement pour l'obtention de YAT. En revanche, même si cette stratégie est moins coûteuse, l'approvisionnement en 4'dEPT, et donc en étoposide serait toujours dépendant d'une espèce végétale.

A plus long terme, la production *de novo* de 4'dEPT en levure permettrait de s'affranchir de cette dépendance. En effet, la voie de biosynthèse de la 4'dEPT à partir d'alcool conifférylique naturellement produit par la levure a entièrement été élucidée. En revanche, le clonage en levure de la protéine dirigeante DIR1 pose problème puisque celle-ci nécessite d'être

co-exprimée avec une laccase/peroxydase nécessitant une glycolysation parfois inefficace chez la levure (Antosova and Sychrova, 2015). Or, le clonage de cette enzyme est nécessaire pour optimiser la synthèse de (+)-pinorésinol puisqu'en absence de protéine dirigeante, de nombreux dimères d'alcool conifférylique sont synthétisés et le (+)-pinorésinol n'est plus le produit majoritaire (Fournand et al., 2003 ; Halls et al., 2004). Il sera donc nécessaire de co-exprimer la protéine dirigeante avec une laccase efficacement glycosylée en levure. L'élaboration d'une souche de levure capable de produire *de novo* de la 4'dEPT impliquera donc de mettre en place de nombreuses stratégies d'optimisation. Notamment, la voie de biosynthèse de la 4'dEPT à partir de l'alcool conifférylique est composée de dix étapes enzymatiques, dont 4 sont catalysées par des cytochromes P450 et 2 par des *O*-méthyltransférases, il sera donc nécessaire d'utiliser des approches d'édition du génome permettant l'assemblage de fragments d'ADN multiples. En effet, la voie de bioconversion mise en place dans ces travaux ne comporte que 3 étapes enzymatiques et nous avons pourtant cloné au total 11 gènes dans la souche Eto7.1, il est donc prévisible qu'il sera aussi nécessaire de multiplier le nombre de copies des différents gènes dans la souche productrice de 4'dEPT *de novo*. Une telle souche pourrait donc contenir un nombre élevé de gènes hétérologues. Il sera alors nécessaire de trouver de nouveaux hot-spots dans le génome de la levure pour l'intégration de gènes à d'autres loci en plus de la multiplication de gènes à chaque locus. De plus, il sera important de travailler sur le recyclage des cofacteurs. En particulier, les cytochromes P450, nombreux dans cette voie de biosynthèse, nécessitent du NADPH pour être fonctionnels, et les *O*-méthyltransférases dépendent de la SAM. La production native de ces cofacteurs sera insuffisante pour assurer l'activité des enzymes hétérologues. La surexpression des gènes codant ZWF1 et POS5 permettant de régénérer le NADPH (Paramasivan and Mutturi, 2017), ou le clonage de METK1 et la surexpression d'ACS2 et MLS1 permettant d'augmenter la production de SAM (Chen et al., 2016) sont des techniques envisageables pour améliorer la production de ces deux cofacteurs. De plus, il serait envisageable d'utiliser les souches construites par Chen et al., 2022 capables de recycler les cofacteurs SAM et NADPH comme châssis cellulaire pour la construction de la souche productrice de 4'dEPT *de novo*. Ainsi, le chemin menant à la mise en place d'une souche productrice *de novo* de 4'dEPT s'avère long et complexe. Il reste cependant totalement envisageable à l'instar des souches développées pour la production de vindoline et catharanthine (Zhang et al., 2022).

En parallèle de ces futurs travaux d'optimisation de la production d'étoposide, la recherche de THCSpd d'intérêt pharmaceutique devra être réalisée en se basant, dans un

premier temps, sur une stratégie de bioconversion couplée à une expression des enzymes employant des plasmides auto-réplicatifs. Cette approche permettra ainsi de contrôler l'ajout des décorations sur la spermidine pour la synthèse de différentes structures selon les orthologues clonés. En effet, le ratio entre chaque acide hydroxycinnamique ajouté au milieu de culture de la souche de levure productrice permet de maîtriser les décorations apportées au noyau spermidine. En revanche, une fois le criblage de structures disposant d'activités intéressantes, il pourra être envisageable de mettre en place des souches de levure productrices de THCSpd à échelle industrielle. Pour cela, il sera donc nécessaire d'intégrer les gènes codant dans le génome de la levure, comme je l'ai fait pour la production de 4'dEPT. Dans un premier temps, il pourra être envisageable de réaliser la montée en échelle de la bioproduction selon une stratégie de bioconversion à partir d'acides hydroxycinnamiques ajoutés dans le milieu de culture. Toutefois, comme pour la souche productrice de 4'dEPT, cette souche nécessitera d'être mise en culture en présence d'acides hydroxycinnamiques relativement purs, ce qui peut avoir un coût non négligeable pour l'industriel. *In fine*, il serait donc optimal de mettre en place des souches de levure capables de produire *de novo* les THCSpd d'intérêt. Des souches capables de synthétiser *de novo* certains acides hydroxycinnamiques et de la spermidine ont déjà été élaborées (Gottardi et al., 2017 ; Qin et al., 2021), une telle stratégie serait donc relativement simple à mettre en place. Par contre, il sera nécessaire de contrôler la synthèse des différents acides hydroxycinnamiques afin d'appréhender les décorations de la spermidine. Là encore, un vaste champ d'investigation s'ouvre et pourra s'appuyer sur les outils développés durant cette thèse.

Dans une perspective à très long terme, il serait envisageable de s'affranchir, de manière générale, de l'utilisation de matériel végétal pour la production de MS végétaux d'intérêt pharmaceutique en mettant en place des stratégies d'ingénierie métabolique en levure (ou autres micro-organismes qui viendraient à être caractérisés) pour la bioproduction de ces composés. En effet, les outils permettant l'optimisation des souches de levure « usines » sont de plus en plus développés et permettent désormais le clonage de nombreux gènes en un nombre d'étapes limité. De plus, les connaissances sur le métabolisme endogène de la levure permettent dorénavant de modifier l'orientation des différentes voies, notamment celles permettant la synthèse ou le recyclage des cofacteurs nécessaires aux enzymes hétérologues, sans altérer le fonctionnement et la croissance des cellules. Produire les molécules pharmaceutiques issues du végétal par ingénierie métabolique en levure, ou autre organisme hétérologue d'intérêt, permettrait d'augmenter les rendements, de diminuer les coûts liés à l'extraction et la

purification de MS à partir de matière végétale souvent plus difficile à mettre en place, et facilement maîtrisable en terme d'accès à partir d'un milieu de culture de levures. Outre ces avantages, cette méthode de production permettrait de s'affranchir de la culture de plantes médicinales, permettant ainsi de libérer de l'espace pour les espèces nécessaires à l'industrie agro-alimentaire, et d'exercer moins de pression sur les écosystèmes végétaux. La population mondiale étant en expansion, il devient urgent d'optimiser les grandes cultures afin de produire de la nourriture en quantité suffisante. Toutefois, une bioproduction de MS d'intérêt pharmaceutique généralisée à l'ingénierie métabolique en levure nécessiterait un fort approvisionnement en source de carbone, notamment en glucose indispensable à la croissance des cellules. Ce sucre étant lui-même extrait de matrice végétale comme la betterave à sucre, la bioproduction de MS d'intérêt serait toujours dépendante des végétaux. Par contre, il est tout à fait envisageable d'utiliser les déchets issus de l'industrie agro-alimentaire comme source de carbone pour l'alimentation des levures, cela permettrait ainsi de valoriser les déchets de l'agriculture tout en apportant le sucre nécessaire à la croissance des levures productrices. En particulier, l'expression du transporteur de cellodextrine CDT-1 issu du champignon *Neurospora crassa* chez *S. cerevisiae* a permis aux cellules de croître avec de la cellobiose (sucre provenant de l'hydrolyse de la cellulose, constituant principal de la paroi cellulaire végétale) comme unique source de carbone (Galazka et al., 2010). Il serait donc envisageable d'exprimer ce transporteur dans les cellules « usines » de levure afin d'utiliser les déchets issus de l'agriculture, souvent riches en cellulose, comme source de carbone pour la culture en bioréacteur des cellules. Dans cette optique, il pourrait aussi être intéressant d'optimiser le métabolisme primaire des levures afin de construire des souches moins nécessiteuses en sucre.

BIBLIOGRAPHIE

A :

Abe, K., Matsuura, H., Ukai, M., Shimura, H., Koshino, H., & Suzuki, T. (2017). *N* 1, *N* 14-diferuloylspermine as an antioxidative phytochemical contained in leaves of *Cardamine fauriei*. *Bioscience, Biotechnology, and Biochemistry*, *81*(10), 1855-1860. https://doi.org/10.1080/09168451.2017.1356214

Al-Husini, N., Medler, S., & Ansari, A. (2020). Crosstalk of promoter and terminator during RNA polymerase II transcription cycle. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1863*(12), 194657. https://doi.org/10.1016/j.bbagrm.2020.194657

Albertsen, L., Chen, Y., Bach, L. S., Rattleff, S., Maury, J., Brix, S., Nielsen, J., & Mortensen, U. H. (2011). Diversion of Flux toward Sesquiterpene Production in *Saccharomyces cerevisiae* by Fusion of Host and Heterologous Enzymes. *Applied and Environmental Microbiology*, *77*(3), 1033-1040. https://doi.org/10.1128/AEM.01361-10

Allevi, P., Anastasia, M., Ciuffreda, P., Bigatti, E., & Macdonald, P. (1993). Stereoselective glucosidation of podophyllum lignans. A new simple synthesis of etoposide. *The Journal of Organic Chemistry*, *58*(15), 4175-4178. https://doi.org/10.1021/jo00067a071

Anand, U., Biswas, P., Kumar, V., Ray, D., Ray, P., Loake, V. I. P., Kandimalla, R., Chaudhary, A., Singh, B., Routhu, N. K., Chen, Z.-S., Proćków, J., & Dey, A. (2022). Podophyllum hexandrum and its active constituents : Novel radioprotectants. *Biomedicine & Pharmacotherapy*, *146*, 112555. https://doi.org/10.1016/j.biopha.2021.112555

Angell, S. M., & Baulcombe, D. C. (1999). Potato virus X amplicon-mediated silencing of nuclear genes. *The Plant Journal*, *20*(3), 357-362.

Antošová, Z., & Sychrová, H. (2016). Yeast Hosts for the Production of Recombinant Laccases : A Review. *Molecular Biotechnology*, *58*(2), 93-116. https://doi.org/10.1007/s12033-015-9910-1

Aravind, L., & Koonin, E. V. (2001). *The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases*. 8.

Ardalani, H., Avan, A., & Ghayour-Mobarhan, M. (2017). Podophyllotoxin : A novel potential natural anticancer agent. *Avicenna Journal of Phytomedicine*, 285-294. https://doi.org/10.22038/ajp.2017.8779

Arendt, P., Miettinen, K., Pollier, J., De Rycke, R., Callewaert, N., & Goossens, A. (2017). An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. *Metabolic Engineering*, *40*, 165-175. https://doi.org/10.1016/j.ymben.2017.02.007

Arumugam, N., & Bhojwani, S. S. (1990). Somatic embryogenesis in tissue cultures of *Podophyllum hexandrum*. *Canadian Journal of Botany*, *68*(3), 487-491. https://doi.org/10.1139/b90-065

Athanasopoulou, K., Boti, M. A., Adamopoulos, P. G., Skourou, P. C., & Scorilas, A. (2021). Third-Generation Sequencing : The Spearhead towards the Radical Transformation of Modern Genomics. *Life*, *12*(1), 30. https://doi.org/10.3390/life12010030

Austel, N., Böttcher, C., & Meiners, T. (2021). Chemical defence in Brassicaceae against pollen beetles revealed by metabolomics and flower bud manipulation approaches. *Plant, Cell & Environment*, *44*(2), 519-534. https://doi.org/10.1111/pce.13949

B :

Badhwar, R. L., & Sharma, B. K. (1963). A note on the germination of Podophyllum seeds. *Indian Forester*, *89*, 445-447.

Bak, S., Beisson, F., Bishop, G., Hamberger, B., Höfer, R., Paquette, S., & Werck-Reichhart, D. (2011). Cytochromes P450. *Arabidopsis Book*, *9*. https://doi.org/10.1199/tab.0144

Bandeira-Melo, C., Serra, M. F., Diaz, B. L., Cordeiro, R. S. B., Silva, P. M. R., Lenzi, H. L., Bakhle, Y. S., Serhan, C. N., & Martins, M. A. (2000). Cyclooxygenase-2-Derived Prostaglandin E_2 and Lipoxin A_4 Accelerate Resolution of Allergic Edema in *Angiostrongylus costaricensis*-Infected Rats : Relationship with Concurrent Eosinophilia. *The Journal of Immunology*, *164*(2), 1029-1036. https://doi.org/10.4049/jimmunol.164.2.1029

Barrangou, R. (2015). The roles of CRISPR–Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, *32*, 36-41. https://doi.org/10.1016/j.coi.2014.12.008

Barton, D., & Nakanishi, K. (1999). Comprehensive Natural Products Chemistry. *Elsevier, Amsterdam*. https://doi.org/10.1002/1521-3773(20010316)40:63.3.CO;2-3

Bassard, J.-E., Ullmann, P., Bernier, F., & Werck-Reichhart, D. (2010). Phenolamides : Bridging polyamines to the phenolic metabolism. *Phytochemistry*, *71*(16), 1808-1824. https://doi.org/10.1016/j.phytochem.2010.08.003

Bedir, E., Tellez, M., Lata, H., Khan, I., Cushman, K. E., & Moraes, R. M. (2006). Post-harvest and scale-up extraction of American mayapple leaves for podophyllotoxin production. *Industrial Crops and Products*, *24*(1), 3-7. https://doi.org/10.1016/j.indcrop.2005.10.001

Bedows, E., & Hatfield, G. M. (1982). An Investigation of the Antiviral Activity of Podophyllum peltatum. *Journal of Natural Products*, *45*(6), 725-729. https://doi.org/10.1021/np50024a015

Berkowitz, David. B., Choi, S., & Maeng, J.-H. (2000). Enzyme-Assisted Asymmetric Total Synthesis of (−)-Podophyllotoxin and (−)-Picropodophyllin. *The journal of Organic Chemistry*, *65*(3), 847-860.

Bernard, G., Buges, J., Delporte, M., Molinié, R., Besseau, S., Bouchereau, A., Watrin, A., Fontaine, J.-X., Mathiron, D., Berardocco, S., Bassard, S., Quéro, A., Hilbert, J.-L., Rambaud, C., & Gagneul, D. (2022). Consecutive action of two BAHD acyltransferases promotes tetracoumaroyl spermine accumulation in chicory. *Plant Physiology*, *189*(4), 2029-2043. https://doi.org/10.1093/plphys/kiac234

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., … Courdavault, V. (2013). A Pair of Tabersonine 16-Hydroxylases Initiates the Synthesis of Vindoline in an Organ-Dependent Manner in *Catharanthus roseus*. *Plant Physiology*, *163*(4), 1792-1803. https://doi.org/10.1104/pp.113.222828

Bianchetti, G. (2019). *Ingénierie métabolique de la production d'un précurseur d'étoposide en système hétérologue* [Mémoire de stage de Master 2].

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology*, *54*(1), 519-546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Bokern, M., Witte, L., Wray, V., Nimtz, M., & Meurer-Grimes, B. (1995). Trisubstituted hydroxycinnamic acid spermidines from *Quercus dentata* pollen. *Phytochemistry*, *39*, 1371-1375.

Bourgaud, F., Hehn, A., Larbat, R., Doerper, S., Gontier, E., Kellner, S., & Matern, U. (2006). Biosynthesis of coumarins in plants : A major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochemistry Reviews*, *5*(2-3), 293-308. https://doi.org/10.1007/s11101-006-9040-2

Bourgaud, F., Gravot, A., Milesi, S., & Gontier, E. (2001). Production of plant secondary metabolites : A historical perspective. *Plant Science*, *161*(5), 839-851. https://doi.org/10.1016/S0168-9452(01)00490-3

Brockington, S. F., Yang, Y., Gandia-Herrero, F., Covshoff, S., Hibberd, J. M., Sage, R. F., Wong, G. K. S., Moore, M. J., & Smith, S. A. (2015). Lineage-specific gene radiations underlie the evolution of novel betalain pigmentation in Caryophyllales. *New Phytologist*, *207*(4), 1170-1180. https://doi.org/10.1111/nph.13441

Bromberg, L. (1998). Polyether-Modified Poly(acrylic acid) : Synthesis and Applications. *Industrial & Engineering Chemistry Research*, *37*(11), 4267-4274. https://doi.org/10.1021/ie980358s

Brown, S., Clastre, M., Courdavault, V., & O'Connor, S. E. (2015). *De novo* production of the plant-derived alkaloid strictosidine in yeast. *Proceedings of the National Academy of Sciences*, *112*(11), 3205-3210. https://doi.org/10.1073/pnas.1423555112

Bucher, M., Brander, K. A., Sbicego, S., Mandel, T., & Kuhlemeier, C. (1995). Aerobic fermentation in tobacco pollen. *Plant Molecular Biology*, *28*(4), 739-750. https://doi.org/10.1007/BF00021197

Bugg, T. D. H. (2003). Dioxygenase enzymes : Catalytic mecanisms and chemical models. *Tetrahedron*, *59*(36), 7075-7101. https://doi.org/10.1016/S0040-4020(03)00944-X

Burhenne, K., Kristensen, B. K., & Rasmussen, S. K. (2003). A New Class of N-Hydroxycinnamoyltransferases : Purification, cloning, and expression of a barley agmatine coumaroyltransferase (EC 2.3.1.64). *Journal of Biological Chemistry*, *278*, 13919-13927.

Bush, E. J., & Jones, D. W. (1998). Asymmetric Total Synthesis of (-)-Podophyllotoxin+. *Journal of the Chemical Society, Perkin Transactions*, *1*, 1489-1492.

Buss, A. D., Cox, B., & Waigh, R. D. (1995). Natural Products as Leads for New Pharmaceuticals. In *Burger's Medicinal Chemistry, Drug Discovery and Development*.

Buta, J. G., & Izac, R. R. (1972). Caffeoylputrescine in *Nicotiana tabacum*. *Phytochemistry*, *11*(3), 1188-1189.

C :

Cabanne, F., Dalebroux, M. A., Martin-Tanguy, J., & Martin, C. (1981). Hydroxycinnamic acid amides and ripening to flower of *Nicotiana tabacum* var. *Xanthi n.c*. *Physiologia Plantarum*, *53*(4), 399-404. https://doi.org/10.1111/j.1399-3054.1981.tb02721.x

Cabello-Hurtado, F., Taton, M., Forthoffer, N., Kahn, R., Bak, S., Rahier, A., & Werck-Reichhart, D. (1999). Optimized expression and catalytic properties of a wheat obtusifoliol 14alpha-demethylase (CYP51) expressed in yeast. Complementation of erg11Delta yeast mutants by

plant CYP51. *European Journal of Biochemistry*, *262*(2), 435-446. https://doi.org/10.1046/j.1432-1327.1999.00376.x

Caldecott, K. W. (2008). Single-strand break repair and genetic disease. *Nature Reviews Genetics*, *9*(8), 619-631. https://doi.org/10.1038/nrg2380

Canel, C., Moraes, R. M., Dayan, F. E., & Ferreira, D. (2000). Podophyllotoxin. *Phytochemistry*, *54*(2), 115-120. https://doi.org/10.1016/S0031-9422(00)00094-7

Cao, L., Huang, J., & Li, J. (2007). Fermentation conditions of *Sinopodophyllum hexandrum* endophytic fungus on production of podophyllotoxin. *Food and Fermentation Industries*, 28-32.

Caputi, L., Franke, J., Farrow, S. C., Chung, K., Payne, R. M. E., Nguyen, T.-D., Dang, T.-T. T., Soares Teto Carqueijeiro, I., Koudounas, K., Dugé de Bernonville, T., Ameyaw, B., Jones, D. M., Vieira, I. J. C., Courdavault, V., & O'Connor, S. E. (2018). Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Science*, *360*(6394), 1235-1239. https://doi.org/10.1126/science.aat4100

Carqueijeiro, I., Dugé de Bernonville, T., Lanoue, A., Dang, T., Teijaro, C. N., Paetz, C., Billet, K., Mosquera, A., Oudin, A., Besseau, S., Papon, N., Glévarec, G., Atehortùa, L., Clastre, M., Giglioli-Guivarc'h, N., Schneider, B., St-Pierre, B., Andrade, R. B., O'Connor, S. E., & Courdavault, V. (2018). A BAHD acyltransferase catalyzing 19- *O* -acetylation of tabersonine derivatives in roots of *Catharanthus roseus* enables combinatorial synthesis of monoterpene indole alkaloids. *The Plant Journal*, *94*(3), 469-484. https://doi.org/10.1111/tpj.13868

Carqueijeiro, I., Masini, E., Foureau, E., Sepúlveda, L. J., Marais, E., Lanoue, A., Besseau, S., Papon, N., Clastre, M., Dugé de Bernonville, T., Glévarec, G., Atehortùa, L., Oudin, A., & Courdavault, V. (2015). Virus-induced gene silencing in *Catharanthus roseus* by biolistic inoculation of tobacco rattle virus vectors. *Plant Biology*, *17*(6), 1242-1246. https://doi.org/10.1111/plb.12380

Chang, Z., Li, G., Liu, J., Zhang, Y., Ashby, C., Liu, D., Cramer, C. L., & Huang, X. (2015). Bridger : A new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biology*, *16*(1), 30. https://doi.org/10.1186/s13059-015-0596-2

Chatterjee, R. (1952). Indian Podophyllum. *Economic Botany*, *6*(4), 342-354. https://doi.org/10.1007/BF02984882

Chattopadhyay, S., Mehra, R., Srivastava, A., Bhojwani, S., & Bisaria, V. (2003). Effect of major nutrients on podophyllotoxin production in *Podophyllum hexandrum* suspension cultures. *Applied Microbiology and Biotechnology*, *60*(5), 541-546. https://doi.org/10.1007/s00253-002-1149-7

Chattopadhyay, S., Srivastava, K., Bhojwani, S. S., & Bisaria', S. (2002a). Production of Podophyllotoxin by Plant Cell Cultures of *Podophyllum hexandrum* in Bioreactor. *Journal of Bioscience and Bioengineering*, *93*(2), 215-220.

Chattopadyay, S., Srivastav, A. K., & Bisaria, V. S. (2002b). Optimization of Culture Parameters for Production of Podophyllotoxin in Suspension Culture of *Podophyllum hexandrum*. *Applied Biochemistry and Biotechnology*, *102-103*(1-6), 381-394. https://doi.org/10.1385/ABAB:102-103:1-6:381

Chattopadhyay, S., Srivastava, K., Bhojwani, S. S., & Bisaria', S. (2001). Development of suspension culture of *Podophyllum hexandrum* for production of podophyllotoxin. *Biotechnology Letters*, *23*(24), 2063-2066.

Chen, A. Y. (1994). DNA Topoisomerases : Essential Enzymes and Lethal Targets. *Annual Review of Pharmacology and Toxicology*, *34*, 28.

Chen, H. C., Wang, H. J., Jin, Y., Tian, X., Zheng, Y.-T., Luo, D. Q., & Tu, Y.-Q. (2007). Synthesis and anti-HIV-1 activities of novel podophyllotoxin derivates. *Bioorganic & Medicinal Chemistry Letters*, *17*(7), 2091-2095. https://doi.org/10.1016/j.bmcl.2006.11.070

Chen, H., Yang, Y., Wang, Z., Dou, J., Wang, H., & Zhou, C. (2016). Elevated intracellular acetyl-CoA availability by acs2 overexpression and mls1 deletion combined with metK1 introduction enhanced SAM accumulation in Saccharomyces cerevisiae. *Biochemical Engineering Journal*, *107*, 26-34. https://doi.org/10.1016/j.bej.2015.11.016

Chen, R., Gao, J., Yu, W., Chen, X., Zhai, X., Chen, Y., Zhang, L., & Zhou, Y. J. (2022). Engineering cofactor supply and recycling to drive phenolic acid biosynthesis in yeast. *Nature Chemical Biology*, *18*(5), 520-529. https://doi.org/10.1038/s41589-022-01014-6

Chen, R., Yang, S., Zhang, L., & Zhou, Y. J. (2020). Advanced Strategies for Production of Natural Products in Yeast. *IScience*, *23*(3), 100879. https://doi.org/10.1016/j.isci.2020.100879

Chen, S., Wang, X., Zhao, B., Yuan, X., & Wang, Y. (2003). *Production of crocin using Crocus sativus callus by two-stage culture system*. *Biotechnology Letters.* 25, 1235-1238.

Chi, P.-T., & Maimone, T. J. (2014). C-H Bond Arylation in the Synthesis of Aryltetralin Lignans : A Short Total Synthesis of Podophyllotoxin. *Angewandte Chemie International Edition*, *53*(12). https://doi.org/10.1002/anie.201311112

Chin, J. X., Chung, B. K.-S., & Lee, D.-Y. (2014). Codon Optimization OnLine (COOL) : A web-based multi-objective optimization platform for synthetic gene design. *Bioinformatics*, *30*(15), 2210-2212. https://doi.org/10.1093/bioinformatics/btu192

Choi, S. W., Lee, S. K., Kim, E. O., Oh, J. H., Yoon, K. S., Parris, N., Hicks, K. B., & Moreau, R. A. (2007). Antioxidant and Antimelanogenic Activities of Polyamine Conjugates from Corn Bran and Related Hydroxycinnamic Acids. *Journal of Agricultural and Food Chemistry*, *55*(10), 3920-3925. https://doi.org/10.1021/jf0635154

Cohen, S. N., Chang, A. C. Y., Boyer, H. W., & Helling, R. B. (1973). Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *Proceedings of the National Academy of Sciences*, *70*(11), 3240-3244. https://doi.org/10.1073/pnas.70.11.3240

Corbin, C., Lafontaine, F., Sepúlveda, L. J., Carqueijeiro, I., Courtois, M., Lanoue, A., Dugé de Bernonville, T., Besseau, S., Glévarec, G., Papon, N., Atehortúa, L., Giglioli-Guivarc'h, N., Clastre, M., St-Pierre, B., Oudin, A., & Courdavault, V. (2017). Virus-induced gene silencing in *Rauwolfia* species. *Protoplasma*, *254*(4), 1813-1818. https://doi.org/10.1007/s00709-017-1079-y

Courdavault, V., O'Connor, S. E., Jensen, M. K., & Papon, N. (2021). Metabolic engineering for plant natural products biosynthesis : New procedures, concrete achievements and remaining limits. *Natural Product Reports*, *38*(12), 2145-2153. https://doi.org/10.1039/D0NP00092B

Courdavault, V., Besseau, S., Oudin, A., Papon, N., & O'Connor, S. E. (2020). Virus-Induced Gene Silencing : Hush Genes to Make Them Talk. *Trends in Plant Science*, *25*(7), 714-715. https://doi.org/10.1016/j.tplants.2020.02.013

Courdavault, V., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., & Burlat, V. (2014). A look inside an alkaloid multisite plant : The Catharanthus logistics. *Current Opinion in Plant*

Biology, *19*, 43-50. https://doi.org/10.1016/j.pbi.2014.03.010

Cravens, A., Payne, J., & Smolke, C. D. (2019). Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nature Communications*, *10*(1), 2142. https://doi.org/10.1038/s41467-019-09848-w

Curran, K. A., Karim, A. S., Gupta, A., & Alper, H. S. (2013). Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic Engineering*, *19*, 88-97. https://doi.org/10.1016/j.ymben.2013.07.001

D :

Dalisay, D. S., Kim, K. W., Lee, C., Yang, H., Rübel, O., Bowen, B. P., Davin, L. B., & Lewis, N. G. (2015). Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed : Integrated Omics and MALDI Mass Spectrometry Imaging. *Journal of Natural Products*, *78*(6), 1231-1242. https://doi.org/10.1021/acs.jnatprod.5b00023

D'Ambrosio, V., Hansen, L. G., Zhang, J., Jensen, E. D., Arsovska, D., Laloux, M., Jakočiūnas, T., Hjort, P., De Lucrezia, D., Marletta, S., Keasling, J. D., & Jensen, M. K. (2022). A FAIR-compliant parts catalogue for genome engineering and expression control in Saccharomyces cerevisiae. *Synthetic and Systems Biotechnology*, *7*(2), 657-663. https://doi.org/10.1016/j.synbio.2022.02.001

Dang, T.-T. T., Franke, J., Tatsis, E., & O'Connor, S. E. (2017). Dual Catalytic Activity of a Cytochrome P450 Controls Bifurcation at a Metabolic Branch Point of Alkaloid Biosynthesis in *Rauwolfia serpentina*. *Angewandte Chemie*, *129*(32), 9568-9572. https://doi.org/10.1002/ange.201705010

Danielsson, O., & Jörnvall, H. (1992). « Enzymogenesis » : Classical liver alcohol dehydrogenase origin from the glutathione-dependent formaldehyde dehydrogenase line. *Proceedings of the National Academy of Sciences*, *89*(19), 9247-9251. https://doi.org/10.1073/pnas.89.19.9247

D'Auria, J. C. (2006). Acyltransferases in plants : A good time to be BAHD. *Current Opinion in Plant Biology*, *9*(3), 331-340. https://doi.org/10.1016/j.pbi.2006.03.016

Davies, M. E., Tsyplenkov, D., & Martin, V. J. J. (2021). Engineering yeast for *de-novo* synthesis of the insect repellent—Nepetalactone. *ACS Synthetic biology 10*(11), 2896-2900.

Davin, L. B., Wang, H.-B., Crowell, A. L., Bedgar, D. L., Martin, D. M., Sarkanen, S., & Lewis, N. G. (1997). Stereoselective Bimolecular Phenoxy Radical Coupling by an Auxiliary (Dirigent) Protein Without an Active Center. *Science, New Series*, *275*(5298), 362-366.

De Luca, V., Salim, V., Levac, D., Atsumi, S. M., & Yu, F. (2012). Discovery and Functional Analysis of Monoterpenoid Indole Alkaloid Pathways in Plants. In *Methods in Enzymology* (Vol. 515, p. 207-229). Elsevier. https://doi.org/10.1016/B978-0-12-394290-6.00010-0

Decoene, T., De Maeseneire, S. L., & De Mey, M. (2019). Modulating transcription through development of semi-synthetic yeast core promoters. *PLOS ONE*, *14*(11), e0224476. https://doi.org/10.1371/journal.pone.0224476

Degreif, D., Kremenovic, M., Geiger, T., & Bertl, A. (2018). Preloading budding yeast with all-in-one CRISPR/Cas9 vectors for easy and high-efficient genome editing. *Journal of Biological Methods*, *5*(3), e98. https://doi.org/10.14440/jbm.2018.254

Delporte, M., Bernard, G., Legrand, G., Hielscher, B., Lanoue, A., Molinié, R., Rambaud, C., Mathiron, D., Besseau, S., Linka, N., Hilbert, J.-L., & Gagneul, D. (2018). A BAHD neofunctionalization promotes tetrahydroxycinnamoyl spermine accumulation in the pollen coats of the Asteraceae family. *Journal of Experimental Botany*. https://doi.org/10.1093/jxb/ery320

Demkura, P. V., Abdala, G., Baldwin, I. T., & Ballaré, C. L. (2010). Jasmonate-Dependent and -Independent Pathways Mediate Specific Effects of Solar Ultraviolet B Radiation on Leaf Phenolics and Antiherbivore Defense. *Plant Physiology*, *152*(2), 1084-1095. https://doi.org/10.1104/pp.109.148999

Djurdjevic, L., Gajic, G., Jaric, S., Kostic, O., Mitrovic, M., & Pavlovic, P. (2013). Analysis of benzoic and cinnamic acid derivatives of some medicinal plants in Serbia. *Archives of Biological Sciences*, *65*(2), 603-609. https://doi.org/10.2298/ABS1302603D

Dudareva, N., D'Auria, J. C., Nam, K. H., Raguso, R. A., & Pichersky, E. (1998). Acetyl-CoA:benzylalcohol acetyltransferase – an enzyme involved in floral scent production in *Clarkia breweri*. *The Plant Journal*, *14*(3), 297-304. https://doi.org/10.1046/j.1365-313X.1998.00121.x

Dugé de Bernonville, T., Carqueijeiro, I., Lanoue, A., Lafontaine, F., Sánchez Bel, P., Liesecke, F., Musset, K., Oudin, A., Glévarec, G., Pichon, O., Besseau, S., Clastre, M., St-Pierre, B., Flors, V., Maury, S., Huguet, E., O'Connor, S. E., & Courdavault, V. (2017). Folivory elicits a strong defense reaction in Catharanthus roseus : Metabolomic and transcriptomic analyses reveal distinct local and systemic responses. *Scientific Reports*, *7*(1), 40453. https://doi.org/10.1038/srep40453

Dugé de Bernonville, T., Clastre, M., Besseau, S., Oudin, A., Burlat, V., Glévarec, G., Lanoue, A., Papon, N., Giglioli-Guivarc'h, N., St-Pierre, B., & Courdavault, V. (2015). Phytochemical genomics of the Madagascar periwinkle : Unravelling the last twists of the alkaloid engine. *Phytochemistry*, *113*, 9-23. https://doi.org/10.1016/j.phytochem.2014.07.023

Durst, F., & Nelson, D. R. (1995). Diversity and Evolution of Plant P450 and P450-Reductases. *Drug Metabolism and Drug Interactions*, *12*(3-4). https://doi.org/10.1515/DMDI.1995.12.3-4.189

E :

Edreva, A. M., Velikova, V. B., & Tsonev, T. D. (2007). Phenylamides in plants. *Russian Journal of Plant Physiology*, *54*(3), 287-301. https://doi.org/10.1134/S1021443707030016

Edreva, A., Yordanov, I., Kardjieva, R., & Gesheva, E. (1998). Heat shock responses of bean plants : Involvement of free radicals, antioxidants and free radical/active oxygen scavenging systems. *Biologia plantarum*, *41*, 185-191.

Edreva, A., Yordanov, I., Kardjieva, R., Hadjiiska, E., & Gesheva, E. (1995). Expression of phenylamides in abiotic stress conditions. *Bulgarian Journal of Plant Physiology*, *21*, 15-23.

Elejalde-Palmett, C., de Bernonville, T. D., Glevarec, G., Pichon, O., Papon, N., Courdavault, V., St-Pierre, B., Giglioli-Guivarc'h, N., Lanoue, A., & Besseau, S. (2015). Characterization of a spermidine hydroxycinnamoyltransferase in *Malus domestica* highlights the evolutionary conservation of trihydroxycinnamoyl spermidines in pollen coat of core Eudicotyledons. *Journal of Experimental Botany*, *66*(22), 7271-7285. https://doi.org/10.1093/jxb/erv423

Empt, U., Alfermann, A., Pras, N., & Petersen, M. (2000). The use of plant cell cultures for the

production of podophyllotoxin and related lignans. *Journal of Applied Botany*, *74*, 145-150.

Erhirhie, E. O., Ikegbune, C., Okeke, A. I., Onwuzuligbo, C. C., Madubuogwu, N. U., Chukwudulue, U. M., & Okonkwo, O. B. (2021). Antimalarial herbal drugs : A review of their interactions with conventional antimalarial drugs. *Clinical Phytoscience*, *7*(1), 4. https://doi.org/10.1186/s40816-020-00242-4

Eyberger, A. L., Dondapati, R., & Porter, J. R. (2006). Endophyte Fungal Isolates from *Podophyllum peltatum* Produce Podophyllotoxin. *Journal of Natural Products*, *69*(8), 1121-1124. https://doi.org/10.1021/np060174f

F :

Favarin, C. D., de Oliveira, R. J., de Oliveira, J. F. C., & de Paula Rogerio, A. (2013). Potential Effects of Medicinal Plants and Secondary Metabolites on Acute Lung Injury. *BioMed Research International*, *2013*, 1-12. https://doi.org/10.1155/2013/576479

Fellenberg, C., & Vogt, T. (2015). Evolutionarily conserved phenylpropanoid pattern on angiosperm pollen. *Trends in Plant Science*, *20*(4), 212-218. https://doi.org/10.1016/j.tplants.2015.01.011

Foureau, E., Carqueijeiro, I., Dugé de Bernonville, T., Melin, C., Lafontaine, F., Besseau, S., Lanoue, A., Papon, N., Oudin, A., Glévarec, G., Clastre, M., St-Pierre, B., Giglioli-Guivarc'h, N., & Courdavault, V. (2016). Prequels to synthetic biology : From candidate gene identification and validation to enzyme subcellular localization in plant and yeast cells. *Methods in Enzymology*, *576*, 167-206. https://doi.org/10.1016/bs.mie.2016.02.013

Fournand, D. (2003). Initial steps of the peroxidase-catalyzed polymerization of coniferyl alcohol and/or sinapyl aldehyde capillary zone electrophoresis study of pH effect. *Phytochemistry*, *62*(2), 139-146. https://doi.org/10.1016/S0031-9422(02)00573-3

Fox, J. M., & Erill, I. (2010). Relative Codon Adaptation : A Generic Codon Bias Index for Prediction of Gene Expression. *DNA Research*, *17*(3), 185-196. https://doi.org/10.1093/dnares/dsq012

Froelich-Ammon, S. J., & Osheroff, N. (1995). Topoisomerase Poisons : Harnessing the Dark Side of Enzyme Mechanism. *Journal of Biological Chemistry*, *270*(37), 21429-21432. https://doi.org/10.1074/jbc.270.37.21429

Fu, X. P., Wu, T., Abdurahim, M., Su, Z., Hou, X. L., Aisa, H. A., & Wu, H. (2008). New spermidine alkaloids from *Capparis spinosa* roots. *Phytochemistry Letters*, *1*(1), 59-62. https://doi.org/10.1016/j.phytol.2008.01.001

Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K., Fukuchi-Mizutani, M., Nakao, M., Fukui, Y., Yamaguchi, M., Ashikari, T., & Kusumi, T. (1998). CDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. *The Plant Journal*, *16*(4), 421-431. https://doi.org/10.1046/j.1365-313x.1998.00312.x

Fuss, E. (2007). *Biosynthesis of Lignans*. Université d'Heintich-Heine.

G :

Galazka, J. M., Tian, C., Beeson, W. T., Martinez, B., Glass, N. L., & Cate, J. H. D. (2010). Cellodextrin Transport in Yeast for Improved Biofuel Production. *Science*, *330*(6000), 84-86. https://doi.org/10.1126/science.1192838

Garabagi, F., Duns, G., & Strommer, J. (2005). Selective recruitment of Adh genes for distinct enzymatic functions in Petunia hybrida. *Plant Molecular Biology*, *58*(2), 283-294. https://doi.org/10.1007/s11103-005-3545-8

Garabagi, F., & Strommer, J. (2004). Distinct Genes Produce the Alcohol Dehydrogenases of Pollen and Maternal Tissues in Petunia hybrida. *Biochemical Genetics volume*, *42*, 199-208. https://doi.org/10.1023/b:bigi.0000026634.69911.2e

Garden, H. J. (2003). *Biotechnological production of podophyllotoxin by Linum album suspension cultures*. Université d'Heinrich-Heine.

Garros, L., Drouet, S., Corbin, C., Decourtil, C., Fidel, T., Lebas de Lacour, J., Leclerc, E., Renouard, S., Tungmunnithum, D., Doussot, J., Abassi, B., Maunit, B., Lainé, É., Fliniaux, O., Mesnard, F., & Hano, C. (2018). Insight into the Influence of Cultivar Type, Cultivation Year, and Site on the Lignans and Related Phenolic Profiles, and the Health-Promoting Antioxidant Potential of Flax (*Linum usitatissimum* L.) Seeds. *Molecules*, *23*(10), 2636. https://doi.org/10.3390/molecules23102636

Gauthier, A., Gulick, P. J., & Ibrahim, R. K. (1998). Characterization of Two cDNA Clones Which EncodeO-Methyltransferases for the Methylation of both Flavonoid and Phenylpropanoid Compounds. *Archives of Biochemistry and Biophysics*, *351*(2), 243-249. https://doi.org/10.1006/abbi.1997.0554

Gellert, M., Mizuuchi, K., O'Dea, M. H., & Nash, H. A. (1976). DNA gyrase : An enzyme that introduces superhelical turns into DNA. *Proceedings of the National Academy of Sciences*, *73*(11), 3872-3876. https://doi.org/10.1073/pnas.73.11.3872

Georgiev, L., Chochkova, M., Ivanova, G., Najdenski, H., Ninova, M., & Milkova, T. (2012). *Radical scavenging and antimicrobial activities of cinnamoyl amides of biogenic monoamines*. 89-91.

Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T., & Halkier, B. A. (2007). USER fusion : A rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Research*, *35*(7), e55-e55. https://doi.org/10.1093/nar/gkm106

Ghosh, S. (2017). Triterpene Structural Diversification by Plant Cytochrome P450 Enzymes. *Frontiers in Plant Science*, *8*, 1886. https://doi.org/10.3389/fpls.2017.01886

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343-345. https://doi.org/10.1038/nmeth.1318

Gietz, R. D., & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, *2*(1), 31-34. https://doi.org/10.1038/nprot.2007.13

Giordano, D., Provenzano, S., Ferrandino, A., Vitali, M., Pagliarani, C., Roman, F., Cardinale, F., Castellarin, S. D., & Schubert, A. (2016). Characterization of a multifunctional caffeoyl-CoA O -methyltransferase activated in grape berries upon drought stress. *Plant Physiology and Biochemistry*, *101*, 23-32. https://doi.org/10.1016/j.plaphy.2016.01.015

Giri, A., & Narasu, M. L. (2000). Production of podophyllotoxin from *Podophyllum hexandrum* : A potential natural product for clinically useful anticancer drugs. *School of Biotechnology, India*.

Glenn, W. S., Nims, E., & O'Connor, S. E. (2011). Reengineering a Tryptophan Halogenase To Preferentially Chlorinate a Direct Alkaloid Precursor. *Journal of the American Chemical*

Society, *133*(48), 19346-19349. https://doi.org/10.1021/ja2089348

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., & Riggs, A. D. (1979). Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences*, *76*(1), 106-110. https://doi.org/10.1073/pnas.76.1.106

Goel, H. C., Prasad, J., Sharma, A., & Singh, B. (1998). Antitumor and radioprotective action of Podophyllum hexandrum. *Indian Journal of Experimental Biology*, *36*(6), 583-587.

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., & Oliver, S. G. (1996). Life with 6000 Genes. *Science*, *274*(5287), 546-567. https://doi.org/10.1126/science.274.5287.546

Gomez-Roldan, V., Fermas, S., Brewer, P. B., Puech-Pagès, V., Dun, E. A., Pillot, J.-P., Letisse, F., Matusova, R., Danoun, S., Portais, J.-C., Bouwmeester, H., Bécard, G., Beveridge, C. A., Rameau, C., & Rochange, S. F. (2008). Strigolactone inhibition of shoot branching. *Nature*, *455*(7210), 189-194. https://doi.org/10.1038/nature07271

Gordaliza, M., García, P. A., Miguel del Corral, J. M., Castro, M. A., & Gómez-Zurita, M. A. (2004). Podophyllotoxin : Distribution, sources, applications and new cytotoxic derivatives. *Toxicon*, *44*(4), 441-459. https://doi.org/10.1016/j.toxicon.2004.05.008

Gorres, K. L., & Raines, R. (2010). Prolyl 4-hydroxylase. *Critical reviews in biochemistry and molecular biology*, *45*(2), 106-124.

Goto, D. B., & Nakayama, J. (2012). RNA and epigenetic silencing : Insight from fission yeast. *Development, Growth & Differentiation*, *54*(1), 129-141. https://doi.org/10.1111/j.1440-169X.2011.01310.x

Gottardi, M., Knudsen, J. D., Prado, L., Oreb, M., Branduardi, P., & Boles, E. (2017). De novo biosynthesis of trans-cinnamic acid derivatives in Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology*, *101*(12), 4883-4893. https://doi.org/10.1007/s00253-017-8220-x

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644-652. https://doi.org/10.1038/nbt.1883

Greenspan, E. M., Lester, J., & Shear, M. J. (1950). Effect of Alpha-Peltatin, Beta-Peltatin, and Podophyllotoxin on Lymphomas and Other Transplanted Tumors2. *Journal of the National Cancer Institute*, *10*(6), 1273-1293.

Grienenberger, E., Pollet, B., Heitz, T., & Legrand, M. (2009). A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *The Plant Journal*, *58*, 14.

Guerriero, G., Berni, R., Muñoz-Sanchez, J., Apone, F., Abdel-Salam, E., Qahtan, A., Alatar, A., Cantini, C., Cai, G., Hausman, J.-F., Siddiqui, K., Hernández-Sotomayor, S., & Faisal, M. (2018). Production of Plant Secondary Metabolites : Examples, Tips and Suggestions for Biotechnologists. *Genes*, *9*(6), 309. https://doi.org/10.3390/genes9060309

Guirimand, G., Kulagina, N., Papon, N., Hasunuma, T., & Courdavault, V. (2021). Innovative Tools and Strategies for Optimizing Yeast Cell Factories. *Trends in Biotechnology*, *39*(5), 488-504. https://doi.org/10.1016/j.tibtech.2020.08.010

Guirimand, G., Guihur, A., Perello, C., Phillips, M., Mahroug, S., Oudin, A., Dugé de Bernonville, T., Besseau, S., Lanoue, A., Giglioli-Guivarc'h, N., Papon, N., St-Pierre, B., Rodríguez-Concepcíon, M., Burlat, V., & Courdavault, V. (2020). Cellular and Subcellular Compartmentation of the 2C-Methyl-D-Erythritol 4-Phosphate Pathway in the Madagascar Periwinkle. *Plants*, *9*(4), 462. https://doi.org/10.3390/plants9040462

Guirimand, G., Courdavault, V., Lanoue, A., Mahroug, S., Guihur, A., Nlanc, N., Giglioli-Guivarc'h, N, St-Pierre, B., & Burlat, V. (2010). Strictosidine activation in Apocynaceae : Towards a "nuclear time bomb"?. *Plant Biology 10*, 182.

Gustafsson, C., Govindarajan, S., & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology*, *22*(7), 346-353. https://doi.org/10.1016/j.tibtech.2004.04.006

H :

Hadadi, N., Hafner, J., Shajkofci, A., Zisaki, A., & Hatzimanikatis, V. (2016). ATLAS of Biochemistry : A Repository of All Possible Biochemical Reactions for Synthetic Biology and Metabolic Engineering Studies. *ACS Synthetic Biology*, *5*(10), 1155-1166. https://doi.org/10.1021/acssynbio.6b00054

Hagel, J. M., & Facchini, P. J. (2010). Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nature Chemical Biology*, *6*(4), 273-275. https://doi.org/10.1038/nchembio.317

Hahlbrock, K., & Scheel, D. (1989). Physiology and Molecular Biology of Phenylpropanoid Metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology 40*, 347-369.

Haigler, B. E., & Gibson, D. T. (1990). Purification and properties of NADH-ferredoxinNAP reductase, a component of naphthalene dioxygenase from Pseudomonas sp. Strain NCIB 9816. *Journal of Bacteriology*, *172*(1), 457-464. https://doi.org/10.1128/jb.172.1.457-464.1990

Hale, G., Clark, M. R., Marcus, R., Winter, G., Dyer, M. J. S., Phillips, J. M., Riechmann, L., & Waldmann, H. (1988). Remission induction in non-Hodgkin lymphoma with reshaped human monoclonal antibody CAMPATH-1H. *The Lancet*, *332*(8625), 1394-1399. https://doi.org/10.1016/S0140-6736(88)90588-0

Hall, B. (1989). Therapy with monoclonal antibodies to CD4 : Potential not appreciated? *American Journal of Kidney Diseases*, *14*, 71-77.

