HAL
open science

Identification et caractérisation fonctionnelle de gènes impliqués dans la voie de biosynthèse des furocoumarines chez les végétaux supérieurs

Guilhem Vialart

To cite this version:

Guilhem Vialart. Identification et caractérisation fonctionnelle de gènes impliqués dans la voie de biosynthèse des furocoumarines chez les végétaux supérieurs. Agronomie. Université de Lorraine, 2012. Français. NNT : 2012LORR0390 . tel-01750069

HAL Id: tel-01750069

https://hal.univ-lorraine.fr/tel-01750069

Submitted on 29 Mar 2018

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

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

UNIVERSITÉ DE LORRAINE

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

UMR 1121, Université de Lorraine-INRA, RP2E
Laboratoire Agronomie et Environnement

Thèse de recherche

Identification et caractérisation fonctionnelle de gènes impliqués dans la voie de biosynthèse des furocoumarines chez les végétaux supérieurs.

Présentée pour l'obtention du titre de

Docteur de l'Université de Lorraine
En Sciences Agronomiques

Guilhem VIALART

Thèse soutenue publiquement le 06 avril 2012, devant un jury composé de :

Mr. P.Hugueney : Directeur de recherche INRA Colmar. Rapporteur.
Mr. H. Schaller : Directeur de recherche au CNRS à IBMP Strasbourg. Rapporteur.
Mme N. Guivarc'h : Professeur. Université François Rabelais Tours. Examinateur.
Mr. E. Gelhaye : Professeur. Université de Lorraine. Nancy. Examinateur
Mr. F.Bourgaud : Professeur, Université de Lorraine, Directeur de thèse. Nancy
Mr. A. Hehn : Maitre de conférences, Université de Lorraine, Co-Directeur. Nancy

Remerciements

Je souhaite remercier Mr Hugueney Philippe, directeur de recherche de INRA Colmar, ainsi que Mr Schaller Hubert, chargés de recherche au CNRS à l'IBMP de Strasbourg, d'avoir accepté d'être les rapporteurs de cette thèse. Je remercie aussi Gelhaye Eric, professeur à la faculté des sciences de l'université de Lorraine, et Mme Guivarc'h Nathalie, professeur à l'université de François Rabelais de Tours, d'avoir accepté d'être les examinateurs de cette thèse. Je les remercie d'avoir jugé mon travail.

Je tiens à remercier l'ancien directeur du Laboratoire Agronomie et Environnement, le Pr. Sylain Plantureux, de m'avoir accueilli au sein de son équipe, lorsqu'il était encore directeur.

Je remercie évidemment le nouveau directeur du Laboratoire Agronomie et Environnement, le Pr Frédéric Bourgaud, qui m'a « conservé » dans son laboratoire. Par cette occasion, j'en profite pour te remercier cette fois-ci en tant que directeur de thèse pour tout ce que tu as fait pour moi. Déjà en tant qu'étudiant de l'école d'ingénieur ENSAIA, j'appréciais les cours « d'art et création » que tu nous prodiguais à travers la musique et la peinture. Et je suis content d'avoir pu te connaître un peu plus durant ces trois années de thèse. Je n'oublierai jamais le fabuleux voyage au pays « du soleil levant », dont tu es en grande partie responsable. De même que les francs moments de surprise et de rigolade qu'on a passé ensemble à découvrir les coutumes de ce pays. Ils restent et resteront un souvenir important à mes yeux. Sans oublier le séminaire à Manchester où l'équipe de choc était au rendez-vous. Bien entendu, je te remercie de m'avoir accordé ta confiance, ton soutien quand les résultats des manips étaient laborieux (microsomes…) et aussi ton enthousiasme quand, au contraire, tout fonctionnait bien. Un grand merci Fred.

Mes remerciements vont aussi à mon codirecteur de thèse, alias Alain Hehn, qui a été d'un très grand soutien durant ces années de thésard. Merci maître Alain pour tous les conseils techniques que tu as pu me donner au laboratoire et pour m'avoir initié à la magie de la biologie moléculaire. Tu as su me transmettre la passion qui t'anime chaque jour au laboratoire, et je dois dire que, si ces trois années m'ont paru aussi rapides, c'est en partie

grâce à toi. Tu as aussi fait partie du voyage au Japon et j'ai beaucoup apprécié ton regard innocent sur cette culture totalement dépaysante. Le goût du voyage va-t-il aussi finir par te tomber dessus ? Dans tous les cas, merci pour toute l'aide que tu m'as apportée, que ce soit sur le plan professionnel comme personnel! J'ai sincèrement beaucoup apprécié de travailler avec toi et d'avoir été ton thésard. Et j'ai même fini dans ton bureau! Si ce n'est pas une preuve ça!

La collaboration effectuée avec l'équipe des professeurs Sugimotto et Mizutani n'aurait jamais pu être aussi riche tant sur un point de vue professionnel que sur un point de vue personnel, sans cet accueil formidable qu'ils m'ont réservé. Je ne remercierai jamais assez toute cette équipe de m'avoir fait découvrir les mystères de la culture japonaise.
Je remercie tout particulièrement le professeur Mizutani pour son humour et sa gentillesse à toute épreuve (je me souviendrai des yeux de poissons…quel farceur!). Je remercie aussi Kyoko qui m'a épaulé dans le Labo, mais aussi dans ma vie de tous les jours. Et franchement, les premiers pas au Japon ne sont pas si faciles pour un non-initié : Merci, merci, merci…Sans toi, je ne serai peut-être jamais revenu du Japon (Passeport égaré)!

Je remercie ensuite la Dream-team des mecs en furie du Labo (Ben, Alex, Romain, Sandro) avec qui j'ai vraiment bien rigolé (et travaillé aussi à l'occasion) : un peu plus matures que moi, selon eux … surtout plus vieux selon moi!! Au niveau des blagues, je ne crois pas être le plus « gamin »…n'est-ce pas Alex ? C'était vraiment cool cette ambiance quotidienne que j'ai un peu de difficulté à décrire d'ailleurs… c'est peut-être mieux ainsi….« Holalalala, tu l'as vu celle-là?...Où ça ?? »

Merci aussi à tous les permanents du labo, ceux qui étaient là avant moi et qui y seront après moi : merci Bernard, Sophie, Séverine, Chhoy, Christophe, Armand, Francoise, Jean, Nadia, Beatrice pour votre accueil.

De même, je remercie aussi tous les vieux thésards de l'époque qui sont devenus maintenant des post-doc et qui m'ont appris plein de choses durant ces années.
Un grand merci particulièrement à Flore qui a eu la patience de supporter mes blagues vaseuses et mon « énergie » qui doit parfois être fatigante.

Évidemment, je remercie tous les thésards qui ont partagé mon bureau pendant un temps. Merci Fazeelat pour toutes les discussions qu'on a pu avoir concernant le Pakistan et ta culture (au fait, ne répète pas tous les gros mots que je t'ai appris!!!) ; merci aussi à Matilde d'avoir supporté mon état électrique dès le matin. Un merci aussi à Celia, la madame plante du bureau, qui a contribué à « végétaliser » notre bureau…

Je remercie aussi tous les autres jeunes thésards, Aline, Simon, Nico, Abdelaak, Charles, et je leur souhaite une bonne fin de thèse.

Je n'oublierai pas non plus toute l'équipe de PAT SAS et notamment Sissi, Cindy, Nath, Paul, Estelle, Léonor, Jean-Paul, Jean-marc, Hinde.

De même je remercie aussi Martine, Ségolène, Aude, Élodie, Claude, pour leur aide précieuse dans la préparation du matériel au laboratoire. Tout comme les différents stagiaires que j'ai pu encadrer durant cette thèse. Merci a Björn et Adeline.

Merci aussi à Thamara de ne pas m'avoir laissé tomber dans les profondeurs obscures des papiers administratifs et de me rappeler, toujours en temps et en heure, les petits trucs que j'avais oublié de signer ou de compléter!

Je souhaiterais aussi remercier la bande d'Hérouville-Saint-Clair! « wouaich »! Les copains et copines de primaire et de collège qui ont toujours été présents durant ces longues années!! Il y en a des souvenirs à raconter jusqu'à nos 30 ans !!!
Merci Greg, Pichon, Flo, Sophie, Yanous, Marioul, Guidoo, Valentine, Laurène, Erwan, Vinz, Stéphane… 13 ans déjà qu'on fête chaque année le réveillon du jour de l'an ensemble, et ce n'est pas près de se terminer!!!

Merci aussi à la Pharma-team avec Romano, Pierro, Lolo, Vivi, Nono, Dalinda, Cahuetos, Franchess, Matmut… mais j'aurai l'occasion de vous remercier à nouveau dans ma prochaine thèse de pharma! La suite dans un prochain épisode donc…

Un grand merci aussi à la carabine-team avec Franky, Marcouillé, Lolo…vivement qu'on se refasse des sorties wakeboard et qu'on retrouve notre cygne!!!

Nancy n’aurait sans doute pas eu la même saveur si je n’avais pas rencontré ici aussi des super amis qui savent faire la fête comme il se doit!!
Un grand merci a vous, Torlot et Jimos, mes anciens colocataires et précieux amis.
Merci aussi à mes compagnons de sortie, de soirée, de ski vosgien, de musique, de pêche, de sport et j’en passe… Merci Tonio, Ben, Olivier, Nico, Aude, Jéjé, Caro, Trusty, Camille…Petite dédicace à Jeannot qui devra me rendre la coupe de squash et de tennis dès que j’aurai soutenu cette thèse! T’es foutu mec!
Merci à tous de m’avoir fait découvrir cette région où il fait bon vivre!!!

Mes prochains remerciements s’adresseront ensuite à ma famille : mes Mamys et Papys que j’embrasse de tout cœur là-haut.
Mes tontons et tatas, cousins et cousines qu’il me tarde de voir très rapidement!
Et bien entendu, merci Maman et Papa de m’avoir toujours soutenu, pour votre réconfort et votre amour! Vous en avez du mérite, car je sais que je ne suis pas un petit ange venu du ciel!! On peut dire que j’en ai fait des bêtises depuis trente ans déjà!! Malgré cela, vous avez toujours eu confiance en moi et vous êtes des parents géniaux! Je vous aime très fort et je suis fier de vous moi aussi! Peut-être que cette thèse me fera grandir un peu qui sait ? Mais, faut pas trop rêver non plus… hi hihihi…
Merci aussi à Stéphane et Josiane, mes beaux-parents depuis déjà si longtemps, et qui ont su apporter réconfort à mon estomac dans les moments les plus difficiles avec le saumon fumé fait maison de folie et la mousse de thon inimitable (ce n’est pas faute d’avoir essayé).
Merci aussi à Alex et Daphné, mon futur beau frère qui « kiffe » le surf et ma future belle sœur!

Enfin, je terminerai par la personne qui a le plus de mérite dans toutes celles déjà citées. Celle qui partage ma vie déjà depuis plus de 13 ans (et ça ne doit pas être de la tarte tous les jours, il parait que j’ai le sang chaud). Celle dont je suis le plus fier et celle dont mon cœur n’a cessé de battre depuis notre première rencontre. Je ne pourrai jamais assez te remercier mon « lapinou » (alias Julie, et oui j’ai balancé ton surnom sur un manuscrit officiel), d’avoir été à mes cotés dans les moments les plus difficiles, de m’avoir encouragé et souvent permis de faire les bons choix tout au long de ces nombreuses années. Tu es ma seconde moitié et je ne serai jamais devenu ce que je suis sans toi! Vivement le 26 mai prochain! Carnaval forever!!!

Sommaire

Liste des figures 15
Liste des tableaux 19
Liste des photos 21
Liste des annexes 22
Abréviations 23
Chapitre I : Synthèse bibliographique 29
I Les furocoumarines 29
I.1 Structures et propriétés chimiques 29
I.2 Distribution dans le règne végétal 30
I.3 Rôle, localisation et stockage 31
I.3.1 Protection de la plante 31
I.3.2 Localisation des furocoumarines dans la plante 32
I.3.3 Stockage des furocoumarines dans la cellule 32
I.4 Biosynthèse 33
I.4.1 Voie de biosynthèse des phénylpropanoïdes 33
I.4.2 Voie de biosynthèse des coumarines 35
I.4.3 Voie de biosynthèse des furocoumarines 40
I.5 Induction de la synthèse des dérivés coumariniques 44
I.6 Propriétés biologiques des furocoumarines 45
I.6.1 Réaction de cyclisation ou photocycloaddition 45
I.6.2 Réaction avec les acides nucléiques 47
I.6.3 Réaction avec les acides gras insaturés 49
I.6.4 Réaction avec les protéines 50
I.7 Utilisations thérapeutiques 50
I.7.1 La photochimiothérapie 50
I.7.2 La photophérèse 52
I.7.3 Traitement antidépresseur 53
I.7.4 Activités bactériostatiques, fongicides et antimycobatérienne. 53
I.7.5 Activité antiproliférative et apoptotique 54
I.8 Toxicités des furocoumarines 54

I.8.1 Photodermatite 54
I.8.2 Mélanome 55
I.8.3 Toxicité alimentaire par inhibition de cytochrome P450 55
II Les cytochromes P450 57
II.1 Présentation 57
II.2 Nomenclature 59
II.3 Classification 59
II.4 Localisation des cytochromes P450 61
II.5 Séquences et domaines conservés 62
II.6 Structure 64
II.7 Cycle catalytique des P450 66
II.8 Rôles et intérêts des P450 67
II.8.1 Un champ de recherche très large 67
II.8.2 Fonctions des P450 chez les végétaux 68
II.8.2.1 Biosynthèse de métabolites secondaires 69
II.8.2.2 Catabolisme de molécules exogènes 69
II.8.3 Les cytochromes P450 à l'origine de l'adaptation entre les plantes et les insectes 71
II.8.3.1 Adaptation et synthèse de composés allélopathiques chez les plantes 72
II.8.3.2 P450 dans le génome des plantes 73
II.8.3.2.1 Organisation 73
II.8.3.2.2 P450 et synthèse de composés allélopathiques 74
II.8.3.3 Détoxication des composés toxiques par les insectes 75
II.8.3.3.1 Diversité, spécificité et régulation des P450 d'insectes 75
II.8.3.3.2 Relations structures-fonctions des p450 d'insectes 78
II.8.3.4 P450 et coévolution 78
III Dioxygénases 81
III.1 Différentes familles de dioxygénases 81
III.2 Les dioxygénases α-cétoglutarates-dépendantes 81
III.2.1 Une diversité de réactions 82
III.2.1.1 Modification de protéines 82
III.2.1.2 Réparation de l'ADN alkylé 83
III.2.1.3 Biosynthèse des antibiotiques 84
III.2.1.4 Le métabolisme des lipides 85

III.2.1.5 Biosynthèse des métabolites secondaires 87
III.2.1.5.1 Les phénylpropanoïdes 87
III.2.1.5.1.1 Les flavonoïdes 87
III.2.1.5.1.2 Les coumarines 89
III.2.1.5.2 Les dérivées terpéniques : les gibbérellines 89
III.2.1.5.3 Les composés azotés : les alcaloïdes 91
III.2.2 Relation structure-activité 93
III.2.2.1 Modélisation tridimensionnelle 93
III.2.2.2 Site actif 95
III.2.2.3 Mécanisme réactionnel 97
III.2.3 Évolution moléculaire des dioxygénases. 100
III.2.3.1 Exemple de la voie de biosynthèse des flavonoïdes 100
III.2.3.2 Exemple des glycosinolates chez *Arabidopsis thaliana* 105
IV Phénomène de duplication génique 109
IV.1 Mécanismes aboutissant à la duplication génique 109
IV.1.1 Enjambement (ou crossing-over) inégal 110
IV.1.2 Rétroposition 112
IV.2 Devenir des gènes dupliqués 113
IV.2.1 Pseudogénisation 113
IV.2.2 Copies redondantes du gène 114
IV.2.3 Acquisition de nouvelles fonctions 114
IV.2.3.1 L'hypothèse de la néo-fonctionnalisation. 114
IV.2.3.2 L'hypothèse de la sub-fonctionnalisation 115
IV.3 Exemples dans le métabolisme secondaire des plantes 117
IV.3.1 Duplications géniques par enjambement inégal 117
IV.3.2 Duplication génique par rétroposition 119
IV.3.3 Les Opérons-like. 121
V Objectif de la thèse 123
Identification et caractérisation de gènes impliqués dans la voie de biosynthèse des furocoumarines. 123
Chapitre II : Matériel et méthodes 125
I Matériel 125
I.1 Matériel végétal 125
I.1.1 Citrus limetta 125

I.1.2 Ruta graveolens 126
I.1.3 Pastinaca sativa 127
I.1.4 Arabidopsis thaliana. 127
I.1.5 Nicotiana benthamiana 129
I.2 Souches bactériennes 129
I.2.1 *E. coli* Top 10 129
I.2.2 *E. coli* XLI Blue 130
I.2.3 *E. coli* GenHogs®[invitrogen] 130
I.2.4 E.coli BM25.8 130
I.2.5 *E.coli* BL21(DE3) 130
I.2.6 *E.coli* HB101 (pRK2013) 130
I.2.7 *E.coli* ccdB survival 131
I.2.8 *Agrobacterium tumefaciens* LBA4404 C58C1RifR (pGV2260) 131
I.3 Souche de levure 131
I.3.1 S. cerevisiae WAT11 131
I.3.2 S. cerevisiae WAT21 132
I.4 Vecteurs et plasmides recombinants 132
I.4.1 pCR8/GW/TOPO® (Invitrogen) 132
I.4.1.1 Description 132
I.4.1.2 Vérification du sens d'insertion du gène dans le vecteur 133
I.4.2 pET-28a 133
I.4.3 pBin-GW 135
I.4.3.1 Présentation 135
I.4.3.2 Recombinaison dans un vecteur de destination pBin (Technologie Gateway) 136
I.4.4 pYeDP6O 136
I.4.5 pTriplEx2 (Clontech) 137
I.4.6 λTriplEx2 (Clontech) 138
I.4.7 pQE30 (Qiagen) 139
II Milieu de culture 140
II.1 Milieu de culture pour bactéries 140
II.1.1 LB liquide 140
II.1.2 LB solide 140
II.1.3 LB $MgSO_4$ solide 140

II.1.4 Gélose molle 140
II.1.5 YEB 141
II.2 Milieu de culture pour levures 141
II.2.1 YPGA 141
II.2.2 SGI 141
II.2.3 Milieu glucosé (YPGE) 142
II.2.4 Milieu galactosé (YPL) 142
II.3 Conditions et milieux de culture pour le végétal 143
II.3.1 Culture de *Ruta graveolens* en condition *in vitro* 143
II.3.2 Murashige et Skoog (MS) 143
II.3.3 Culture des plantes en sol 144
II.3.4 Terreau 144
II.4 Antibiotiques 144
III Outils de biologie moléculaire 145
III.1 Extraction d'ADN génomique de plante 145
III.2 Extraction D'ARN de plante 145
III.3 Amplification par PCR d'un fragment d'ADN. 145
III.3.1 Amplification classique 145
III.3.2 PCR par essais 147
III.4 Reverse Transcriptase PCR 147
III.5 Digestion par des enzymes de restrictions 148
III.6 Ligations 148
III.6.1 PCR8/TOPO/TA/GW® 148
III.6.2 Dans les autres vecteurs 149
III.7 Électrophorèse sur gel d'agarose 149
III.8 Électrophorèse sur gel polyacrylamide en condition dénaturante ou SDS PAGE 150
III.9 Extraction d'ADN à partir d'un gel d'agarose 150
III.10 Préparation de bactéries compétentes 151
III.10.1 Bactéries électrocompétentes 151
III.10.2 Bactéries chimiocompétentes 151
III.11 Transformation de bactéries compétentes 151
III.11.1 Électroporation 151
III.11.2 Choc thermique 152

III.11.3 Extraction d'ADN plasmidique 152
III.12 Construction et criblage d'une banque d'ADNc 152
III.12.1 Construction de la banque d'ADNc 152
III.12.2 Titration de la banque 153
III.12.3 Étalement de la banque d'ADNc 153
III.12.4 Transfert et fixation sur membrane du nylon 153
III.12.5 Synthèse de la sonde 154
III.12.6 Hybridation 154
III.12.7 Lavage des membranes 155
III.12.8 Révélation 155
III.12.9 Récupération des plages de lyse 156
III.12.10 Criblage secondaire 156
III.12.11 Excision des plasmides 157
III.13 Banque de cDNA normalisée 157
III.14 Séquençage 157
III.15 Induction des plantes aux UV 157
III.16 Inhibition de la synthèse de furocoumarines par le Regalis 158
III.17 Modélisation tridimensionnelle des dioxygénases 159
IV Système d'expression hétérologue 159
IV.1 Système d'expression dans les procaryotes 160
IV.1.1 La bactérie, un système inductible 160
IV.1.1.1 Expression *in vitro* 160
IV.1.1.1.1 Culture 160
IV.1.1.1.2 Purification de la protéine 160
IV.2 Système d'expression dans les eucaryotes type levure 161
IV.2.1.1 Préparation de levures compétentes 161
IV.2.1.2 Transformation de levures 161
IV.2.1.3 Expression *in vitro* 161
IV.2.1.3.1 Culture 162
IV.2.1.3.2 Préparation de microsomes 162
IV.3 Système d'expression hétérologue dans les plantes 163
IV.3.1 Expression constitutive 163
IV.3.1.1 Obtention des plasmides recombinants pBin-X 163
IV.3.1.2 Conjugaison triparentale 163

IV.3.1.3 Inoculum d'agrobactérie utilisé pour la transformation génétique de plantes 164
IV.3.1.4 Transformation génétique de *Ruta graveolens* 165
IV.3.1.4.1 Germination des graines de *Ruta graveolens L* 165
IV.3.1.4.2 Transformation d'hypocotyles de *Ruta graveolens* 165
IV.3.1.5 Transformation d'*Arabidopsis thaliana* 167
IV.3.1.5.1 Germination des graines 167
IV.3.1.5.2 Transformation des inflorescences d'*Arabidopsis thaliana* 167
IV.3.2 Expression transitoire 168
IV.3.2.1 Principe de l'expression transitoire 168
IV.3.2.2 Préparation de l'inoculum bactérien 168
IV.3.2.3 Agroinfiltration de *Nicotiana benthamiana* 169
V Méthode d'analyse biochimique 170
V.1 Dosage de protéines 170
V.2 Quantification des P450 par spectre CO 170
V.3 Mesure d'activités enzymatiques 171
V.3.1 Détermination des constantes cinétiques 171
V.3.1.1 Dioxygénase 171
V.3.1.2 P450 171
V.3.2 Test et mesure de l'inhibition par les furocoumarines 171
V.4 Analyse en HPLC 172
V.4.1 Esters de CoA 172
V.4.2 Furocoumarines 172
V.4.3 Extraits coumariniques 172
V.5 Analyse en spectrométrie de masse 173
V.6 Synthèses des substrats 173
V.6.1 Formation des esters de CoA 173
V.6.1.1 Production de la 4- coumarate CoA ligase 173
V.6.1.2 Réaction de synthèse 174
V.6.2 Suivi cinétique de la réaction chimique 174
VI Méthode analytique 175
VI.1 Extraction des composés coumariniques 175
VI.2 Quantification du niveau d'expression par PCR en temps réel 175
VI.2.1 Quantification relative 175

VI.2.2 Préparation du matériel 176
VI.2.3 Mélange réactionnel 176
VII Traitement statistique des données 177
Chapitre III : Résultats 179
I Identification des gènes codant pour les dioxygénases chez les différentes plantes modèles 179
I.1 Recherche de gènes candidats 179
I.1.1 Approche bibliographique 180
I.1.2 Approche bio-informatique et recherche d'orthologues des F6'H sur les banques informatiques publiques et privées par tBlastn. 182
I.1.3 Approche expérimentale 189
I.1.3.1 Approche PCR 189
I.1.3.1.1 Diox Citrus 189
I.1.3.1.2 Pastinaca sativa 190
I.1.3.2 Criblage de la banque de *Ruta graveolens* 191
I.1.3.2.1 Choix des sondes et méthodologie 192
I.1.3.2.2 Caractéristiques des gènes isolés 192
I.1.3.2.2.1 Comparaison avec les gènes déjà isolés 192
I.1.3.2.2.2 Multiplicité de dioxygénases chez la Rue : Réalitée ou artefact ? 195
I.2 Discussion 203
I.3 Conclusion 206
II Caractérisation de gènes codant pour des dioxygénases dans des systèmes d'expressions hétérologues. 209
II.1 Expression hétérologue dans un système procaryote : *E.coli* BL21(DE3) 209
II.1.1 Clonage des dioxygénases dans le plasmide pET28a 210
II.1.2 Production et purification des protéines recombinantes 211
II.1.3 Tests d'incubations 212
II.2 Caractérisation de la Diox 4 215
II.2.1 Constantes cinétiques de la D4 (ou C2'H) 215
II.2.1.1 Paramètres cinétiques 215
II.2.1.2 Inhibition compétitive de la D4 217
II.2.2 Relation entre la teneur en coumarines et furocoumarines et la Diox 4 (C2'H). 218

II.2.2.1 Analyse de la composition en furocoumarines chez des plantes de Rue 219
II.2.2.2 PCR en temps réel sur des ARN extraits de Rue 220
II.2.2.3 Inhibition par le Regalis® 221
II.3 Caractérisation de la Diox 4 dans un système eucaryote *in planta* 223
II.3.1 Expression transitoire 223
II.3.1.1 Transformation de Nicotiana benthamiana 223
II.3.1.1.1 Test préliminaire avec la Green Fluorescent Protein ou GFP 223
II.3.1.1.1.1 Clonage de la GFP dans le pBin 223
II.3.1.1.1.2 Préparation de l'inoculum d'agrobactérie 223
II.3.1.1.1.3 Infiltration et visualisation en binoculaire 224
II.3.1.1.1.4 Co-infiltration de la GFP avec la protéine P19 225
II.3.1.1.2 Infiltration de la Diox 4 226
II.3.1.2 Extraction coumarinique et identification par SM 226
II.3.1.3 Dosage par HPLC 227
II.3.2 Expression stable 229
II.3.2.1 Transformation de *Ruta graveolens* 229
II.3.2.2 Transformation Arabidopsis thaliana 230
II.4 Discussion 231
II.5 Conclusion 239
III Identification et tentatives de caractérisation des cytochromes P450. 241
III.1 Recherche de P450 candidats 241
III.1.1 Les collaborations 241
III.1.1.1 Collaboration avec l'équipe du professeur Matern 241
III.1.1.1.1 Présentation 241
III.1.1.1.2 Méthode d'identification des P450 candidats 241
III.1.1.1.3 Modification de l'extrémité N-terminale de CYP71D98 et CYP71AZ1. 242
III.1.1.2 Collaboration avec l'équipe du professeur Simonsen 243
III.1.2 Les bonnes raisons pour étudier ces différents P450 candidats 243
III.1.2.1.1 CYP71AJ5, le meilleur candidat. 243
III.1.2.1.2 CYP71AZ1, un grand potentiel. 248
III.1.2.1.3 CYP71D97 et CYP71D98, les outsiders 251
III.1.2.1.4 CYP73A43 (C4H) et CYP98A21 (C3'H), une alternative à la C2'H ? 252

III.1.2.1.5 CYP82H1 et CYP76B8, les peu probables 253
III.2 La levure : un système plébiscité, mais pas optimisé pour des P450 de plantes « exotiques »..... 254
III.2.1 Clonage dans le plasmide pYeDP60 254
III.2.2 Transformations des levures..... 255
III.2.3 Amélioration de l'expression des P450 dans la levure 255
III.2.3.1 Amélioration du niveau d'expression..... 256
III.2.3.1.1 T°C induction 256
III.2.3.1.2 Durée induction 257
III.2.3.2 Amélioration de la préparation de microsomes..... 257
III.2.3.2.1 Broyage mécanique automatisé des levures 257
III.2.3.2.2 Précipitation par ultracentrifugation..... 258
III.2.3.3 Composition des milieux..... 258
III.2.3.4 Analyse de l'expression fonctionnelle des P450 260
III.2.3.5 Criblage métabolique..... 262
III.2.3.5.1 CYP71D97..... 262
III.2.3.5.2 CYP71AJ5..... 263
III.3 Discussion 264
III.4 Conclusion..... 269
Chapitre IV : Conclusion générale et perspectives 271
I Conclusion générale 271
II Perspectives..... 274
Chapitre V : Références bibliographiques 277
Chapitre VI : Annexes..... 319

Liste des figures

Figure 1 : Structure chimique de certaines furocoumarines linéaires et angulaires. 29
Figure 2: Voie simplifiée de la biosynthèse des phénylpropanoïdes et de ses principaux dérivés. 35
Figure 3 : Représentations des différents dérivés de coumarines. 36
Figure 4 : Voie de biosynthèse d'une coumarine, la scopolétine. 39
Figure 5 : Formation de l'umbelliférone et de la scopolétine. 42
Figure 6 : Comparaison de la voie de biosynthèse de la scopolétine et de l'hypothétique voie de biosynthèse des furocoumarines linéaires et angulaires. 43
Figure 7 : Photodimérisation des composés coumariniques (Kitamura et al., 2005). 46
Figure 8 : Photocycloaddition du psoralène avec un groupement éthylène substitué. 47
Figure 9 : Illustration de la formation de liaisons chimiques entre le 8-metoxypsoralen (8-MOP) et les bases pyrimidiques des différents brins de l'ADN (Kitamura et al., 2005). 48
Figure 10 : Mécanisme proposé d'activation de la mélanogénèse par la photoaddition des furocoumarines sur les acides gras insaturés (Zarebska et al., 2000). 49
Figure 11 : Exemples de diversité réactionnelle de P450 (Mansuy, 1998). 58
Figure 12 : Partenaires de transfert d'électrons des Cytochromes P450 (Paine et al., 2005)... 60
Figure 13 : Représentation schématique d'un cytochrome P450 microsomale d'eucaryote et de la NADPH P450 réductase (Werck-Reichhart and Feyereisen, 2000). 61
Figure 14 : Séquences protéiques consensus distales (G/A)Gx(E/D)T et proximale FxxGx(R/H)xCxG « cys-pocket » conservées chez tous les P450s 62
Figure 15 : Représentation tridimensionnelle des différentes régions SRS. 64
Figure 16 : Représentation simplifiée d'une coupe transversale de la structure d'un Cytochrome P450. 65
Figure 17 : Représentation du noyau protoporphyrine IX. 66
Figure 18 : Mécanisme réactionnel des cytochromes P450. 67
Figure 19 : Domaines de recherches des P450 et les différents champs d'applications. 68
Figure 20 : Schéma simplifié du processus de détoxication. 70
Figure 21 : Détoxication des furocoumarines par des P450 identifiés chez des insectes. 77
Figure 22 : Réaction générale effectuée par une dioxygénase α-cétoglutarate dépendante. 82

Figure 23 : La première dioxygénase Fe(II)/α-cétoglutarate dépendante identifiée sur le plan moléculaire : La Prolyl 4-hydroxylase. 83
Figure 24 : Intervention de l'AlkB, une dioxygénase Fe(II)/α-cétoglutarate dépendante, dans le processus de réparation de l'ADN alkylé. 84
Figure 25 : Hydroxylation du β-lactam par la clavaminate synthase. 85
Figure 26 : Réaction de cyclisation (A) et de désaturation (B) par la clavaminate synthase dans la synthèse de l'acide proclavaminique. 85
Figure 27 : Intervention de la phytanoyl CoA -2-hydroxylase dans le catabolisme de l'acide phytanique. 86
Figure 28 : Représentation d'une partie de la voie de biosynthèse des flavonoïdes. 88
Figure 29 : Synthèse de la scopolétine par la Féruloyle CoA 6'- hydroxylase (F6'H). 89
Figure 30 : Oxydation du carbone C7 par une hydroxylase Fe(II)/α-cétoglutarate dépendante : la gibbérelline 7-oxydase. 90
Figure 31 : Synthèse de gibbérelline par une dioxygénase Fe(II)/α-cétoglutarate dépendante multifonctionnelle : la gibbérelline 20-oxydase. 90
Figure 32 : Etape de la synthèse de vindoline par une dioxygénase Fe(II)/α-cétoglutarate dépendante. 91
Figure 33 : Intervention de la hyoscyamine 6β-hydroxylase (H6B) dans la synthèse de scopolétine chez *Datura metel L.* 92
Figure 34 : Voies de biosynthèses des alcaloïdes morphiniques chez *Papaver somniferum L.* 93
Figure 35 : Modelisation des dioxygénases α-cétoglutarates-dépendantes. 94
Figure 36 : Séquences consensus H-X-D/E-X_n-H et R-X-S de trois dioxygénases Fe(II)/α-cétoglutarate- dépendantes. 95
Figure 37 : Localisation tridimensionnelle des séquences consensus H-X-D/E-X_n-H et R-X-S sur la dioxygénase α-cétoglutarate-dépendante (At3g13610). 96
Figure 38 : Structure du site actif « in-line » A et « off-line » B. 97
Figure 39 : Mécanisme d'action proposé pour les dioxygénases Fe(II)/α-cétoglutarate dépendante 98
Figure 40 : Mécanisme d'action proposé pour les dioxygénases Fe(II)/α-cétoglutarate dépendantes 99
Figure 41 : Voie de biosynthèse des flavonoïdes. 101
Figure 42 : Arbre phylogénétique des dioxygénases α-cétoglurates-dépendantes identifiées dans la voie de biosynthèse des flavonoïdes. 103

Figure 43 : Arbre phylogénétique des dioxygénases α-cétoglutarates-dépendantes (AOP1, AOP2 et AOP3) isolées chez différents écotypes *d'Arabidopsis thaliana* (Ler, Cvi , Col) et *A. lyrata* (Kliebenstein et al., 2001). 106
Figure 44 : Représentation simplifiée d'un enjambement et d'un enjambement inégal entre deux chromosomes homologues. 111
Figure 45 : Duplication génétique par rétroposition. 113
Figure 46 : Modèle d'évolution pour l'acquisition de nouvelles fonctions enzymatiques par l'hypothèse de néo-fonctionnalisation (Inspiré de (Ober, 2010)). 115
Figure 47 : Modèle d'évolution pour l'acquisition de nouvelles fonctions enzymatiques par l'hypothèse de sub-fonctionnalisation (Inspiré de (Ober, 2010)). 117
Figure 48 : Duplication génique du gène codant à l'origine pour la sous unité α du complexe tryptophane synthase chez le maïs. 118
Figure 49 : Construction d'un arbre phylogénétique à partir des différents CYP98A incluant CYP98A8 et CYP98A9 (Matsuno et al., 2009). 120
Figure 50 : Système de ligation dans le PCR8 (Invitrogen)..... 132
Figure 51 : Polylinker du plasmide pCR8/GW/TOPO®..... 133
Figure 52 : Carte physique du plasmide pET-28a (Novagen)..... 134
Figure 53 : Carte physique du plasmide pBin-GW (Biteau, 2009)..... 135
Figure 54 : Schéma représentant la LR recombinaison..... 136
Figure 55 : Carte physique du plasmide pYedP60..... 137
Figure 56 : Carte physique du pTriplEx2 (Clontech)..... 138
Figure 57 : Carte physique du plasmide pQE30 (Clontech). 139
Figure 58 : Métabolisation du CDP-Star..... 156
Figure 59 : suivi de la synthèse cinétique du p-coumaroyle CoA (A), féruloyle CoA (B), cinnamoyle CoA (C) et caféoylee CoA (D). 174
Figure 60 : Comparaison des mécanismes réactionnels mis en jeu entre la synthèse d'umbelliférone et de scopolétine. 181
Figure 61 : Alignement de la F6'H1 d'Arabidopsis thaliana (AT3G13610) et des deux fragments de gène (GenBank CX667249 and CX667579) identifiés chez un hybride *Citrus paradisi x Poncirus trifoliata.* 183
Figure 62 : Alignement des trois gènes candidats (Diox Panais, Diox Panais2 et Diox Panais3) identifiés dans notre banque d'EST privée de *Pastinaca sativa* avec les F6'H isolée chez *Arabidposis.*..... 184

Figure 63 : Alignement des séquences peptidiques des dioxygénases isolées chez *Arabidopsis thaliana* (F6'H1 et F6'H2) avec les séquences identifiées par tBlastn sur les banques publiques et privées. 187
Figure 64 : Construction d'un arbre phylogénétique à partir de différentes dioxygénases cétoglutarates-dépendantes caractérisées. 188
Figure 65 : Alignement des séquences peptidiques des Diox 1/2/4/8 isolées chez *Ruta graveolens*. 194
Figure 66 : Construction d'un arbre phylogénétique à partir de différents membres appartenant à la famille des dioxygénases cétoglutarates-dépendantes déjà caractérisées. ... 195
Figure 67 : Identification de deux acides aminés différents sur la Diox 2. 196
Figure 68 : Identification d'un acide aminé différent sur la Diox 1 197
Figure 69 : Position de l'intron dans les différentes dioxygénases identifiées chez *Citrus limetta*, *Ruta graveolens* et *Pastinaca sativa*. 199
Figure 70 : Séquences consensus des sites d'épissages des introns d'eucaryotes. 200
Figure 71 : Alignement des différentes séquences codantes (notées ORF) et génomiques (notées intron) sur ClustalX de toutes les dioxygénases potentielles identifiées chez *Citrus limetta*, *Ruta graveolens* et *Pastinaca sativa*. 202
Figure 72 : Caractérisation cinétique de la Diox 2 nvl (A) et de la Diox 8 (B) isolées chez *Ruta graveolens*. 213
Figure 73 : Caractérisation fontionnelle de la Diox 4 isolée chez *Ruta graveolens*. 215
Figure 74 : Représentation des courbes cinétiques de type michaeliennes de la Diox 4 en fonction de la concentration (5 à 200 µM) en féruloyle CoA (trait plein) et en *p*-coumaroyle CoA (pointillé). 216
Figure 75 : Estimation du Ki du psoralène par la représentation de Kitz Wilson. 218
Figure 76 : Analyse quantitative de la teneur en umbelliférone dans les feuilles de Ruta graveolens. 220
Figure 77 : Quantification du niveau d'expression de la Diox 4 réalisée sur les feuilles, les pétioles, les tiges et les racines de *Ruta graveolens* ayant subit ou non un traitement de 24 h aux UV. 221
Figure 78 : Expression transitoire de la GFP dans les feuilles de *Nicothiana benthamiana*. 224
Figure 79 : Co-expression transitoire de la GFP et de la protéine P19 dans les feuilles de *Nicothiana benthamiana*. 225
Figure 80 : Identification de l'umbelliférone et de la scopolétine en spectrométrie de masse chez *Nicotiana benthamiana* surexprimant la Diox 4. 227

Figure 81 : Dosage de la quantité d'umbelliférone et de scopolétine exprimé en mg.g^{-1} de matière fraiche chez les feuilles de *Nicotiana benthamiana* issues de l'expression transitoire. 228
Figure 82 : Modélisation tridimensionelle de la Diox 8 à l'aide du logiciel MOE. 235
Figure 83 : Arbre phylogénétique de différents P450 caractérisés incluant les gènes candidats appartenant à la famille CYP71AJ et CYP71AZ. 246
Figure 84 : Alignement des séquences peptidiques des différents membres de la famille des CYP71AJ1-4 et du nouveau CYP71AJ5 isolé chez *Thapsia garganica*. 247
Figure 85 : Comparaison du mécanisme de cyclisation lors de la synthèse de menthofurane et de la marmésine. 249
Figure 86 : Comparaison du niveau d'expression relatif de *CYP71AJ1* par rapport au profil d'accumulation du psoralène après induction du milieu de culture cellulaire d'*Ammi majus* par *Phytophtora megasperma (Larbat et al., 2007).* 251
Figure 87 : Représentation des hydroxylations des acides cinnamiques catalysées par la cinnamate 4-hydroxylase (C4H), la p-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H) et la ferulate 5-hydroxylase (F5H). 252
Figure 88 : Spectres différentiels au CO enregistrés avec les microsomes de levures transformés avec pYe71D97 (A) et (B) et pYeAZ1-nt (C). 261
Figure 89 : Chromatogramme obtenu après incubation de CYP71D97 avec la 5-0Hflavone. 263
Figure 90 : Chromatogramme obtenu après incubation de CYP71AJ5 avec le bergapol et le bergaptène. 264

Liste des tableaux

Tableau 1 : Exemple de P450 impliqués dans la biosynthèse de métabolites secondaires chez les végétaux. 69
Tableau 2 : P450 exprimés en fonction du régime alimentaire pratiqué par l'espèce d'insectes (Hung et al., 1995; Petersen et al., 2001; Wen et al., 2006b; Mao et al., 2007; Mao et al., 2008). 76
Tableau 3 : Composition du milieu LB 140
Tableau 4 : Composition du milieu YEB. 141
Tableau 5: Composition des milieux de culture pour levures 142