Halls, S. C., Davin, L. B., Kramer, D. M., & Lewis, N. G. (2004). Kinetic Study of Coniferyl Alcohol Radical Binding to the (+)-Pinoresinol Forming Dirigent Protein. *Biochemistry*, *43*(9), 2587-2595. https://doi.org/10.1021/bi035959o

Hande, K. R. (1998). Etoposide : Four decades of development of a topoisomerase II inhibitor. *European Journal of Cancer*, *34*(10), 1514-1521. https://doi.org/10.1016/S0959-8049(98)00228-7

Handrick, V., Vogt, T., & Frolov, A. (2010). Profiling of hydroxycinnamic acid amides in *Arabidopsis thaliana* pollen by tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*,

398(7-8), 2789-2801. https://doi.org/10.1007/s00216-010-4129-2

Hano, C. F., Dinkova-Kostova, A. T., Davin, L. B., Cort, J. R., & Lewis, N. G. (2021). Editorial : Lignans: Insights Into Their Biosynthesis, Metabolic Engineering, Analytical Methods and Health Benefits. *Frontiers in Plant Science*, *11*, 630327. https://doi.org/10.3389/fpls.2020.630327

Hartwell, J. L., & Schrecker, A. W. (1985). The chemistry of Podophyllum. *Progress in the chemistry of organic natural products*, *15*, 83-166.

Hawkins, K. M., & Smolke, C. D. (2008). Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Nature Chemical Biology*, *4*(9), 564-573. https://doi.org/10.1038/nchembio.105

Hayaishi, O. (1966). Crystalline oxygenases of pseudomonads. *Bacteriological Reviews 30*, 4, 720-731.

Hayashi, K., Sakaki, T., Kominami, S., Inouye, K., & Yabusaki, Y. (2000). Coexpression of Genetically Engineered Fused Enzyme between Yeast NADPH P450 Reductase and Human Cytochrome P450 3A4 and Human Cytochrome b5 in Yeast. *Archives of Biochemistry and Biophysics*, *381*(1), 164-170.

Hedberg, C., Hesse, M., & Werner, C. (1996). Spermine and spermidine hydroxycinnamoyl transferases in *Aphelandra tetragona*. *Plant Science*, *113*(2), 149-156. https://doi.org/10.1016/0168-9452(95)04298-9

Hoballah, M. E., Stuurman, J., Turlings, T. C. J., Guerin, P. M., Connétable, S., & Kuhlemeier, C. (2005). The composition and timing of flower odour emission by wild Petunia axillaris coincide with the antennal perception and nocturnal activity of the pollinator Manduca sexta. *Planta*, *222*(1), 141-150. https://doi.org/10.1007/s00425-005-1506-8

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J. C., Zeitler, B., Cherone, J. M., Meng, X., Hinkley, S. J., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Jaenisch, R. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology*, *29*(8), 731-734. https://doi.org/10.1038/nbt.1927

Holm, T. (1899). *Podophyllum peltatum*, a morphological study. *Botanical Gazette 27*(6), 419-433.

Hu, W., & Hesse, M. (1996). Synthese der *p*-Cumaroylspermidine. *Helvetica Chimica Acta*, *79*(2), 548-559. https://doi.org/10.1002/hlca.19960790222

Huang, J.-X., Zhang, J., Zhang, X.-R., Zhang, K., Zhang, X., & He, X.-R. (2014). *Mucor fragilis* as a novel source of the key pharmaceutical agents podophyllotoxin and kaempferol. *Pharmaceutical Biology*, *52*(10), 1237-1243. https://doi.org/10.3109/13880209.2014.885061

Humphreys, J. M., & Chapple, C. (2002). Rewriting the lignin roadmap. *Current Opinion in Plant Biology*, *5*(3), 224-229. https://doi.org/10.1016/S1369-5266(02)00257-1

Huttanus, H. M., & Feng, X. (2017). Compartmentalized metabolic engineering for biochemical and biofuel production. *Biotechnology Journal*, *12*(6), 1700052. https://doi.org/10.1002/biot.201700052

Hwang, G. S., Phuong, N. T., Park, K. R., Kim, Y. H., Kim, K. H., & Kang, J.-S. (2004). Distribution of (-)yatein in cupressaceae family analysed by high performance liquid chromatrography. *Archives of Pharmacal Research*, *27*(1), 35-39. https://doi.org/10.1007/BF02980042

I :

Ibrahim, R. K., De Luca, V., Khouri, H., Latchinian, L., Brisson, L., & Charest, P. M. (1987). Enzymology and compartmentation of polymethylated flavonol glucosides in *Chrysosplenium americanum*. *Phytochemistry*, *26*(5), 1237-1245. https://doi.org/10.1016/S0031-9422(00)81789-6

International Agency for Research on Cancer (Éd.). (2000). *Some antiviral and antineoplastic drugs, and other pharmaceutical agents. 12 - 19 October 1999*. IARC.

Isogai, S., Okahashi, N., Asama, R., Nakamura, T., Hasunuma, T., Matsuda, F., Ishii, J., & Kondo, A. (2021). Synthetic production of prenylated naringenins in yeast using promiscuous microbial prenyltransferases. *Metabolic Engineering Communications*, *12*, e00169. https://doi.org/10.1016/j.mec.2021.e00169

J :

Jackson, D. E., & Dewick, P. M. (1984). Aryltetralin lignans from *Podophyllum hexandrum* and *Podophyllum peltatum*. *Phytochemistry*, *23*(5), 1147-1152. https://doi.org/10.1016/S0031-9422(00)82628-X

Jakočiūnas, T., Jensen, E. D., Jensen, M. K., & Keasling, J. D. (2018). Assembly and Multiplex Genome Integration of Metabolic Pathways in Yeast Using CasEMBLR. *Methods in Molecular Biology*, *1671*, 185-201. https://doi.org/10.1007/978-1-4939-7295-1_12

Jakočiūnas, T., Jensen, M. K., & Keasling, J. D. (2016). CRISPR/Cas9 advances engineering of microbial cell factories. *Metabolic engineering*, *34*. https://doi.org/10.1016/j.ymben.2015.12.003

Jensen, K., & Møller, B. L. (2010). Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry*, *71*(2-3), 132-141. https://doi.org/10.1016/j.phytochem.2009.10.017

Jensen, N. B., Strucko, T., Kildegaard, K. R., David, F., Maury, J., Mortensen, U. H., Forster, J., Nielsen, J., & Borodina, I. (2014). EasyClone : Method for iterative chromosomal integration of multiple genes Saccharomyces cerevisiae. *FEMS Yeast Research*, *14*(2), 238-248. https://doi.org/10.1111/1567-1364.12118

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816-821. https://doi.org/10.1126/science.1225829

Jessop-Fabre, M. M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M. K., Keasling, J. D., & Borodina, I. (2016). EasyClone-MarkerFree : A vector toolkit for marker-less integration of genes into *Saccharomyces cerevisiae* via CRISPR-Cas9. *Biotechnology Journal*, *11*(8), 1110-1117. https://doi.org/10.1002/biot.201600147

Julsing, M. K., Vasilev, N. P., Schneidman-Duhovny, D., Muntendam, R., Woerdenbag, H. J., Quax, W. J., Wolfson, H. J., Ionkova, I., & Kayser, O. (2008). Metabolic stereoselectivity of cytochrome P450 3A4 towards deoxypodophyllotoxin : In silico predictions and experimental validation. *European Journal of Medicinal Chemistry*, *43*(6), 1171-1179. https://doi.org/10.1016/j.ejmech.2007.09.005

K :

Kartal, M., Konuklugil, B., Indrayanto, G., & Alfermann, A. W. (2004). Comparison of different extraction methods for the determination of podophyllotoxin and 6-methoxypodophyllotoxin in Linum species. *Journal of Pharmaceutical and Biomedical Analysis*, *35*(3), 441-447. https://doi.org/10.1016/j.jpba.2004.01.016

Kawai, Y., Ono, E., & Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, *78*(2), 328-343. https://doi.org/10.1111/tpj.12479

Keasling, J. D. (2010). Manufacturing Molecules Through Metabolic Engineering. *Science*, *330*(6009), 1355-1358. https://doi.org/10.1126/science.1193990

Keller, H., Hohlfeld, H., Wray, V., Hahlbrock, K., Scheel, D., & Strack, D. (1996). Changes in the accumulation of soluble and cell wall-bound phenolics in elicitor-treated cell suspension cultures and fungus-infected leaves of *Solanum tuberosum*. *Phytochemistry*, *42*(2), 389-396. https://doi.org/10.1016/0031-9422(95)00866-7

Kellner, F., Kim, J., Clavijo, B. J., Hamilton, J. P., Childs, K. L., Vaillancourt, B., Cepela, J., Habermann, M., Steuernagel, B., Clissold, L., McLay, K., Buell, C. R., & O'Connor, S. E. (2015). Genome-guided investigation of plant natural product biosynthesis. *The Plant Journal*, *82*(4), 680-692. https://doi.org/10.1111/tpj.12827

Kennedy, E. P. (1961). Biosynthesis of Complex Lipids. *Federation Proceedings*, *20*, 934-940.

Kim, J.-E., Jang, I.-S., Son, S.-H., Ko, Y.-J., Cho, B.-K., Kim, S. C., & Lee, J. Y. (2019). Tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metabolic Engineering*, *56*, 50-59. https://doi.org/10.1016/j.ymben.2019.08.013

Kim, J.-E., Jang, I.-S., Sung, B. H., Kim, S. C., & Lee, J. Y. (2018). Rerouting of NADPH synthetic pathways for increased protopanaxadiol production in Saccharomyces cerevisiae. *Scientific Reports*, *8*(1), 15820. https://doi.org/10.1038/s41598-018-34210-3

Kite, G. C., Larsson, S., Veitch, N. C., Porter, E. A., Ding, N., & Simmonds, M. S. J. (2013). Acyl Spermidines in Inflorescence Extracts of Elder (*Sambucus nigra L.*, Adoxaceae) and Elderflower Drinks. *Journal of Agricultural and Food Chemistry*, *61*(14), 3501-3508. https://doi.org/10.1021/jf304602q

Klapheck, S. (1983). Polyamines and Cinnamoyl-Putrescines in Normal and Sulfur-Starved Suspension Cultures of *Nicotiana tabacum*. *Zeitschrift Für Pflanzenphysiologie*, *112*(3), 275-279. https://doi.org/10.1016/S0044-328X(83)80141-X

Köhler, U., Kubny, S., & Wiesemann, J. (2006). Eurofins' adaption and optimisation software "GENEius" in comparison to other optimisation algorithms. *Gene Synthesis Department of Eurofins Genomics, Ebersberg*.

Königshofer, H., & Lechner, S. (2002). Are polyamines involved in the synthesis of heat-shock proteins in cell suspension cultures of tobacco and alfafa in response to high-temperature stress? *Plant Physiology and Biochemistry*, *40*, 51-59.

Koudounas, K., Carqueijeiro, I., Lemos Cruz, P., Perrin, J., Lanoue, A., Oudin, A., Besseau, S., & Courdavault, V. (2022). A Rapid and Efficient Vacuum-Based Agroinfiltration Protocol for Transient Gene Overexpression in Leaves of Catharanthus roseus. In V. Courdavault & S.

Besseau (Éds.), *Catharanthus roseus* (Vol. 2505, p. 263-279). Springer US. https://doi.org/10.1007/978-1-0716-2349-7_19

Koudounas, K., Guirimand, G., Hoyos, L. F. R., Carqueijeiro, I., Cruz, P. L., Stander, E., Kulagina, N., Perrin, J., Oudin, A., Besseau, S., Lanoue, A., Atehortùa, L., St-Pierre, B., Giglioli-Guivarc'h, N., Papon, N., O'Connor, S. E., & Courdavault, V. (2022). Tonoplast and Peroxisome Targeting of γ-tocopherol *N* -methyltransferase Homologs Involved in the Synthesis of Monoterpene Indole Alkaloids. *Plant and Cell Physiology*, *63*(2), 200-216. https://doi.org/10.1093/pcp/pcab160

Kour, A., Shawl, A. S., Rehman, S., Sultan, P., Qazi, P. H., Suden, P., Khajuria, R. K., & Verma, V. (2008). Isolation and identification of an endophytic strain of Fusarium oxysporum producing podophyllotoxin from Juniperus recurva. *World Journal of Microbiology and Biotechnology*, *24*(7), 1115-1121. https://doi.org/10.1007/s11274-007-9582-5

Kulagina, N., Besseau, S., Godon, C., Goldman, G. H., Papon, N., & Courdavault, V. (2021a). Yeasts as Biopharmaceutical Production Platforms. *Frontiers in Fungal Biology*, *2*, 733492. https://doi.org/10.3389/ffunb.2021.733492

Kulagina, N., Besseau, S., Papon, N., & Courdavault, V. (2021b). Peroxisomes : A New Hub for Metabolic Engineering in Yeast. *Frontiers in Bioengineering and Biotechnology*, *9*, 659431. https://doi.org/10.3389/fbioe.2021.659431

Kulagina, N., Guirimand, G., Melin, C., Lemos-Cruz, P., Carqueijeiro, I., De Craene, J., Oudin, A., Heredia, V., Koudounas, K., Unlubayir, M., Lanoue, A., Imbault, N., St-Pierre, B., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., Marc, J., Besseau, S., & Courdavault, V. (2021c). Enhanced bioproduction of anticancer precursor vindoline by yeast cell factories. *Microbial Biotechnology*, *14*(6), 2693-2699. https://doi.org/10.1111/1751-7915.13898

Kulagina, N., Perrin, J., Besseau, S., & Courdavault, V. (2021d). Efficient Terpene Production by Marine Thraustochytrids : Shedding Light on the Thermodynamic Driving Force. *MBio*, *12*(5), e01976-21. https://doi.org/10.1128/mBio.01976-21

Kumagai, M. H., Donson, J., della-Cioppa, G., Harvey, D., Hanley, K., & Grill, L. K. (1995). Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proceedings of the National Academy of Sciences*, *92*(5), 1679-1683. https://doi.org/10.1073/pnas.92.5.1679

Kumar, P., Pal, T., Sharma, N., Kumar, V., Sood, H., & Chauhan, R. S. (2015). Expression analysis of biosynthetic pathway genes vis-à-vis podophyllotoxin content in *Podophyllum hexandrum* Royle. *Protoplasma*, *252*(5), 1253-1262. https://doi.org/10.1007/s00709-015-0757-x

Kusari, S., Lamshöft, M., & Spiteller, M. (2009). *Aspergillus fumigatus* Fresenius, an endophytic fungus from *Juniperus communis* L. Horstmann as a novel source of the anticancer pro-drug deoxypodophyllotoxin. *Journal of Applied Microbiology*, *107*(3), 1019-1030. https://doi.org/10.1111/j.1365-2672.2009.04285.x

Kyselka, J., Bleha, R., Dragoun, M., Bialasová, K., Horáčková, Š., Schätz, M., Sluková, M., Filip, V., & Synytsya, A. (2018). Antifungal Polyamides of Hydroxycinnamic Acids from Sunflower Bee Pollen. *Journal of Agricultural and Food Chemistry*, *66*(42), 11018-11026. https://doi.org/10.1021/acs.jafc.8b03976

L :

Lam, K. C., Ibrahim, R. K., Behdad, B., & Dayanandan, S. (2007). Structure, function, and evolution of plant *O* -methyltransferases. *Genome*, *50*(11), 1001-1013. https://doi.org/10.1139/G07-077

Lanza, A. M., Curran, K. A., Rey, L. G., & Alper, H. S. (2014). A condition-specific codon optimization approach for improved heterologous gene expression in Saccharomyces cerevisiae. *BMC Systems Biology*, *8*(1), 33. https://doi.org/10.1186/1752-0509-8-33

Larbat, R., Paris, C., Le Bot, J., & Adamowicz, S. (2014). Phenolic characterization and variability in leaves, stems and roots of Micro-Tom and patio tomatoes, in response to nitrogen limitation. *Plant Science*, *224*, 62-73. https://doi.org/10.1016/j.plantsci.2014.04.010

Lau, W., & Sattely, E. S. (2015). Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science*, *349*(6253), 1224-1228. https://doi.org/10.1126/science.aac7202

Laverty, T. M., & Plowright, R. C. (1988). Fruit and seed set in Mayapple (*Podophyllum peltatum*) : Influence of intraspecific factors and local enhancement near *Pedicularis canadensis*. *Canadian Journal of Botany*, *66*(1), 173-178. https://doi.org/10.1139/b88-027

Leander, K., & Rosen. (1988). Medicinal use for podophyllotoxin. *U.S. patent*, *4:788:216*.

Lee, P. C., Yoon, Y. geol, & Schmidt-Dannert, C. (2009). Investigation of cellular targeting of carotenoid pathway enzymes in Pichia pastoris. *Journal of Biotechnology*, *140*(3-4), 227-233. https://doi.org/10.1016/j.jbiotec.2009.01.019

Lemos Cruz, P., Carqueijeiro, I., Koudounas, K., Bomzan, D. P., Stander, E. A., Abdallah, C., Kulagina, N., Oudin, A., Lanoue, A., Giglioli-Guivarc'h, N., Nagegowda, D. A., Papon, N., Besseau, S., Clastre, M., & Courdavault, V. (2022). Identification of a second 16-hydroxytabersonine-O-methyltransferase suggests an evolutionary relationship between alkaloid and flavonoid metabolisms in Catharanthus roseus. *Protoplasma*. https://doi.org/10.1007/s00709-022-01801-x

Lemos Cruz, P., Kulagina, N., Guirimand, G., De Craene, J.-O., Besseau, S., Lanoue, A., Oudin, A., Giglioli-Guivarc'h, N., Papon, N., Clastre, M., & Courdavault, V. (2021). Optimization of Tabersonine Methoxylation to Increase Vindoline Precursor Synthesis in Yeast Cell Factories. *Molecules*, *26*(12), 3596. https://doi.org/10.3390/molecules26123596

Lemos Cruz, P., Restrepo, M. I., Dugé de Bernonville, T., Oudin, A., Munsch, T., Lanoue, A., Besseau, S., Atehortùa, L., Giglioli-Guivarc'h, N., Papon, N., Clastre, M., Carqueijeiro, I., & Courdavault, V. (2020). A Biolistic-Mediated Virus-Induced Gene Silencing in Apocynaceae to Map Biosynthetic Pathways of Alkaloids. *Methods in Molecular Biology*, *2172*, 93-110. https://doi.org/10.1007/978-1-0716-0751-0_8

Lepeillet, J. (2018). *Ingénierie métabolique de la production de podophyllotoxine en système hétérologue* [Mémoire de stage de Master 2].

Lerndal, T. (2000). A clinical study of CPH 82 vs methotrexate in early rheumatoid arthritis. *Rheumatology*, *39*(3), 316-320. https://doi.org/10.1093/rheumatology/39.3.316

Levac, D., Murata, J., Kim, W. S., & De Luca, V. (2008). Application of carborundum abrasion for investigating the leaf epidermis : Molecular cloning of Catharanthus roseus 16-hydroxytabersonine-16-O-methyltransferase: 16-methoxytabersonine biosynthesis

inCatharanthus. *The Plant Journal*, *53*(2), 225-236. https://doi.org/10.1111/j.1365-313X.2007.03337.x

Lewis, N. G. & Yamamoto, E. (1990). Lignin : Occurrence, Biogenesis and Biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology 41*(1), 455-496.

Lewis, N. G., Davin, L. B. & Sarkanen, S. (1998). Lignin and Lignan Biosynthesis : distinctions and reconciliations. *ACS Symposium Series, American Chemical Society 697*, 1-27. https://doi.org/10.1021/bk-1998-0697

Li, M., Kildegaard, K. R., Chen, Y., Rodriguez, A., Borodina, I., & Nielsen, J. (2015). *De novo* production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. *Metabolic Engineering*, *32*, 1-11. https://doi.org/10.1016/j.ymben.2015.08.007

Li, M., Zhou, L., Yang, D., Li, T., & Li, W. (2012). Biochemical composition and antioxidant capacity of extracts from *Podophyllum hexandrum* rhizome. *BMC Complementary and Alternative Medicine*, *12*(1), 263. https://doi.org/10.1186/1472-6882-12-263

Li, W., Li, M., Yang, D., Xu, R., & Zhang, Y. (2009). Production posophyllotoxin by root culture of *Podophyllum hexandrum* Royle. *Electronic Journal of Biology*, *5*(2), 6.

Li, Y., Li, S., Thodey, K., Trenchard, I., Cravens, A., & Smolke, C. D. (2018). Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proceedings of the National Academy of Sciences*, *115*(17). https://doi.org/10.1073/pnas.1721469115

Li, Z., Zhao, C., Zhao, X., Xia, Y., Sun, X., Xie, W., Ye, Y., Lu, X., & Xu, G. (2018). Deep Annotation of Hydroxycinnamic Acid Amides in Plants Based on Ultra-High-Performance Liquid Chromatography–High-Resolution Mass Spectrometry and Its In Silico Database. *Analytical Chemistry*, *90*(24), 14321-14330. https://doi.org/10.1021/acs.analchem.8b03654

Liao, M., Faouzi, S., Karyakin, A., & Almira Correia, M. (2006). Endoplasmic Reticulum-Associated Degradation of Cytochrome P450 CYP3A4 in *Saccharomyces cerevisiae* : Further Characterization of Cellular Participants and Structural Determinants. *Molecular Pharmacology*, *69*, 1897-1904. https://doi.org/10.1124/mol.105.021816

Lin, S., & Mullin, C. A. (1999). Lipid, Polyamide, and Flavonol Phagostimulants for Adult Western Corn Rootworm from Sunflower (*Helianthus annuus* L.) Pollen. *Journal of Agricultural and Food Chemistry*, *47*(3), 1223-1229. https://doi.org/10.1021/jf980858w

Liscombe, D. K., & O'Connor, S. E. (2011). A virus-induced gene silencing approach to understanding alkaloid metabolism in Catharanthus roseus. *Phytochemistry*, *72*(16), 1969-1977. https://doi.org/10.1016/j.phytochem.2011.07.001

Liu, C.-J., Deavours, B. E., Richard, S. B., Ferrer, J.-L., Blount, J. W., Huhman, D., Dixon, R. A., & Noel, J. P. (2006). Structural Basis for Dual Functionality of Isoflavonoid *O* -Methyltransferases in the Evolution of Plant Defense Responses. *The Plant Cell*, *18*(12), 3656-3669. https://doi.org/10.1105/tpc.106.041376

Liu, G.-S., Li, T., Zhou, W., Jiang, M., Tao, X.-Y., Liu, M., Zhao, M., Ren, Y.-H., Gao, B., Wang, F.-Q., & Wei, D.-Z. (2020a). The yeast peroxisome : A dynamic storage depot and subcellular factory for squalene overproduction. *Metabolic Engineering*, *57*, 151-161. https://doi.org/10.1016/j.ymben.2019.11.001

Liu, H., Liao, J.-X., Hu, Y., Tu, Y.-H., & Sun, J.-S. (2016). A Highly Efficient Approach To Construct (*epi*)-Podophyllotoxin-4- *O* -glycosidic Linkages as well as Its Application in Concise

Syntheses of Etoposide and Teniposide. *Organic Letters*, *18*(6), 1294-1297. https://doi.org/10.1021/acs.orglett.6b00216

Liu, R., Liu, L., Li, X., Liu, D., & Yuan, Y. (2020b). Engineering yeast artificial core promoter with designated base motifs. *Microbial Cell Factories*, *19*(1), 38. https://doi.org/10.1186/s12934-020-01305-4

Liu, T., Gou, Y., Zhang, B., Gao, R., Dong, C., Qi, M., Jiang, L., Ding, X., Li, C., & Lian, J. (2022). Construction of ajmalicine and sanguinarine de novo biosynthetic pathways using stable integration sites in yeast. *Biotechnology and Bioengineering*, *119*(5), 1314-1326. https://doi.org/10.1002/bit.28040

Liu, T., Zhang, C., & Lu, W. (2018). Heterologous production of levopimaric acid in Saccharomyces cerevisiae. *Microbial Cell Factories*, *17*(1), 114. https://doi.org/10.1186/s12934-018-0964-1

Long, B., & Minocha, A. A. (1983). Inhibition of topoisomerase-II by VP-16-213 (etoposide) VM-26 (teniposide), and structural congeners as an explanation for in vivo DNA breakage and cytotoxicity. *). Proceedings of the American Association for Cancer Research, PA 19106-4404 USA : Amer Assoc Cancer Reasearch*, 321.

Lopez-Nieves, S., Yang, Y., Timoneda, A., Wang, M., Feng, T., Smith, S. A., Brockington, S. F., & Maeda, H. A. (2018). Relaxation of tyrosine pathway regulation underlies the evolution of betalain pigmentation in Caryophyllales. *New Phytologist*, *217*(2), 896-908. https://doi.org/10.1111/nph.14822

Lou, Y.-R., Bor, M., Yan, J., Preuss, A. S., & Jander, G. (2016). Arabidopsis NATA1 acetylates putrescine and decreases defense-related hydrogen peroxide accumulation. *Plant Physiology 171*, 1443-1455. https://doi.org/10.1104/pp.16.00446

Lu, R., Martin-Hernandez, A. M., Peart, J. R., Malcuit, I., & Baulcombe, D. C. (2003). Virus-induced gene silencing in plants. *Methods*, *30*(4), 296-303. https://doi.org/10.1016/S1046-2023(03)00037-9

Luo, H., Hansen, A. S. L., Yang, L., Schneider, K., Kristensen, M., Christensen, U., Christensen, H. B., Du, B., Özdemir, E., Feist, A. M., Keasling, J. D., Jensen, M. K., Herrgård, M. J., & Palsson, B. O. (2019). Coupling S-adenosylmethionine–dependent methylation to growth : Design and uses. *PLOS Biology*, *17*(3), e2007050. https://doi.org/10.1371/journal.pbio.2007050

Luo, J., Fuell, C., Parr, A., Hill, L., Bailey, P., Elliott, K., Fairhurst, S. A., Martin, C., & Michael, A. J. (2009). A Novel Polyamine Acyltransferase Responsible for the Accumulation of Spermidine Conjugates in *Arabidopsis* Seed. *The Plant Cell*, *21*(1), 318-333. https://doi.org/10.1105/tpc.108.063511

M :

Ma, C.-M., Nakamura, N., & Hattori, M. (2001). Inhibitory Effects on HIV-1 Protease of Tri-p-coumaroylspermidine from *Artemisia caruifolia* and Related Amides. *Chemical and Pharmaceutical Bulletin*, *49*(7), 915-917. https://doi.org/10.1248/cpb.49.915

Ma, H., Kunes, S., Schatz, P. J., & Botstein, D. (1987). Plasmid construction by homologous recombination in yeast. *Gene*, *58*(2-3), 201-216. https://doi.org/10.1016/0378-1119(87)90376-3

MacRae, W. D., & Towers, G. H. N. (1984). Biological activities of lignans. *Phytochemistry*, *23*(6), 1207-1220. https://doi.org/10.1016/S0031-9422(00)80428-8

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., & Koonin, E. V. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, *13*(11), 722-736. https://doi.org/10.1038/nrmicro3569

Maqbool, M. (2011). Mayapple : A review of the literature from a horticultural perspective. *Journal of Medicinal Plants Research 5*, 1037-1045.

Marques, J. V., Kim, K.-W., Lee, C., Costa, M. A., May, G. D., Crow, J. A., Davin, L. B., & Lewis, N. G. (2013). Next Generation Sequencing in Predicting Gene Function in Podophyllotoxin Biosynthesis. *Journal of Biological Chemistry*, *288*(1), 466-479. https://doi.org/10.1074/jbc.M112.400689

Martin-Tanguy, J. (1997). Conjugated polyamines and reproductive development : Biochemical, molecular and physiological approaches. *Physiologia Plantarum*, *100*, 675-688.

Martin-Tanguy, J. (1985). The occurrence and possible function of hydroxycinnamoyl acid amides in plants. *Plant Growth Regulation*, *3*, 381-399.

Martin-Tanguy, J., Martin, C., Gallet, M., & Vernoy, R. (1976). Sur de puissants inhibiteurs de multiplication du virus de la mosaïque du tabac. *Académie des Sciences*, *282*, 2231-2234.

Martin-Tanguy, J., Martin, C., & Gallet, M. (1973). Présence de composés aromatiques liés à la putrescine dans divers Nicotiana virosés. *Académie des Sciences*, *276*, 1433-1435.

Matsuno, M., Compagnon, V., Schoch, G. A., Schmitt, M., Debayle, D., Bassard, J.-E., Pollet, B., Hehn, A., Heintz, D., Ullmann, P., Lapierre, C., Bernier, F., Ehlting, J., & Werck-Reichhart, D. (2009). Evolution of a Novel Phenolic Pathway for Pollen Development. *Science*, *325*(5948), 1688-1692. https://doi.org/10.1126/science.1174095

Mbadiwe, A. I. (1973). Caffeoylputrescine from *Pentaclethra macrophylla*. *Phytochemical Reports*, *12*, 2546.

Medema, M. H., & Osbourn, A. (2016). Computational genomic identification and functional reconstitution of plant natural product biosynthetic pathways. *Natural Product Reports*, *33*(8), 951-962. https://doi.org/10.1039/C6NP00035E

Meresse, P., Dechaux, E., Monneret, C., & Bertounesque, E. (2004). Etoposide : Discovery and Medicinal Chemistry. *Current Medicinal Chemistry*, *11*(18), 2443-2466. https://doi.org/10.2174/0929867043364531

Meurer, B., Wiermann, R., & Strack, D. (1988). Phenylpropanoid patterns in Fagales pollen and their phylogenetic relevance. *Phytochemistry*, *27*(3), 823-828. https://doi.org/10.1016/0031-9422(88)84100-1

Meurer, B., Wray, V., Grotjahn, L., Wiermann, R., & Strack, D. (1986). Hydroxycinnamic acid spermidine amides from pollen of *Corylus avellana L*. *Phytochemistry*, *25*(2), 433-435. https://doi.org/10.1016/S0031-9422(00)85496-5

Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., & Halkier, B. A. (2012). Microbial production of indolylglucosinolate through engineering of a

multi-gene pathway in a versatile yeast expression platform. *Metabolic Engineering*, *14*(2), 104-111. https://doi.org/10.1016/j.ymben.2012.01.006

Miyata, R., Hoshino, S., Ahn, M.-R., & Kumazawa, S. (2022). Chemical Profiles of Korean Bee Pollens and Their Catechol-O-methyltransferase Inhibitory Activities. *Journal of Agricultural and Food Chemistry*, *70*(4), 1174-1181. https://doi.org/10.1021/acs.jafc.1c07778

Mizutani, M., & Sato, F. (2011). Unusual P450 reactions in plant secondary metabolism. *Archives of Biochemistry and Biophysics*, *507*(1), 194-203. https://doi.org/10.1016/j.abb.2010.09.026

Moraes-Cerdeira, R., Burandt, C., Bastos, J., Nanayakkara, N., & McChesney, J. (1998). *In Vitro* Propagation of *Podophyllum peltatum*. *Planta Medica*, *64*(01), 42-45. https://doi.org/10.1055/s-2006-957363

Moreau, R. A., Nuñez, A., & Singh, V. (2001). Diferuloylputrescine and *p*-coumaroyl-feruloylputrescine, abundant polyamine conjugates in lipid extracts of maize kernels. *Lipids*, *36*(8), 839-844. https://doi.org/10.1007/s11745-001-0793-6

Moses, T., Pollier, J., Shen, Q., Soetaert, S., Reed, J., Erffelinck, M.-L., Van Nieuwerburgh, F. C. W., Vanden Bossche, R., Osbourn, A., Thevelein, J. M., Deforce, D., Tang, K., & Goossens, A. (2015). OSC2 and CYP716A14v2 Catalyze the Biosynthesis of Triterpenoids for the Cuticle of Aerial Organs of *Artemisia annua*. *The Plant Cell*, *27*(1), 286-301. https://doi.org/10.1105/tpc.114.134486

Moss, G. P. (2000). Nomenclature of lignans and neolignans. *Pure and Applied Chemistry*, *72*(8), 31.

Muroi, A., Ishihara, A., Tanaka, C., Ishizuka, A., Takabayashi, J., Miyoshi, H., & Nishioka, T. (2009). Accumulation of hydroxycinnamic acid amides induced by pathogen infection and identification of agmatine coumaroyltransferase in *Arabidopsis thaliana*. *Planta*, *230*(3), 517-527. https://doi.org/10.1007/s00425-009-0960-0

Murphy, M. B., Kumar, P., Bradley, A. M., Barton, C. E., Deweese, J. E., & Mercer, S. L. (2020). Synthesis and evaluation of etoposide and podophyllotoxin analogs against topoisomerase IIα and HCT-116 cells. *Bioorganic & Medicinal Chemistry*, *28*(22), 115773. https://doi.org/10.1016/j.bmc.2020.115773

Murray, E. E., Lotzer, J., & Eberle, M. (1989). Codon usage in plant genes. *Nucleic Acids Research*, *17*(2), 477-498. https://doi.org/10.1093/nar/17.2.477

N :

Nadeem, M., Ram, M., Alam, P., Ahmad, M. M., Mohammad, A., Al-Qurainy, F., Khan, S., & Abdin, M. Z. (2012). *Fusarium solani*, P1, a new endophytic podophyllotoxin-producing fungus from roots of *Podophyllum hexandrum*. *African Journal of Microbiology Research*, *6*(10). https://doi.org/10.5897/AJMR11.1596

Nadeem, M., Palni, L. M. S., Purohit, A. N., Pandey, H., & Nandi, S. K. (2000). Propagation and conservation of Podophyllum hexandrum Royle : An important medicinal herb. *Biological Conservation*, *92*(1), 121-129. https://doi.org/10.1016/S0006-3207(99)00059-2

Negrel, J., & Jeandet, P. (1987). Metabolism of tyramine and feruloyltyramine in TMV inoculated leaves of *Nicotiana tabacum*. *Phytochemistry*, *26*, 2185-2190.

Negrel, J., & Martin, C. (1984). The biosynthesis of feruloyltyramine in *Nicotiana tabacum*. *Phytochemistry*, *23*, 2797-2801.

Nelson, D. R., Schuler, M. A., Paquette, S. M., Werck-Reichhart, D., & Bak, S. (2004a). Comparative Genomics of Rice and Arabidopsis. Analysis of 727 Cytochrome P450 Genes and Pseudogenes from a Monocot and a Dicot. *Plant Physiology*, *135*(2), 756-772. https://doi.org/10.1104/pp.104.039826

Nelson, D. R., Zeldin, D. C., Hoffman, S. M., Maltais, L. J., Wain, H. M., & Nebert, D. W. (2004b). Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants: *Pharmacogenetics*, *14*(1), 1-18. https://doi.org/10.1097/00008571-200401000-00001

Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W., Gunsalus, I. C., & Nebert, D. W. (1996). P450 superfamily : Update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, *6*(1), 1-42. https://doi.org/10.1097/00008571-199602000-00002

Newman, M.-A., von Roepenack-Lahaye, E., Parr, A., Daniels, M. J., & Dow, J. M. (2001). Induction of Hydroxycinnamoyl-Tyramine Conjugates in Pepper by *Xanthomonas campestris* , a Plant Defense Response Activated by *hrp* Gene-Dependent and *hrp* Gene-Independent Mechanisms. *Molecular Plant-Microbe Interactions®*, *14*(6), 785-792. https://doi.org/10.1094/MPMI.2001.14.6.785

Nikel, P. I., & de Lorenzo, V. (2018). *Pseudomonas putida* as a functional chassis for industrial biocatalysis : From native biochemistry to trans-metabolism. *Metabolic Engineering*, *50*, 142-155. https://doi.org/10.1016/j.ymben.2018.05.005

O :

Okimoto, R., Sachs, M. M., Porter, E. K., & Freeling, M. (1980). Patterns of polypeptide synthesis in various maize organs under anaerobiosis. *Planta*, *150*(1), 89-94. https://doi.org/10.1007/BF00385619

Olson, J. A., & Hayaishi, O. (1965). The enzymatic cleavage of β-carotene into vitamin A by soluble enzymes of rat liver and intestine. *Proceedings of the National Academy of Sciences*, *54*(5), 1364-1370. https://doi.org/10.1073/pnas.54.5.1364

Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, I. T., & Galis, I. (2012). MYB8 Controls Inducible Phenolamide Levels by Activating Three Novel Hydroxycinnamoyl-Coenzyme A:Polyamine Transferases in *Nicotiana attenuata*. *Plant Physiology*, *158*(1), 389-407. https://doi.org/10.1104/pp.111.187229

Ortiz de Montellano, P. R., & De Voss, J. J. (2005). Substrate Oxidation by Cytochrome P450 Enzymes. *Cytochrome P450*, 183-245.

P :

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., … & Newman, J. D. (2013). High-level semi-synthetic production of the potent antimalarial

artemisinin. *Nature*, *496*(7446), 528-532. https://doi.org/10.1038/nature12051

Paquette, S. M., Bak, S., & Feyereisen, R. (2000). Intron–Exon Organization and Phylogeny in a Large Superfamily, the Paralogous Cytochrome P450 Genes of *Arabidopsis thaliana*. *DNA and Cell Biology*, *19*(5), 307-317. https://doi.org/10.1089/10445490050021221

Parage, C., Foureau, E., Kellner, F., Burlat, V., Mahroug, S., Lanoue, A., Dugé de Bernonville, T., Londono, M. A., Carqueijeiro, I., Oudin, A., Besseau, S., Papon, N., Glévarec, G., Atehortùa, L., Giglioli-Guivarc'h, N., St-Pierre, B., Clastre, M., O'Connor, S. E., & Courdavault, V. (2016). Class II Cytochrome P450 Reductase Governs the Biosynthesis of Alkaloids. *Plant Physiology*, *172*(3), 1563-1577. https://doi.org/10.1104/pp.16.00801

Paramasivan, K., & Mutturi, S. (2017). Regeneration of NADPH Coupled with HMG-CoA Reductase Activity Increases Squalene Synthesis in *Saccharomyces cerevisiae*. *Journal of Agricultural and Food Chemistry*, *65*(37), 8162-8170. https://doi.org/10.1021/acs.jafc.7b02945

Park, S., Song, K., & Kim, Y. (2017). Tetra-cis/trans-Coumaroyl Polyamines as NK1 Receptor Antagonists from *Matricaria chamomilla*. *Planta Medica International Open*, *4*(02), e43-e51. https://doi.org/10.1055/s-0043-106742

Partow, S., Siewers, V., Bjørn, S., Nielsen, J., & Maury, J. (2010). Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. *Yeast*, *27*(11), 955-964. https://doi.org/10.1002/yea.1806

Payne, R. M. E., Xu, D., Foureau, E., Teto Carqueijeiro, M. I. S., Oudin, A., Bernonville, T. D. de, Novak, V., Burow, M., Olsen, C.-E., Jones, D. M., Tatsis, E. C., Pendle, A., Ann Halkier, B., Geu-Flores, F., Courdavault, V., Nour-Eldin, H. H., & O'Connor, S. E. (2017). An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. *Nature Plants*, *3*(2), 16208. https://doi.org/10.1038/nplants.2016.208

Pearce, D. K., & Thieret, J. W. (1993). Persimmon (*Diospyros virginiana*, Ebenaceae) and mayapple (*Podophyllum peltatum*, Berberidaceae) : Proximate analysis of their fruits. *ransactions of the Kentucky Academy of Science (USA)*, *54*(1-2), 30-31.

Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., & Vickers, C. E. (2015). Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift : A comparison of yeast promoter activities. *Microbial Cell Factories*, *14*(1), 91. https://doi.org/10.1186/s12934-015-0278-5

Peng, H., Yang, T., Whitaker, B. D., Trouth, F., Shangguan, L., Dong, W., & Jurick, W. M. (2016). Characterization of spermidine hydroxycinnamoyl transferases from eggplant (*Solanum melongena L.*) and its wild relative *Solanum richardii* Dunal. *Horticulture Research*, *3*(1), 16062. https://doi.org/10.1038/hortres.2016.62

Perrin, J., Besseau, S., Papon, N., & Courdavault, V. (2022). Boosting lignan-precursor synthesis in yeast cell factories through co-factor supply optimization. *Frontiers in Bioengineering and Biotechnology*, *10*, 1079801. https://doi.org/10.3389/fbioe.2022.1079801

Perrin, J., Kulagina, N., Unlubayir, M., Munsch, T., Carqueijeiro, I., Dugé de Bernonville, T., De Craene, J.-O., Clastre, M., St-Pierre, B., Giglioli-Guivarc'h, N., Gagneul, D., Lanoue, A., Courdavault, V., & Besseau, S. (2021). Exploiting Spermidine *N* - Hydroxycinnamoyltransferase Diversity and Substrate Promiscuity to Produce Various Trihydroxycinnamoyl Spermidines and Analogues in Engineered Yeast. *ACS Synthetic Biology*, *10*(2), 286-296. https://doi.org/10.1021/acssynbio.0c00391

Petersen, S. D., Zhang, J., Lee, J. S., Jakočiūnas, T., Grav, L. M., Kildegaard, H. F., Keasling, J. D., & Jensen, M. K. (2018). Modular 5′-UTR hexamers for context-independent tuning of protein expression in eukaryotes. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gky734

Pichersky, E., & Gershenzon, J. (2002). The formation and function of plant volatiles : Perfumes for pollinator attraction and defense. *Current Opinion in Plant Biology*, *5*(3), 237-243. https://doi.org/10.1016/S1369-5266(02)00251-0

Pickens, L. B., Tang, Y., & Chooi, Y.-H. (2011). Metabolic Engineering for the Production of Natural Products. *Annual Review of Chemical and Biomolecular Engineering*, *2*(1), 211-236. https://doi.org/10.1146/annurev-chembioeng-061010-114209

Podwyssotzki, V. (1881). The active constituents of podophyllotoxin. *Pharmaceutical Journal and Transactions.*, *12*, 217-218.

Polturak, G., & Aharoni, A. (2019). Advances and future directions in betalain metabolic engineering. *New Phytologist*, *224*(4), 1472-1478. https://doi.org/10.1111/nph.15973

Pott, D. M., Osorio, S., & Vallarino, J. G. (2019). From Central to Specialized Metabolism : An Overview of Some Secondary Compounds Derived From the Primary Metabolism for Their Role in Conferring Nutritional and Organoleptic Characteristics to Fruit. *Frontiers in Plant Science*, *10*, 835. https://doi.org/10.3389/fpls.2019.00835

Procopio-Melino, R., Kotch, F. W., Prashad, A. S., Gomes, J. M., Wang, W., Arve, B., Dawdy, A., Chen, L., Sperry, J., Hosselet, C., He, T., Kriz, R., Lin, L., Marquette, K., Tchistiakova, L., Somers, W., Rouse, J. C., & Zhong, X. (2022). Cysteine metabolic engineering and selective disulfide reduction produce superior antibody-drug-conjugates. *Scientific Reports*, *12*(1), 7262. https://doi.org/10.1038/s41598-022-11344-z

Promdonkoy, P., Mhuantong, W., Champreda, V., Tanapongpipat, S., & Runguphan, W. (2020). Improvement in D-xylose utilization and isobutanol production in *S. cerevisiae* by adaptive laboratory evolution and rational engineering. *Journal of Industrial Microbiology and Biotechnology*, *47*(6-7), 497-510. https://doi.org/10.1007/s10295-020-02281-9

Puri, S. C., Nazir, A., Chawla, R., Arora, R., Riyaz-ul-Hasan, S., Amna, T., Ahmed, B., Verma, V., Singh, S., Sagar, R., Sharma, A., Kumar, R., Sharma, R. K., & Qazi G. N. (2006). The endophytic fungus *Trametes hirsuta* as a novel alternative source of podophyllotoxin and related aryl tetralin lignans. *Journal of Biotechnology*, *122*(4), 494-510. https://doi.org/10.1016/j.jbiotec.2005.10.015

Pyne, M. E., Kevvai, K., Grewal, P. S., Narcross, L., Choi, B., Bourgeois, L., Dueber, J. E., & Martin, V. J. J. (2020). A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nature Communications*, *11*(1), 3337. https://doi.org/10.1038/s41467-020-17172-x

Q :

Qin, J., Krivoruchko, A., Ji, B., Chen, Y., Kristensen, M., Özdemir, E., Keasling, J. D., Jensen, M. K., & Nielsen, J. (2021). Engineering yeast metabolism for the discovery and production of polyamines and polyamine analogues. *Nature Catalysis*, *4*(6), 498-509. https://doi.org/10.1038/s41929-021-00631-z

Qiu, D., Wang, M., Zhou, C., Zhao, J., & Zhang, G. (2022). De novo biosynthesis of vanillin in engineered Saccharomyces cerevisiae. *Chemical Engineering Science*, *263*, 118049. https://doi.org/10.1016/j.ces.2022.118049

R :

Rabeharindranto, H., Castaño-Cerezo, S., Lautier, T., Garcia-Alles, L. F., Treitz, C., Tholey, A., & Truan, G. (2019). Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in S. cerevisiae. *Metabolic Engineering Communications*, *8*(e00086). https://doi.org/10.1016/j.mec.2019.e00086

Rabin, R. C. (2019). Faced With a Drug Shortfall, Doctors Scramble to Treat Children With Cancer. *New York Times*. https://www.nytimes.com/2019/10/14/health/cancer-drug-shortage.html

Rajkumar, A. S., Özdemir, E., Lis, A. V., Schneider, K., Qin, J., Jensen, M. K., & Keasling, J. D. (2019). Engineered Reversal of Function in Glycolytic Yeast Promoters. *ACS Synthetic Biology*, *8*(6), 1462-1468. https://doi.org/10.1021/acssynbio.9b00027

Rajkumar, A. S., Liu, G., Bergenholm, D., Arsovska, D., Kristensen, M., Nielsen, J., Jensen, M. K., & Keasling, J. D. (2016). Engineering of synthetic, stress-responsive yeast promoters. *Nucleic Acids Research*, *44*(17), e136-e136. https://doi.org/10.1093/nar/gkw553

Renault, H., Bassard, J.-E., Hamberger, B., & Werck-Reichhart, D. (2014). Cytochrome P450-mediated metabolic engineering : Current progress and future challenges. *Current Opinion in Plant Biology*, *19*, 27-34. https://doi.org/10.1016/j.pbi.2014.03.004

Riley, L. A., & Guss, A. M. (2021). Approaches to genetic tool development for rapid domestication of non-model microorganisms. *Biotechnology for Biofuels*, *14*(1), 30. https://doi.org/10.1186/s13068-020-01872-z

Ro, D.-K., Ouellet, M., Paradise, E. M., Burd, H., Eng, D., Paddon, C. J., Newman, J. D., & 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 Biotechnology*, *8*(1), 83. https://doi.org/10.1186/1472-6750-8-83

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., & Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, *440*(7086), 940-943. https://doi.org/10.1038/nature04640

Ross, J. J., Tivendale, N. D., Davidson, S. E., Reid, J. B., Davies, N. W., Quittenden, L. J., & Smith, J. A. (2012). A mutation affecting the synthesis of 4-chloroindole-3-acetic acid. *Plant Signaling & Behavior*, *7*(12), 1533-1536. https://doi.org/10.4161/psb.22319

Ross, W., Rowe, T., Glisson, B., Yalowich, J., & Liu, L. (1984). Role of Topoisomerase II in Mediating Epipodophyllotoxin-induce DNA Cleavage. *Cancer Research 44*(12), 5857-5860.