Tableau 6 : Composition du milieu MS. 143
Tableau 7 : Concentration finale des antibiotiques utilisés. 144
Tableau 8 : Les différentes étapes d'une PCR classique. 146
Tableau 9 : Les différentes étapes d'une PCR par essais. 147
Tableau 10 : Compositions des tampons utilisés pour le criblage de banque. 154
Tableau 11 : Composition des tampons de lavages. 155
Tableau 12 : Composition des tampons pour la préparation de microsomes de levures. 163
Tableau 13 : Amorces utilisées pour évaluer l'expression des différentes dioxygénases chez la Rue. 177
Tableau 14 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différentes dioxygénases isolées chez *Arabidopsis thaliana* (F6'H1 et F6'H2) par Kai et ses collaborateurs (Kai et al., 2008), et les candidats identifiés dans les banques d'EST publiques [Diox Citrus (EST)] et privées [Diox Panais (EST)]. 185
Tableau 15 : Comparaison des séquences peptidiques entre 4 dioxygénases Fe/α-cétoglutarates-dépendantes. 186
Tableau 16 : Amorces spécifiques générées à partir de la séquence identifiée *in silico* chez l'hybride *Citrus paradisi x Poncirus trifoliata*. 189
Tableau 17 : Amorces utilisées pour amplifier la séquence codante pour une dioxygénase Fe/α-cétoglutarate-dépendante chez *Pastinaca sativa*. 191
Tableau 18 : Pourcentage d'identité nucléotidique entre les différentes dioxygénases candidates identifiées. 193
Tableau 19 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différents clones de dioxygénases candidates identifiées. 193
Tableau 20 : Amorces utilisées pour récupérer les clones de la Diox 1 et de la Diox 2 sans mutation chez *Ruta graveolens*. 197
Tableau 21 : Amorces utilisées pour amplifier les séquences génomiques des différentes dioxygénases identifiées chez *Ruta graveolens* (Diox 1/2/4/8), chez *Citrus limetta* (Diox Citrus) et chez *Pastinaca sativa* (Diox Panais). 198
Tableau 22: Identification des acides aminés flanquant les introns dans chacune des dioxygénases identifiées. 199
Tableau 23 : Amorces utilisées pour amplifier les séquences codantes des dioxygénases Fe/α-cétoglutarates-dépendantes identifiées chez *Citrus limetta* (Diox Citrus), *Ruta graveolens* (Diox Ruta) et *Pastina sativa* (Diox Panais) 210

Tableau 24 : Récapitulatif des différentes protéines produites par le système bactérien *E. coli* BL21(DE3). 211
Tableau 25 : Résultats des incubations des différentes dioxygénases purifiées (Diox Citrus, Diox 8/4/2) avec les différents acides cinnamiques et leurs formes estérifiées. 212
Tableau 26 : Constantes cinétiques de la Diox 4 calculées à l'aide du logiciel Sigmaplot.... 216
Tableau 27 : Constantes cinétiques de la Diox 4 calculées à l'aide du logiciel Sigmaplot.... 217
Tableau 28 : Analyse quantitative en $mg.g^{-1}$ de la teneur en composés furocoumariniques chez *Ruta graveolens* après traitement par le Regalis®. 222
Tableau 29 : Couple d'amorces utilisé pour la mutagenèse dirigée contre l'acide aminé valine (Val^{218}) et leucine (Glu^{104}) de la Diox 8. 236
Tableau 30 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différents clones isolés de la famille des CYP71AJ. 245
Tableau 31 : Pourcentage d'identité des séquences peptidiques codant pour les SRS des différents membres de la famille des CYP71AJ. 248
Tableau 32 : Comparaison des identités et homologies peptidiques des clones CYP71AZ2 et CYP71AZ3 identifiés dans la banque d'EST de *Pastinaca sativa* et du clone CYP71AZ1 identifié chez *Ammi majus*. 250
Tableau 33 : Amorces utilisées pour amplifier les séquences codantes pour les cytochromes P450 ; CYP71D97, CYP71D98-nt et CYP71AZ1-nt. 255
Tableau 34 : Composition des milieux de culture selon le protocole décrit par Pompon pour les levures transformées avec le plasmide pYeDP60. 258
Tableau 35 : Composition des milieux de culture du SGI, du SGI modifié et du YPL modifié suivant les recommandations de Jiang et Morgan (Jiang and Morgan, 2004). 259
Tableau 36 : Conditions de culture permettant l'expression de CYP71D97 et CYP71AZ1-nt. 261

Liste des photos

Photo 1 : Exemples de plantes productrices de furocoumarines. 30
Photo 2 : Indications de la puvathérapie. 51
Photo 3 : *Citrus limetta.* 125
Photo 4 : *Ruta graveolens.* 126
Photo 5 : *Pastinaca sativa.* 127
Photo 6 : *Arabidopsis thaliana* 128

Photo 7 : Nicotiana benthamiana. 129
Photo 8 : Induction des plantes aux UV. 158
Photo 9 : Vaporisation d'un inhibiteur spécifique des dioxygénases (le Regalis®) sur les parties aériennes des plantes de *Ruta graveolens*. 159
Photo 10 : Germination de graine de *Ruta graveolens*. 165
Photo 11 : Transformation d'hypocotyles de *Ruta graveolens*. 166
Photo 12 : Repiquage des plantes transformées. 167
Photo 13 : *Arabidopsis thaliana*. 168
Photo 14 : *Nicotiana benthamiana*. 169
Photo 15 : Expression transitoire chez *Nicotiana benthamiana*. 170
Photo 16 : Sélection des explants d'hypocotyles de *Ruta graveolens* sur milieu sélectif contenant de la kanamycine pendant un à deux mois. 229
Photo 17 : Plantule de *Ruta graveolens in vitro* potentiellement transformée. 230

Liste des annexes

Annexe 1 : Extraction de l'ADN de plantes suivant le protocole de « DNeasy plant minik de Qiagen » 319
Annexe 2 : Extraction d'ARN de plantes suivant le protocole « RNeasy plant minikit de Qiagen ». 320
Annexe 3 : Préparation de cDNA par reverse transcription suivant le protocol « high capacity RNA-to-cDNA master mix » d'Applied Biosystems 321
Annexe 4 : RT-PCR en une seule étape suivant le protocole « SuperScript™ One-Step RT-PCR with Platinum® *Taq* » d'Invitrogène 321
Annexe 5 : Extraction d'ADN à partir d'agarose suivant le protocle QIAquick Gel Extaraction Kit de Qiagen 322
Annexe 6 : Extraction plasmidique suivant le protocol du « kit Genelute™ plamside miniprep » de Sigma-Aldrich 323
Annexe 7 : Synthèse de la sonde d'hybridation suivant le protocole décrit dans le kit «Amersham Gene Images AlkPhos Direct Labelling and Detection System ». 324
Annexe 8 : Purification des protéines suivant le protocole décrit dans le kit « Ni-NTA Purification System » de Qiagen (Annexe 8). 325
Annexe 9 : Criblage métabolique des différents P450 candidats 326
Annexe 10 : Publications scientifiques 327

Abréviations

2HN : 2-hydroxy-1-naphtoate

3PA : acide trans-3-(pyrid-4-yl)-acrylique

4CL : 4-hydroxycinnamoyle CoA ligase

5-MOP : 5-méthoxypsoralène ; bergaptène

5-8-MOP : 5-8-méthoxypsoralène ; isopimpinelline

8-MOP : 8-méthoxypsoralène ; xanthotoxine

ADN : acide désoxyribonucléique

ADNc : ADN complémentaire

ANS : antocyanidine synthase

ARN : acide ribonucléique

AS : Angélicine synthase

ATP : adénine tri-phosphate

BAP : 6- Benzylaminopurine

BMT : bergaptol O-methyltransférase

BX : benzoxazinoides

C2'H : cinnamate/*p*-coumarate-2-hydroxylase

C3'H : p-coumaroyleshikimate/quinate 3'-hydrolxylase

C4H : cinnamate-4-hydroxylase

CAS : clavaminate synthase

CCM : chromatographie sur couche mince

CCoAOMT1 : caféoylee CoA o-methyltransférase

CO : monoxyde de carbone

CoA : coenzyme A

CODM : codeine O-demethylase

CPR : cytochromes P450 réductase

CYP : cytochrome P450

DAG : diacylglycérol

DAOCS : deacetoxycephalosporine C synthase

DACS : deacetocephalosporine C synthase

DHF : dihydroflavonol

DIBOA : 2,4-dihydroxy-2H-1,4-benzoxazin-3(4H)-one

DIMBOA : 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one

DIG : digoxygénine

DMAPP :diméthyl-allyl-pyrophosphate

DMS : déméthylsubérosine

DMSO : diméthylsulfoxyde

DO : Densité optique

EDTA : acide ethylène diamine tétraacétique

EGF : epidermal growth factor

EOR : espèce oxygénée réactive

EST : expressed sequence tag : étiquette de séquneces exprimées

F2H : Férulate-2-hydroxylase

F3'H : flavonoïde-3'-hydroxylase

F3'5'H : flavonoïde-3'-5'-hydroxylase

F5H : férulate-5-hydroxylase

F6'H : féruloyle CoA 6-hydroxylase

FAD : flavine adénine dinucléotide

FHT : flavanone 3β- hydroxylase

FLS : flavonol synthase

FMN : flavine mononucléotide

FNS I : flavone synthase I

FNS II : flavone synthase II

Flav6OH : flavonoïde-6-hydroxylase

HPLC : High performance liquid chromatography / chromatographie liquide haute performance

HHS : hypothalamo-hypophyso-surrénalien

IGL : indole-3-glycerol phosphate lyase

IGP : indole-3-glycerol phosphate

IPNS : isopenicillin N synthase

IPP : isopentényl-pyrophosphate

IPTG : isopropyl-beta-D-thiogalactopyranoside

LB : Luria Bertani

MAO : monoamine oxidase

MBI : mechanism based inactivation

MS : Murashige et Skoog

NADH : nicotinamide adénine dinucléotide, forme réduite

NADP : nicotinamide adénine dinucléotide phosphate, forme oxydée

NADPH : nicotinamide adénine dinucléotide phosphate, forme réduite

NO : oxyde d'azote

PAL : phénylalanine ammonia-lyase

PCR : polymerase chain raction

PDB : protein data bank

PEG : polyéthylène glycol

PKC : protéine kinase C

PLA_2 : phospholipase A2

PS : psoralen synthase

P-UVA : psoralène-UVA-thérapie

EOR : Espèces réactives de l'oxygène.

RE : Réticulum endoplasmique

RDX : cyclotriméthylènetrinitramine

RMN : résonance magnétique nucléaire

SDS : sodium dodécyl sulfate

SRS : substrate recognition site. site de reconnaissance du substrat

T6ODM : thebaine 6-O-déméthylase

TAL : tyrosine ammonia-lyase

TEA : tris acétate EDTA

TMP : 4,5',8- trimethylpsoralène

UV : Ultra violet

Répertoire des espèces citées

(et noms vernaculaires lorsqu'ils existent)

Acacia mangium : Acacia

Agrobacterium tumefaciens : Agrobactérie (responsable de la galle du collet)

Allium cepa : Echalotte

Ammi majus : Queen Anne's Lace

Angelica archangelica : Angélique officinale, herbe aux anges, herbe du Saint-Esprit

Apis mellifera (insectes) : Abeille à miel

Apium graveolens : Céleri

Arabidopsis thaliana : Arabette des dames

Arabidopsis lyrata : Arabette à feuilles lyrées

Aspergillus niger

Avena strigosa : Avoine rude, avoine maigre

Brassica napus : Chou-navet blanc

Candida albicans

Catharanthus roseus : Pervenche de Madagascar

Cinnamomum verrum : Cannelier de Ceylan, cannelle de Ceylan, cinnamome

Citrus bergamia : Bergamotier

Citrus limetta : Citronnier doux, limettier, bergamotier de Tunisie

Citrus paradisi : Pamplemoussier

Citrus sinensis : Oranger

Citrus unshiu : Mandarinier Satsuma

Coronilla varia: Coronille

Coumarouna odorata : Gaiac de Cayenne

Cryptococcus neoformans :

Cucurbita maxima : Potiron

Datura metel : Métel, herbe diable

Dianthus caryophyllus : Oeillet commun, œillet des fleuristes

Drosophila melanogaster : Mouche du vinaigre

Drosophila pastinacella

Escherichia coli : Colibacille

Ficus carica : Figuier

Gibberella fujikuroi

Glycine max : Soja

Heracleum lanatum : Berce laineuse

Heracleum mantegazzianum : Berce du Caucase

Helianthus annuus : Tournesol, grand-soleil, soleil des jardins, soleil commun

Helianthus tuberosus : Topinambour

Helicoverpa zea : Ver de l'épi du maïs

Hierochloe odorata : Avoine odorante, hiérochloé odorant

Hydrangea macrophylla : Hortensia à tête plate

Hyoscyamus niger : Jusquiame noire

Lotus japonicus : Lotier corniculé

Lactuca sativa : Laitue, Laitue cultivée

Medicago sativa : Luzerne cultivée

Melilotus alba : Mélilot blanc

Melilotus officinalis : Mélilot jaune

Mycobacterium fortuitum

Mycobacterium phlei

Mycobacterium smegmatis

Nicotiana benthamiana

Notopterygium incisum

Oryza sativa : Riz

Papaver somniferum : Pavot somnifère

Papilio glaucus : Papillon glauque

Papilio polyxenes : Papillon du céleri

Pastinaca sativa : Panais

Petroselinum crispum : Persil

Petunia hybrida : Pétunia

Phaseolus aureus : Soja

Pythium aphanidermatum

Phytophtora megasperma

Pseudomaonas aeruginosa : Bacille pyocyanique

Psoralea canescens

Psoralea cinerea : Psoralée

Psoralea corylifolia : Graine de Psoralea, thé malais

Radix Astragali

Ruta graveolens : Rue officinale

Saccharomyces cerevisiae : Levure de boulanger

Salmonella typhimurium : Salmonelle

Steptomyces antibioticus

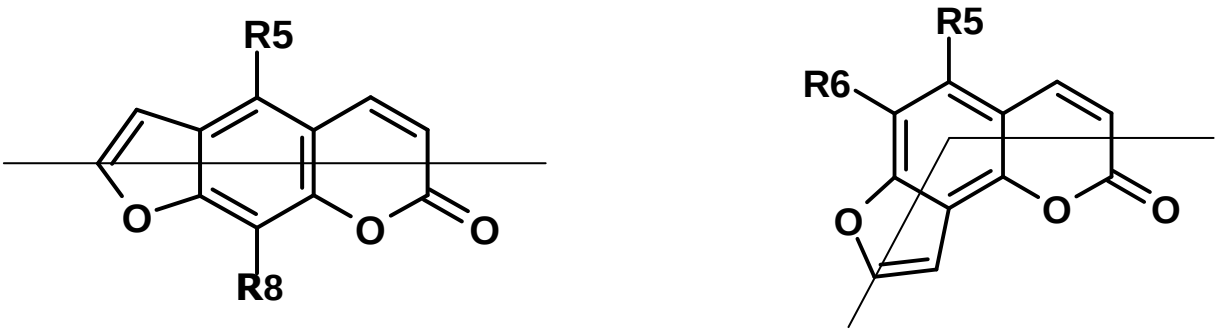
Zea mays : Maïs

Chapitre I : Synthèse bibliographique

I Les furocoumarines

I.1 Structures et propriétés chimiques

Les furocoumarines sont des molécules issues du métabolisme secondaire des plantes appartenant à la famille des phénylpronaoïdes. Ces molécules résultent de la condensation de deux hétérocycles : coumarines (benzo-α-pyrone) et furanes. Cette structure tricyclique est caractérisée par la position du cycle furane qui permet de distinguer deux types de furocoumarines : les furocoumarines linéaires ou angulaires (Figure 1). De plus, la présence de différents groupements hydroxyle, alkyle, méthoxyle sur les hétérocycles confèrent un large panel de molécules différentes à cette famille (Bourgaud et al., 2006).



Psoralène: R5, R8 : H
Bergaptène (5-MOP) : OMe
Xanthotoxine (8-MOP) : R8: OMe
Isopimpinelline (5-8-MOP) :R5,R8: OMe

Angelicine : R5, R6: H
Sphondine : R6: OMe
Pimpinelline : R5, R6: OMe

Figure 1 : Structure chimique de certaines furocoumarines linéaires et angulaires.

Les cycles aromatiques constituant ces molécules attribuent à celle-ci une forte absorbance des longueurs d'onde comprises entre 250 et 320 nm. Ces molécules présentent une meilleure solubilité dans les solvants aprotiques apolaires tels que le chloroforme, ainsi que dans certains solvants aprotiques polaires tels que le demethylesufoxide (DMSO).

Par contre, elles sont très peu solubles dans les solvants protiques polaires tels que l'eau, le méthanol ou l'éthanol.

Il est possible de caractériser de manière précise les différentes furocoumarines par des techniques de chromatatographie sur couche mince (CCM), de chromatographie liquide haute performance (CLHP) ou de spectrométrie de masse (SM) (Spencer et al., 1987; Kang et al., 2010) .

I.2 Distribution dans le règne végétal

Les furocoumarines se retrouvent principalement dans 4 familles de plante: Les Fabacées, les Apiacées, les Rutacées et les Moracées (Pathak et al., 1962) (Photo 1).

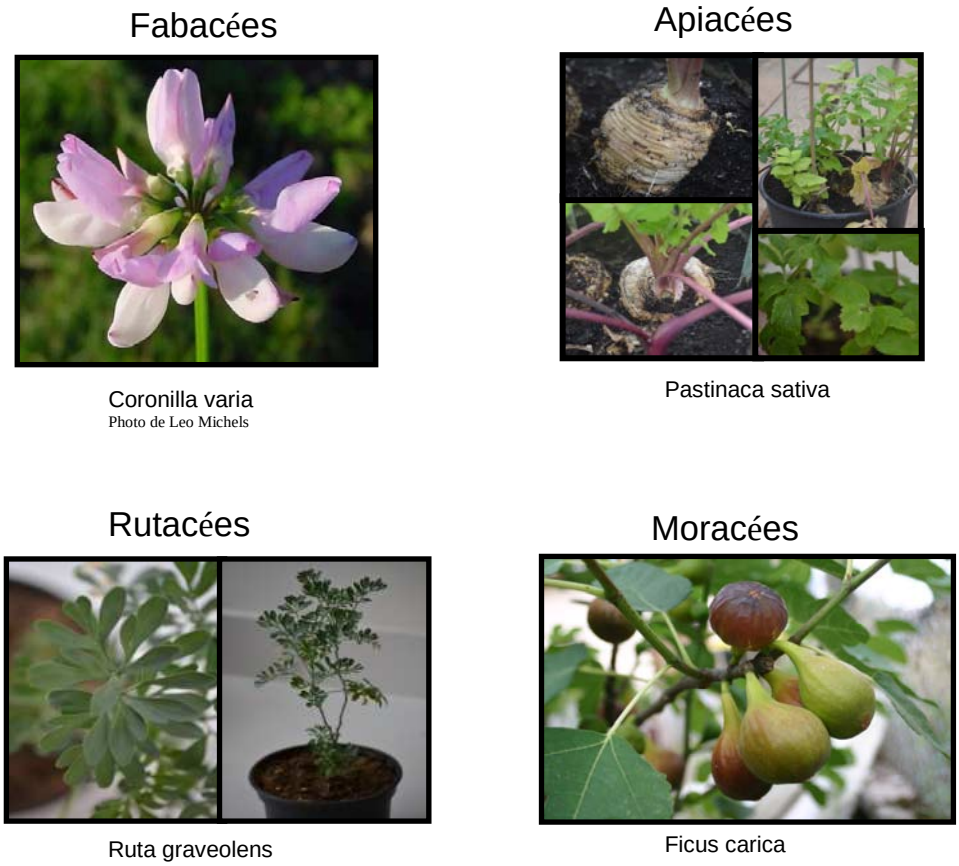


Photo 1 : Exemples de plantes productrices de furocoumarines.
Ces espèces appartiennent à la famille des Fabacées (*Coronilla varia)*, des Apiacées (*Pastinaca sativa*), des Rutacées (*Ruta graveolens*) et des Moracées (*Ficus corilla*).

Certaines familles de plantes ne synthétisent que des furocoumarines linéaires telles que les Moracées (*Coronilla varia*, *Ficus carica*), les Rutacées (*Ruta graveolens*, *Citrus bergamia*…)

et certaines espèces de la famille des Apiacées (*Ammi majus*) (Damjanic and Akacic, 1974; Abyshev et al., 1992; Gattuso et al., 2007; Pokrovskii et al., 2009). D'autres familles de plantes synthétisent des furocoumarines linéaires et angulaires comme la famille des légumineuses (genres *Psoralea et Coronilla*) et certaines espèces de la famille des Apiaceés (*Apium graveolens*, *Pastinaca sativa*) (Bourgaud et al., 1994; Innocenti et al., 1997; Sah et al., 2006).

Les concentrations et les différentes formes de furocoumarines sont variables d'une famille et d'une espèce à l'autre (Girennavar et al., 2008a). À titre d'exemples, les Rutacées ne synthétisent que des furocoumarines linéaires dont les molécules principalement retrouvées sont le bergaptène, la xanthotoxine et l'isopimpinelline. Au contraire, les Fabacées dont *Psoralea canescens* synthétisent principalement 2 types de molécules : une furocoumarine linéaire, le psoralène, et une furocoumarine angulaire, l'angélicine (Bourgaud F., 1989; Innocenti et al., 1997).

I.3 Rôle, localisation et stockage

I.3.1 Protection de la plante

Les furocoumarines sont des phytoalexines. La phototoxicité naturelle des furocoumarines est principalement à l'origine de la protection des plantes contre les herbivores, les insectes et autres microorganismes pathogènes. De par leurs propriétés de phototoxicités, la plante s'est dotée au cours des processus d'évolution d'un ensemble de molécules variées (furocoumarines linéaires et angulaires) de protection disponible contre les prédateurs herbivores. En effet, elles sont anti-appétentes pour les ravageurs. Elles sont phototoxiques pour les prédateurs de par leurs activités mutagènes. En effet, elles sont capables de réaliser des adduits et de se complexer avec l'ADN. Après photoactivation par un rayonnement UV, elles établissent des liaisons covalentes reliant les thymines des deux brins d'ADN, et bloquent ainsi la réplication et la traduction de l'ADN (Dall'Acqua et al., 1978b; Dardalhon et al., 1998). De plus, elles inhibent les cytochromes P450 des insectes et autres organismes qui sont habituellement impliqués dans la détoxication de composés toxiques (McCloud et al., 1992; Hadacek et al., 1994; Neal, 1994; Ojala et al., 2000). Cet ensemble d'interactions biochimiques existant entre les plantes et d'autres organismes constitue les interactions allélopathiques.

Dans la compétition des espèces végétales pour l'acquisition de l'espace et des nutriments, les furocoumarines confèrent un avantage indéniable contre les plantes voisines et concurrentes,

de par cette activité allélopathique. En effet, il a été démontré que les furocoumarines possèdent une action sur le développement des plantes. À titre d'exemple, les furocoumarines (bergaptène et xantotoxine) extraites de *Ruta graveolens* sont des inhibiteurs de germination et de croissance de la laitue (*Lactuca sativa*) ou de l'oignon (*Allium cepa*) (Baskin et al., 1967; Hale et al., 2004; Razavi, 2011).

I.3.2 Localisation des furocoumarines dans la plante

Les furocoumarines sont principalement localisées au niveau des jeunes feuilles, des fruits et des graines (Pathak et al., 1962; Cappelletti et al., 1984; Zobel and Brown, 1990b; Zobel et al., 1990; Zobel et al., 1991). Le rôle de protection des furocoumarines explique principalement cette localisation au niveau des organes sujets à la prédation (Milesi, 2001). Cette synthèse de furocoumarines est inductible chez certaines espèces comme le céleri (*Apium graveolens*). Un stress environnemental ou un bioagresseur engendre rapidement la synthèse de furocoumarines chez le céleri (Chaudhary et al., 1985). Chez d'autres plantes, la synthèse de furocoumarines se fait de manière constitutive, comme chez la rue. (Eiler, 1989; Milesi et al., 2001)

I.3.3 Stockage des furocoumarines dans la cellule

Les furocoumarines sont des molécules très toxiques pour la cellule de par leur propriété mutagène qui inhibe la réplication cellulaire.

La présence de ces molécules dans les tissus de la plante peut donc devenir néfaste pour elle-même si elle accumule une trop grande quantité de composés toxiques.

Des mécanismes de stockages spécifiques des furocoumarines sont donc mis en place afin de protéger la plante. Les métabolites sont glycosylés par des glycosyl-transférases présentes dans le cytosol afin de neutraliser la toxicité naturelle des furocoumarines. Le stockage de ces molécules glycosylées se fait dans la vacuole (Zobel and Brown, 1990b; Nguyen, 1992).

Lorsque la cellule est blessée par un bioagresseur, les furocoumarines glycosylées sont immédiatement catabolisées et la forme aglycone est libérée, retrouvant ainsi son activité toxique. Cette réaction a lieu dans le cytosol par des *O*-glycosidases. Cette stratégie de défense contre une agression mécanique permet de libérer des composés toxiques de façon rapide et localisée. De plus, il est possible de trouver des furocoumarines stockées sous forme libre à l'extérieur de la cellule, dans les espaces intercellulaires ou à la surface des feuilles de

Ruta graveolens (Zobel and Brown, 1990b). Cette optimisation de la localisation de molécules de défense dans les organes reproducteurs et les parties exposées à la prédation (Brown, 1970), est interprétée comme un phénomène évolutif dans le cadre de la théorie de la « défense optimale » (Nitao and Zangerl, 1987; Milesi, 2001).

Par ailleurs, le psoralène et la xanthotoxine sont étroitement associés aux vaisseaux du xylème chez *Ruta graveolens* d'après des recherches par immunohistolocalisation (Massot et al., 2000). Cette localisation pourrait être un avantage pour une diffusion systémique des furocoumarines en réponse à une agression.

Le paradoxe de cette voie de biosynthèse provient du fait que les furocoumarines sont des inhibiteurs de P450 et que les principales enzymes permettant la synthèse des furocoumarines sont justement des P450 (Brown et al., 1960; Brown, 1970).

I.4 Biosynthèse

I.4.1 Voie de biosynthèse des phénylpropanoïdes

Les phénylpropanoïdes sont une classe de composés organiques, biosynthétisés à partir d'un acide aminé, la phénylalanine ou la tyrosine. Ce sont des métabolites secondaires comprenant un cycle benzène additionné de 3 atomes de carbone (structure C6-C3). Cette voie de biosynthèse est commune à tous les végétaux supérieurs. La synthèse de nombreux métabolites secondaires dérive directement de cette voie de biosynthèse comme l'acide salicylique, les monomères de lignines, les coumarines, les furocoumarines et les flavonoïdes (Figure 2) (Werck-Reichhart, 1995a; Boudet, 2007).

La voie de biosynthèse des phénylpropanoïdes est initiée par la conversion de la phénylalanine en cinnamate, par l'action de la phénylalanine amonia lyase (PAL). Cette enzyme clef de l'embranchement de la voie de synthèse des phénylpropanoïdes et des acides aminés aromatiques appartient à la famille des lyases, et plus spécifiquement, à la famille des enzymes catalysant les réactions de rupture de liaison carbone-azote (Bauer et al., 2011).

Le cinnamate subit ensuite une hydroxylation par la cinnamate-4-hydroxylase (C4H) pour donner le *p*-coumarate. Cette enzyme membranaire du réticulum endoplasmique appartient à la famille des cytochromes P450. Elle a été isolée, purifiée puis identifiée pour la première fois au début des années 1990 à partir de tubercule de topinambour (*Helianthus tuberosus*) (Gabriac et al., 1991) et de plantules de soja (*Phaseolus aureus*) (Mizutani et al., 1993a; Mizutani et al., 1993b). Elle forme une nouvelle famille au regard de la nomenclature des cytochromes P450, les CYP73. Depuis, de nombreux autres clones ont été identifiés chez une

grande variété de plantes supérieures (Xu et al., 2010; Park et al., 2011). La C4H est une enzyme exprimée de façon constitutive dans les tissues où la voie de biosynthèse des phénylpropanoïdes intervient. La régulation de son expression est principalement contrôlée au niveau transcriptionnel (Weisshaar and Jenkinst, 1998). On trouve d'ailleurs 28 allèles différents de la C4H chez *Acacia mangium* dont 8 engendrent une modification de la séquence peptidique (Nur Fariza et al., 2008). Les C4H sont induites par des facteurs de stress environnementaux tels que les infections, les blessures ou la lumière (Fahrendorf and Dixon, 1993; Hubner et al., 2003).

Il existe néanmoins des régulations de la C4H au niveau enzymatique : les inhibiteurs compétitifs comme le trans-3-(pyrid-4-yl)-acrylic acide (3PA) ou le 2-hydroxy-1-naphtoate (2HN) et les inactivateurs autocatalytiques (Schalk et al., 1998). Ces derniers inactivent l'enzyme après être oxydé par le P450 en établissant une liaison covalente avec l'apoprotéine.

Gravot a montré durant sa thèse (Gravot, 2002) que les C4H de plantes productrices de furocoumarines, telles que CYP73A32 de Rue ou CYP73A10 de persil, étaient peu sensibles à l'inactivation catalytique du psoralène ou du 8-MOP. Au contraire, chez les plantes non productrices de furocoumarines tel que le topinambour, la C4H (CYP73A1) est très sensible a cette inactivation autocatalytique. Il en tire deux hypothèses sur le plan évolutif, et suppose que cette inhibition catalytique résiduelle aurait soit pour fonction de permettre une activité de la C4H en présence de psoralène ou de 8-MOP, soit un rôle de régulation de la synthèse des furocoumarines par un effet seuil (Gravot et al., 2004).

Chez les monocotylédones, la déammination de la tyrosine, catalysée par la tyrosine amonia lyase (TAL), aboutit directement à la formation de *p*-coumarate (Nishiyama et al., 2010).

La molécule de *p*-coumarate obtenue par ces deux voies va ensuite subir une estérification avec du Coenzyme A (CoA), *via* la 4-hydroxycinnamoyle CoA ligase (4CL), pour donner le *p*-coumaroyle CoA. Ce composé estérifié est un précurseur de nombreuses autres voies de métabolites secondaires telles que les flavonoïdes, les furocoumarines et les monomères de lignine (Figure 2) (Rani et al., 2009).

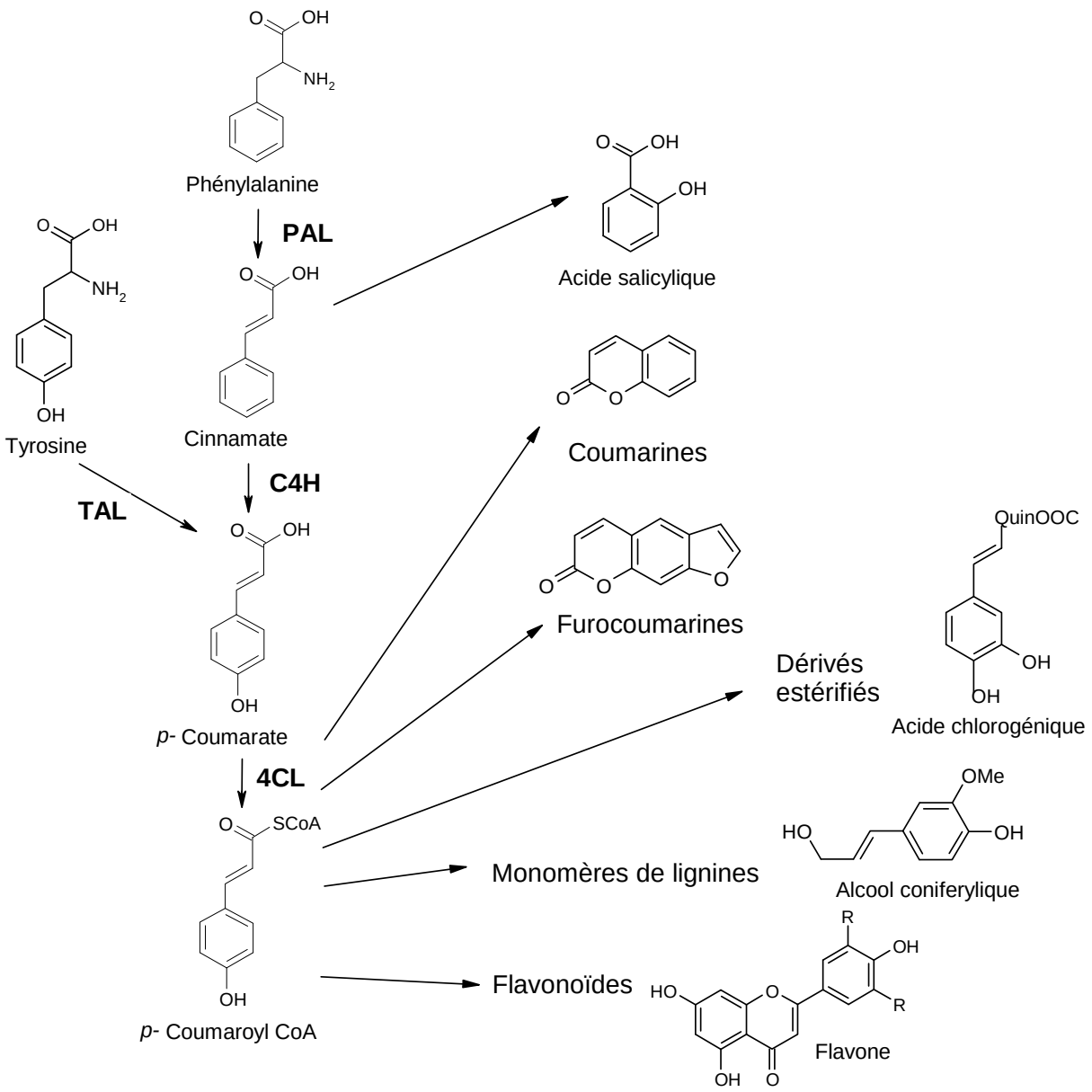


Figure 2: Voie simplifiée de la biosynthèse des phénylpropanoïdes et de ses principaux dérivés.
(TAL = Tyrosine amonia lyase, PAL = Phenyl alanine amonia lyase, C4H = Cinnamate 4 hydroxylase, 4CL = 4-Coumarate : CoA ligase)

I.4.2 Voie de biosynthèse des coumarines

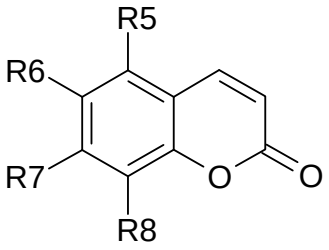
Les coumarines (ou 2H-1-benzopyran-2-one) sont des molécules distribuées dans l'ensemble des végétaux supérieurs. On les retrouve notamment chez les Apiacées, Fabacées, Rutacées, Umbelliferacées, Solanacées, Rubiacées, Asteracées, Saxifragacées, Rosacées (Keating, 1997; Zobel, 1997).

Cette famille de molécule est largement utilisée en médecine thérapeutique. En effet, les coumarines possèdent un large spectre d'activités. De nombreuses coumarines présentent des activités anticoagulante, antitumorale, antivirale, antiinflammatoire ou antimicrobienne (Riveiro et al., 2010). Elles sont impliquées dans des mécanismes de défense des plantes qui font suite à différents types de stress engendrés par les bioagresseurs, et possèdent entre autre une activité antimicrobienne (Tietjen et al., 1983; Tietjen and Matern, 1983; Hamerski and Matern, 1988b; Matern, 1991; Valle et al., 1997; Kitamura et al., 1998; Carpinella et al., 2005; Kai et al., 2008). Certaines coumarines telles que la scopolétine sont probablement impliquées dans la réduction du stress oxydatif cellulaire (Chong et al., 2002).

Elles sont également responsables d'un goût légèrement sucré chez de nombreuses variétés de plantes comme le mélilot (Kosuge, 1961; Gestetner and Conn, 1974; Oba et al., 1981; Zobel, 1997), le tabac (Stanfill et al., 2006) et de nombreux aliments (Zobel, 1997; Zhou et al., 2002).

Les coumarines dégagent une odeur vanillée que l'on retrouve chez certaines plantes appartenant à la famille des Fabacées (*Coumarouna odorata* ou *Melilotus alba*) (Ehlers et al., 1995) ou des Lauracées (*Cinnamonum verrum*) (Miller et al., 1996).

De nombreuses molécules sont déjà décrites dans la littérature, telles que la coumarine, la scopolétine, l'herniarine, la scoparone, l'esculetine, la fraxetine, l'isofraxedine et bien d'autres (Figure 3).



Coumarine: R5, R6, R7, R8: H
Scopolétine: R6: OMe, R7: OH,
Scopoline: R6: OMe, R7: O-Gl
Umbelliférone: R7: OH
Herniarine: R7: OMe
Scoparone: R6: R7: OMe
Esculetine: R6, R7: OH
Fraxetine: R6: OMe, R7, R8: OH
Isofraxidine: R6, R8: OMe, R7: OH
Isoscopolétine: R6: OH, R7: OMe
Daphnetine: R7, R8: OH
Hydrangetine: R7: OH, R8: OMe

Figure 3 : Représentations des différents dérivés de coumarines.

La voie de biosynthèse des coumarines dérive de la voie de biosynthèse des phénylpropanoïdes (Figure 2) (Brown, 1962; Kai et al., 2008; Umezawa, 2010). Les acides

cinnamiques subissent de nombreuses hydroxylations en différentes positions du noyau phénol. Après l'hydroxylation en position 4 du cinnamate par la C4H, celle réalisée en position 3' et 5 du p-coumarate et de férulate ont été démontré comme étant aussi P450 dépendantes (Gabriac et al., 1991; Meyer et al., 1996; Schoch et al., 2001b). L'hydroxylation en position 2 du *p*-coumarate, du cinnamate, du ferulate et du caféate permet de s'orienter vers la voie de biosynthèse des coumarines. Cette réaction d'hydroxylation n'a cependant jamais été véritablement caractérisée avant les travaux réalisés par Kai en 2008. En effet, il était supposé que cette réaction soit catalysée par une enzyme appartenant aussi à la famille des P450. D'ailleurs les premiers travaux relatant une activité 2-hydroxyalase ont été publiés par Kindl en 1971 chez *Hydrangea macrophylla* (Kindl, 1971). Dans cette publication, Kindl identifie une enzyme dans la fraction chloroplastique capable de catalyser la conversion du p-coumarate et du ferulate respectivement en umbelliférone et scopolétine. Une autre étude réalisée en 1974 rapporte l'activité d'*ortho*-hydroxylation du cinnamate dans les chloroplastes de *Melilotus alba* (Gestetner and Conn, 1974). Cette dernière étude est sujette à controverse car cette observation n'a jamais pu être renouvelée d'une part, et d'une autre part les techniques analytiques de l'époque ne permettaient pas de séparer les différents acides cinnamiques (U. Matern, communication personnelle). Mais au regard des enzymes déjà caractérisées impliquées dans l'hydroxylation du noyau phénolique tel que la cinnamate 4-hydroxylase (C4H), la p-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H) ou la ferulate 5-hydroxylase (F5H), on pourrait supposer que l'hydroxylation en 2' du noyau phénolique soit aussi réalisée par une enzyme appartenant à la famille des P450.

Néanmoins, deux études récentes ont ébranlé cette théorie. Une équipe japonaise s'est intéressée à la voie de biosynthèse spécifique de la scopolétine. La scopolétine et la scopoline (7-β-D glucoside de la scopolétine) sont décrites chez de nombreuses espèces végétales telles que l'arbre à caoutchouc (Silva et al., 2002), le manioc (Gomez-Vasquez et al., 2004), la carotte (Coxon et al., 1973) et le coton (Zeringue Jr, 1984). Elles ont cependant été principalement étudiées chez le tabac (Fraissinet-Tachet et al., 1998; Maier et al., 2000) et le tournesol (Cabello-Hurtado et al., 1998). Des expériences de traçage par marquage au carbone ^{14}C du ferulate et d'autres intermédiaires ont mis en évidence la formation de scopolétine à partir de la voie de biosynthèse des phénylpropanoïdes (Fritig et al., 1970).

Kai et ses collaborateurs se sont intéressés à la voie de biosynthèse de la scopolétine. Ils ont d'abord démontré le rôle primordial de la *p*-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H) et de la caféoylee CoA O-methyltransférase 1 (CCoAOMT1) chez *Arabidopsis thaliana* (Kai et al., 2008). Pour cela, ils ont réalisé des analyses quantitatives de composés

coumariniques effectuées sur des mutants d'*Arabidopsis thaliana* (Kai et al., 2008) n'exprimant pas la C3'H (mutants d'insertion). Les analyses ont révélé une importante diminution de scopolétine et scopoline dans les racines en comparaison aux plantes sauvages. Cette enzyme, qui joue un rôle très important dans la voie de biosynthèse des phénylpropanoïdes et dans la synthèse de lignine (Anterola et al., 2002; Franke et al., 2002a; Franke et al., 2002b), semble donc aussi fortement impliquée dans la synthèse de scopolétine et de scopoline chez *Arabidopsis thaliana* (Kai et al., 2008). D'autres analyses métaboliques effectuées sur des mutants d'insertion n'exprimant pas la caféoylee CoA o-methyltransférase 1 (CCoAOMT1) chez *Arabidopsisi thaliana*, ont également montré une très forte diminution en scopolétine et scopoline (Kai et al., 2008). Cette enzyme qui transforme exclusivement le caféoyle CoA en féruloyle CoA est donc très probablement impliquée dans la voie de biosynthèse de la scopolétine et de la scopoline. Ce qui suggère que la scopolétine proviendrait de l'hydroxylation en 2' du féruloyle CoA et non pas de l'acide férulique, au moins chez Arabidopsis (Kai et al., 2008).

Une seconde étude basée sur l'identification et la caractérisation de l'enzyme impliquée dans la conversion du féruloyle CoA en scopolétine a été menée en 2008 par l'équipe du professeur Mizutani (Kai et al., 2008). Des expériences de traçage par marquage en $^{18}O_2$ réalisées sur les racines de patate douce ont révélé que l'atome d'oxygène en position 1 de l'umbelliférone et de la scopolétine provenait de l'oxygène moléculaire. Seules trois grandes familles d'oxygénases réalisent une catalyse oxydative à partir de l'O_2 moléculaire : Les P450, les dioxygénases Fe(II)/α-cétoglutarate dépendante et les flavines mono-oxygenases.

La scopolétine et la scopoline sont principalement retrouvées dans les racines d'*Arabidopsis thaliana*. Cependant après traitement de la plante avec de l'acide 2,4-dichlorophenoxyacetique, l'accumulation de ces coumarines se situe aussi dans les pousses. En se basant sur la localisation de l'expression des gènes chez *Arabidopsis thaliana*, (base MPSS) deux candidats ont été identifiés. L'un appartenant à la famille des P450, l'autre à la famille des dioxygénases (Kai et al., 2008).