Röttig, A., & Steinbüchel, A. (2013). Acyltransferases in Bacteria. *Microbiology and Molecular Biology Reviews*, *77*(2), 277-321. https://doi.org/10.1128/MMBR.00010-13

Roumani, M., Besseau, S., Gagneul, D., Robin, C., & Larbat, R. (2021). Phenolamides in plants : An update on their function, regulation, and origin of their biosynthetic enzymes. *Journal of Experimental Botany*, *72*(7), 2334-2355. https://doi.org/10.1093/jxb/eraa582

Runguphan, W., Qu, X., & O'Connor, S. E. (2010). Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature*, *468*(7322), 461-464. https://doi.org/10.1038/nature09524

Ruprecht, C., Vaid, N., Proost, S., Persson, S., & Mutwil, M. (2017). Beyond Genomics : Studying

Evolution with Gene Coexpression Networks. *Trends in Plant Science*, *22*(4), 298-307. https://doi.org/10.1016/j.tplants.2016.12.011

Ryabinin, A. A., & Ilina, E. M. (1949). The alkaloid of *Salsola subaphylla*. *Doklady Akademiya Nauk 67*, 513.

S :

Sachs, M. M., Freeling, M., & Okimoto, R. (1980). The anaerobic proteins of maize. *Cell*, *20*(3), 761-767. https://doi.org/10.1016/0092-8674(80)90322-0

Sadowska, A., Wiweger, M., Lata, B., & Obidoska, G. (1997). *In vitro* propagation of *Podophyllum peltatum* L. by the cultures of embrya and divided embryo. *Biologia plantarum*, *39*(3), 331-336.

Sakaki, T., Kominami, S., Takemori, S., Ohkawa, H., Akiyoshi-Shibata, M., & Yabusaki, Y. (1994). Kinetic Studies on a Genetically Engineered Fused Enzyme between Rat Cytochrome P4501A1 and Yeast NADPH-P450 Reductase. *Biochemistry*, *33*(16), 4933-4939. https://doi.org/10.1021/bi00182a023

Salvachúa, D., Rydzak, T., Auwae, R., De Capite, A., Black, B. A., Bouvier, J. T., Cleveland, N. S., Elmore, J. R., Furches, A., Huenemann, J. D., Katahira, R., Michener, W. E., Peterson, D. J., Rohrer, H., Vardon, D. R., Beckham, G. T., & Guss, A. M. (2019). Metabolic engineering of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin. *Microbial Biotechnology*, *13*(1), 290-298. https://doi.org/10.1111/1751-7915.13481

Schenkman, J. B., & Jansson, I. (2003). The many roles of cytochrome. *Pharmacology and Therapeutics*, *97*(2), 139-152.

Schmidt, A., Scheel, D., & Strack, D. (1998). Elicitor-stimulated biosynthesis of hydroxycinnamoyltyramines in cell suspension cultures of *Solanum tuberosum*. *Planta*, *205*(1), 51-55.

Schuler, M. A., & Werck-Reichhart, D. (2003). Functional genomics of P450s. *Annual review of plant biology*, *54*, 629-667. https://doi.org/10.1146/annurev.arplant.54.031902.134840

Schwab, W., Davidovich-Rikanati, R., & Lewinsohn, E. (2008). Biosynthesis of plant-derived flavor compounds. *The Plant Journal*, *54*(4), 712-732. https://doi.org/10.1111/j.1365-313X.2008.03446.x

Schwartz, D. (1970). Genetic control of alcohol dehydrogenase – a competition model for regulation of gene action. *Genetics*, *67*(3), 411-425. https://doi.org/10.1093/genetics/67.3.411

Schwartz, S. H., Qin, X., & Zeevaart, JanA. D. (2001). Characterization of a Novel Carotenoid Cleavage Dioxygenase from Plants. *Journal of Biological Chemistry*, *276*(27), 25208-25211. https://doi.org/10.1074/jbc.M102146200

Schwarz, D. (1991). Rational motion and membrane topology of the microsomal cytochrome P450 systems analyzed by saturation EPR. *Frontiers in biotechnology*, *5*, 94-137.

Shafqat, J., El-Ahmad, M., Danielsson, O., Martínez, M. C., Persson, B., Parés, X., & Jornvall, H. (1996). Pea formaldehyde-active class III alcohol dehydrogenase : Common derivation of the plant and animal forms but not of the corresponding ethanol-active forms (classes I and P). *Proceedings of the National Academy of Sciences*, *93*(11), 5595-5599.

https://doi.org/10.1073/pnas.93.11.5595

Shah, Z., Gohar, U. F., Jamshed, I., Mushtaq, A., Mukhtar, H., Zia-UI-Haq, M., Toma, S. I., Manea, R., Moga, M., & Popovici, B. (2021). Podophyllotoxin : History, Recent Advances and Future Prospects. *Biomolecules*, *11*(4), 603. https://doi.org/10.3390/biom11040603

Shao, Z., Zhao, H., & Zhao, H. (2009). DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Research*, *37*(2), 10.

Sharp, P. M., Tuohy, T. M. F., & Mosurski, K. R. (1986). Codon usage in yeast : Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Research*, *14*(13), 5125-5143. https://doi.org/10.1093/nar/14.13.5125

Shevyakova, N. I., Rakitin, V. Yu., Stetsenko, L. A., Aronova, E. E., & Kuznetsov, V. V. (2006). Oxidative stress and fluctuations of free and conjugated polyamines in the halophyte *Mesembryanthemum crystallinum* L. under NaCl salinity. *Plant Growth Regulation*, *50*(1), 69-78. https://doi.org/10.1007/s10725-006-9127-1

Sibirny, A. A. (2016). Yeast peroxisomes : Structure, functions and biotechnological opportunities. *FEMS Yeast Research*, *16*(4), 1-14. https://doi.org/10.1093/femsyr/fow038

Sicilia, T., Niemeyer, H. B., Honig, D. M., & Metzler, M. (2003). Identification and Stereochemical Characterization of Lignans in Flaxseed and Pumpkin Seeds. *Journal of Agricultural and Food Chemistry*, *51*(5), 1181-1188. https://doi.org/10.1021/jf0207979

Siminszky, B., Gavilano, L., Bowen, S. W., & Dewey, R. E. (2005). Conversion of nicotine to nornicotine in *Nicotiana tabacum* is mediated by CYP82E4, a cytochrome P450 monooxygenase. *Proceedings of the National Academy of Sciences*, *102*(41), 14919-14924. https://doi.org/10.1073/pnas.0506581102

Sinkule, J. A. (1984). Etoposide : A Semisynthetic Epipodophyllotoxin Chemistry, Pharmacology, Pharmacokinetics, Adverse Effects and Use as an Antineoplastic Agent. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, *4*(2), 61-71. https://doi.org/10.1002/j.1875-9114.1984.tb03318.x

Sobolev, V. S., Sy, A. A., & Gloer, J. B. (2008). Spermidine and Flavonoid Conjugates from Peanut (*Arachis hypogaea*) Flowers. *Journal of Agricultural and Food Chemistry*, *56*(9), 2960-2969. https://doi.org/10.1021/jf703652a

Stähelin, H. F., & Von Wartburg, A. (1991). The Chemical and Biological Route from Podophyllotoxin Glucoside to Etoposide : Ninth Cain Memorial Award Lecture. *Cancer Research*, *51*(1), 5-15.

Stander, E. A., Dugé de Bernonville, T., Papon, N., & Courdavault, V. (2022). Chromosome-scale genomes throw light on plant drug biosynthesis. *Trends in Pharmacological Sciences*, *43*(7), 542-545. https://doi.org/10.1016/j.tips.2022.01.007

Stander, E. A., Sepúlveda, L. J., Dugé de Bernonville, T., Carqueijeiro, I., Koudounas, K., Lemos Cruz, P., Besseau, S., Lanoue, A., Papon, N., Giglioli-Guivarc'h, N., Dirks, R., O'Connor, S. E., Atehortùa, L., Oudin, A., & Courdavault, V. (2020). Identifying Genes Involved in Alkaloid Biosynthesis in *Vinca minor* through Transcriptomics and Gene Co-Expression Analysis. *Biomolecules*, *10*(12), 1595. https://doi.org/10.3390/biom10121595

Stavrinides, A., Tatsis, E. C., Caputi, L., Foureau, E., Stevenson, C. E. M., Lawson, D. M., Courdavault, V., & O'Connor, S. E. (2016). Structural investigation of heteroyohimbine alkaloid synthesis reveals active site elements that control stereoselectivity. *Nature Communications*, *7*(1), 12116.

https://doi.org/10.1038/ncomms12116

Steyer, D., Erny, C., Claudel, P., Riveill, G., Karst, F., & Legras, J.-L. (2013). Genetic analysis of geraniol metabolism during fermentation. *Food Microbiology*, *33*(2), 228-234. https://doi.org/10.1016/j.fm.2012.09.021

St-Pierre, B., & De Luca, V. (2000). Evolution of Acyltransferase Genes : Origin and Diversification of the BAHD Superfamily of Acyltransferases Involved in Secondary Metabolism. In *Recent Advances in Phytochemistry 34*, 285-315. Elsevier. https://doi.org/10.1016/S0079-9920(00)80010-6

St-Pierre, B., Laflamme, P., Alarco, A.-M., D, V., & De Luca, V. (1998). The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. *The Plant Journal*, *14*(6), 703-713. https://doi.org/10.1046/j.1365-313x.1998.00174.x

Strack, D., Eilert, U., Wray, V.; Wolff, J. & Jaggy H. (1990). Tricoumaroylspermidine in flowers of rosaceae. *Phytochemistry*, *29*, 2893−2896.

Strommer, J. (2011). The plant ADH gene family : The plant ADH gene family. *The Plant Journal*, *66*(1), 128-142. https://doi.org/10.1111/j.1365-313X.2010.04458.x

Sun, J., Shao, Z., Zhao, H., Nair, N., Wen, F., Xu, J.-H., & Zhao, H. (2012). Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in Saccharomyces cerevisiae. *Biotechnology and Bioengineering*, *109*(8), 2082-2092. https://doi.org/10.1002/bit.24481

Sung, Y.-C., Lin, C.-P., & Chen, J.-C. (2014). Optimization of virus-induced gene silencing in *Catharanthus roseus*. *Plant Pathology*, *63*(5), 1159-1167. https://doi.org/10.1111/ppa.12186

Syed, T. A., Cheema, K. M., Khayyami, M., Ahmad, S. A., Ahmad, S. H., & Ahmad, S. (1995). Human leukocyte interferon-a versus podophyllotoxin in cream for the treatment of genital warts in males. A placebo-controlled, double-blind, comparative study. *Dermatology*, *191*(2).

Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D., & Dumas, B. (2003). Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nature Biotechnology*, *21*(2), 143-149. https://doi.org/10.1038/nbt775

T :

Takeuchi, K., & Reue, K. (2009). Biochemistry, physiology, and genetics of GPAT, AGPAT, and lipin enzymes in triglyceride synthesis. *American Journal of Physiology-Endocrinology and Metabolism*, *296*(6), 1195-1209. https://doi.org/10.1152/ajpendo.90958.2008

Tatsis, E. C., Carqueijeiro, I., Dugé de Bernonville, T., Franke, J., Dang, T.-T. T., Oudin, A., Lanoue, A., Lafontaine, F., Stavrinides, A. K., Clastre, M., Courdavault, V., & O'Connor, S. E. (2017). A three enzyme system to generate the *Strychnos* alkaloid scaffold from a central biosynthetic intermediate. *Nature Communications*, *8*(1), 316. https://doi.org/10.1038/s41467-017-00154-x

Tebayashi, S., Horibata, Y., Mikagi, E., Kashiwagi, T., Mekuria, D. B., Dekebo, A., Ishihara, A., & Kim, C.-S. (2007). Induction of Resistance against the Leafminer, *Liriomyza trifolii* , by

Jasmonic Acid in Sweet Pepper. *Bioscience, Biotechnology, and Biochemistry*, *71*(6), 1521-1526. https://doi.org/10.1271/bbb.70033

Tempest, P. R., Bremner, P., Lambert, M., Taylor, G., Furze, J. M., Carr, F. J., & Harris, W. J. (1991). Reshaping a human monoclonal antibody to inhibit human respiratory syncytial virus infection in vivo. *Biotechnology*, *9*, 266-271. https://doi.org/10.1038/nbt0391-266

Tiburcio, A. F., Kaur-Sawhneyz, R., & Galston, A. W. (1993). Spermidine biosynthesis as affected by osmotic stress in oat leaves. *Plant Growth Regulation*, *13*(1). https://doi.org/10.1007/BF00207599

Timoneda, A., Sheehan, H., Feng, T., Lopez-Nieves, S., Maeda, H. A., & Brockington, S. F. (2018). Redirecting Primary Metabolism to Boost Production of Tyrosine- Derived Specialised Metabolites in Planta. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-33742-y

Trill, J. J., Shatzman, A. R., & Subinay, G. (1995). Production of monoclonal antibodies in COS and CHO cells. *Current Opinion in Biotechnology*, *6*(5), 553-560. https://doi.org/10.1016/0958-1669(95)80092-1

Tuller, T., Waldman, Y. Y., Kupiec, M., & Ruppin, E. (2010). Translation efficiency is determined by both codon bias and folding energy. *Proceedings of the National Academy of Sciences*, *107*(8), 3645-3650. https://doi.org/10.1073/pnas.0909910107

Tuominen, L. K., Johnson, V. E., & Tsai, C.-J. (2011). Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among *Populus paralogues*. *BMC Genomics*, *12*(1), 236. https://doi.org/10.1186/1471-2164-12-236

U :

Umezawa, T., Ragamustari, S. K., Nakatsubo, T., Wada, S., Li, L., Yamamura, M., Sakakibara, N., Hattori, T., Suzuki, S., & Chiang, V. L. (2013). A lignan *O*-methyltransferase catalyzing the regioselective methylation of matairesinol in Carthamus tinctorius. *Plant Biotechnology*, *30*(2), 97-109. https://doi.org/10.5511/plantbiotechnology.12.1230a

Umezawa, T. (2010). The cinnamate/monolignol pathway. *Phytochemistry Reviews*, *9*(1), 1-17. https://doi.org/10.1007/s11101-009-9155-3

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., & Pompon, D. (1997). Cloning, Yeast Expression, and Characterization of the Coupling of Two Distantly Related *Arabidopsis thaliana* NADPH-Cytochrome P450 Reductases with P450 CYP73A5. *Journal of Biological Chemistry*, *272*(31), 19176-19186. https://doi.org/10.1074/jbc.272.31.19176

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*, *11*(9), 636-646. https://doi.org/10.1038/nrg2842

V :

Van Dijken, J. P., Bauer, J., Brambilla, L., Duboc, P., Francois, J. M., Gancedo, C., Giuseppin, M. L. F., Heijnen, J. J., Hoare, M., Lange, H. C., Madden, E. A., Niederberger, P., Nielsen, J., Parrou, J. L., Petit, T., Porro, D., Reuss, M., van Riel, N., Rizzi, M., … Pronk, J. T. (2000). An

interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme and Microbial Technology*, *26*(9-10), 706-714. https://doi.org/10.1016/S0141-0229(00)00162-9

Van Eldik, G. J., Ruiter, R. K., van Herpen, M. M. A., Schrauwen, J. A. M., & Wullems, G. J. (1997). Induced ADH gene expression and enzyme activity in pollinated pistils of Solanum tuberosum. *Sexual Plant Reproduction*, *10*(2), 107-109. https://doi.org/10.1007/s004970050075

Van Uden, W., Woerdenbag, H. J., & Pras, N. (1994). Cyclodextrins as a useful tool for bioconversions in plant cell biotechnology. *Plant Cell, Tissue and Organ Culture*, *38*(2), 103-113.

Van Uden, W., Pras, N., & Malingré, T. M. (1990). On the improvement of the podophyllotoxin production by phenylpropanoid precursor feeding to cell cultures of Podophyllum hexandrum Royle. *Plant Cell, Tissue and Organ Culture*, *23*(3), 217-224. https://doi.org/10.1007/BF00034435

Van Uden, W., Pras, N., Visser, J. F., & Malingré, T. M. (1989). Detection and identification of podophyllotoxin produced by cell cultures derived from Podophyllum hexandrum Royle. *Plant Cell Reports*, *8*(3), 165-168.

Vasilev, N. P., Julsing, M. K., Koulman, A., Clarkson, C., Woerdenbag, H. J., Ionkova, I., Bos, R., Jaroszewski, J. W., Kayser, O., & Quax, W. J. (2006). Bioconversion of deoxypodophyllotoxin into epipodophyllotoxin in E. coli using human cytochrome P450 3A4. *Journal of Biotechnology*, *126*(3), 383-393. https://doi.org/10.1016/j.jbiotec.2006.04.025

Vazquez-Flota, F., De Carolis, E., Alarco, A.-M., & De Luca, V. (1997). Molecular cloning and characterization of desacetyloxyvindoline-4-hyroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus*. *Plant Molecular Biology*, *34*(6), 935-948. https://doi.org/10.1023/A:1005894001516

Verpoorte, R., Van der Heijden, R., Hoopen, H. J. C., & Memelink J. (1999). Metabolic engineering of plant secondary metabolite pathways for the production of fine chemicals. *Biotechnology Letters 21*, 467-479.

Verwaal, R., Buiting-Wiessenhaan, N., Dalhuijsen, S., & Roubos, J. A. (2018). CRISPR/Cpf1 enables fast and simple genome editing of *Saccharomyces cerevisiae* : CRISPR/Cpf1-mediated genome editing of *Saccharomyces cerevisiae*. *Yeast*, *35*(2), 201-211. https://doi.org/10.1002/yea.3278

Vidal. (2017, juin 6). *ETOPOSIDE TEVA 200 mg/10 mL injectable : Rupture de stock et mesures provisoires*. https://www.vidal.fr/actualites/21497/etoposide_teva_200_mg_10_ml_injectable_rupture_de_stock_et_mesures_provisoires/

Vigalok, A. (2015). Electrophilic Halogenation–Reductive Elimination Chemistry of Organopalladium and -Platinum Complexes. *Accounts of Chemical Research*, *48*(2), 238-247. https://doi.org/10.1021/ar500325x

Vogt, T. (2018). Unusual spermine-conjugated hydroxycinnamic acids on pollen : Function and evolutionary advantage. *Journal of Experimental Botany*, *69*(22), 5311-5315. https://doi.org/10.1093/jxb/ery359

Vogt, T. (2010). Phenylpropanoid Biosynthesis. *Molecular Plant*, *3*(1), 2-20. https://doi.org/10.1093/mp/ssp106

Von Heimendahl, C. B. I., Schäfer, K. M., Eklund, P., Sjöholm, R., Schmidt, T. J., & Fuss, E. (2005).

Pinoresinol–lariciresinol reductases with different stereospecificity from *Linum album* and *Linum usitatissimum*. *Phytochemistry*, *66*(11), 1254-1263. https://doi.org/10.1016/j.phytochem.2005.04.026

W :

Walters, D., Meurer-Grimes, B., & Rovira, I. (2001). Antifungal activity of three spermidine conjugates. *FEMS Microbiology Letters*, *201*(2), 255-258. https://doi.org/10.1111/j.1574-6968.2001.tb10765.x

Wang, C., Poborsky, M., Crocoll, C., Nødvig, C. S., Mortensen, U. H., & Halkier, B. A. (2022). Comparison of Genome and Plasmid-Based Engineering of Multigene Benzylglucosinolate Pathway in Saccharomyces cerevisiae. *Applied and Environmental Microbiology*, e00978-22. https://doi.org/10.1128/aem.00978-22

Wang, J. C. (1977). Interaction between DNA and *Escherichia coli* protein omega. Formation of a complex between single-stranded DNA and omega protein. *Journal of Biological Chemistry*, *253*(2), 511-518. https://doi.org/10.1016/S0021-9258(17)38239-X

Wang, L., Deng, A., Zhang, Y., Liu, S., Liang, Y., Bai, H., Cui, D., Qiu, Q., Shang, X., Yang, Z., He, X., & Wen, T. (2018). Efficient CRISPR–Cas9 mediated multiplex genome editing in yeasts. *Biotechnology for Biofuels*, *11*(1), 277. https://doi.org/10.1186/s13068-018-1271-0

Wang, L., & Hatzimanikatis, V. (2006). Metabolic engineering under uncertainty—II : Analysis of yeast metabolism. *Metabolic Engineering*, *8*(2), 142-159. https://doi.org/10.1016/j.ymben.2005.11.002

Wang, Y., Shi, Y., Li, K., Yang, D., Liu, N., Zhang, L., Zhao, L., Zhang, X., Liu, Y., Gao, L., Xia, T., & Wang, P. (2021). Roles of the 2-Oxoglutarate-Dependent Dioxygenase Superfamily in the Flavonoid Pathway : A Review of the Functional Diversity of F3H, FNS I, FLS, and LDOX/ANS. *Molecules*, *26*(21), 6745. https://doi.org/10.3390/molecules26216745

Wang, Y., Halls, C., Zhang, J., Matsuno, M., Zhang, Y., & Yu, O. (2011). Stepwise increase of resveratrol biosynthesis in yeast Saccharomyces cerevisiae by metabolic engineering. *Metabolic Engineering*, *13*(5), 455-463. https://doi.org/10.1016/j.ymben.2011.04.005

Ward, R. S. (1992). Synthesis of podophyllotoxin and related compounds. *Synthesis*, 719-730.

Wattelet, V. (2022). *Rôle des acides lysophosphatidiques acyltransférases dans la voie sécrétoire chez les plantes*. Mémoire de thèse de l'Université de Bordeaux.

Werck-Reichhart, D., Bak, S., & Paquette, S. (2002). Cytochromes P450. *The Arabidopsis Book*, *1*, e0028. https://doi.org/10.1199/tab.0028

Werck-Reichhart, D., & Feyereisen, R. (2000). Cytochromes P450 : A success story. *Genomic biology 1*(6).

Werner, C., Hu, W., Lorenzi-Riatsch, A., & Hesse, M. (1995). Di-coumaroylspermidines and tri-coumaroylspermidines in anthers of different species of the genus *Aphelandra*. *Phytochemistry*, *40*(2), 461-465. https://doi.org/10.1016/0031-9422(95)00288-I

Who, G. (2000). *Some Antiviral and Antineoplastic Drugs, and Other Pharmaceutical Agents*. *76*.

Wiese, S., Wubshet, S. G., Nielsen, J., & Staerk, D. (2013). Coupling HPLCSPE- NMR with a microplate-based high-resolution antioxidant assay for efficient analysis of antioxidants in food—Validation and proof-of-concept study with caper buds. *Food Chemistry*, *141*(4), 4010-4018.

Wink, M. (2010). Introduction : Biochemistry, Physiology and Ecological Functions of Secondary Metabolites. In M. Wink (Éd.), *Biochemistry of Plant Secondary Metabolism 40*,1-19. Wiley-Blackwell. https://doi.org/10.1002/9781444320503.ch1

Winzer, T., Gazda, V., He, Z., Kaminski, F., Kern, M., Larson, T. R., Li, Y., Meade, F., Teodor, R. I., Vaistij, F., Walker, C., Bowser, T. A., & Graham, I. A. (2012). A Papaver somniferum 10-Gene Cluster for Synthesis of the Anticancer Alkaloid Noscapine. *Science*, *336*(6089), 1704-1708. https://doi.org/10.1126/science.1220757

X :

Xia, Z.-Q., Costa, M. A., Pélissier, H. C., Davin, L. B., & Lewis, N. G. (2001). Secoisolariciresinol Dehydrogenase Purification, Cloning, and Functional Expression. *Journal of Biological Chemistry*, *276*(16), 12614-12623. https://doi.org/10.1074/jbc.M008622200

Xu, W.-H., Zhao, P., Wang, M., & Liang, Q. (2018). Naturally occurring furofuran lignans : Structural diversity and biological activities. *Natural Product Research*, *33*(9), 1357-1373. https://doi.org/10.1080/14786419.2018.1474467

Y :

Yamamoto, A., Nakamura, K., Furukawa, K., Konishi, Y., Ogino, T., Higashiura, K., Yago, H., Okamoto, K., & Otsuka, M. (2002). A New Nonpeptide Tachykinin NK1 Receptor Antagonist Isolated from the Plants of Compositae. *Chemical and Pharmaceutical Bulletin*, *50*(1), 47-52. https://doi.org/10.1248/cpb.50.47

Yang, Q., Reinhard, K., Schiltz, E., & Matern, U. (s. d.). Characterization and heterologous expression of hydroxycinnamoyl / benzoyl-CoA : anthranilate N-hydroxycinnamoyl /benzoyltransferase from elicited cell cultures of carnation, Dianthus caryophyllus L. *Plant Molecular Biology*, *35*, 777-789.

Yang, X., Gao, S., Guo, L., Wang, B., Jia, Y., Zhou, J., Che, Y., Jia, P., Lin, J., Xu, T., Sun, J., & Ye, K. (2021). Three chromosome-scale Papaver genomes reveal punctuated patchwork evolution of the morphinan and noscapine biosynthesis pathway. *Nature Communications*, *12*(1), 6030. https://doi.org/10.1038/s41467-021-26330-8

Yingyongnarongkul, B., Apiratikul, N., Aroonrerk, N., & Suksamrarn, A. (2008). Synthesis of bis, tris and tetra(dihydrocaffeoyl)polyamine conjugates as antibacterial agents against VRSA. *Archives of Pharmacal Research*, *31*(6), 698-704. https://doi.org/10.1007/s12272-001-1215-4

Yoshihara, T., Yamaguchi, K., Takamatsu, S., & Sakamura, S. (1981). A new lignan amide, grossamide, from bell pepper (*Capsicum annuum* var. *grossurri*). *Agricultural and Biological Chemistry*, *45*(11), 2593-2598.

Youhnovski, N., Werner, C., & Hesse, M. (2001). N,N',N''-triferuloylspermidine, a new UV absorbing polyamine derivative from pollen of *Hippeastrum x hortorum*. *Zeitschrift für Naturforschung -*

Section C Journal of Biosciences, *56*(7), 526-530.

Yu, X., Che, Z., & Xu, H. (2016). Recent Advances in the Chemistry and Biology of Podophyllotoxins. *Chemistry - A European Journal*, *23*(19), 4467-4526. https://doi.org/10.1002/chem.201602472

Z :

Zacarés, L., López-Gresa, M. P., Fayos, J., Primo, J., Bellés, J. M., & Conejero, V. (2007). Induction of *p* -Coumaroyldopamine and Feruloyldopamine, Two Novel Metabolites, in Tomato by the Bacterial Pathogen *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions*®, *20*(11), 1439-1448. https://doi.org/10.1094/MPMI-20-11-1439

Zhang, A., Mernitz, K., Wu, C., Xiong, W., He, Y., Wang, G., & Wang, X. (2021). ATP Drives Efficient Terpene Biosynthesis in Marine Thraustochytrids. *MBio*, *12*(3), e00881-21. https://doi.org/10.1128/mBio.00881-21

Zhang, J., Hansen, L. G., Gudich, O., Viehrig, K., Lassen, L. M. M., Schrübbers, L., Adhikari, K. B., Rubaszka, P., Carrasquer-Alvarez, E., Chen, L., D'Ambrosio, V., Lehka, B., Haidar, A. K., Nallapareddy, S., Giannakou, K., Laloux, M., Arsovska, D., Jørgensen, M. A. K., Chan, L. J. G., Kristensen, M., Christensen, H. B., Sudarsan, S., Stander, E. A., Baidoo, E., Petzold, C. J., Wulff, T., O'Connor, S. E., Courdavault, V., Jensen, M. K., & Keasling, J. D. (2022). A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature*, *609*(7926), 341-347. https://doi.org/10.1038/s41586-022-05157-3

Zhang, J., Liu, Y.-Q., Yang, L., & Feng, G. (2010). Podophyllotoxin Derivatives Show Activity against *Brontispa Longissima* Larvae. *Natural Product Communications*, *5*(8), 1934578X1000500. https://doi.org/10.1177/1934578X1000500820

Zhang, S., Guo, F., Yan, W., Dai, Z., Dong, W., Zhou, J., Zhang, W., Xin, F., & Jiang, M. (2020). Recent Advances of CRISPR/Cas9-Based Genetic Engineering and Transcriptional Regulation in Industrial Biology. *Frontiers in Bioengineering and Biotechnology*, *7*, 459. https://doi.org/10.3389/fbioe.2019.00459

Zhao, G., Qin, G.-W., Gai, Y., & Guo, L.-H. (2010). Structural Identification of a New Tri-p-coumaroylspermidine with Serotonin Transporter Inhibition from Safflower. *Chemical and Pharmaceutical Bulletin*, *58*(7), 950-952. https://doi.org/10.1248/cpb.58.950

Zhao, W., Cong, Y., Li, H.-M., Li, S., Shen, Y., Qi, Q., Zhang, Y., Li, Y.-Z., & Tang, Y.-J. (2021). Challenges and potential for improving the druggability of podophyllotoxin-derived drugs in cancer chemotherapy. *Natural Product Reports*, *38*(3), 470-488. https://doi.org/10.1039/D0NP00041H

Zheng, M., Zhang, L., Tang, M., Liu, J., Liu, H., Yang, H., Fan, S., Terzaghi, W., Wang, H., & Hua, W. (2020). Knockout of two *Bna MAX 1* homologs by CRISPR /Cas9-targeted mutagenesis improves plant architecture and increases yield in rapeseed (*Brassica napus* L.). *Plant Biotechnology Journal*, *18*(3), 644-654. https://doi.org/10.1111/pbi.13228

Zhu, M., Wang, C., Sun, W., Zhou, A., Wang, Y., Zhang, G., Zhou, Xi., Huo, Y., & Li, C. (2018). Boosting 11-oxo-β-amyrin and glycyrrhetinic acid synthesis in Saccharomyces cerevisiae via pairing novel oxidation and reduction system from legume plants. *Metabolic Engineering*, *45*, 43-50. https://doi.org/10.1016/j.ymben.2017.11.009

ANNEXES

Au cours de ma thèse, j'ai eu l'opportunité de rédiger ou de participer à la rédaction d'articles de commentaire (Spotlight ou Highlight) destinés à mettre en lumière des résultats de très forts intérêts publiés par différents groupes de recherche dans le domaine de l'ingénierie métabolique. Deux de ces articles, un dont je suis le premier auteur et un second dont je suis le deuxième auteur, seront présentés ci-après.

En outre, j'ai également participé à des travaux de l'équipe menés par des doctorants et post-doctorants du laboratoire sur des thématiques éloignées de la mienne. J'ai ainsi souhaité contribuer à la réalisation de ces études pour étendre mes compétences techniques et mes connaissances scientifiques de façon plus générale. Il en résulte deux articles dont je suis également co-auteur.

Annexe I: Commentaire " Efficient Terpene Production by Marine Thraustochytrids : Shedding Light on the Thermodynamic Driving Force "

par Kulagina N, Perrin J, Besseau S, Courdavault V. publié en 2021 dans la revue **mBio**.

Il s'agit du premier article de type commentaire auquel j'ai participé et qui commente les travaux de Zhang et al., 2021 (https://doi.org/10.1128/mBio.00881-21) publiés dans la revue mBio. Dans cette étude, Zhang et son équipe ont rapporté la potentielle utilisation de la thermodynamique métabolique des thraustochytrides marins, eucaryotes unicellulaires hétérotrophes, pour l'optimisation de la synthèse des tepènes en levure. Ce commentaire discute des effets métaboliques induits par le sodium chez les thraustochytrides marins permettant de doubler la synthèse de squalène dans ces organismes, permettant d'avancer dans la compréhension de la dynamique du métabolisme impliqué dans la synthèse des terpènes. Ces résultats mettent, en effet, en lumière l'importance de la compréhension de la thermodynamique du métabolisme pour l'ingénierie métabolique et l'optimisation des rendements de biosynthèse.

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

COMMENTARY

Efficient Terpene Production by Marine Thraustochytrids: Shedding Light on the Thermodynamic Driving Force

Natalja Kulagina,[a] Jennifer Perrin,[a] Sébastien Besseau,[a] Vincent Courdavault[a]

[a]Université de Tours, EA2106 "Biomolécules et Biotechnologies Végétales," Tours, France

ABSTRACT Terpenoids, such as squalene, are valuable compounds for cosmetic and drug industries, the supply of which is often limited by natural sources. Alternative production strategies have been investigated for decades but remain challenging due to low yields. In a recent study, Zhang and coworkers (A. Zhang, K. Mernitz, C. Wu, W. Xiong, et al., mBio 12:e0088121, 2021, https://doi.org/10.1128/mBio.00881-21) report the potential use of marine thraustochytrid metabolic thermodynamics in effective terpene engineering. Through comparative proteomics and metabolomics, as well as thermodynamic modeling, the authors demonstrated sodium-induced changes in thraustochytrid metabolism leading to a twofold increase in squalene accumulation. The differential abundances of the metabolic enzymes and metabolites, as well as higher respiration, indicated the metabolic shift from carbohydrate to lipid oxidation and increased ATP input to the mevalonate pathway and squalene synthesis. This breakthrough provides new important insights into microbial terpene metabolic engineering but above all displays thermodynamics as a valuable tool in metabolic engineering.

KEYWORDS thermodynamics, metabolic flux, squalene, terpene engineering, *Aurantiochytrium* sp., thermodynamics

Terpenoids represent a wide range of primary and secondary metabolites produced by plants, animals, and microbes. In some of these organisms, the mevalonate (MVA) pathway supplies isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are further condensed to form a variety of terpene compounds, such as mono-, sesqui-, di-, tri-, and tertraterpenoids (1). Many of these molecules possess beneficial properties employed in various sectors of human activity, including the cosmetic and drug industries. For instance, the triterpenoid squalene is a well-known compound displaying a high potential for therapeutic applications, either used as an emulsion-like carrier for improved vaccine and drug delivery or due to its chemopreventive and immunostimulating effects (2). Currently, squalene is mainly extracted from shark liver and some vegetal sources (3), and its growing demand requires alternative supply strategies, such as microbe metabolic engineering. However, obtaining high production yields via this approach remains challenging due to limiting endogenous metabolic fluxes, competing pathways, and underperforming enzymes. In addition to the common optimization of metabolic fluxes and enzyme activity via gene expression fine-tuning, thermodynamics and understanding of carbon flux regulation between native primary and secondary metabolism were recently demonstrated as a valuable tool for efficient terpene engineering (4).

In this recent article (4), the authors exploited the marine protist thraustochytrids (here *Aurantiochytrium* sp.), known to accumulate high levels of squalene and carotenoids in their natural sodium-rich environment (5, 6). In particular, this study reported a twofold increase in squalene accumulation, reaching 123.6 mg/liter under sodium (NaCl) supplementation growth conditions (4). To determine the mechanism behind

Citation Kulagina N, Perrin J, Besseau S, Courdavault V. 2021. Efficient terpene production by marine thraustochytrids: shedding light on the thermodynamic driving force. mBio 12:e01976-21. https://doi.org/10.1128/mBio.01976-21.

Copyright © 2021 Kulagina et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Vincent Courdavault, vincent.courdavault@univ-tours.fr.

For the article discussed, see. https://doi.org/10.1128/mBio.00881-21.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

Published 28 September 2021

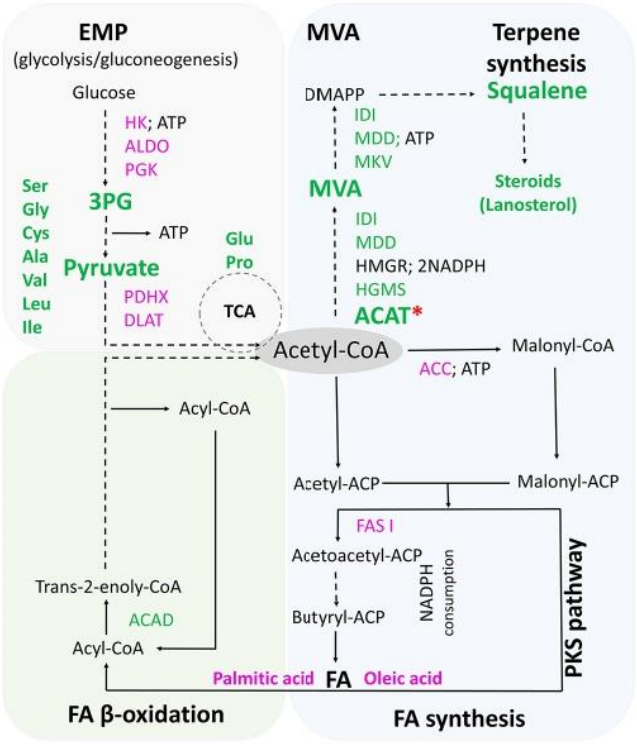


FIG 1 Sodium-induced changes in *Thraustochytrium* sp. metabolism (adapted from reference 4). The schematic pathways illustrate terpene synthesis from glucose or fatty acid (FA) β-oxidation, and the changes that occur in *Thraustochytrium* sp. metabolism under sodium (NaCl) supplementation growth conditions. Significantly increased levels of metabolites and enzymes are shown in green, while significantly decreased levels of metabolites or enzymes, determined by comparative proteomics and metabolomics, are shown in pink. For clarity, numerous steps were omitted (dashed arrows) to highlight the identified changes. The red asterisk indicates the rate-limiting enzyme in both growth conditions, determined by *in silico* analysis of local thermodynamics. EMP, Embden-Meyerhof-Parnas pathway; HK, hexokinase; ALDO, fructose-biphosphate aldolase; PGK, phosphoglycerate kinase; ACC, acetyl-CoA carboxylase; FAS I, fatty acid synthase I; ACAD, acyl-CoA dehydrogenase; MVA, mevalonate pathway; ACAT, acetyl-CoA acetyltransferase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; MVK, mevalonate kinase; MDD, mevalonate diphosphate decarboxylase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; IDI, 3-ketoacyl-CoA thiolase; 3PG, 3-phospho-D-glyceroyl phosphate; MVA, mevalonate; DMAPP, dimethylallyl diphosphate; PDHX, pyruvate dehydrogenase protein X component; DLAT, dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex; TCA, tricarboxylic acid cycle; PKS, polyketide synthase pathway.

the sodium-driven squalene accumulation, comparative proteomics were performed and highlighted differential abundances of metabolism-related enzymes between NaCl-supplemented and control growth conditions (Fig. 1). For instance, the main enzymes of the Embden-Meyerhof-Parnas (EMP) pathway (e.g., hexokinase [HK], fructose-biphosphate aldolase [ALDO], and phosphoglycerate kinase [PGK]) and the fatty acid (FA) synthesis pathway (acetyl coenzyme A [acetyl-CoA] carboxylase [ACC] and fatty acid synthase I [FAS I]) had significantly lower abundance under the high sodium growth condition, while the key enzyme of the β-oxidation pathway acyl-CoA dehydrogenase (ACAD) abundance increased 2.2-fold. These changes thus suggest a sodium-induced shift from carbohydrate to lipid oxidation, a lower fatty acid synthesis, and higher ATP/acetyl-CoA formation. On the other hand, the MVA pathway enzymes (e.g., acetyl-CoA acetyltransferase [ACAT], 3-hydroxy-3-methylglutaryl-CoA synthase [HMGS], mevalonate kinase [MVK], and mevalonate diphosphate decarboxylase [MDD]) showed a significantly increased abundance, demonstrating *per se* an enhanced metabolic flux toward terpene biosynthesis. Consistent with the proteomics analysis, the comparative metabolomic study displayed a 1.6- and 1.9-fold increase in mevalonate and squalene levels, respectively. In addition, despite a lower abundance of EMP pathway key enzymes, higher accumulations of glyceraldehyde

3-phosphate (3PG) and pyruvate (as well as of several downstream amino acids) were observed (Fig. 1) suggesting their synthesis via gluconeogenesis. Moreover, although the total fatty acid content did not significantly differ between the two growth conditions, several fatty acids, such as the most abundant palmitic acid and less abundant oleic acid decreased in the presence of supplemented sodium by 1.5- and 3.2-fold, respectively. These results support the rationale that in the presence of sodium, the cellular energy is preferentially generated from lipid degradation rather than glycolysis, which probably leads to a higher ATP/acetyl-CoA supply, thus favoring terpene biosynthesis. To strengthen their findings and identify the potential bottleneck(s) in squalene biosynthesis from glucose or short-chain fatty acid hexanoyl-CoA, the authors performed thermodynamic modeling and *in silico* analysis using the pathway analysis tool PathParser (7), which can predict the thermodynamically restrained metabolic step(s). Consequently, the thermodynamically highly unfavorable ACAT-mediated reaction, which requires acetyl-CoA as a substrate, was identified *in silico* as a rate-limiting step in both standard and high sodium growth conditions. In contrast, the measurement of acetyl-CoA levels at different time points between the two growth modes did not reveal any differences. Surprisingly, ATP levels were also found to be similar in the presence or absence of NaCl, suggesting an energy homeostasis-based regulation. Thus, cellular respiration, evaluated via the measurement of oxygen consumption, was significantly increased after 24 h in the sodium-supplemented cultures and pointed out ATP as a thermodynamic driving force for higher terpene production. In sum, the tremendous advances reported by Zhang and colleagues demonstrate that *Aurantiochytrium* sp. employs a sodium-induced mechanism to rewire its metabolism and adapt to the high sodium growth conditions. Such a process could be efficiently exploited as an additional tool to provide valuable and more specific insights for microbe metabolic engineering.

In this respect, the identified bottlenecks in heterologous production are routinely addressed via gene copy balance, the use of more active orthologs or engineered enzymes. In parallel, primary metabolism can be rewired to provide a higher cofactor supply and reduce the competition for common substrates (8). For instance, in yeast terpene biosynthesis, 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), which requires two NADPHs, is generally regarded as a limiting step instead of ACAT. However, for both enzymes as well as for HMGS, the expression of more active orthologs from *Enterococcus faecalis* was already shown to result in enhanced squalene titer (9). In addition to MVA pathway overexpression, yeast terpene engineering starts to make use of spatial reconfiguration of biosynthetic pathways, notably via enzyme peroxisome targeting, given the peroxisome intrinsic pool of fatty acid β-oxidation-derived acetyl-CoA (9, 10). For instance, after targeting the entire squalene biosynthetic pathway to *Saccharomyces cerevisiae* peroxisomes via the peroxisomal targeting signal PTS1, the squalene titer was improved 68-fold (650 mg/liter) (10), while another study demonstrated an improved biosynthesis of heterologous monoterpenoids via artificial peroxisomal localization of the MVA pathway and downstream heterologous enzymes in *S. cerevisiae* (9). Furthermore, the overexpression of several native genes involved in ATP import into the peroxisome, NADPH generation, and acetyl-CoA synthesis led to an additional 1.65-fold increase in squalene synthesis up to 1.3g/liter (10), which is coherent with the thraustochytrid thermodynamics. These strategies could additionally benefit from the thermodynamics-based data to reinforce rational pathway design. Indeed, implementing an identified thermodynamic driving force to direct the whole metabolism toward the desired compound production would undoubtedly lead to enhanced titers. Thus, in addition to the classical metabolic engineering strategies, the metabolic shift and local thermodynamic constraints identified in marine thraustochytrids have the potential to beneficially assist and refine rational terpene engineering. For instance, the enhancement of yeast FA β-oxidation and acetyl-CoA formation, which takes place in the peroxisome, could be assessed via the overexpression of corresponding enzymes to increase FA uptake from the FA-supplemented growth media (e.g., yeast protein kinase Ypk1p), activation (e.g., acyl-CoA synthases Fat1p and Faa1p/Faa2p/Faa3p/Faa4p) and degradation (e.g., acyl-CoA oxidase Fox1p) (11). In

parallel, yeast respiratory metabolism could be favored to limit fermentation and increase ATP supply via altering glucose import (e.g., hexose transporter mutagenesis) (12). Globally, advances in thermodynamics data, in line with predictive engineering, systems metabolic engineering, and rational pathway design, presage a facilitated and more efficient native/heterologous pathway construction and optimization, and high-scale production.

ACKNOWLEDGMENTS

We acknowledge funding from the ARD2020 Biopharmaceutical program of the Région Centre Val de Loire (BioPROPHARM, CatharSIS and ETOPOCentre projects), EU Horizon 2020 research and innovation program (MIAMi project-grant agreement 814645), and ANR MIACYC (ANR-20-CE43-0010).

REFERENCES

1. Wang X, Ort DR, Yuan JS. 2015. Photosynthetic terpene hydrocarbon production for fuels and chemicals. Plant Biotechnol J 13:137–146. https://doi.org/10.1111/pbi.12343.
2. Kim S-K, Karadeniz F. 2012. Biological importance and applications of squalene and squalane. Adv Food Nutr Res 65:223–233. https://doi.org/10.1016/B978-0-12-416003-3.00014-7.
3. Lozano-Grande MA, Gorinstein S, Espitia-Rangel E, Dávila-Ortiz G, Martínez-Ayala AL. 2018. Plant sources, extraction methods, and uses of squalene. Int J Agron 2018:1829160. https://doi.org/10.1155/2018/1829160.
4. Zhang A, Mernitz K, Wu C, Xiong W, He Y, Wang G, Wang X. 2021. ATP drives efficient terpene biosynthesis in marine thraustochytrids. mBio 12:e0088121. https://doi.org/10.1128/mBio.00881-21.
5. Kaya K, Nakazawa A, Matsuura H, Honda D, Inouye I, Watanabe MM. 2011. Thraustochytrid Aurantiochytrium sp. 18W-13a accumulates high amounts of squalene. Biosci Biotechnol Biochem 75:2246–2248. https://doi.org/10.1271/bbb.110430.
6. Aasen IM, Ertesvåg H, Heggeset TMB, Liu B, Brautaset T, Vadstein O, Ellingsen TE. 2016. Thraustochytrids as production organisms for docosahexaenoic acid (DHA), squalene, and carotenoids. Appl Microbiol Biotechnol 100:4309–4321. https://doi.org/10.1007/s00253-016-7498-4.
7. Wu C, Jiang H, Kalra I, Wang X, Cano M, Maness PC, Yu J, Xiong W. 2020. A generalized computational framework to streamline thermodynamics and kinetics analysis of metabolic pathways. Metab Eng 57:140–150. https://doi.org/10.1016/j.ymben.2019.08.006.
8. Guirimand G, Kulagina N, Papon N, Hasunuma T, Courdavault V. 2021. Innovative tools and strategies for optimizing yeast cell factories. Trends Biotechnol 39:488–504. https://doi.org/10.1016/j.tibtech.2020.08.010.
9. Dusséaux S, Wajn WT, Liu Y, Ignea C, Kampranis SC. 2020. Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. Proc Natl Acad Sci U S A 117:31789–31799. https://doi.org/10.1073/pnas.2013968117.
10. Liu GS, Li T, Zhou W, Jiang M, Tao XY, Liu M, Zhao M, Ren YH, Gao B, Wang FQ, Wei DZ. 2020. The yeast peroxisome: a dynamic storage depot and subcellular factory for squalene overproduction. Metab Eng 57:151–161. https://doi.org/10.1016/j.ymben.2019.11.001.
11. Klug L, Daum G. 2014. Yeast lipid metabolism at a glance. FEMS Yeast Res 14:369–388. https://doi.org/10.1111/1567-1364.12141.
12. Otterstedt K, Larsson C, Bill RM, Ståhlberg A, Boles E, Hohmann S, Gustafsson L. 2004. Switching the mode of metabolism in the yeast Saccharomyces cerevisiae. EMBO Rep 5:532–537. https://doi.org/10.1038/sj.embor.7400132.