Une quantification des coumarines est effectuée sur des mutants d'*Arabidopsis thaliana* n'exprimant pas ces gènes candidats. Les résultats montrent une importante diminution en scopolétine et scopoline dans les racines d'*Arabidospsis thaliana* pour les mutants en

At3g13610, ORF correspondant à une dioxygénase Fe(II)/oxoglutarate dépendante. Aucune autre modification du phénotype ou d'accumulation de composés secondaires n'est identifiée. Ce résultat suggère que la teneur en scopolétine est étroitement corrélée à l'expression de l'ORF At3g13610.

L'enzyme correspondante a été nommée féruloyle CoA-6'-hydroxylase (F6'H), et hydroxyle en position 6' un composé estérifié, le féruloyle CoA (Kai et al., 2008). La F6'H intervient dans la voie de biosynthèse de la scopolétine après la CCoAOMT (Figure 4). Cette F6'H est une dioxygénase appartenant à la famille des dioxygénases Fe(II)/α-cétoglutarate dépendante.

La voie préférentielle de biosynthèse des coumarines fait donc intervenir successivement une cinnamate-4-hydroxylase (C4H), une 4-hydroxycinnamoyle CoA ligase (4CL), une *p*-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H), une caffeoyl CoA O-methyltransferase (CCoAOMT) et une féruloyle CoA-6'-hydroxylase (F6'H) (Figure 4).

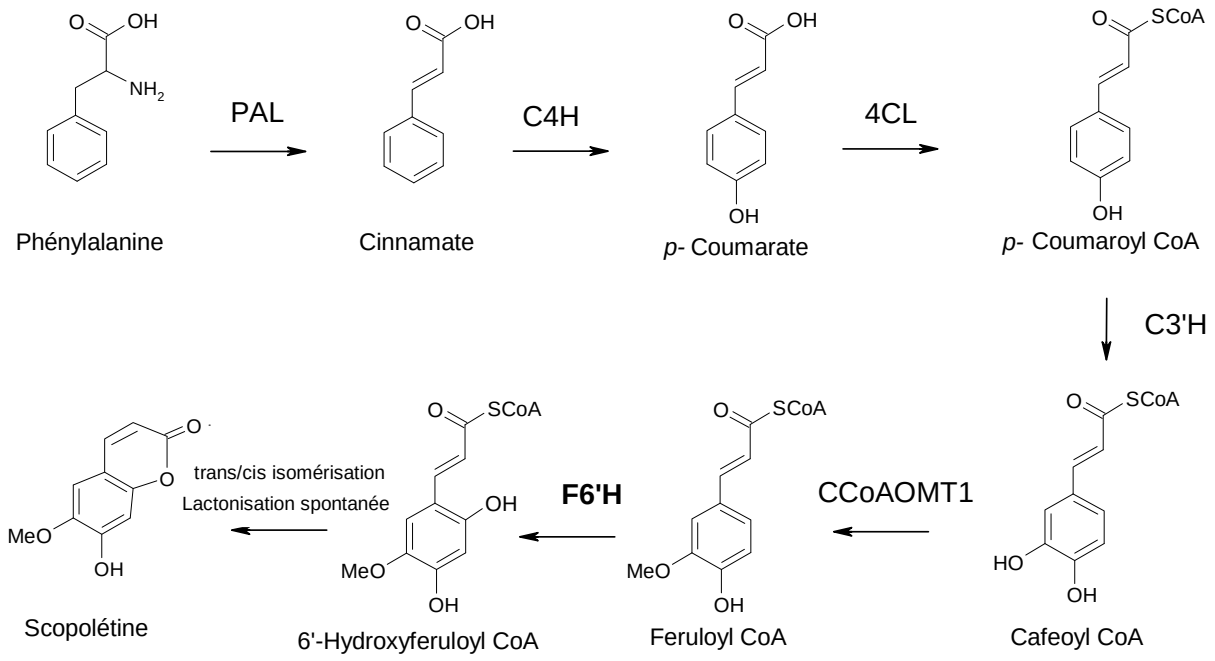


Figure 4 : Voie de biosynthèse d'une coumarine, la scopolétine.

PAL = Phenyl alanine amonia lyase, C4H = Coumarate 4 hydroxylase, 4CL = 4-Coumarate CoA ligase, C3'H = *p*-coumaroyleshikimate/quinate 3'-hydrolxylase, CCoAOMT1 = Caffeoyl CoA O-methyltransferase, F6'H = Féruloyle CoA-6'-hydroxylas

I.4.3 Voie de biosynthèse des furocoumarines

La voie de biosynthèse spécifique des furocoumarines commence à partir de la formation d'umbelliférone (7-hydroxycoumarine). Ce composé est à l'origine de la synthèse de nombreux dérivés complexes comprenant les polyhydroxy, methoxy et prenyl-coumarines, les furocoumarines et les pyranocoumarines (Bourgaud et al., 2006). Des premières expériences, basées sur le suivi et l'évolution de précurseurs isotopiques dans différents modèles de plantes, ont fourni des informations concernant la formation d'umbelliférone. En effet, les premiers résultats obtenus sur *Hierochloe odorate* et *Melilotus officinalis* (Brown et al., 1960) et un travail complémentaire effectué sur Melilotus, (Ellis and Amrhein, 1971) ont démontré que l'acide cinnamique était transformé en coumarine et en umbelliférone. De plus, d'autres analyses de marquage moléculaire réalisées sur *Lavandula officinalis* ont mis en évidence la nécessité du précurseur *para*-hydroxylé de l'acide cinnamique avant d'aboutir à l'*ortho*-hydroxylation de celui-ci (Brown, 1962). L'acide 2-4-dihydroxycinnamique provenant de l'*ortho*-hydroxylation du *cis-p*-coumarate, se transforme ensuite en umbelliférone par une réaction spontanée non enzymatique dépendante, dans des conditions de pH acide ou neutre entrainant la fermeture du cycle de la lactone. Cette réaction est très similaire à celle rencontrée dans la synthèse d'un composé coumarinique, la scopolétine (Figure 5).

La comparaison du mécanisme réactionnel mis en jeu pour la synthèse de l'umbelliférone avec celui décrit par Kai et ses collaborateurs (Kai et al., 2008) dans la formation de la scopolétine, me fait supposer qu'une étape de ligation avec un ester CoA est réalisée par une 4CL en amont de la formation de l'umbelliférone. L'*ortho*-hydroxylation se ferait sur le *para*-coumaroyle CoA (Figure 6) (3) et non sur l'acide *para*-coumarique (2) comme décrit auparavant et représenté par la flèche verte sur la Figure 6. Ensuite, une étape de prénylation par ajout d'une sous unité isoprène (diméthylallyl pyrophosphate) réalisée sur le carbone C6 ou C8 de l'umbelliférone oriente la synthèse respectivement vers les furocoumarines linéaires (6) ou angulaires (14). Une étape de cyclisation entre le groupement prényl et le groupement hydroxyle en C7 aboutit à la synthèse soit de marmésine (7) pour les furocoumarines linéaires, soit de columbianétine (15) pour les furocoumarines angulaires. La réaction catalysée par un P450 et aboutissant à la libération d'une molécule d'acétone entraine la formation de psoralène (8) ou de l'angélicine (16) (Larbat et al., 2007; Larbat et al., 2009). La marmésine synthase et la psoralène synthase sont deux cytochromes P450 localisés dans le réticulum endoplasmique chez *Ammi majus* (Hamerski and Matern, 1988a) . Une hydroxylation en position 5 ou 8 des furocoumarines linéaires et angulaires (Hamerski and

Matern, 1988b), est immédiatement suivie par la O-méthylation de ces composés hydroxylés. Seules les séquences codant pour la bergaptol O-methyltransférase (BMT), la psoralène synhase et l'angélicine synthase ont été identifiées à ce jour à partir de cellules d'*Ammi majus* (Hehmann et al., 2004) et de *Pastinaca sativa* (Larbat et al., 2007; Larbat et al., 2009).

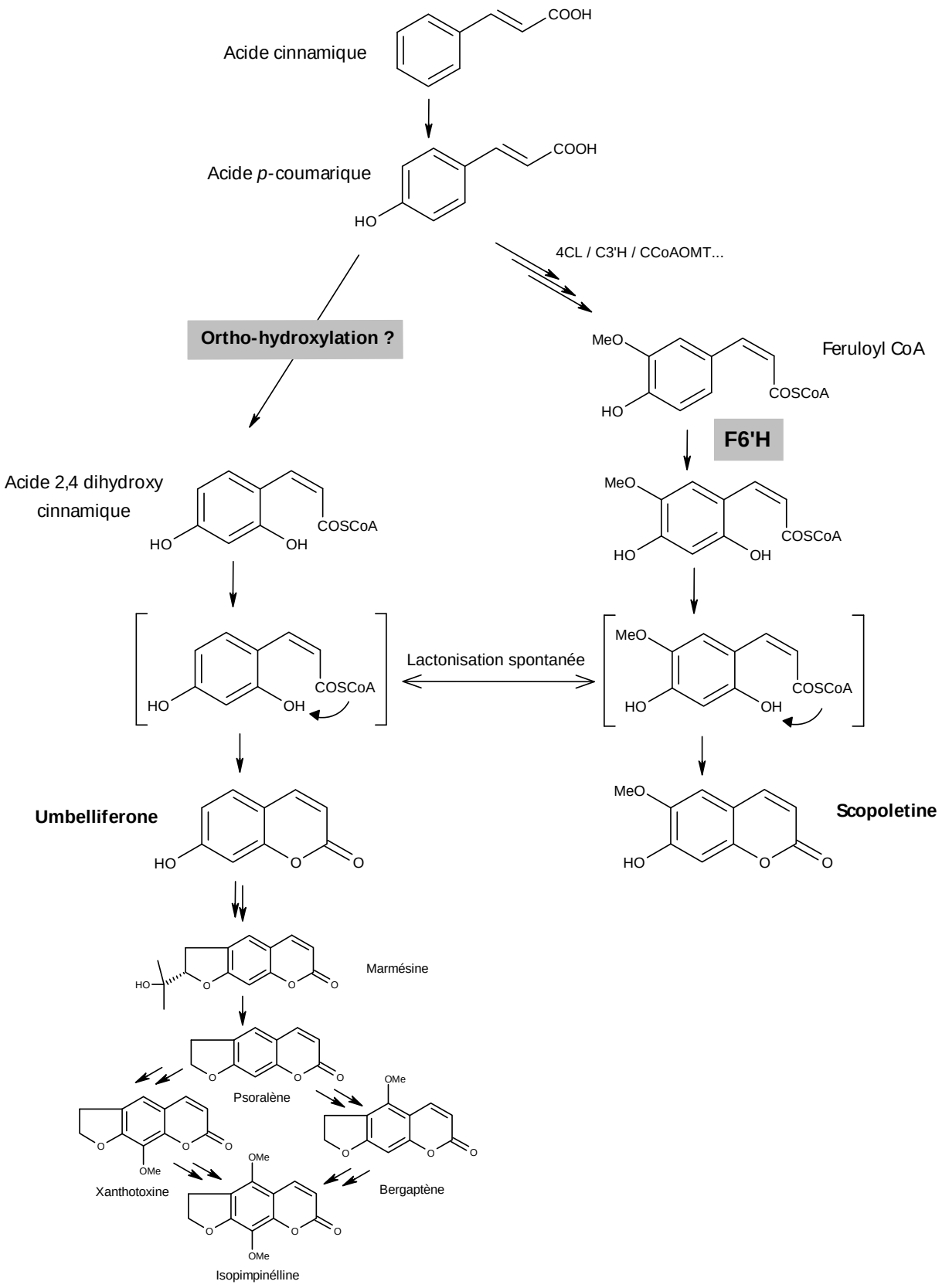


Figure 5 : Formation de l'umbelliférone et de la scopolétine.

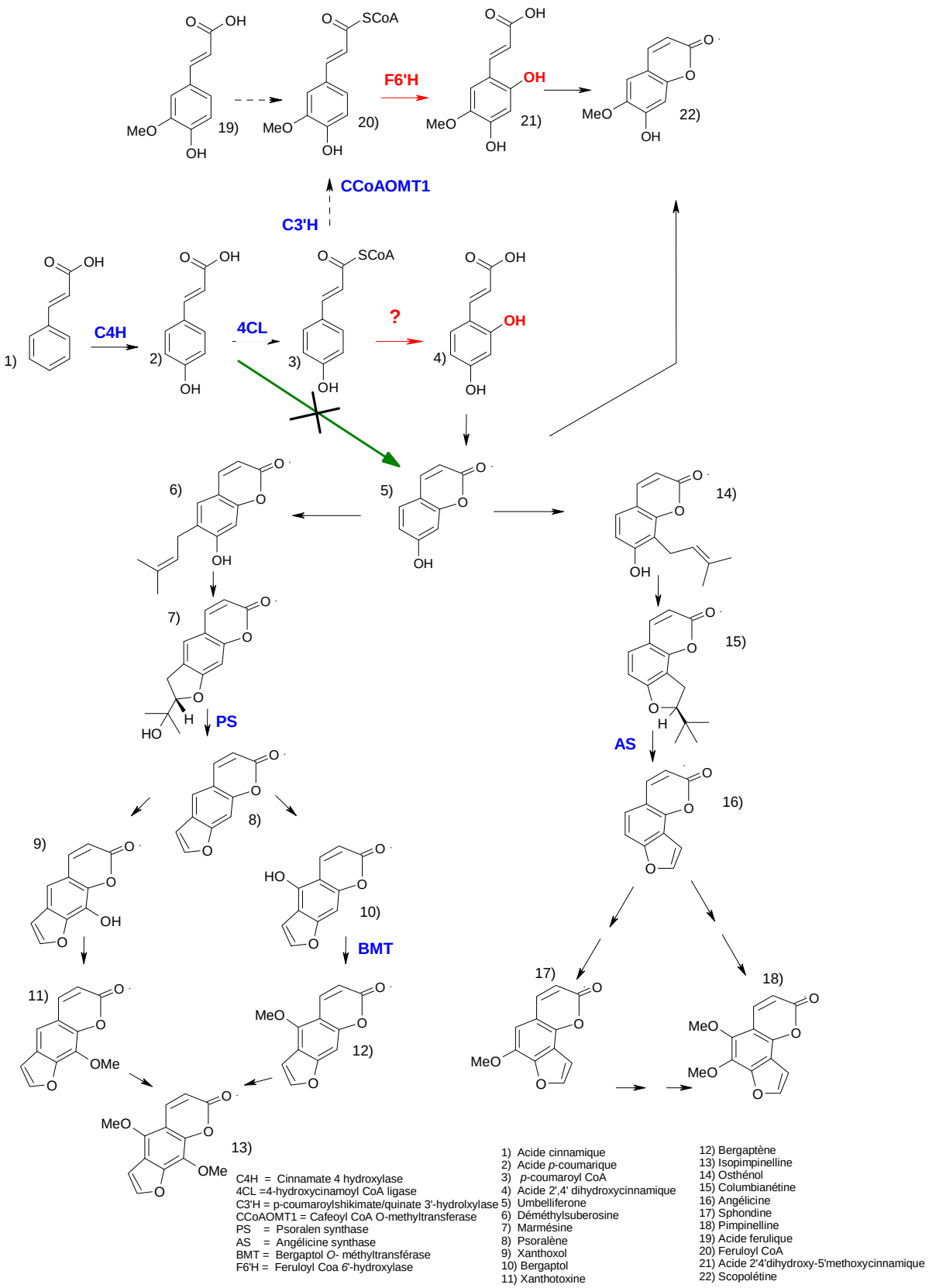


Figure 6 : Comparaison de la voie de biosynthèse de la scopolétine et de l'hypothétique voie de biosynthèse des furocoumarines linéaires et angulaires.

I.5 Induction de la synthèse des dérivés coumariniques

Comme la plupart des métabolites secondaires, les dérivés coumariniques sont synthétisés pour se protéger contre les stress biotiques ou abiotiques comme les infections, la présence d'herbivores, les blessures, les radiations UV, les expositions à des teneurs anormales en ozone, en polluants ou en nutriments (Zangerl and Berenbaum, 1987; Hamerski and Matern, 1988b; Dercks et al., 1990; Matern, 1991; Kitamura et al., 1998; Kai et al., 2008).

Chez certaines plantes, il est possible de modifier *in vitro* et/ou *in vivo* la teneur en furocoumarines par l'ajout dans le milieu de culture d'un pathogène (extrait de *Phytophtora megasperma*), ou par irradiation des plantes par des rayons UV (366 nm) (Zobel and Brown, 1993). Dans le cas de cultures cellulaires de *Petroselinum crispum*, la présence d'un extrait de pathogène fongique dans la culture engendre une augmentation de synthèse en furocoumarines et en pyranocoumarines par les cellules en suspension (Tietjen and Matern, 1983). L'élicitation des cultures cellulaires d'*Ammi majus* par *Phytophthora megasperma f.sp glycinea* entraine aussi une augmentation de la teneur en furocoumarines (Hamerski and Matern, 1988b).

L'efficacité de l'élicitation ne dépend pas uniquement de l'éliciteur utilisé mais résulte d'une association de différents facteurs. La surproduction de coumarines, chez des cultures cellulaires de *Angelica archangelica L.*, est augmentée d'un facteur 15 en combinant une déplétion du milieu de culture en hormone de croissance, une culture des cellules dans l'obscurité et une élicitation par des extraits de *Pythium aphanidermatum* (Siatka and KaÅ¡parovÃ¡, 2009).

L'élicitation et la surproduction de certains composés naturels se traduisent en général par une surexpression des gènes impliqués dans la voie de biosynthèse, et une régulation de l'expression des autres gènes de la cellule ou de la plante. Ce résultat, qui n'est pas encore démontré sur la voie de synthèse des furocoumarines, a notamment été mis en évidence sur la voie de synthèse du chorismate. En effet, les blessures ou les infections des plantes induisent une surexpression des gènes de la voie du shikimate, aboutissant à la synthèse du chorismate (Weaver and Herrmann, 1997).

L'élicitation par les ultraviolets de *Radix Astragali* induit une nette augmentation de la synthèse des isoflavonoïdes. Cette surproduction d'isoflavonoïdes est corrélée à une surexpression des gènes impliqués dans la voie de biosynthèses des flavonoïdes telles que la cinnamate-4- hydroxylase (Xu et al., 2011). Zobel et Brown ont montré qu'une exposition des feuilles de *Ruta graveolens* à une faible onde lumineuse (λ=366 nm) entraine une

augmentation de 20 % de la concentration en furocoumarines totale au niveau des feuilles les plus anciennes. Il est suggéré que cette réponse pourrait augmenter la protection de la plante contre la pénétration des UV dans le tissu végétal (Zobel and Brown, 1993). Larbat a de même montré durant sa thèse (Larbat, 2006) une augmentation de synthèses des composés furocoumariniques après une induction de 24 h aux UV (312 nm).

L'induction de la synthèse des furocoumarines telles que la xanthotoxine, le bergaptène et le psoralène au niveau des feuilles de *Heracleum lanatum* est directement corrélée aux changements saisonniers. En effet, la concentration en furocoumarines augmente jusqu'au milieu de mai ,avec un pic maximal le 25 avril, puis décroit rapidement avec l'élargissement des feuilles (Zobel and Brown, 1990a). Le profil métabolique en furocoumarine est donc différent en fonction du stade végétatif de la plante.

I.6 Propriétés biologiques des furocoumarines

Les furocoumarines sont des molécules qui interagissent avec un large spectre de macromolécules telles que les acides nucléiques, les lipides ou les protéines. Ces molécules se retrouvent la plupart du temps liées et forment des adduits avec les macromolécules (Van Henegouwen et al., 1989).

I.6.1 Réaction de cyclisation ou photocycloaddition

La dimérisation des coumarines et furocoumarines se fait sur les doubles liaisons C3=C4 après une induction aux UV aboutissant à la formation de 4 isomères : *trans*-head-to-head (tHH), *trans*-head-to-tail (tHT), *cis* head-to-head (cHH) et *cis*-head-to-tail (cHT) (Figure 7). La dimérisation du psoralène entraine majoritairement une structure *trans*-head-to-head (Schönberg, 1968). Ces réactions de photocyloadition peuvent être à l'origine de la formation de nouveaux médicaments de synthèse (Sinkel et al., 2011).

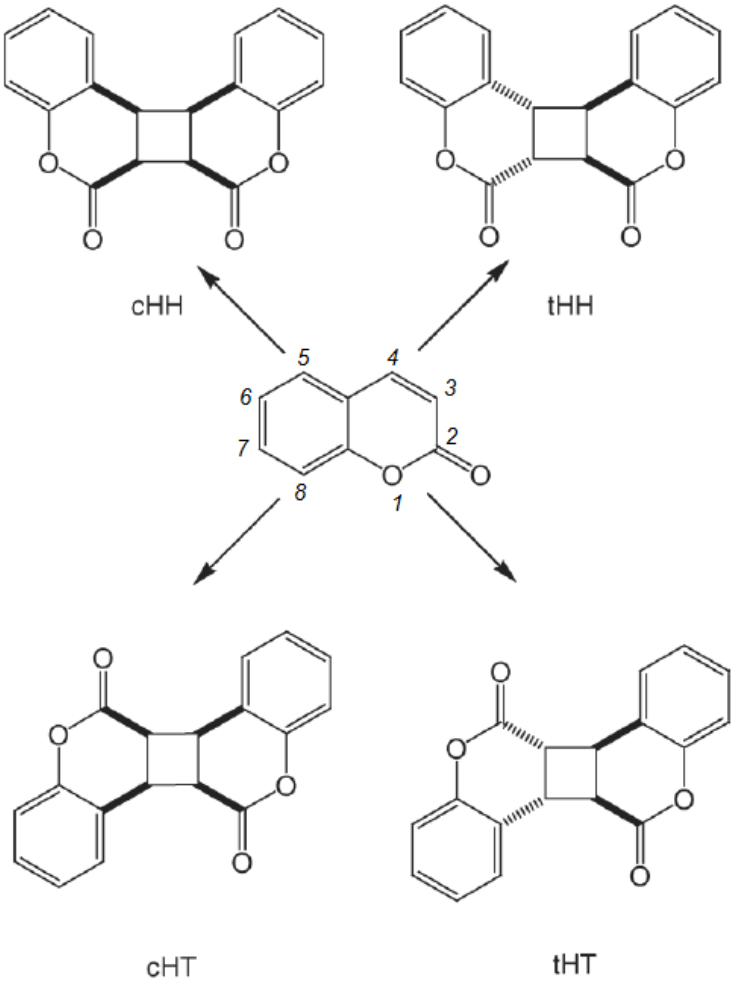


Figure 7 : Photodimérisation des composés coumariniques (Kitamura et al., 2005).
(tHH), *trans*-head-to-head ; (tHT), *trans*-head-to-tail ; (cHH), *cis* head-to-head ; (cHT), *cis*-head-to-tail

La photocycloaddition du psoralène se fait aussi avec les molécules insaturées possédant une ou plusieurs doubles liaisons. Elle a lieu sur les carbones C3=C4 et/ou sur les C4'=C5'. Huit isomères peuvent être générés par formation de mono-adduit sur le psoralène (Figure 8). Il est également possible de générer des di-adduits sur le psoralène. En comptabilisant les différentes formes *syn* et *anti* ainsi que *cis* et *trans*, 16 isomères peuvent alors être générés à partir d'un groupement éthylène substitué H_2C=CHX (Kagan.J, 1993). Cette propriété biologique est à l'origine de la photoaddition des furocoumarines sur les acides gras insaturés.

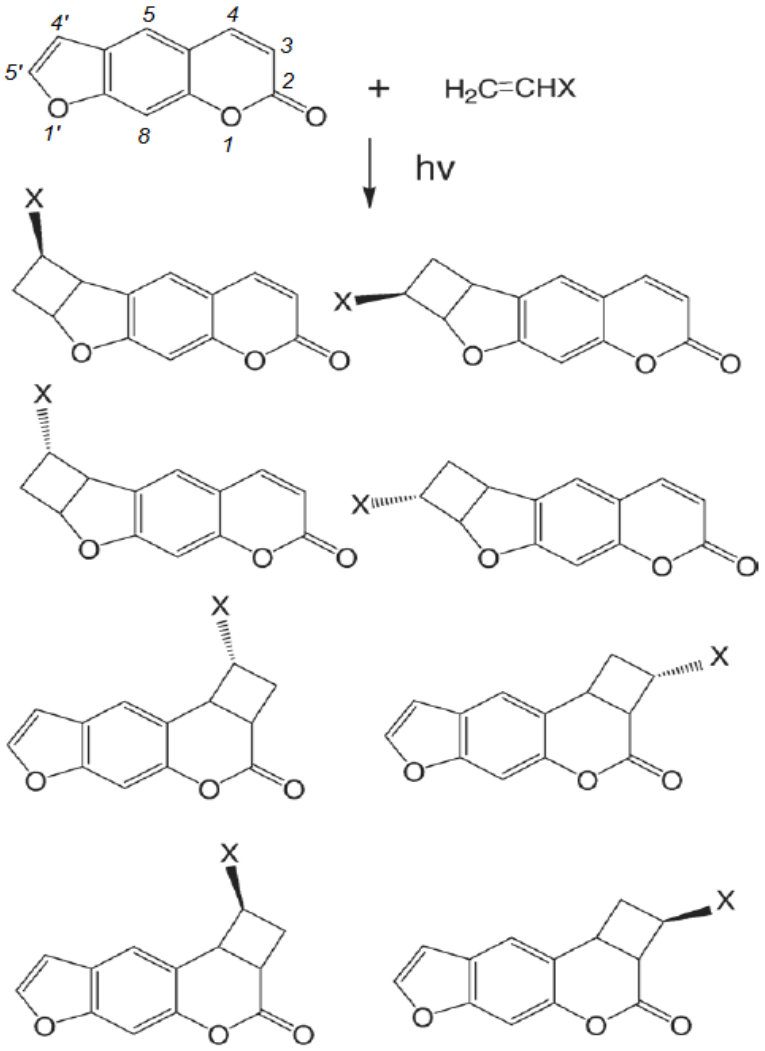


Figure 8 : Photocycloaddition du psoralène avec un groupement éthylène substitué.

I.6.2 Réaction avec les acides nucléiques

Les réactions photochimiques des furocoumarines avec les acides nucléiques sont connues depuis longtemps. Musajo était le premier à décrire ce type d'interaction avec l'ADN (Musajo et al., 1965).

La majeure partie de la réaction de photosensibilisation se déroule dans le noyau de la cellule (Sasaki et al., 1988; Cadet et al., 1992). Dans l'obscurité, les furocoumarines se complexent avec l'ADN par formation de liaisons hydrogènes. Suite à une irradiation aux UVA (315-400 nm) ou à la lumière visible, la photoaddition par liaison covalente se fait sur les carbones C4'= C5' du noyau furane et/ou sur les carbones C3=C4 du noyau pyrane (Figure 9 A).

Suivant la conformation de l'ADN et l'agencement des bases nucéotidiques, une deuxième photoaddition peut se produire. En effet, l'adduit sur le C4'=C5' du cycle furane peut absorber un second photon et ainsi, aboutir à la formation d'une seconde liaison covalente du cycle pyrane C3=C4 avec une autre base pyrimidique adjacente (Edelson, 1988; Serrano-Pérez et al., 2008) (Figure 9 B).

Dans le cas de la formation d'une liaison bivalente, la furocoumarine se comporte comme un pont moléculaire reliant les thymines des deux brins d'ADN, et bloque ainsi la réplication et la traduction de l'ADN (Figure 9 C). Cette propriété explique le caractère mutagène des furocoumarines (Dardalhon et al., 1998)

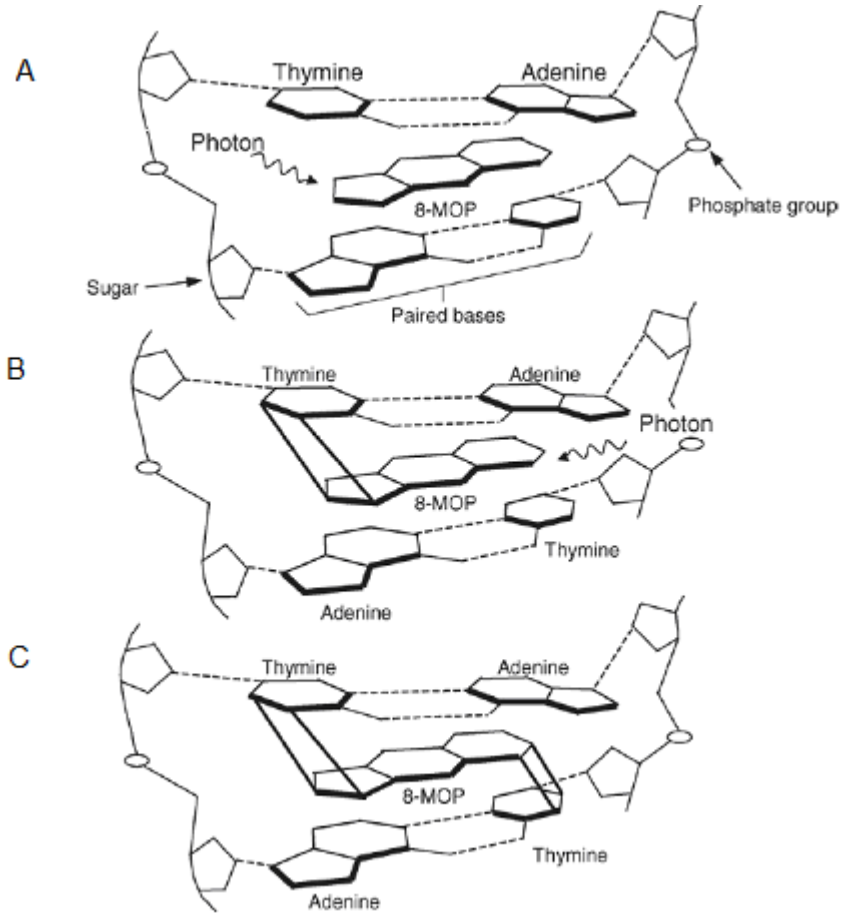


Figure 9 : Illustration de la formation de liaisons chimiques entre le 8-metoxypsoralen (8-MOP) et les bases pyrimidiques des différents brins de l'ADN (Kitamura et al., 2005).

La formation de photocycloadduit des furocoumarines sur les bases pyrimidiques se fait préférentiellement avec la thymine plutôt qu'avec l'adénine. Les régions comportant une

séquence poly[dA-dT].poly[dA-dT] sont les plus favorables à la formation de liaisons avec les furocoumarines (Kang et al., 1992).

I.6.3 Réaction avec les acides gras insaturés

L'établissement d'une liaison covalente du C3=C4 du cycle pyrane et/ou C4'=C5' du cycle furane avec un acide gras, est rendu possible après photoactivation des furocoumarines. Tout comme pour la réaction avec les acides nucléiques, cette réaction est indépendante de la présence d'oxygène. Cette interaction entre les furocoumarines et les acides gras pourrait être à l'origine de la stimulation de la mélanogénèse (Dall'Acqua and Martelli, 1991; Zarebska et al., 2000).

En effet, la photoaddition des furocoumarines sur les phospholipides, entraine une dégradation de ce complexe par la phospholipase A2 (PLA_2) qui génère l'accumulation d'acides gras liés aux furocoumarines. Cette augmentation d'acides gras photoliés active la protéine kinase C (PKC) par le biais du diacylglycérol (DAG). La phosphorylation de protéines spécifiques aboutit à l'activation de la voie de signalisation entrainant la stimulation de la mélanogénèse (Figure 10).

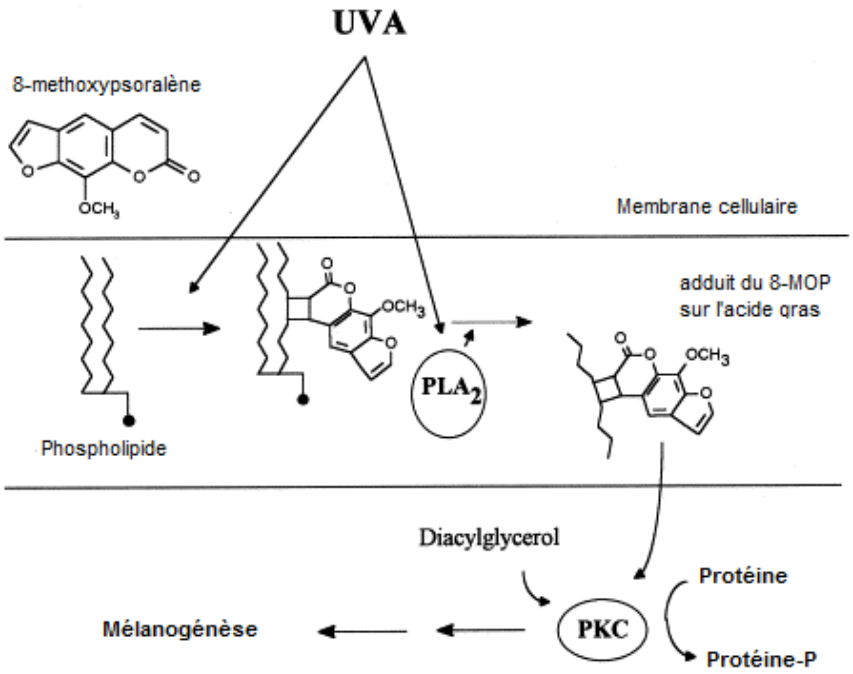


Figure 10 : Mécanisme proposé d'activation de la mélanogénèse par la photoaddition des furocoumarines sur les acides gras insaturés (Zarebska et al., 2000).

I.6.4 Réaction avec les protéines

La photoaddition des furocoumarines se fait sur un large éventail de protéines, telles que les ribonucléases, l'albumine bovine, les histones, les lysozymes (Schmitt et al., 1994).
Expérimentalement, des cellules de rats traitées avec du 8-MOP et exposées aux UV montrent une forte proportion de 8-MOP liée aux protéines (57 %), contre (26 %) liée aux lipides et seulement (17 %) liée à l'ADN (Van Henegouwen et al., 1989).

I.7 Utilisations thérapeutiques

Depuis toujours, la médecine traditionnelle a utilisé les plantes produisant des furocoumarines pour le traitement de nombreuses maladies. Dans l'Égypte Ancienne, on soignait les leucodermies avec des extraits d'*Ammi majus*, une plante qui poussait sur les berges du delta du Nil. Cependant, nos ancêtres ignoraient les raisons expliquant que la consommation d'*Ammi majus* associée à une exposition au soleil, permettait de soigner des maladies de peau.
La pharmacognosie, l'extraction et la purification de dérivés du psoralène ont abouti dans les années 1950 à l'utilisation plus large de ces composés dans notre médecine moderne en particulier pour le traitement du psoriasis et vitiligo (SchÃ¶nberg and Sina, 1950; Lerner et al., 1953; Musajo et al., 1954; Fitzpatrick and Pathak, 1959; Benedetto, 1977; Pathak and Fitzpatrick, 1992)

I.7.1 La photochimiothérapie

La photochimiothérapie utilise simultanément l'association de médicaments et de lumière afin de traiter certaines maladies de peau (Parrish et al., 1974). La Puvathérapie décrit dans sa dénomination, une thérapie associant le psoralène et des UVA (320-340 nm).
À l'état excité, le psoralène peut réagir avec un grand nombre de molécules différentes, dont les acides nucléiques. Pour cette raison, le psoralène a longtemps été utilisé comme principe actif dans le traitement de maladie de peau liée à une hyperprolifération cellulaire. Il apparait maintenant que son utilisation n'entraîne pas uniquement un effet antiprolifératif comme on l'a pensé pendant une cinquantaine d'années, mais induirait d'autres effets tels que l'apoptose ou des réactions antigéniques (Godar, 1999).

La photochimiothérapie consiste à associer la prise orale ou topique de furocoumarines avec une exposition aux UVA (320-340 nm) (J.Kagan, 1993; Santana et al., 2004). Le 8-

methoxypsoralène (8-MOP), le 5-methoxypsoralène (5-MOP) et le 4,5',8- trimethylpsoralène (TMP) sont les furocoumarines les plus souvent utilisées (Marzano et al., 2002).

Les indications de la Puvathérapie sont nombreuses et permettent le traitement de nombreuses maladies incluant le psoriasis, le lichen plantaire, la mycose fongoïde, l' urticaire pigmentaire, la lucite polymorphe, la pelade et le vitiligo (Photo 2) (Llano et al., 2003). La plupart de ces maladies sont caractérisées par une hyperprolifération cellulaire.

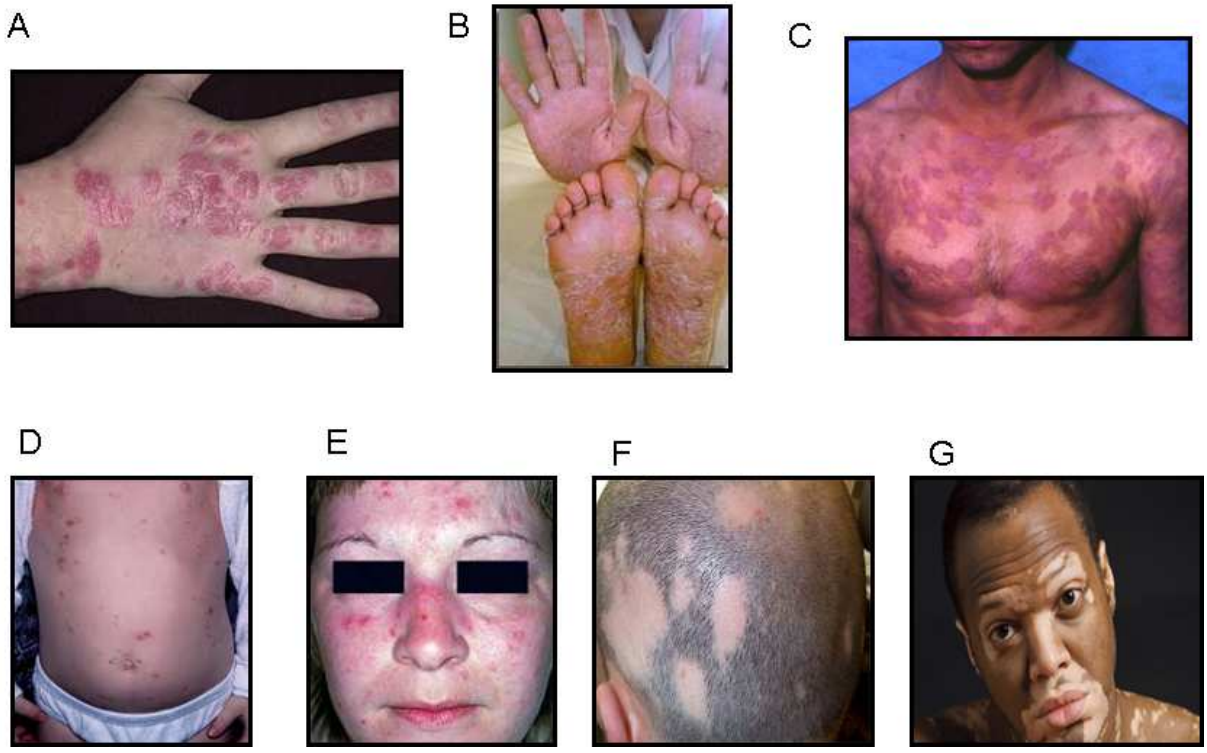


Photo 2 : Indications de la puvathérapie.

A) Psoriasis B) Lichen plantaire C) Mycose fongoïde D) Urticaire pigmentaire E) Lucite polymorphe F) Pelade G) Vitiligo

Les lésions caractéristiques du psoriasis proviennent de l'hyperprolifération des kératinocytes. La formation de photoadduit entre le psoralène et l'ADN inhibe la réplication cellulaire et entraine une amélioration des symptômes de la maladie (Sasaki et al., 1988).

Le vitiligo, lui, est caractérisé par une diminution de la production des pigments de la peau (la mélanine), provenant certainement d'une perte de la prolifération naturelle des mélanocytes (Bordin et al., 1993).
Les dernières études montrent que le 5-MOP serait le meilleur photosensibilisant utilisable dans la Puvathérapie (Serrano-Pérrez et al., 2008).

L'utilisation des furocoumarines *via* la Puvathérapie dans le traitement du psoriasis diminue cependant progressivement du fait de la découverte de nouvelles molécules d'immunodépresseurs ayant des effets thérapeutiques plus importants et plus ciblés au niveau cellulaire (Krueger, 2002; Nograles et al., 2010)

I.7.2 La photophérèse

La photochimiothérapie extracorporelle ou photophérèse est une « PUVA extracorporelle des leucocytes ». Ils sont préalablement séparés des globules rouges, puis irradiés par des UVA après traitement externe par le 8 MOP, avant d'être réinjectés au malade (Van Iperen and Beijersbergen Van Henegouwen, 1997; Chiesa-Fuxench and González-Chávez, 2010). Cette approche thérapeutique est utilisée pour le traitement des lymphomes cutanés à cellules T, qui est une maladie à prolifération maligne des cellules helper T (AdiÅŸen and GÃ¼rer, 2010). La photophorèse, comparée aux immunosuppresseurs et à la corticothérapie générale, présente peu d'effets secondaires et constitue un élément important en faveur de son indication.
Cette forme de traitement constitue une relativement bonne thérapie immunomodulatrice, et d'autres applications thérapeutiques de la photophérèse voient le jour.
Les résultats obtenus dans la maladie du greffon contre l'hôte (Jagasia et al., 2009), et dans la prévention du rejet de greffe d'organe (rein, cœur, foie, poumon) (Hivelin et al., 2009), ont incité à élargir les indications thérapeutiques de la photophorèse. C'est le cas pour les maladies médiées par les lymphocytes T dans des formes réfractaires aux traitements classiques telles que la sclérodermie, le lupus, la sclérose en plaques, la dermatite atopique, la polyarthrite rhumatoïde (Guyot et al., 2007) et plus récemment le diabète de type I et la maladie de Crohn (Abreu et al., 2009).
D'autres essais cliniques semblent montrer que les produits photolysés du psoralène à faible concentration, mis en contact avec les précurseurs des cellules érythroides, stimulent la

formation d'hémoglobine A et F. Cette découverte pourrait être à l'origine d'une alternative dans le traitement de la β-thalassémie (Viola et al., 2008).