Annexe II: Commentaire " Boosting lignan-precursor synthesis in yeast cell factories through co-factor supply optimization "

par Perrin J, Besseau S, Papon N, Courdavault V. publié en 2022 dans la revue **Frontiers in Bioengineering and Biotechnology.**

Ce second article de commentaire, pour lequel j'ai pris une place centrale dans la rédaction commente le manuscrit, publié dans la revue Nature Chem Biol par Chen et al., 2022 (https://doi.org/10.1038/s41589-022-01014-6). En effet, les travaux menés R. Chen et ses collaborateurs permettent aux souches de levures développées de produire et de recycler de façon remarquable les cofacteurs requis par les enzymes hétérologues surexprimées. Comme décrit précédemment dans ce manuscrit de thèse, l'optimisation des rendements de biosynthèse de MS d'intérêt en levure nécessite de recycler les cofacteurs nécessaires aux enzymes hétérologues, mais aussi de réécrire le métabolisme primaire de la levure. Dans ce commentaire, nous avons ainsi décrit et discuté les stratégies employées par Chen et son équipe afin de synthétiser et recycler les cofacteurs NADPH, FADH2 et S-adénosyl-L-méthionine nécessaires à la synthèse de MS, et en particulier des lignanes.

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Opinion
PUBLISHED 02 December 2022
DOI 10.3389/fbioe.2022.1079801

Check for updates

OPEN ACCESS

EDITED BY
Farshad Darvishi,
Alzahra University, Iran

REVIEWED BY
Yongjun Wei,
Zhengzhou University, China
Arman Beyraghdar Kashkooli,
Tarbiat Modares University, Iran

*CORRESPONDENCE
Vincent Courdavault,
vincent.courdavault@univ-tours.fr

SPECIALTY SECTION
This article was submitted to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and Biotechnology

RECEIVED 25 October 2022
ACCEPTED 21 November 2022
PUBLISHED 02 December 2022

CITATION
Perrin J, Besseau S, Papon N and Courdavault V (2022), Boosting lignan-precursor synthesis in yeast cell factories through co-factor supply optimization.
Front. Bioeng. Biotechnol. 10:1079801.
doi: 10.3389/fbioe.2022.1079801

COPYRIGHT
© 2022 Perrin, Besseau, Papon and Courdavault. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Boosting lignan-precursor synthesis in yeast cell factories through co-factor supply optimization

Jennifer Perrin[1], Sébastien Besseau[1], Nicolas Papon[2] and Vincent Courdavault[1]*

[1]Biomolécules et Biotechnologies Végétales, Université de Tours, Tours, France, [2]University Angers, University Brest, IRF, SFR ICAT, Angers, France

KEYWORDS

cofactor, yeast cell factories, metabolic engineering, lignan, NADPH

Introduction

Plants constitute one of the largest sources of natural products (NPs) whose biological activities have been exploited by humans for agronomic, cosmetic or pharmacological purposes. For instance, many of these NPs became essential components of the therapeutic arsenal deployed against cancers since many years, such as the lignan-derived podophyllotoxin produced by the mayapple *Podophyllum hexandrum* (Buss et al., 1995). However, NPs are usually accumulated in minute amounts in plants and their complex structures hinder total chemical synthesis at least at an industrial scale. While extraction and purification procedures have been improved and combined to semi-synthetic approaches, many plant drugs still exhibit high production costs and a supply mostly based on the exploitation of the natural resources. This engenders a huge pressure on the NP producing plants, many of which suffer from overexploitation, causing a recurring shortage of drugs currently used in the clinic. In order to secure NP supply, new strategies of production have emerged such as those relying on the transfer of genes in heterologous organisms and the metabolic engineering (Carqueijeiro et al., 2020).

Based on the pioneering production of the semi-synthetic production of the antimalarial artemisinin (Paddon and Keasling, 2014), metabolic engineering approaches relying on the creation of microbial cell factories have been continuously developed over the last 10 years. Through the transfer of plant genes into diverse heterologous organisms such as yeast, successful productions of valuables NPs have been achieved including resveratrol, a grape wine polyphenol displaying antioxidant properties (Li et al., 2015), ginsenosides Rh2 and R3 with promising properties against cancers (Wang et al., 2015), cannabinoids Δ9-tetrahydrocannabinol (THC) and cannabidiol (CBD) from hemp known for their beneficial psychoactive effects on human health (Gülck et al., 2020), and monoterpene indole alkaloids (MIAs) from the Madagascar periwinkle whose prominent members vincristine and vinblastine exhibit anticancer activities (Brown et al., 2015). While most of these NPs results from *de novo* synthesis, bioconversion approaches, relying on the feeding of the cell factories with an easily accessible precursor, has been also explored with success to produce vindoline used

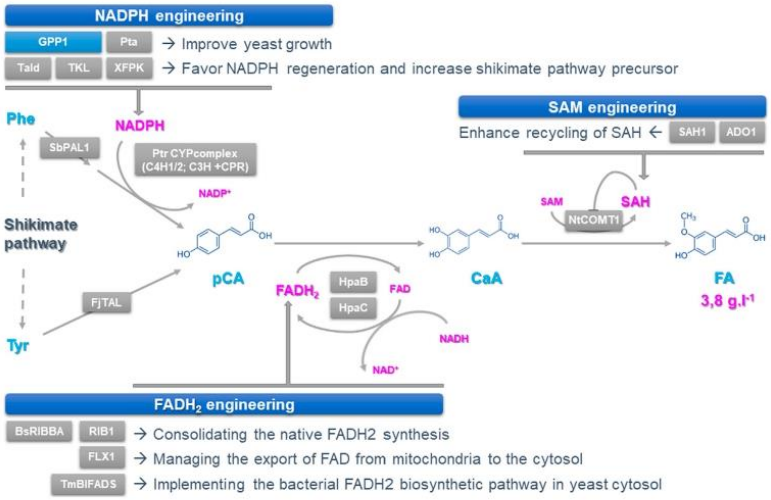


FIGURE 1
Improving the synthesis and regeneration of cofactors to increase the synthesis of podophyllotoxin precursors in yeast. Podophyllotoxin biosynthesis can be achieved through a three-step route involving deamination of tyrosine and phenylalanine to form p-coumaric acid (pCA), which is then hydroxylated and *O*-methylated to from caffeic acid (CaA) and ferulic acid (FA), respectively. Each step catalyzed by heterologous enzymes requires cofactors including NADPH, FAD (H_2) and SAM whose supplies have been engineered by optimizing their synthesis and/or recycling. Overexpressed and deleted genes are indicated with grey and blue boxes respectively. ADO1: Adenosine kinase; BsRIBBA: *Bacillus subtilis* bifunctional DHBP synthase/GTP cyclohydrolase II; CaA: Caffeic Acid; CPR: Cytochrome P450 Reductase; FA: Ferulic Acid; FAD (H2): Flavine Adenine Dinucleotide; FjTAL: *Flavobacterium* johnsoniae tyrosine ammonia lyase; FLX1: mitochondrial FAD exporter; GPP1: GAP Phosphatase; HpaB: FAD-dependent 4-hydroxyphenylacetate-3-monooxygenase; HpaC: NADH-flavin oxidoreductase; NADP(H): Nicotinamide Adenine Dinucleotide Phosphate; NtCOMT1: *Nicotiana tabacum* CaA O-methyltransferase 1; pCA: p-Coumaric Acid; Phe: Phenylalanine; Pta: phosphotransacetylase; PtrC3H: *Populus trichocarpa* coumarate-3-hydroxylase; PtrC4H1/2: *P. trichocarpa* cinnamic acid hydroxylase ½; RIB1: GTP cyclohydrolase II; SAH: S-adenosy-L-homocysteine; SAH1: SAH hydrolase; SAM: S-adenosyl-l-methionin; SbPAL1: *Isatis indigotica* phenylalanine ammonia lyase 1; Tald: transaldolase; TKL: transketolase; TmBIFADS: *Thermotoga maritima* bifunctional riboflavin kinase/FAD synthase; Tyr: Tyrosine; XFPK: phosphoketolase.

for vincristine synthesis (Kulagina et al., 2021a). One of the main challenges of metabolic engineering is obtaining high production titers compatible with industrial productions. While a production of tetrahydroisoquinoline alkaloids reaching up to 4.6 g^{-1} has been reported in yeast (Pyne et al., 2020), many NP syntheses by engineered yeast still require dramatic increases in production yields to replace or supplement plant-based extractions. Many of the previously reported improvements involves modifications of the molecular scaffold of the yeast as well as the transferred plant genes (Guirimand et al., 2021). In an elegant study (Chen et al., 2022), the Zhou's group recently describes a distinct approach relying on the efficient recycling of enzyme co-factors in yeast to fulfill the synthesis of caffeic acid (CaA) and ferulic acid (FA), two precursors of podophyllotoxin (Figure 1). Through this commentary article, we highlight the outstanding work of this research group and discuss how the cofactor recycling approach will undoubtedly pave the way for future high titer production of valuable NPs in yeast.

State of the art of strategies of natural product bioproduction improvement

While the initial production of hydrocortisone in yeast reached the outstanding yield of 25 g L^{-1} (Szczebara et al., 2003), most of the plant NPs resulting from complex biosynthetic routes were only produced at a µg or mg per liter titer, thus highlighting the need of improving biosynthesis in the engineered yeasts. In recent years, several works highlighted new strategies for improving these bioproduction yields. Obviously, controlling gene expression level is one of the evident levers to operate by selecting the best yeast genomic environment such as hot-spot (Mikkelsen et al., 2012), by using strong promoters and adapted terminators (Wang et al., 2019; Liu R. et al., 2020) or by improving the translation efficiency of the heterologous genes through expression of yeast-optimized codon sequences (Lanza et al., 2014). Handling the copy ratio of the transferred gene also constitutes a powerful option as recently reported for vindoline production. By increasing the copy number of specific genes

along the pathway, production bottlenecks have been alleviated as well as the synthesis of by-product resulting from enzyme promiscuity (Kulagina et al., 2021a). In addition, the identification of more efficient enzyme isoforms or engineered variant has also provided great improvements as reported for benzylisoquinoline alkaloids (Payne et al., 2021). In parallel, adapting yeast primary metabolism is also a major prerequisite as reported with the expression of truncated HMGCoA reductase gene, which encodes a non–feedback regulated rate-limiting enzyme of the mevalonate pathway for terpenoid production (Ro et al., 2006; Brown et al., 2015) or by knocking out genes encoding yeast enzymes from the Ehrlich pathway which divert tetrahydroisoquinoline alkaloid precursor from the synthesis of the final expected compound (Payne et al., 2021). Finally, spatial reconfiguration/adaptation of parts of biosynthetic pathways can be achieved as reported for transporters of tropane alkaloids (Srinivasan and Smolke, 2020) or the transfer of the whole mevalonate pathway in peroxisome for monoterpene production. Indeed, sequestration of enzymes in the peroxisomes can also increase the proximity between enzymes and their substrates in this specific cell compartment (Liu G.S. et al., 2020; Kulagina et al., 2021b) or limit cytotoxicity that some overexpressed enzyme may display (Grewal et al., 2021). Based on a similar reasoning, improving extracellular transports can limit the poisonous effects of heterologous produced metabolites and thus act as a detoxification mechanism. This is of importance when metabolites can not diffuse through membranes or when endogenous yeast transporters are unable to proceed thus requiring additional transporter overexpression. For instance, overexpressing the malic transporter Mae1 from *Schizosaccharomyces pombe* results in an increase of dicarboxylic acids production in *S. cerevisiae* (Darbani et al., 2019). Finally, playing with organelles size can also be helpful as reported for peroxisome (Kulagina et al., 2021b) and endoplasmic reticulum whose size can be increased through INO2, a regulatory gene in yeast phospholipid biosynthesis, overexpression (Kim et al., 2019) or PAH1 mutation (Arendt et al., 2017), resulting in an optimized triterpene biosynthesis.

Boosting the cofactor supply

Almost all biosynthetic pathways of plant natural products encompass enzymes needing co-factors such as the cytochromes P450 and alcohol dehydrogenases requiring nicotinamide adenine dinucleotide phosphate (NADPH), *N*- and *O*-methyltransferases requiring *S*-adenosyl-L-methionine (SAM) as well as acyltransferases requiring acetyl-Coenzyme A for instance. All these cofactors thus constitute limiting elements especially when numerous heterologous enzymes are overexpressed in cell factories. So far, strategies aiming at optimizing co-factor supply have already been described such as for producing squalene, an ergosterol pathway intermediate of interest in pharmaceutical and cosmetic industries (Paramasivan and Mutturi, 2017). To compensate the high demand of NADPH of HMGCoA reductase, authors overexpressed regenerating enzyme encoding genes such as ZWF1, a glucose-6-phosphate dehydrogenase and POS5, an NAPH kinase, which resulted in an increase of the bioconversion yields. Similarly, the effect of several NADPH-regenerating enzymes has been tested in yeast for the production of tropane alkaloids (Srinivasan and Smolke, 2020). Similar overexpressions of biosynthetic enzymes have been also carried out to increase SAM amounts inside yeast cells (Zhao et al., 2016).

For the production of podophyllotoxin precursors CaA and FA, Chen and coworkers first recreated a biosynthetic route starting from *p*-coumaric acid (*p*CA) whose synthesis is achieved from deamination of the aromatic amino acid tyrosine and phenylalanine (Li et al., 2015) (Figure 1). Several strategies have been applied to increase *p*CA synthesis up to 130.8 ± 17.0 mg L^{-1}. This first includes the improvement of the phenylpropanoid precursors availability through both strengthening the shikimate pathway by overexpressing alternative biosynthetic enzymes and reducing aromatic amino acid consumption *via* knocking-out genes encoding the phenylpyruvate decarboxylase (ARO10) and the pyruvate decarboxylase (PDC5). This was also combined to the direct exploitation of the tyrosine and phenylalanine precursors through overexpression of tyrosine ammonia lyase (TAL) and phenylalanine ammonia lyase (PAL) as well as the heterologous expression of a plant P450s complex containing two cinnamic acid 4-hydroxylases (C4H), a coumarate-3-hydroxylase (C3H) associated with a cytochrome P450 reductase (CPR). *In planta*, C3H mainly catalyzes the hydroxylation of shikimate esters but can also ensure conversion of *p*CA into CaA at low efficiency as reported in yeast. To overcome this limit, authors co-expressed an alternative pCA to CaA direct conversion step from bacteria composed of a FAD (H_2)-dependent 4-hydroxyphenylacetate 3-monooxygenase (HpaB) and a NADH-dependent flavin oxidoreductase (HpaC), which enabled a CaA production of 157.5 ± 1.2 mg L^{-1}. Next, besides overexpressing upstream genes of the pathway such as those encoding shikimate dehydrogenase and chorismate synthase, higher yields of CaA were obtained by increasing the amount of available NADPH fulfilled by the pentose phosphate pathway (Figure 1). Instead of overexpressing NADPH regenerating enzymes which are not always limiting, optimization was achieved through pulling the biosynthetic flux of the pentose phosphate pathway by overexpressing the downstream genes in order to favor NADPH regeneration. This "tour-de-force" was achieved by expressing phosphoketolase (Xfpk) splitting sugars such as fructose 6-phosphate and or xylose 5-phosphate into acetylphosphate and E4P for instance. Furthermore, this last compound has been also efficiently converted to acetyl-CoA by phosphotransacetylase (Pta) with the deletion of GAP

phosphatase (Gpp1) to favor growth as previously reported (Liu et al., 2019). A last optimization step was achieved by expressing a native transaldolase (Tald) to further drag flux toward E4P. These combined approaches increased CaA synthesis by 45%, reaching more than 360 mg L^{-1}, definitively confirming that optimizing NADPH regeneration *via* pentose phosphate pathway pulling is a key strategy to improve CaA titers.

Besides NADPH, expressing HpaB and HpaC also requires boosting the supply of FAD (H_2) to improve CaA synthesis. Authors thus deployed three strategies for enhancing FAD (H_2) availability: 1) consolidating the native FAD (H_2) synthesis that operates in both the cytosol and mitochondria, 2) managing the export of FAD from mitochondria to the cytosol and 3) implementing the bacterial FAD (H_2) biosynthetic pathway in yeast cytosol (Figure 1). Firstly, the cytosolic GTP hydrolase II (Rib1) which is the first-rate limiting step of FAD (H_2) synthesis as well the mitochondrial transporter FLX1 ensuring FA export to the cytosol have been overexpressed resulting in a 51,7% increase in the CaA titer. Then, to alleviate the allosteric inhibition of Rib1 by FAD (H_2), the bacterial RibBA displaying both guanosine triphosphate (GTP) cyclohydrolase II and 3,4-dihydroxy-2-butanone-4-phosphate synthase activities was implemented in yeast together with the bifunctional riboflavin kinase/FAD synthase (BiFads) that catalyze the conversion of riboflavin into FAD (H_2) in bacteria. Both expressions resulted in a 54% improvement of the CaA production. When implemented in the strain displaying an enhanced NADPH supply, those strategies were also successful regarding CaA production and were further enhanced by using strong and constitutive promoters for driving gene expression as well as the yeast cultivation in rich medium. The resulting improvement in the accumulation of CaA and its biosynthetic intermediate definitive confirms that the improvement of both the NADPH and FAD (H_2) supply is of prime importance to boost NP production.

The synthesis of FA relies on the *O*-methylation of CaA which can be catalyzed by different orthodiphenol-*O*-methyltransferases (Omt) including the highly efficient CaA *O*-methyltransferase from *Nicotiana benthamiana* (NtCOMT1) using SAM as a cofactor (Figure 1). While increasing the NtCOMT1 gene copy number together with NADPH and FAD (H_2) enhanced supply resulted in doubling the FA production, the remaining amount of CaA suggested that the amount of SAM could be limiting in strains highly producing CaA. Interestingly, the classical strategies for optimizing SAM supply such as the expression of rate-limiting methionine adenosyltransferase did not show any huge effect on effect on FA synthesis. By contrast, SAM recycling opened new avenues. Keeping in mind that *S*-adenosyl-L-homocysteine (SAH), which is a SAM by-product resulting from transmethylation, is a potent inhibitor of Omt activity, Chen and coworkers induced the degradation of SAH by overexpression of the SAH hydrolase (Sah1) thus producing homocysteine and adenosine (Figure 1). Since the thermodynamic equilibrium of Sah1 may favor the synthesis of SAH instead of its degradation, the consumption of adenosine has been enhanced through the co-expression of an adenosine kinase (Ado1), thus resulting in a 39% increase in FA production and a 64% conversion of CaA. Finally, such a concrete improvement of the three cofactors supply resulted in a production of 3.8 ± 0.3 g L^{-1} of FA when cultivating yeasts in fed-batch fermentation, thus constituting the best production reported to date.

Concluding remarks

While many metabolite productions in microbial cell factories did not reach industrial scales so far, the optimization of the biosynthetic fluxes represent one of the key steps in the deployment of the metabolic engineering-based approaches for the sustainable supply of valuables NPs. While many strategies to this aim have been previously developed, Chen and coworkers have demonstrated the prime importance of managing the supply of the cofactors used by the heterologous expressed enzymes. Their integrated strategy based on the synthesis and recycling of three distinct cofactors and ranging from recycling enzyme overexpression, subcellular pathway reorganization as well as inhibitory by-product degradation, resulted in dramatic increase of the CaA and FA synthesis. Besides paving the way for the high production of podophyllotoxin and other lignan derived compounds, this strategy combined to the classical optimization approaches will undoubtedly be used to improve the synthesis of many other NPs requiring co-factors in the near future.

Perspectives

In line with this recent improvement, yeast cell factories have recently developed for the *de novo* production of the highly valuable alkaloid vinblastine used in several chemotherapy (Zhang et al., 2022). This represents the longest biosynthetic pathway to be transferred from plant to a microorganism. Several additional genetic edits have been also performed to favor alkaloid synthesis including deletion of genes responsible for syphoning biosynthetic intermediates to shunt products. Interestingly, this biosynthetic pathway includes no less than seven different cytochromes P450 and 3 methyltransferases. It is thus temping to speculate that a combined enhancement of both NADPH and SAM supply would greatly increase the production of alkaloids in newly engineered yeast cells. In addition, besides being used as a potential platform for lignan production, the yeast strain developed by the Zhou's group can be readily valorized for the synthesis of vanillin the most popular flavor compound in the world. Indeed, vanillin synthesis only relies on the metabolization of FA by a hydratase/lyase type enzyme designated vanillin synthase (Gallage

et al., 2014). It is therefore obvious that overexpressing this enzyme in the engineered yeast strain would results in high vanillin production amounts. Finally, from a broader point of view, considering non-model organisms for the transfer of biosynthetic pathways could also be of high interest. This is notably the case for *Corynebacterium glutamicum* that can produce high levels of *p*CA (Mhatre et al., 2022).

Author contributions

JP wrote the manuscript. Corrections, formatting and editorial changes were done by SB, NP, and VC.

Funding

This work was supported by EU Horizon 2020 research and innovation program (MIAMi project-Grant agreement N°814645); ANR MIACYC project (ANR-20-CE43-0010) and ARD CVL Biopharmaceutical program of the Région Centre-Val de Loire (ETOPOCentre project). We also acknowledge funding from le Studium Institute (Consortium fellowship).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

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. doi:10.1016/j.ymben.2017.02.007

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. U. S. A.* 112 (11), 3205–3210. doi:10.1073/pnas.1423555112

Buss, A. D., Cox, B., and Waigh, R. D. (1995). in *Natural products as leads for new pharmaceuticals*. Editor M. E. Wolff (New York: Wiley). Chapter 24.

Carqueijeiro, I., Langley, C., Grzech, D., Koudounas, K., Papon, N., O'Connor, S. E., et al. (2020). Beyond the semi-synthetic artemisinin: metabolic engineering of plant-derived anti-cancer drugs. *Curr. Opin. Biotechnol.* 65, 17–24. doi:10.1016/j.copbio.2019.11.017

Chen, R., Gao, J., Yu, W., Chen, X., Zhai, X., Chen, Y., et al. (2022). Engineering cofactor supply and recycling to drive phenolic acid biosynthesis in yeast. *Nat. Chem. Biol.* 18, 520–529. doi:10.1038/s41589-022-01014-6

Darbani, B., Stovicek, V., van der Hoek, S. A., and Borodina, I. (2019). Engineering energetically efficient transport of dicarboxylic acids in yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 116 (39), 19415–19420. doi:10.1073/pnas.1900287116

Gallage, N., Hansen, E. H., Kannangara, R., Olsen, C. E., Motawia, M. S., Jorgensen, K., et al. (2014). Vanillin formation from ferulic acid in Vanilla planifolia is catalysed by a single enzyme. *Nat. Commun.* 5, 4037. doi:10.1038/ncomms5037

Grewal, P. S., Samson, J. A., Baker, J. J., Choi, B., and Dueber, J. E. (2021). Peroxisome compartmentalization of a toxic enzyme improves alkaloid production. *Nat. Chem. Biol.* 17, 96–103. doi:10.1038/s41589-020-00668-4

Guirimand, G., Kulagina, N., Papon, N., Hasunuma, T., and Courdavault, V. (2021). Innovative tools and strategies for optimizing yeast cell factories. *Trends Biotechnol.* 39, 488–504. doi:10.1016/j.tibtech.2020.08.010

Gülck, T., Booth, J. K., Carvalho, A., Khakimov, B., Crocoll, C., Motawia, M. S., et al. (2020). Synthetic iology of cannabinoids and cannabinoid glucosides in *Nicotiana benthamiana* and *Saccharomyces cerevisiae*. *J. Nat. Prod.* 83, 2877–2893. doi:10.1021/acs.jnatprod.0c00241

Kim, J. E., Jang, I-S., Son, S-H., Ko, Y-J., Cho, B-K., Kim, S. C., et al. (2019). Tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metab. Eng.* 56, 50–59. doi:10.1016/j.ymben.2019.08.013

Kulagina, N., Besseau, S., Papon, N., and Courdavault, V. (2021b). Peroxisomes: A new hub for metabolic engineering in yeast. *Front. Bioeng. Biotechnol.* 9, 659431. doi:10.3389/fbioe.2021.659431

Kulagina, N., Guirimand, G., Melin, C., Lemos-Cruz, P., Carqueijeiro, I., De Craene, J-O., et al. (2021a). Enhanced bioproduction of anticancer precursor vindoline by yeast cell factories. *Microb. Biotechnol.* 14, 2693–2699. doi:10.1111/1751-7915.13898

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.BMC Syst. Biol.* 8, 33–43. doi:10.1186/1752-0509-8-33

Li, M., Kildegaard, K. R., Chen, Y., Rodriguez, A., and Nielsen, J. (2015). Borodina I. and Nielsen J. *De novo* production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* 32, 1–11. doi:10.1016/j.ymben.2015.08.007

Liu, G. S., Li, T., Zhou, W., Jiang, M., Tao, X-Y., Liu, M., et al. (2020). The yeast peroxisome: a dynamic storage depot and subcellular factory for squalene overproduction. *Metab. Eng.* 57, 151–161. doi:10.1016/j.ymben.2019.11.001

Liu, Q., Yu, T., Li, X., Chen, Y., Campbell, K., Nielsen, J., et al. (2019). Rewiring carbon metabolism in yeast for high-level production of aromatic chemicals. *Nat. Commun.* 10, 4976. doi:10.1038/s41467-019-12961-5

Liu, R., Liu, L., Li, X., Liu, D., and Yuan, Y. (2020). Engineering yeast artificial core promoter with designated base motifs. *Microb. Cell Fact.* 19, 38. doi:10.1186/s12934-020-01305-4

Mhatre, A., Shinde, S., Jha, A. K., Rodriguez, A., Wardak, Z., Jansen, A., et al. (2022). Corynebacterium glutamicum as an efficient omnivorous microbial host for the bioconversion of lignocellulosic biomass. *Front. Bioeng. Biotechnol.* 10, 827386. doi:10.3389/fbioe.2022.827386

Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., et al. (2012). Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* 14, 104–111. doi:10.1016/j.ymben.2012.01.006

Paddon, C. J., and Keasling, J. D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* 12, 355–367. doi:10.1038/nrmicro3240

Paramasivan, K., and Mutturi, S. (2017). Regeneration of NADPH coupled with HMG-CoA reductase activity increases squalene synthesis in *Saccharomyces cerevisiae*. *Saccharomyces cerevisiaeJ. Agric. Food Chem.* 65, 8162–8170. doi:10.1021/acs.jafc.7b02945

Payne, J. T., Valentic, T. R., and Smolke, C. D. (2021). Complete biosynthesis of the bisbenzylisoquinoline alkaloids guattegaumerine and berbamunine in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2112520118. doi:10.1073/pnas.2112520118

Pyne, M. E., Kevvai, K., Grewal, P. S., Narcross, L., Choi, B., Bourgeois, L., et al. (2020). A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nat. Commun.* 11, 3337. doi:10.1038/s41467-020-17172-x

Ro, D-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943. doi:10.1038/nature04640

Srinivasan, P., and Smolke, C. D. (2020). Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* 585, 614–619. doi:10.1038/s41586-020-2650-9

Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., et al. (2003). Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat. Biotechnol.* 18, 143–149. doi:10.1038/nbt775

Wang, P., Wei, Y., Fan, Y., Liu, Q., Wei, W., Yang, C., et al. (2015). Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts. *Metab. Eng.* 29, 97–105. doi:10.1016/j.ymben.2015.03.003

Wang, Z., Wei, L., Sheng, Y., and Zhang, G. (2019). Yeast synthetic terminators: fine regulation of strength through linker sequences. *ChemBioChem* 20, 2383–2389. doi:10.1002/cbic.201900163

Zhang, J., Hansen, L. G., Gudich, O., Viehrig, K., Lassen, L. M. M., Schrubbers, L., et al. (2022). A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* 609, 341–347. doi:10.1038/s41586-022-05157-3

Zhao, W., Shi, F., Hang, B., Huang, L., Cai, J., and Xu, Z. (2016). The improvement of SAM accumulation by integrating the endogenous methionine adenosyltransferase gene SAM2 in genome of the industrial *Saccharomyces cerevisiae* strain. *Appl. Biochem. Biotechnol.* 178, 1263–1272. doi:10.1007/s12010-015-1943-1

Annexe III: Article "Tonoplast and Peroxisome Targeting of γ-tocopherol *N*-methyltransferase Homologs Involved in the Synthesis of Monoterpene Indole Alkaloids"

par Koudounas K, Guirimand G, Hoyos LFR, Carqueijeiro I, Cruz PL, Stander E, Kulagina N, Perrin J, Oudin A, Besseau S, Lanoue A, Atehortùa L, St-Pierre B, Giglioli-Guivarc'h N, Papon N, O'Connor SE, Courdavault V. publié en 2022 dans la revue **Plant and Cell Physiology**.

De nombreux AIM issus d'espèces végétales de la famille des Apocynacées, des Loganiacées et des Rubiacées sont exploités pour leur propriétés pharmacologiques intéressantes. Les voies de synthèse des AIM *in planta* sont complexes de part la multitude d'enzymes impliquées et l'organisation cellulaire et subcellulaire sophistiquée. En particulier, les *N*-méthyltransférases (NMT) représentent une famille d'enzymes indispensables à la synthèse des AIM, mais leur caractérisation, et en particulier, leur localisation *in planta*, est toujours restée difficile. Ces travaux présentent une analyse approfondie de la localisation subcellulaire de NMT issues de quatre plantes médicinales différentes (*C. roseus*, *P. somniferum*, *R. serpentina* et *V. minor*) par fusion des NMT à des protéines fluorescentes. Ce travail a, de plus, permis la caractérisation de deux NMT fonctionnelles chez *C. roseus.* Ces résultats soulignent la complexité dans l'organisation subcellulaire de la voie de biosynthèse des AIM en introduisant le tonoplaste et le peroxysome comme nouveaux acteurs dans la biosynthèse de ces molécules d'intérêt. Pour cette étude, j'ai plus spécifiquement participé au clonage des séquences codantes de certaines candidates *N*-méthyltransférases dans les vecteurs d'expression GFP et contribué à l'analyse des résultats.

PCP
PLANT & CELL PHYSIOLOGY

Tonoplast and Peroxisome Targeting of γ-tocopherol *N*-methyltransferase Homologs Involved in the Synthesis of Monoterpene Indole Alkaloids

Konstantinos Koudounas[1], Gregory Guirimand[1,2,†], Luisa Fernanda Rojas Hoyos[3], Ines Carqueijeiro[1], Pamela Lemos Cruz[1], Emily Stander[1], Natalja Kulagina[1], Jennifer Perrin[1], Audrey Oudin[1], Sébastien Besseau[1], Arnaud Lanoue[1], Lucia Atehortùa[4], Benoit St-Pierre[1], Nathalie Giglioli-Guivarc'h[1], Nicolas Papon[5], Sarah E. O'Connor[6] and Vincent Courdavault[1,*]

[1]EA2106 Biomolécules et Biotechnologies Végétales, Université de Tours, 31 Av. Monge, Tours 37200, France
[2]Graduate School of Sciences, Technology and Innovation, Kobe University, 1-1 Rokkodai, Nada, Kobe 657-8501, Japan
[3]Grupo de Biotransformación—Escuela de Microbiología, Universidad de Antioquia, Calle 70 No 52-21, A.A 1226, Medellín, Colombia
[4]Laboratorio de Biotecnología, Sede de Investigación Universitaria, Universidad de Antioquia, Medellin 50010, Colombia
[5]GEIHP, SFR ICAT, University of Angers, Université de Bretagne Occidentale, 4 rue de Larrey - F49933, Angers 49000, France
[6]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, Jena 07745, Germany
[†]LE STUDIUM RESEARCH FELLOW, Loire Valley Institute for Advanced Studies.

*Corresponding author: E-mail, vincent.courdavault@univ-tours.fr

(Received 2 July 2021; Accepted 31 October 2021)

Many plant species from the Apocynaceae, Loganiaceae and Rubiaceae families evolved a specialized metabolism leading to the synthesis of a broad palette of monoterpene indole alkaloids (MIAs). These compounds are believed to constitute a cornerstone of the plant chemical arsenal but above all several MIAs display pharmacological properties that have been exploited for decades by humans to treat various diseases. It is established that MIAs are produced *in planta* due to complex biosynthetic pathways engaging a multitude of specialized enzymes but also a complex tissue and subcellular organization. In this context, *N*-methyltransferases (NMTs) represent an important family of enzymes indispensable for MIA biosynthesis but their characterization has always remained challenging. In particular, little is known about the subcellular localization of NMTs in MIA-producing plants. Here, we performed an extensive analysis on the subcellular localization of NMTs from four distinct medicinal plants but also experimentally validated that two putative NMTs from *Catharanthus roseus* exhibit NMT activity. Apart from providing unprecedented data regarding the targeting of these enzymes *in planta*, our results point out an additional layer of complexity to the subcellular organization of the MIA biosynthetic pathway by introducing tonoplast and peroxisome as new actors of the final steps of MIA biosynthesis.

Keywords: Alkaloids • *Catharanthus roseus* • *N*-methyltransferase • Peroxisome • Subcellular localization • Tonoplast

Regular Paper

Introduction

Monoterpene indole alkaloids (MIAs) are plant specialized metabolites synthetized in thousands of plants from the Apocynaceae, Loganiaceae and Rubiaceae families (St-Pierre et al. 2013). Nowadays, more than 2,500 distinct MIAs have been described, which almost all derive from the common biosynthetic intermediate strictosidine (O'Connor and Maresh 2006). While their specific functions have not been firmly established to date, a growing number of evidences point out their involvement in responses of plants to environmental constraints and notably in defense processes to face bioagressor challenges (Dugé de Bernonville et al. 2017). Their high biological activities also confer to MIAs a prominent place in human pharmacopoeia with premium anticancer agents vinblastine and vincristine from the Madagascar periwinkle (*Catharanthus roseus*), the antimalarial quinine produced in the cinchona bark (*Cinchona officinalis*), the antiarrhythmic ajmaline synthesized by the devil pepper (*Rauwolfia serpentina*) or the nootropic vincamine accumulated in the dwarf periwinkle (*Vinca minor*) (O'Connor and Maresh 2006). Over the last decades, numerous studies established that MIAs are produced *in planta* thanks to complex routes encompassing up to 30–40 enzymatic steps in the aforementioned plants (Guirimand et al. 2010b, Courdavault et al. 2014, Caputi et al. 2018). These pathways involve enzymes from several protein families including P450, glucosyltransferases, β-glucosidases, dioxygenases, BAHD acyltransferases, decarboxylases, short- and medium-chain dehydrogenases/reductases,

Plant Cell Physiol. 00(0): 1–17 (2021) doi:https://doi.org/10.1093/pcp/pcab160, Advance Access publication on 10 November 2021, available online at https://academic.oup.com/pcp
© The Author(s) 2021. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.
All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

'Pictet-Spenglerases' or *O*-methyltransferases and *N*-methyltransferases (NMTs) as non-exhaustive examples (Dugé de Bernonville et al. 2015a). Furthermore, these manifold reactions are accompanied by a complex organization of the MIA biosynthetic pathways *in planta* distributed in no less than four different cell types and six distinct subcellular compartments in *C. roseus* (Courdavault et al. 2014).

In this context, the quest for identification of new actors of these biosynthetic pathways has always remained challenging in the past decades for many research groups. This was particularly the case for the *S*-adenosyl-L-methionine (SAM)-dependent NMT involved in the biosynthesis of vindoline, a MIA specific to *C. roseus*. This enzyme, named 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase (CrNMT), catalyzes the *N*-methylation of 16-methoxy-2,3-dihydro-3-hydroxytabersonine to desacetoxyvindoline in the biosynthesis of vindoline, one of the two precursors of vinblastine and vincristine (**Fig. 1**). While pioneering descriptions of this protein were published 30 years ago (De Luca et al. 1987, De Luca and Cutler 1987, Dethier and De Luca 1993), the corresponding gene was cloned nearly 20 years later (Liscombe et al. 2010). This elucidation was further accompanied by identifications of distinct NMTs involved in the synthesis of other MIAs in other medicinal plant species, including picrinine NMTs from both *V. minor* (VmPiNMT) and *R. serpentina* (RsPiNMT) yielding ervincine (**Fig. 1**) (Levac et al. 2016) but also norajmaline *N*-methyltransferase (RsNNMT) and ajmaline N_β-methyltransferase (RsANMT) from the ajmaline biosynthetic pathway in *R. serpentina* (**Fig. 1**) (Cázares-Flores et al. 2016). Importantly, phylogenetic analysis revealed that all these proteins originated from a common ancestor related to γ-tocopherol *C*-methyltransferase (γTMT), a broadly distributed enzyme catalyzing the formation of α-tocopherol within the vitamin E biosynthetic pathway. Accordingly, this phylogenetic relationship between all these plant NMTs gave rise to the so-called γ-tocopherol methyltransferase homolog group (γTMTh) (Liscombe et al. 2010, Cázares-Flores et al. 2016, Levac et al. 2016).

γ-Tocopherol methylation is common to all plants and was first characterized at the genetic level in Arabidopsis, cyanobacterium *Synechocystis* PCC6803 and pepper *Capsicum annuum*, revealing remarkable substrate specificities of γTMT, as observed later for MIA NMTs (Shintani and DellaPenna 1998, Koch et al. 2003). Interestingly, while the localization of the Arabidopsis γTMT (also referred to as VTE4) has been firmly established to be anchored in the plastid envelope membranes (Soll et al. 1985, Zbierzak et al. 2009), the subcellular localization of CrNMT still remains a matter of debate. Indeed, first cellular fractioning experiments associated CrNMT to the thylakoid membranes of chloroplasts (De Luca et al. 1987, De Luca and Cutler 1987, Dethier and De Luca 1993). However, analyses of the sequences predicted from cloned cDNAs failed to identify any chloroplast targeting peptide (cTP), leaving open the question of CrNMT subcellular localization (Liscombe et al. 2010, Cázares-Flores et al. 2016). Furthermore, with the exception of

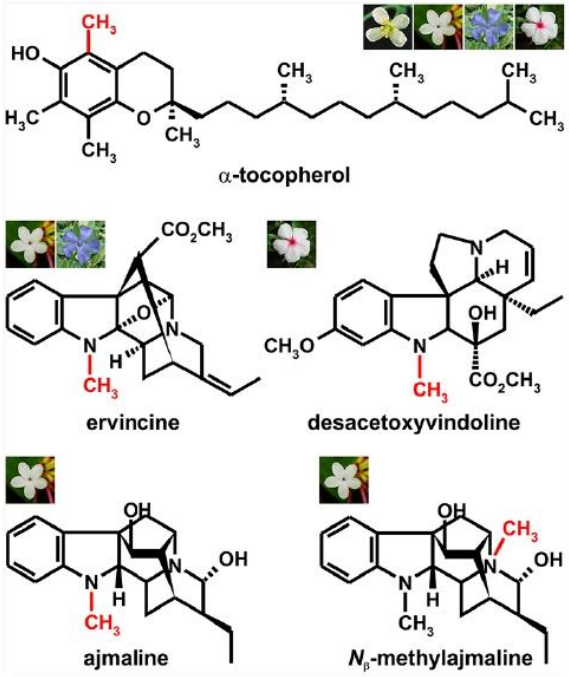


Fig. 1 Products resulting from *N*- or *C*-methyltransferase activities studied in this work. α-Tocopherol results from γ-tocopherol methylation by γTMT, ervincine from picrinine methylation by VmPiNMT or RsPiNMT, desacetoxyvindoline from 16-methoxy-2,3-dihydro-3-hydroxytabersonine methylation by CrNMT, ajmaline from norajmaline methylation by RsNNMT and N_β-methylajmaline by ajmaline methylation by RsANMT. The methyl group highlighted in red shows methylation position while flower pictures indicate methylated compound accumulation in *A. thaliana*, *R. serpentina*, *V. minor* and/or *C. roseus*.

dwarf periwinkle VmPiNMT that exhibits a potential secretory peptide, no other information regarding *Rauwolfia* RsPiNMT, RsANMT and RsNNMT localizations has been obtained to date. In opium poppy (*Papaver somniferum*), two NMTs involved in the biosynthesis of sanguinarine, namely, (*R*, *S*)-coclaurine *N*-methyltransferase (PsCNMT) and tetrahydroprotoberberine *N*-methyltransferase (PsTNMT), have been experimentally validated to be localized in the cytosol despite the existence of a potential tri-Lys (KKK) ER retention signal at the *C*-terminus (Hagel and Facchini 2012, Morris and Facchini 2019).

Overall, despite the importance of NMTs in the biosynthesis of numerous prominent alkaloids, little is known about the localization of these enzymes in plant cells. In this study, by expressing fusions with fluorescent proteins, we performed an extensive analysis of the subcellular localization of NMTs from four medicinal plants. Additionally, two previously unknown NMTs from *C. roseus* were found to catalyze *N*-methylation of picrinine, thus encoding functional NMTs *in planta*. Finally, beyond providing unprecedented information regarding NMT localization in plant cells, we added a new layer of complexity

to the subcellular organization of the MIA biosynthetic pathway by introducing tonoplast and peroxisome as new crucial compartments for the final steps of MIA synthesis.

Results

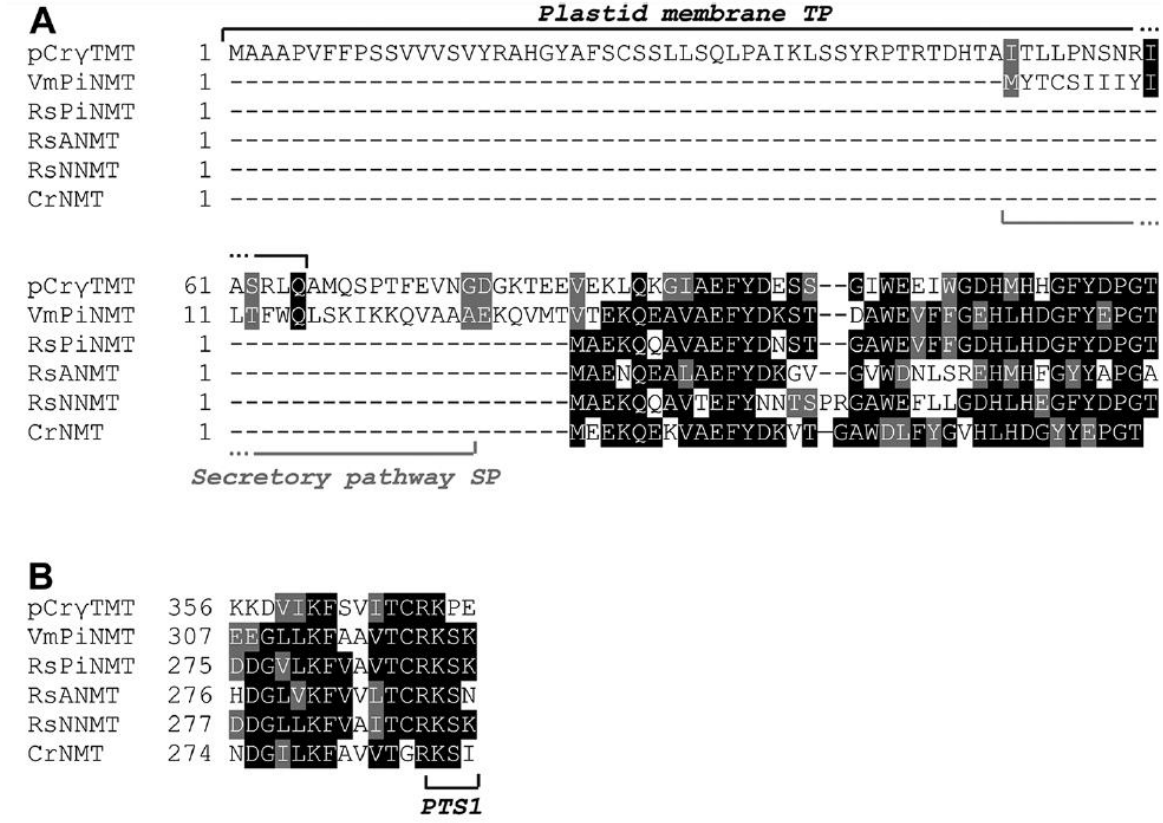
Identification of a putative C. roseus γTMT and prediction of γTMTh localizations

We first initiated this study by searching for an ortholog of the common ancestor of γTMThs in plants synthesizing MIAs to ensure that no divergence of γTMT localization occurs with previously published results (Zbierzak et al. 2009). The *C. roseus* CDF97 consensus transcriptome was thus screened by a BLAST analysis with VTE4 to identify the most similar transcript from periwinkle (SRR1144633_TR29246_c0_g1_i7_len = 1611; Dugé de Bernonville et al. 2015b). This sequence was further compared to the *C. roseus* genome to check for completeness and lack of other orthologs of high identity (Kellner et al. 2015). The protein deduced from the corresponding sequence (named hereafter putative *C. roseus* γTMT, pCrγTMT) exhibits an overall identity of 59% with VTE4 that reaches up to 71% without the highly divergent N-terminal end and includes the Rubisco large subunit methyltransferase motif regulating γTMT activity (Magnani et al. 2007) (**Supplementary Fig. S1**). Predictions of subcellular localizations performed using online bioinformatics predictors confirmed the presence of a putative cTP in the first 65 residues that could target this protein to plastids (**Fig. 2A, C**). The prediction was also applied to the five already cloned γTMThs. While this confirmed the previous identification of a putative secretory peptide at the N-terminal end of VmPiNMT (Levac et al. 2016) (**Fig. 2A, C**), no consensus in the prediction of its localization emerged from bioinformatics analyses that predicted targeting of protein to the endoplasmic reticulum (ER) or apoplast (**Fig. 2C**). A similar inconsistency was also obtained for RsPiNMT, RsANMT, RsNNMT and CrNMT, which were predicted to be located to peroxisomes, nucleus, cytosol or cytoskeleton although lacking any potential targeting sequences (**Fig. 2C**). For instance, prediction of RsNNMT and CrNMT to nucleus relies on the presence of weak monopartite nuclear localization sequences (NLS) (65-KKPR-68 in CrNMT) while possible targeting of RsPiNMT and RsANMT to peroxisomes mostly involve putative peroxisome targeting sequence type 1 (PTS1; 235-SRL-237 in RsANMT) located within the protein and are likely non-functional. Nevertheless, careful examination of γTMTh C-terminal extremities revealed the presence of a potential atypical PTS1 in CrNMT, according to previous studies (Reumann et al. 2016) (**Fig. 2B**). While this tripeptide (KSI) does not meet the exact atypical PTS1 definition, it includes one canonical (I) and two non canonical (KS) residues when two canonical and a single non-canonical residue are usually required. However and importantly, variations around PTS1 definition have been already described notably in *C. roseus* (Guirimand et al. 2012b). By contrast, C-terminal tripeptides of all other γTMThs only bear two canonical residues and are thus less prone to peroxisome targeting. In conclusion, all these discrepancies in the *in silico* prediction of localizations thus prompted us to experimentally investigate the subcellular localization of pCrγTMT and the five γTMThs involved in MIA biosynthesis.

The putative C. roseus γTMT is localized to membrane of plastid envelope

C. roseus cells are now well established as an appropriate host system for analyzing localization of MIA biosynthetic enzymes targeted to vacuoles (Guirimand et al. 2010a, Caputi et al. 2018, Carqueijeiro et al. 2018b), tonoplast (Payne et al. 2017) and plasma membrane (Héricourt et al. 2013), plastids (Guirimand et al. 2009, 2020, Thabet et al. 2012, Simkin et al. 2013), mitochondria (Guirimand et al. 2012a), ER (Besseau et al. 2013, Parage et al. 2016), nucleus (Stavrinides et al. 2015, 2016) and peroxisomes (Simkin et al. 2011, Guirimand et al. 2012b, Carqueijeiro et al. 2018a). On this basis, to compare localizations of AtγTMT (VTE4) and pCrγTMT, the VTE4 coding sequence was fused to YFP to express the VTE4-YFP fusion protein and to maintain cTP accessibility. By contrast, the pCrγTMT coding sequence was independently fused to both extremities of YFP to generate pCrγTMT-YFP and YFP-pCrγTMT to avoid masking of any other potential targeting peptide that would not have been identified by *in silico* prediction. After sequencing, each recombinant plasmid was co-transformed in *C. roseus* cells with plasmids expressing distinct subcellular compartment markers. In transiently transformed cells, pCrγTMT-YFP displayed a circular shaped fluorescence signal surrounding plastids (**Fig. 3A–H**). This signal did not overlap with the signal emitted by the CFP marker of the outer membrane of the plastid envelope, OEP7 (**Fig. 3I–L**) but perfectly merged with that of the marker of inner membrane of the plastid envelope, TIC40-CFP (**Fig. 3M–P**). Interestingly, VTE4-YFP displayed a similar distribution with circular shaped fluorescence signals co-localizing with the inner membrane of the plastid envelope (**Fig. 3Q–T**), strongly suggesting that it shares a similar localization with pCrγTMT-YFP. By contrast, the YFP-pCrγTMT exhibited a nucleocytosolic localization as confirmed by superimposition of the YFP fluorescent signal with that of the nucleocytosolic CFP marker (**Fig. 3U–X**), indicating that the N-terminal end of pCrγTMT must be accessible to drive this protein to plastids.