I.7.3 Traitement antidépresseur

La dépression est une maladie qui peut avoir pour origine une altération du système hypothalamo-hypophyso-surrénalien (HHS), un dérèglement du fonctionnement de la monoamine oxidase (MAO) et des systèmes oxydatifs.

Des marqueurs de dépression sont couramment utilisés tels que le suivi de l'activité des MAO-A et MAO-B, la mesure du niveau de cortisol ou de malondialdehyde et l'évaluation de l'activité de la superoxyde dismutase (Abdalla et al., 1986; Bilici et al., 2001; Bao et al., 2005).

L'ethnopharmacologie est à l'origine de la découverte de certaines propriétés antidépressive des furocoumarines. En effet, depuis des décennies les graines de *Psoralea corylifolia*, appelées aussi Buguzhi, sont utilisées dans la médecine chinoise pour traiter différents symptômes liés au vieillissement de la population (Zhang et al., 2005a). Or, les furocoumarines sont les composés secondaires les plus représentés dans cette espèce végétale. Chen et ses collaborateurs ont évalué l'activité potentielle antidépressive des furocoumarines présentes chez *Psoralea corylifolia* en suivant l'évolution des marqueurs de dépression après administration des furocoumarines.

Les résultats suggèrent que la prise de furocoumarines entraine des propriétés antidépressives sur les populations vieillissantes par action sur la MAO, l'axe HHS et sur les systèmes oxydatifs (Chen et al., 2007).

I.7.4 Activités bactériostatiques, fongicides et antimycobatérienne.

Les furocoumarines sont capables d'inhiber l'activité des molécules impliquées dans la communication intra et interspécifique (N-acylhomoserine lactones, diester de boronat) entre bactéries Gram positif et Gram négatif et dans l'expression de gènes pathogènes.

De plus, les furocoumarines inhibent la formation de biofilms synthétisés par les bactéries telles que *E.coli O157, Salmonella typhimurium* et *Pseudomaonas aeruginosa* impliqué dans la protection des bactéries (Girennavar et al., 2008b). Elles sont donc bactériostatiques et

peuvent présenter un intérêt dans la course au développement de nouvelles molécules thérapeutiques contre les infections bactériennes.

De plus, les furocoumarines telles que l'angélicine et ses dérivés, possèdent aussi des activités antifongiques. Elles révèlent une toxicité sur *Candida albicans*, *Cryptococcus neoformans*, *Saccharomyces cerevisiae* et *Aspergillus niger* (Sardari et al., 1999) pouvant montrer un intérêt sur la prise en charge de nombreuses mycoses.

Adams et ses collaborateurs ont démontré l'activité antimycobactérienne des composés furocoumariniques « geranylifié » contre *Mycobacterium fortuitum*, *Mycobacterium smegmatis* et *Mycoabacterium phlei*. Ces composés ont une forte activité tout en possédant une concentration minimale inhibitrice faible (Adams et al., 2006). Ces molécules pourraient constituer une alternative de traitements aux infections de certaines mycobactéries.

I.7.5 Activité antiproliférative et apoptotique

Des essais d'activité ont été conduits sur des cellules cancéreuses humaines et de rats, avec différentes furocoumarines linéaires issues de *Notopterygium incisum*, une plante décrite dans la pharmacopée traditionnelle chinoise. Une activité antiproliférative et apoptotique de ces molécules sur les cellules cancéreuses est mise en évidence.

Les relations structures-fonctions de ces furocoumarines expliquant leurs propriétés apoptotiques sont en cours d'étude. Les premiers résultats indiquent que la présence des groupements libres hydroxyles sur la chaine lipophile attachée en C5 des furocoumarines linéaires est indispensable à l'activité anticancéreuse (Wu et al., 2010). Les mécanismes réactionnels mis en jeu restent à découvrir, et permettront peut-être l'essor d'une nouvelle famille de molécules anticancéreuses.

I.8 Toxicités des furocoumarines

I.8.1 Photodermatite

L'application sur la peau d'une décoction d'herbes médicinales chinoises contenant des furocoumarines est souvent responsable de phytophotodermatite aiguë.

Les lésions apparaissent après 5 à 7 jours d'exposition au soleil, et se caractérisent par un érythème cutané au niveau du site d'application de la décoction ainsi que l'apparition de macules et de taches sur les jambes (Zhu, 2010).

Un simple contact avec des plantes appartenant à la famille des Apiacées ou des Rutacées suivi d'une exposition au soleil suffit à provoquer des lésions cutanées (Janda et al., 2008; Polat et al., 2008). Dans ce cas, les lésions se caractérisent la plupart du temps par une hyperpigmentation de la peau et l'apparition de vésicules bulleuses sur les mains et autour de la bouche (Kung et al., 2009).

I.8.2 Mélanome

Des études statistiques, effectuées par Sayre et Dowdy (Sayre and Dowdy, 2008) sur les patients ayant subi une Puvathérapie, semblent montrer une certaine augmentation de l'incidence des mélanomes chez ces patients ayant suivi le traitement associant 8 MOP et radiation UVA. De plus, ils constatent que le taux d'incidence des mélanomes dans une population, augmente de la même manière que le taux de consommation de citrus dans cette même population. Une étude récente démontrerait la corrélation positive, entre l'incidence d'apparition des mélanomes et la consommation de jus d'orange d'après Sayre et Dowdy. Cependant, cette étude statistique est vivement contestée par Godar (Godar, 2008).

I.8.3 Toxicité alimentaire par inhibition de cytochrome P450

Les furocoumarines largement représentées dans les jus de citron et de pamplemousse peuvent être à l'origine d'intoxication. En effet, il est démontré que le catabolisme des principes actifs des médicaments chez l'homme, se fait dans 60 % des cas par l'intervention d'enzymes appartenant à la famille des cytochromes P450. Or, les furocoumarines telles que le bergaptol ou la bergamottine sont des inhibiteurs de P450, et plus particulièrement de CYP3A4, CYP2C9 et CYP2D6. Ces cytochromes interviennent dans la dégradation de nombreux médicaments cardiovasculaires ou des benzodiazépines (Girennavar et al., 2007; Gorgus et al., 2010). Une intoxication par effet de surdosage est donc possible en cas d'associations médicamenteuses avec la prise de jus de citron ou de pamplemousse.

II Les cytochromes P450

II.1 Présentation

Les cytochromes P450 font actuellement partie des sujets de recherche les plus en vogue de ce dernier demi-siècle, que ce soit sur une composante biochimique ou moléculaire. Le terme général de cytochrome P450 désigne une superfamille multigénique de protéines enzymatiques dite « hemo-thiolates » à propriété redox.

Cette famille d'enzymes a été découverte dans les années 60 dans des cellules de foie de mammifères (Omura and Sato, 1962). Ce sont des enzymes ubiquitaires que l'on retrouve dans tous les organismes vivants : végétaux ou animaux, procaryotes ou eucaryotes. Elles sont aussi désignées par le terme de mono-oxygénases capables d'activer les molécules de dioxygènes en des entités hautement réactives (ROS), afin d'insérer l'oxygène moléculaire au niveau d'un atome de carbone, d'azote ou de soufre.

En 1958, Garfinkel et Kligenberg ont décrit l'apparition d'un pigment jaune orangé dans des fractions microsomales hépatiques de rat et de cochon en présence de monoxyde de carbone. Ce pigment fut d'abord caractérisé comme une hémoprotéine, identifiée à un cytochrome de type *b* rarissime. Finalement, il sera désigné sous le terme de « P450 » (Klingenberg, 1958; Omura and Sato, 1962). La dénomination « P450 » a été attribuée à ce type d'enzyme, car elle présente une absorption intense dans le spectre visible à 450 nm lorsque le fer de l'hème, à l'état réduit, est complexé au monoxyde de carbone. Cette particularité d'absorption lui confère son appellation actuelle : P signifiant Pigment et 450 correspondant à la longueur d'onde pour le maximum d'absorption.

Depuis la découverte de cette famille d'enzymes et l'essor dans les années 1990 de la biologie moléculaire, le boum du séquençage a permis d'isoler plus de 11 000 séquences de P450, dont 4000 séquences identifiées chez les plantes. (http://drnelson.utmen.edu/CytochromeP450.html).

Les P450 sont des monooxygénases impliquées dans de nombreuses réactions aussi diverses que variées, telles que : l'hydroxylation, la déalylation, l'oxydation, la dimérisation, la désamination ou l'époxydation (Mansuy, 1998) (Figure 11).

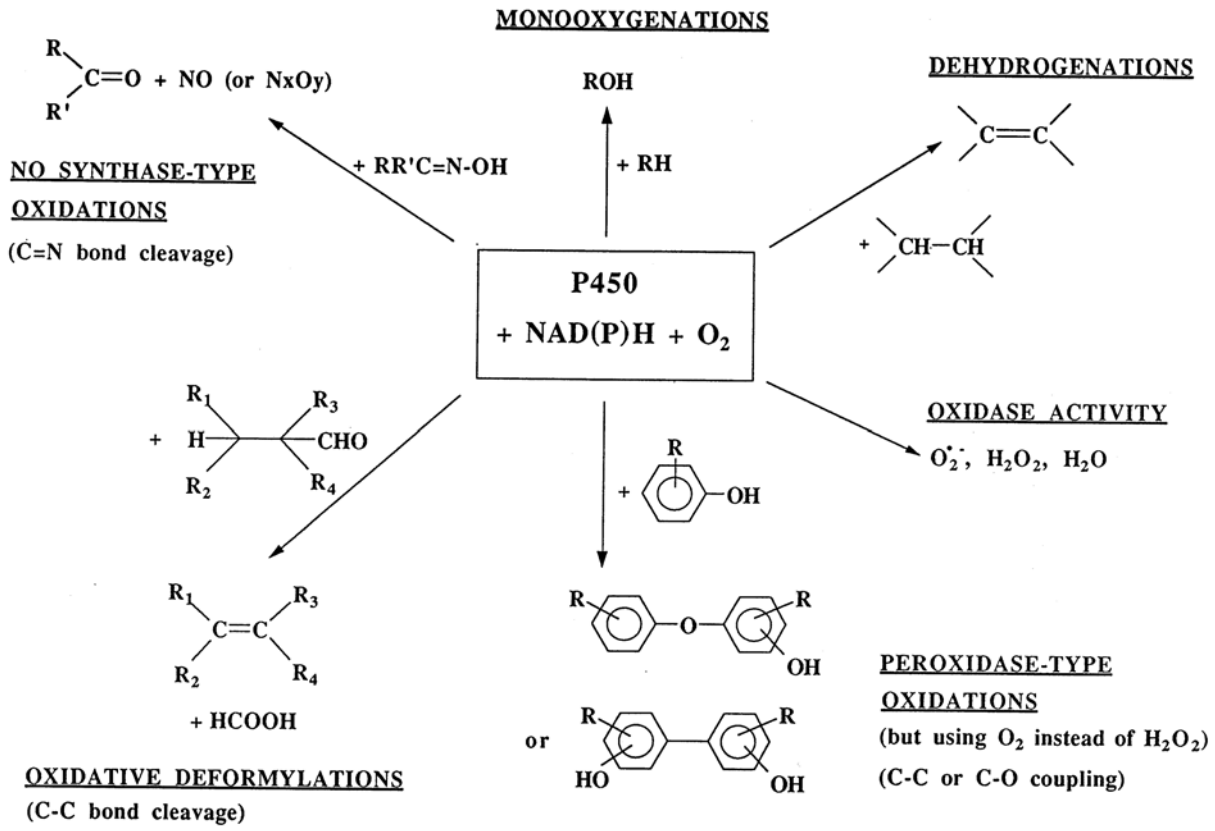


Figure 11 : Exemples de diversité réactionnelle de P450 (Mansuy, 1998).

Cette grande diversité de réactions explique qu'on retrouve les P450 dans de nombreuses voies de biosynthèses. La réaction d'hydroxylation est la plus couramment rencontrée. Cette monooxygénation se fait à partir du dioxygène et nécessite un pouvoir réducteur fourni dans la plupart des cas par du NADPH ou du NADP.

Les origines phylogénétiques et les liens qui unissent l'ensemble des P450 ne sont pas totalement déterminés et demeurent encore discutables. Toutefois, on pense que le système des P450 aurait évolué depuis plus de 3,5 milliards d'années, avant même l'arrivée de l'oxygène atmosphérique (Wickramasinghe and Villee, 1975). Nelson émet l'hypothèse, d'après l'élaboration d'alignement de séquences et d'arbres phylogénétiques d'évolution, qu'un gène ancestral de P450 serait apparu il y a 3,5 milliards d'années chez la bactérie (Nelson et al., 1993; Nelson et al., 1996; Nelson, 1999). Ce gène ancestral serait très proche du seul P450 retrouvé chez tous les organismes : CYP51 qui catalyse la 14α-déméthylation du

lanostérol chez les bactéries ou du obtusifoliol chez les plantes. Ce sont des intermédiaires intervenant dans les premières phases de la synthèse des stérols. La diversification des cytochromes P450 pourrait provenir de la complexification de la biosynthèse des stérols au cours de l'évolution. La pression de sélection a permis par la suite de sélectionner les P450 les plus adaptés aboutissant, à l'heure actuelle, à une incroyable diversité d'enzymes appartenant à cette famille.

II.2 Nomenclature

Du fait de la grande quantité de gènes décrits, il a été rapidement nécessaire d'établir une nomenclature et une classification. La classification se base sur l'alignement de séquences peptidiques. Elle est disponible sur le site internet du Prof. David Nelson (http://drnelson.utmem.edu/cytochromeP450.html). La nomenclature a été adoptée pour l'ensemble des cytochromes P450. La nature de la protéine est décrite par les trois lettres CYP (pour **Cy**tochrome **P**450). Le chiffre suivant les trois lettres correspond à la famille. Toutes les protéines appartiennent à une même famille dès lors qu'elles possèdent plus de 40 % d'identité entre elles. Suit ensuite une lettre qui définit la sous-famille du gène. Les gènes appartenant à une même sous-famille ont plus de 55 % d'identité. Enfin, un second nombre permet de distinguer les différents membres appartenant à la même sous-famille.

En raison du nombre de familles devenu très important, un niveau de regroupement supplémentaire est requis. On parle alors de clans qui rassemblent les familles appartenant à un même groupe (issu d'un même gène ancestral) d'après les nombreux arbres phylogénétiques établis auparavant par Nelson (Nelson, 1999). Les clans sont désignés par le plus petit chiffre des familles qu'ils regroupent ou par le chiffre de la famille majoritaire.

II.3 Classification

Les Cytochromes P450 nécessitent l'apport d'électrons au substrat afin de pouvoir réaliser des réactions de mono-oxygénation. Ces électrons sont fournis par le NADPH ou le NADH. La classification des P450 permet de différencier la façon dont les électrons sont transportés du NAD(P)H au site catalytique.

On distingue deux classes majoritaires de transfert d'électrons. Deux autres classes sont rencontrées, mais de façon beaucoup plus rare (Paine et al., 2005).

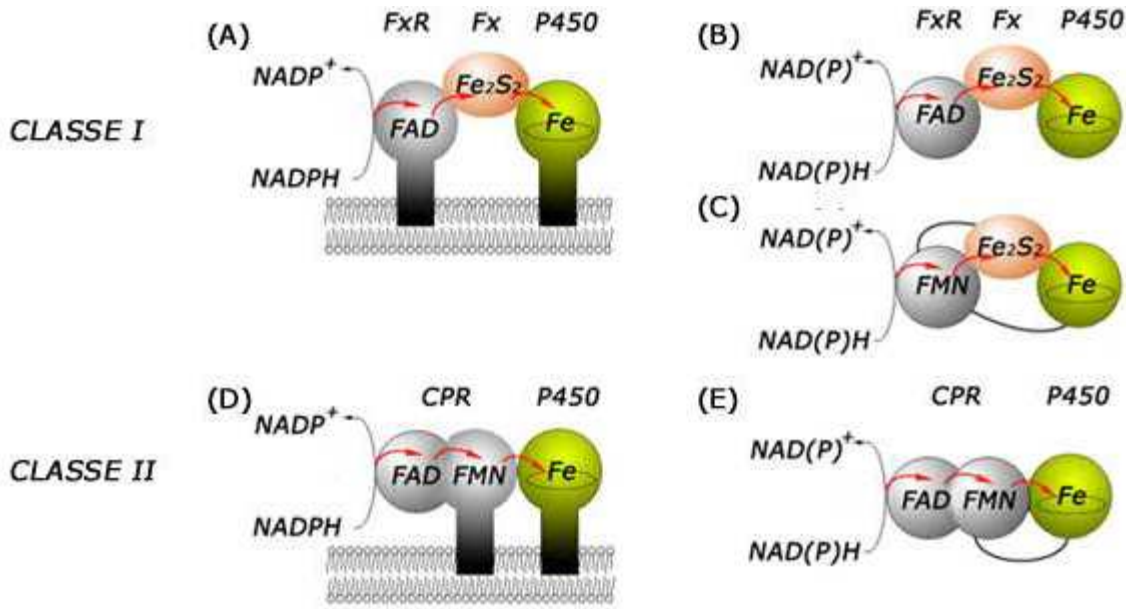


Figure 12 : Partenaires de transfert d'électrons des Cytochromes P450 (Paine et al., 2005).
La classe I comprend les P450 mitochondriaux (A), la plupart des P450 solubles (B) et certains systèmes fusionnés (C). La classe II comprend les P450 du RE (D) et les P450 qui incluent la ferredoxine à deux flavines. Fx : ferredoxine, FxR : ferredoxine réductase, CPR : CYP réductase

- La **classe I** comprend dans la chaîne de transfert d'électrons, des enzymes à centre fer-soufre. Les électrons du NAD(P)H sont captés par une ferrodoxine réductase à flavine (flavine adénine dinucléotide ou FAD) et sont transférés au P450 par une ferrodoxine à centre fer-soufre Fe_2S_2 (Figure 12.A et .B). On retrouve dans cette classe la plupart des P450 bactériens et des P450 mitochondriaux. On trouve aussi un système de transfert comprenant une protéine de fusion contenant une réductase flavine et FMN (flavine mononucléotide), une protéine type ferredoxine à centre fer-souffre et le cytochrome P450 (Figure 12.C).

- La **classe II** requiert une NADPH-cytochrome P450 réductase (CPR) contenant les groupements prosthétiques FAD et FMN. Cette classe est majoritaire chez les eucaryotes. On les retrouve au niveau du réticulum endoplasmique (Figure 12.D) (Figure 13). Dans le système microsomal de P450, on retrouve aussi un autre système non spécifique des P450 qui fait intervenir un cytochrome b5. Ce système est capable aussi de transférer des électrons au

P450 après avoir reçu les électrons, soit par la CPR, soit par la NADH-cytochrome b5 réductase (Figure 12.E)

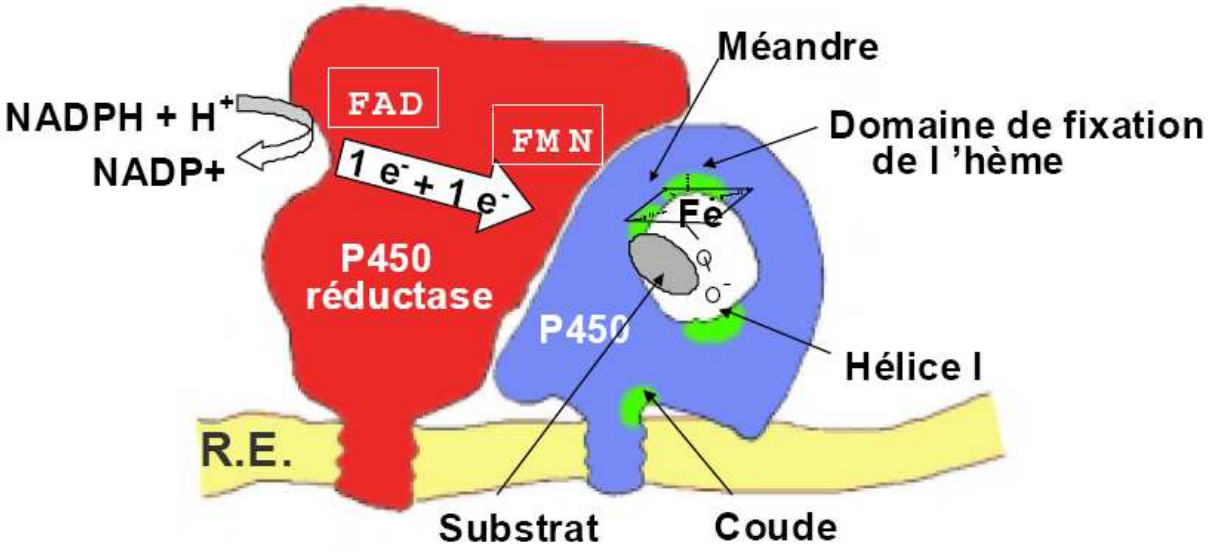


Figure 13 : Représentation schématique d'un cytochrome P450 microsomale d'eucaryote et de la NADPH P450 réductase (Werck-Reichhart and Feyereisen, 2000).

- La **classe III** ne nécessite pas d'apport d'électrons extérieurs ou d'oxygène moléculaire, car les substrats transformés sont généralement riches en électrons comme des hydroperoxydes ou des endoperoxydes. Ces P450 catalysent des isomérisations ou des déshydratations.

- La **classe IV** contient un représentant, le **$P450_{nor}$** qui réduit le monoxyde d'azote NO et reçoit les électrons du NADH directement, sans faire appel à des protéines intermédiaires.

II.4 Localisation des cytochromes P450

On retrouve les cytochromes P450 dans tous les organismes vivants à différents niveaux de localisation cellulaire. En effet, chez les procaryotes les cytochromes P450 sont des protéines cytosoliques, alors que chez les eucaryotes ils sont membranaires et distribués dans différents compartiments cellulaires. On distingue de ce fait, les P450 mitochondriaux ancrés dans la membrane des mitochondries, les P450 microsomaux ancrés dans la membrane du réticulum

endoplasmique et les cytochromes plastidiaux ancrés au niveau des chloroplastes chez les végétaux.

II.5 Séquences et domaines conservés

L'acquisition de nombreuses données structurales cristallographiques de P450 cytosoliques et membranaires depuis une trentaine d'années a permis de mettre en évidence une architecture globalement similaire entre les différents P450 de procaryotes et d'eucaryotes bien que les homologies de séquences peptidiques soit relativement éloignées. Malgré l'importante diversité génétique des cytochromes, cette famille est caractérisée par des séquences protéiques consensus :

- Une séquence consensus située du côté proximal de l'hème (FxxGx(R/H)xCxG). Ce domaine est appelé « Cys-pocket » ; il contient la cystéine, ligand du fer, et est considéré comme étant la signature peptidique de la famille des P450. Ce motif est aussi dénommé « signature cytchrome P450- hème » et contribue à la fixation et au maintien de l'hème au sein de la protéine.
- Une autre séquence consensus distale (G/A)Gx(E/D)T, moins bien conservé au sein des P450 jouant un rôle dans l'activation de l'oxygène moléculaire lors de la catalyse (Figure 14).

```
                 335         345        ...  ...      475         485         495
CYP71AJ4     EECIKAIVWD MILGGTETTS ... ... DYKGLHYEYL PFGGGRRGCP GIQFAMAVNE
CYP71D9      DESIKAVIWD IFGGGSDTSS ... ... EYKSNSFEFI PFGAGRRMCP GLTFGLSNVE
CYP75B2      DTEIKALLLN LFVAGTDTSS ... ... DVRGNDFEVI PFGAGRRICA GMNLGIRMVQ
CYP6B3       DGVISAQMFI FYAAGYETSA ... ... VKDRHPCAYL PFSAGPRNCL GMRFAKWQSE
CYP98A       ETTVIGLLWD MITAGMDTTA ... ... DIKGHDYRLL PFGAGRRICP GAQLGLNMVQ
```

Figure 14 : Séquences protéiques consensus distales (G/A)Gx(E/D)T et proximale FxxGx(R/H)xCxG « cys-pocket » conservées chez tous les P450s

L'ancrage des P450 d'eucaryotes à la membrane du réticulum endoplasmique (RE) se fait par l'intermédiaire d'une hélice hydrophobe, immédiatement suivie par la succession de résidus basiques délimitant l'ancre membranaire. Des séquences riches en prolines se retrouvent sur cette partie charnière entre l'ancre transmembranaire qui fixe la protéine à la membrane et la partie globulaire de la protéine. La séquence consensus riche en prolines délimite un coude qui permet au reste de l'enzyme de s'accoler à la membrane du RE. Cette séquence de proline permet donc à la protéine de se déplacer autour d'une articulation et ainsi de se positionner

correctement par rapport à la réductase et à la surface de la membrane réticulum endoplasmique (Kemper, 2004). Cette région est également importante pour le bon repliement de la protéine (Chen and Kemper, 1996). Les régions spécifiques de l'enzyme impliquées dans le positionnement du substrat à proximité du site actif sont dénommées SRS (Substrate Recognition Site). Elles ont été identifiées pour la première fois pour la famille CYP2 par analogie avec des données de P450 bactériens (Gotoh, 1992). Ces SRS sont numérotées de 1 à 6 et correspondent à des régions spécifiques conservées chez tous les P450 (Williams et al., 2000). La position de ces SRS semble être bien conservée entre les P450 en particulier pour les SRS 1, 4, 5 et 6 (Figure 15). Le SRS1 contenant l'Arg^{104} chez CYP71AJ1 semble être impliqué en collaboration avec l'Arg^{434}, dans la stabilisation de l'hème (Larbat et al., 2009). Les SRS 2 et 3 semblent former un canal impliqué dans l'accessibilité du substrat au site actif (Wilmouth et al., 2002). Le triplet HPP situé sur le SRS5 est retrouvé sur de nombreux membres de la famille des CYP71, mais d'après les travaux de Larbat *et al* sur la modélisation de CYP71AJ, celui-ci est remplacé par YFT et semble être important dans la conversion de la marmésine en psoralène chez CYP71AJ1-3 (Larbat et al., 2009). En effet, chez l'angélicine synthase (CYP71AJ4), celui-ci est remplacé par le triplet YIT. De même, la Thr^{479} située au niveau du SRS6 chez CYP71AJ1 parait essentiel à l'activité catalytique du cytochrome (Larbat et al., 2009).

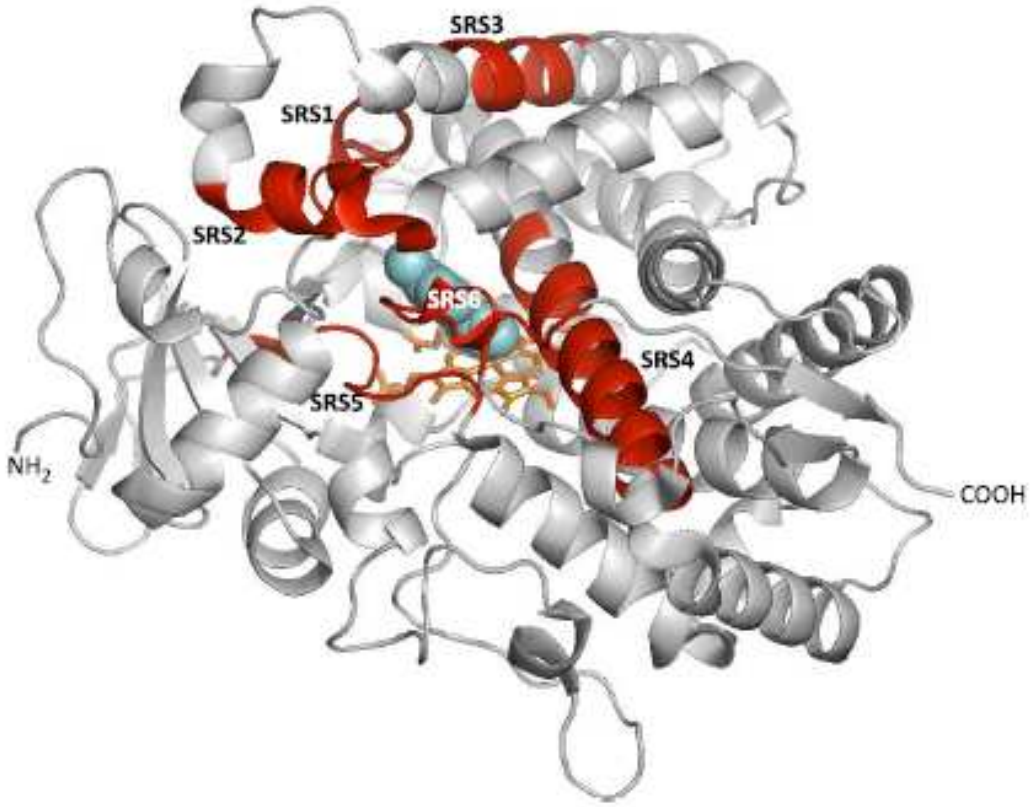


Figure 15 : Représentation tridimensionnelle des différentes régions SRS.
Les séquences SRS sont colorées en rouge sur CYP2B4, un cytochrome P450 de mammifère. Les SRS entourent le substrat, coloré en bleu, et le groupement héminique, coloré en orange (De Graaf et al., 2005).

II.6 Structure

Tous les P450 identifiés à ce jour désignent des hémoprotéines constituées d'un hème relié à l'apoprotéine par un groupement thiolate complexant le fer du groupe prosthétique (Figure 16). On représente schématiquement les P450 comme étant une grosse hémoprotéine constituée d'une chaîne polypeptidique unique, l'apoprotéine (d'un poids moléculaire compris entre 45 et 60 kDa) et d'un groupement prosthétique central, l'hème (catalyseur de la réaction enzymatique). Celui-ci est lié de façon « non covalente » par l'intermédiaire d'une cystéine à l'apoprotéine. Le substrat se loge dans la partie hydrophobe de l'enzyme qui constitue le site catalytique.

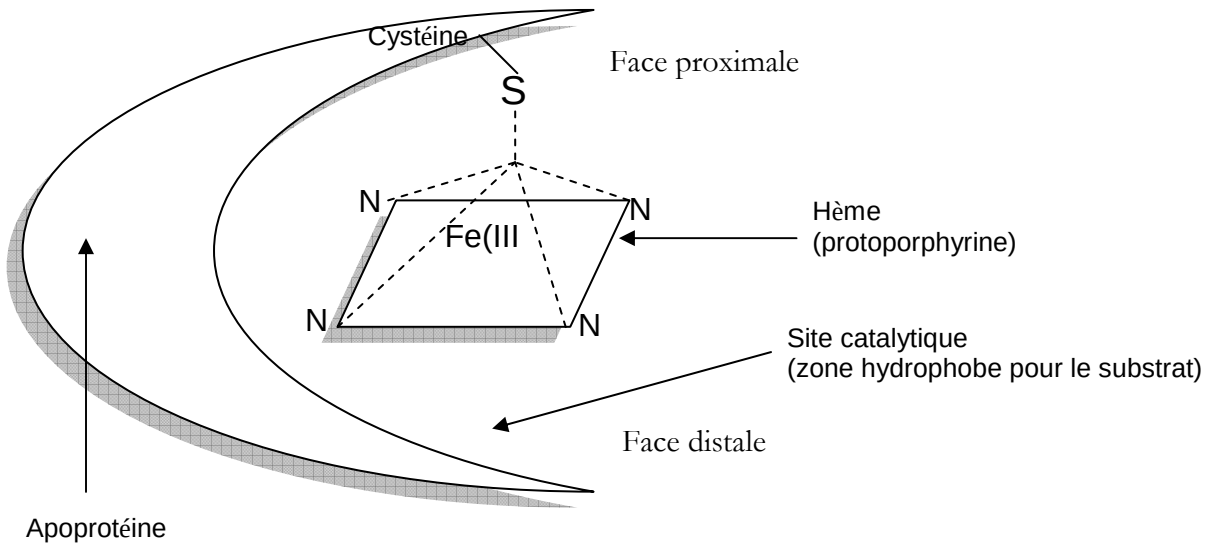


Figure 16 : Représentation simplifiée d'une coupe transversale de la structure d'un Cytochrome P450.

L'hème des cytochromes P450 est constitué d'un noyau protoporphyrine IX. Ce noyau protoporphyrine IX est localisé au centre du cytochrome P450 et constitue une partie totalement invariable. Quatre cycles pyrole renferment en leur centre un atome de fer qui peut effectuer 6 liaisons de coordinence dont quatre permettent la liaison avec le tetrapyrole (Figure 17). La cinquième liaison de coordinence du fer permet de stabiliser la protoporphyrine dans le site actif du P450 par une liaison avec le groupement thiolate de la cystéine présente dans la région C-terminale de l'apoprotéine. Le fer peut aussi être hexacoordonné, la sixième liaison de coordinence etant établie avec une molécule d'H_2O, d'O_2 ou de CO.

Le plan de l'hème définit donc deux régions dans le P450 : la face proximale et la face distale contenant le site actif de l'enzyme.

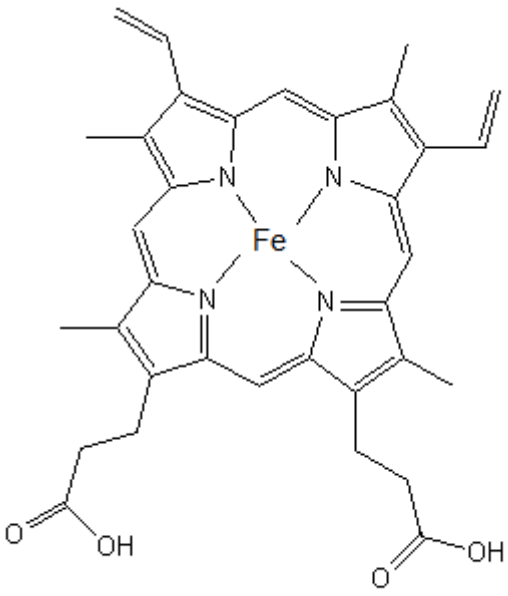


Figure 17 : Représentation du noyau protoporphyrine IX.

II.7 Cycle catalytique des P450

Le substrat établit une liaison avec l'atome de fer après s'être introduit dans la poche délimitée par les hélices K et I. L'arrivée concomitante d'un électron fourni par le complexe NADPH réductase permet la fixation de l'oxygène moléculaire après transformation du fer de l'état ferrique (Fe^{3+}) à l'état ferreux (Fe^{2+}). Un second électron aboutit ensuite à la formation d'eau et de substrat sous forme oxydée (Zangar et al., 2004). Le produit formé est ensuite relargué et le fer se retrouve à l'état initial (Figure 18). En condition ordinaire, un cycle catalytique complet s'accompagne de la consommation d'une molécule d'oxygène et d'une molécule de NADPH fournissant les deux électrons nécessaires. Il existe cependant des réactions de découplage entre la consommation de NADPH et la formation de produit aboutissant à la formation d'espèces oxygénées réactives (EOR), telles que H_2O_2 et $O_2^{\cdot}$

Les EOR sont des molécules hautement réactives et très toxiques pour les cellules. Elles sont dans certains cas, impliquées dans la signalisation de l'apoptose (Zhou et al., 2011).

Le cycle catalytique des P450 peut être une des sources principales de stress oxydant avec la génération importante d'EOR (Zangar et al., 2004). La cellule contrôle la production de ces EOR par régulation de l'activité des P450 d'une part, mais aussi par régulation transcriptionnelle d'autre part. En effet, la répression de l'expression de certains P450 et l'induction de l'expression d'enzyme antioxydante telle que la catalase contribuent à moduler

la quantité d'EOR produite par la cellule (Barouki and Morel, 2001; MarÃ- and Cederbaum, 2001).

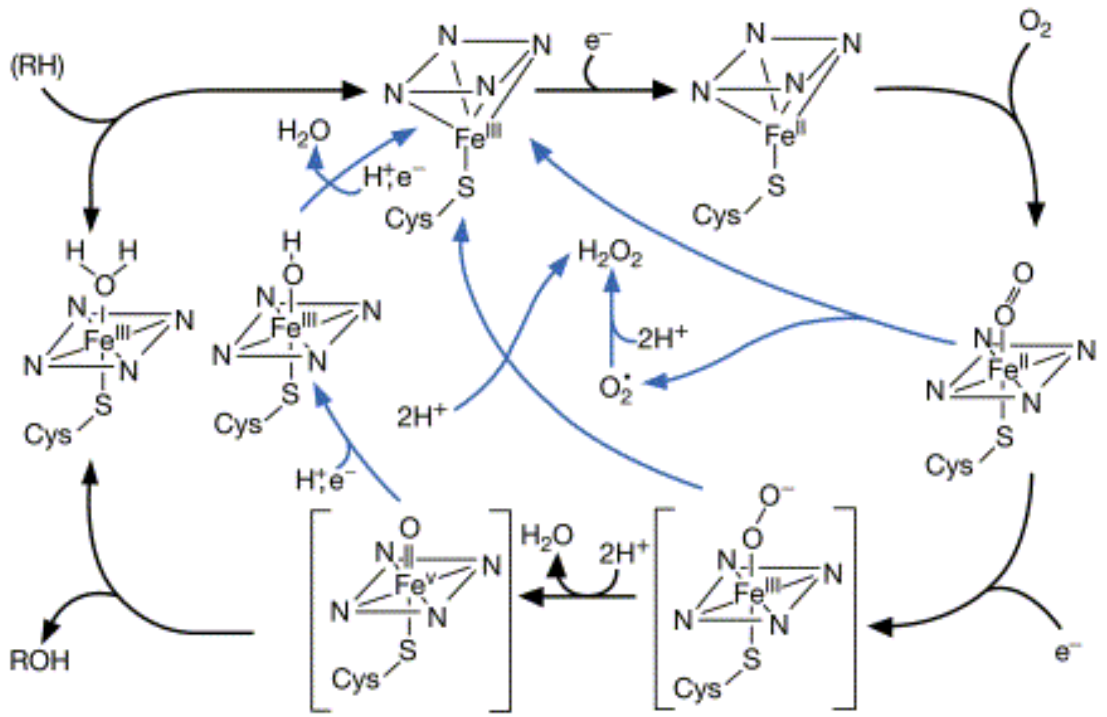


Figure 18 : Mécanisme réactionnel des cytochromes P450.

Les flèches bleues représentent les trois voies de découplage amenant à la formation d'EOR. (Zangar et al., 2004)

II.8 Rôles et intérêts des P450

II.8.1 Un champ de recherche très large

La variété de réactions catalysées confère aux P450 un rôle très important aussi bien dans la transformation de molécules endogènes et exogènes. Ils n'ont pas le même rôle suivant l'organisme dans lequel il est exprimé. Chez les procaryotes, certains P450 contribuent à la biosynthèse de composés endogènes tels que les stérols ou des antibiotiques afin d'assurer une fonction de défense du microorganisme. Chez les eucaryotes, les P450 assurent la synthèse d'hormones stéroïdiennes, de rétinoïdes, de certaines vitamines, de métabolites primaires et secondaires. De plus, les P450 sont les principales enzymes impliquées dans le métabolisme

des xénobiotiques et participent à l'élimination de composés exogènes comme des médicaments, des toxines ou des polluants.

Le domaine d'application des P450 étant très étendu, le nombre de publications sur les P450 (ou relative aux P450) ne cesse d'augmenter depuis les dernières décennies. L'étude des P450 se fait dans des domaines de recherche aussi nombreux que variés (Bernhardt, 1995, , 2006) (Figure 19).

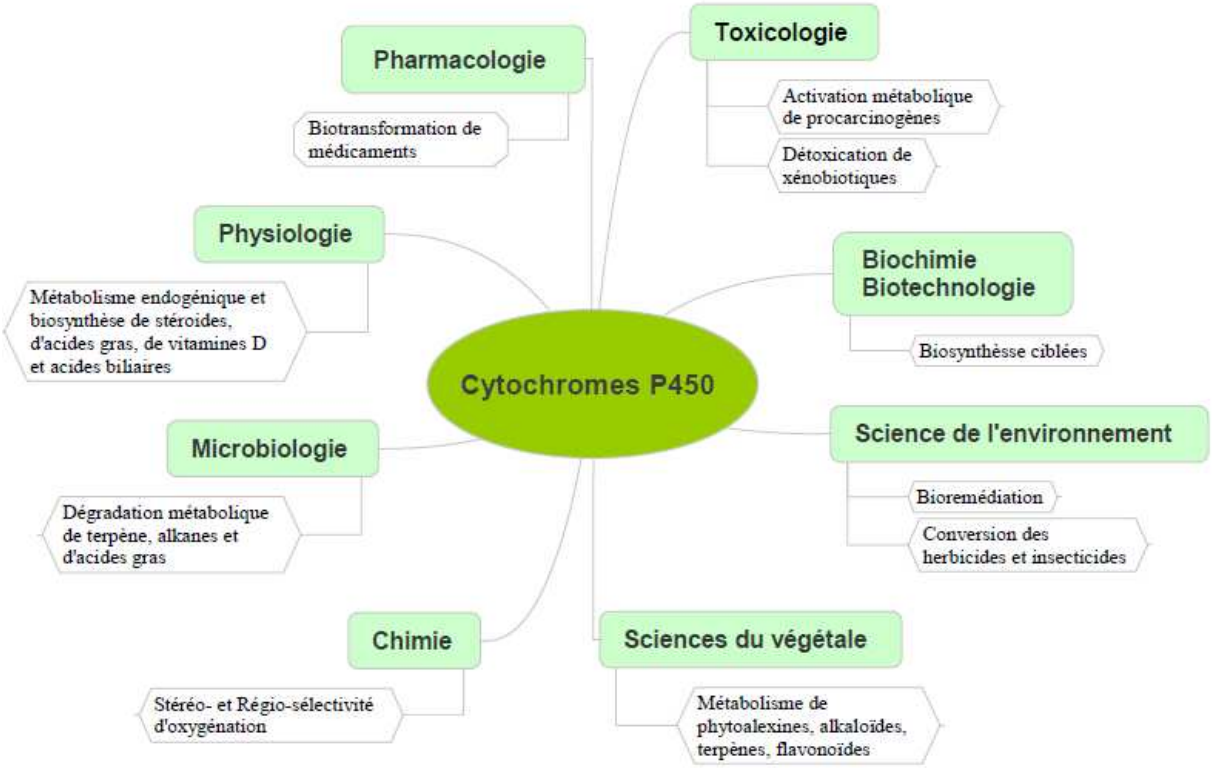


Figure 19 : Domaines de recherches des P450 et les différents champs d'applications.