To further confirm that *C. roseus* cells are a suitable platform to study the subcellular localization of γTMThs from other species, we generated N- and C-termini fused YFP constructs of PsCNMT and PsTNMT from opium poppy. Transient co-expression of PsCNMT-YFP, YFP-PsCNMT, PsTNMT-YFP or YFP-PsTNMT constructs in *C. roseus* cells confirmed that the two NMTs exhibit a cytosolic localization since a perfect co-localization with the cytosolic marker (CFP-GUS) was observed (**Fig. 4**). As expected, these results are in agreement with previous subcellular localization studies of PsCNMT and PsTNMT in opium poppy cells (Hagel and Facchini 2012) and comfort the use of *C. roseus* cells for localization studies of heterologous enzymes.



C

	PSORTII	WolfPSORT	ChloroP	TargetP	Predotar
pCrγTMT	OUT	PLAST	cTP (65AA)	cTP (65AA)	PLAST
VmPiNMT	ER	OUT	-	SP (25AA)	ER
RsPiNMT	PER	CYTO	-	-	ELSW
RsANMT	PER	CYSK	-	-	ELSW
RsNNMT	NUC	CYTO	-	-	ELSW
CrNMT	NUC	CYTO	-	-	ELSW

Fig. 2 Alignment of the N-terminal (A) and C-terminal (B) extremities of the putative pCrγTMT with the five γTMThs and prediction of localizations (C). Black lines highlight potential plastid targeting peptide (cTP) in pCrγTMT (A) and the putative peroxisome targeting signal type 1 (PTS1) in CrNMT (B) while gray line indicates a secretory pathway signal peptide (SP) in VmPiNMT (A). (C) Predictions were performed with the PSORTII, WolfPSORT and Predotar online predicators and identification of potential targeting sequences were achieved with TargetP and ChloroP. Only localization with the highest probability is shown as follow, CYTO, cytosol; CYSK, cytoskeleton; ELSW, elsewhere; NUC, nucleus; PER, peroxisome; PLAST, plastid; OUT, outside.

VmPiNMT and RsPiNMT exhibit differential subcellular localization

To initiate this comprehensive localization study of NMTs, we first investigated the subcellular targeting of VmPiNMT by creating YFP fusions at both N- and C- termini of VmPiNMT. Since TargetP predicted the existence of a signal peptide (sp) with a 54% likelihood at aa 23–27 (-VAAAE-), we also generated two truncated versions encompassing the first 30 (spVmPiNMT-YFP) or first 80 (sp2VmPiNMT-YFP) amino acids of VmPiNMT fused with YFP at the C-terminal. Transient co-expression of either VmPiNMT-YFP or YFP-VmPiNMT in *C. roseus* cells revealed that VmPiNMT perfectly co-localized with the cytosolic marker (CFP-GUS) (**Fig. 5A–H**). This localization was further validated in onion cells (i.e. a non-Apocynaceae species) in which VmPiNMT also resulted in cytosolic localization (**Fig. 5I–L**). Plasmolysis of cells showed

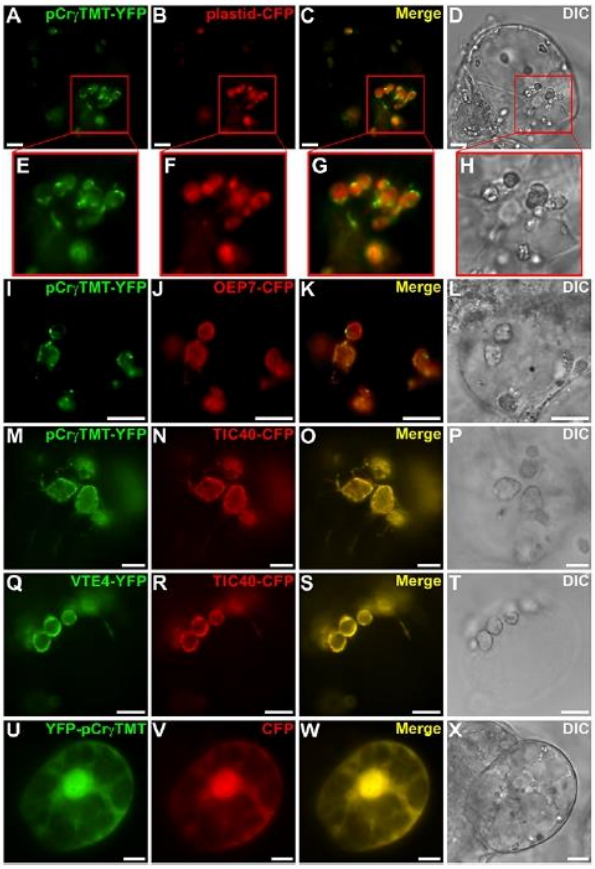


Fig. 3 pCrγTMT and VTE4-YFP are targeted to plastid envelop membrane. *C. roseus* cells were transiently co-transformed with plasmids expressing pCrγTMT (A, E, I, M, U) or VTE4-YFP (Q) and a plastid CFP marker (B, F), the marker of outer membrane of the plastid envelop OEP7-CFP (J), the marker of inner membrane of plastid envelop TIC40-CFP (N, R); or YFP-pCrγTMT (U) and a CFP nucleocytosolic marker (V). Co-localization of the two fluorescence signals appears on the merged images (C, G, K, O, S, W). Cell morphology (D, H, L, P, T, X) was observed with differential interference contrast (DIC). Bars, 10 μm.

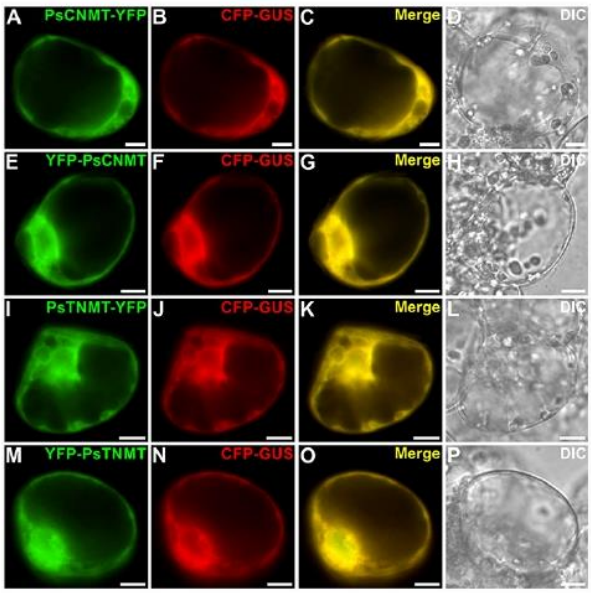


Fig. 4 Validation of the proper localization of PsCNMT and PsTNMT in *C. roseus* cells. *C. roseus* cells were transiently co-transformed with plasmids expressing PsCNMT-YFP (A) or YFP-PsCNMT (E) or PsTNMT-YFP (I) or YFP-PsTNMT (M) and a cytosolic marker (CFP-GUS; B, F, J and N). Co-localization of the two fluorescence signals appears on the merged images (C, G, K and O). Cell morphology (D, H, L and P) was observed with differential interference contrast (DIC). Bars, 10 μm.

that VmPiNMT remained in the cytosol thus confirming the lack of secretion (**Fig. 5M–N**). Additionally, both spVmPiNMT-YFP and sp2VmPiNMT-YFP constructs revealed a nucleocytosolic localization since the fluorescence signal was identical to the one from a nucleocytosolic CFP marker (**Fig. 5O–V**). Taken together, these results suggest that VmPiNMT does not encompass any signal peptide and is localized in the cytosol.

We next questioned whether the isoenzyme from *R. serpentina* (RsPiNMT) shares the same localization with VmPiNMT. Interestingly, transient expression of RsPiNMT-YFP in *C. roseus* cells revealed a potential tonoplastic-like localization, with fluorescence emanating from a large central vacuole (**Fig. 6A1–A2**) or from several small ones (**Fig. 6A3–A4**), which may also correspond to bulbs caused by YFP dimerization according to Segami et al. (2014). Such profiles of localization were also observed in onion cells, confirming that fusion protein targeting was independent from the cell type used for observations (**Fig. 6B1–C2**). However, many of our attempts to co-express RsPiNMT-YFP with a tonoplast marker (TPK-CFP) (Carqueijeiro et al. 2018b) resulted in the unexpected localization of both proteins and in the formation of aggregated foci during the first 24 h of observation (**Fig. 6D1–D4**). After longer incubations (36 h), a partial restoration of the tonoplastic localization was observed for both RsPiNMT-YFP and TPK-CFP thus suggesting that foci formation probably resulted from technical artifacts (**Fig. 6E1–E4**). Furthermore, co-transformation of RsPiNMT-YFP with a vacuolar marker (STR-CFP) allowed to confirm this tonoplastic localization as revealed by the YFP fluorescent signal enclosing that of the vacuole marker (**Fig. 6F1–F4**). In addition, expression of RsPiNMT N-terminally fused to YFP (YFP-RsPiNMT) revealed a cytosolic localization of the enzyme highlighted by the perfect co-localization with the cytosolic signal from the nucleocytosolic CFP marker and the loss of the tonoplastic targeting (**Fig. 6G1–G4**). This suggested that the N-terminal extremity of RsPiNMT needs to be accessible for proper localization of this enzyme. Surprisingly, the examination of the N-terminal amino acid sequence revealed that RsPiNMT does not encompass the typical tonoplastic targeting signal (D/E)$X_{3\text{-}5}$L(L/I) (Pedrazzini et al. 2013, Payne et al. 2017) or any other known tonoplast targeting peptide, and we were only able to identify a potentially atypical motif (133-DgtfdLV-139) in which the dibasic sequence encompasses a valine instead of a leucine or isoleucine residue. Taken together,

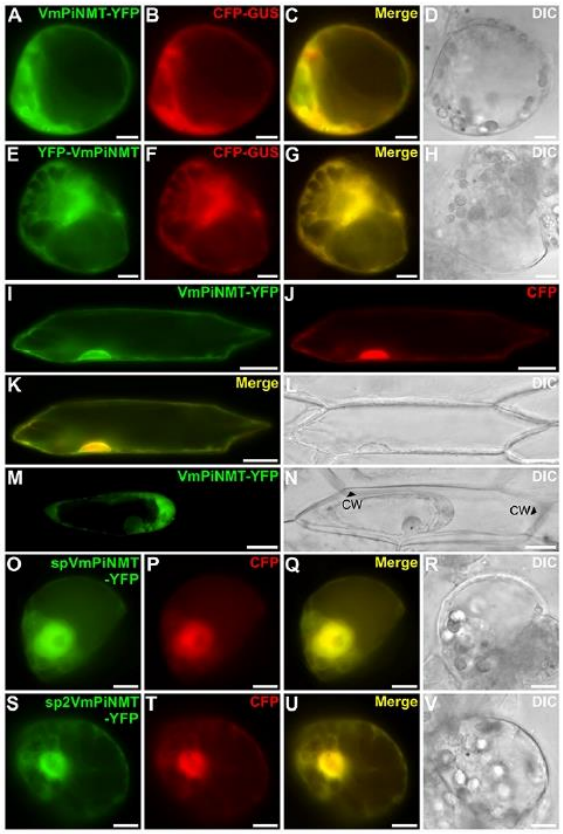


Fig. 5 VmPiNMT is localized in the cytosol. *C. roseus* cells (A-H, O-V) or onion cells (I-N) were transiently co-transformed with plasmids expressing VmPiNMT-YFP (A, I) or YFP-VmPiNMT (E) and a cytosolic marker (CFP-GUS; B, F,) or a nucleocytosolic marker (CFP; J). The plasmolysis of onion cells was performed to distinguish the signal from VmPiNMT-YFP and cell wall (CW, arrowheads). The existence of the predicted N-terminal signal peptide was tested by transiently co-transforming *C. roseus* cells with either spVmPiNMT-YFP (O) or tp2VmPiNMT-YFP (S) constructs and a nucleocytosolic CFP marker (P, T). Co-localization of the two fluorescence signals appears on the merged images (C, G, K, Q, U). Cell morphology (D, H, L, N, R, V) was observed with differential interference contrast (DIC). Bars, 10 μm (A–H, M–T) and 100 μm (I–N).

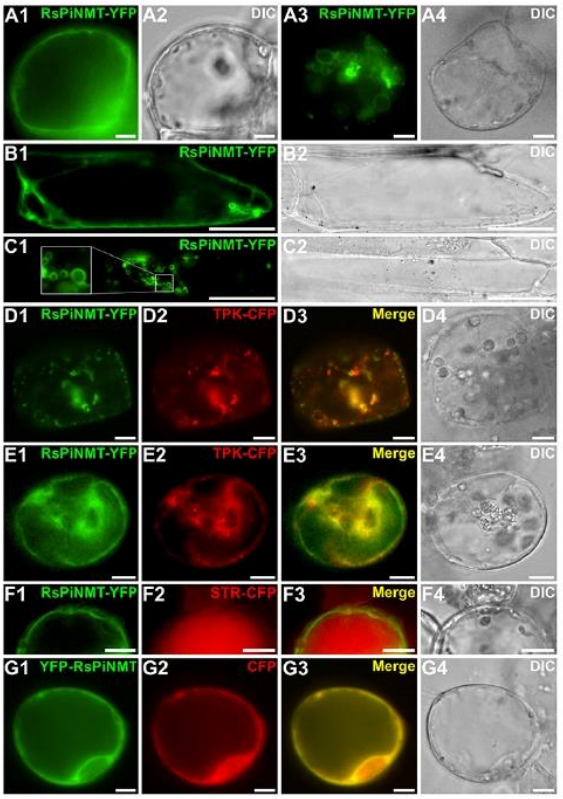


Fig. 6 RsPiNMT exhibits a tonoplast-like localization. *C. roseus* cells (A1, A3, D1, E1, F1 and G1) or onion cells (B1, C1) were transiently transformed with a plasmid expressing RsPiNMT-YFP (A1, A3, B1, C1), co-transformed with plasmids expressing either RsPiNMT-YFP and the tonoplastic marker (TPK-CFP; D2, E2) or the vacuolar marker (STR-CFP; F2); or YFP-RsPiNMT (G1) and a nuclocytosolic CFP marker (G2). Co-localization of the two fluorescence signals appears on the merged images (D3,E3,F3 and-G3). Cell morphology (A2, A4, B2, C2, D4, E4, F4 and G4) was observed with differential interference contrast (DIC). Bars, 10 μm (A1–A4, D1–G4) and 100 μm (B1–C2).

these results suggest that RsPiNMT displays a tonoplast localization via a—yet to be characterized—non-canonical N-terminal sequence.

RsNNMT and RsANMT exhibit a pre-vacuolar compartment-arrested localization

To gain more insight into the subcellular localizations of NMTs, additional localizations of enzymes from *R. serpentina* were achieved notably for RsNNMT and RsANMT from the ajmaline biosynthetic branch specific to *Rauwolfia sp*. Interestingly, the expression of RsNNMT and RsANMT fused to the N-terminal end of YFP (Rs-NNMT-YFP, RsANMT-YFP) revealed a punctuated pattern of fluorescence in *C. roseus* cells, but distinct from the one observed for RsPiNMT (**Fig. 7A–D**). Accordingly, such a pattern was also observed in onion cells for RsNNMT-YFP (**Fig. 7E–F**). However, we noticed that the fluorescence foci observed for RsNNMT-YFP and RsANMT-YFP perfectly superimposed with the signal emitted by pre-vacuolar compartment (PVC) marker RFP-VSR2 (Shen et al. 2013) (**Fig. 7G–J, O–R**). By contrast, expressing RsNNMT and RsANNMT fused with YFP in the opposite orientation (YFP-RsNNMT and YFP-RsANMT) led to the loss of this punctuated signal and resulted in a cytosolic localization as revealed by co-transforming the cells with the cytosolic marker (CFP-GUS) (**Fig. 7K–N, S–V**). Overall, these results thus suggest that RsNNMT and RsANMT contains an active targeting signal at their N-terminus thus suggesting localization of the native proteins in PVCs. However, it is worth

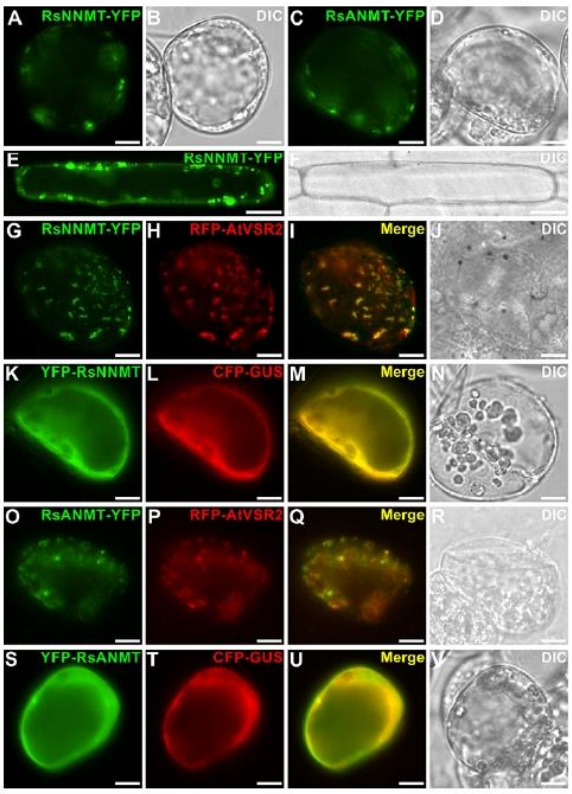


Fig. 7 RsNNMT and RsANMT exhibit a PVC-arrested localization. *C. roseus* cells (A–D, G–V) and onion cells (E–F) were transiently co-transformed with a plasmid expressing either RsNNMT-YFP (A–B; E–J), or RsANMT (C–D; O–R) or YFP-RsNNMT (K–N) or YFP-RsANMT (S–V) and either with the pre-vacuolar compartment (PVC) marker (RFP-VSR2; G–J; O–R) or the cytosolic marker (CFP-GUS; K–N; S–V) or without a marker (A–F). Co-localization of the two fluorescence signals appears on the merged images (I, M, Q, U). Cell morphology (B, D, F, J, N, R, V) was observed with differential interference contrast (DIC). Bars, 10 μm (A–D, G–V) and 100 μm (E–F).

noting that while PVC-arrested localizations were observed for both fusion proteins, no definitive vacuolar/tonoplast targeting was observed after prolonged observation as described for RsPiNMT.

CrNMT is targeted to peroxisomes in *C. roseus* cells

Since studies on NMTs subcellular localization revealed heterogeneous results—even among isoenzymes from different species (i.e. VmPiNMT and RsPiNMT), we next performed a more detailed analysis of the γTMThs homologs from *C. roseus*. Although CrNMT is well known to be engaged in the biosynthetic pathway of vindoline (De Luca et al. 1987, De Luca and Cutler 1987, Dethier and De Luca 1993, Liscombe et al. 2010), the subcellular localization of this enzyme has not been explored by direct methodologies (i.e. microscopic analysis of YFP-fused constructs) and the predicted plastid/thylakoid localization was only based on cell fractionation in sucrose gradient (De Luca and Cutler 1987, Dethier and De Luca 1993). Therefore, to further characterize the localization of this enzyme, we created fusions between CrNMT and the N- or C-terminal end of YFP (CrNMT-YFP and YFP-CrNMT) for transient expression in *C. roseus* cells. Expressing CrNMT-YFP together with a nucleo-cytoplasmic marker (CFP) suggested a nucleocytosolic localization of CrNMT-YFP as revealed by the perfect superimposition of the two fluorescence signals (**Fig. 8A1–A4**). By contrast, the inverted construct (YFP-CrNMT), displaying an accessible CrNMT C-terminal end, exhibited a more complex localization (**Fig. 8B1–B4**). While a substantial part of the fluorescence signal emerged from the cytosol, more discrete fluorescent punctuations detached from the cytosolic signal were also observed (**Fig. 8B1–B4**). Interestingly, co-transformation of the YFP-CrNMT expressing plasmid with the plasmid encoding the peroxisome CFP marker (perox-CFP) show a partially similar pattern of fluorescence and a perfect merging of the fluorescence signal emitted by punctuations (**Fig. 8C1–C4**). Besides showing that PSORTII and WolfPSORT predictions were not accurate as for most of the NMT studied in this work (**Fig. 2C**), this result suggested that CrNMT was at least partially targeted to peroxisomes in agreement with the potential presence of a non-canonical PTS1 (**Fig. 2B**).

To gain insight into this particular localization, additional transformations were performed and accompanied by cell treatments by cycloheximide. By inhibiting translation of newly produced fluorescent proteins, this treatment was shown to efficiently improve detection of peroxisomes in transiently transformed cells through limiting the overflow of the peroxisome import machinery and allowing peroxisome targeting of the remaining already produced fusion proteins (Guirimand et al. 2012b). Indeed, while YFP-CrNMT exhibited a similar mixed localization in both peroxisomes and cytosol at the beginning of the cycloheximide treatment (**Fig. 8D1–D4**), a complete targeting to peroxisomes was observed after 4 h, suggesting that peroxisomes may represent the exclusive destination of CrNMT (**Fig. 8E1–F4**). Closer observations highlighted the entire superimposition of the YFP-CrNMT fluorescence signal with that of the peroxisomal-CFP marker targeted to peroxisome lumen suggesting that CrNMT is rather located to this peroxisomal compartment. In addition, we also generated a truncated version of CrNMT deleted of the C-terminal PTS1 and at the C-terminal end of YFP (YFP-CrNMTΔ). Interestingly, this fusion protein totally lost the targeting to peroxisomes and only exhibited a nucleo-cytoplasmic localization (**Fig. 8G1–G4**). This last result combined to previous ones strongly suggest that CrNMT is targeted at least partially to peroxisomes through a C-terminal tripeptide (-KSI) corresponding to a non-canonical PTS1 signal that is necessary for proper localization *in planta*.

Subcellular organization of other *C. roseus* γTMThs and validation of their *N*-methyltransferase activity

The unexpected localization of CrNMT prompted us to study the localization of other γTMThs closely related to pCrγTMT.

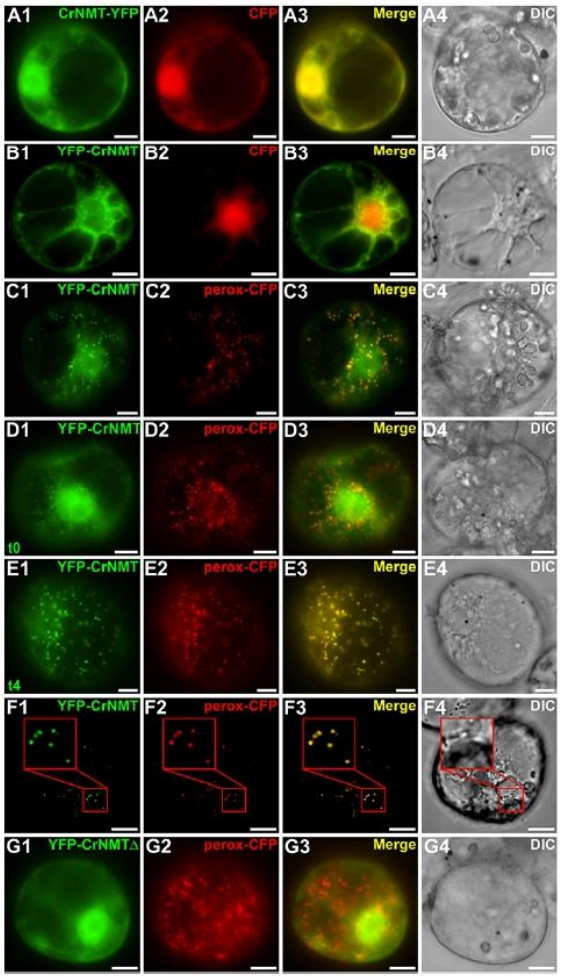


Fig. 8 CrNMT is localized in the peroxisomes. *C. roseus* cells were transiently co-transformed with a plasmid expressing either CrNMT-YFP (A1) or YFP-CrNMT (B1-F1) or YFP-CrNMTΔ (G1) and either the nucleocytosolic marker (CFP; A2 and B2) or the peroxisome marker (perox-CFP; C2–G2). To enhance peroxisome visualization, cells were cultivated 14 h prior harvesting and treatment with 2 μM cycloheximide. Cells were observed immediately (*t*0, D1) or 4 h (*t*4, E1) after treatment. A magnification of the signal emitted by YFP-CrNMT and the perox-CFP marker is presented in F1–F4. Co-localization of the two fluorescence signals appears on the merged images (A3–G3). Cell morphology (A4–G4) was observed with differential interference contrast (DIC). Bars, 10 μm.

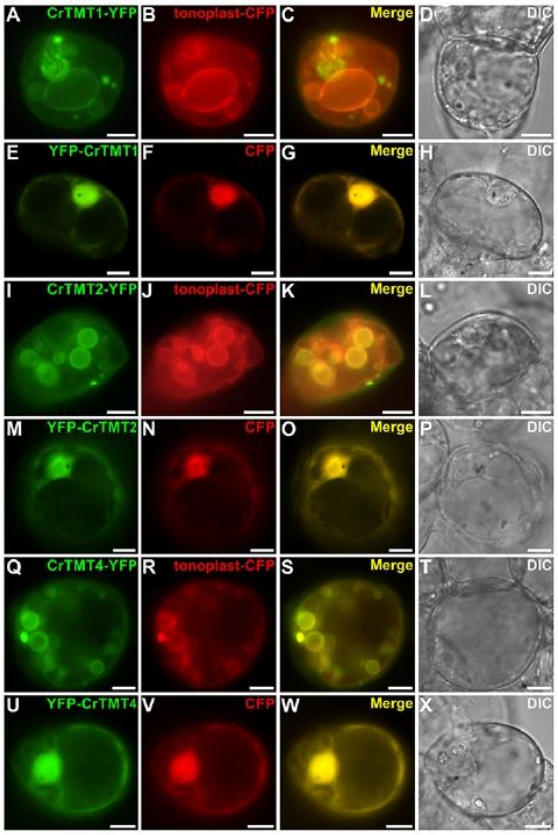


Fig. 9 The three *C. roseus* γTMThs (TMT1, TMT2, TMT4) are targeted to the tonoplast. *C. roseus* cells were transiently co-transformed with a plasmid expressing either CrTMT1-YFP (A–D) or CrTMT2-YFP (I–L) or CrTMT4-YFP (Q–T) or YFP-CrTMT1 (E-H), YFP-CrTMT2 (M–P) YFP-CrTMT4 (U–X) and either the tonoplast marker (TPK-CFP; B, J, R) or the nucleocytosolic marker (CFP; F, N, V). Co-localization of the two fluorescence signals appears on the merged images (C, G, K, O, S, W). Cell morphology (D, H, L, P, T, X) was observed with differential interference contrast (DIC). Bars, 10 μm.

Blasting the *C. roseus* transcriptome and genome allowed to identify three additional candidates namely CRO_113801 (CrTMT1) and the two previously reported CRO_111268 (CrTMT2) and CRO_11267 (CrTMT4) identified as Cr3710017 and Cr7756 by (Liscombe et al. 2010) (**Supplementary Figs. S2 and S3A**). All the corresponding genes were expressed in leaves, flowers and roots (**Supplementary Fig. S3B**). For localization studies, the three corresponding proteins were thus transiently expressed in *C. roseus* cells as fusions with either the N- or C-terminal end of YFP (**Fig. 9A–X**). Interestingly, on most of the observed cells, all the three γTMThs fused to the N-terminal end of YFP (CrTMT1-YFP, CrTMT2-YFP, CrTMT4-YFP) exhibited a fluorescence profile perfectly merging with that of the tonoplast marker (TPK-CFP) 24 h after transformation, with also potential bulbs (**Fig. 9A–D, I–L, Q–T**). By contrast, the constructs fused with YFP in the inverted orientation (YFP-CrTMT1, YFP-CrTMT2, YFP-CrTMT4) were not targeted to the tonoplast as revealed by the superimposition of fluorescence signal with the nucleocytosolic marker (CFP; **Fig. 9E–H**, M–P, U–X). All these results thus suggested that CrTMT1, CrTMT2 and CrTMT4 are targeted to the tonoplast as previously observed for RsPiNMT (**Fig. 6**). However, it is worth noting that none of the three γTMThs displays the consensus tonoplast targeting signal (D/E)$X_{3\text{-}5}$L(L/I) (Pedrazzini et al. 2013, Payne et al. 2017,

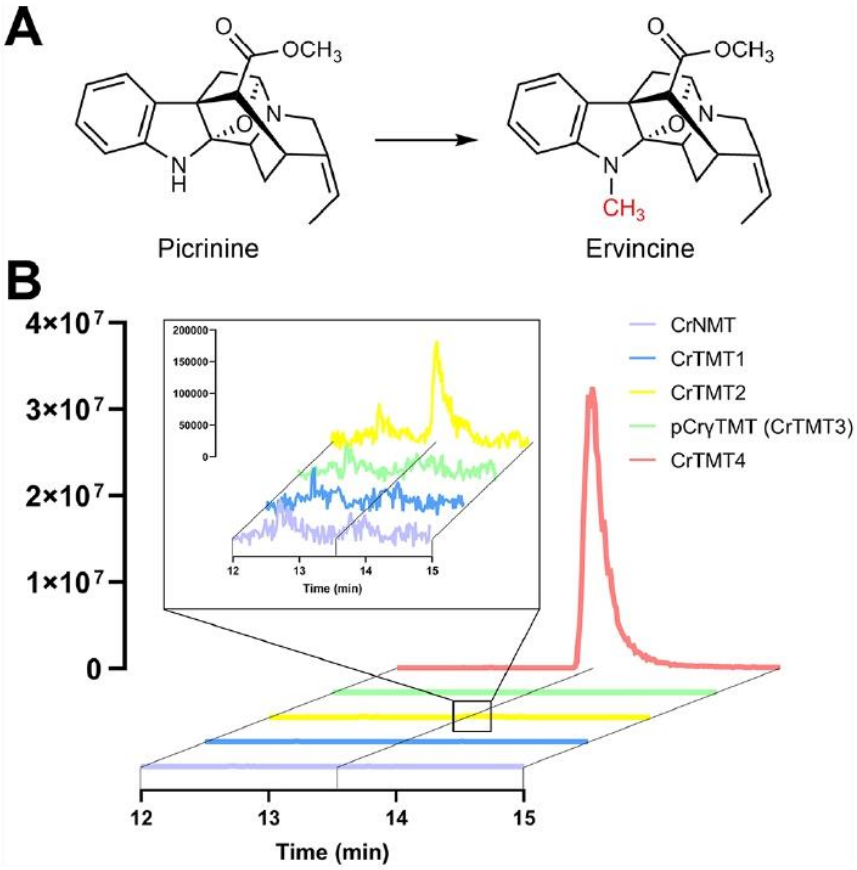


Fig. 10 *In vitro* enzymatic assays of recombinant CrTMTs against picrinine. (A) The chemical structures of picrinine (substrate) and ervincine (product). (B) UPLC-MS chromatograms at *m/z* 353 among enzymatic assays. Ervincine production was screened after incubating the respective purified recombinant proteins with picrinine. *Y*-axis corresponds to AU. A new peak at the *m/z* corresponding to *N*-methylated picrinine was detected only in the case of CrTMT4 and CrTMT2 (depicted in the inset).

Wojcik and Kriechbaumer 2021) and no closely related pattern could be identified within their N-terminal region.

To get more insight into the potential functions of these three additional γTMThs, we analyzed the genomic organization of the corresponding genes in the recently released draft genome of *C. roseus* (Franke et al. 2019). Interestingly, while *CrTMT1* (CRO_113801) and *pCrγTMT* (*CrTMT3*;CRO_130556) are distributed on two distinct scaffolds (cro_v2_scaffold_186 and cro_v2_scaffold_51, respectively), *CrTMT2* (CRO_111268) and *CrTMT4* (CRO_111267) are located in the same genomic region (cro_v2_scaffold_163) and in close proximity with *CrNMT* (CRO_111273) (**Supplementary Fig. S4**). This genomic organization thus suggests that *CrTMT2*, *CrTMT4* and *CrNMT* may originate from local gene duplication and possibly encode related enzymes. To investigate whether the putative γTMThs from *C. roseus* encode for functional NMTs, we generated recombinant enzymes in *E. coli* (**Supplementary Fig. S5**) and performed *in vitro* enzymatic assays. Given the sequence identity and similarity of localization between CrTMT2, CrTMT4 and RsPiNMT, *in vitro* assays were conducted with picrinine as a substrate to potentially produce ervincine (**Fig. 10A**). As expected, the γ-tocopherol C-methyltransferase pCrγTMT (CRO_130556) did not exhibit any *N*-methylation activity (**Fig. 10B**). In addition, no *N*-methylated product was observed when CrNMT (CRO_111273) or CrTMT1 (CRO_113801) was incubated with picrinine. By contrast, incubation of CrTMT2 (CRO_111268) or CrTMT4 (CRO_111267) with picrinine, resulted in a new peak at *m/z* 353 indicative of a potential conversion of picrinine into ervincine (*N*-methylpicrinine). The conversion was more than 100 times higher with CrTMT4 compared to CrTMT2. These results thus showed that both CrTMT2 and CrTMT4 exhibit alkaloid NMT activity and supported a potential evolutionary relationship with CrNMT.

Discussion

Numerous species from the Apocynaceae family produce MIAs that display high biological activities (O'Connor and Maresh 2006). Among the enzymes engaged in the corresponding biosynthetic pathways, alkaloid NMTs constitute a recently described subgroup of γTMThs, also involved in the synthesis of other alkaloid types such as benzylisoquinoline alkaloids

(BIAs) for instance (Liscombe et al. 2010, Cázares-Flores et al. 2016, Levac et al. 2016). Despite pioneering studies aiming at functionally characterizing alkaloid NMTs, little is known hitherto about their subcellular localization. Previous subcellular localization studies of alkaloid NMTs performed by direct methodologies (e.g. microscopic observation of fluorescent protein (FP)-fused chimeric constructs) have been limited to a few species including opium poppy and the coffee (*Coffea arabica*) and all revealed a cytoplasmic distribution of these enzymes (Kodama et al. 2008, Hagel and Facchini 2012). The cytoplasmic localization seems to be basically a common feature of methyltransferases since also reported for alkaloid *O*-methyltransferases involved in the biosynthesis of the BIA sanguinarine like (*R*,*S*)-3′-hydroxy-*N*-methylcoclaurine 4′-*O*-methyltransferase (4′OMT) and (*R*,*S*)-norcoclaurine 6-*O*-methyltransferase (6OMT) (Hagel and Facchini 2012) or the synthesis of vindoline (16 hydroxytabersonine 16-*O*-methyltransferase, 16OMT) (Guirimand et al. 2011b). Furthermore, other *O*-methyltransferases involved in iridoid, phenylpropanoid/flavonoid or anthocyanin/flavonol biosynthesis have been also located to cytoplasm (Hugueney et al. 2009, Guirimand et al. 2011a, Shaipulah et al. 2016). However, besides this cytoplasmic localization, plastids are known to host other methyltransferases such as the γTMT engaged in γ-tocopherol methylation (Zbierzak et al. 2009). In addition, *N*-, *O*- or *S*-methyltransferases not directly related to specialized metabolisms are also located to the nucleus (Choi et al. 2014), the ER membrane (Bracha-Drori et al. 2008), the Golgi apparatus (Lee et al. 2012) and the chloroplast (Ravanel et al. 2004, Tamiru et al. 2016). Therefore, the apparent highly diverse distribution of methyltransferases prompted us at deepening the study of their subcellular localization by notably focusing on NMTs involved in MIA synthesis.

We validated the use of *C. roseus* cells as a suitable platform for subcellular localization studies by confirming the localization of already characterized *C*-methyltransferases or NMTs. For instance, Arabidopsis AtγTMT (VTE4) correctly targeted the plastid membranes (**Fig. 3**) and the two opium poppy *N*-methyltransferases (PsTNMT and PsCNMT) in the cytosol, as previously reported (**Fig. 4**) (Zbierzak et al. 2009, Hagel and Facchini 2012). We also showed that the *C. roseus* homolog of AtγTMT (pCrγTMT, CrTMT3) was located on the inner membrane of the plastid envelope in agreement with its putative role in vitamin E synthesis (Muñoz and Munné-Bosch 2019) (**Fig. 3**).

Based on this system, we next highlighted differences of localization between two orthologues of PiNMT. In contrast to *in silico* predictions (**Fig. 2C**), we first described cytosolic localization of VmPiNMT invalidating the prediction of a potential targeting peptide at the N-terminal end (**Fig. 5**). This result was further confirmed in onion cells as well as with two truncated constructs bearing the first 30 or 80 amino acids fused in frame with YFP. Such a cytoplasmic localization is consistent with the distribution of PsTNMT, PsCNMT and NMTs involved in caffeine biosynthesis (Kodama et al. 2008). It also matches with the cytoplasmic localization of many other MIA biosynthetic enzymes such as tryptophan decarboxylase, iridoid synthase, 7-deoxyloganetic acid glucosyltransferase, loganic acid *O*-methyltransferase as well as desacetoxyvindoline 4-hydroxylase and deacetylvindoline 4-*O*-acetyltransferase (Miettinen et al. 2014, Dugé de Bernonville et al. 2015a) making of the cytosol a common crossroad for MIA production. However, we secondly showed that the ortholog from *R. serpentina* (RsPiNMT) that is devoid of any predicted targeting peptide displayed unexpected tonoplast localization (**Fig. 6**). Although not common, differences in ortholog localization has been already reported for the NMT theobromine synthase that exhibits distinct subcellular targeting among *C. arabica* and *Camellia gymnogyna* (Kodama et al. 2008, Teng et al. 2020). Different localization of RsPiNMT and VmPiNMT may thus point out to a differential compartmentalization for a similar reaction among the Rauvolfiinae and Vincinae subtribes. Interestingly, apart from RsPiNMT, we also showed that the three γTMThs identified in *C. roseus* (CrTMT1, CrTMT2 and CrTMT4) also exhibited a tonoplast localization when their N-terminal end was accessible (**Fig. 9**). These localizations suggest that the tonoplast localization may be observed for different methyltransferases as also previously suggested by proteomic approaches of tonoplast-enriched fractions for methyltransferase homologues (Endler et al. 2006, Shen et al. 2014). Although not common, other examples of specialized metabolisms have been associated with the tonoplast by direct methodologies such as enzymes involved in flavonoid biosynthesis (Saslowsky and Winkel-Shirley 2001, Toda et al. 2012). However, this tonoplast localization raised some important questions about the corresponding targeting mechanisms and membrane anchoring processes. Regarding targeting, we first noted that the tonoplast localization was disrupted when YFP was fused to the N-terminal of the enzymes since a nucleocytosolic localization was observed, thus confirming that N-terminal accessibility of RsPiNMT, CrTMT1, CrTMT2 and CrTMT4 is critical for a proper localization. In plants, localization of proteins to the tonoplast has been associated to the targeting signal $(D/E)_{X3\text{-}5}L(L/I)$ (Pedrazzini et al. 2013, Payne et al. 2017, Wojcik and Kriechbaumer 2021). However, no closely related motif could be identified in the N-terminal amino acid sequences of the four enzymes, except a potentially atypical motif in RsPiNMT. This still lets open the question of how these methyltransferases are targeted to the tonoplast but the lack of clearly identifiable sequences seems to be not prohibitive since two calcium sensors from Arabidopsis (CBL2 and CBL3) are known to be localized in the tonoplast *via* N-terminal domains while not exhibiting any known tonoplast targeting signal (Tang et al. 2012).

Regarding tonoplast anchoring, no evident potential transmembrane helices could be predicted in the primary sequences of RsPiNMT, CrTMT1, CrTMT2 and CrTMT4 albeit some discrepancies exist between the MemBrain and TMHMM servers (**Supplementary Table S2**). Similarly, no known post-translational protein modification involved in membrane anchoring could be identified with a high confidence

(**Supplementary Table S3**). This outcome resembles the Arabidopsis VTE4 and its rice homologue that are located to the inner membrane of the plastid envelope without any identified transmembrane domains (Bréhélin et al. 2007, Zbierzak et al. 2009, Jo and Hyun 2011). In addition, it should be noted that the characterization of interactions among flavonoid biosynthetic enzymes assembling in varying metabolons suggested that the observed localizations could be due to interactions between proteins (Nakayama et al. 2019). Such a mechanism could thus be also considered for tonoplast association of alkaloid NMTs. Besides this mechanistic consideration, the tonoplast targeting of RsPiNMT, CrTMT1, CrTMT2 and CrTMT4 also questioned about possible roles in MIA synthesis. Interestingly, the anabolism of some MIA precursors in *C. roseus* is known to depend on several trans-tonoplast translocations of intermediates (Guirimand et al. 2011a) and at least two enzymes involved in MIAs biosynthesis (i.e. STR; strictosidine synthase and PRX1; peroxidase 1) are known to be localized in vacuoles where strictosidine synthesis occurs (Costa et al. 2008, Guirimand et al. 2010a). It is thus possible to hypothesize that the tonoplast localization of alkaloid NMTs would favor the methylation reaction of MIAs that are ultimately stored in the vacuole. Interestingly, we also showed that both RsNNMT and RsANMT exhibited a PVC localization since the corresponding YFP-fused proteins co-localized with the PVC marker RFP-VSR2 (Shen et al. 2013) (**Fig. 7**). However, numerous observations suggested that the localization was PVC-arrested since no ultimate vacuolar targeting was observed. This is thus slightly different from the precondylocarpine acetate synthase which is progressively secreted to the vacuole by ER-derived vesicles (Caputi et al. 2018). This is also different from other enzymes that reach the tonoplast through PVCs as illustrated for AtβFructosidase 4 (Jung et al. 2011). On the other hand, this result is in agreement with older observations that associated enzymes involved in MIAs precursor biosynthesis with vesicles in *C. roseus* (Madyastha et al. 1977) as well as for columbamine-*O*-methyltransferase (Rueffer et al. 1986), stilbene synthase (Fornara et al. 2008) and nicotianamine synthase (Nozoye et al. 2014).

While the 16-methoxy-2,3-dihydro-3-hydroxytabersonine-*N*-methyltransferase activity has been described more than 30 years ago in *C. roseus* and the corresponding enzyme identified more recently (De Luca et al. 1987, De Luca and Cutler 1987, Dethier and De Luca 1993, Liscombe et al. 2010), the subcellular localization of CrNMT has still remained a matter of debate. In this work, we showed that CrNMT is at least partially targeted to peroxisomes via a non-canonical PTS1 (**Figs. 2B**, **8**). Most of the proteins targeted to peroxisomes encompass a PTS1 tripeptide at their extreme C-terminal end, which mostly correspond to the canonical consensus [SA][RK][LMI] sequences (Reumann et al. 2016). However, several studies have validated that various combinations of tripeptides encode for functional but non-canonical PTS1s such as the [SAPCFVGTLKIQ][RKSNLMHGETFPQCYDA][LMIVYF] sequences (Reumann et al. 2016). The putative non-canonical PTS1 (-KSI) at the C-terminal of CrNMT falls within the reported range of residues that constitute the non-canonical PTS1 tripeptides. This has been experimentally validated by deletion of this tripeptide which resulted in the loss of the peroxisome targeting of YFP-CrNMTΔ (**Fig. 8**). While CrNMT has been originally associated to the thylakoid membranes of chloroplasts by subcellular fractionation (De Luca et al. 1987, De Luca and Cutler 1987, Dethier and De Luca 1993), no plastid targeting sequences has been identified after identification of the corresponding gene (Liscombe et al. 2010). The previously reported plastid localization may thus be due to the initial experimental procedure used for enzyme characterization since peroxisomes are known to sediment similarly to chloroplasts (Abas and Luschnig 2010) resulting in cross organellar contamination in isolated fractions (Eubel et al. 2007). From a physiological point of view, the targeting of CrNMT to peroxisomes reinforces the involvement of this subcellular compartment in the biosynthesis of MIAs since it also hosts enzymes involved in the synthesis of the monoterpenoid precursor of MIAs (Simkin et al. 2011, Thabet et al. 2011, Guirimand et al. 2012a). In addition, other enzymes involved in alkaloid synthesis have been previously reported to be localized in peroxisomes such as the *N*-methylputrescine oxidase and its close homolog NtDAO1 engaged in nicotine biosynthesis in tobacco (Naconsie et al. 2014). To date, Arabidopsis HARMLESS TO OZONE LAYER 3 (AT2G43940) is the only example of a plant methyltransferase known to be localized in peroxisomes, although the *in planta* function of this enzyme remains unknown (Pan et al. 2018). Interestingly, COBALAMIN-INDEPENDENT METHIONINE SYNTHASE1 (ATMS1 or METS1; AT5G17920) an enzyme involved in the production of SAM—the universal donor of most methyltransferases—was also identified as a peroxisomal by N-terminal FP fused constructs confirming the capacity of methyltransferases to function in this compartment (Quan et al. 2013, González and Vera 2019). Other transferases (EC: 2.-) involved in primary metabolism that exhibit a peroxisomal localization include acyltransferases (Pracharoenwattana et al. 2005, Carrie et al. 2007, Tanabe et al. 2011), an aminotransferase (ASP3; At5g11520) (Eubel et al. 2008) and glutathione S-transferase T1 (GSTT1; At5g41210) (Reumann et al. 2007) supporting the idea that peroxisome is a hub for many metabolic pathways. In the tabersonine to vindoline pathway, while all other reactions occur in the cytosol, even for the ER-anchored P450, the peroxisomal targeting of CrNMT may constitute a regulatory step in this biosynthetic branch, besides adding a new layer of complexity to the architecture of the MIA biosynthetic pathway (**Fig. 11**).