II.8.2 Fonctions des P450 chez les végétaux

Plus de 4000 séquences de P450 ont déjà été recensées chez les végétaux. Relativement peu de ces séquences ont des fonctions identifiées. L'aire de la génomique et la démocratisation des outils bio-informatiques entraîne une augmentation perpétuelle du corpus de connaissances des P450. La caractérisation de ces gènes par expression hétérologue permet d'affiner la compréhension et les interactions coexistantes au sein de cette grande famille de gènes.

II.8.2.1 Biosynthèse de métabolites secondaires

Les cytochromes P450 sont impliqués dans de nombreuses voies aboutissant à la biosynthèse de métabolites avec des fonctions aussi complexes que variées. Le maintien physique de la plante (lignine), la coloration des fleurs (flavonoïdes, anthocyane), la protection contre les bioagresseurs (furocoumarines, alcaloïdes) ou la régulation hormonale (auxines, gibbérellines) sont autant de fonctions assurées par des métabolites qui nécessitent une ou plusieurs interventions de P450 pour leur synthèse. Les voies de biosynthèse de phénylpropanoïdes, des terpénoïdes, des glucosinolates ou des alcaloïdes font toutes intervenir des cytochromes P450 (Tableau 1).

Phénylpropanoïdes	Cinnamate-4-hydroxylase	**CYP73A**	Park *et al*. 2010
	Psoralen synthase	**CYP71AJ**	Larbat *et al.* 2007
	p-Coumaroyle-3'-hydroxylase	**CYP98A**	Morand *et al*. 2002
	Flavonoïdes 3'-hydroxylase	**CYP75B**	Seitz *et al*. 2007
	Flavone synthase II	**CYP93B**	Fliegman *et al*, 2010
	Naringenine-3'-hydroxylase	**CYP76B**	Sung *et al*, 2011
Terpènes	Geraniol-10-hydroxylase	**CYP76B**	Sung *et al* 2011
	Abscisic acid 8'-hydroxylase	**CYP707A**	Mizutani *et al*.2006
	Limonene-3-hydroxylase	**CYP71D**	Wust *et al*.2001
Glucosinolates	p-Hydroxymandelonitrile synthase	**CYP71E**	Kirsten *et al*.2008
	1-acid-nitro-2-indoyl-éthane synthase	**CYP83B**	Naur *et al*.2003
Alcaloïdes	Tabersonine-19-hydroxylase	**CYP71BJ**	Giddins *et al*. 2011

Tableau 1 : Exemple de P450 impliqués dans la biosynthèse de métabolites secondaires chez les végétaux.

II.8.2.2 Catabolisme de molécules exogènes

Chez les plantes, la fonction des P450 dans les différentes voies de biosynthèse de composés secondaires est nettement mieux documentée que leurs rôles dans la détoxication de composés endogènes ou exogènes (Schuler, 1996b, 1996a). Cependant les P450 de plantes représentent un atout majeur dans la décontamination des polluants environnementaux caractérisés pour la plupart par un faible poids moléculaire d'une part et une hydrophobicité d'autre part

(Sandermann Jr, 1994). Ces P450 pourraient être utilisés en biotechnologie dans la sélectivité et la tolérance de certaines plantes à certains herbicides et constituer un fort potentiel dans le développement de la biorémédiation (détoxication de composés organiques issus des déchets de l'industrie) (Frear, 1995; Werck-Reichhart, 1995b). Le CYP76B1 est un P450 isolé chez *Helianthus tuberosus* qui montre une capacité à détoxifier un large éventail de xénobiotiques incluant les alkoxycoumarines, les alkosyresorufines et les herbicides de la classe des phénylurées (Robineau et al., 1998).

De manière générale, les P450 participent à la dégradation ou à l'inactivation par la plante, des herbicides, des pesticides et des xénobiotiques. Une étude récente a même démontré qu'un de ces P450 pouvait métaboliser le cyclotriméthylènetrinitramine (RDX), un explosif utilisé par les militaires (Bhushan et al., 2003).

Les réactions de biotransformation des xénobiotiques, catalysées par les cytochromes P450 s'inscrivent dans un processus de détoxication évitant l'accumulation de substances potentiellement toxiques pour la plante.

Le processus de détoxication se déroule en trois phases qui mettent en jeu différentes enzymes de biotransformation (Phase I et II) et des transporteurs (Phase III). Cet ensemble forme un système appelé enzyme du métabolisme des xénobiotiques (Figure 20).

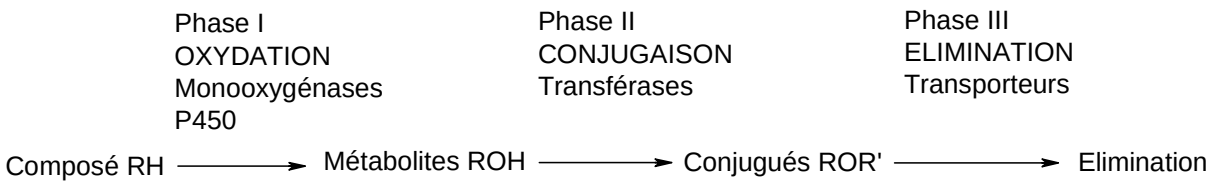


Figure 20 : Schéma simplifié du processus de détoxication.

Le processus de détoxication fait intervenir une première étape dite de fonctionnalisation (Phase I). Cette étape consiste à supprimer ou à rajouter des groupements OH, NH_2 ou COOH à la molécule, afin d'augmenter la polarité et la solubilité des molécules à détoxifier. Il s'en suit une étape dite de conjugaison (Phase II) avec d'autres molécules (sucre, glutathion, malonate, acides aminés ou macromolécules). La dernière étape consiste à éliminer les métabolites ainsi transformés (Phase III) par l'intermédiaire de transporteurs. Dans la plupart

des cas, les métabolites désactivés sont stockés dans les vacuoles ou encore dans la paroi végétale. (Schuler and Werck-Reichhart, 2003) Ils sont ainsi dépourvus de toxicité pour la cellule.

II.8.3 Les cytochromes P450 à l'origine de l'adaptation entre les plantes et les insectes

Une étroite relation existe entre les plantes et les insectes. Les insectes utilisent les plantes pour leur apporter l'alimentation nécessaire à leur survie. Ils trouvent chez celles-ci la nourriture essentielle pour leur développement, telle que les acides aminés, la plupart des vitamines du groupe B, les stérols et les minéraux indispensables (Fraenkel, 1959).
En contrepartie, certaines plantes utilisent les insectes pour coloniser l'espace et se reproduire. Cette étroite collaboration, non symbiotique, a fait émerger des phénomènes adaptatifs et évolutifs chez les plantes, mais aussi chez les insectes. Des modifications du métabolisme basal de part et d'autre sont apparues. Les plantes synthétisent de nouveaux composés de défense, dits allélopathiques, et les insectes acquièrent différentes stratégies pour lutter contre les molécules toxiques des plantes.

Un des phénomènes adaptatifs observé chez les insectes concerne leur régime alimentaire. En effet, certains insectes se nourrissent de plantes de manière très sélective et ne vont prédater qu'une seule espèce de plante : ce sont les monophages. A l'inverse, d'autres sont capables de se nourrir de plusieurs plantes appartenant à la même famille : il s'agit des oligophages. Enfin, les polyphages dévorent une grande variété de plantes (Schuler, 2011).

Il convient de se demander quelles sont les pressions de sélections à l'origine d'une évolution du régime alimentaire de certains insectes. Serait-ce la composition nutritionnelle des plantes ? Ou cette évolution serait-elle liée à la présence, chez certaines plantes, de composés métaboliques plus ou moins répulsifs ou attractifs ?

Dans les années 1960, Fraenkel étaye cette 2ème hypothèse en découvrant la présence de nombreux composés dits « secondaires », tels que les glucosides, les tanins, les saponines, les alcaloïdes, les acides organiques. Ces composés, non essentiels à la survie de la plante, sont présents dans des familles ou sous-familles de plantes bien distinctes (Fraenkel, 1959).

En essayant de comprendre comment les insectes rivalisent et s'adaptent aux composés de défense développés par les plantes, Fraenkel entrouve un vaste sujet de recherche concernant la perpétuelle adaptation des plantes et des insectes pour coexister. Les réponses apportées par Frankel relèvent de l'observation. Nous verrons par la suite que de nombreux travaux tentent d'expliquer les phénomènes évolutifs et adaptatifs apparus sur une composante moléculaire (Zumwalt and Neal, 1993; Berenbaum, 2002; Cianfrogna et al., 2002b; David et al., 2006; Wen et al., 2006b; Bono et al., 2008).

Il est maintenant admis que la capacité et la manière d'un organisme, de se défendre contre les métabolites toxiques issus de son environnement, sont directement corrélées à la variété et à la quantité de cytochromes P450 exprimées chez ce même organisme (Schuler, 2011).

II.8.3.1 Adaptation et synthèse de composés allélopathiques chez les plantes

Les plantes ont évolué depuis des millénaires dans leur environnement écologique, soumises à des interactions biochimiques avec les organismes agresseurs tels que les insectes herbivores. Les végétaux n'ont pas la capacité de fuir face à un environnement hostile ou à une prédation importante. De ce fait, les pressions de sélection et l'évolution ont abouti à la protection des plantes par une stratégie chimique. En effet, l'une des fonctions du métabolisme secondaire consiste à protéger les plantes par la synthèse de molécules chimiques contre un environnement inadapté. Une grande variété de composés chimiques a d'ailleurs été identifiée chez les plantes. Plus de 300 000 variétés de plantes sont déjà recensées dans le monde et seraient à l'origine de la production de plus de 200 000 composés naturels (Wu and Chappell, 2008).

Parmi tous ces composés, les terpénoïdes représentent la plus grande famille de molécules synthétisées impliquées dans des fonctions de défense avec plus de 40 000 structures déjà identifiées (limonene, myrcène, cira)(Zulak and Bohlmann, 2010). Les alcaloïdes constituent aussi une large classe de composés de défense avec plus de 12 000 structures recensées (morphine, tubocurarine, vincristine, vinblastine)(Facchini, 2001; Ziegler and Facchini, 2008). D'autres métabolites secondaires, tels que les phénylpropanoïdes et les dérivés de cette voie de biosynthèse comme les flavonoïdes, les anthocyanes ou les phénylpropanoïdes dimériques (Dixon and Paiva, 1995; Murray, 2000; Ververidis et al., 2007), possèdent aussi une activité allélopathique. On retrouve aussi les furocoumarines dont la diversité est moins

importante, mais qui présentent cependant un niveau de protection très important avec plus de 200 structures identifiées (bergaptène, psoralène, angélicine)(Curini et al., 2006).
Enfin, d'autres composés moins représentés, tels que les glucosinolates, les pipéramides, les benzoxazinoïdes, interviennent aussi dans la protection des plantes contre l'environnement (Gierl and Frey, 2001; Wilmouth et al., 2002; Halkier and Gershenzon, 2006).

Toutes ces molécules allélopathiques ne sont évidemment pas présentes chez les mêmes familles ou espèces de plantes. Leur synthèse reflète différentes stratégies employées pour protéger les plantes face à l'environnement. Certaines plantes produisent une grande variété de composés faiblement toxiques. D'autres produisent une seule famille de molécules, mais bien plus toxique. À titre d'exemple, la synthèse simultanée de différentes furocoumarines a pour effet d'augmenter le niveau de protection. De même, la synthèse d'une seule furocoumarine en présence concomitante d'octyle butyrate chez certaines plantes de la famille des Apiacées entraîne un effet synergique de toxicité sur les insectes prédateurs (Cianfrogna et al., 2002a). L'évolution a donc permis aux plantes de se doter d'une variété très importante de composés allélopathiques, mais dont la distribution est bien spécifique d'une espèce à l'autre.

II.8.3.2 P450 dans le génome des plantes

Comme nous l'avons vu précédemment les cytochromes P450 sont très largement impliqués dans la synthèse de métabolites secondaires et plus spécifiquement dans la synthèse de composés de défense dits allélopathiques. En effet, toutes ces molécules ont en commun de faire intervenir au cours de leur biosynthèse, des enzymes appartenant à la famille des P450.

II.8.3.2.1 Organisation

Des programmes de recherche de séquençage de P450 ont été entrepris afin de mieux comprendre les mécanismes moléculaires mis en place par les plantes, contribuant à développer cette formidable diversité de métabolites secondaires.
Ils ont permis de révéler un très grand nombre de P450 exprimé chez les plantes. À ce jour plus de 4000 séquences de P450 sont recensées dans des bases de données. Cette famille d'enzymes constitue l'une des plus grandes superfamilles de gène appartenant au génome de plante (http//drnelson.utmen.edu/CytochromeP450.html). À titre d'exemple, *Arabidospsis*, le

riz, le peuplier et la vigne contiennent respectivement 246, 356, 312, et 457 gènes appartenant à la famille des P450. Ce qui représenterait plus de 1 % du nombre total de gènes présents chez ces espèces (Nelson et al., 2004). L'analyse transcriptomique a permis de révéler l'importance de cette superfamille, mais la majeure partie des fonctions de tous ces P450 recensés, reste inconnue à ce jour. Chez *Arabidopsis*, seulement une soixantaine de gènes est caractérisée par expression hétérologue ou par analyse de mutant sur les 276 P450 identifiés (Chapple, 1998; Schuler and Werck-Reichhart, 2003).

Tout comme dans le génome bactérien, de récentes études ont mis en évidence la présence de clusters de gène chez les eucaryotes comme les levures, les insectes et les plantes. Ces gènes fonctionneraient et se réguleraient de la même manière que des opérons bactériens (Blanco et al., 2008; Michalak, 2008; Osbourn and Field, 2009).

Chez les plantes, les régions du génome regroupées en clusters de P450 sont mises en évidence dans des voies de biosynthèse bien spécifiques : ce sont les gènes impliqués dans la synthèse du thalianol chez *Arbidopsis (Wilmouth et al., 2002)*, des benzoxazinoides chez le maïs (Frey et al., 1997), de l'avénacine chez l'avoine (Qi et al., 2004) et le momilactone chez le riz (Shimura et al., 2007; Okada et al., 2009).

L'importante diversité de gènes codant pour des P450 présents dans le génome des végétaux supérieurs, provient certainement de phénomènes multifactoriels tels que la duplication génique, les mutations aléatoires des séquences nucléotidiques, la néo et la sub-fonctionnalisation (Schuler, 2011). La duplication génique par enjambement (ou crossing-over) inégal semble être à l'origine de la présence de clusters de gènes dans le génome de plante. La présence habituelle d'introns situés à la même position dans certains clusters renforce également l'hypothèse de cette duplication génique (Paquette et al., 2000).

II.8.3.2.2 P450 et synthèse de composés allélopathiques

Les enzymes appartenant à la famille des cytochromes P450 sont largement impliquées dans la synthèse des molécules allélopathiques. Certaines interviennent dans l'hydroxylation des cycles aromatiques des cinnamates qui donneront une grande variété de composés phénoliques tels que les flavonoïdes ou les coumarines (Liu et al., 2009), d'autres hydroxylent les acides gras (Kandel et al., 2006) indispensables pour se protéger des rayons UV et des pathogènes (Wellesen et al., 2001), ou enore interviennent dans l'oxydation des diterpenoïdes qui protègent les conifères des herbivores (Keeling and Bohlmann, 2006).

Certaines familles de P450 ont conservé leurs fonctions et leurs spécificités à travers l'évolution dans la plupart des plantes supérieures. On les retrouve chez les algues vertes, mais aussi chez toutes les plantes vasculaires. C'est le cas pour la famille de CYP51 (que l'on retrouve dans tous les règnes) et CYP710, toutes deux impliquées dans la synthèse de stérols (Bak et al., 1997; Morikawa et al., 2006). C'est aussi le cas de la famille CYP97 qui métabolise les xanthophylles (Morikawa et al., 2006). Il en est de même pour les familles CYP86, CYP94 ou CYP703, qui sont impliquées dans l'hydroxylation des acides gras et que l'on retrouve dans la plupart des plantes vasculaires (Mizutani and Ohta, 2010).

Hormis ces exemples, les familles et sous-familles de P450 impliquées dans la synthèse de composés allélochimiques se retrouvent dans un groupe de plantes bien plus restreint (Schuler, 2011). La nature et la distribution des composés allélochimiques sont bien spécifiques d'une espèce de plante. Il en est de même pour les familles et sous-familles de gènes impliqués dans leur métabolisme. Pour ne citer que quelques exemples qui illustrent la spécificité de ces P450 en fonction des familles de plantes, le CYP79F1 et le CYP79F2 intervenant dans la synthèse de composés aliphatiques glucosinolés, sont identifiés uniquement chez les brassicacées, dont *Arabidopsis thaliana* (Chen et al., 2003). Chez certaines Fabacées telles que *Lotus japonicus*, la synthèse de linamarine fait intervenir CYP79D (Forslund et al., 2004). Chez les Apiacées telles qu'*Ammi majus*, l'intervention de CYP71AJ dans la synthèse de psoralène ou d'angélicine a été caractérisée par Larbat *et al* (Larbat et al., 2007; Larbat et al., 2009).

II.8.3.3 Détoxication des composés toxiques par les insectes

Malgré cette formidable spécialisation des plantes pour les synthèses de ces composés secondaires, les insectes ont réussi à subsister et à s'adapter à la présence de métabolites plus ou moins toxiques dans leur alimentation. Par quel(s) mécanisme(s) y sont-il parvenus et pour combien de temps ? Telles sont les questions qui alimentent encore la recherche.

II.8.3.3.1 Diversité, spécificité et régulation des P450 d'insectes

Parmi les enzymes capables d'inactiver ces métabolites toxiques, les P450 d'insectes sont des enzymes clefs qui effectuent une hydroxylation et/ou une époxidation de ces molécules afin d'atténuer leur toxicité et de les éliminer (Figure 21). Un nombre important de gènes codant pour des P450 est recensé aujourd'hui chez les insectes. Cependant, la quantité de gènes

codant pour les P450 n'est pas comparable à celle rencontrée dans le génome de plante. À titre d'exemple, chez *Drosophilia melanogaster*, 87 gènes codant pour des P450 ont été répertoriés et 46 gènes chez *Apis mellifera* (Tijet et al., 2001; Schuler, 2011) ; ces chiffres sont beaucoup moins importants que le nombre de P450 rencontrés chez *Arabidopsi*s ou la vigne.

Des recherches portant sur la nature et la spécificité des P450 rencontrés chez différents types d'insectes, ont été entreprises afin de mieux comprendre la divergence et les évolutions génétiques qui ont permis aux insectes de s'adapter à la présence de composés toxiques. Pour cela, des recherches ont porté sur des insectes possédant un régime alimentaire bien spécifique (Berenbaum, 1990; Schuler, 1996b; Wen et al., 2006a; Schuler, 2011). Trois régimes alimentaires ont été déterminés : les insectes monophages, comprenant *D. pastinacella*, qui se nourrissent de plantes ne synthétisant qu'un seul type de furocoumarine (imperatorine) ; les insectes oligophages comportant *P. polyxenes*, qui sont capables de détoxifier l'ensemble des furocoumarines linéaires ; et enfin les insectes polyphages avec *P. glaucus* qui métabolisent à la fois les furocoumarines linéaires et angulaires.
Il a été identifié, pour ces trois types d'insectes, des P450 capables de détoxifier de manière très spécifique les furocoumarines (Tableau 2). Les insectes monophages ne possèdent pas le même équipement enzymatique (P450) que les insectes polyphages ; cependant la plupart de ces P450 appartiennent à la même sous famille les CYP6B (Tableau 2). Ce sont cette diversité et cette spécificité d'activité des P450 « détoxifiant » qui sont à l'origine de cette divergence de régime alimentaire rencontrée chez les insectes (Berenbaum, 1990; Wilmouth et al., 2002; Wen et al., 2003; Wen et al., 2006b).

Régime alimentaire	**Espèces**	**P450 étudié**
Monophages	*Drosophilia pastinacella*	CYP6AB3
Oligophages	*Papilio polyxenes*	CYP6B1, CYP6B3
Polyphages	*Papilio glaucus*	CYP6B4, CYP6B17, CYP6B33
	Helicoverpa zea	CYP6B8

Tableau 2 : P450 exprimés en fonction du régime alimentaire pratiqué par l'espèce d'insectes (Hung et al., 1995; Petersen et al., 2001; Wen et al., 2006b; Mao et al., 2007; Mao et al., 2008).

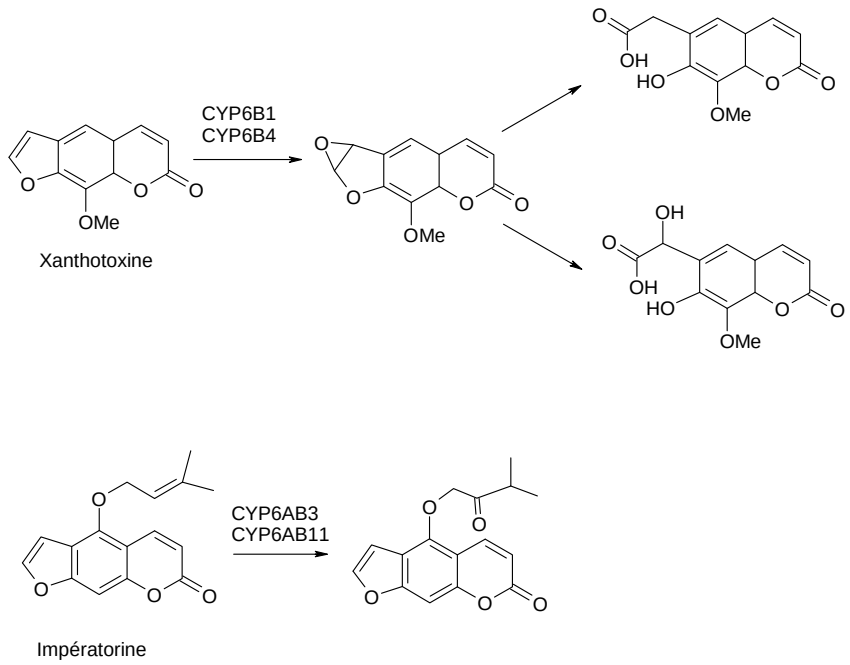


Figure 21 : Détoxication des furocoumarines par des P450 identifiés chez des insectes.

La régulation du niveau d'expression de certains cytochromes P450 d'insectes provient d'une élicatation par des métabolites toxiques. Cette régulation constitue une réponse rapide face à la défense biochimique mise en place par les plantes. A titre d'exemple, CYP6B1 et CYP6B3 deux P450 exprimés chez *P.polyxenes* ne sont pas induite par les mêmes éliciteurs. En effet, l'induction de CYP6B1 est engendrée par la présence de deux furocoumarines uniquement : la xanthotoxine (un furocoumarine linéaire) et dans une bien moindre mesure par l'angélicine (une furocoumarine angulaire) (Hung et al., 1995). Inversement l'expression de CYP6B3 est induite en présence de nombreuses furocoumarines linéaires et angulaires (Petersen et al., 2001).

Chez un autre polyphage comme *H. zea*, la synthèse de CYP6B8, qui métabolise un large panel de furocoumarines et autres insecticides (cypermetrine, diazino), est induite par la majeure partie des dérivées phénylpropanoïdes dont les furocoumarines linéaires et angulaires (Wen et al., 2009).

La régulation de l'expression des cytochromes P450 est donc en partie régulée par la présence d'éliciteurs. Néanmoins, il est démontré que les cytochromes P450 des insectes polyphages sont sensibles à une plus grande variété de composés que les insectes monophages (Cohen et al., 1992; Wen et al., 2009; Zeng et al., 2009). De nombreux rétrocontrôles directs et/ou

indirects sont donc mis en place par le métabolisme des insectes afin de s'adapter le mieux possible à la défense biochimique des plantes.Cette complexité de régulation du niveau d'expression des P450 explique en partie le niveau de défense disparate entre les insectes.

II.8.3.3.2 Relations structures-fonctions des p450 d'insectes

La résolution de structure cristalline de différents P450 a permis de réaliser des analyses tridimensionnelles des P450 membranaires d'insecte. Ceci permet de comprendre les relations fonctionnelles associées à la structure des différents P450 impliqués dans le catabolisme des composés allélopathiques. En comparant les sites catalytiques du CYP6B1 de *P.polyxenes*, un oligophage, et du CYP6B8 de *H.zea*, un polyphage, il est constaté que de nombreux résidus aromatiques, situés sur les SRS2, SRS5 et SRS6 de CYP6B1, sont remplacés par des résidus non aromatiques chez CYP6B8 (Wilmouth et al., 2002). Cette modification entraîne un changement de polarité autour du site actif, modifiant le positionnement des substrats. Ces changements expliquent la faible spécificité de CYP6B8 vis-à-vis des composés toxiques et la capacité de ce cytochrome P450 de métaboliser un plus large panel de substrats toxiques (Wilmouth et al., 2002; Baudry et al., 2003).

De même, il est démontré que CYP6B4 de *P.glaucus*, insecte polyphage métabolisant un grand nombre de produits toxiques, possède un site actif beaucoup plus grand pour reçevoir le substrat que celui du CYP6B1. Ceci s'explique par le fait que les résidus volumineux qui constituent le site actif de CYP6B1 ont été remplacés par des résidus plus petits (Wilmouth et al., 2002).

Au fil de l'évolution, il semblerait que certains insectes se soient ainsi adaptés au nouvel arsenal de molécules mis en place par les plantes.

II.8.3.4 P450 et coévolution

Le terme de coévolution est introduit pour la première fois en 1964 par Ehrlich et Raven dans une étude basée sur l'interaction entre les plantes et les papillons (Ehrlich and Raven, 1964).

Puis Janzen en 1969 redéfinit le terme comme un changement évolutif d'un trait chez une espèce en réponse à un trait chez une deuxième espèce, suivi d'une adaptation évolutive de cette deuxième espèce (Janzen and Muller, 1969).

Malgré le rôle prédominant du phénomène de coévolution entre les insectes et les plantes dans l'acquisition par les végétaux de nouvelles molécules de défense, relativement peu d'études

ont été effectuées sur une vision quantitative et qualitative de l'intensification des composés allélopathiques (Becerra et al., 2009). Dans la famille des Apiacées, certaines plantes ne synthétisent que des furocoumarines linéaires (*Ammi majus*) (Pokrovskii et al., 2009) et d'autres produisent des furocoumarines linéaires et angulaires (*Pastina sativa*) (Stein and Posocco, 1984). À ce jour, aucune espèce décrite ne produit que des furocoumarines angulaires. Ce constat semblerait confirmer l'hypothèse selon laquelle les furocoumarines linéaires sont apparues avant les furocoumarines angulaires. La pression de sélection et l'adaptation des insectes aux premières furocoumarines linéaires ont certainement fait émerger la synthèse de furocoumarines angulaires chez les plantes afin de se protéger contre les insectes les mieux adaptés (Berenbaum and Feeny, 1981). Cette évolution et adaptation des plantes sur un point de vue moléculaire est en partie rendue possible *via* l'émergence de nouvelles fonctions des cytochromes P450 qui font suites aux nombreuses pressions de sélection.

Nous avons donc bien dans cette relation plante-insecte, un phénomène de coévolution suivant la définition de Janzen (Janzen and Muller, 1969). En effet, les insectes se sont adaptés une première fois à la défense métabolique des plantes (furocoumarines linéaires), celles-ci ont alors évolué pour former des métabolites plus complexes (furocoumarines angulaires), et certains insectes ont encore une fois évolué afin de réussir à détoxifier ces nouvelles molécules angulaires, par l'émergence une fois encore de nouvelles fonctions de leurs cytochromes P450.

Un autre exemple de coévolution est caractérisé avec *Bursera*, un arbre de la famille des Burseracées, et *Blepharida* un coléoptère herbivore prédateur spécifique de *Bursera* (Becerra, 2001; Becerra et al., 2009)

Bursera synthétise des terpènes et principalement des mono et sesquiterpènes (Evans et al., 2000; Evans and Becerra, 2006). Cependant certaines espèces de *Bursera,* ne produisent qu'un seul composé terpénique dominant alors que d'autres espèces en synthétisent plus de 35 differents (Evans and Becerra, 2006; Becerra, 2007). De même, certaines espèces de *Blepharida* sont des espèces monophages et n'attaquent qu'une seule espèce de *Bursera* alors que d'autres sont oligophages ou polyphages pouvant être associées parfois à 12 espèces de Bursera différentes (Becerra, 2007).

L'étude réalisée par Beccera et son équipe en 2009 (Becerra et al., 2009) démontre statistiquement qu'il existe une forte corrélation entre l'apparition des nouvelles espèces de *Blepharida* et l'apparition de nouveaux composés de défense (diversité et complexité) synthétisés par *Bursera* au cours du temps.

Cette récente étude épidémiologique relate donc un autre phénomène de coévolution entre un arbre et un insecte qui aboutit à la sélection de nouvelles molécules allélopathiques et à la sélection de nouvelles espèces prédatrices. Quand on connait l'importance du rôle des cytochromes P450 dans la synthèse des composés secondaires et plus spécifiquement dans la synthèse des molécules de défense de type terpénique, il serait intéressant, par la suite, de suivre l'évolution moléculaire de ces enzymes P450 au sein de ces différentes espèces d'arbres et d'insectes.

III Dioxygénases

Il existe deux types d'enzymes qui sont capables d'utiliser le dioxygène atmosphérique afin de réaliser des réactions enzymatiques : les oxydases qui utilisent l'oxygène comme oxydant afin de réduire le dioxygène en peroxyde d'hydrogène ou en H_2O, et les oxygénases qui incorporent un atome d'oxygène, provenant du dioxygène, dans un substrat. Dans cette dernière catégorie se trouvent des mono-oxygénases catalysant l'addition d'un atome d'oxygène dans le substrat, tel que les P450 et les dioxygénases.

III.1 Différentes familles de dioxygénases

Les dioxygénases regroupent différentes familles d'enzymes qui sont impliquées dans de nombreux types de réactions :

- les catéchol dioxygénases. Elles sont impliquées dans la dégradation de composés aromatiques chez les bactéries, par clivage oxydatif du groupement catéchol (Bugg and Winfield, 1998) ;

- les arènes dioxygénases qui dégradent les hydrocarbures aromatiques par une *cis*-dihydroxylation du cycle aromatique (Djokic et al., 2011) ;

- les lipoxygénases qui catalysent la conversion d'acides gras polyinsaturés tels que l'acide linoléique chez les plantes, en hydroperoxydes de lipide (Frazee et al., 1998).

- les dioxygénases α-cétoglutarates dépendantes qui catalysent des mono-hydroxylations, des désaturations, des épimérisations de composés très divers en utilisant comme cosubstrat l'α-cétoglutarate ;

- de nombreuses autres dioxygénases métallo dépendantes ont aussi été caractérisées, telles que l'indoléamine 2,3-dioxygénase, la quercétine 2,3 dioxygénase, la β-carotène dioxygénase (Tranchimand et al., 2008; Ilg et al., 2009; Yuasa et al., 2011).

III.2 Les dioxygénases α-cétoglutarates-dépendantes

Parmi les différentes familles de dioxygénases, les α-cétoglutarates-dépendantes, comme leur nom l'indique, nécessitent la présence d'un cosubstrat: l' α-cétoglutarate. Celui-ci subit une décomposition oxydative pour former du CO_2 et du succinate. Cette oxydation est couplée à l'hydroxylation d'un substrat. Dans la majorité des cas, elle requiert la présence de Fer(II)

pour être fonctionnelle. On parle alors de dioxygénase Fe(II)/α-cétoglutarate dépendante. Ces dioxygénases fonctionnent la plupart du temps avec trois substrats : l' α-cétoglutarate, l'oxygène moléculaire et un substrat spécifique de l'enzyme (Purpero and Moran, 2007) (Figure 22).

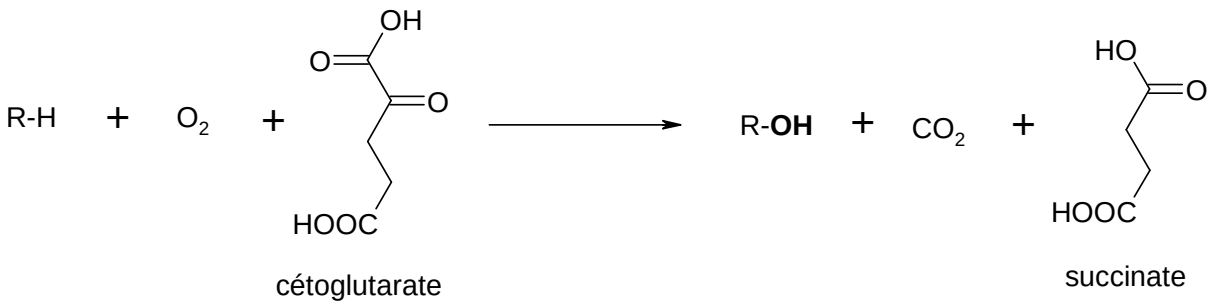


Figure 22 : Réaction générale effectuée par une dioxygénase α-cétoglutarate dépendante.

III.2.1 Une diversité de réactions

Les dioxygénases α-cétoglutarate-dépendantes sont des enzymes qui jouent un rôle très important dans les fonctions physiologiques de la cellule. Elles sont impliquées dans de nombreuses réactions d'hydroxylation mais aussi d'époxydation, d'épimérisation ou d'halogénation (Purpero and Moran, 2007).

III.2.1.1 Modification de protéines

La prolyl 4-hydroxylase est la première dioxygénase Fe(II)/α-cétoglutarate dépendante à être identifiée (Hutton Jr et al., 1966). Cette dioxygénase hydroxyle en position *trans* le résidu aminoacyle prolyle (Figure 23). Chez les plantes, elle a pour rôle d'hydroxyler des composés clefs de la matrice extracellulaire et des tissus reproductifs tels que les protéines arabinogalactan ou les glycoprotéines riches en séquences prolines (Wu et al., 2001). Récemment, leurs rôles ont été mis en évidence d'un point de vue moléculaire dans la sénescence des fleurs de *Dianthus caryophyllus* (Vlad et al., 2010).

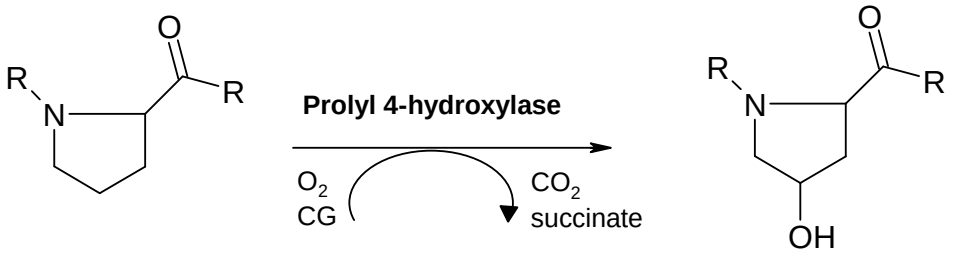


Figure 23 : La première dioxygénase Fe(II)/α-cétoglutarate dépendante identifiée sur le plan moléculaire : La Prolyl 4-hydroxylase.

D'autres dioxygénases Fe(II)/α-cétoglutarates-dépendantes sont impliquées dans la modification de protéines comme la lysyl hydroxylase qui hydroxyle le triplet X-Lys-Gly du collagène afin de permettre la fixation du galactose ou du gucosylgalactose (Kivirikko and Pihlajaniemi, 1998; Passoja et al., 1998) ou comme l'aspartyl β-hydroxyalse qui modifie l'acide aspartique et qui a été clonée à partir de différents organismes y compris l'homme (Jia et al., 1992; Korioth et al., 1994).

L'intérêt pour ces enzymes s'est intensifié lors de la découverte en 2001 de leurs rôles importants dans les réponses à des phénomènes d'hypoxies lors de la formation de tumeurs. En effet, la prolyl 4-hydroxylase intervient directement dans l'hydroxylation de facteur d'hypoxie exprimé dans les tissus ischémiés, de même que l'aspartyl β-hydroxyalse (Ivan et al., 2001; Song et al., 2010; Yang et al., 2010).

III.2.1.2 Réparation de l'ADN alkylé

Des composants endogènes, tels que la S-adénosylméthionine, et des produits chimiques, tels que les agents S_N1 (N-methyl-N'-nitro-N-nitrosuguanidine et méthylnitrosourée) ou les agents S_N2 (méthylméthanesulfonate), sont impliqués dans des processus d'alkylation de l'ADN (Sedgwick and Lindahl, 2002). Cet ADN alkylé est néfaste pour la cellule parce qu'il perturbe les interactions protéine-ADN existantes au sein de la cellule, et nécessaires à la bonne régulation de l'expression des gènes. Afin de se protéger contre les effets mutagènes et toxiques de ces alkylations nucléotidiques, les cellules synthétisent une variété d'enzymes impliquées dans la déalkylation des bases azotées. AlkB code pour une dioxygénase Fe(II)/α-

cétoglutarate dépendante qui répare l'ADN en éliminant les groupements 1-alkyle de l'adénine (Figure 24) et 3-alkyle de la cytosine (Trewick et al., 2002; Aas et al., 2003).

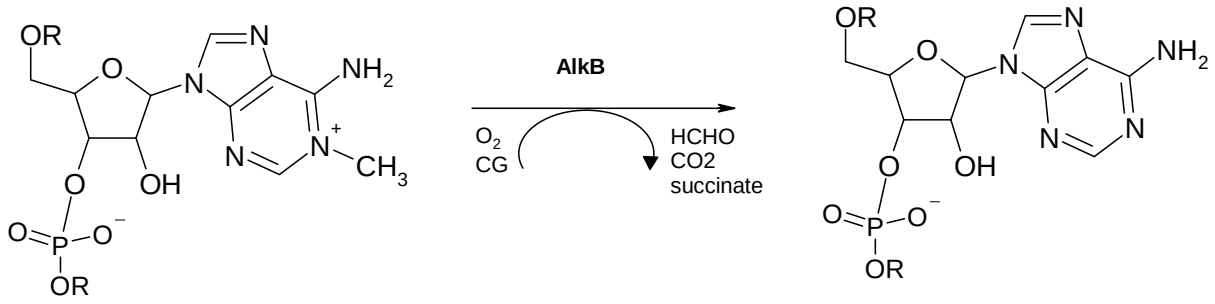


Figure 24 : Intervention de l'AlkB, une dioxygénase Fe(II)/α-cétoglutarate dépendante, dans le processus de réparation de l'ADN alkylé.

III.2.1.3 Biosynthèse des antibiotiques

Les antibiotiques sont des molécules qui détruisent ou bloquent la croissance des bactéries. Elles sont produites la plupart du temps par des micro-organismes, des champignons ou bactéries. La synthèse de nombreux antibiotiques, tels que l'acide clavulanique, les céphalosporines, les carbapénèmes, la fosfomycine font intervenir des dioxygénases Fe(II)/α-cétoglutarate-dépendantes. Ces enzymes interviennent à différents stades de la synthèse et réalisent des réactions d'hydroxylations, de désaturations, ou de cyclisations (Chen et al., 2000; Yin et al., 2003).

À titre d'exemple, la clavaminate synthase (CAS) identifiée chez *Steptomyces antibioticus* catalyse la synthèse d'acide clavulanique en intervenant dans trois différentes étapes oxydatives (Janc et al., 1995). Cette enzyme trifonctionnelle intervient d'abord en hydroxylant le β-lactam (Figure 25), puis la CAS réalise une réaction de cyclisation (Figure 26.A) et de désaturation de l'acide proclavaminique (Figure 26.B) (Lloyd et al., 1999).

De même, la deacétoxycéphalosporine C synthase (DAOCS) et la deacétocéphalosporine C synthase (DACS) sont des dioxygénses Fe(II)/α-cétoglutarates-dépendantes qui catalysent la formation de céphalosporine à partir de pénicilline N (Baldwin et al., 1987; Dotzlaf and Yeh, 1989).

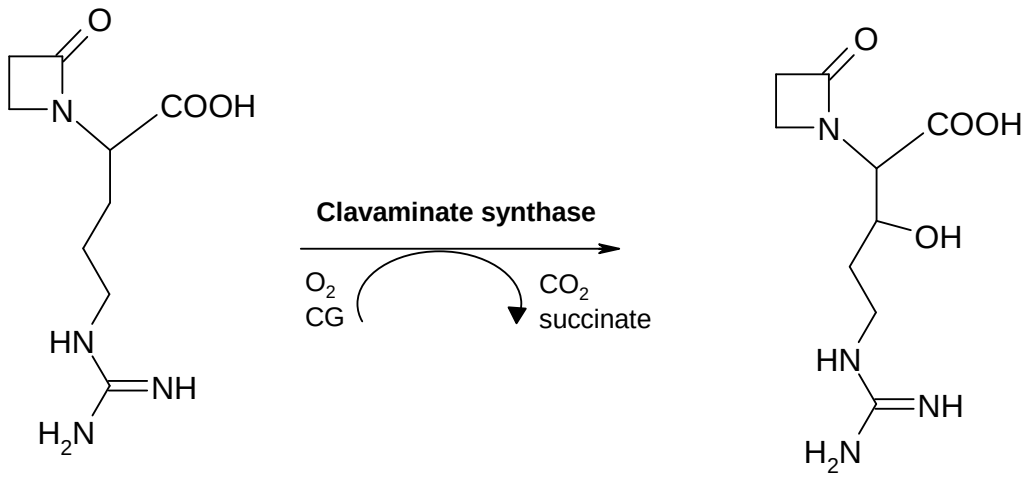


Figure 25 : Hydroxylation du β-lactam par la clavaminate synthase.