Finally, the analysis of the genomic organization encompassing CrNMT showed that CrTMT2 and CrTMT4 are clustered in close proximity with CrNMT suggesting they may be functionally related genes (Nützmann et al. 2016). Physically clustered homologs most likely result from tandem gene duplication events and the expanded gene pools can potentially give rise to sub- or neo-functionalization due to absence of selective pressure on the new copies (Moore and Purugganan 2005,

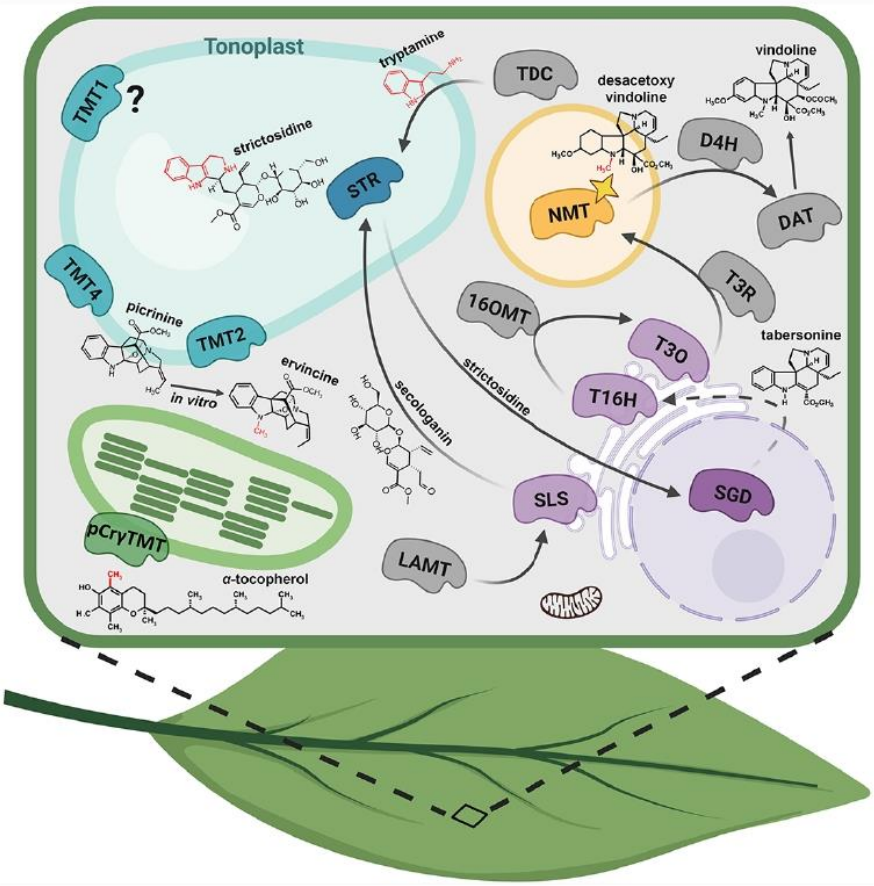


Fig. 11 Schematic illustration of the subcellular organization of *C. roseus* γTMThs and MIA-related enzymes in cells. Enzyme localization is color-coded: green—plastid envelope membrane; light blue—tonoplast; dark blue—vacuole; orange—peroxisome; grey—cytosol; light purple—ER; dark purple—nucleus. The yellow star represents the identified atypical peroxisome targeting sequence (PST1). Grey arrows show pathway enzymatic steps. The dark grey dashed arrow indicates that several enzymatic steps were omitted for scheme simplification. LAMT: loganic acid *O*-methyltransferase; SLS: secologanin synthase; TDC: tryptophan decarboxylase; STR: strictosidine synthase; SGD: strictosidine β-glucosidase; T16H: tabersonine 16-hydroxylase; 16OMT: tabersonine 16-*O*-methyltransferase, T3O: tabersonine 3-oxygenase, T3R: tabersonine 3-reductase, NMT: 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase, D4H: deacetoxyvindoline 4-hydroxylase, DAT: deacetylvindoline *O*-acetyltransferase, TMT: γ-tocopherol methyltra–nsferase homolog.

Panchy et al. 2016, Lichman et al. 2020). Genomic clustering of paralogs has been reported for some MIA biosynthetic genes (Kellner et al. 2015, Zhao et al. 2017, Franke et al. 2019, Rai et al. 2021) and, similarly to the three aforementioned NMTs in *C. roseus*, genomic clustering of *O*-methyltransferases involved in flavonoid biosynthesis has been reported in Fabaceae and Lamiaceae species (De Vega et al. 2015, Mochida et al. 2017, Zhao et al. 2019). To support the concept of functional relationship, we conducted enzymatic assays with recombinant enzymes and picrinine as a substrate given that RsPiNMT and CrTMT4 are closely related (**Supplementary Fig. S3**, **Fig. 10**). Through the production of ervincine, these assays specifically revealed that CrTMT2 and CrTMT4 encode functional alkaloid NMTs capable of picrinine methylation. Interestingly, we observed that RsPiNMT, CrTMT2 and CrTMT4 share similar localization and activity. By contrast, CrTMT1 that is not clustered with CrNMT within the genome of *C. roseus* does not display the picrinine NMT activity suggesting a more distant function. Although we used picrinine as a substrate to screen for enzymatic activity among the candidates, this alkaloid is not known to be produced in *C. roseus*. Therefore, our assay only confirmed that CrTMT2 and CrTMT4 are *bona fide* alkaloid NMTs but their roles *in planta* as well as their respective actual substrates remain to be addressed. Most likely these enzymes are involved in *N*-methylation of minor alkaloids that are closely related to the structure of picrinine like vincorine (Jossang et al. 1998) or cathafoline (Rasoanaivo et al. 1973, Jacquier et al. 1980) that has been identified in *C. roseus* and *C. longifolius*, respectively. Other minor *N*-methylated alkaloids identified in *C. roseus* that could potentially be substrates of CrTMT2 and

CrTMT4 include cathovaline (Rahman and Chaudhry 1985) and bannucine (Rahman et al. 1986). The lack of commercial sources and even standards of minor alkaloids hampers the functional characterization of these enzymes. However, reconstitution of biosynthetic pathways in heterologous systems could be an attractive approach to produce tailor-made precursors in high quantities and thus facilitate further functional studies (Carqueijeiro et al. 2020, Courdavault et al. 2020, Guirimand et al. 2021).

In conclusion, we performed an extensive localization study to shed light on the subcellular logistics of alkaloid NMTs leading to the association of tonoplast and peroxisomes to MIA synthesis. Beyond filling a gap into the controversial localization of CrNMT, our analysis also expands our knowledge of the metabolic channeling within the plant cells. In the developing context of natural product engineering in heterologous hosts (Courdavault et al. 2020), such compartments could inspire new metabolic engineering approaches as recently described for yeast peroxisomes (Dusséaux et al. 2020, Kulagina et al. 2021).

Materials and Methods

Plant and cell growth conditions

Seeds of *C. roseus* and *R. serpentina* from our greenhouse were germinated according to Carqueijeiro et al. (2015) and Corbin et al. (2017), respectively. The resulting plants were then grown at 28°C under a 16 h light/8 h dark cycle. *V. minor* leaves and roots were collected in the botanical garden of Tours. All these samples were used for RNA extraction. The *C. roseus* C20A cell suspension culture used for subcellular localization studies was propagated in Gamborg's B5 medium (Duchefa) at 24°C under continuous shaking (100 rpm) for 7 days as described previously (Guirimand et al. 2009).

RNA extraction and reverse-transcription

RNAs were extracted from young leaves and roots of *C. roseus*, *R. serpentina* and *V. minor* using the NucleoSpin RNA Plus Kit (Macherey-Nagel) according to manufacturer's recommendations. First-strand cDNA was synthesized from 0.5 μg of total RNA using oligo(dT)18 primers (0.5 μM) and 15 units of Thermoscript reverse transcriptase (Invitrogen). Following reverse transcription, complementary RNA was removed by treatment with *Escherichia coli* RNase H (Invitrogen) for 20 min at 37°C as described previously (Parage et al. 2016).

Transcriptomic resources and sequence alignment

Identification of γTMT orthologs was achieved by interrogating the consensus *C. roseus* transcriptome CDF97 (Dugé de Bernonville et al. 2015b) (http://bbv-ea2106.sciences.univ-tours.fr/images/files_to_download/BAFC94.fasta) using VTE4 as query. Sequence alignments were performed using Clustal Omega and graphically displayed with BoxShade (https://embnet.vital-it.ch/software/BOX_form.html).

Bioinformatics prediction of subcellular localizations

Whole amino-acid sequences of the putative *C. roseus* γTMTh, VmPiNMT, RsPiNMT, RsANMT, RsNNMT and CrNMT were used for bioinformatics predictions performed using the PSORTII, WolfPSORT, Predotar, ChloroP 1.1 and TargetP 1.1 servers. For all predictions, default 'cut-off' and plant were selected. Prediction of transmembrane helices was conducted with TMHMM (Krogh et al. 2001) and MemBrain (Yin et al. 2017) servers. Prediction of protein post-translational lipid modifications was conducted with GPS-Lipid (Xie et al. 2016).

Constructs for subcellular localization studies

Coding sequences were amplified with the corresponding primers described in **Supplementary Table S1** and the PCR products were digested with *Spe*I. In order to generate constructs fused in frame with YFP at the C-terminal, each product was ligated in a *Spe*I-digested pSCA-YFP plasmid (Guirimand et al. 2009). To generate constructs fused in frame with YFP at the N-terminal, each product was ligated in a *Nhe*I-digested pSCA-YFP plasmid. The subcellular compartment markers have been described in previous publications: plastid-CFP, OEP7-CFP and TIC40-CFP (Guirimand et al. 2020); cytosolic CFP and nucleocytosolic CFP-GUS (Guirimand et al. 2011a), TPK-CFP (Héricourt et al. 2013), STR-CFP (Guirimand et al. 2010a), RFP-AtVSR2 (Shen et al. 2013) and peroxisome (perox)-CFP (Guirimand et al. 2012a).

Biolistic transformation of C. roseus cells and epifluorescence microscopy

Transient transformation of *C. roseus* and onion cells by particle bombardment and fluorescence imaging were performed following the procedures previously described (Guirimand et al. 2009, Foureau et al. 2016, Carqueijeiro et al. 2018b). *C. roseus* or onion cells were bombarded with DNA-coated gold particles (1 μm) and 1,100 psi rupture disc at a stopping-screen-to-target distance of 6 cm, using the Bio-Rad PDS1000/He system and 400 ng of each plasmid per transformation. Cells were cultivated for 16–38 h before being harvested and observed. Cell plasmolysis was performed according to Carqueijeiro et al. (2018b). The subcellular localization was determined using an Olympus BX-51 epifluorescence microscope equipped with an Olympus DP-71 digital camera and a combination of YFP (Chroma#31040, 500–520 nm excitation filter, 540–580 nm band pass emission filter) and CFP (Chroma#31044v2, 426–446 nm excitation filter, 460–500 nm band pass emission filter) filters. Patterns of localization are representative of around 100 observed cells.

Cycloheximide treatment of transiently transformed cells

Transiently transformed cells were treated with cycloheximide (2 μM) for 0 or 4 h prior to observation as described in (Guirimand et al. 2012b).

Production of recombinant C. roseus TMTs in E. coli and enzymatic assays

The coding sequences of CrTMT1, CrTMT2, pCrγTMT, CrTMT4 and CrNMT were amplified with the corresponding primers (**Supplementary Table S1**) and cloned in either pRSET-A (CrTMT1, pCrγTMT and CrNMT) or pQE-30 (CrTMT2 and CrTMT4) expression vectors using the BamHI restriction site. Recombinant protein production was performed in either BL21(DE3) or Top10 *E. coli* strains for pRSET-A and pQE-30 constructs, respectively. Cultures of an $OD_{600} = 0.5$ were induced for 5 h at 28°C with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Protein purification was performed with a Co^{2+} resin (TALON Metal Affinity Resin, Cat. No. 635502, Clontech) according to the manufacturer's instructions. Purified proteins were desalted and dialyzed in 50 mM potassium phosphate buffer pH 7.5 using PD-10 columns (Cat. No. 17085101, GE healthcare). Concentration of proteins was determined by a standard Bradford assay, aliquots in 10% glycerol were flash frozen in liquid nitrogen and stored at −80°C. Enzymatic assays were conducted with 1 μg of purified recombinant proteins in 200 μL activity buffer (100 mM Tris-HCL pH 7.5, 2 mM SAM, 14 mM β-mercaptoethanol and 5 μg picrinine) at 37°C for 120 min under continuous shaking. The reactions were stopped by adding an equal volume of methanol (100%), followed by centrifugation at 16,000 g for 10 min and the supernatants

were analyzed by UPLC-MS. UPLC-MS analyses were performed as described previously (Dugé de Bernonville et al. 2017). Data collection was carried out in selected ion monitoring mode for picrinine (m/z 339; Rt = 11.29) and ervincine (m/z 353; Rt = 13.54).

Supplementary Data

Supplementary data are available at *PCP* online.

Data Availability

The cDNA sequences of the enzymes studied in this article can be accessed in the GenBank database (https://www.ncbi.nlm.nih.gov/genbank/) under accession no. KC708450 (VmPiNMT), KC708448 (RsPiNMT), KC708445 (RsANMT), KC708449 (RsNNMT), HM584929 (CrNMT), MZ367707 (CrTMT1), MZ367708 (CrTMT2), MZ367709 (pCrγTMT), MZ367710 (CrTMT4), AF104220 (AtVTE4; AT1G64970), AY217336 (PsCNMT), DQ028579 (PsTNMT).

Funding

The authors acknowledge funding from the ARD2020 Biopharmaceutical program of the Région Centre Val de Loire (BioPROPHARM, CatharSIS and ETOPOCentre projects), and ANR MIACYC (ANR-20-CE43-0010). GG and VC acknowledge 'le STUDIUM' for a STUDIUM Research Fellowship and STUDIUM Research Consortium grant, respectively.

Acknowledgements

The authors would like to acknowledge Prof. Peter Facchini (University of Calgary) for kindly providing plasmids encompassing the PsCNMT and PsTNMT sequences, Prof. Vincenzo De Luca (Brock University) for kindly providing plasmids encompassing the RsNNMT and RsANMT sequences and Prof. Liwen Jiang (Chinese University of Hong Kong) for kindly providing the RFP-VSR2 marker.

Disclosures

The authors have no conflicts of interest to declare.

References

Abas, L. and Luschnig, C. (2010) Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. *Anal. Biochem.* 401: 217–227.

Besseau, S., Kellner, F., Lanoue, A., Thamm, A.M.K., Salim, V., Schneider, B., et al. (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.

Bracha-Drori, K., Shichrur, K., Lubetzky, T.C. and Yalovsky, S. (2008) Functional analysis of *Arabidopsis* postprenylation caax processing enzymes and their function in subcellular protein targeting. *Plant Physiol.* 148: 119–131.

Bréhélin, C., Kessler, F. and van Wijk, K.J. (2007) Plastoglobules: versatile lipoprotein particles in plastids. *Trends Plant Sci.* 12: 260–266.

Caputi, L., Franke, J., Farrow, S.C., Chung, K., Payne, R.M.E., Nguyen, T.-D., et al. (2018) Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Science* 360: 1235–1239.

Carqueijeiro, I., Dugé de Bernonville, T., Lanoue, A., Dang, -T.-T., Teijaro, C.N., Paetz, C., et al. (2018a) A BAHD acyltransferase catalyzing 19-*O*-acetylation of tabersonine derivatives in roots of *Catharanthus roseus* enables combinatorial synthesis of monoterpene indole alkaloids. *Plant J.* 94: 469–484.

Carqueijeiro, I., Langley, C., Grzech, D., Koudounas, K., Papon, N., O'Connor, S.E., et al. (2020) Beyond the semi-synthetic artemisinin: metabolic engineering of plant-derived anti-cancer drugs. *Curr. Opin. Biotechnol.* 65: 17–24.

Carqueijeiro, I., Masini, E., Foureau, E., Sepúlveda, L.J., Marais, E., Lanoue, A., et al. (2015) Virus-induced gene silencing in *Catharanthus roseus* by biolistic inoculation of tobacco rattle virus vectors. *Plant Biol.* 17: 1242–1246.

Carqueijeiro, I., Sepúlveda, L.J., Mosquera, A., Payne, R., Corbin, C., Papon, N., et al. (2018b) Vacuole-targeted proteins: ins and outs of subcellular localization studies. *In* Plant Vacuolar Trafficking: Methods and Protocols. Edited by Pereira, C. pp. 33–54. Springer New York, New York.

Carrie, C., Murcha, M.W., Millar, A.H., Smith, S.M. and Whelan, J. (2007) Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in *Arabidopsis thaliana* are targeted either to peroxisomes or cytosol but not to mitochondria. *Plant Mol. Biol.* 63: 97–108.

Cázares-Flores, P., Levac, D. and De Luca, V. (2016) *Rauvolfia serpentina* *N*-methyltransferases involved in ajmaline and N_{β}-methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol *C*-methyltransferase. *Plant J.* 87: 335–342.

Choi, S.C., Lee, S., Kim, S.-R., Lee, Y.-S., Liu, C., Cao, X., et al. (2014) Trithorax group protein *Oryza sativa* Trithorax1 controls flowering time in rice via interaction with Early heading date3. *Plant Physiol.* 164: 1326–1337.

Corbin, C., Lafontaine, F., Sepúlveda, L.J., Carqueijeiro, I., Courtois, M., Lanoue, A., et al. (2017) Virus-induced gene silencing in *Rauwolfia species*. *Protoplasma* 254: 1813–1818.

Costa, M.M.R., Hilliou, F., Duarte, P.C., Pereira, L.S.G., Almeida, I., Leech, M., et al. (2008) Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol.* 146: 413–417.

Courdavault, V., O'Connor, S.E., Oudin, A., Besseau, S. and Papon, N. (2020) Towards the microbial production of plant-derived anticancer drugs. *Trends Cancer* 6: 444–448.

Courdavault, V., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B. and Burlat, V. (2014) A look inside an alkaloid multisite plant: the *Catharanthus* logistics. *Curr. Opin. Plant Biol.* 19: 43–50.

De Luca, V., Balsevich, J., Tyler, R.T. and Kurz, W.G.W. (1987) Characterization of a novel *N*-methyltransferase (NMT) from *Catharanthus roseus* plants. *Plant Cell Rep.* 6: 458–461.

De Luca, V. and Cutler, A.J. (1987) Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol.* 85: 1099–1102.

De Vega, J.J., Ayling, S., Hegarty, M., Kudrna, D., Goicoechea, J.L., Ergon, Å., et al. (2015) Red clover (*Trifolium pratense* L.) draft genome provides a platform for trait improvement. *Sci. Rep.* 5: 17394.

Dethier, M. and De Luca, V. (1993) Partial purification of an *N*-methyltransferase involved in vindoline biosynthesis in *Catharanthus roseus*. *Phytochemistry* 32: 673–678.

Dugé de Bernonville, T., Carqueijeiro, I., Lanoue, A., Lafontaine, F., Sánchez Bel, P., Liesecke, F., et al. (2017) Folivory elicits a strong defense reaction in *Catharanthus roseus*: metabolomic and transcriptomic analyses reveal distinct local and systemic responses. *Sci. Rep.* 7: 40453.

Dugé de Bernonville, T., Clastre, M., Besseau, S., Oudin, A., Burlat, V., Glévarec, G., et al. (2015a) Phytochemical genomics of the Madagascar periwinkle: unravelling the last twists of the alkaloid engine. *Phytochemistry* 113: 9–23.

Dugé de Bernonville, T., Foureau, E., Parage, C., Lanoue, A., Clastre, M., Londono, M.A., et al. (2015b) Characterization of a second secologanin synthase isoform producing both secologanin and secoxyloganin allows enhanced de novo assembly of a *Catharanthus roseus* transcriptome. *BMC Genom.* 16: 619.

Dusséaux, S., Wajn, W.T., Liu, Y., Ignea, C. and Kampranis, S.C. (2020) Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl. Acad. Sci.* 117: 31789–31799.

Endler, A., Meyer, S., Schelbert, S., Schneider, T., Weschke, W., Peters, S.W., et al. (2006) Identification of a vacuolar sucrose transporter in barley and *Arabidopsis* mesophyll cells by a tonoplast proteomic approach. *Plant Physiol.* 141: 196–207.

Eubel, H., Lee, C.P., Kuo, J., Meyer, E.H., Taylor, N.L. and Millar, A.H. (2007) Technical advance: free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J.* 52: 583–594.

Eubel, H., Meyer, E.H., Taylor, N.L., Bussell, J.D., O'Toole, N., Heazlewood, J.L., et al. (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol.* 148: 1809–1829.

Fornara, V., Onelli, E., Sparvoli, F., Rossoni, M., Aina, R., Marino, G., et al. (2008) Localization of stilbene synthase in *Vitis vinifera* L. During berry development. *Protoplasma* 233: 83–93.

Foureau, E., Carqueijeiro, I., Dugé de Bernonville, T., Melin, C., Lafontaine, F., Besseau, S., et al. (2016) Prequels to synthetic biology: from candidate gene identification and validation to enzyme subcellular localization in plant and yeast cells. *In* Methods in Enzymology. Edited by O'Connor, S.E. pp. 167–206. Academic Press, Cambridge, MA, USA.

Franke, J., Kim, J., Hamilton, J.P., Zhao, D., Pham, G.M., Wiegert-Rininger, K., et al. (2019) Gene discovery in *Gelsemium* highlights conserved gene clusters in monoterpene indole alkaloid biosynthesis. *ChemBioChem* 20: 83–87.

González, B. and Vera, P. (2019) Folate metabolism interferes with plant immunity through 1C methionine synthase-directed genome-wide DNA methylation enhancement. *Mol. Plant* 12: 1227–1242.

Guirimand, G., Burlat, V., Oudin, A., Lanoue, A., St-Pierre, B. and Courdavault, V. (2009) Optimization of the transient transformation of *Catharanthus roseus* cells by particle bombardment and its application to the subcellular localization of hydroxymethylbutenyl 4-diphosphate synthase and geraniol 10-hydroxylase. *Plant Cell Rep.* 28: 1215–1234.

Guirimand, G., Courdavault, V., Lanoue, A., Mahroug, S., Guihur, A., Blanc, N., et al. (2010a) Strictosidine activation in Apocynaceae: towards a "nuclear time bomb"? *BMC Plant Biol.* 10: 182.

Guirimand, G., Courdavault, V., St-Pierre, B. and Burlat, V. (2010b) Biosynthesis and regulation of alkaloids. *In* Plant Developmental Biology - Biotechnological Perspectives: Volume 2. Edited by Pua, E.C. and Davey, M.R. pp. 139–160. Springer Berlin Heidelberg, Berlin, Heidelberg.

Guirimand, G., Guihur, A., Ginis, O., Poutrain, P., Héricourt, F., Oudin, A., et al. (2011a) The subcellular organization of strictosidine biosynthesis in *Catharanthus roseus* epidermis highlights several trans-tonoplast translocations of intermediate metabolites. *FEBS J.* 278: 749–763.

Guirimand, G., Guihur, A., Poutrain, P., Héricourt, F., Mahroug, S., St-Pierre, B., et al. (2011b) Spatial organization of the vindoline biosynthetic pathway in *Catharanthus roseus*. *J. Plant Physiol.* 168: 549–557.

Guirimand, G., Guihur, A., Phillips, M.A., Oudin, A., Glévarec, G., Melin, C., et al. (2012a) A single gene encodes isopentenyl diphosphate isomerase isoforms targeted to plastids, mitochondria and peroxisomes in *Catharanthus roseus*. *Plant Mol. Biol.* 79: 443–459.

Guirimand, G., Simkin, A.J., Papon, N., Besseau, S., Burlat, V., St-Pierre, B., et al. (2012b) Cycloheximide as a tool to investigate protein import in peroxisomes: a case study of the subcellular localization of isoprenoid biosynthetic enzymes. *J. Plant Physiol.* 169: 825–829.

Guirimand, G., Guihur, A., Perello, C., Phillips, M., Mahroug, S., Oudin, A., et al. (2020) Cellular and subcellular compartmentation of the 2*C*-methyl-*D*-erythritol 4-phosphate pathway in the Madagascar periwinkle. *Plants* 9: 462.

Guirimand, G., Kulagina, N., Papon, N., Hasunuma, T. and Courdavault, V. (2021) Innovative tools and strategies for optimizing yeast cell factories. *Trends Biotechnol.* 39: 488–504.

Hagel, J.M. and Facchini, P.J. (2012) Subcellular localization of sanguinarine biosynthetic enzymes in cultured opium poppy cells. *Vitro Cell. Dev. Biol. Plant* 48: 233–240.

Héricourt, F., Chefdor, F., Bertheau, L., Tanigawa, M., Maeda, T., Guirimand, G., et al. (2013) Characterization of histidine-aspartate kinase HK1 and identification of histidine phosphotransfer proteins as potential partners in a *Populus* multistep phosphorelay. *Physiol. Plant.* 149: 188–199.

Hugueney, P., Provenzano, S., Verriès, C., Ferrandino, A., Meudec, E., Batelli, G., et al. (2009) A novel cation-dependent *O*-methyltransferase involved in anthocyanin methylation in grapevine. *Plant Physiol.* 150: 2057–2070.

Jacquier, M.J., Vercauteren, J., Massiot, G., Le Men-Olivier, L., Pussett, J. and Sevenet, T. (1980) Alkaloids of *Alstonia plumosa*. *Phytochemistry* 21: 2973–2977.

Jo, Y. and Hyun, T.K. (2011) Genome-wide identification of antioxidant component biosynthetic enzymes: comprehensive analysis of ascorbic acid and tocochromanols biosynthetic genes in rice. *Comput. Biol. Chem.* 35: 261–268.

Jossang, A., Fodor, P. and Bodo, B. (1998) A new structural class of bisindole alkaloids from the seeds of *Catharanthus roseus*: vingramine and methylvingramine. *J. Org. Chem.* 63: 7162–7167.

Jung, C., Lee, G.-J., Jang, M., Lee, M., Lee, J., Kang, H., et al. (2011) Identification of sorting motifs of AtβFruct4 for trafficking from the ER to the vacuole through the Golgi and PVC. *Traffic* 12: 1774–1792.

Kellner, F., Kim, J., Clavijo, B.J., Hamilton, J.P., Childs, K.L., Vaillancourt, B., et al. (2015) Genome-guided investigation of plant natural product biosynthesis. *Plant J.* 82: 680–692.

Koch, M., Lemke, R., Heise, K.-P. and Mock, H.-P. (2003) Characterization of γ-tocopherol methyltransferases from *Capsicum annuum* L and *Arabidopsis thaliana*. *Eur. J. Biochem.* 270: 84–92.

Kodama, Y., Shinya, T. and Sano, H. (2008) Dimerization of *N*-methyltransferases involved in caffeine biosynthesis. *Biochimie* 90: 547–551.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305: 567–580.

Kulagina, N., Besseau, S., Papon, N. and Courdavault, V. (2021) Peroxisomes: a new hub for metabolic engineering in yeast. *Front. Bioeng. Biotechnol.* 9: 659431.

Lee, C., Teng, Q., Zhong, R., Yuan, Y., Haghighat, M. and Ye, Z.-H. (2012) Three arabidopsis DUF579 domain-containing GXM proteins are methyltransferases catalyzing 4-*O*-methylation of glucuronic acid on xylan. *Plant Cell Physiol.* 53: 1934–1949.

Levac, D., Cázares, P., Yu, F. and De Luca, V. (2016) A picrinine *N*-methyltransferase belongs to a new family of γ-tocopherol-like methyltransferases found in medicinal plants that make biologically active monoterpenoid indole alkaloids. *Plant Physiol.* 170: 1935–1944.

Lichman, B.R., Godden, G.T. and Buell, C.R. (2020) Gene and genome duplications in the evolution of chemodiversity: perspectives from studies of Lamiaceae. *Curr. Opin. Plant Biol.* 55: 74–83.

Liscombe, D.K., Usera, A.R. and O'Connor, S.E. (2010) Homolog of tocopherol C methyltransferases catalyzes *N* methylation in anticancer alkaloid biosynthesis. *Proc. Natl. Acad. Sci.* 107: 18793–18798.

Madyastha, K.M., Ridgway, J.E., Dwyer, J.G. and Coscia, C.J. (1977) Subcellular localization of a cytochrome P-450-dependent monogenase in vesicles of the higher plant *Catharanthus roseus*. *J. Cell Biol.* 72: 302–313.

Magnani, R., Nayak, N.R., Mazarei, M., Dirk, L.M.A. and Houtz, R.L. (2007) Polypeptide substrate specificity of *Ps*LSMT: a set domain protein methyltransferase. *J. Biol. Chem.* 282: 27857–27864.

Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., et al. (2014) The seco-iridoid pathway from *Catharanthus roseus*. *Nat. Commun.* 5: 3606.

Mochida, K., Sakurai, T., Seki, H., Yoshida, T., Takahagi, K., Sawai, S., et al. (2017) Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *Plant J.* 89: 181–194.

Moore, R.C. and Purugganan, M.D. (2005) The evolutionary dynamics of plant duplicate genes. *Curr. Opin. Plant Biol.* 8: 122–128.

Morris, J.S. and Facchini, P.J. (2019) Molecular origins of functional diversity in benzylisoquinoline alkaloid methyltransferases. *Front. Plant Sci.* 10: 1058.

Muñoz, P. and Munné-Bosch, S. (2019) Vitamin E in plants: biosynthesis, transport, and function. *Trends Plant Sci.* 24: 1040–1051.

Naconsie, M., Kato, K., Shoji, T. and Hashimoto, T. (2014) Molecular evolution of *N*-methylputrescine oxidase in tobacco. *Plant Cell Physiol.* 55: 436–444.

Nakayama, T., Takahashi, S. and Waki, T. (2019) Formation of flavonoid metabolons: functional significance of protein-protein interactions and impact on flavonoid chemodiversity. *Front. Plant Sci.* 10: 821.

Nozoye, T., Nagasaka, S., Bashir, K., Takahashi, M., Kobayashi, T., Nakanishi, H., et al. (2014) Nicotianamine synthase 2 localizes to the vesicles of iron-deficient rice roots, and its mutation in the YXXφ or LL motif causes the disruption of vesicle formation or movement in rice. *Plant J.* 77: 246–260.

Nützmann, H.-W., Huang, A. and Osbourn, A. (2016) Plant metabolic clusters – from genetics to genomics. *New Phytolo.* 211: 771–789.

O'Connor, S.E. and Maresh, J.J. (2006) Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat. Prod. Rep.* 23: 532–547.

Pan, R., Reumann, S., Lisik, P., Tietz, S., Olsen, L.J. and Hu, J. (2018) Proteome analysis of peroxisomes from dark-treated senescent *Arabidopsis* leaves. *J. Integr. Plant Biol.* 60: 1028–1050.

Panchy, N., Lehti-Shiu, M. and Shiu, S.-H. (2016) Evolution of gene duplication in plants. *Plant Physiol.* 171: 2294–2316.

Parage, C., Foureau, E., Kellner, F., Burlat, V., Mahroug, S., Lanoue, A., et al. (2016) Class II cytochrome P450 reductase governs the biosynthesis of alkaloids. *Plant Physiol.* 172: 1563–1577.

Payne, R.M.E., Xu, D., Foureau, E., Teto Carqueijeiro, M.I.S., Oudin, A., Bernonville, T.D.D., et al. (2017) An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. *Nat. Plants* 3: 16208.

Pedrazzini, E., Komarova, N.Y., Rentsch, D. and Vitale, A. (2013) Traffic routes and signals for the tonoplast. *Traffic* 14: 622–628.

Pracharoenwattana, I., Cornah, J.E. and Smith, S.M. (2005) *Arabidopsis* peroxisomal citrate synthase is required for fatty acid respiration and seed germination. *Plant Cell* 17: 2037–2048.

Quan, S., Yang, P., Cassin-Ross, G., Kaur, N., Switzenberg, R., Aung, K., et al. (2013) Proteome analysis of peroxisomes from etiolated *Arabidopsis* seedlings identifies a peroxisomal protease involved in β-oxidation and development. *Plant Physiol.* 163: 1518–1538.

Rahman, A., Ali, I. and Chaudhry, M.I. (1985) Isolation and 13c-nmr studies on cathovaline, an alkaloid from the leaves of *Catharanthus roseus*. *Planta Med.* 51: 447–448.

Rahman, A., Ali, I. and Chaudhary, M.I. (1986) Bannucine – a new dihydroindole alkaloid from *Catharanthus roseus* (L.) G. Don. *J. Chem. Soc. Perkin Trans.* 1: 923–926.

Rai, A., Hirakawa, H., Nakabayashi, R., Kikuchi, S., Hayashi, K., Rai, M., et al. (2021) Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis. *Nat. Commun.* 12: 405.

Rasoanaivo, P., Langlois, N., Potier, P. and Bladon, P. (1973) Plantes malgaches x (1) - nouvel alcaloide du *Catharanthus longifolius* Pich. (Apocynacees). *Tetrahedron Lett.* 14: 1425–1428.

Ravanel, S., Block, M.A., Rippert, P., Jabrin, S., Curien, G., Rébeillé, F., et al. (2004) Methionine metabolism in plants: chloroplasts are autonomous for *de novo* methionine synthesis and can import *S*-adenosylmethionine from the cytosol. *J. Biol. Chem.* 279: 22548–22557.

Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G.E., et al. (2007) Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19: 3170–3193.

Reumann, S., Chowdhary, G. and Lingner, T. (2016) Characterization, prediction and evolution of plant peroxisomal targeting signals type 1 (PTS1s). *Biochim. Biophys. Acta Mol. Cell Res.* 1863: 790–803.

Rueffer, M., Amann, M. and Zenk, M.H. (1986) S-Adenosyl-L-methionine: columbamine-*O*-methyl transferase, a compartmentalized enzyme in protoberberine biosynthesis. *Plant Cell Rep.* 5: 182–185.

Saslowsky, D. and Winkel-Shirley, B. (2001) Localization of flavonoid enzymes in *Arabidopsis* roots. *Plant J.* 27: 37–48.

Segami, S., Makino, S., Miyake, A., Asaoka, M. and Maeshima, M. (2014) Dynamics of vacuoles and H^+-pyrophosphatase visualized by monomeric green fluorescent protein in *Arabidopsis*: artifactual bulbs and native intravacuolar spherical structures. *Plant Cell* 26: 3416–3434.

Shaipulah, N.F.M., Muhlemann, J.K., Woodworth, B.D., Van Moerkercke, A., Verdonk, J.C., Ramirez, A.A., et al. (2016) *CCoAOMT* down-regulation activates anthocyanin biosynthesis in petunia. *Plant Physiol.* 170: 717–731.

Shen, H., He, Z., Yan, H., Xing, Z., Chen, Y., Xu, W., et al. (2014) The fronds tonoplast quantitative proteomic analysis in arsenic hyperaccumulator *Pteris vittata* L. *J. Proteomics* 105: 46–57.

Shen, J., Suen, P.K., Wang, X., Lin, Y., Lo, S.W., Rojo, E., et al. (2013) An *in vivo* expression system for the identification of cargo proteins of vacuolar sorting receptors in *Arabidopsis* culture cells. *Plant J.* 75: 1003–1017.

Shintani, D. and DellaPenna, D. (1998) Elevating the vitamin E content of plants through metabolic engineering. *Science* 282: 2098–2100.

Simkin, A.J., Guirimand, G., Papon, N., Courdavault, V., Thabet, I., Ginis, O., et al. (2011) Peroxisomal localisation of the final steps of the mevalonic acid pathway. *In Planta. Planta* 234: 903.

Simkin, A.J., Miettinen, K., Claudel, P., Burlat, V., Guirimand, G., Courdavault, V., et al. (2013) Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. *Phytochemistry* 85: 36–43.

Soll, J., Schultz, G., Joyard, J., Douce, R. and Block, M.A. (1985) Localization and synthesis of prenylquinones in isolated outer and inner envelope

membranes from spinach chloroplasts. *Arch. Biochem. Biophys.* 238: 290–299.

Stavrinides, A., Tatsis, E.C., Caputi, L., Foureau, E., Stevenson, C.E.M., Lawson, D.M., et al. (2016) Structural investigation of heteroyohimbine alkaloid synthesis reveals active site elements that control stereoselectivity. *Nat. Commun.* 7: 12116.

Stavrinides, A., Tatsis, E.C., Foureau, E., Caputi, L., Kellner, F., Courdavault, V., et al. (2015) Unlocking the diversity of alkaloids in *Catharanthus roseus*: nuclear localization suggests metabolic channeling in secondary metabolism. *Chem. Biol.* 22: 336–341.

St-Pierre, B., Besseau, S., Clastre, M., Courdavault, V., Courtois, M., Crèche, J., et al. (2013) Deciphering the evolution, cell biology and regulation of monoterpene indole alkaloids. *In* Advances in Botanical Research. Edited by Giglioli-Guivarc'h, N. pp. 73–109. Academic Press, Burlington.

Tamiru, M., Takagi, H., Abe, A., Yokota, T., Kanzaki, H., Okamoto, H., et al. (2016) A chloroplast-localized protein lesion and lamina bending affects defence and growth responses in rice. *New Phytolo.* 210: 1282–1297.

Tanabe, Y., Maruyama, J.-I., Yamaoka, S., Yahagi, D., Matsuo, I., Tsutsumi, N., et al. (2011) Peroxisomes are involved in biotin biosynthesis in *Aspergillus* and *Arabidopsis*. *J. Biol. Chem.* 286: 30455–30461.

Tang, R.-J., Liu, H., Yang, Y., Yang, L., Gao, X.-S., Garcia, V.J., et al. (2012) Tonoplast calcium sensors CBL2 and CBL3 control plant growth and ion homeostasis through regulating V-ATPase activity in *Arabidopsis*. *Cell Res.* 22: 1650–1665.

Teng, J., Yan, C., Zeng, W., Zhang, Y., Zeng, Z. and Huang, Y. (2020) Purification and characterization of theobromine synthase in a theobromine-enriched wild tea plant (*Camellia gymnogyna* Chang) from Dayao Mountain, China. *Food Chem.* 311: 125875.

Thabet, I., Guirimand, G., Courdavault, V., Papon, N., Godet, S., Dutilleul, C., et al. (2011) The subcellular localization of periwinkle farnesyl diphosphate synthase provides insight into the role of peroxisome in isoprenoid biosynthesis. *J. Plant Physiol.* 168: 2110–2116.

Thabet, I., Guirimand, G., Guihur, A., Lanoue, A., Courdavault, V., Papon, N., et al. (2012) Characterization and subcellular localization of geranylgeranyl diphosphate synthase from *Catharanthus roseus*. *Mol. Biol. Rep.* 39: 3235–3243.

Toda, K., Kuroiwa, H., Senthil, K., Shimada, N., Aoki, T., Ayabe, S.-I., et al. (2012) The soybean F3′H protein is localized to the tonoplast in the seed coat hilum. *Planta* 236: 79–89.

Wojcik, S. and Kriechbaumer, V. (2021) Go your own way: membrane-targeting sequences. *Plant Physiol.* 185: 608–618.

Xie, Y., Zheng, Y., Li, H., Luo, X., He, Z., Cao, S., et al. (2016) GPS-Lipid: a robust tool for the prediction of multiple lipid modification sites. *Sci. Rep.* 6: 28249.

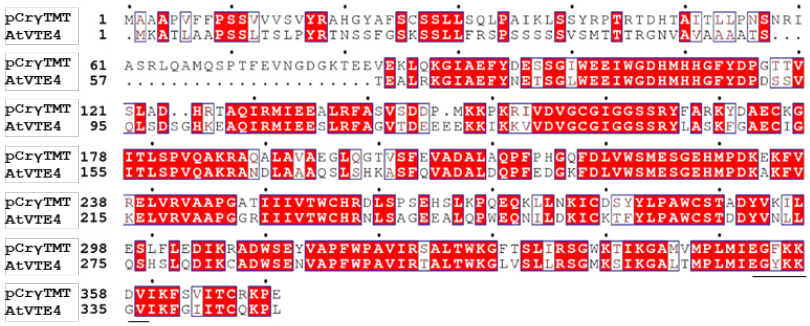
Yin, X., Yang, J., Xiao, F., Yang, Y. and Shen, H.-B. (2017) MemBrain: an easy-to-use online webserver for transmembrane protein structure prediction. *Nano-Micro Lett.* 10: 2.

Zbierzak, A.M., Kanwischer, M., Wille, C., Vidi, P.-A., Giavalisco, P., Lohmann, A., et al. (2009) Intersection of the tocopherol and plastoquinol metabolic pathways at the plastoglobule. *Biochem. J.* 425: 389–399.

Zhao, D., Hamilton, J.P., Pham, G.M., Crisovan, E., Wiegert-Rininger, K., Vaillancourt, B., et al. (2017) *De novo* genome assembly of *Camptotheca acuminata*, a natural source of the anti-cancer compound camptothecin. *GigaScience* 6: 1–7.

Zhao, Q., Yang, J., Cui, M.-Y., Liu, J., Fang, Y., Yan, M., et al. (2019) The reference genome sequence of *Scutellaria baicalensis* provides insights into the evolution of wogonin biosynthesis. *Mol Plant* 12: 935–950.

A



B

	pCrγTMT	AtVTE4
pCrγTMT	371 0 0	59% 71% 7%
AtVTE4	224 268 29	348 0 0

C

	pCrγTMT_truncated	AtVTE4_truncated
pCrγTMT_truncated	289 0 0	71% 83% 1%
AtVTE4_truncated	208 245 3	292 0 0

Figure S1. (**A**) Alignment of the amino acid sequences of pCrγTMT and AtVTE4. Identical amino acids are highlighted in red. Red letters indicate identity and blue boxes show similarity. The Rubisco large subunit methyltransferase motif regulating γTMT activity (Magnani et al., 2007) is underlined. Graphical representation of the alignment was generated with the ESPript server (https://espript.ibcp.fr). (**B**) The alignment matrix of (**A**). (**C**) The alignment matrix of the two truncated proteins after removing the first 82 or 56 amino acids of pCrγTMT or AtVTE4 respectively. The first number represents the identical residues, the second number represents the similar residues or conservative substitutions, and the third number represents the residues lined up with a gap character. Each number is expressed as a percentage and as a count on different sides of the matrix diagonal.

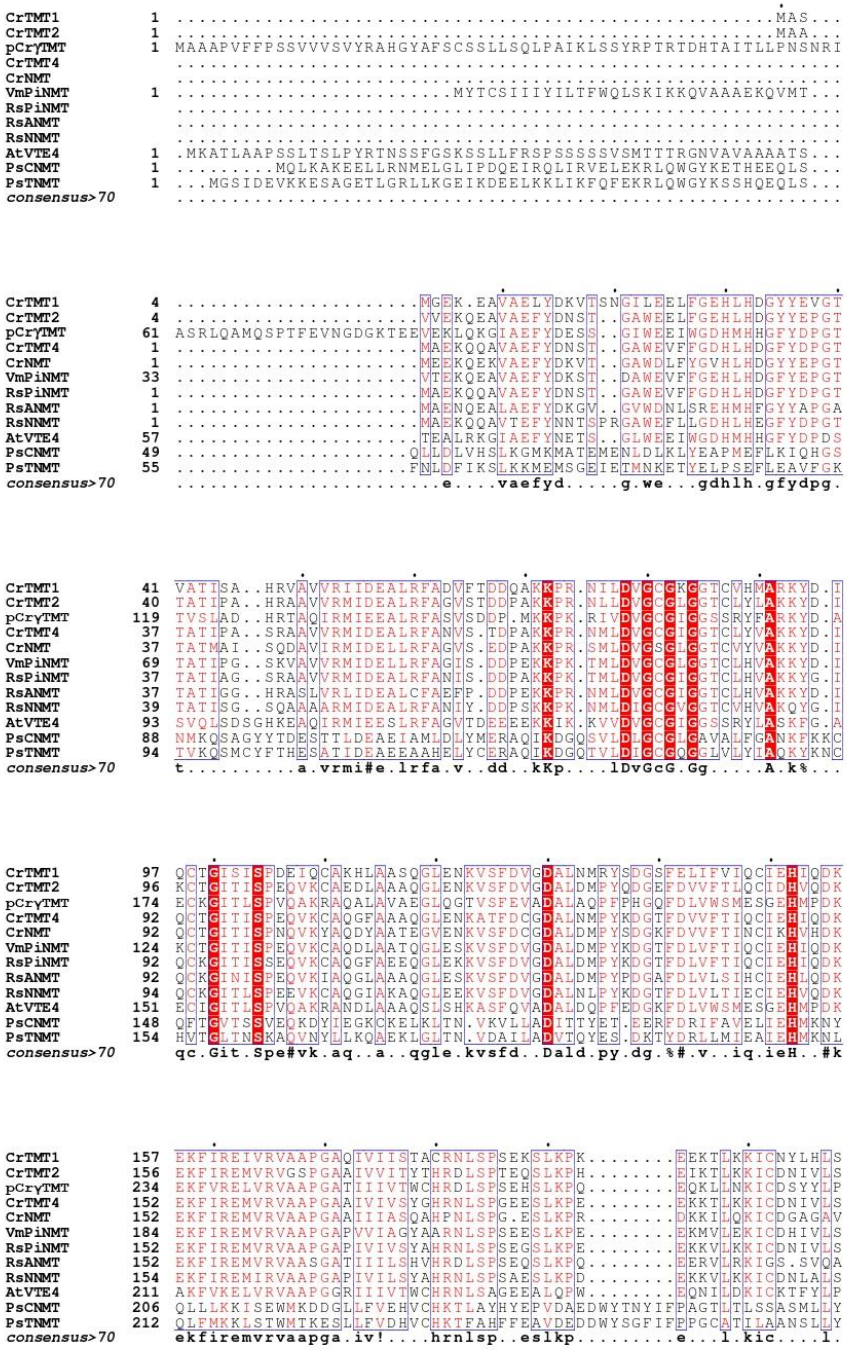


Figure S2. Alignment of the amino acid sequences of the respective methyltransferases used in this study. Identical amino acids are highlighted in red. Red letters indicate homology and blue boxes show similarity. The overall sequence consensus (bottom line) represents identity (uppercase) or consensus level > 0.7 (lowercase). Residues I or V are represented as "!", F or Y as "%" and N, D, Q or E as "#". Graphical representation of the alignment was generated with the ESPript server (https://espript.ibcp.fr).

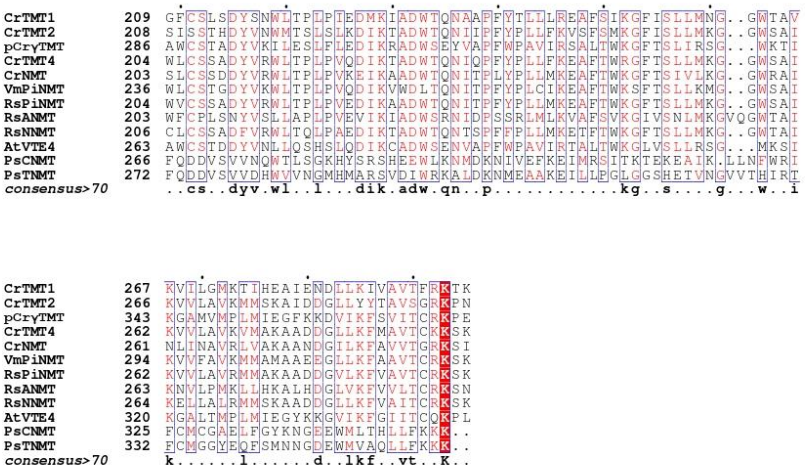


Figure S2. (continued).