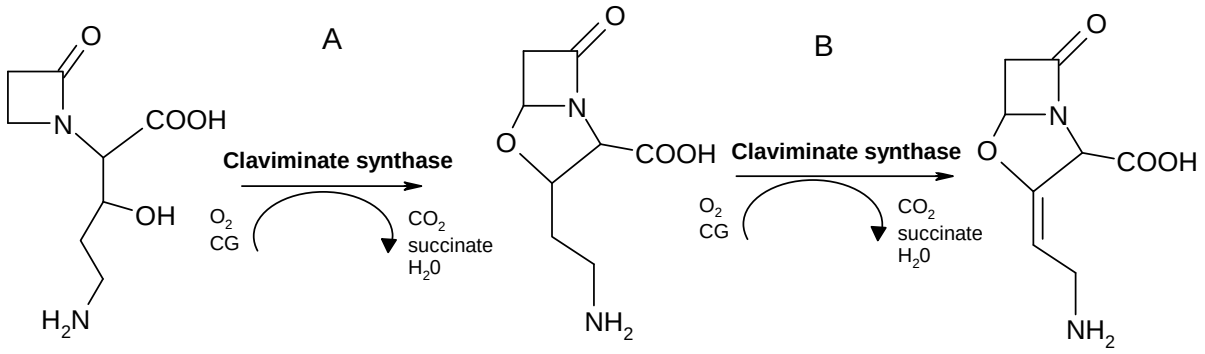


Figure 26 : Réaction de cyclisation (A) et de désaturation (B) par la clavaminate synthase dans la synthèse de l'acide proclavaminique.

III.2.1.4 Le métabolisme des lipides

Les dioxygénases Fe(II)/α-cétoglutarate-dépendantes jouent un rôle primordial dans le métabolisme de certains lipides chez les bactéries, les plantes et les hommes.

Ces lipides proviennent, en totalité ou en partie, de condensations de thioesters basées sur des carbanions et/ou de condensations d'unités d'isoprène basées sur des carbocations. Ceci définit deux types de voies de biosynthèse suivant le squelette de base utilisé : un thioester ou un isoprène.

L'acide phytanique est un acide gras issu des aliments. C'est un lipide qui dérive du phytol et qui provient principalement de la dégradation de la chlorophylle. Comme tous les acides gras, l'acide phytanique est dégradé par β-oxydation dans le peroxysome. Cependant l'acide phytanique doit d'abord être activé en ester de coenzyme A par une coumaroyle CoA ligase pour donner le phytanoyl CoA (Watkins et al., 1994). Ce composé estérifié est ensuite hydroxylé en position C2 par une dioxygénase Fe(II)/α-cétoglutarate, la phytanoyl-CoA hydroxylase (Figure 27) avant d'être dégradé en acide pristanique par une aldéhyde deshydrogénase et une thiamine pyrophosphatase (Jansen et al., 1996).

La déficience de cette dioxygénase dans le métabolisme de l'acide phytanique est à l'origine de la maladie de Refsum provoqué par l'accumulation d'acide phytanique entraînant des lésions au niveau de la rétine et du système nerveux central (Wanders et al., 2011).

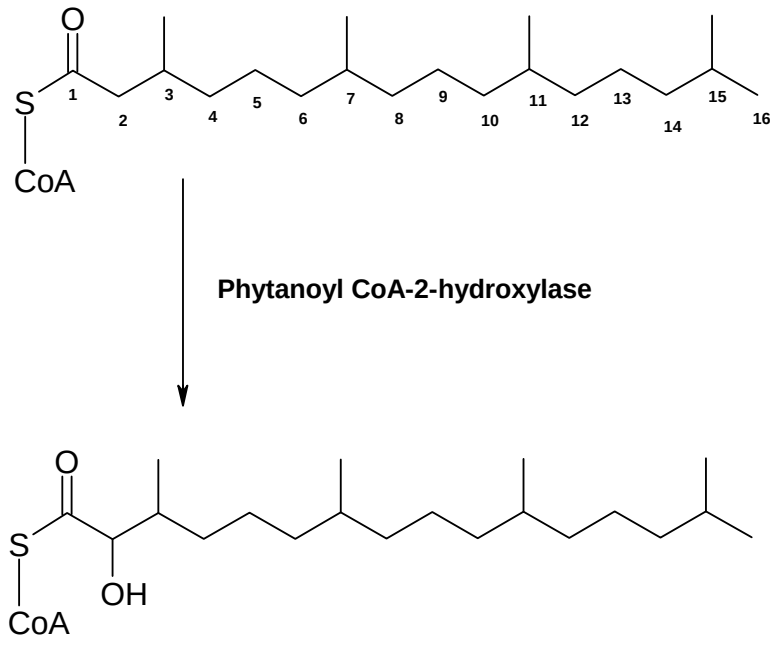


Figure 27 : Intervention de la phytanoyl CoA -2-hydroxylase dans le catabolisme de l'acide phytanique.

Pour citer un autre exemple, deux dioxygénases Fe(II)/α-cétoglutarates dépendantes interviennent dans la synthèse de la carnitine, un transporteur des acides gras à longue chaîne à l'intérieur des mitochondries (Ruetschi et al., 1993; Swiegers et al., 2002)

III.2.1.5 Biosynthèse des métabolites secondaires

De nombreux composés secondaires font intervenir des dioxygénases Fe(II)/α-cétoglutarates dépendantes lors de leur synthèse, comme les flavonoïdes, les coumarines, les gibbérellines, ou les alcaloïdes.

III.2.1.5.1 Les phénylpropanoïdes

III.2.1.5.1.1 Les flavonoïdes

Les dioxygénases Fe(II)/α-cétoglutarates-dépendantes impliquées dans la synthèse des flavonoïdes sont les dioxygénases les plus étudiées. Ces flavonoïdes tels que les flavones, les isoflavones, les flavonols, les anthocyanes ont pour fonction de colorer la plante, de la protéger contre les rayons ultraviolets ou contre les organismes phytopathogènes (Cushnie and Lamb, 2011; Piri et al., 2011). La voie de biosynthèse des flavonoïdes fait intervenir à ce jours trois hydroxylases Fe(II)/α-cétoglutarate-dépendantes : la flavanone 3β- hydroxylase (FHT), la flavonol synthase (FLS) et la flavone synthase I (FNS I).

La flavanone 3β- hydroxylase isolée en premier chez *Petunia hybrida*, catalyse la conversion du flavanone en *trans*-dihydroflavonol (Figure 28) (LukaÄ□in and Britsch, 1997).

La flavanol synthase, isolée pour la première fois dans les pétales de *Petunia hybrida*, introduit une double liaison entre le carbone C2 et C3 du *trans*-dihydroflavonol (Holton et al., 1993) pour former le flavonol (Figure 28).

Un autre clone de la FLS, isolée chez *Citrus unshiu*, a révélé une seconde activité identique à la FHT. En effet, la FLS de *C. unshiu* hydroxyle, en plus de son activité déjà décrite, hydoxyle la flavanone en position C3 pour former le *trans*-dihydroflavonol (LukaÄ□ina et al., 2003). Cette hydroxylase Fe(II)/α-cétoglutarate dépendante présente donc une double activité (désaturation et hydroxylation) sur des substrats différents (*trans*-dihydroflavanol et flavanone).

La flavone synthase I, isolée et clonée à partir du persil, transforme le flavanone en flavone par insertion d'une double liaison entre les carbones C2 et C3 (Bredebach et al., 2011).

Il est intéressant de constater que cette dernière étape est aussi catalysée par une enzyme appartenant à la famille des P450, la flavone synthase II. Cette enzyme appartient à la famille CYP93B. Récemment, une flavone synthase II (CYP93B16) a été identifiée et caractérisée chez le soja *(Glycine max L.)* (Fliegmann et al., 2010). Cette enzyme se retrouve dans la plupart des végétaux supérieurs qui synthétisent des flavonoïdes (Bredebach et al., 2011), alors que la flavone synthase I, principalement rencontrée chez les Apiacées, a été dernièrement mise en évidence chez une monocotylédone, le riz (Jeong et al., 2008).

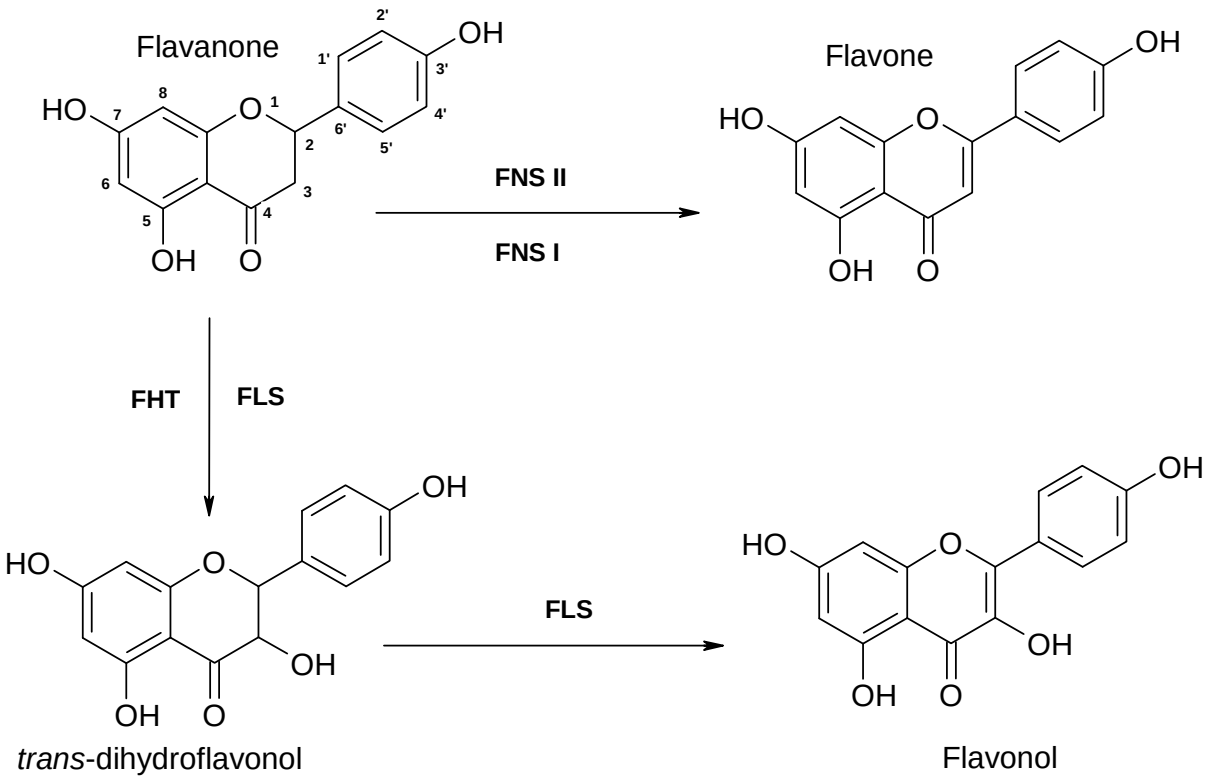


Figure 28 : Représentation d'une partie de la voie de biosynthèse des flavonoïdes.

FNS I = Flavone synthase I, FNS II = Flavone synthase II, FHT = Flavanone 3β-hydroxylase, FLS = Flavonol synthase.

De nombreuses autres dioxygénases Fe(II)/α-cétoglutarate-dépendantes, telles que l'anthocyanidine synthase (ANS) ou la flavonol 6-hydroxylase (F6H), interviennent dans la voie de biosynthèse des flavonoïdes (Wilmouth et al., 2002; Anzellotti and Ibrahim, 2004).

III.2.1.5.1.2 Les coumarines

La scopolétine est une coumarine qui dérive de la voie de biosynthèse des phénylpropanoïdes et que l'on retrouve notamment chez *Arabidopsis thaliana*. Elle est impliquée dans des mécanismes de défense de la plante contre les herbivores. Comme cité dans le chapitre sur les coumarines, Kai et ses collaborateurs ont récemment décrit l'intervention d'une dioxygénase Fe(II)/α-cétoglutarate dépendante, la féruloyle CoA 6'-hydroxylase (F6'H) dans la synthèse de scopolétine chez *A. thaliana* (Kai et al., 2008). À ce jour, il s'agit de la seule dioxygénase à avoir été identifiée dans cette voie de biosynthèse.

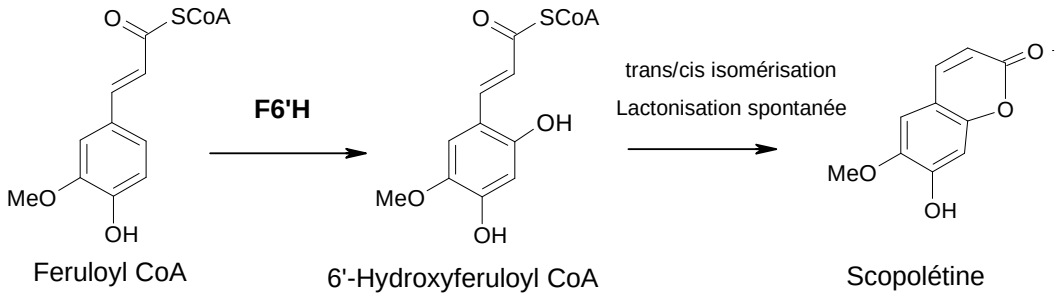


Figure 29 : Synthèse de la scopolétine par la Féruloyle CoA 6'- hydroxylase (F6'H).

III.2.1.5.2 Les dérivées terpéniques : les gibbérellines

Les gibbérellines sont une famille de phytohormones (diterpènes tetracyclique) qui régulent la croissance et le développement des plantes (Hedden and Phillips, 2000). Elles furent mises en évidence pour la première fois par le phytopathologiste Eiichi Kurosawa en 1926 chez *Gibberella fujikuroi*, un ascomycète parasite du riz, qui allonge exagérément les tiges (Grennan, 2006). Les gibbérellines sont nommées G ou Ga suivi d'un chiffre (de 1 à 110).

La première dioxygénase Fe(II)/α-cétoglutarate-dépendante identifiée et caractérisée dans la voie de biosynthèse des gibbérellines est la gibbérelline 7-oxydase, isolée chez *Curcubita maxima*. Elle réalise une oxydation sur le carbone C7 (Figure 30) (Lange et al., 1994b).

Il semblerait que chez d'autres espèces de plantes, cette étape puisse être catalysée par des enzymes appartenant à la famille des cytochromes P450 (Hedden and Phillips, 2000).

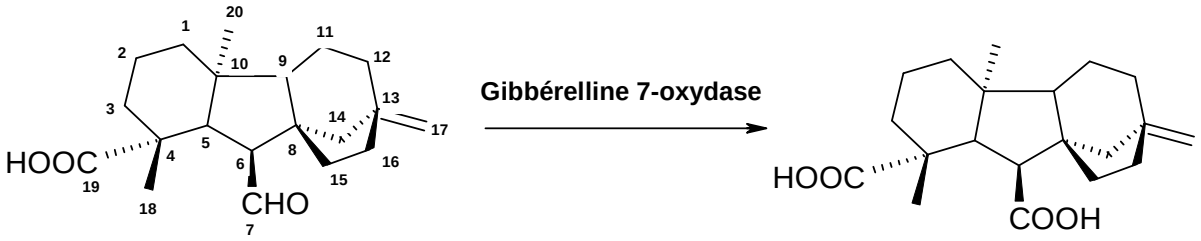


Figure 30 : Oxydation du carbone C7 par une hydroxylase Fe(II)/α-cétoglutarate dépendante : la gibbérelline 7-oxydase.

Un autre exemple, la gibbérelline 20-oxidase est une dioxygénase Fe(II)/α-cétoglutarate dépendante multifonctionnelle qui transforme successivement le groupement méthyl du carbone C20 en alcool (Figure 31.A), en aldéhyde (Figure 31.B) et en acide carboxylique (Figure 31.C) (Hedden, 1999). Cette multifonctionnalité a été caractérisée sur un gène isolé du *Cucurbita maxima* (Lange et al., 1994a). Récemment, 2 autres gibbérellines 20-oxidase ont été identifiées chez *Helianthus annuus* (Carzoli et al., 2008).

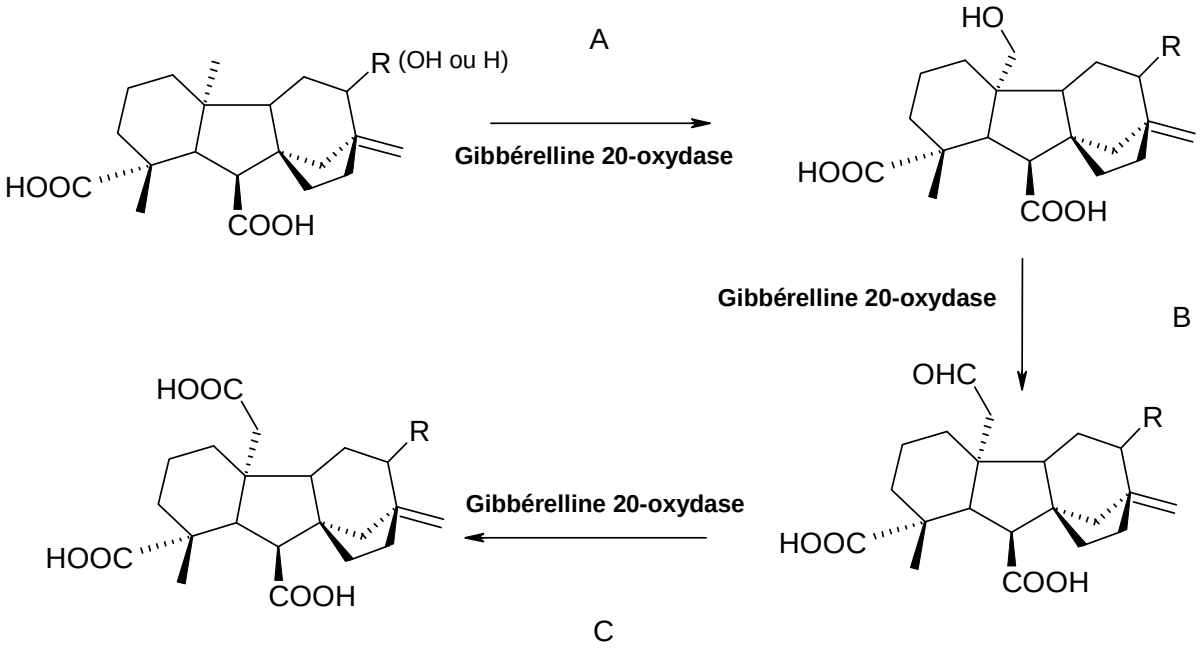


Figure 31 : Synthèse de gibbérelline par une dioxygénase Fe(II)/α-cétoglutarate dépendante multifonctionnelle : la gibbérelline 20-oxydase.

Comme dans la voie de biosynthèse des flavonoïdes, on retrouve des dioxygénases Fe(II)/α-cétoglutarates-dépendantes qui possèdent de multiples fonctions et qui sont capables de métaboliser différents substrats. En outre, on retrouve aussi en fonction de l'espèce une dualité entre les P450 et les dioxygénases pour une réaction donnée (Hedden and Phillips, 2000).

III.2.1.5.3 Les composés azotés : les alcaloïdes

Les alcaloïdes sont des molécules organiques hétérocycliques azotées basiques possédant généralement une activité pharmacologique. La vinblastine et la vincristine sont deux vinca alcaloïdes extraits du *Catharanthus roseus*, la pervenche de Madagascar, que l'on utilise respectivement dans le traitement de la maladie de Hodgkin et des leucémies (Bazarbachi et al., 2011; Björkholm et al., 2011).

La vindoline est un précurseur de la synthèse de vinblastine. L'hydroxylation du 2,3-dihydro-3- hydroxy-N(1)-methyltabersonine est réalisée par une hydroxylase Fe(II)/α-cétoglutarate dépendante sur le carbone C4 (Figure 32) aboutissant à la formation de vindoline (De Carolis et al., 1990).

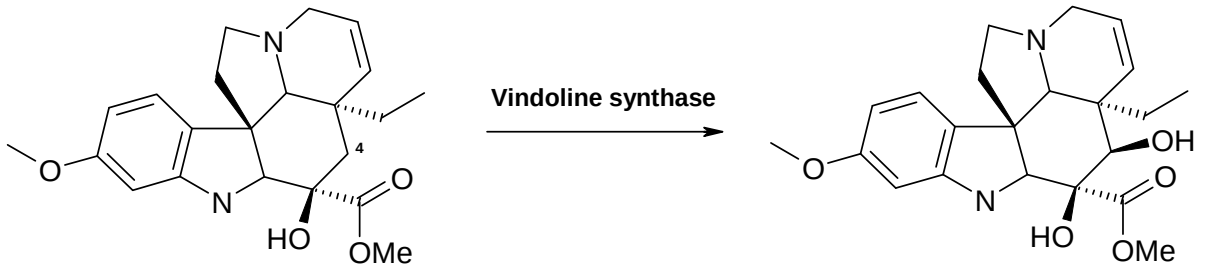


Figure 32 : Etape de la synthèse de vindoline par une dioxygénase Fe(II)/α-cétoglutarate dépendante.

La synthèse de scopolamine, un alcaloïde tropanique présent chez les solanacées telles que *Hyoscyamus niger*, fait aussi intervenir une dioxygénase Fe(II)/α-cétoglutarate dépendante qui possède une double fonctionnalité. En effet, cette hyoscyamine 6β-hydroxylase (H6B) réalise successivement une réaction d'hydroxylation puis une réaction d'époxidation (Hashimoto and Yamada, 1987).

Récemment, une H6B a été identifiée et caractérisée chez *Datura metel L.* (Figure 33) (Pramod et al., 2010).

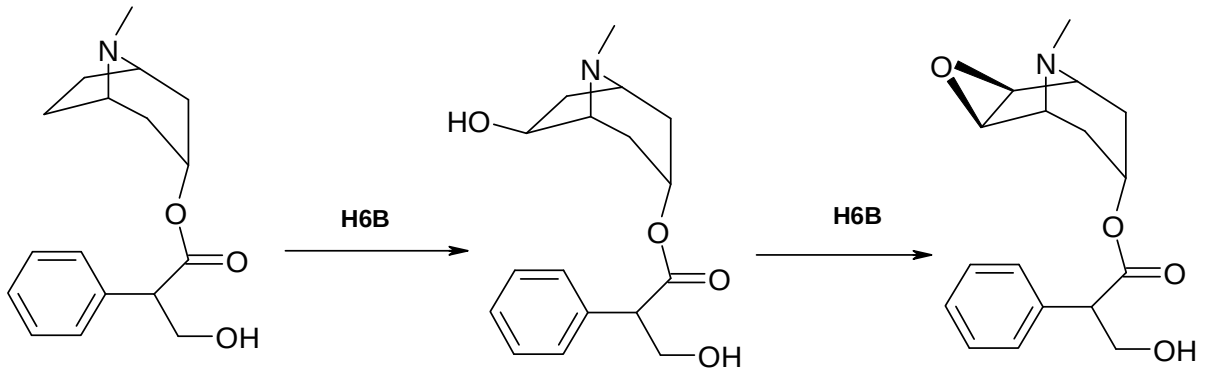


Figure 33 : Intervention de la hyoscyamine 6β-hydroxylase (H6B) dans la synthèse de scopolétine chez *Datura metel L.*

Un autre exemple concerne la voie de biosynthèse de la morphine, un alcaloïde de l'opium utilisé comme analgésique dans le traitement de la douleur.

Hagel et Facchini ont identifié et caractérisé l'intervention de deux dioxygénases Fe(II)/α-cétoglutarate-dépendantes, la thébaïne 6-O-déméthylase (T6ODM) et la codéine O-demethylase (CODM), qui catalysent toutes les deux, une réaction de O-déméthylation lors de la synthèse de morphine chez *Papaver somniferum L.* (Hagel and Facchini, 2010).

La T6ODM réalise une 6-O-déméthylation alors que la CODM réalise une 3-O-déméthylation de la thébaine et de ses dérivés (oripavine et codéine) (Figure 34).

Chez les humains, l'enzyme qui intervient dans la 3-O-déméthylation de la thébaïne et de la codéine appartient à la famille des cytochromes P450, CYP2D6 (Zhu, 2008).

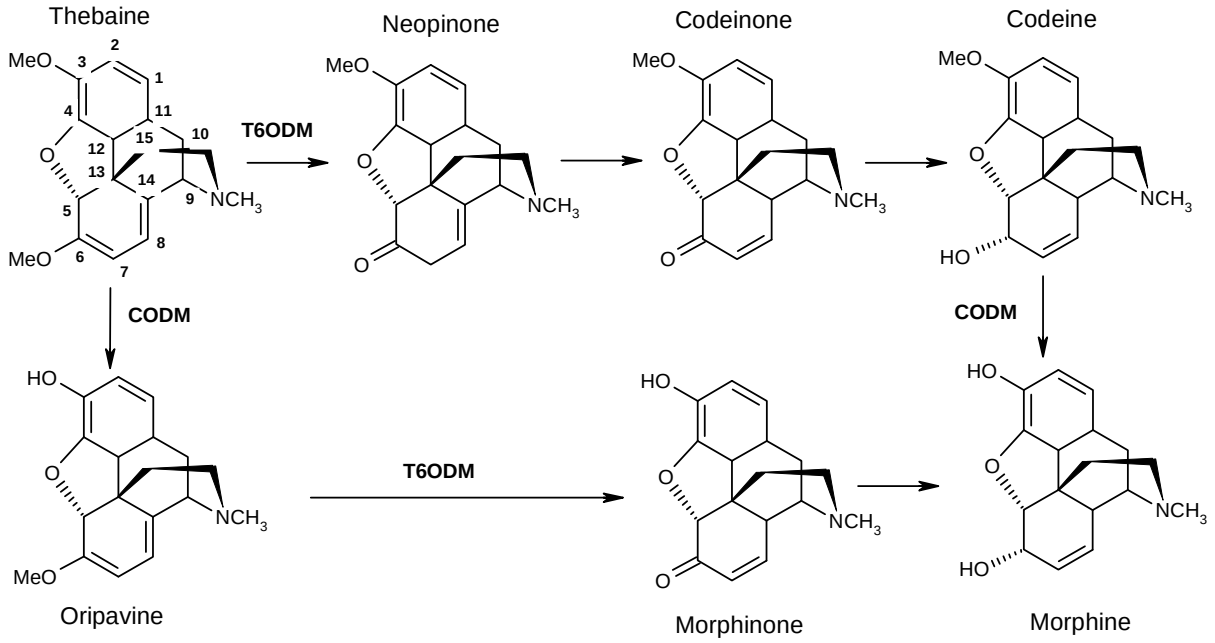


Figure 34 : Voies de biosynthèses des alcaloïdes morphiniques chez *Papaver somniferum L.*
Ces voies décrivent deux chemins de la thébaïne à la morphine chez Papaver somniferum. Deux dioxygénases Fe(II)/α-cétoglutarates dépendantes interviennent dans cette voie de biosynthèse : la thébaïne 6-O-demethylase (T6ODM) et la codéine O-demethylase (CODM).

III.2.2 Relation structure-activité

III.2.2.1 Modélisation tridimensionnelle

Les dioxygénases sont des molécules solubles qui permettent de ce fait une cristallisation beaucoup plus facile que les protéines membranaires. L'acquisition de nombreuses données cristallographiques a permis la modélisation tridimensionnelle des différentes dioxygénases. Ceci a conduit à déterminer l'agencement spatial des acides aminés afin de mieux comprendre les relations structure-fonction. La première résolution de structure tridimensionnelle de dioxygénase provient de la modélisation d'une dioxygénase non α-cétoglutarate dépendante, l'isopenicillin N synthase (IPNS), impliquée dans la synthèse de la pénicilline (Roach et al., 1995). Depuis, l'acquisition et le traitement de nombreuses autres données cristallographiques de dioxygénases ont permis de mettre en évidence des structures communes aux enzymes appartenant à la famille des dioxygénases Fe(II)/α-cétoglutarate dépendante.

Le corps de l'enzyme est constitué par 8 feuillets β formant une poche contenant le site actif (Figure 35.A). D'autres feuillets β, hélices α, et des structures protéiques en boucles sont

additionnés autour de cette cavité. Leurs fonctions consistent à stabiliser la structure de la dioxygénase, à participer aux liaisons interprotéines et à modifier le site de liaison du substrat (Figure 35.B) (Hausinger, 2004).

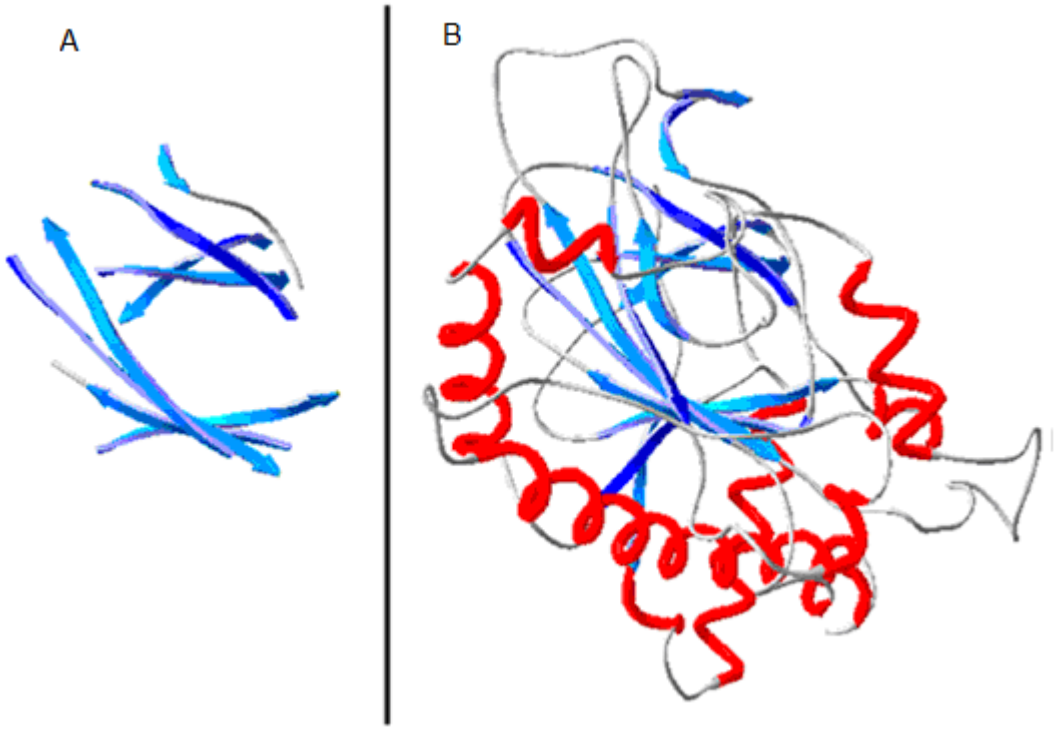


Figure 35 : Modelisation des dioxygénases α-cétoglutarates-dépendantes.

A) Structure du corps de l'enzyme commune à toutes les hydroxylases Fe(II)/α-cétoglutarates dépendantes. B) Résolution de la structure de l'isopenicilline N synthase (IPNS) comprenant le corps de l'enzyme additionné des hélices α, des feuillets β et boucles protéiques indispensables à la stabilité et spécificité de l'enzyme (Roach et al., 1995).

Les dioxygénases sont des enzymes mono ou multimériques. On les retrouve sous différentes conformations mono, di, tétra et même hexamériques. À titre d'exemple, l'isopenicilline N synthase, la clavaminate synthase et l'anthocyanidine synthase sont des dioxygénases monomériques (Roach et al., 1995; Zhang et al., 2000; Wilmouth et al., 2002). La proline 3-hydroxylase est une dioxygénase dimérique (Wilmouth et al., 2002) et l'alkylsulfatase une dioxygénase tétramérique (Müller et al., 2004).

III.2.2.2 Site actif

Le site actif des dioxygénases Fe(II)/α-cétoglutarate-dépendantes ont tous une similarité. En effet, trois acides aminés sont toujours conservés sur la séquence protéique codant pour l'enzyme et permettent la fixation du fer au niveau du site actif. Ce sont les motifs His^1-X-Asp/Glu-X_n-His^2 (Que Jr et al., 1987; LukaÄ□in and Britsch, 1997; Wilmouth et al., 2002; Bugg, 2003) .

```
                 245                 295         305         315
At3g13610    GVGRHSDVSS … … QIMSNGLYKS VEHRVLANGY NNRISVPIFV
AAX89400     GLKRHTDPGT … … HFLSNGRFKN ADHQAVVNSN SSRLSIATFQ
AAB64346     GLIHHEDANC … … QVLSNKKFKS ATHRVVRQKG KERYSFAFFR
```

Figure 36 : Séquences consensus H-X-D/E-X_n-H et R-X-S de trois dioxygénases Fe(II)/α-cétoglutarate-dépendantes.
Une féruloyle CoA hydroxylase (At3g13610) ; une flavonone 3-β hydroxylase (AAX89400) ; une gibbérelline 7-oxidase (AAB64346).

Il existe deux catégories de sites actifs : « in-line » et « off- line ». Cette distinction provient du positionnement du cosubstrat, l'α-cétoglutarate, par rapport au site actif.

Dans le site « in-line », le positionnement du cosubstrat chélaté avec le fer, oriente la formation d'un site de liaison de coordinence avec l'oxygène moléculaire juste à côté de l'His^1 (Figure 38.A) (Elkins et al., 2003; Müller et al., 2004).

Dans le site « off-line », le positionnement du cosubstrat oriente le site de liaison de coordinence avec l'oxygène moléculaire, non pas à côté, mais en face de l'acide aminé His^1 (Figure 38.B)(Wilmouth et al., 2002).

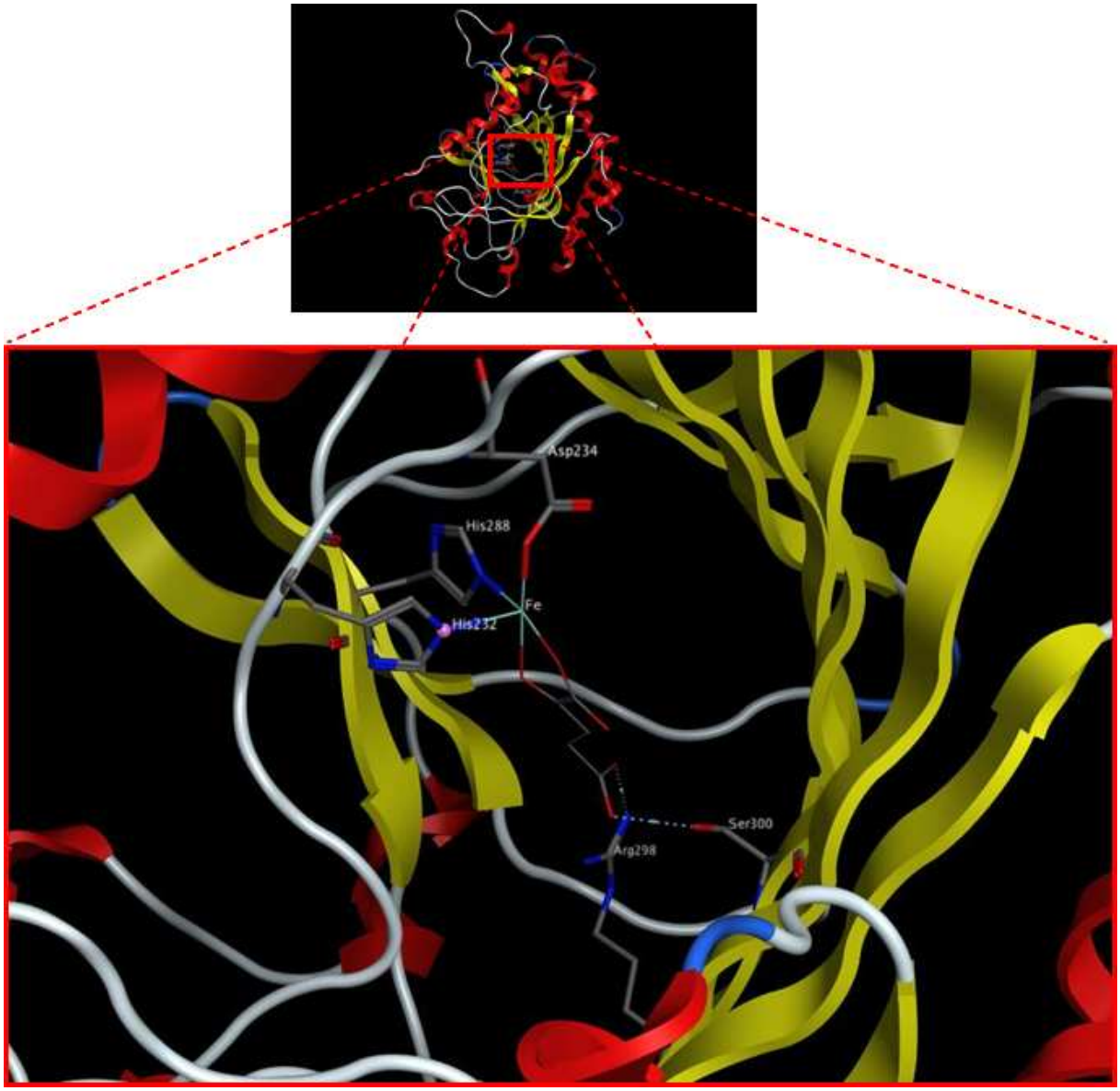


Figure 37 : Localisation tridimensionnelle des séquences consensus H-X-D/E-X_n-H et R-X-S sur la dioxygénase α-cétoglutarate-dépendante (At3g13610).

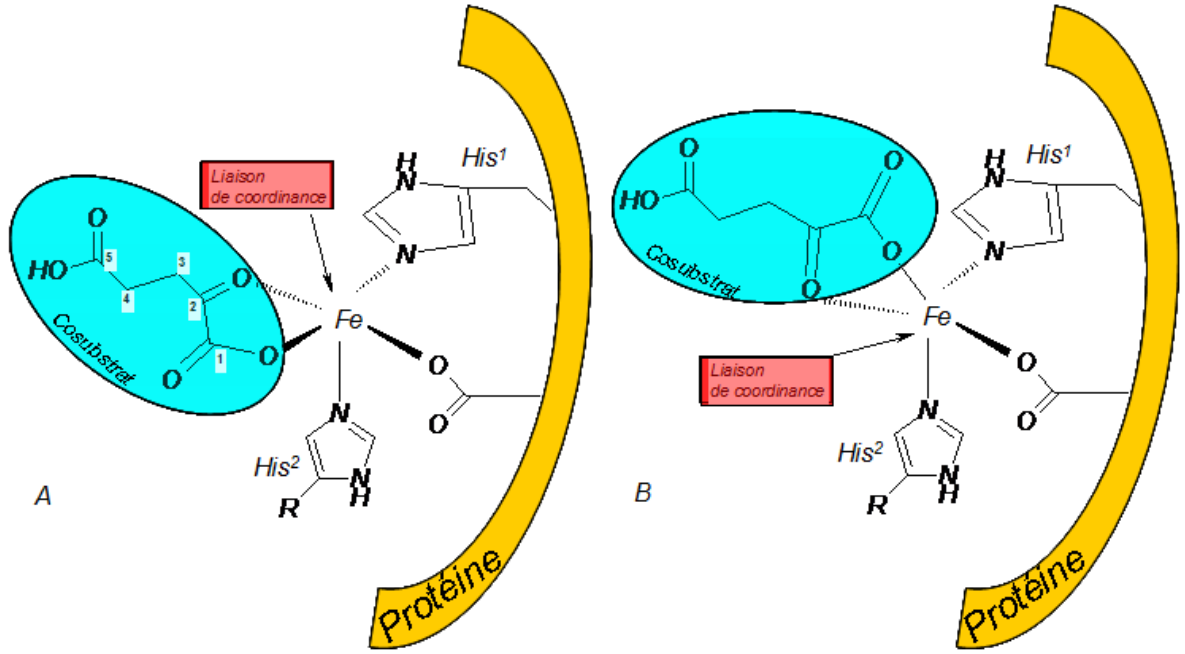


Figure 38 : Structure du site actif « in-line » A et « off-line » B.

Une autre similarité réunit les dioxygénases Fe(II)/α-cétoglutarate-dépendantes, c'est la présence des acides aminés Arg-X-Ser au niveau du site actif afin de lier le carbone C5 du α-cétoglutarate (Lukačin et al., 2000; Wilmouth et al., 2002) (Figure 36).

III.2.2.3 Mécanisme réactionnel

Différents mécanismes réactionnels sont proposés suivant la fonction attribuée à la dioxygénase Fe(II)/α-cétoglutarate-dépendante (réaction d'époxydation, de désaturation, de cyclisation, d'hydroxylation). Ces mécanismes réactionnels sont proposés après l'analyse de différentes expériences de cristallographies, de spectroscopies et de marqueurs isotopiques.

Pour les réactions d'hydroxylations, deux mécanismes sont proposés suivant qu'il s'agit d'une hydroxylation « in-line » ou « off-line ».

Dans le cas de l'hydroxylation « in-line », le site actif de la dioxygénase fixe le Fe au niveau des trois acides aminés His^1-X-Asp/Glu-X_n-His^2 retrouvés chez toutes les hydroxylases Fe(II)/α-cétoglutarates-dépendantes. L'analyse des structures cristallographiques révèle la

présence dans cette conformation, de trois liaisons de coordinence avec H_2O (Figure 39.A) (Valegard et al., 2004).