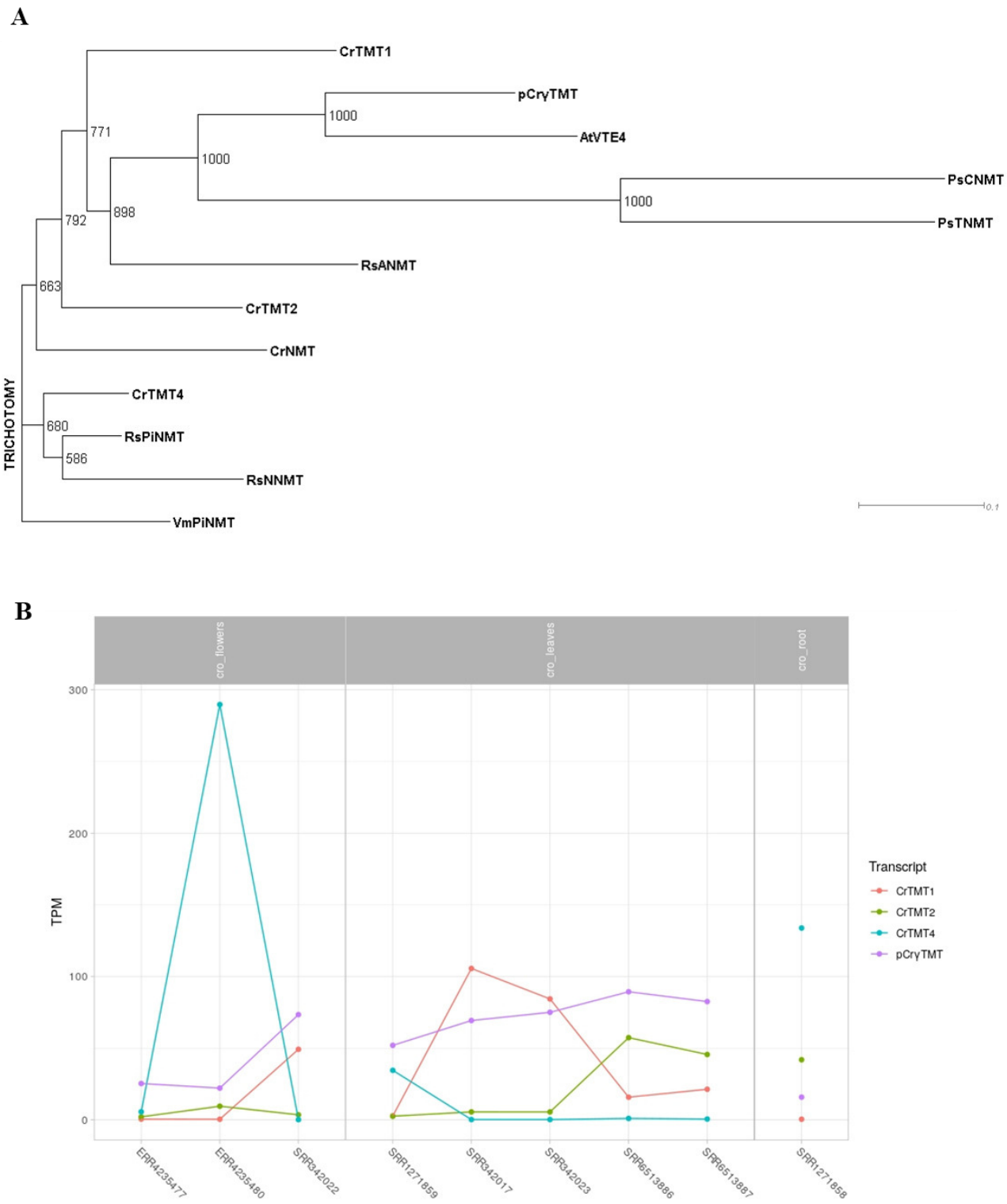


Figure S3. (A) Phylogenetic tree of the respective methyltransferases used in this study. Numbers on the internal branches represent the bootstrap reproducibility values (1000 permutations). Bar, 0.1 amino acid substitutions per site. Graphical representation of the phylogenetic tree was generated with Dendroscope 3 (http://ab.inf.uni-tuebingen.de/software). (B) Expression of the four identified TMTs. The relative expression of CrTMTs was quantified as transcripts per million (TPM) after mapping RNA-seq reads from *C. roseus* leaves (SRR342017, SRR342022, SRR342023, SRR6513886, SRR6513887), flowers (SRR1271859, ERR4235480, ERR4235477) and root (SRR1271858) samples to *C. roseus* V2 transcript sequences (cro_v2.transcripts.fasta).

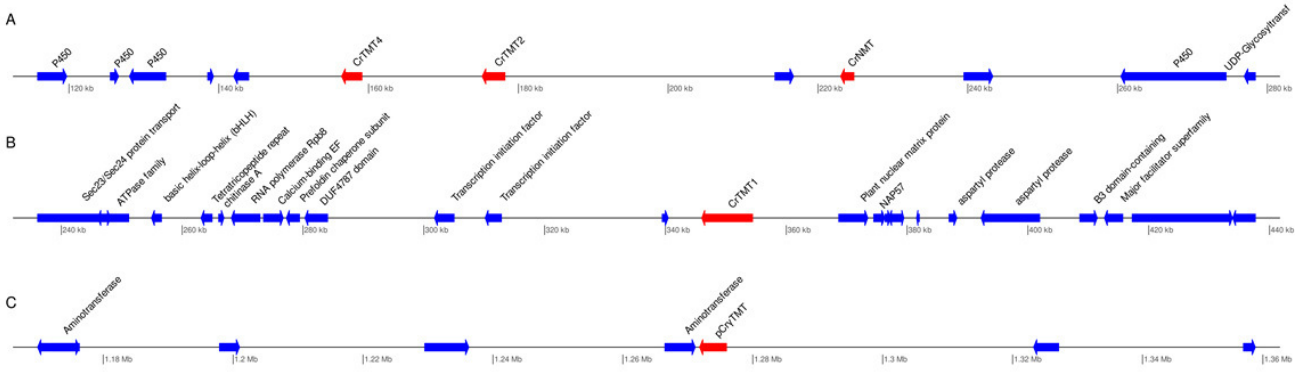


Figure S4. Genomic organization of CrTMT1, CrTMT2, pCrγTMT, CrTMT4 and CrNMT in the genomic scaffolds of *C. roseus*. (A) cro_v2_scaffold_163, (B) cro_v2_scaffold_186 and (C) cro_v2_scaffold_51.

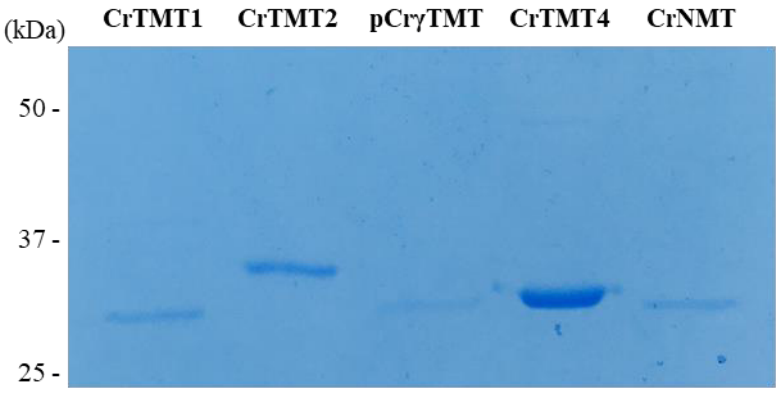


Figure S5. CBB staining of purified proteins (500ng) heterologously expressed in *E. coli*. Numbers on the left indicate molecular weights.

Primer	Sequence (5' – 3')
AtTMT_F	CTGAGAACTAGTATGAAAGCAACTCTAGCAGCACC
AtTMT_R	CTGAGAACTAGTGAGTGGCTTCTGGCAAGTGATGAT
TMT1_F	CTGAGAACTAGTATGGCCTCAATGGGAGAGAAG
TMT1_Rstop	CTGAGAACTAGTTCATTTAGTTTTGCGAAATGTAACTGC
TMT1_R	CTGAGAACTAGTTTTAGTTTTGCGAAATGTAACTGCAAC
TMT2_F	CTGAGAACTAGTATGGCTGCCGTGGTAGAGAAGCA
TMT2_Rstop	CTGAGAACTAGTTCAATTAGGTTTCCTGCCTGAAACCG
TMT2_R	CTGAGAACTAGTATTAGGTTTCCTGCCTGAAACCGCA
pCRγTMT_F	CTGAGAACTAGTATGGCCGCTGCTCCTGTTTTTTTTC
pCRγTMT_R	CTGAGAACTAGTCTCAGGTTTTCGGCATGTAATGACAG
pCRγTMT_Rstop	CTGAGAACTAGT TTACTCAGGTTTTCGGCATGTAATGACAG
pCRγTMT_F2	CTGAGAACTAGTATGGCCGCTGCTCCTGTTTTTTTTC
pCRγTMT_R2	CTGAGAACTAGTCTCAGGTTTTCGGCATGTAATGACAGAAA
TMT4_F	CTGAGAACTAGTATGGCAGAGAAGCAACAAGCAGTT
TMT4_Rstop	CTGAGAACTAGTTTATTTTGATTTCTTGCACGTAACAGCC
TMT4_R	CTGAGAACTAGTTTTTGATTTCTTGCACGTAACAGCCATA
PsTNMT_F	CTGAGAACTAGTATGGGTTCAATAGATG
PsTNMT_Rstop	CTGAGAACTAGTCTACTTCTTCTTGAAAAG
PsTNMT_R	CTGAGAACTAGTCTTCTTCTTGAAAAG
PsCNMT_F	CTGAGAACTAGTATGCAGTTAATGGCAAAGG
PsCNMT_Rstop	CTGAGAACTAGTTCATTTTTTCTTGAAGAGAAGATGG
PsCNMT_R	CTGAGAACTAGTTTTTTTCTTGAAGAGAAGATGG
VmPiNMT_F	CTGAGAACTAGTATGTACACTTGTTCAATTATAATATATATTTTAACG
VmPiNMT_R	CTGAGAACTAGTTTTAGATTTGCGGCATGTAACTG
VmPiNMT_Rstop	CTGAGAACTAGTTTATTTAGATTTGCGGCATGTAACT
VmPiNMT_R2	CTGAGAACTAGTGACTTGCTTTTCCGCCGCTGC
VmPiNMT_R3	CTGAGAACTAGTAACTACGGCGACTTTGCTACCA
RsPiNMT_F	CTGAGAACTAGTATGGCAGAGAAGCAGCAGGCC
RsPiNMT_R	CTGAGAACTAGTTTTGGATTTCCTGCATGTAACTGCAACA
RsPiNMT_Rstop	CTGAGAACTAGTTTATTTGGATTTCCTGCATGTAACTGCAACA
RsNNMT_F	CTGAGAACTAGTATGGCAGAGAAGCAGCAGGCAG
RsNNMT_R	CTGAGAACTAGTTTTTGATTTCCTGCATGTAATTGCAACG
RsNNMT_Rstop	CTGAGAACTAGTTTATTTTGATTTCCTGCATGTAATTGCAACG
RsANMT_F	CTGAGAACTAGTATGGCAGAGAACCAGGAGGCAC
RsANMT_R	CTGAGAACTAGTATTAGATTTCCGGCAAGTCAGTACAACA
RsANMT_Rstop	CTGAGAACTAGTTTAATTAGATTTCCGGCAAGTCAGTACAACA
NMT-GFP-for2	CTGAGAACTAGTATGGAAGAGAAGCAGGAGAAGGTG
NMT-GFP-rev	GCACTAGTTATTGATTTTCGTCCCGTAACTACAGC
NMT-GFP-revstop	CTGAGAACTAGTTCATATTGATTTTCGTCCCGTAACTACAG
NMT-GFP-mut	CTGAGAACTAGTTTATCGTCCCGTAACTACAGCGAACTTAAG
TMT1_Fec	CTGAGAGGATCCATGGCCTCAATGGGAGAGAAG
TMT1_Rec	CTGAGAGGATCCTCATTTAGTTTTGCGAAATGTAACTGC
TMT2_Fec	CTGAGAGGATCCATGGCTGCCGTGGTAGAGAAGCAG
pCRγTMT_Fec	CTGAGAGGATCCATGCAGAGTCCGACGTTCGAG
pCRγTMT_Rec	CTGAGAGGATCCTTACTCAGGTTTTCGGCATGTAATGAC

TMT4_Fec	CTGAGAGGATCCATGGCAGAGAAGCAACAAGCAGTTGCT
TMT4_Rec	CTGAGAGGATCCTTATTTTGATTTCTTGCACGTAACAGCCA
NMT_Fec	CTGAGAGGATCCATGGAAGAGAAGCAGGAGAAGGTG
NMT_Rec	CTGAGAGGATCCTCATATTGATTTTCGTCCCGTAACTACA
Subcellular localization studies	
Construct	**Combination of primers**
VTE4-YFP	AtTMT_F & AtTMT_R
TMT1-YFP	TMT1_F & TMT1_R
YFP-TMT1	TMT1_F & TMT1_Rstop
TMT2-YFP	TMT2_F & TMT2_R
YFP-TMT2	TMT2_F & TMT2_Rstop
pCRγTMT-YFP (TMT3-YFP)	pCRγTMT_F & pCRγTMT_R
YFP-pCrγTMT (YFP-TMT3)	pCRγTMT_F & pCRγTMT_R
tp pCRγTMT-YFP	pCRγTMT_F2 & pCRγTMT_R2
TMT4-YFP	TMT4_F & TMT4_R
YFP-TMT4	TMT4_F & TMT4_Rstop
PsTNMT-YFP	PsTNMT_F & PsTNMT_R
YFP-PsTNMT	PsTNMT_F & PsTNMT_Rstop
PsCNMT-YFP	PsCNMT_F & PsCNMT_R
YFP-PsCNMT	PsCNMT_F & PsCNMT_Rstop
VmPiNMT-YFP	VmPiNMT_F & VmPiNMT_R
YFP-VmPiNMT	VmPiNMT_F & VmPiNMT_Rstop
spVmPiNMT-YFP	VmPiNMT_F & VmPiNMT_R2
sp2VmPiNMT-YFP	VmPiNMT_F & VmPiNMT_R3
RsPiNMT-YFP	RsPiNMT_F & RsPiNMT_R
YFP-RsPiNMT	RsPiNMT_F & RsPiNMT_Rstop
RsNNMT-YFP	RsNNMT_F & RsNNMT_R
YFP-RsNNMT	RsNNMT_F & RsNNMT_Rstop
RsANMT-YFP	RsANMT_F & RsANMT_R
YFP- RsANMT	RsANMT_F & RsANMT_Rstop
CrNMT-YFP	NMT-GFP-for2 & NMT-GFP-rev
YFP- CrNMT	NMT-GFP-for2 & NMT-GFP-revstop
YFP-CrNMTΔ	NMT-GFP-for2 & NMT-GFP-mut
Expression in *E. coli*	
CrTMT1	TMT1_Fec & TMT1_Rec
CrTMT2	TMT2_Fec & TMT2_Rec
pCRγTMT	pCRγTMT_Fec & pCRγTMT_Rec
CrTMT4	TMT4_Fec & TMT4_Rec
CrNMT	NMT_Fec & NMT_Rec

Table S1. Primers used in this study.

Protein	Prediction	
	TMHMM	**MemBrain** (AAs position)
CrTMT1	No predicted TMHs	46-55 TMH (half) 105-116 TMH 273-278 TMH (half)
CrTMT2	No predicted TMHs	107-114 TMH (half) 194-201 AH
pCrγTMT	No predicted TMHs	10-21 TMH 122-129 TMH (half) 269-280 AH 343-356 TMH
CrTMT4	No predicted TMHs	42-50 TMH (half) 82-87 TMH (half) 101-106 TMH (half) 189-198 AH 260-273 TMH
RsPiNMT	No predicted TMHs	40-49 TMH (half) 102-109 TMH (half)
CrNMT	No predicted TMHs	40-46 TMH (half) 189-196 AH 258-272 TMH (half)
RsNNMT	No predicted TMHs	190-198 AH 261-275 TMH
RsANMT	No predicted TMHs	101-111 TMH 257-275 TMH
PsCNMT	No predicted TMHs	21-29 AH 104-110 TMH (half) 135-142 TMH (half) 315-333 TMH
PsTNMT	No predicted TMHs	9-20 TMH 50-63 AH 141-148 TMH (half) 325-337 TMH
VmPiNMT	No predicted TMHs	2-17 TMH 77-84 TMH (half) 134-139 TMH (half) 221-228 AH 290-303 TMH
AtVTE4	No predicted TMHs	160-167 TMH (half) 320-333 TMH

Table S2. Prediction of transmembrane helices in the respective methyltransferases used in this study. TMH: Transmembrane Helix; AH: Amphipathic Helix

Annexe IV: Article "A Rapid and Efficient Vacuum-Based Agroinfiltration Protocol for Transient Gene Overexpression in Leaves of *Catharanthus roseus*"

par Koudounas K, Carqueijeiro I, Lemos Cruz P, Perrin J, Lanoue A, Oudin A, Besseau S, Courdavault V, publié en 2022 dans un volume de la série **Methods in Molecular Biology**.

La recherche de gènes candidats impliqués dans la synthèse de certains MS peut être entravée chez les plantes non modèles et peu réceptives à la transformation génétique telle que *Catharanthus roseus*. Des outils de répression transitoire génique ont été développés pour cette plante, telle que la technique VIGS, mais les outils permettant la surexpression de gènes sont rares. C'est pourquoi, dans ces travaux, nous avons développé une méthode efficace de surexpression transitoire dans les feuilles de *C. roseus*. Cette approche combine une méthode d'agro-infiltration sous vide d'*Agrobacterium tumefaciens* transformées avec un vecteur viral amélioré. En effet, le clonage des transgènes ou cassettes d'intérêt dans les vecteurs binaires pEAQ-HT composés de gènes viraux déconstruits permet de stabiliser l'ARNm produit, et donc d'améliorer la production des protéines recombinantes. La méthode décrite dans cet article est robuste, simple à mettre en œuvre et permettra d'améliorer la caractérisation de la fonction des gènes *in planta*, voire même d'introduire *in planta* de nouvelles activités enzymatiques. J'ai ainsi directement participé à la mise au point de ce protocole de transformation que j'ai par la suite exploité pour la surexpression de facteurs de transcription.



Chapter 19

A Rapid and Efficient Vacuum-Based Agroinfiltration Protocol for Transient Gene Overexpression in Leaves of *Catharanthus roseus*

Konstantinos Koudounas, Ines Carqueijeiro, Pamela Lemos Cruz, Jennifer Perrin, Arnaud Lanoue, Audrey Oudin, Sébastien Besseau, and Vincent Courdavault

Abstract

Functional genomics analyses in planta can be hampered in non-model plants that are recalcitrant to the genetic transformation such as the medicinal plant *Catharanthus roseus* (Apocynaceae). No stable transformation and regeneration of plantlets have been achieved with a high efficiency in this plant to date. In addition, while virus-mediated transient gene silencing has been reported a decade ago in *C. roseus*, tools for transient overexpression remain scarce. Here, we describe an efficient and reliable methodology for transiently overexpressing any gene of interest in *C. roseus* leaves. This protocol combines a vacuum-based Agroinfiltration approach and the high translational efficiency of a deconstructed virus-based binary vector (pEAQ-HT). The described methodology is robust, easy to perform, and results in high amount of transient expression in *C. roseus*. This protocol is expected to serve as valuable tool to enhance the in planta characterization of gene functions or even transiently knock-in novel enzymatic activities.

Key words *Catharanthus roseus*, Madagascar periwinkle, Apocynaceae, *Agrobacterium tumefaciens*, Agroinfiltration, Vacuum infiltration, Transient expression, Functional genomics, Cowpea mosaic virus HT (CPMV-HT), Recombinant proteins

1 Introduction

The Madagascar periwinkle (*Catharanthus roseus*; Apocynaceae) is an emblematic medicinal plant biosynthesizing a plethora of bioactive monoterpene indole alkaloids (MIAs). While these MIAs serve as defense compounds in planta, some of them are also of importance for human pharmacopoeia such as the anticancer drugs vinblastine and vincristine [1, 2]. Over years, researchers from several laboratories have continuously contributed to the elucidation of MIA biosynthetic pathways and to the discovery of the different enzymes composing the MIA biosynthetic machinery. As a

Vincent Courdavault and Sébastien Besseau (eds.), *Catharanthus roseus: Methods and Protocols*,
Methods in Molecular Biology, vol. 2505, https://doi.org/10.1007/978-1-0716-2349-7_19,
© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2022

non-model plant, functional genomic approaches remained scarce until recently. The characterization of candidate genes/enzymes was thus mostly based on expression in heterologous hosts and was limited to in vitro assays, providing limited knowledge about the actual in planta role.

Over the last decade, a tremendous boost in functional genomics studies was achieved in *C. roseus* through the transient silencing of candidate genes (usually named Gene of Interest, GOI) via a virus-induced gene silencing approach (VIGS). Such a transient silencing was obtained after delivering genetic constructs encompassing the bipartite RNA genome of Tobacco Rattle Virus (TRV) in planta either by *Agrobacterium*-mediated inoculation [3] or by biolistic transformation [4]. VIGS thus enables the in planta validation of gene functions, up to the elucidation of the biosynthetic pathway of the anti-cancer drugs vinblastine and vincristine [5]. Ideally, besides gene silencing, gene function should be also validated and characterized by GOI overexpression in planta. However, stable transformation and overexpression of a GOI are hampered in *C. roseus* since this species is recalcitrant to genetic transformation and regeneration protocols [6]. As a consequence, most of the genetic transformation approaches developed so far in *C. roseus* have been performed in hairy roots or cell suspensions [7] while these protocols are either heavily time-consuming or need experienced personnel.

Only a few successful transient transformations have been reported in *C. roseus* notably in seedlings [8, 9], petals [10] but remained limited in leaves [11, 12]. The recent advances in plant virology and the identification of specific sequences enhancing the replication and/or expression within the plant cells have resulted in the development of novel deconstructed virus-based vectors that enable transient overexpression at levels overtaking stable transformation approaches [13]. One prominent example is the pEAQ-HT vector which encompasses the nonreplicating Cowpea Mosaic Virus-based *HyperTrans* (CPMV-HT) system comprised of a modified 5′-untranslated region (UTR) and the 3′-UTR from CPMV RNA-2 under the control of the Cauliflower mosaic virus (CaMV) 35S promoter and the nopaline synthase (nos) terminator (Fig. 1). Additionally, an expression cassette for the P19 suppressor of gene silencing has been introduced within the T-DNA borders resulting in high levels of transient expression [14] allowing the plasmid to be used in several studies [15]. It is worth noting that an autonomously replicating viral expression system based on *Catharanthus* yellow mosaic virus (CaYMV) has been lately reported [16].

We have recently developed an optimized vacuum-based Agroinfiltration protocol for *C. roseus* plants resulting in high levels of transient expression in aerial parts of the plant by recruiting the pEAQ-HT vector (Fig. 1). This methodology has already been successfully applied in order to overexpress candidate genes in

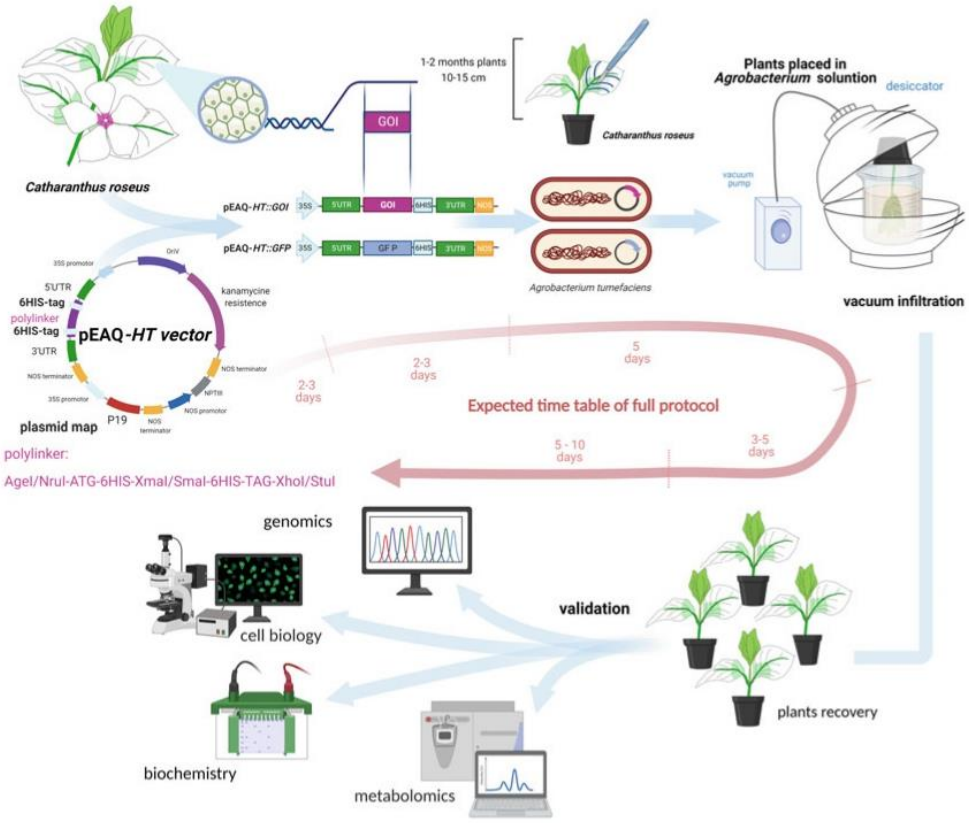


Fig. 1 Schematic overview of the vacuum-based Agroinfiltration protocol in *C. roseus* plants. A GOI is introduced in the polylinker site of the pEAQ-HT vector encompassing the modified 5′ and 3′ UTRs of CPMV resulting in increased translation of GOI. The plasmid is transformed in *A. tumefaciens* and delivered to *C. roseus* plants by vacuum-infiltration. After recovery, samples can be collected and used for microscopic, metabolics, genomics, or biochemical analyses

C. roseus [17] and is expected to boost functional genomics studies complementing the VIGS-based approach. Additionally, this approach sets the bases to initiate in planta metabolic engineering of *C. roseus* and explores novel chemotypes. Through transient expression of enzymes from distinct species and using the endogenous pool of MIA biosynthetic intermediates as leverage, production of tailor-made or/and new-to-nature MIAs can be investigated. In this chapter, we described in a detailed manner the critical steps of our transient transformation protocol to allow the *C. roseus* community to perform a controlled transient overexpression of genes in planta.

2 Materials

2.1 Vectors Construction

1. Sterilized tips, 0.2 and 1.5 mL tubes.
2. Agarose gel apparatus (gel tank, gel tray, well combs, and casting tray) and relative equipment (UV transilluminator).
3. 6× Gel loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, 40% sucrose.
4. Tris-acetate buffer (TAE): TAE can be prepared as a 50× stock solution by mixing 242 gr Tris-base, 57.1 mL glacial acetic acid, and 100 ml of 0.5 M EDTA (pH 8.0) in one Liter of ddH_2O.
5. Powdered agarose.
6. 10 mg/mL Ethidium bromide (stored at room temperature and protected from light). Working solution is 0.5 μg/mL.
7. Restriction enzymes (depending on the selected cloning strategy) and the corresponding buffers (*see* Subheading 3.1).
8. Alkaline phosphatase.
9. Heat blocks or water baths.
10. pEAQ-HT vector (GenBank Acc. nr.: GQ497234) [14].
11. pEAQ-HT:GFP-6HIS vector.
12. Kit for RNA extraction.
13. Spectrophotometer.
14. DNase I treatment kit and RNase inhibitor.
15. cDNA synthesis kit including oligo$(dT)_{17}$.
16. Proofreading DNA polymerase and buffer.
17. 10 mM dNTPs set.
18. GOI-specific primers including desired restriction sites (depending on the selected cloning strategy, *see* Subheading 3.1).
19. PCR machine.
20. T4 DNA ligase.
21. PCR purification and miniprep kits.
22. Chemically competent *Escherichia coli* (Top10) cells.
23. Luria-Bertani broth medium: 1% tryptone, 0.5% yeast extract and 1% NaCl. In case of solid medium, agar is added in a final concentration of 1.4% before autoclaving.
24. 50 g/L Kanamycin solution prepared in ddH_2O and stored at −20 °C (1000× stock solution).
25. Orbital shakers.

26. Centrifuges.
27. Incubators at 37 °C.

2.2 Handling of Agrobacterium Cells

1. *Agrobacterium tumefaciens* strain C58C1 Rif^R (GV3101) containing the T-DNA deficient Ti plasmid pMP90.
2. LB medium.
3. Antibiotic stock solutions: 5 g/L rifampicin resuspended in methanol (100×), 50 g/L gentamicin resuspended in ddH_2O (1000×), 50 g/L kanamycin resuspended in ddH_2O (1000×). All antibiotic stocks are stored at −20 °C. Working solution for each antibiotic is 50 mg/L.
4. Electroporator and electroporation cuvettes of a 0.1 cm gap.
5. Refrigerated centrifuge, centrifugation bottles, and tubes.
6. Erlenmeyer flasks, shakers, and incubators at 28 °C.
7. Dilution buffer: prepare stock solutions of 100 mM 2-N-morpholino-ethanesulfonic acid (MES) pH 5.6 and 100 mM $MgCl_2$. Prepare fresh dilution buffer (10 mM MES, pH 5.6, 10 mM $MgCl_2$) prior Agroinfiltration.
8. 100 mM acetosyringone stock solution in dimethyl sulfoxide (DMSO).
9. Sterile glycerol (100%).

2.3 Preparation of Catharanthus roseus Plants

1. *Catharanthus roseus* seeds (cv. Little Bright Eyes).
2. Soil composed of a 4:1 volume ratio of loam and vermiculite.
3. Greenhouse under controlled conditions (28 °C, 50% relative humidity, 16/8 h light/dark cycle, and 70 μmol m^{-2} s^{-1} light intensity).
4. Small trays for seedling cultivation.
5. Square pots (9 cm).
6. Boxes to contain small trays that can be covered with transparent plastic domes to achieve high relative humidity.
7. Razor blades or scalpels.
8. Piece of cardboard or plastic wrap to hold the soil in the pot.

2.4 Vacuum Equipment

1. Vacuum pumping unit.
2. Vacuum desiccator unit of 25 cm inner diameter.
3. Tall-form beakers of 150 mL volume capacity.

2.5 Macro/ Microscopic Analysis

1. UV lamp for macroscopic observation of GFP.
2. Epifluorescence microscope equipped with a proper light source (mercury short arc HBO or LED) and a specific filter set for GFP fluorescence observation (e.g., 460–480 nm excitation filter, 500–520 nm band-pass emission filter). The microscope should be equipped with a digital camera and a proper imaging software.
3. Coverslips (24 × 50 mm) and microscope slides (76 × 26 mm).

2.6 Extraction of Proteins and Western Blot

1. Liquid nitrogen.
2. Mortar and pestle.
3. Protein extraction buffer (100 mM Tris–HCl, pH 8.8, 100 mM EDTA, 1 mM PMSF, 100 mM KCL, 100 mM glycine).
4. Heat block.
5. 30% Acrylamide/Bis-acrylamide solution (29:1).
6. 1.5 M Tris–HCl, pH 8.8.
7. 0.5 M Tris–HCl, pH 6.8.
8. 10% sodium dodecyl sulfate (SDS).
9. 10% ammonium persulfate (APS). Prepare in ddH_2O and store aliquots of 200 μL in −20 °C. Do not thaw and refreeze an aliquot.
10. *N*,*N*,*N*′,*N*′-tetramethylethylene diamine (TEMED).
11. SDS-PAGE running buffer: 1× Tris-glycine electrophoresis buffer contains 25 mM Tris base, 250 mM glycine, 0.1% SDS, and can be prepared as 5× or 10× stock solutions.
12. 4× SDS gel-loading loading buffer: 200 mM Tris–HCl, pH 6.8, 8% SDS, 40% glycerol, 0.4% bromophenol blue, and 5% β-mercaptoethanol.
13. Protein electrophoresis system.
14. CBB Staining Solution: 45.5% methanol, 9.2% acetic acid, and 0.125% Coomassie Brilliant Blue R-250.
15. CBB Destaining Solution: 5% methanol, 7.5% acetic acid, 2% glycerol.
16. Bradford protein assay kit.
17. Prestained protein ladder.
18. Blotter for semidry transfer.
19. Western blot boxes.
20. Polyvinylidene fluoride (PVDF) transfer membranes.
21. Blotting filter paper.
22. Methanol.

23. 10× Transfer Buffer: 480 mM Tris base, 390 mM glycine, and 0.375% SDS. Prepare working solution (1×) including 20% methanol.
24. TBST buffer: 10 mM Tris–HCl pH 8, 150 mM NaCl, 0.1% Tween-20.
25. Blocking solution: 5% nonfat dry milk in TBST. Store at −20 °C. Can be reused several times.
26. Antibodies: Anti-6xHis tag primary antibody and a horseradish peroxidase-conjugated secondary antibody, prepared in blocking solution.
27. Orbital shaker.
28. Horseradish peroxidase (HRP) chemiluminescent substrate.
29. Imaging system for chemiluminescence detection.

3 Methods

3.1 Construction of Vectors for Transient Expression

In this protocol, the pEAQ-HT vector [14] is used for transient overexpression in *C. roseus* plants. This plasmid enables the expression of either N- or C- terminal 6xHIS tagged proteins depending on the polylinker sites that the insert will be introduced (*see* **Note 1**).

1. Extract RNA from an appropriate source (i.e., plant tissue) using an RNA extraction kit, treat the sample with DNase following the manufacturer's instructions and determine the quantity and the quality of the extracted DNA-free RNAs using a spectrophotometer and agarose-gel electrophoresis, respectively.
2. Perform cDNA synthesis using a commercial kit and following the manufacturer's instructions.
3. Amplify the GOI by PCR on cDNA with a proofreading polymerase and a set of primers that introduce specific restriction sites at both extremities according to the selected cloning strategy (i.e., N- or C-terminal tagged protein) (*see* **Note 1**).
4. Verify by agarose-gel electrophoresis that the amplification resulted in a single band of the expected size.
5. Purify the PCR product by a commercial kit and digest both the product and the pEAQ-HT plasmid with the respective restriction enzymes following the manufacturer's instructions. Perform the double digestions sequentially if needed and heat-inactivate the enzymes.
6. Dephosphorylate the linearized vector with an alkaline phosphatase following the manufacturer's instructions and purify the digestions using a commercial kit.

7. Determine the concentrations of the digested insert and vector by a spectrophotometer.
8. Use a commercial T4 ligase to ligate the insert and the vector following the manufacturer's instructions.
9. Transform the reaction in chemically competent *E. coli* cells by a heat-shock transformation protocol.
10. Plate the cell suspension in LB plates supplemented with 50 mg/L kanamycin, grow overnight at 37 °C, and screen the colonies by colony-PCR for the presence of the insert.
11. Inoculate the positive clones in 5 mL liquid LB medium supplemented with 50 mg/L kanamycin, grow at 37 °C under continuous shaking (200–250 rpm).
12. Use a commercial kit to extract the plasmids, validate the presence of the insert after digestion with different combinations of restriction enzymes and verify the constructs by Sanger sequencing. Prepare glycerol stocks for archiving the strains at −80 °C.

3.2 Agroinfiltration of C. roseus Plants

3.2.1 Preparation of C. roseus Plants

1. Germinate the seeds of *C. roseus* (cv. Little Bright Eyes) in soil composed of a 4:1 volume ratio of loam and vermiculite at 28 °C in the greenhouse.
2. When the cotyledons are fully expanded, transfer the plantlets in individual pots and cultivate under the same controlled conditions.
3. Use plants 1–2 months old for vacuum infiltration with 3–5 fully developed pairs of leaves.

3.2.2 Preparation of Agrobacterium Strains for Agroinfiltration

The *Agrobacterium tumefaciens* GV3101 (pMP90) strain is used in this methodology. Follow standard protocols [18] for preparing electrocompetent *Agrobacterium* cells and electroporating the strain with the pEAQ-HT:GFP-6HIS or pEAQ-HT:GOI-6HIS plasmids.

1. Streak the GV3101 (pMP90) *Agrobacterium* strains transformed with the pEAQ-HT:GFP-6HIS or pEAQ-HT:GOI-6HIS plasmids in LB plates supplemented with 50 mg/L rifampicin, 50 mg/L gentamicin, and 50 mg/L kanamycin and grow for 2 days at 28 °C.
2. Inoculate single colonies from each strain in 3 mL of liquid LB medium supplemented with 50 mg/L rifampicin, 50 mg/L gentamicin, and 50 mg/L kanamycin and grow at 28 °C for ~24 h under continuous shaking (200–250 rpm).
3. Initiate 50 mL cultures from each strain in liquid LB supplemented with 50 mg/L rifampicin, 50 mg/L gentamicin, and 50 mg/L kanamycin by diluting 1/250 and grow overnight (~14–16 hrs) under continuous shaking (200–250 rpm).

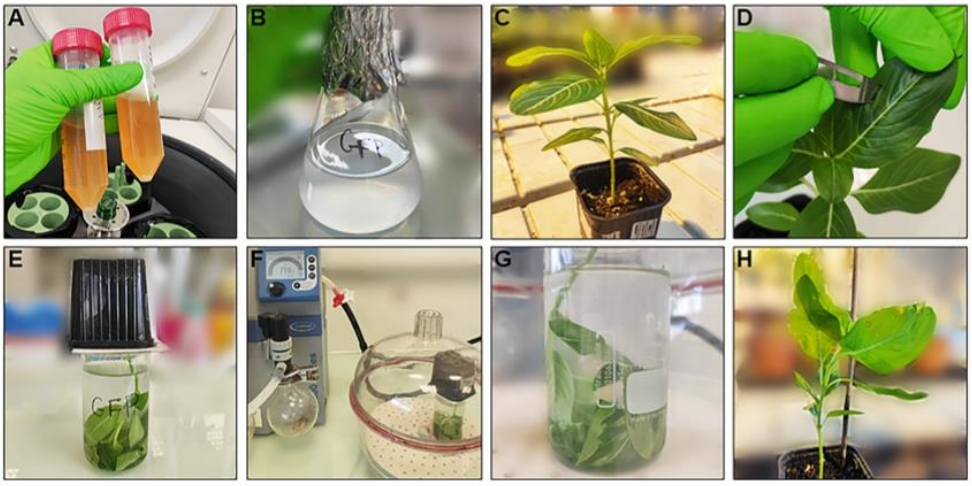


Fig. 2 Vacuum-based Agroinfiltration procedure of *C. roseus* plants. Cultures of *A. tumefaciens* cells are centrifuged (**A**), resuspended in dilution buffer at a final $OD_{600} = 0.25$ and incubated at 28 °C for 3 hrs with acetosyringone (**B**). *C. roseus* plants with 3–5 fully developed pairs of leaves (**C**) are carefully traumatized with a razor blade between the leaf veins (**D**) and fully submerged in *Agrobacterium* cell suspensions (**E**). The desiccator is sealed and vacuum is applied (**F**). After restoring the atmospheric pressure, darkened leaf areas appear indicating a successful Agroinfiltration (**G**). Five rounds of vacuum infiltration ensure that leaves are homogenously infiltrated. Plants are recovered in the greenhouse under controlled conditions (**H**)

4. Centrifuge the *Agrobacterium* cultures for 10 min at 1000 g at room temperature (Fig. 2A).
5. Wash each pellet with 20 mL of dilution buffer without disrupting the pellet in order to remove any remaining LB medium.
6. Gently resuspend the pellet of each culture in 20 mL dilution buffer.
7. Measure the OD_{600} of a 1/10 dilution with a spectrophotometer.
8. Dilute the cell suspensions in dilution buffer and combine GOI and GFP strains so that each has a final $OD_{600} = 0.25$ (*see* **Note 2**) in a final volume of 150 mL supplemented with acetosyringone to a final concentration of 150 μM. Use Erlenmeyer flasks (Fig. 2B).
9. Incubate the cell suspensions at 28 °C for 3 h under slow agitation.

3.2.3 *Vacuum Infiltration*

1. Select *C. roseus* plants with 3–5 fully developed pairs of leaves (i.e., 1–2 months old) (Fig. 2C).
2. Expose the plants to high humidity conditions 16–24 hours before Agroinfiltration. Use a box that can be covered with a

transparent plastic dome, add water in the bottom, and put the tray with the plants inside the box so that the tray floats (*see* **Note 3**).

3. Gently traumatize the fully developed leaves by scarifying with a razor blade or a scalpel. Apply 8–10 cuttings of 0.5 cm evenly distributed to a fully developed leaf. Traumatize only in leaf areas between the veins of the leaf (Fig. 2D).
4. Transfer the *Agrobacterium* cell suspensions in 150 mL tall-form beakers (Fig. 2E).
5. Invert the pot with a *C. roseus* plant and submerge the aerial parts into the beaker containing the *Agrobacterium* cell suspensions (Fig. 2E). Ensure that all the leaves are submerged in the solution. Use a piece of cardboard or plastic wrap to retain the soil in the pot during vacuum infiltration.
6. Place the beaker in the center of the vacuum desiccator unit, and properly seal the desiccator unit (Fig. 2F).
7. Initiate the vacuum pump and apply a negative pressure of 50 mbar, retain this pressure for 1 minute and gently break the vacuum until atmospheric pressure is reached. Repeat this step for a total of five rounds to ensure that leaves are evenly infiltrated. Typically, after the first or second round, darkening of leaf areas can be observed which indicates a successful infiltration (Fig. 2G).
8. Remove the pots from the beakers (Fig. 2H), place the infiltrated plants in a tray, and return the plants to the greenhouse. Ensure sufficient watering every 2–3 days.

3.3 Validation of Transient Expression

3.3.1 Macroscopic Observation of GFP Fluorescence

1. Observe the infiltrated *C. roseus* plants for GFP fluorescence with a UV lamp at 3 days post infiltration (*see* **Note 4**) in a dark room.
2. Collect two or three representative samples (i.e., by cutting regions of approx. 1 cm and avoiding the veins) for confirmation of GFP expression under an epifluorescence microscope (*see* Subheading 3.3.2).
3. Collect leaves (or leaf areas) exhibiting high fluorescence (Fig. 3A) and freeze immediately in liquid nitrogen (*see* **Note 5**).
4. Store the samples at −80 °C until needed.

3.3.2 Microscopic Observation of GFP Fluorescence

1. Deposit the samples collected for microscopic observation on a microscope slide. The abaxial side of the leaf sample should face the microscope objective lens.
2. Add a drop of distilled water and shield with a coverslip.

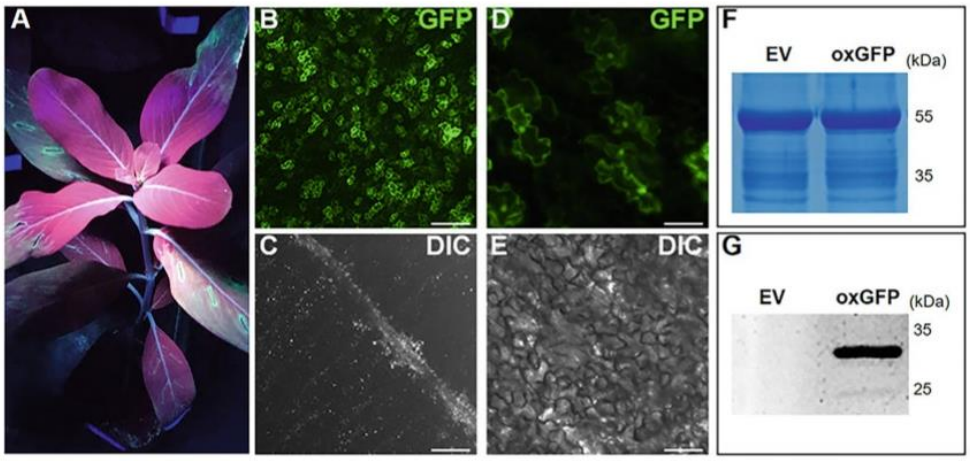


Fig. 3 Validation of transient overexpression of GFP in *C. roseus* leaves. Plants are examined macroscopically in a dark room under a UV lamp at 3 dpi and leaf areas exhibiting GFP fluorescence are collected (**A**). Overexpression of GFP is further confirmed under an epifluorescence microscope (**B–E**) and by immunodetection in soluble protein extracts from the leaf tissues (**F**, **G**). Bars 20 μm (**B**, **C**) or 100 μm (**D**, **E**), DIC, differential interference contrast. Protein extracts (20 μg) of *C. roseus* leaves Agroinfiltrated with either pEAQ-HT (EV) or pEAQ-HT:GFP-6HIS (oxGFP) constructs were resolved in 10% polyacrylamide gels and bands were visualized after CBB staining (**F**) or Western blotting (**G**) after immunodetection of the 6xHIS epitope. Numbers on the right (**F**, **G**) indicate molecular weights in kDa

3. Observe the slide under an epifluorescence microscope and acquire images with different objective lenses to ensure that a homogenous expression of GFP was achieved (Fig. 3B–E).

3.3.3 *Protein Extraction*

1. Prechill a mortar and a pestle with liquid nitrogen.
2. Transfer the harvested infiltrated *C. roseus* leaves from −80 °C to the cold mortar and grind the samples into a fine powder in the presence of liquid nitrogen. Supplement liquid nitrogen periodically to avoid thawing.
3. Transfer 500 mg of the ground sample in a sterile prechilled microcentrifuge tube and immediately add 500 μL of protein extraction buffer. Keep the remaining sample at −80 °C in case needed.
4. Mix thoroughly by vortexing the sample for 2 min.
5. Centrifuge the sample at 13.000 g for 10 min at 4 °C, collect the supernatant, and transfer to a sterile prechilled microcentrifuge tube. Keep the sample on ice.
6. Determine the concentration of the extracted soluble protein by a Bradford protein assay kit and a spectrophotometer following the manufacturer's instructions. Proceed with SDS-PAGE analysis or store the sample in −20 °C until needed.

3.3.4 SDS-PAGE

1. Assemble the glass plates according to the manufacturer's instructions. Prepare two identical gels, one for Coomassie staining and one for Western blotting.
2. Prepare 10 mL of 10% polyacrylamide resolving gel (*see* **Note 6**). Mix 4 mL of ddH_2O, 3.3 mL of Acrylamide/Bis-acrylamide solution (30%), 2.5 mL of 1.5 M Tris–HCl (pH 8.8), 100 μL of SDS (10%) and 100 μL of APS (10%) (*see* **Note 7**). Add 5 μL of TEMED, mix gently and immediately pour the acrylamide solution into the gap between the glass plates of the vertical cassette provided with 0.75 mm thick spacers. Leave enough space (~2.5 cm) for the stacking gel. Gently overlay the gel with ddH_2O.
3. After polymerization is complete, pour off the overlay from the cassette and use the edge of a paper towel to completely remove any remaining ddH_2O.
4. Prepare 5 mL of 5% polyacrylamide stacking gel. Mix 2.8 mL of ddH_2O, 830 μL of Acrylamide/Bis-acrylamide solution (30%), 1.26 mL of 0.5 M Tris–HCl (pH 6.8), 50 μL of SDS (10%), and 50 μL of APS (10%) (*see* **Note 7**). Add 5 μL of TEMED, mix gently, and immediately pour the acrylamide solution onto the surface of the polymerized resolving gel. Insert 10-well combs and avoid trapping air bubbles.
5. After polymerization of the stacking gel is complete, assemble the electrophoresis apparatus according to the manufacturer's instructions. Add 1× SDS-PAGE running buffer to both electrode compartments, remove the well combs, and use a pipette to wash 2–3 times the wells with running buffer in order to remove any unpolymerized acrylamide.
6. Dilute 20 μg of soluble protein extracts in protein extraction buffer up to a total volume of 15 μL. Add 5 μL of 4× SDS gel-loading buffer, mix gently, denature the proteins by incubating at 90 °C for 5 minutes and keep the samples on ice.
7. Load each protein sample in the wells and load prestained protein ladder in an empty well.
8. The electrophoresis is conducted with a constant current of 20 mA.
9. Stop the electrophoresis when the tracking dye front (bromophenol blue) is approximately 0.5 cm above the bottom of the glass plates.
10. Use one gel for Coomassie staining in order to observe the protein bands and ensure the integrity of the protein extracts. Immerse the resolving gel in 30 mL CBB staining solution for 20 minutes under gentle agitation. Remove the solution and soak the gel in an equal volume of CBB Destaining Solution for 20 minutes. Repeat this step three or four times until the bands are visible (Fig. 3F).

3.3.5 Western Blot and Chemiluminescence Detection

1. Use the second gel for Western blotting. Equilibrate the resolving gel in 50 mL of 1× Transfer Buffer for 10 minutes.
2. Cut two extra-thick blot papers and the PVDF membrane on the exact dimensions of the resolving gel. Activate the PVDF membrane by rinsing in methanol for ~1 minute and soak the membrane and the blotting papers in 1× Transfer Buffer.
3. Assemble the Western blot sandwich in the blotter for semidry transfer. Place one extra-thick blot paper on the anode, then the PVDF membrane, then the resolving gel, and finally the second extra-thick blot paper. Roll out any air bubbles with a 5 mL pipette tip and close the blotter with the cathode and the lid.
4. Connect the blotter to the power supply and apply a constant voltage of 10 V for 60 minutes (*see* **Note 8**).
5. After the blotting is finished, disassemble the sandwich and confirm that the prestained protein ladder was evenly migrated to the PVDF membrane. Immediately transfer the membrane into a western blot box containing 25 mL of blocking solution. The side of the membrane that was in contact with the resolving gel should be facing up. Incubate the membrane in blocking solution for 60 minutes by gently agitating at room temperature. The membrane should be fully submerged in all steps to avoid drying.
6. Discard the blocking solution and incubate the membrane in the primary antibody solution for 90 minutes by gently agitating at room temperature (*see* **Note 9**).
7. Remove the primary antibody solution and wash the PVDF membrane with TBST for 3×10 minutes. The primary antibody solution can be stored at −20 °C and reused several times.
8. Incubate the membrane in secondary antibody solution for 45 minutes by gently agitating at room temperature (*see* **Note 10**). Remove the secondary antibody solution and wash the PVDF membrane with TBST for 3×10 minutes.
9. Prepare the working solution of the substrate of HRP for chemiluminescence detection in a dark room. Mix (1:1) the two components of the substrate in a total volume of 1.5 mL.
10. Take the membrane out of the TBST solution and remove the excess of TBST by vertically touching one edge of the membrane on a paper towel.
11. Place the membrane on a glass plate, distribute evenly the HRP substrate on the membrane and incubate for 2 minutes at room temperature. Remove the HRP substrate by vertically touching one edge of the plate on a paper towel.

12. Transfer the plate with the membrane in the imaging system for chemiluminescence detection and acquire images at different times of exposure until the immunodetected bands are visible (Fig. 3G).

4 Notes

1. In order to generate a N-terminal 6xHIS tagged protein, amplify the GOI with a forward primer which introduces a *Xma*I or *Sma*I restriction site and a reverse primer that introduces a *Xho*I or *Stu*I restriction site and ligate in the respective sites of the plasmid's polylinker. In order to generate a C-terminal 6xHIS tagged protein, amplify the GOI with a forward primer which introduces a *Nru*I or *Age*I restriction site and a reverse primer (without a stop codon) which introduces a *Xma*I or *Sma*I restriction site and ligate in the respective sites of the plasmid's polylinker. If a non 6xHIS tagged protein is desired, amplify the GOI with a forward primer which introduces a *Nru*I or AgeI restriction site and a reverse primer which introduces a *Xho*I or *Stu*I restriction site and ligate in the respective sites of the plasmid's polylinker. *Nru*I, *Sma*I, and *Stu*I are restriction enzymes that generate blunt ends. *Age*I and *Xma*I are restriction enzymes that generate cohesive ends which are compatible. Whatever the cloning strategy (blunt-end or cohesive end) used, it is important to dephosphorylate the linearized vector to reduce recyclization of the plasmid during ligation without an insert. Compatibility of the ends though increases the options of the restriction sites to be incorporated by the primers in case one occurs in the coding sequence of the GOI. Note that the primers should be designed with respect to the frame of the 6xHIS epitope (s) encoded by the pEAQ-HT plasmid.
2. The optimal final OD_{600} of each culture may be adjusted depending on specific experimental needs. Due to high translational efficiency of the pEAQ-HT vector, an OD600 of 0.25 is enough for high expression and co-infiltration of *Agrobacterium* cells harboring different constructs (i.e., GFP and one or more GOI). Note that if the OD_{600} of the combined *Agrobacterium* cell suspension is very high, necrotic symptoms may be observed.
3. High humidity promotes stomatal opening and thus *Agrobacterium* cell suspensions can more easily access the mesophyll intercellular spaces elevating the susceptibility of the leaves to Agroinfiltration.
4. Peak of transient expression must be validated for each GOI. Typically, the highest protein accumulation is observed at 3–4

dpi but depending on the experiment, more days may be needed until a phenotype is observed. In our experience 7 dpi is usually sufficient to observe alterations on the metabolic content of the leaves and immunodetect the GOI product but, if needed, transient expression of the GOI is immunodetected even at 16 dpi [17].

5. Avoid pooling all the leaves together. Leaves should be collected separately or, if needed, pool only leaves of the same node. The susceptibility to Agroinfiltration, the efficiency, and the stability of the transient expression hugely vary as per age of leaves (i.e., leaves harvested at different nodes in the same plant).
6. The percentage of acrylamide in the resolving gel can be adjusted depending on the size of the protein of interest and the desired resolution. Typically, the percentage of acrylamide in resolving gels ranges from 6 up to 14%. A 10% Acrylamide/Bis-acrylamide (29:1) gel has a linear range of separation for proteins of 20 up to 80 kDa.
7. Unpolymerized acrylamide is neurotoxic. Wear gloves and handle with care. APS and TEMED should be added immediately before pouring the gels.
8. The time needed for complete transfer depends on the size of the protein and the percentage of acrylamide in the resolving gel.
9. The primary antibody may be either monoclonal or polyclonal, either raised against a specific epitope (i.e., 6xHIS, FLAG or other) or against a specific recombinant protein. The optimum dilution to be prepared needs to be validated in each case. Typically, primary antibodies are diluted in blocking solution in a range of 1:500 up to 1:5000. Store at − 20 °C. Primary antibody solution can be used several times.
10. Select a secondary antibody in respect to the primary antibody that will be used. If the primary antibody is produced, for example, in mouse, the secondary antibody should be an anti-mouse produced in another animal (i.e., goat, rabbit, or other). Secondary antibody is diluted in blocking solution and typically in a range of 1:1000 up to 1:3000. The optimum dilution to be prepared needs to be validated in each case. Store at − 20 °C. Avoid using the secondary antibody more than 4–5 times.

Acknowledgments

We acknowledge funding from the ARD2020 and ARD CVL Biopharmaceutical programs of the Région Centre Val de Loire

(BioPROPHARM, CatharSIS, and ETOPOCentre projects), EU Horizon 2020 research and innovation program (MIAMi project-grant agreement N°814645), and ANR MIACYC (ANR-20-CE43-0010).

References

1. O'Connor SE, Maresh JJ (2006) Chemistry and biology of monoterpene indole alkaloid biosynthesis. Nat Prod Rep 23(4):532–547. https://doi.org/10.1039/B512615K
2. Dugé de Bernonville T, Clastre M, Besseau S, Oudin A, Burlat V, Glévarec G, Lanoue A, Papon N, Giglioli-Guivarc'h N, St-Pierre B, Courdavault V (2015) Phytochemical genomics of the Madagascar periwinkle: unravelling the last twists of the alkaloid engine. Phytochemistry 113:9–23. https://doi.org/10.1016/j.phytochem.2014.07.023
3. Liscombe DK, O'Connor SE (2011) A virus-induced gene silencing approach to understanding alkaloid metabolism in *Catharanthus roseus*. Phytochemistry 72(16):1969–1977. https://doi.org/10.1016/j.phytochem.2011.07.001
4. Carqueijeiro I, Masini E, Foureau E, Sepúlveda LJ, Marais E, Lanoue A, Besseau S, Papon N, Clastre M, Dugé de Bernonville T, Glévarec G, Atehortùa L, Oudin A, Courdavault V (2015) Virus-induced gene silencing in *Catharanthus roseus* by biolistic inoculation of tobacco rattle virus vectors. Plant Biol 17(6):1242–1246. https://doi.org/10.1111/plb.12380
5. Caputi L, Franke J, Farrow SC, Chung K, Payne RME, Nguyen T-D, Dang T-TT, Soares Teto Carqueijeiro I, Koudounas K, Dugé de Bernonville T, Ameyaw B, Jones DM, Vieira IJC, Courdavault V, O'Connor SE (2018) Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. Science 360(6394):1235–1239. https://doi.org/10.1126/science.aat4100
6. Facchini PJ, De Luca V (2008) Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J 54(4):763–784. https://doi.org/10.1111/j.1365-313X.2008.03438.x
7. Sharma A, Amin D, Sankaranarayanan A, Arora R, Mathur AK (2020) Present status of *Catharanthus roseus* monoterpenoid indole alkaloids engineering in homo- and heterologous systems. Biotechnol Lett 42(1):11–23. https://doi.org/10.1007/s10529-019-02757-4
8. Weaver J, Goklany S, Rizvi N, Cram EJ, Lee-Parsons CWT (2014) Optimizing the transient FAST agro-mediated seedling transformation (FAST) method in *Catharanthus roseus* seedlings. Plant Cell Rep 33(1):89–97. https://doi.org/10.1007/s00299-013-1514-2
9. Mortensen S, Bernal-Franco D, Cole LF, Sathitloetsakun S, Cram EJ, Lee-Parsons CWT (2019) EASI transformation: an efficient transient expression method for analyzing gene function in *Catharanthus roseus* seedlings. Front Plant Sci 10(755). https://doi.org/10.3389/fpls.2019.00755
10. Schweizer F, Colinas M, Pollier J, Van Moerkercke A, Vanden Bossche R, de Clercq R, Goossens A (2018) An engineered combinatorial module of transcription factors boosts production of monoterpenoid indole alkaloids in *Catharanthus roseus*. Met Eng 48:150–162. https://doi.org/10.1016/j.ymben.2018.05.016
11. Di Fiore S, Hoppmann V, Fischer R, Schillberg S (2004) Transient gene expression of recombinant terpenoid indole alkaloid enzymes in *Catharanthus roseus* leaves. Plant Mol Biol Report 22(1):15–22. https://doi.org/10.1007/BF02773344
12. Sharma A, Verma P, Mathur A, Mathur AK (2018) Genetic engineering approach using early Vinca alkaloid biosynthesis genes led to increased tryptamine and terpenoid indole alkaloids biosynthesis in differentiating cultures of *Catharanthus roseus*. Protoplasma 255(1):425–435. https://doi.org/10.1007/s00709-017-1151-7
13. Peyret H, Lomonossoff GP (2015) When plant virology met agrobacterium: the rise of the deconstructed clones. Plant Biotechnol J 13(8):1121–1135. https://doi.org/10.1111/pbi.12412
14. Sainsbury F, Thuenemann EC, Lomonossoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol J 7(7):682–693. https://doi.org/10.1111/j.1467-7652.2009.00434.x
15. Peyret H, Lomonossoff GP (2013) The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. Plant Mol Biol

83(1):51–58. https://doi.org/10.1007/s11103-013-0036-1

16. Mortimer C, Dugdale B, Waterhouse P (2020) Development of an autonomously replicating viral expression system tailored for Catharanthus roseus. Plant Biotechnol J 18(5): 1115–1117. https://doi.org/10.1111/pbi.13274
17. Carqueijeiro I, Koudounas K, Dugé de Bernonville T, Sepúlveda LJ, Mosquera A, Bomzan DP, Oudin A, Lanoue A, Besseau S, Lemos Cruz P, Kulagina N, Stander EA, Eymieux S, Burlaud-Gaillard J, Blanchard E, Clastre M, Atehortùa L, St-Pierre B, Giglioli-Guivarc'h N, Papon N, Nagegowda DA, O'Connor SE, Courdavault V (2021) Alternative splicing creates a pseudo-strictosidine β-D-glucosidase modulating alkaloid synthesis in *Catharanthus roseus*. Plant Physiol 185(3): 836–856. https://doi.org/10.1093/plphys/kiaa075
18. Koudounas K, Thomopoulou M, Angeli E, Tsitsekian D, Rigas S, Hatzopoulos P (2020) Virus-induced gene silencing in olive tree (Oleaceae). In: Courdavault V, Besseau S (eds) Virus-induced gene silencing in plants: methods and protocols. Springer US, New York, pp 165–182. https://doi.org/10.1007/978-1-0716-0751-0_13

DONNÉES SUPPLÉMENTAIRES

actagttcggtcatgggttctacagcacccctaaggcttccagttatagatttatccatgaagaacttgaagcctggaacaacttcttggaa
ctcggtacgcacccaggtacgggaggcactggaagaatacggttgctttgaagctgtgatcgatgctgtgtctccagagctgcagaag
gcagtatgtaacaaaggacacgagctgcttaatcttccattagaaaccaagatgttgaacggaaacaaaccagaatatgatggatttacg
tcaataccaaacctcaacgaaggcatgggagtcggcagaataacagatttggaaaaagttgagaggttcactaatcttatgtggcccga
ggggaataaggatttctgtgaaactgtgtattcttatggcaaacgaatggcggaggtggaccacatattgaaaatgatggttttcgagagt
tttggaatggagaagcacttcgactcgttctgcgaatcaacaaattaccttctccatttcatgagataccaacaaccagggaaggatgga
cgttcacctgctctttcgttgcataaggacaagagcatcttgaccatagtaaaccaaaatgatgtcaagggattggaatttgaaaccaagg
atggagaatggattttacctacagctgacaaccatattgttcttctaggagactgcttcatggcatggagcaatggtagattacatagtcct
cttcaccgggtcacgttggtcgcgaaccaggcgaggttatctacatcatcgttttcgtttccaaaggacataatagagacccctgcagag
ctggtggatgaagagcatcctttgctatttaatccctttgagataacggagttgcttgcttactgtttcacaaaagagggtgcaaaggcggt
gtgtgacctcaagcaatacaaggcgtacacaggtgcatgaactagt

Séquence codant la 2ODD de *P. hexandrum*

atggagttcctttcatttcccttatcatccgctttgctcatcatcttgttgttcatgctggtgaagaaagcaatccaaaacaaaaattcaaagct
tccccccagggccacggaagctacctgtcattggaagtttgcaccatttgataggcgatggccttccacatcacgcactgagaaatctag
ccaaaaaacatggcccccttatgcacctacaacttggtgaaaattccactcttgttgtatcctcatcgaaaatggccagagcaatcatgtc
aacacacgacctcatgtttgcaaacaggactgtccagtttgcaacagacattttgttatatggtggtaaaggcattggtttggcaccttatg
gtgattattggaggcaaataagaaagatttgcgtcttagagctcctaagtgcgaagcgtgtgcaatcattccaaacagtacgggaagaa
gagatatcaaatctcattcgtagtatttcggctgaatcatcaaaagttaacttcagtgaaatgatcacatcattaacaaatgacataatcgcc
agggctgcctttggcgaaaaatgcaaagataagcacgcattcttatccttgattgaagaaggaattcagttagctggagggtttgatattg
ctggtttatttccttcactaaaactccttcatgtcatcactgggaaaaagcgccaattggagagggtacacaatgagttggacagaatcctc
gtagacatcataaaagaaaacaaagagaaaagaagagcaagtaaacttaacgaggacaaatatgaagagaatctagtagatgtacttg
tgcgcgtccaagagaatactcagcttgaggttccacttgcaaatgacaatcttaaagccgtaatactagacatctttgttgccggaagtga
aacatcatccactacaatagaatgggcaatgtcagaaatgctgaaaaatccaaaagttatgaagaaggcacaagctgaggtgagacga
gtgttcagtgggaacaaaaagatcaacgagacggaaattcaagaattgagttacttaaaattagttctcaaagaaactctcaggttacatg
ctccagctccactgttgatcccccgagaatgtagggagagttgtgagattgatggctatgagataccgaaaaagacgatggttatggtta
atgcatgggcaattggaagagatcctgagaactggagtaacggggataggtttgagccagagaggtttgatggaatttctgtggattac
aaggggacaaactttgaatacatccctttggagcaggtagaagaatatgtccaggcatgttgtttgggatggcaaacgttgaggttcca
ctagctcgcttactataccacttcgactgggaactccccaacggcatcaaaccagaagagctggacatggatgagacttttggtacaac
ggttcgaaggaaaaatcacttatacttaatccccactgtggtgcatgagctggaaagatcgactacaaaagaatag

Séquence codant CYP71BE54 de *P. hexandrum*

atggattccctgcactgcctagaaactctgttgcttggcttctttgtgttactcccatgtttttctactttgtatggaagaagccaaataacaaa
attaaagagccaccccaaccagctggagcatggccgatcatcggccacctccatctactagcgcgaggtgatctcccccacaaaatttt
gtcttcctttgcagacaagaatggcccggttttcaagatacaacttggtgtacaccaagcactggtggtaaataattctgagatagccaag
gaatgtttcaccacaaatgacaggttcttttttaaatcgtccttccggggtagcagcaaagatcatgggttacaactatgttatgttaggagta
gcccccttatgggccttattggcgtgacatgcgtaaaataatcatgttggagttcctctccaaccgtcggctgcagtctctaaagcatgtatg
gcactcagaaatatcaatatcttccaaagaattgtataaactgtgggaaacccaaaatatagatttttgcctagttgatatgaagcaatggct
ggctgatctcacactcaatatgtcagttaagatggtggttgggaagagatttttcggttcagcttctgcttctgcttgtgaagaaactgagtc
aagcaattgtccaaaaacattgagaaatatgtttagacttatgggttcgtttgtgctgtcagactatcttccttatctaaggtggttggacttgg
gaggccatgagaaggagatgaagagaacagtgaaagaacttgacattttattcaagggatggttggatgagcataaaagaaagaggc
tatctggtgggaaggaagacgacgaccaggattttatggacgtaatgctttccatcctagaagaatcaaaattgggaaatgatgttgatac
catcaacaagactgcatgcttggcaataatactaggtggagctgatactacatgggccaccttgacatgggcactttcattgttattgaac

aatccaaatgcgttgaaaaaggcccaagacgagttggacttacatgttggcagggatagaaacgtggacgaatcagacctcgtgaagt
tgacatacatcgatgccattatcaaggaaacattgcgtctatacccacctggtccgttactaggtccccgtgtggtcacagaggattgcac
catagctgggtatcatgttcgagcaggcactcgactaattgtgaatgcttggaagatacaacgagacccactagtttggtcgcaacctca
tgagtaccaacctgagagatttcttgagagagatgtggacatgaaaggtcaacattttgagttaatcccatttggatcgggtagacgggc
atgcccagcgatctctttagcacttcaagtgttgcctttgacgttggctcatatattgcatggttttgagcttaggaccccaaaccagaacaa
ggtggatatgactgagacaccaggtatagttcatgcaaaggcaactcctttggaagtacttgtcgccccacgcatctctccaaaatgtttt
gtctag

Séquence codant CYP82D61 de *P. hexandrum*

atgaaaaacatgatgaattataaattaaaactctgttctgtctcaaaaaactcaaaaggagtctctctctcacctacaccacacctaaccaa
acccccctacgattcacacagagagagatcttcttcttccttcttcttccttcttctttcttcttctttcttcttctagctacaacatctacaacgcca
tgtcctcttcttcttcttcgtcaacctccatgatcgatctcatggcagcaatcatcaaaggagagcctgtaattgtctccgacccagctaatg
cctccgcttacgagtccgtagctgctgaattatcctctatgcttatagagaatcgtcaattcgccatgattgttaccacttccattgctgttctt
attggttgcatcgttatgctcgtttggaggagatccggttctgggaattcaaaacgtgtcgagcctcttaagcctttggttattaagcctcgt
gaggaagagattgatgatgggcgtaagaaagttaccatctttttcggtacacaaactggtactgctgaaggttttgcaaaggctttaggag
aagaagctaaagcaagatatgaaaagaccagattcaaaatcgttgatttggatgattacgcggctgatgatgatgagtatgaggagaaa
ttgaagaaagaggatgtggctttcttcttcttagccacatatggagatggtgagcctaccgacaatgcagcgagattctacaaatggttca
ccgaggggaatgacagaggagaatggcttaagaacttgaagtatggagtgtttggattaggaaacagacaatatgagcattttaataag
gttgccaaagttgtagatgacattcttgtcgaacaaggtgcacagcgtcttgtacaagttggtcttggagatgatgaccagtgtattgaag
atgactttaccgcttggcgagaagcattgtggcccgagcttgatacaatactgagggaagaaggggatacagctgttgccacaccatac
actgcagctgtgttagaatacagagtttctattcacgactctgaagatgccaaattcaatgatataaacatggcaaatgggaatggttacac
tgtgtttgatgctcaacatccttacaaagcaaatgtcgctgttaaaagggagcttcatactcccgagtctgatcgttcttgtatccatttggaa
tttgacattgctggaagtggacttacgtatgaaactggagatcatgttggtgtactttgtgataacttaagtgaaactgtagatgaagctctta
gattgctggatatgtcacctgatacttatttctcacttcacgctgaaaaagaagacggcacaccaatcagcagctcactgcctcctcccttc
ccaccttgcaacttgagaacagcgcttacacgatatgcatgtcttttgagttctccaaagaagtctgctttagttgcgttggctgctcatgca
tctgatcctaccgaagcagaacgattaaaacaccttgcttcacctgctggaaaggatgaatattcaaagtgggtagtagagagtcaaag
aagtctacttgaggtgatggccgagtttccttcagccaagccaccacttggtgtcttcttcgctggagttgctccaaggttgcagcctaggt
tctattcgatatcatcatcgcccaagattgctgaaactagaattcacgtcacatgtgcactggtttatgagaaaatgccaactggcaggatt
cataagggagtgtgttccacttggatgaagaatgctgtgccttacgagaagagtgaaaactgttcctcggcgccgatatttgttaggcaat
ccaacttcaagcttccttctgattctaaggtaccgatcatcatgatcggtccagggactggattagctccattcagaggattccttcaggaa
agactagcgttggtagaatctggtgttgaacttgggccatcagttttgttctttggatgcagaaaccgtagaatggatttcatctacgagga
agagctccagcgatttgttgagagtggtgctctcgcagagctaagtgtcgccttctctcgtgaaggacccaccaaagaatacgtacagc
acaagatgatggacaaggcttctgatatctggaatatgatctctcaaggagcttatttatatgtttgtggtgacgccaaaggcatggcaag
agatgttcacagatctctccacacaatagctcaagaacaggggtcaatggattcaactaaagcagagggcttcgtgaagaatctgcaaa
cgagtggaagatatcttagagatgtatggtaa

Séquence codant AtATR2 d'*A. thaliana*

atgtattcactgcctttcctagaaactctgttgcttggcttctttgtgttactatgtttttctacttagtatggaagaagccaaataacaaaatta
aagagccacccccaaccagctggagcatggccgatcatcggccacctccatctactagtacgaggtgacctcccccacctaactttgtct
gccatggccgataagtatggcccggttttcaaaatacaccttggtgtacagcaagcactggtggtaaataattctgagatagccaaggaa
tgtttcaccacaaatgacaggttctttttaaatcgtacttccggggtagcagtaaagatcatgagttacaactatgttatgtttggatttgcccc
ttatgggccttattggcgtgatatgcgtaaaataatcatgttggacctcctctcgaaacgtcggctgcagtctctaaagcatgtatggcact
ctgaaatatcaatatcttccaaagaattgtataaactgtgggaaacccaaaagatagatttttgcctagtagatatgaagcaatggcttgct

gatctgacactcaatatgtcagttaagatggtggttggaaagagagattttggttcagcttctgcttgtgaggaaacggagtcaaggaagt
gtcaaaaagcaatgagaagtatgtttagatttgtgggtttgtttatgctgtcagactatcttccttatctagggtggttggacttgggaggcta
tgagaaggagatgaagagaacaatgaaagaacttgatattttgtttcagggatggttggatgagcataaaacaaagaggctgtctggtg
cgaaggaagacgacgaccaggattttatggacgtaatgctttccatcctagaagaatcaaaattgggatatgatgttgataccatcaaca
agtctgcatgcatgggaatattattaggtggaactgatactacatgggccaccttgacatgggcactttcattgttattgaacaacccaaat
gcgttaaaaaaggcccacgatgagctggacttccacgttggcagagatagaaacgtggacgaatcagaccttgtgaagttgacataca
tcgacgccattatcaaggaaacattgcgtctatacccagcttctccgttactaggcctccgtatggtcacagaggactgcaccatagctg
ggtatcatgttcgagcaggcacacggctaattgtgaatggttggaagattcaacgagacccactggtttggtcgcaacctcatgagttcc
aacctgagagatttctcgagagagatgtggacatgaaaggtcaaaattttgagttaatcccatttggttcaggtaggcgggcatgcccag
gatctctttagcacttcaagtgttgcctttgacgttggctcatatattgcatggttttgagcttaggaccccaaaccagaataaggtggata
tgaccgagacaccaggtatggttcacgcaaagacaactcctttggaagtacttgttgccccacgcatctctccaaaatgttttgtctag

Séquence codant DN717 de *P. hexandrum*

atggattccctgcactgcctagaaactctgttgcttggcttctttgtgtcactatgttttttctacttagtatggaagaaaccaaataacaaactt
aaagagccaccccaaccagctggagcatggccgatcatcggccacttacatctactagcgcgaggtaatctcctccacaaaactttgtc
ttccttggccgacaagtatggcccggttttcaagatacaacttggtgtacaccaagcactggtggtaaataattctgagatagccaagga
atgtttcaccacaaatgacaggttctttttaaatcgtccttccggggtagcagcaaagatcatgggttacaactatgttatgtttggagtagc
cccttatggtccttattggcgtgaaatgcgtaaaataatcatgttggagttcctctccaaccgtcggctgcagtctctaaagcatgtatggc
actcagaaatatcaatatcttccaaagaattgtatcaactgtgggaaacccaaaatatagattttgcctagttgatatgaagcaatggcttg
ctgatctcacactcaatatgtcagttaagatggtggttgggaagagatttttcggttcagcttctgcttctgcttctgaggaaactgagtcaa
gtaagtgtccaaaaacattgaaatatatgtttagacttatgggttcgtttgtgctgtcagactattttccttatctaaggtggttggacttggga
ggccctgagaaggagatgaagagaacagtgaaagaagttgatattttgttcaagggatggttggatgagcataaaagaaagaggctgt
ctggtgggaaggaagacgacgaccaggattttatggacgtaatgctttcctgcctagaagaatcaaaattgggacatgatgttgatacca
tcaacaagactgcatgcttggcaatagttctaggtggagctgaaactacatgggccaccttgacatgggcactttcattgttattgaataat
ccaaatgcgttgaaaaaggcccaagacgagttggacttccatgttggcacggatagaaacgtggacgaatctgaccttgtgaagttgac
atacatcgacgccattatcaaggaaacattgcgcctatacccacctggtcccctactaggcccccgtgtggtcacagaggactgcacca
tagctgggtatcatgttcgagcaggcacacgactacttgtgaatgcttggaagatgcaacgagacccactagtttggtcgcagcctcatg
agttccaacctgagagatttctcgagagagatgtggacatgaaaggtcaacattttgagttaatcccatttggatcgggtagacgggcat
gcccagggatctctttagcacttcaagtgttgcctttgacgttggctcatatattgcatggttttgagcttaggaccccataccagaacaca
gtggatatgactgagacaccaggcatagttcatgtaaagacaactcctttggaagtacttgttgctccacgcatctctccaaaatgttttgtc
tag

Séquence codant CYP82D61 de *P. spotty dotty*

atggattccctgcactgcctagaaactctgttgcttggcttctttgtgttactatgttttttctacttagtatggaagaaaccaaataacaaaatt
aaagaggcaccccaaccagctggagcatggccgatcatcggccacctccatctactagcgcgaggtgatctcccccacaaaactttgt
cttccttggccgacaagtatggtccggttttcaagatacaacttggtgtacaccaagcactggtggtaaataattctgagatagccaagga
atgtttcaccacaaatgacaggttctttttaaatcgtccttccggggtagtaccaaagatcatgggttacaactatgttatgtttggagtagcc
ccttatggtccttattggcgtgatatgcgtaaaataatcatgttggagttcctctccaaccgtccgctgcagtctctaaagcatgtatggcac
tcagaaatatcaatatcgtccaaagaattgtacaaactgtgggaaacccaaaatatagattttgcctagttgatatgaagcaatggcttgct
gatctcacactcaatatgttagttaagatggtggttgggaagaagatttttcggttcagcttcggcttctgcttctgcttgtgaggaaactga
gtcaagcaattgtccaaaaacattgagaaatatgtttagacttatgggttcgtttgtgctgtcagactatcttccttatctaaggtggttggact
tgggaggccatgagaaggagatgaagagaacagtgaaagaacttgacattttattcaagggatggttggatgagcataaaagaaaga
ggctatctggtgggaaggaagacgacgaccaggattttatggacgtaatgctttccatcctagaagaatcaaaattgggacatgatgttg

ataccatcaacaagactgcatgcttggcaatagtactaggtggagctgatactacatgggccaccttgacatgggcactttcattgttattg
aacaatccaaatgcgttgaaaaaggcacaagacgagttggacttccatgttggcagggatagaaacgtggacgaatcaggccttgtga
agttgacatacatcgacctcattatcaaggaaacattgcgtctatacccacctggtccattattaggtcgtgaggtcacagaggactgcac
catagctgggtatcatgttcgagcaggcacacgactacttatgaatgcttggaagatgcaacgagacccactagtttggtcgcaacctca
tgagttccaacctgagagatttctcgagagagatgtggacatgaaaggtcaacattttgagttaatcccatttggttcaggtagacgggca
tgcccagggatctctttagcacttcaagtgttgcctttgacattggctcgtatattgcatagttttgagcttaggaccccaaaccagaacaag
gtggatatgaccgagacaccaggcatagttcatgcaaagacaactcctttggaagtacttgtcgccccacgcatctctccaaaatgttttg
tctag

Séquence codant CYP82D61 de *P. peltatum*

atggctctcatcccagacttggccatggaaacctggcttctcctggctgtcagcctggtgctcctctatctatatggaacccattcacatgg
acttttaagaagcttggaattccagggcccacacctctgccttttttgggaaatattttgtcctaccataagggcttttgtatgtttgacatgg
aatgtcataaaaagtatggaaaagtgtggggcttttatgatggtcaacagcctgtgctggctatcacagatcctgacatgatcaaaacagt
gctagtgaaagaatgttattctgtcttcacaaaccggaggccttttggtccagtgggatttatgaaaagtgccatctctatagctgaggatg
aagaatggaagagattacgatcattgctgtctccaaccttcaccagtggaaaactcaaggagatggtccctatcattgcccagtatggag
atgtgttggtgagaaatctgaggcgggaagcagagacaggcaagcctgtcaccttgaaagacgtctttggggcctacagcatggatgt
gatcactagcacatcatttggagtgaacatcgactctctcaacaatccacaagacccctttgtggaaaacaccaagaagcttttaagattt
gatttttttggatccattctttctctcaataacagtctttccattcctcatcccaattcttgaagtattaaatatctgtgtgtttccaagagaagttac
aaatttttttaagaaaatctgtaaaaaggatgaaagaaagtcgcctcgaagatacacaaaagcaccgagtggatttccttcagctgatgatt
gactctcagaattcaaaagaaactgagtcccacaaagctctgtccgatctggagctcgtggcccaatcaattatctttatttttgctggctat
gaaaccacgagcagtgttctctccttcattatgtatgaactggccactcaccctgatgtccagcagaaactgcaggaggaaattgatgca
gttttacccaataaggcaccacccacctatgatactgtgctacagatggagtatcttgacatggtggtgaatgaaacgctcagattattccc
aattgctatgagacttgagagggtctgcaaaaaagatgttgagatcaatgggatgttcattcccaaaggggtggtggtgatgattccaag
ctatgctcttcaccgtgacccaaagtactggacagagcctgagaagttcctccctgaaagattcagcaagaagaacaaggacaacata
gatccttacatatacacacccttggaagtggacccagaaactgcattggcatgaggtttgctctcatgaacatgaaacttgctctaatcag
agtccttcagaacttctccttcaaaccttgtaaagaaacacagatccccctgaaattaagcttaggaggacttcttcaaccagaaaaaccc
gttgttctaaaggttgagtcaagggatggcaccgtaagtggagcctga

Séquence codant la CYP3A4 humain

atggcagagcagtcggacgaggccgtgaagtactacaccctagaggagattcagaagcacaaccacagcaagagcacctggctgat
cctgcaccacaaggtgtacgatttgaccaaatttctggaagagcatcctggtggggaagaagttttaagggaacaagctggaggtgac
gctactgagaactttgaggatgtcgggcactctacagatgccagggaaatgtccaaaacattcatcattggggagctccatccagatga
cagaccaaagttaaacaagcctccggaaactcttatcactactattgattctagttccagttggtggaccaactgggtgatccctgccatct
ctgcagtggccgtcgccttgatgtatcgcctatacatggcagaggactga

Séquence codant HsCB5 humain

atgccgtttggaatagacaacaccgacttcactgtcctggcggggctagtgcttgccgtgctactgtacgtaaagagaaactccatcaag
gaactgctgatgtccgatgacggagatatcacagctgtcagctcgggcaacagagacattgctcaggtggtgaccgaaaacaacaag
aactacttggtgttgtatgcgtcgcagactgggactgccgaggattacgccaaaaagttttccaaggagctggtggccaagttcaaccta
aacgtgatgtgcgcagatgttgagaactacgactttgagtcgctaaacgatgtgcccgtcatagtctcgatttttatctctacatatggtgaa
ggagacttccccgacggggcggtcaactttgaagactttatttgtaatgcggaagcgggtgcactatcgaacctgaggtataatatgtttg
gtctgggaaattctacttatgaattctttaatggtgccgccaagaaggccgagaagcatctctccgccgcgggcgctatcagactaggc

aagctcggtgaagctgatgatggtgcaggaactacagacgaagattacatggcctggaaggactccatcctggaggttttgaaagacg
aactgcatttggacgaacaggaagccaagttcacctctcaattccagtacactgtgttgaacgaaatcactgactccatgtcgcttggtga
accctctgctcactatttgccctcgcatcagttgaaccgcaacgcagacggcatccaattgggtcccttcgatttgtctcaaccgtatattg
cacccatcgtgaaatctcgcgaactgttctcttccaatgaccgtaattgcatccactctgaatttgacttgtccggctctaacatcaagtact
ccactggtgaccatcttgctgtttggccttccaacccattggaaaaggtcgaacagttcttatccatattcaacctggaccctgaaaccattt
ttgacttgaagcccctggatcccaccgtcaaagtgcccttcccaacgccaactactattggcgctgctattaaacactatttggaaattaca
ggacctgtctccagacaattgttttcatctttgattcagttcgcccccaacgctgacgtcaaggaaaaattgactctgctttcgaaagacaa
ggaccaattcgccgtcgagataacctccaaatatttcaacatcgcagatgctctgaaatatttgtctgatggcgccaaatgggacaccgt
acccatgcaattcttggtcgaatcagttccccaaatgactcctcgttactactctatctcttcctcttctctgtctgaaaagcaaaccgtccat
gtcacctccattgtggaaaactttcctaacccagaattgcctgatgctcctccagttgttggtgttacgactaacttgttaagaaacattcaat
tggctcaaaacaatgttaacattgccgaaactaacctacctgttcactacgatttaaatggcccacgtaaacttttcgccaattacaaattgc
ccgtccacgttcgtcgttctaacttcagattgccttccaacccttccaccccagttatcatgatcggtccaggtaccggtgttgccccattcc
gtgggtttatcagagagcgtgtcgcgttcctcgaatcacaaaagaagggcggtaacaacgtttcgctaggtaagcatatactgttttatgg
atcccgtaacactgatgatttcttgtaccaggacgaatggccagaatacgccaaaaaattggatggttcgttcgaaatggtcgtggccca
ttccaggttgccaaacaccaaaaaagtttatgttcaagataaattaaaggattacgaagaccaagtatttgaaatgattaacaacggtgcat
ttatctacgtctgtggtgatgcaaagggtatggccaagggtgtgtcaaccgcattggttggcatcttatcccgtggtaaatccattaccact
gatgaagcaacagagctaatcaagatgctcaagacttcaggtagataccaagaagatgtctggtaa

Séquence codant la ScCPR2 de *S. cerevisiae*

atgcaatcaagctcagctttcgatctaatgtcggcgattctccaaggaaaatccgatccattcgatgtatcaaaccgagaggttttgatgat
cttcacgacgtcgatcgctgttttgatcggttgtgtggttgtatttgtctggagaagatcctccggttcgaaatcgagtaaacctgttgagcc
ggttagaccgttgattgttagagaagctgaaccagatgttgctgatggtaagaagaaggtttctatcttctttggtactcaaaccggtaccg
ccgaaggcttcgctaaggcgttagcggatgaggccaaagctaggtacgataaggcggtgtttaaagtaatcgatctggatgattatgca
gccgacgatgatgagtacgagacgaagttgaagaaggaatctatcgccatttttcttttggctacatatggggatggtgaaccaactgac
aattgtgcgagattttataaatggtttactgaagaaggaaaagagaaaggagcttggcttcagaatcttcagtatggcgtattcgggctgg
gcaacaggcagtatgagcatttcaataagattgcgaaggtggtggatgaacaacttactgtgcaaggtgcaaagcgtctggttccagtg
ggtattggtgatgatgatcaatgcattgaagatgacttcactgcgtggagagaattggtgtggccagaattggatcaattactccgagac
gaagatgatgcaccaagtgcttctactccttatactgctgctgtaccagaatatcgggttgtattccatgaccctgaggatgctatgctaca
ggaaaagaactggggtaatgcaaatggacatgctgttcatgatattaaccatccatgccgggctaatgtagctgttaagagggagcttca
tactccagcttctgatcgctcttgtactcatctggaattcgacatctctggcactggacttgcatacgaaacaggagaccatgttggggtct
attgtgaaaatcatattgaaaatgtggaggaggcagagattttattgggttttgcaccagagactttcttctcagtctactctgaaaaagatg
atggcacaccaattggtggttccttacctcctcctttcccatctccctgcactgttagagcagcacttacccgatatgcagatcttttgaattc
ccccaaaaaggctgctctagttgctttagctgctcatgcatctgatccaagcgaagctgatcgattgagatttcttgcatctcctgctggaa
aggatgaatattctcagtggattgttgcaagtcagagaagtctactggaagtcatggctgcatttccttcagccaaacctcctctaggagtt
ttctttgcagcaatagcgcctcgcttgcagccaagatactattcaatctcatcctccccaaggatggcaccatctagaattcatgtcacatg
tgcattagtgtatgagaaaacacctagtggtcggattcacaaaggagtgtgctctacttggatgaagaatgctgttccattagaagaaagc
cgcgattgcagctgggcacccatttttgtgaggcaatccaacttcaaattgccagctgatactaaaacacctattgtcatgattggccctg
gtacaggattggctccttttaggggatttttgcaggaacggttaggcttgaaagaagctggtgcagaattgggacctgcaattctcttcttc
ggatgcagaaaccgtaggatggattttatatatgaagacgagttgaatggctttgttgaagcgggtgcaatctcagagttggtaattgcttt
ctcgagagagggacccaacaaggaatatgtgcaacataagatgatggagagggcctcagatgtatgggacatgatctcgaagggtg
gttatctatacgtctgtggtgatgccaaaggcatggccagggatgttcaccggactcttcacacgattgcccaagagcagggattattgg
acagcagcaaggctgagagcatggtgaagaatctgcagatggatgggagatatctccgtgacgtatggtga

Séquence codant la PhCPR2 de *P. hexandrum*

atggattctagctcggagaagttgtcgccgttcgaattgatgagcgcgatcttgaagggagctaaattagatgggtctaactcttcagatt
ctggcgtagctgtgtcgccggcagttatggctatgttgttggagaataaggagttagtgatgattttgactacttcagtggcggttttgatcg
gttgtgtcgtagttttgatatggcggcgatcttccggatcgggtaaaaaagtcgtggagcctccgaagctcatagtgcctaaatctgttgta
gaaccggaggaaattgatgaagggaagaagaaatttaccatatttttttggaacacaaactggaacagctgaaggcttcgctaaggctct
agctgaggaagccaaagctcgatatgaaaaggcagttatcaaagtgattgatatagatgattatgcggctgatgatgaagaatacgagg
agaaattcagaaaagagaccttggcatttttcatcttggccacgtatggagatggtgagccaaccgacaatgctgcaaggttctacaaat
ggtttgtagagggaaatgatagaggggactggctaaagaatctgcaatatggagtttttggccttggtaacagacaatatgagcatttcaa
caagattgctaaagtggtggatgagaaagttgctgaacagggtggtaagcggattgttccattggttctgggagacgatgaccagtgca
ttgaagatgactttgctgcatggcgtgagaatgtatggcctgagttggataacttgctccgggatgaggatgatacaactgtttctacaac
ctacactgctgctattccagaatatcgtgttgtgttccctgacaaatcagattcacttatttcagaagcaaatggccatgccaatggttatgct
aatggcaacaccgtatatgatgcccagcatccttgcagatctaatgttgcagtgaggaaggagcttcatactccagcatctgatcgttctt
gcacccatttggattttgacattgctggcactggcctttcatatggaactggagatcatgttggagtgtactgtgataatctatctgaaaccg
tggaggaggctgagagattactgaatttacccccagaaacttatttctcgcttcatgctgataaagaggatggaaccccacttgctggga
gctcattgcctcctcctttcccaccttgtactctaagaaccgccctcactcgttatgcagatctcttaaatactcctaagaagtctgctttgtta
gctctagcagcttatgcatctgatccaaatgaggccgatcgtctaaaatatcttgcttctccagccggaaaggatgaatatgctcagtcact
agttgcaaatcagagaagcctcctcgaggtcatggctgaatttccatcagcaaagcctcctcttggagtattctttgcagcaattgctccac
gcctccaacccagattctattctatatcgtcttctccaaggatggcaccatctagaattcatgtcacttgtgcacttgtttatgaaaaaacacc
tggaggacgaattcacaagggtgtgtgttcgacatggatgaagaatgccattccattggaggaaagccgtgactgcagctgggctcct
atctttgtcaggcagtctaacttcaaactccctgccgatcctaaagtgcctgttataatgatcggccctggtactggactagctcccttcag
aggattccttcaggaaagattagctctgaaggaagaaggagctgaacttggtactgcagttttcttttttggatgcaggaaccgcaaaatg
gattacatctatgaagatgagctaaaccatttccttgaaattggtgcactttccgagctacttgttgctttctcacgtgagggacccactaag
cagtatgtgcaacacaagatggcagaaaaggcttctgatatttggaggatgatttctgatggagcatatgtttacgtctgcggtgatgcca
aaggcatggccagggatgtccacagaactctccacaccattgctcaagagcagggatcgatggatagcacacaggctgagggttttgt
gaagaatctgcaaatgaccggaaggtatctccgagatgtctggtga

Séquence codant la CrCPR2 de *C. roseus*

Jennifer PERRIN

Approches d'ingénierie métabolique en levure pour la sécurisation de la production de l'anticancéreux Etoposide en Région Centre-Val de Loire.

Résumé

L'ETOPOSIDE est un lignane, dérivé de la podophyllotoxine, très couramment employé en chimiothérapie. Actuellement, sa production ne dépend que du sourcing en podophyllotoxine extraite des rhizomes de *Podophyllum hexandrum*. La pression exercée sur les ressources naturelles qui en découle occasionne des pénuries d'approvisionnement de cette molécule. Dans ce contexte, ces travaux de thèse CIFRE, menés avec le groupe industriel Axyntis, visent à développer une alternative de synthèse par des approches d'ingénierie métabolique via la création de cellules usines de levures productrices d'un précurseur d'ETOPOSIDE. Ils reposent plus spécifiquement sur la création d'une souche de levure capable de bioconvertir la yatéine (YAT, un lignane précurseur de podophyllotoxine) hautement biodisponible chez de nombreuses espèces végétales en (-)-4'-déméthyl-épipodophyllotoxine (4'dEPT) directement utilisable pour la production d'ETOPOSIDE. Pour ce faire, les trois gènes assurant la conversion de YAT en 4'dEPT chez *P. hexandrum* et codant une dioxygénase et deux cytochromes P450 (CYP71BE54 et CYP82D61) ont été intégrés, via le système CRISPR/Cas9, dans le génome de la souche *Saccharomyces cerevisiae* CEN-PK713D compatible avec une utilisation industrielle. L'optimisation de l'activité de bioconversion de la souche a été axée sur l'ajustement de la balance des copies de gènes transférés et sur l'amélioration de l'activité catalytique de CYP82D61 décrite pour catalyser l'étape d'hydroxylation terminale de la 4'dEPT à faible rendement. Parmi les stratégies testées, l'utilisation de gènes de synthèse à codons optimisés pour favoriser la traduction en levure et la co-expression d'une cytochrome P450 réductase spécifique de *P. hexandrum* ont permis une optimisation majeure du flux métabolique permettant la bioconversion de la YAT fournie aux levures en culture en 4'dEPT. En parallèle, un protocole d'extraction et de purification de YAT à partir des parties vertes de plants de *Juniperus* a été mis en place et est en cours de transfert à échelle industrielle au sein d'Axyntis. A terme, la souche de levure lui sera également transférée pour assurer des cultures en bioréacteurs de grande capacité afin de sécuriser l'approvisionnement en ETOPOSIDE via la relocalisation de production en région Centre Val de Loire. Enfin, une souche productrice de tri-hydroxycinnamoyl spermidines, phénolamides accumulés dans le manteau pollinique, a également été générée et permettra d'étudier plus profondément les fonctions et les applications potentielles de ce type de composés.

Mots-Clefs : Ingénierie métabolique, levure, bioconversion, podophyllotoxine, étoposide, bioréacteur

Summary

ETOPOSIDE is a lignan derived from podophyllotoxin, which is widely used in chemotherapy. Currently, its production only relies on the podophyllotoxin extracted from *Podophyllum hexandrum* rhizomes. The pressure on natural resources resulting from this supply causes recurrent shortages of this molecule. In this context, this CIFRE PhD thesis, carried out with the industrial group Axyntis, aims at developing an alternative approach of synthesis through metabolic engineering and the creation of yeast factory cells producing an ETOPOSIDE precursor. More specifically, this work is based on the creation of a yeast strain capable of bioconverting yatein (YAT, a lignan precursor of podophyllotoxin) highly bioavailable in many plant species into (-)-4'-desmethyl-epipodophyllotoxin (4'dEPT) directly usable for the ETOPOSIDE production. To his end, the three genes ensuring the conversion of YAT into 4'dEPT in *P. hexandrum*, encoding a dioxygenase and two cytochromes P450 (CYP71BE54 and CYP82D61) were integrated via the CRISPR/Cas9 system into the genome of the *Saccharomyces cerevisiae* strain CEN-PK713D compatible with an industrial use. To optimize bioconversion efficiency of this strain, we have mainly focused on adjusting the balance of transferred genes copies and on the improvement of CYP82D61 catalytic activity described to catalyze the terminal hydroxylation step of 4'dEPT with a low yield. Among the tested strategies, the use of codon-optimized synthetic genes to promote translation in yeast and the co-expression of a cytochrome P450 reductase specific to *P. hexandrum* allowed a major improvement of the metabolic flow thus allowing a very efficient bioconversion of the YAT supplied to the yeast in culture into 4'dEPT. On the other hand, a protocol for the extraction and the purification of YAT from the green parts of *Juniperus* plants has been implemented and is being transferred to an industrial scale at Axyntis. At midterms, the strain will also be transferred to Axyntis to perform the yeast culture in high capacity bioreactors in order to secure the ETOPOSIDE supply by relocating the production to the Centre-Val de Loire region. Finally, a last yeast strain synthesizing tri-hydroxycinnamoyl spermidines, which correspond to phenolamides compounds accumulated in the pollen coat, was also generated in order to initiate the detailed study of their *in planta* functions but also their potential industrial applications.

Keywords : Metabolic engineering, yeast, bioconversion, podophyllotoxin, etoposide, bioreactor