La chélation de l'α-cétoglutarate avec le fer aboutit à l'état transitoire B. L'enzyme est alors liée avec le fer et l'α-cétoglutarate. La perte d'une molécule d'H_2O crée un site de liaison de coordinence proche du substrat (C). Cette liaison permet de lier le fer avec l'oxygène moléculaire (D). Le fer se trouve en équilibre entre trois espèces Fe(II)-O2, Fe(III)-superoxo et Fe(IV)-peroxo (D). Une attaque nucléophile du groupement cétone de l'α-cétoglutarate aboutit à la formation du complexe bicyclique Fe(III) ou Fe(IV) peroxyhemicétal (E). Le cosubstrat α-cétoglutarate, lié avec le fer en état d'équilibre [Fe(III) ou Fe (IV)], est additionné d'un atome d'oxygène entre les carbones C1 et C2 (F). L'oxygène activé, situé proche du substrat, prélève un atome d'hydrogène. Cette réaction aboutit à la formation d'une espèce Fe(III)-hydroxylée et à la transformation du substrat sous forme radicalaire (G).

La dernière étape consiste à hydroxyler le substrat radicalaire avec le groupement hydroxy du Fe(III). Après liaison avec trois molécules d'eau, le Fer se retrouve à l'état oxydé initial sous la forme Fe(II), l' α-cétoglutarate est transformé en succinate et le substrat est hydroxylé (A). Un autre cycle catalytique peut alors se mettre en place.

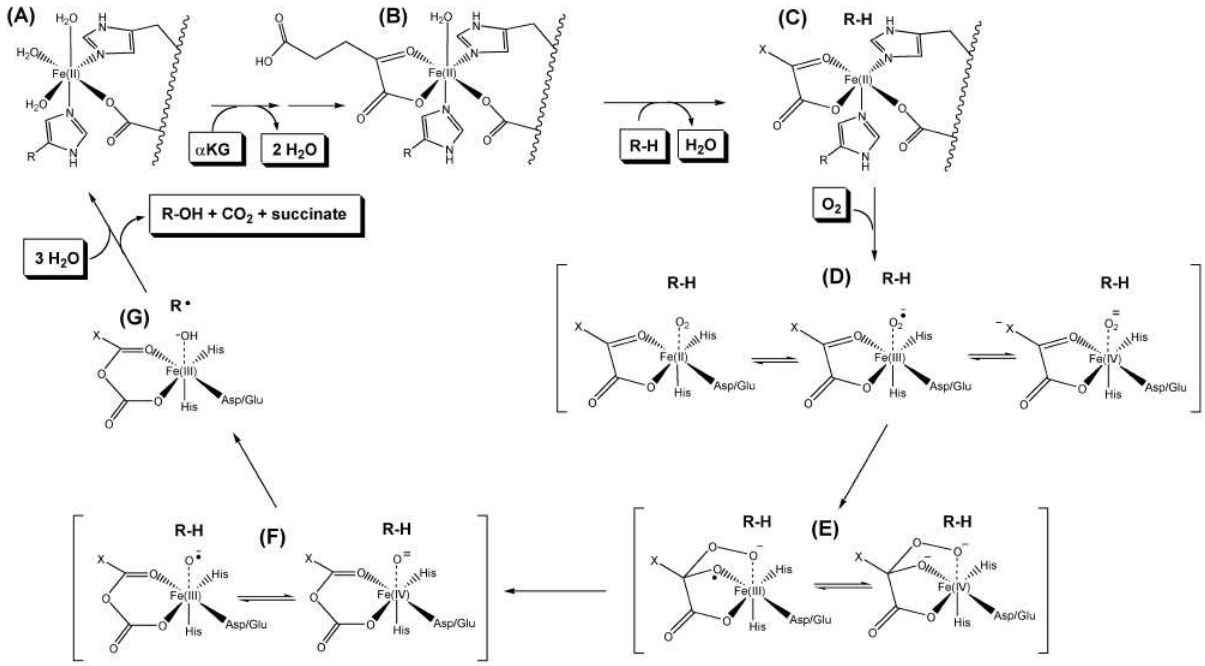


Figure 39 : Mécanisme d'action proposé pour les dioxygénases Fe(II)/α-cétoglutarate dépendante dans le cas d' hydroxylations « in line » (Hausinger, 2004).

Dans le cas de l'hydroxylation « off-line », deux hypothèses sont avancées pour expliquer de quelle manière l'atome d'oxygène activé se retrouve positionné à côté du substrat. Cette condition est indispensable afin de réaliser le transfert de groupement hydroxyl du fer sur le substrat.

La première hypothèse suppose que l' α-cétoglutarate est suffisamment flexible pour se réorienter en position « in-line » avant que le cycle réactionnel ne débute.

La deuxième hypothèse suppose que le début du mécanisme réactionnel est identique à celui proposé pour les hydroxylations « in-line », à la différence que la conformation de Fe(IV)-oxo est éloignée du substrat (Figure 40.A à D). Ensuite, la perte d'un groupement CO_2 entraîne l'apparition d'un site de liaison de coordinence qui peut se lier avec de l'eau (D à F). Enfin, la libération d'un groupement H_2O provenant du métal ferrique dihydroxylé permet un changement conformationnel du site actif. L'atome d'oxygène se positionne près du substrat et peut réagir avec la forme radicalaire comme décrit dans l'hydroxylation « in-line ».

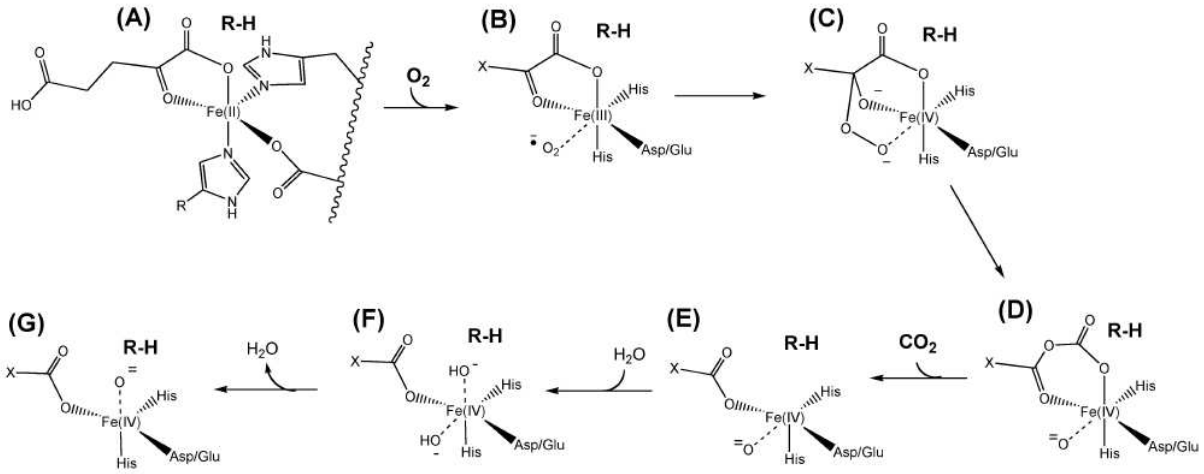


Figure 40 : Mécanisme d'action proposé pour les dioxygénases Fe(II)/α-cétoglutarate dépendantes dans le cas d' hydroxylations « off-line » (Hausinger, 2004).

III.2.3 Évolution moléculaire des dioxygénases.

III.2.3.1 Exemple de la voie de biosynthèse des flavonoïdes

Comme nous avons pu le voir pécedément, la voie de biosynthèse des flavonoïdes chez les Apiacées comporte de nombreuses dioxygénases et constitue donc un bon outil d'analyse pour émettre des hypothèses concernant leur évolution moléculaire. En effet, l'existence de nombreuses dioxygénases permet de réaliser une étude phylogénétique et donc d'étudier l'évolution des gènes (Gebhardt et al., 2005). Cette voie de biosynthèse comprend : la flavanone 3β hydroxyalse (FHT), une dioxygénase α-cétoglutarate dépendante qui hydroxyle de manière stéréospécifique le (2S)-flavanone en (2R,3R)-dihydroflavonol (DHF) (Figure 41) ; l'antocyanidine synthase (ANS), une autre dioxygénase α-cétoglutarate dépendante qui convertit le flavan-cis-3,4-diol (leucoanthocyanidin) en anthocyane ; la flavone synthase I (FNS I) et la flavonol synthase qui (FLS) métabolisent la (2S)-flavanone et le DHF respectivement en flavone et flavonol (Figure 41). L'expression de la FNS I est spécifique des plantes appartenant à la famille des Apiacées. Il est intéressant de noter que chez d'autres plantes cette même réaction est réalisée par la flavone synthase II (FNS II), une enzyme appartenant à la famille des cytochromes P450 (Springob et al., 2003).

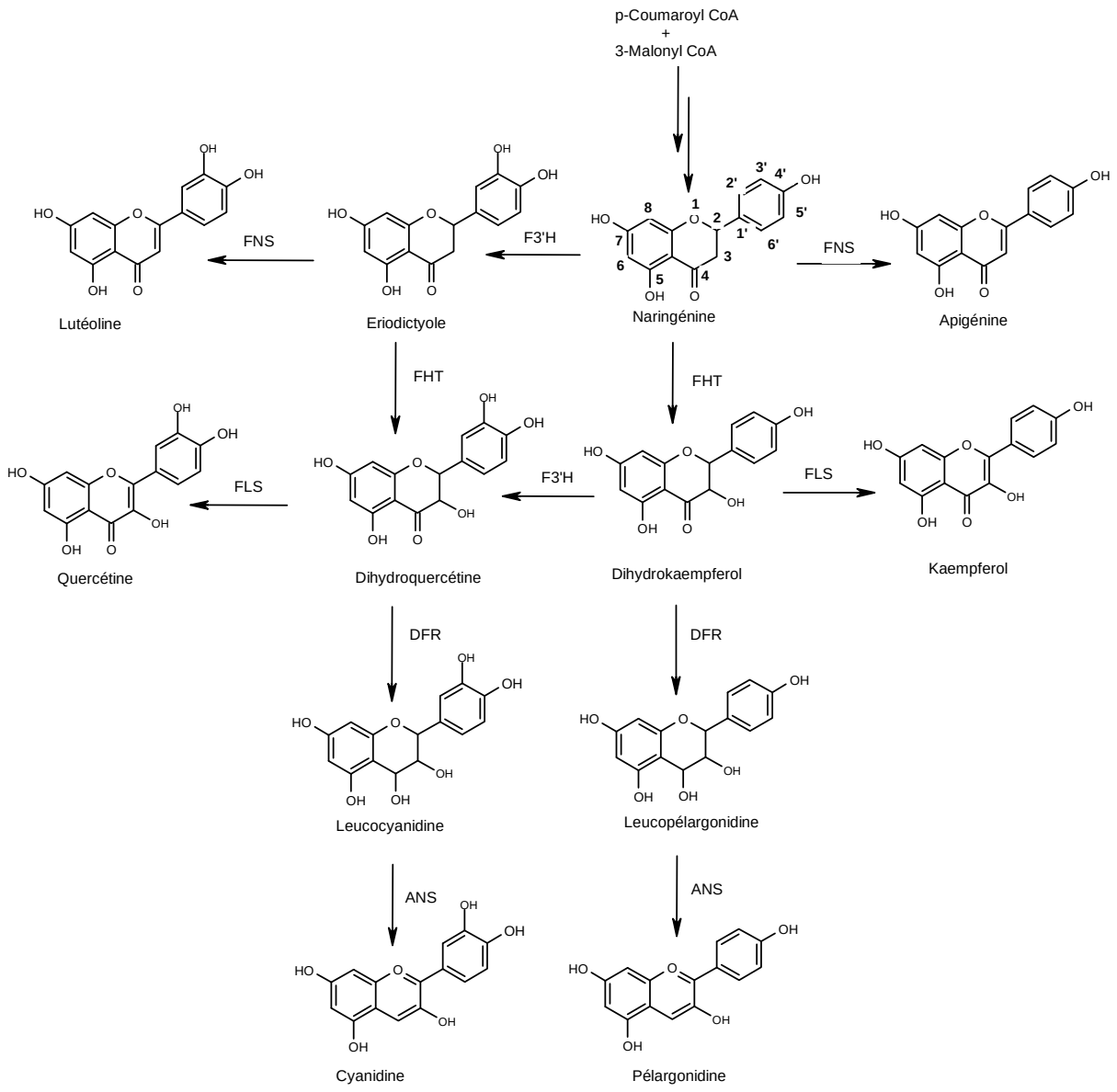


Figure 41 : Voie de biosynthèse des flavonoïdes.

Dioxygenases α-oxoglutarates-dépendantes : Flavone synthase I (Apiacées) (FHT), Flavonol synthase (FLS), Anthocyanidin synthase (ANS). Les autres enzymes : Flavonoid 3'-hydroxylase (F3'H), flavone synthase II (FNS II)

Gebhardt et ses collaborateurs ont isolé différents clones de FHT et de FNS I chez *Ammi*, *Anethum*, *Pimpinella* et *Conium* (Gebhardt et al., 2005). Toutes les séquences conservent les domaines caractéristiques des dioxygénases Fe/α-cétoglutarates-dépendantes (les sites de liaison au fer et à l'α-cétoglutarate). L'alignement des séquences peptidiques entre la FNS I et la FHT isolés chez ces différentes espèces révèle une très forte similitude (80-83 %) (Gebhardt et al., 2005). Par contre, une grande divergence peptidique existe entre ces deux

dioxygénases (FNS I et FHT) et l'ANS (23-31 %) ou avec les autres dioxygénases rencontrées dans les autres voies de biosynthèse (<32 % par exemple avec la 1-aminno-cyclopropan-4-carboxylate, ACCO). Dans la littérature, les séquences codantes pour les FNS I et les FHT extraites de feuilles de persils présentent aussi cette forte similitude entre elles (81 % d'identité) (Bredebach et al., 2011).

Un arbre phylogénétique a été construit à partir des FHT, FLS, FNS I et ANS présents chez les plantes produisant des flavonoïdes. D'après cet arbre phylogénétique, les FHT, FLS et ANS constituent chacun un embranchement spécifique. De plus, les FNS I et FHT issues de la famille des Apiacées forment un sous-groupe bien distinct à l'intérieur de l'embranchement des FHT provenant des autres familles de plantes (Figure 42).

Les processus d'évolution de ces différentes dioxygénases décrits à ce jour semblent correspondre à deux scénarios différents. La première hypothèse suggère qu'une seule enzyme a développé un caractère multifonctionnel et métabolise des produits différents à partir d'un ou plusieurs substrats. Ce scénario d'évolution est notamment attribué à la FLS et à l'ANS (Wilmouth et al., 2002; LukaÄ□ina et al., 2003; Bredebach et al., 2011).

La deuxième hypothèse suggère que de nouvelles fonctionnalités de l'enzyme apparaissent suite à un processus de duplication génique associé à des mutations évolutives ; c'est ce qu'on appelle la dérive génétique. L'augmentation du nombre de copies du même gène qui en résulte et la pression de sélection entraînent l'apparition de nouvelles fonctionnalités enzymatiques tout en conservant la fonction primaire du gène ancestral (Lynch and Force, 2000). C'est notamment le cas pour la FNS I et la FHT qui acceptent le même substrat avec une étroite spécificité (Bredebach et al., 2011) mais qui forment des produits différents. Le regroupement des FNS I en un sous groupe de gènes issus de l'embranchement des FHT (Figure 42) pourrait indiquer que ce sous-groupe de gènes possède un gène ancestral commun avec les FHS. De plus, au vu des fortes homologies peptidiques entre les FNS I et les FHT, il semblerait que le gène codant pour la FHT soit le gène ancestral ayant évolué pour donner la FNS I (Gebhardt et al., 2005).

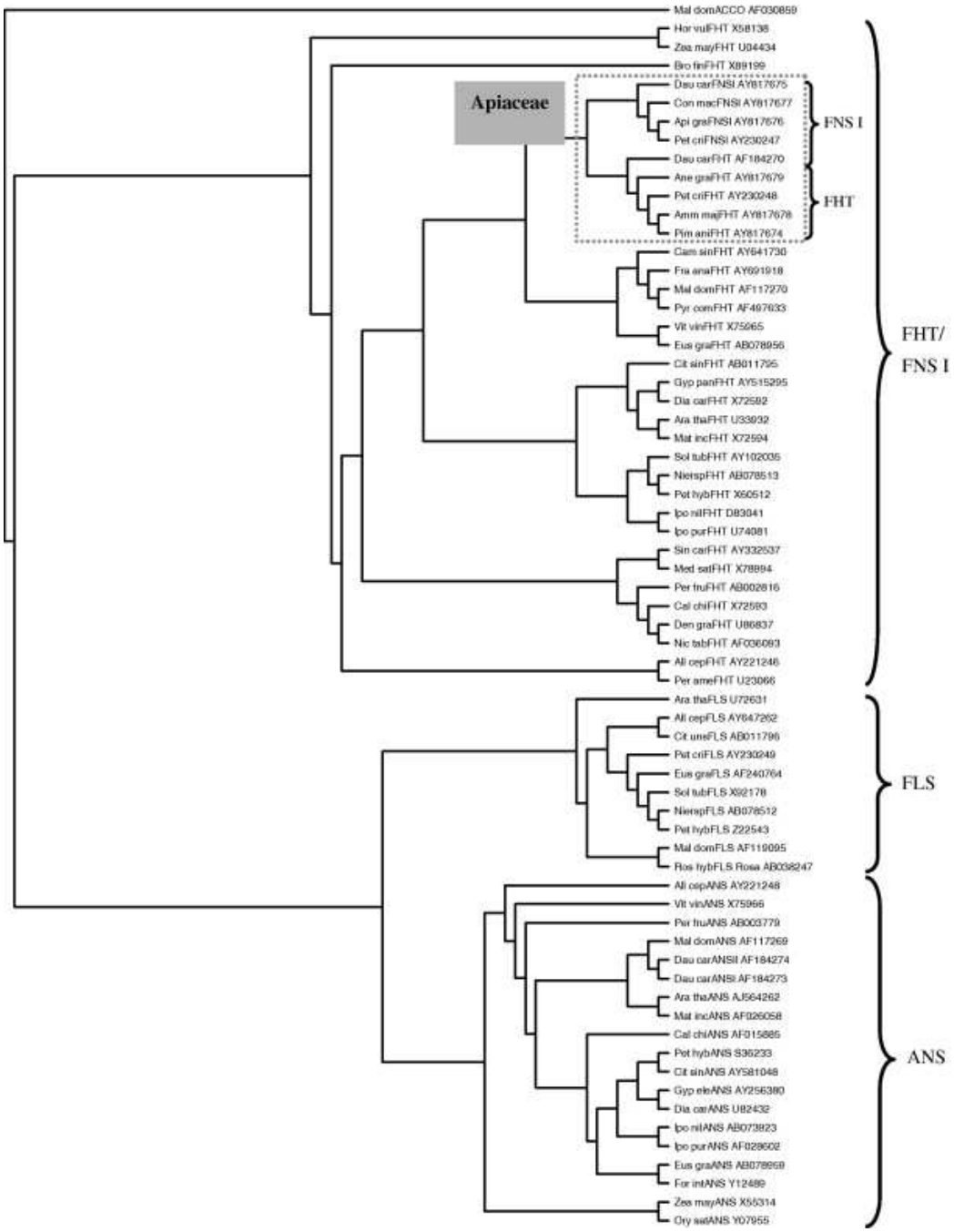


Figure 42 : Arbre phylogénétique des dioxygénases α-cétoglurates-dépendantes identifiées dans la voie de biosynthèse des flavonoïdes.

Les séquences isolées dans la famille des Apiacées sont encadrées par un rectangle en pointillé (Gebhardt et al., 2005).

La production de flavones par les Apiacées constitue probablement un avantage sélectif pour cette famille (Harborne, 1971). C'est pourquoi la nouvelle fonctionnalité (FNS I) apportée par la duplication génique de la FHT a été préservée durant l'évolution chez les Apiacées. Chez

les autres familles de plantes productrices de flavonoïdes, il est possible qu'elle ait perdu la capacité d'exprimer la FNS I à partir du moment où la FNS II (P450) fut en mesure de synthétiser des flavones (Middleton and Teramura, 1993; Logemann et al., 2000; Gebhardt et al., 2005). Gebhardt avance donc l'hypothèse que la haute protection des flavones vis-à-vis du rayonnment UV chez les Apiacées serait le moteur de l'évolution et de la conservation de la FNS I (Gebhardt et al., 2005). Cette hypothèse est renforcée par le fait que certaines espèces d'Apiacées ont supplanté leurs synthèses de flavonoles par de la flavone (Harborne, 1971).

Parrallèlement à cette étude phylogénétique, l'analyse des séquences génomiques met en évidence la présence d'introns chez les FNS I isolées de *Petroselinum* et *Daucus* (des Apiacées). Les deux séquences génomiques de la FNS I présentent deux introns différents mais situés exactement au même emplacement sur la séquence gènomique. Ils sont situés en avale des séquences de fixation du fer et de l'oxoglutarate.
Les FHT isolées chez ces mêmes espèces ne contiennent pas d'intron. Néanmoins, des FHT caractérisées chez *Dianthus caryophyllus* et *Medicago sativa* présentent des introns situés exactement au même endroit que les introns des FNS I (Charrier et al., 1995; Charrier et al., 1996). L'absence d'introns chez les FHT issues de la famille des Apiacées, peut s'expliquer par la perte de celui-ci comme il a déjà été décrit dans d'autres familles d'enzymes tels que les peroxydases chez *Arabidospis thaliana* (Tognolli et al., 2002) ou les catalases chez les angiospermes (Frugoli et al., 1998). Ce processus de perte d'intron est expliqué par des phénomènes de duplication génique par rétroposition (rétrotranscription des ARNm épissés qui sont intégrés secondairement au génome de manière aléatoire) (Roy and Gilbert, 2005).

De nombreux travaux proposent ce type d'évolution moléculaire enzymatique par phénomène de duplication génique. Ainsi, pour de nombreux autres gènes de la voie de biosynthèse de certains flavonoïdes tels que la chalcone synthase chez *Ipomea* ou *Gerbera*, la chalcone isomérase chez *Lotus* ou la dihydroflavonol-4-reductase chez *Ipomea*, il est proposé un phénomène d'évolution par duplication génique (Helariutta et al., 1996; Shimada et al., 2003). Il en est de même pour d'autres voies de biosynthèses, telles que la voie des glucosinolates chez *Arabidopsis thaliana*, qui regroupent des dioxygenases α-cétoglutarates-dépendantes ayant acquis de nouvelles fonctions par ce même processus (Kliebenstein et al., 2001).

III.2.3.2 Exemple des glycosinolates chez *Arabidopsis thaliana*

Les glycosinolates sont des métabolites secondaires recensés notamment chez *Arabidopsis thaliana* avec plus d'une trentaine de molécules décrites (Hogge et al., 1988). Ils sont synthétisés à partir de nombreux précurseurs tels que la méthionine, le tryptophane ou la phénylalanine. La première étape de cette voie de biosynthèse est catalysée par une enzyme appartenant à la famille des cytochromes P450 qui réalise une oxydation du groupe amine pour former une oxime (Wittstock and Halkier, 2000). Après formation d'un intermédiaire thiohydroximate, le transfert d'un résidu glucose et sulfate forme le squelette glucosinolate basique (Halkier and Du, 1997). Ensuite, par une succession de modifications de chaines latérales on obtient une grande variété de composés glucosinolates.

Deux enzymes intervenant dans cette voie et appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes ont été caractérisées chez *Arabidopsis thaliana* : l'une catalyse la formation d'alkenyl-glycosinolates (AOP2) et l'autre forme de l'hydroxyalkyl-glycosinolate (AOP3), chacune à partir de methylsulfinylalkyl glucosinate. Une autre enzyme appartenant à la même famille de dioxygénases (AOP1), est identifiée chez *Arabidopsis thaliana* et *Arabidopsis lyrata* mais dont le rôle catalytique est encore incertain (Kliebenstein et al., 2001).

Les gènes codant pour ces AOP ont été identifiés chez différents écotypes d'*Arabidopsis thaliana* présentant des phénotypes particuliers. L'écotype L*er* accumule principalement du 3-hydroxypropyl glucosinate, l'écotype C*vi* des allyl glucosinolates et l'écotype C*ol* produit uniquement le 4-methylsulfinylbutyl glycosinolate : précurseur des deux glucosinolates précedents.

L'analyse par BlastN révèle que les séquences codantes pour les AOP1, AOP2 et AOP3 sont des homologues presentants 70 % d'identité. Un arbre phylogénétique est construit à partir de AOP isolés chez les différents écotypes (Figure 43). Les travaux réalisés par Kliebenstein montrent une regroupement des différentes séquences codantes pour les AOP qui sembleraient se différencier suite à deux phénomènes de duplication génique.

La première duplication génique a séparé les AOP1 des progéniteurs AOP2/3, puis une seconde duplication génique a entrainé une différenciation des AOP2 et AOP3. Suite à ces duplications géniques, il semblerait que les copies de gènes aient évolué vers des fonctions enzymatiques différentes (Kliebenstein et al., 2001).

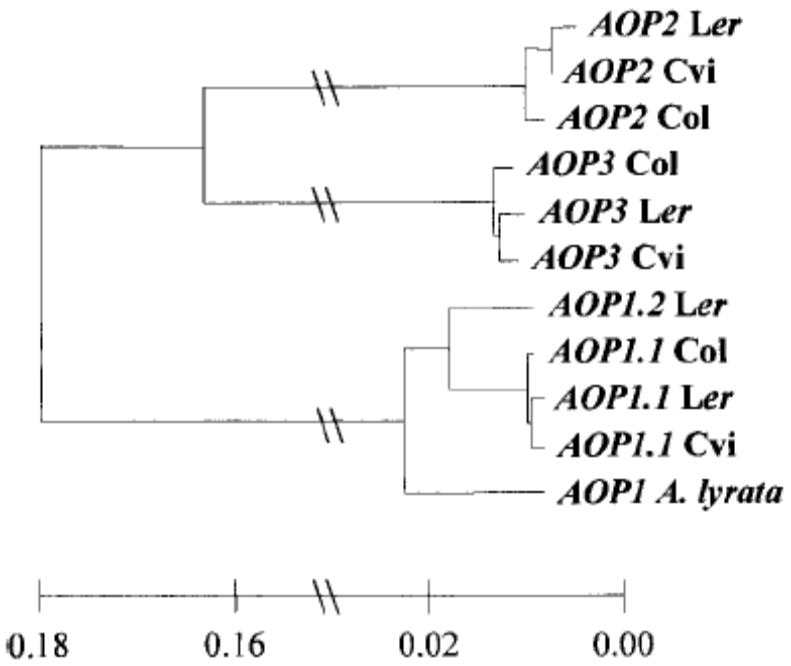


Figure 43 : Arbre phylogénétique des dioxygénases α-cétoglutarates-dépendantes (AOP1, AOP2 et AOP3) isolées chez différents écotypes *d'Arabidopsis thaliana* (Ler, Cvi , Col) et *A. lyrata* (Kliebenstein et al., 2001).

Pour compléter cette étude phylogénétique, les auteurs ont cherché à connaitre le profil d'expression des ces gènes dans les différents écotypes. Ils ont identifié les gènes codant pour les différents AOP dans le génome de tous les écotypes étudiés. Néanmoins, ils ont démontré par des analyses de RT-PCR quantitative que l'expression de AOP2 et AOP3 est respectivement corrélée à la présence d'alkenyl-glycosinolates et d'hydroxyalkyl-glycosinolates chez les différents écotypes (Cvi et Ler) d'*Arabidopsis thaliana.*
Toutefois, les AOP2 et AOP3 ne sont jamais coexprimés ensemble dans le même écotype d'*Arabidopsis thaliana* alors que AOP1 est retrouvé chez tous les écotypes. Le mécanisme majeur aboutissant à la variation phénotypique des différents écotypes d'*Arabidopsis thaliana* semble donc provenir d'une modification des promoteurs associés à chaque gène.

Dans cet exemple, les dioxygénases α-cétoglutarates-dépendantes semblent avoir évolué par duplication génique en tandem suite à un enjambement (crossing-over) inégal. Les mutations évolutives semblent avoir eu lieu à la fois sur les séquences codantes attribuant de nouvelles

fonctions enzymatiques mais aussi sur les parties régulatrices du gène contrôlant le niveau d'expression de chacun. D'autres études sur les promoteurs de chaque AOP aideront à mieux comprendre les mécanismes qui régulent l'expression de ces gènes dupliqués (Gebhardt et al., 2005).

Ces deux exemples d'évolution moléculaire, établis pour les dioxygénases α-cétoglutarates-dépendantes issues de la voie de biosynthèse des flavonoïdes ou des glucosinolates, caractérisent deux types de duplication génique : La duplication par transposition et la duplication par enjambement inégal.
La résultante de cette évolution est l'apparition de nouvelles fonctions enzymatiques qui contribuent à une plus grande spécificité réactionnelle.

IV Phénomène de duplication génique

Le métabolisme secondaire des plantes est caractérisé par une très large diversité de composés chimiques. La plupart de ces composés ainsi que leurs voies de biosynthèse sont restreints à des espèces de plantes bien spécifiques. L'évolution a donc permis la création et la sélection de nouveaux gènes afin de parvenir à cette diversité chimique.

Le phénomène de duplication génique est un mécanisme central dans l'évolution des gènes et du génome. Il semble être le mécanisme majeur mis en place dans l'acquisition de nouveaux gènes et contribue de ce fait énormément à l'évolution des espèces végétales (Ober, 2005; Roth et al., 2007). Contrairement à l'idée reçue, la duplication génique est un phénomène qui se produit fréquemment à l'échelle de l'évolution. En effet, l'analyse de nombreuses bases de données génomiques regroupant diverses espèces d'eucaryotes suggère que la fréquence de duplication d'un gène avoisinerait en moyenne 0,01 gène par million d'années (Lynch and Conery, 2000). Cette fréquence de mutation très importante à l'échelle de l'évolution est pondérée cependant par une demi-vie relativement courte d'un gène dupliqué, de l'ordre de 4 millions d'années. D'ailleurs, ces duplications géniques laissent des traces dans le génome des plantes, et de nombreuses régions ont été identifiées comme telles chez *A. thaliana.* (The *Arabidopsis* Genome Initiative 2000). Malgré tout, l'apparition de nouvelles fonctions après duplication génique et sélection naturelle reste un évènement excessivement rare (Lynch and Conery, 2003).

IV.1 Mécanismes aboutissant à la duplication génique

La duplication génique est un processus qui peut mettre en jeu un fragment ou un gène entier, un chromosome ou encore un génome complet dans le cas de polyploïdisation. Dans l'histoire de l'évolution du règne végétal, la duplication complète du génome est fréquente chez de nombreuses espèces. Il est d'ailleurs maintenant admis que la majorité des angiospermes sont des organismes polyploïdes (Soltis, 2005).

Dans le cas de la duplication d'un fragment ou d'un gène complet, les mécanismes les plus souvent impliqués sont les duplications par enjambement inégal (ou crossing-over inégal) et les rétropositions (Long et al., 2003; Zhang, 2003).

IV.1.1 Enjambement (ou crossing-over) inégal

L'enjambement inégal est un phénomène génétique qui a lieu lors de la méiose et qui contribue au brassage génétique lors de la reproduction. Lors de la prophase I de la méiose, les chromosomes homologues échangent des fragments de chromatides, entraînant un brassage intrachromosomique des allèles (Figure 44.A). On dit que les chromosomes sont recombinés. Lorsque l'échange de chromosomes ne se fait pas exactement au même niveau, il en résulte des chromosomes avec des longueurs de bras différents. Par conséquent, la région d'un des chromosomes est dupliquée et forme alors une répétition du gène en tandem (Figure 44.B). L'enjambement inégal est favorisé par les régions du chromosome qui possèdent déjà des séquences répétitives en tandem (Figure 44.C) (Achaz et al., 2001; Ober, 2010).

Les superfamilles comprenant plusieurs centaines de membres telles que la famille des récepteurs kinases ou des cytochromes P450, sont constituées de gènes pour la plupart issus d'une duplication de gènes en tandem (Cannon et al., 2004). À titre d'exemple, dans le génome d'*Arabidopsis thaliana*, 1528 régions contenant des répétitions en tandem sont identifiées sur 4140 gènes individuels. Près de 17 % des gènes chez *Arabidopsis thaliana* sont organisés en répétitions en tandem et proviendraient d'une duplication génétique par un enjambement inégal (The *Arabidopsis* Genome Initiative 2000).

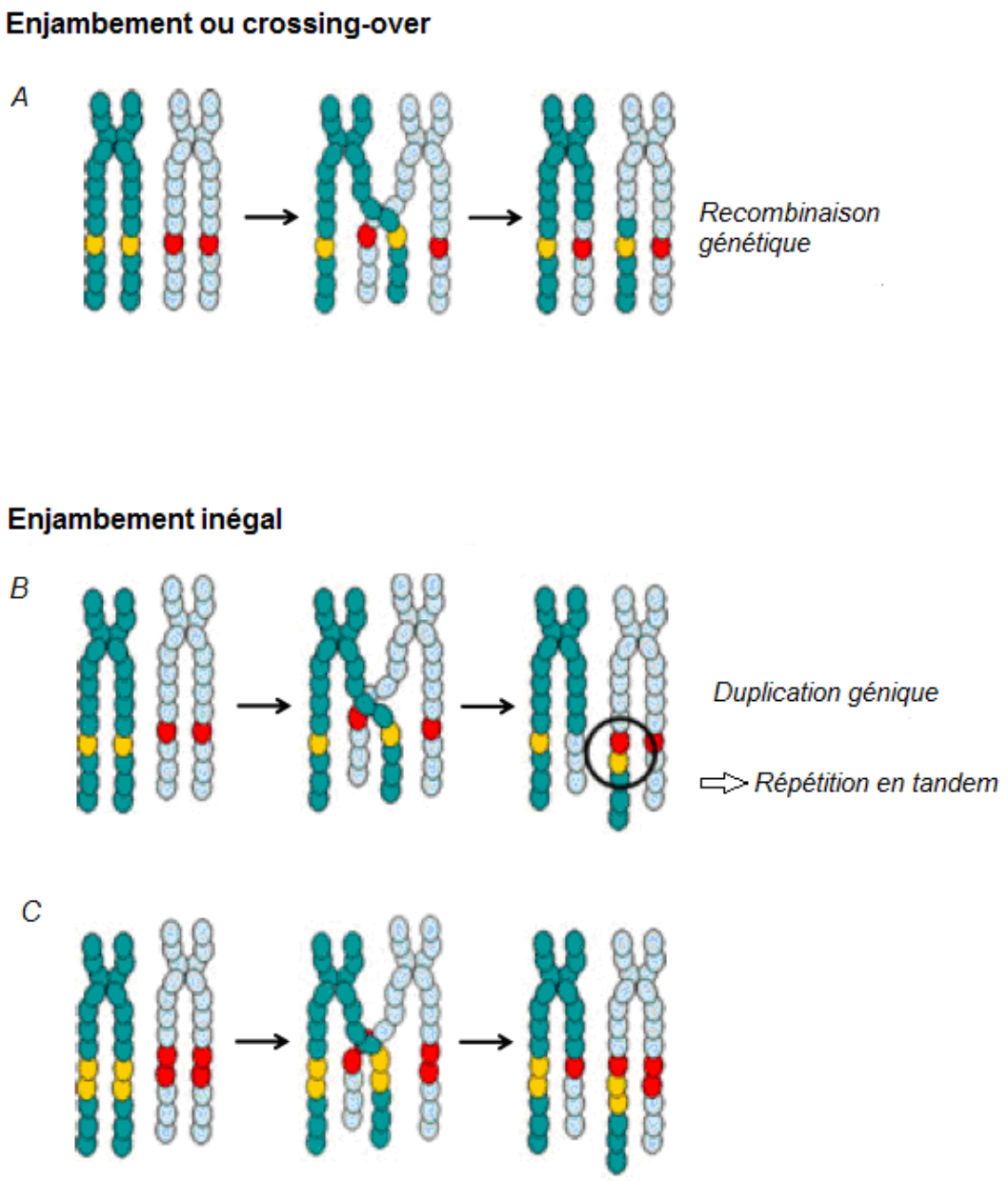


Figure 44 : Représentation simplifiée d'un enjambement et d'un enjambement inégal entre deux chromosomes homologues.

Représentation de deux chromosomes homologues lors de la prophase I de la méiose. Les couleurs marquent des segments chromosomiques. L'enjambement (ou crossing over) est caractérisé par un échange de fragment de chromatide entrainant un brassage intrachromosomique des allèles (A). L'enjambement inégal entraine des répétitions en tandem quand l'échange chromosomique ne se fait pas exactement au même endroit (B). Les régions en tandem présentes sur les chromosomes favorisent le phénomène d'enjambement inégal (C) (Ober, 2010).

IV.1.2 Rétroposition

Dans le cas de rétroposition, l'ARNm issu de la transcription et de l'épissage de la séquence génomique, subit une rétrotranscription afin de former de l'ADNc double brin. Cet ADNc peut alors être inséré de manière aléatoire dans le génome (Wang et al., 2006) (Figure 45). De ce fait, les rétrogènes, issus du phénomène de duplication génétique par rétroposition, sont caractérisés par l'absence d'intron et d'élément régulateur (promoteur). Une fois intégrés dans le génome, ces gènes deviennent la plupart du temps silencieux (c'est-à-dire qu'ils ne s'expriment pas), entraînant la formation et l'accumulation de pseudogènes qui seront éliminés au cours de l'évolution (Zhang et al., 2005b). Cependant, si l'insertion du retrogène se fait à proximité d'un gène actif, il peut utiliser les facteurs de régulation de ce dernier afin de rétablir sa transcription. Néanmoins, la proportion de régions codantes dans le génome est très faible, et donc diminue les chances pour un rétrogène de pouvoir être exprimé (Zhang et al., 2005b).

Il est possible d'identifier les gènes apparus par rétroposition car ils présentent des caractéristiques particulières au niveau de leurs séquences. En effet, les rétroposons sont caractérisés par la présence en 5' et 3' de régions non traduites et d'un segment riche en adénine. Ces régions spécifiques proviennent de l'époque où la séquence codait pour un ARNm (avant la rétrotranscription). De plus, 7 à 21 paires de bases dupliquées se retrouvent de part et d'autre du site d'insertion du retrogène qui caractérise un mécanisme ancien de reverse-transcription de l'ARNm en ADNc (Brosius, 1991).

Toutefois, l'intégration du rétrogène dans le génome se fait de manière aléatoire, ce qui explique le faible pourcentage de gènes pouvant être exprimés. Des études récentes caractérisent l'activité de rétroposons et démontrent néanmoins que ce processus fait aussi partie de la dynamique d'évolution des gènes (Brosius, 2003; Zhang et al., 2005b; Parker et al., 2009).

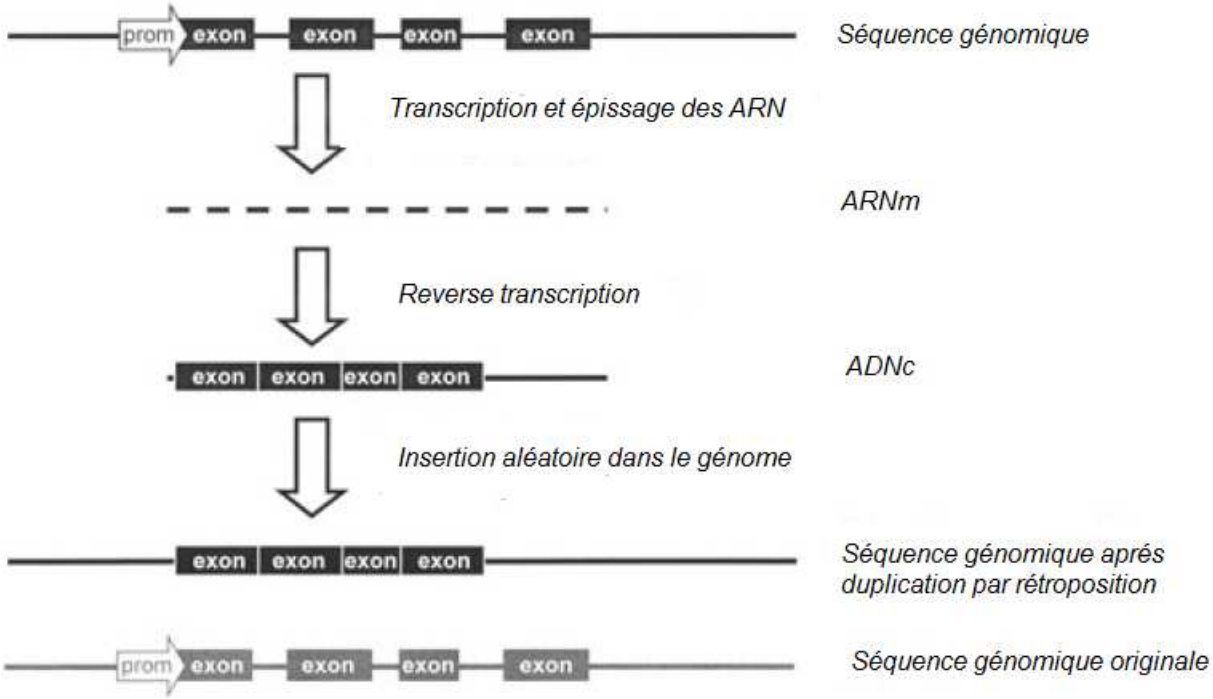


Figure 45 : Duplication génétique par rétroposition.
Ce processus requiert au préalable une transcription et un épissage de la séquence génomique. L'ARNm subit alors une rétro-transcription en ADNc double brin qui sera à nouveau inséré dans le génome. L'insertion du nouveau gène se fait de manière aléatoire dans la séquence génomique.

IV.2 Devenir des gènes dupliqués

Le phénomène de duplication génique entraine la formation immédiate de deux copies du même gène. Cependant, les copies du gène peuvent avoir des avenirs différents. Trois possibilités se présentent :

- Soit l'une des copies du gène devient un pseudogène (i)
- Soit les deux copies sont rigoureusement identiques (ii)
- Soit les gènes évoluent par mutation et peuvent alors développer une divergence de fonctionnalité (iii).

IV.2.1 Pseudogénisation

La pseudogénisation (i) est la possibilité la plus fréquente quant à l'avenir des gènes dupliqués (Zhang, 2003). En effet, l'insertion d'une mutation inhibant l'expression du néogène ou l'intégration du retrogène sans promoteur dans une partie du génome non traduite

entraine la formation de gènes inactifs. Dans ce cas, le pseudogène sera la plupart du temps éliminé au cours de l'évolution car il n'apporte pas d'avantage évolutif (Zhang, 2003).

IV.2.2 Copies redondantes du gène

Après duplication génique, il est possible d'obtenir dans le génome deux copies rigoureusement identiques (ii). Cette possibilité est rencontrée lorsque la présence d'une grande quantité d'un composé représente un avantage sélectif. En effet, la redondance d'une même fonction peut permettre l'acquisition d'une meilleure activité. Ce qui aboutit à la conservation de gènes paralogues. Ces gènes sont protégés contre les mutations altérant la fonction des gènes du fait de l'avantage bénéfique pour l'organisme (Zimmer et al., 1980; Zhang et al., 2005b). À titre d'exemple, les gènes codant pour la traduction de l'ARNr en ARNt ainsi que ceux codant pour les histones sont généralement présents dans les génomes sous différentes copies identiques. En effet, lors de la phase de réplication de l'ADN, ces protéines doivent pouvoir être mobilisables et présentes en grande quantité. Les organismes présentant plusieurs copies du même gène codant pour ces protéines en tirent donc un avantage. C'est habituellement cet « effet dose » qui est à l'origine de la conservation des nombreux gènes paralogues (Graur and Li, 1999).

IV.2.3 Acquisition de nouvelles fonctions

Le développement de nouvelles fonctions enzymatiques par duplications génétiques (iii), est défini selon deux hypothèses : la néo-fonctionalisation et la sub-fonctionalisation.

IV.2.3.1 L'hypothèse de la néo-fonctionnalisation.

Ohno (Ohno, 1970) est le premier à émettre l'hypothèse que la duplication génétique peut avoir pour conséquence la création de nouveaux gènes avec de nouvelles fonctions attribuées. L'acquisition de deux copies du même gène garantit la pérennité de la fonction associée au gène ancestral alors que l'autre copie du gène est libérée des contraintes fonctionnelles engendrées par la pression de sélection. Ainsi l'évolution permet à ce dernier d'accumuler des mutations : soit les mutations font perdre la fonction initiale du gène (non-fonctionnalisation), soit elles entrainent l'apparition de nouvelles fonctions (néo-fonctionnalisation) (Figure 46) et ont alors des chances de perdurer dans la population. Cette étape de néo-fonctionnalisation est

toujours précédée par une période où la pression de sélection n'intervient plus, et où la copie du gène se comporte exactement comme un pseudogène (Ohno, 1973). De nombreuses mutations entrainent de ce fait des altérations délétères du gène initial, avant que la pression de sélection ne permette au néogène de se stabiliser (Rodin and Riggs, 2003).

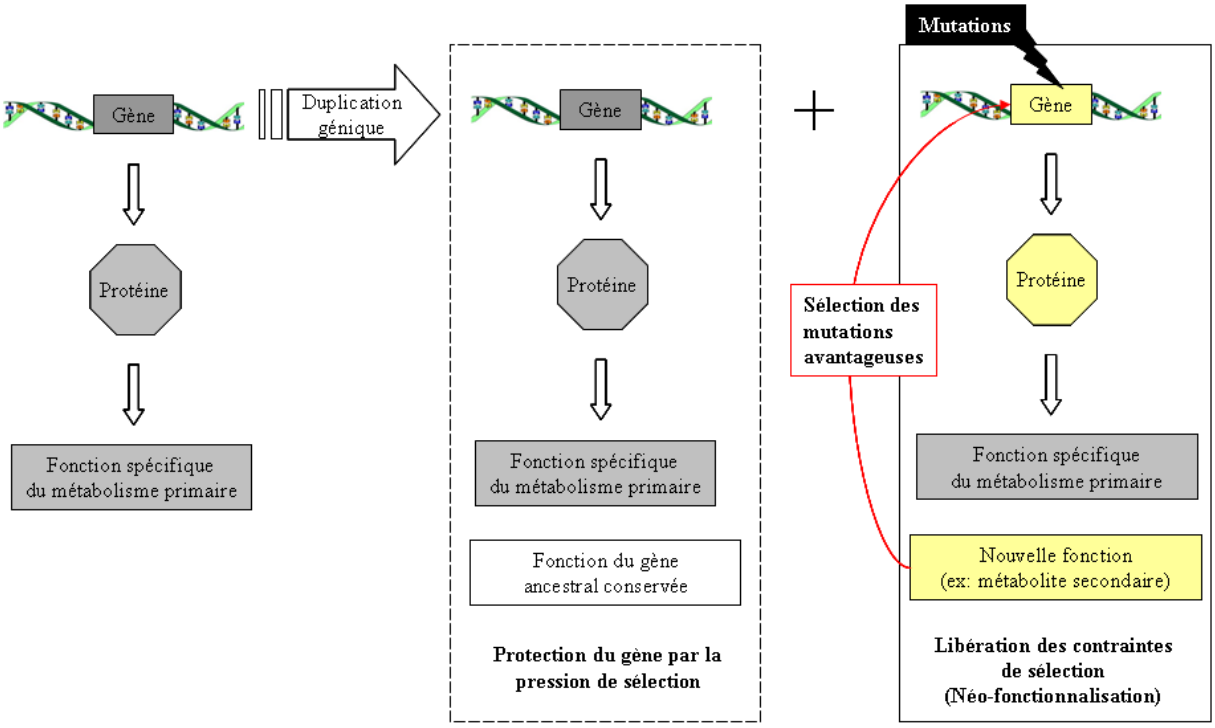


Figure 46 : Modèle d'évolution pour l'acquisition de nouvelles fonctions enzymatiques par l'hypothèse de néo-fonctionnalisation (Inspiré de (Ober, 2010)).

IV.2.3.2 L'hypothèse de la sub-fonctionnalisation

Dans ce modèle suggéré par Hughes puis par Lynch, les fonctions associées au gène ancestral sont partagées dans chacune des copies du gène. La résultante de l'ensemble des activités des copies du gène initial regroupe l'ensemble des activités assurées par l'expression de l'unique gène ancestral (Hughes, 1994; Lynch and Force, 2000; Hughes, 2002) (Figure 47). Dans ce cas, les copies de gène sont soumises à la pression de sélection de manière permanente. L'une des copies du gène ancestral peut perdre l'une des fonctions, et l'autre copie du gène perdre la fonction complémentaire. Ainsi l'organisme est capable d'exprimer les deux fonctions, alors assurées par deux protéines différentes. Cette sub-fonctionnalisation peut être sélectionnée, notamment si les deux fonctions présentent une certaine incompatibilité (expression dans des

tissus différents, par exemple) (Kai et al., 2008). Ce modèle est donc basé sur l'accumulation de mutations dégénératives qui se déroulent après la duplication (Roth et al., 2007). Force a introduit le modèle de « Duplication-Dégénération-Complémentarité » (Force et al., 1999). Une des conséquences de la sub-fonctionnalisation d´un gène dupliqué peut être la fission d´un gène en deux gènes différents. En effet, dans le cas où le gène originel code pour une protéine avec deux domaines fonctionnels distincts notés I et II, l´un des gènes dupliqués accumule des mutations (dégénération) dans la région correspondant au domaine II alors que le second gène accumule des mutations dans la région correspondant au domaine I. Il en résulte un gène codant pour une protéine avec le domaine I et un second gène codant une protéine avec le domaine II. La perte de fonction peut donc être avantageuse dans la mesure où elle participe au processus de spécialisation fonctionnelle. Cette spécialisation entraine généralement une meilleure activité de la protéine du fait de la diminution des contraintes pléiotropes exercées sur le gène (Hughes and Hughes, 1993; Rodin and Riggs, 2003). C'est notamment le cas du gène codant à l'origine pour la sous unité α du complexe tryptophane synthase chez le maïs (Frey et al., 2000; Ober, 2010).
La sub-fonctionalisation peut aussi faire apparaitre des modifications de la structure du gène et notamment par accumulation de mutations au niveau des éléments régulateurs (Hinman and Davidson, 2007). D'ailleurs une récente étude effectuée sur le génome de 17 champignons, révèle que les gènes dupliqués divergent plus souvent des éléments régulateurs plutôt que des fonctions biologiques (Wapinski et al., 2007).

Le modèle de sub-fonctionalisation ne s'applique donc que sur des gènes qui possèdent à l'origine une multifonctionnalité. Après duplication génique, les copies du gène sont séparées dans le génome et peuvent facilement se spécialiser pour acquérir des fonctions individuelles qui pourront être optimisées. Par ce mécanisme, les organismes peuvent s'adapter plus facilement à des conditions environnementales, ce qui favoriseraient l'acquisition de nouvelles fonctions enzymatiques.

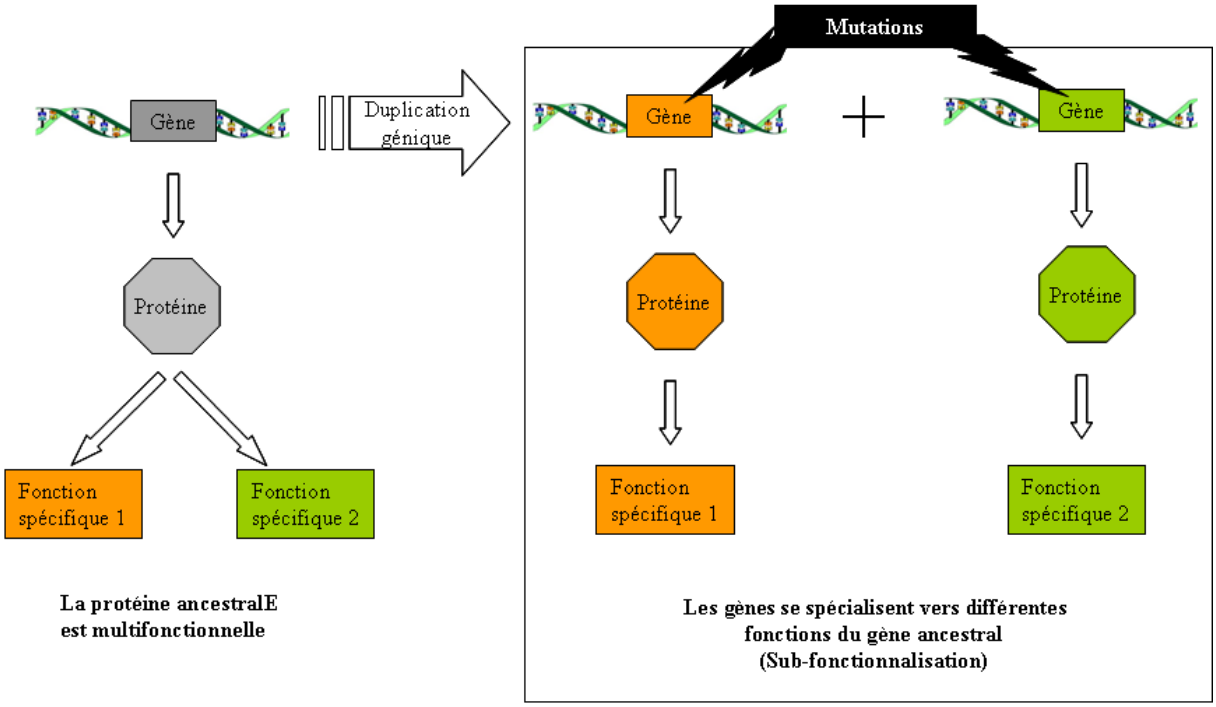


Figure 47 : Modèle d'évolution pour l'acquisition de nouvelles fonctions enzymatiques par l'hypothèse de sub-fonctionnalisation (Inspiré de (Ober, 2010)).

IV.3 Exemples dans le métabolisme secondaire des plantes

IV.3.1 Duplications géniques par enjambement inégal

L'identification et la caractérisation de nombreux gènes impliqués dans le métabolisme secondaire des plantes illustrent l'acquisition de nouvelles fonctions enzymatiques par le processus de duplication génique, suivi d'une spécialisation des copies du gène ancestral. Ces gènes appartiennent pour la plupart à la famille des terpènes synthases (Tholl, 2006), des chalcones synthases (Durbin et al., 2000), des glucosidases (Kliebenstein et al., 2005), des cytochromes P450 (Bak et al., 2006). Dans certains cas, les différents gènes dupliqués en tandem ont été identifiés : il s'agit de la terpène synthase, la dioxygénase 2-oxoglutarate-dépendante et la methylthioalkylmalate synthase chez *Arabidopsis thaliana* (Kliebenstein et al., 2001; Chen et al., 2004; Benderoth et al., 2006) ou des P450 dans la synthèse de benzoxazinoïdes chez le maïs (Frey et al., 2009).

L'exemple de la voie de biosynthèse des benzoxazinoïdes est particulièrement intéressant car il illustre bien l'évolution de la fonction enzymatique par le phénomène de duplication génique. Les benzoxazinoides (BX) comprenant le DIBOA (2,4-dihydroxy-2H-1,4-benzoxazin-3(4H)-one) et les formes glycosylées des dérivées methoxy, le DIMBOA, sont des pesticides naturels rencontrés chez le maïs (Sicker et al., 2000). Le précurseur de la voie de biosynthèse du DIMBOA-glc est l'indole-3-glycerol phosphate (IGP) (Figure 48). Cet IGP est métabolisé par l'enzyme BX1 en indole et en glyceraldehyde-3-phosphate. L'indole sera ensuite transformé successivement par quatre enzymes appartenant toutes à la même famille des cytochromes P450. (BX2-BX5) (Figure 48) (Frey et al., 1997).

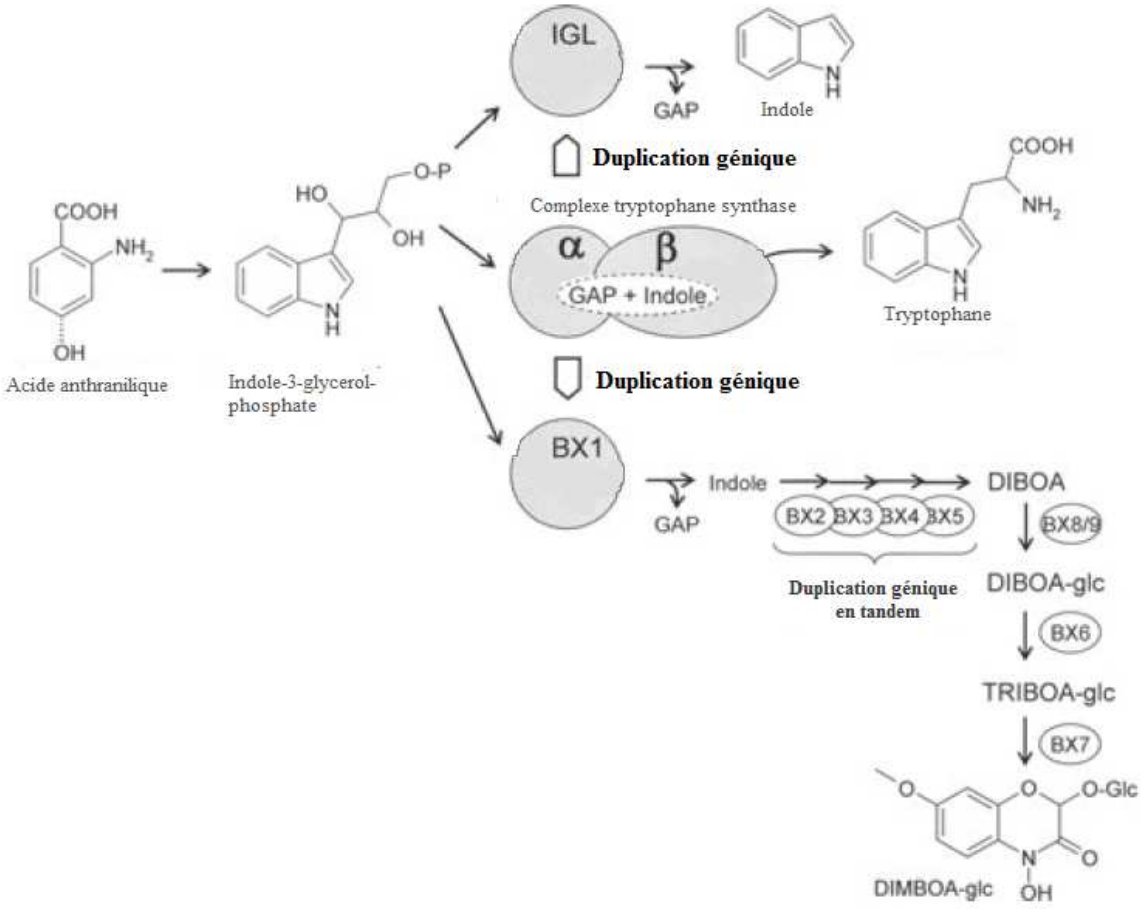


Figure 48 : Duplication génique du gène codant à l'origine pour la sous unité α du complexe tryptophane synthase chez le maïs.

Deux duplications indépendantes sont à l'origine des gènes codant pour l'IGL et la BX1. BX2-BX9 sont les enzymes impliquées dans la voie de biosynthèse des benzoxazinoides. DIBOA , 2,4-dihydroxy-2H-1,4-benzoxazin-3(4H)-one ; DIBOA-glc, DIBOA glucoside ; DIMBOA-glc, 2,4-dihydroxy-7-methoxy-2H-1,4benzoxazin-3(4H)-one glucoside ; GAP, glyceraldehyde-3-phosphate ; IGL, indole-3-glycerol-phosphate lyase ; TRIBOA, 2,4,7-trihydroxy-2H-1,4-benzoxazin-3-(4H)-one glucoside (Ober, 2010).

Plusieurs éléments font penser que ces quatre cytochromes P450 proviennent d'un unique gène ancestral qui aurait subi de nombreuses duplications géniques. Tout d'abord, ce sont des gènes homologues très proches qui présentent 45-65 % d'identité peptidique entre eux ; d'autre part les structures introns/exons sont conservées ; par ailleurs, ces gènes se clustérisent sur le bras court du chromosome 4 du maïs (Frey et al., 1997).

En plus d'être transformé par BX1-BX9 en DIMBOA, l'IGP est aussi le substrat de l'indole-3-glycerol phosphate lyase (IGL) qui catalyse la formation d'indole libre comme composé volatil impliqué dans les systèmes de défense indirecte du maïs (Frey et al., 2000).
Il est aussi le substrat du complexe tryptophane synthase qui produit du tryptophane impliqué cette fois dans le métabolisme primaire du maïs (Figure 48). Ce complexe est un hétérodimère constitué de deux sous unité α et β (Miles, 2001).
L'analyse phylogénétique et les comparaisons de séquences des gènes codant pour le BX1 et l'IGL indiquent vraisemblablement qu'ils dérivent tous les deux du gène codant pour la sous unité α de la tryptophane synthase. La spécialisation de chacune de ces trois enzymes serait survenue suite à des événements indépendants de duplication génique. D'ailleurs, l'expression hétérologue du monomère α de la tryptophane synthase de maïs a permis de montrer que cette sous unité est aussi capable de transformer l'IGP en indole, mais avec une activité bien moindre que lorsqu'elle se complexe en hétérodimère avec la sous unité β (Kriechbaumer et al., 2008). Cette observation suggère que le gène ancestral codait probablement pour une protéine (sous unité α) capable de catalyser le tryptophane et qui aurait été recrutée secondairement pour la biosynthèse de composés secondaires de défense (Frey et al., 1997).

Dans cette voie de biosynthèse, les homologies de séquences, les relations phylogénétiques et la localisation clusterisée des enzymes BX2/3/4/5 permettent aussi de suggérer une évolution moléculaire enzymatique par un processus de duplication en tandem (Figure 48) ayant abouti à l'acquisition de nouvelles fonctions enzymatiques des cytochromes P450 (Frey et al., 1997; Frey et al., 2009; Ober, 2010).

IV.3.2 Duplication génique par rétroposition

Les phenolamides sont des molécules antioxydantes et antimicrobiennes (Bienz et al., 2005) que l'on retrouve dans le pollen et les inflorescences de la plupart des plantes à fleurs. Elles

ont aussi un rôle dans la fertilisation et le développement des fruits (Kakkar and Rai, 1993; Meurer-Grimes, 1995).

De récentes études semblent montrer qu'une voie de biosynthèse des phenolamides, avec pour composé majeur du pollen l'hydroxyféruloylespermidine, est uniquement présente chez les Brassicacées (Fellenberg et al., 2008; Grienenberger et al., 2009). Cette voie de biosynthèse typique des Brassicacées fait notamment intervenir les gènes codant pour CYP98A8 et CYP98A9 (Matsuno et al., 2009). Ils sont situés sur le chromosome 1 d'*Arabidopsis thaliana* et les enzymes correspondantes hydroxylent le N^1,N^5,N^{10}-tricaffeoyl spermidine. CYP98A8 révèle en plus une activité d'hydroxylation avec N^1,N^5,N^{10}-triféruloyle spermidine. Ils dérivent tous les deux d'un même gène ancestral, CYP98A3 (Figure 49), qui métabolise la *meta*-hydroxylation des *p*-coumaroyle shikimate/quinate pour former des monomères de lignine (Schoch et al., 2001b).

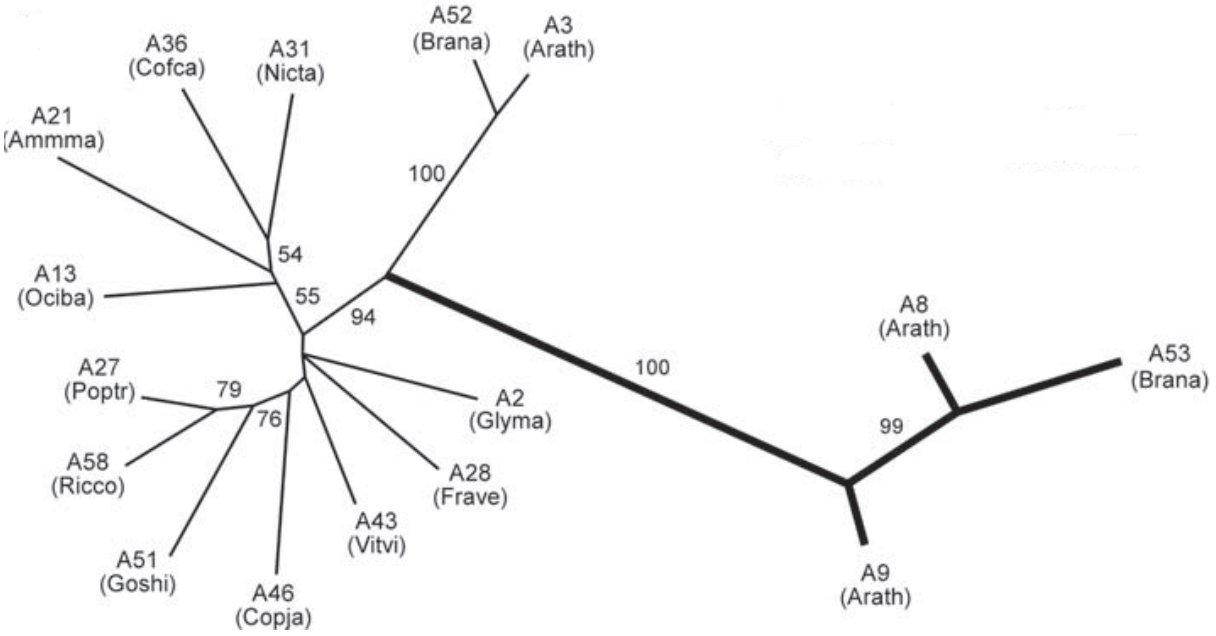


Figure 49 : Construction d'un arbre phylogénétique à partir des différents CYP98A incluant CYP98A8 et CYP98A9 (Matsuno et al., 2009).

L'analyse des séquences de CYP98A8 et CYP98A9 montre seulement 50 % d'identité peptidique avec CYP98A3. En revanche, les trois P450 se situent phylogénétiquement dans le

groupe des CYP98A incluant tous les P450 déjà caractérisés impliqués dans l'hydroxylation des *p*-coumaroyle shikinates/quinates.

Matsuno et ses collaborateurs suggèrent que CYP98A8 et CYP98A9 auraient émergé récemment et évolué de manière très rapide. En effet, aucun autre orthologue n'a pour l'instant pu être identifié dans les banques de données TIGR Plant Transcript Assemblie et PlantGDB, hormis un seul transcript CYP98A53 isolée chez *Brassica napus*. Les deux CYP98A8/A9 auraient perdu l'activité hydroxylase du *p*-comaroylshikimate/quinate de CYP98A3. De même, CYP98A3 ne métabolise pas la triféruloylespermidine mais possède une faible activité sur le tri(*p*-coumaroyle)spermidine. La dérive génétique dans cet exemple aurait abouti à l'acquisition de nouvelles fonctions (Matsuno et al., 2009).

En parallèle de cette étude phylogénétique, Matsuno et ses collaborateurs ont montré que CYP98A8 et CYP98A9 ne possèdent pas d'intron alors que tous les autres membres de la famille des CYP98A identifiés chez les végétaux supérieurs possèdent deux introns situés au même endroit. Cette absence d'intron est caractéristique du phénomène de rétroposition.

L'étude phylogénétique effectuée par Matsuno *et al* montre que l'apparition de cette voie de biosynthèse résulterait d'une néofonctionalisation à partir d'un gène ancestral CYP98A3 par phénomène de duplication génique par rétroposition (Matsuno et al., 2009).

Chez les primates, l'acquisition de nouvelles fonctions enzymatiques par rétroposition se vérifie la plupart du temps dans les organes impliqués dans la spermatogénèse (Kaessmann et al., 2009). En effet, les gènes impliqués dans la reproduction des animaux et des plantes sont réputés pour évoluer plus rapidement que les autres (Clark et al., 2006). Il semblerait que CYP98A8 et CYP98A9 constituent un exemple similaire quant à l'acquisition de nouvelles fonctions enzymatiques par rétroposition dans les organes sexuels chez les plantes.

IV.3.3 Les Opérons-like.

Un opéron est un groupement de gènes et de séquences régulatrices du génome. On le retrouve principalement chez les procaryotes et quelques nématodes. L'opéron comprend une région régulatrice ou opérateur, un promoteur commun, et plusieurs gènes qui sont contrôlés de manière coordonnée. L'ensemble de ces gènes est co-transcrit sous forme d'un ARN messager polycistronique. Ils ont été décrits la première fois par François Jacob et Jacques Monod en 1960 (Jacob et al., 2005).

Hormis quelques exceptions chez les nématodes (Blumenthal and Gleason, 2003), on ne retrouve pas d'opérons chez les eucaryotes. Cependant, on observe une organisation en cluster

de certains gènes non homologues, possédant des fonctions liées. On parle alors d'opérons-like. Contrairement aux opérons, ces gènes sont transcrits indépendamment mais présentent une organisation similaire aux opérons et sont généralement co-régulés (Osbourn and Field, 2009). Ils sont rencontrés dans le génome des animaux, des champignons et de certaines plantes (Wilmouth et al., 2002; Hurst et al., 2004; Hoffmeister and Keller, 2007). Le premier cluster de gène a été identifié dans la voie de biosynthèse des acides hydroxamiques chez le maïs (*Zea mays*) en 1997 (Frey et al., 1997). Depuis d'autres clusters de gène impliqués dans le métabolisme secondaire des plantes ont été identifiées dans la voie de biosynthèse d'un triterpène (avénicine) chez l'avoine (*Avena strigosa*) (Qi et al., 2004) ainsi que dans la voie de biosynthèse de différents diterpènes (momilactones et phytocassanes) chez le riz (*Oryza sativa*) (Shimura et al., 2007). Ces quatre clusters sont impliqués dans la synthèse de molécules de défense. Ils sont généralement exprimés en réponse à un stress de la plante (Wang et al., 2011). Récemment, Field et ses collaborateurs ont mis en évidence la présence de deux opérons-like chez *Arabidopsis thaliana*. L'un est impliqué dans la synthèse de thalianol, un dérivé triterpénique, et l'autre dans la synthèse de marneral, un autre dérivé triterpenique (Wilmouth et al., 2002). La découverte de ces deux operons-like chez la même espèce végétale, *Arabidopsis thaliana,* a permis de comparer leurs évolutions au sein de cette même espèce. D'après leurs résultats, il semblerait que le cluster de thalianol et de marneral proviennent d'un même gène ancestral OSC/CYP705. Leurs résultats montrent des ressemblances avec les clusters impliqués dans le métabolisme secondaire des champignons filamenteux. En effet, ils se situent dans une région dynamique du chromosome, la partie subtélomérique. Ils sont constitués de régions riches en éléments transposables et sont régulés par la chromatine (Palmer and Keller, 2010). Comme décrit pour le cluster de la voie de biosynthèse du trichothecène chez *Fusarium*, les auteurs pensent que les deux opérons like d'*Arabidposis* auraient évolué de manière indépendante, et qu'ils se seraient formés par des phénomènes de duplication, de néofunctionalisation et de relocalisations géniques (Wilmouth et al., 2002; Proctor et al., 2009).

V Objectif de la thèse

Identification et caractérisation de gènes impliqués dans la voie de biosynthèse des furocoumarines.

Les furocoumarines sont des molécules qui présentent un intérêt majeur dans le traitement thérapeutique de nombreuses dermopathologies. À partir des années 1960, la voie de synthèse de ces molécules a été élucidée en partie, d'un point de vue biochimique. Néanmoins, la compréhension d'un point de vue moléculaire de la synthèse des furocoumarines est relativement peu documentée. Cette ambivalence entre l'intérêt grandissant que présentent les furocoumarines en médecine et le manque général de connaissances a contribué à l'élaboration d'un nouvel axe de recherche dans notre laboratoire.

La thématique « furocoumarines » mise en place en 1998 au laboratoire Agronomie et Environnement de l'UMR 1121 Nancy université-INRA Colmar a très nettement évolué au fil des années. Les premiers travaux portaient sur une approche agronomique avec l'élaboration de techniques de production de furocoumarines par Milesi (Milesi et al., 2001) et Massot (Massot et al., 2000). Avec l'acquisition de techniques de clonage au sein du laboratoire, la thématique « furocoumarines » s'est peu à peu tournée vers l'étude de la composante moléculaire de cette voie de biosynthèse. Cette exploration fonctionnelle a commencé avec la caractérisation de la C4H de *Ruta graveolens* réalisé par Antoine Gravot au sein de notre laboratoire (Gravot et al., 2004). Puis dans le même temps, les travaux de Karine Lièvre ont permis de mettre en place un protocole de transformation génétique de *Ruta graveolens* à partir d'hypocotyles (Lièvre et al., 2005). La compréhension moléculaire de la voie de biosynthèse des furocoumarines s'est encore améliorée avec la caractérisation par Romain Larbat des premières enzymes directement impliquées dans la synthèse de psoralène (la psoralène synthase) et d'angélicine (l'angélicine synthase) (Larbat et al., 2007; Larbat et al., 2009). Néanmoins, de nombreuses autres enzymes de cette voie de biosynthèse restent à être identifiées dans un premier temps, et à caractériser dans un second temps. Les objectifs de ma thèse étaient donc de poursuivre la découverte moléculaire des différentes étapes conduisant à la synthèse des furocoumarines.

Le premier objectif de ma thèse était d'une part à identifier des séquences nucléotidiques susceptibles de constituer de bons candidats codant pour des gènes permettant la synthèse de

furocoumarines chez les plantes, et d'autre part à caractériser ces gènes candidats. Les différentes approches bibliographiques, bio-informatiques et expérimentales que j'ai mises en place pour identifier de nouveaux gènes candidats, sont présentées dans la première partie du chapitre III : « Résultat ». La seconde partie de ce travail, l'étape de caractérisation fonctionnelle des gènes candidats, est présentée dans une deuxième partie du chapitre III. J'expose les résultats obtenus lors de la caractérisation *in vitro* des gènes candidats dans un système d'expression hétérologue bactérien, mais aussi les résultats acquis après caractérisation *in planta*.

Le second objectif de ma thèse consistait à exprimer et caractériser des cytochromes P450 candidats susceptibles d'intervenir dans la voie de biosynthèse des furocoumarines. L'isolement de ces gènes a été réalisé au préalable par l'équipe du Professeur Matern de l'université de Marburg en Allemagne et par l'équipe du Professeur Simonsen de l'université de Copenhague. Dans cette troisième partie du chapitre « Résultat », je développe les différentes modifications et les résultats obtenus après mises au point des conditions d'expression des différents P450 candidats dans le système hétérologue levure.

Chapitre II : Matériel et méthodes

I Matériel

I.1 Matériel végétal

I.1.1 Citrus limetta

Le matériel végétal utilisé pour la recherche de dioxygénases provient de feuilles prélevées d'un citronnier, *Citrus limetta*, situé au Jardin botanique de Nancy.

Citrus limetta communément appelé limettier doux ou limette est un arbre fruitier appartenant à la famille des Rutacées. Cet arbuste peut atteindre huit mètres de haut. Le limettier se rencontre principalement en Amérique du Sud, en Inde et en Afrique. Le fruit du limettier se nomme la limette. Il est de couleur vert-jaune, ovoïde, très charnu et très juteux. Les feuilles sont alternes de couleur verte brillante avec des tiges ligneuses (Photo 3). Les rameaux sont épineux.

Photo 3 : *Citrus limetta.*

I.1.2 Ruta graveolens

Le matériel végétal utilisé pour cette étude provient de la germination de graines de *Ruta graveolens* provenant de la société Conrard Appel, Samen und Pfalanzen GmbH, BismarckstraBe 59 , 64 293 Darmstadt, Allemagne.

Ruta graveolens L. appelée communément Rue des jardins, Rue fétide, herbe de grâce ou Rue officinale est un arbrisseau appartenant à la famille des Rutacées (Photo 4). Cette plante, originaire du sud-est de l'Europe (Ukraine, Bulgarie, Albanie), s'est largement naturalisée dans toute l'Europe, et constitue maintenant une plante dite méditerranéenne semi-arbustive. Cet arbrisseau de 70 cm de haut environ est très ramifié et ligneux à la base. Les feuilles sont vert pâle, semi-persistantes, alternes et de consistance un peu charnue. La floraison commence de mai jusqu'à août, et présente des petites fleurs de couleur jaune-verdâtre regroupées en corymbe. Les fruits sont des capsules déhiscentes produisant des petites graines noirâtres se dispersant par gravité à proximité de la plante mère.

Ruta graveolens dégage une odeur forte et pénétrante, provenant des huiles contenues dans les poches schizolysogènes à la surface des feuilles.

Photo 4 : *Ruta graveolens.*

I.1.3 Pastinaca sativa

Les plantes utilisées pour la construction de la banque d'EST normalisée, sont issues de la germination de graines de *Pastinaca sativa L. subsp sativa* variété Demi-Long de Guernesey. *Pastinaca sativa* communément appelée panais, pastenade ou racine blanche est une plante bisannuelle de 50 cm à un mètre de haut appartenant à la famille des Apiacées (Photo 5). On la trouve spontanément dans toute l'Europe, dans le Caucase et la Sibérie. Cette plante bisannuelle présente des inflorescences en ombelles de couleur jaune-or. Les feuilles sont grandes, pennées, à segments lobés dentés. La racine principale est pivotante et ligneuse.

Photo 5 : ***Pastinaca sativa.***

I.1.4 Arabidopsis thaliana.

Les plantes utilisées pour la transformation génétique proviennent de la germination de graines d'*Arabidopsis thaliana columbia 0* qui nous ont été données par l'équipe du D.WERCK de l'Institut de Biologie Moléculaire des Plantes (IBMP) de Strasbourg.

Arabidopsis thaliana communément appellé arabette des dames ou arabette de thalius est une espèce de plante appartenant à la famille des Brassicacées. Cette plante annuelle, originaire d'Europe, d'Asie et du nord-ouest de l'Afrique, est de petite taille (25 cm de hauteur). Elle présente des feuilles vertes, légèrement violacées, formant une rosette à sa base (Photo 6). À la floraison, de petites fleurs blanches disposées en corymbe formeront des siliques contenant en moyenne 50 graines.

Arabidopsis thaliana est un organisme modèle pour la recherche génétique dans le monde végétal (Fink, 1998). C'est devenu un organisme de référence du fait de son cycle de vie rapide de 6 semaines (de graines à graines), de sa petite taille qui permet de cultiver un millier de pieds au mètre carré, de sa résistance, de sa capacité à s'autoféconder, et surtout, du fait que cette plante possède l'un des plus petits génomes connus dans le monde végétal (Bennett et al., 2003).

De plus, *Arabidopsis thaliana* peut facilement être transformé par l'utilisation d'*Agrobacterium tumefaciens* avec la méthode « floral dip », relativement rapide et facile à mettre en place (Zhang et al., 2006).

Photo 6 : ***Arabidopsis thaliana***

I.1.5 Nicotiana benthamiana

Les plantes utilisées pour l'agro-infiltration transitoire proviennent de la germination des graines de *Nicotiana benthamiana* fourni par Étienne Herrbach de l'INRA de Colmar.

Nicotiana benthamiana appartient à la famille des Solanacées. Cette plante herbacée vit parmi les rochers, sur les collines et falaises du nord de l'Australie. Elle possède une hauteur très variable (de 20 cm jusqu'à 1,5 m), et présente des fleurs blanches à la floraison (Photo 7). C'est une espèce modèle en biologie végétale, et notamment pour l'étude de l'expression transitoire par agroinfiltration.

Photo 7 : Nicotiana benthamiana.

I.2 Souches bactériennes

I.2.1 *E. coli* Top 10

Cette souche de bactérie est utilisée pour l'amplification des constructions plasmidiques. Elle permet d'amplifier et de répliquer de manière stable des plasmides à fort nombre de copies.

Génotype : [F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG]

I.2.2 *E. coli* XLI Blue

Cette souche est utilisée pour la propagation des phages λTriplEx2 lors du criblage de banque d'ADNc de Rue.

Génotype : [recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′proAB lacIqZΔM15 Tn10] (Tetr)]

I.2.3 *E. coli* GenHogs®[invitrogen]

Cette souche dérive de *E.coli* DH10™. Elle est transformée en souche électrocompétente dans notre laboratoire et spécialement utilisée pour l'amplification de plasmide.

Génotype : [F- mcrA D(mrr-hsdRMS-mrcBC) f80lacZDM15 DlacX74 deoR recA1 endA1 araD139 D(ara-leu)7697 galU galK l rpsL nupG].

I.2.4 E.coli BM25.8

Cette souche est utilisée lors du criblage de banque pour l'excision et la recircularisation du plasmide pTriplEx2 à partir du phagemide λTriplEx2 grâce à l'activité Cre-recombinase.

Génotype : SupE44, thi (lac-proAB) [F' traD36, proAB+, lacqZ M15]λimm434 (kanR) P1 (camR) hsdR (rK12-mK12-)

I.2.5 *E.coli* BL21(DE3)

Cette souche est utilisée pour l'expression de protéines recombinantes. Elle contient le lysogène DE3 comprenant le gène codant pour l'ARN polymérase T7 sous le contrôle du promoteur *lac*UV5. L'expression de l'ARN polymérase T7 est induite par l'IPTG.

Génotype : F– ompT hsdSB(rB–, mB–) gal dcm (DE3)

I.2.6 *E.coli* HB101 (pRK2013)

Cette souche contient le plasmide helper pRK2013 qui permet la conjugaison entre bactéries. Le transfert des transgènes dans *Agrobacterium tumefaciens* se fait par conjugaison des plasmides.

Génotype : supE44, hsdS20 (rB- mB-), recA13, ara-14, proA2, lacY1, galK2, rpsL20,xyl-5, mtl-1.

I.2.7 *E.coli* ccdB survival

Cette souche est utilisée pour l'amplification et le maintien de vecteurs contenant le gène ccdB. Ce gène présente des effets toxiques pour toutes les bactéries usuelles. *E. coli* ccdB présente une forte résistance aux effets délétères de ce gène de sélection.
Genotype : F- mcrA Δ(mrr-hsdRMS-mcrBC) 80lacZΔM15 ΔlacX74 recA1 araΔ139 D(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG tonA::Ptrc –ccdA

I.2.8 *Agrobacterium tumefaciens* LBA4404 C58C1RifR (pGV2260)

Cette souche désarmée est utilisée comme vecteur de transformation génétique pour *Ruta graveolens, Arabidopsis thaliana* et *Nicotiana benthamiana.* Un gène de résistance à la rifampicine est intégré dans le matériel génétique bactérien. Cette souche contient le plasmide pGV2260 désarmé qui exprime un gène de résistance à la carbénicilline et qui permet l'infection des plantes par la présence de gènes de virulences.

I.3 Souche de levure

I.3.1 S. cerevisiae WAT11

La souche *S. cerevisiae* WAT11 dérive de la souche *S. cerevisiae* W303. Le gène endogène de la réductase a été remplacé par recombinaison homologue par le gène de réductase d'*Arabidopsis thaliana*, AthR1, sous le contrôle de GAL10-CYC1. Ce promoteur est réprimé par le glucose, et induit par le galactose. Lorsque cette souche est transformée avec le plasmide d'expression pYeDP60, contenant un gène d'intérêt codant pour un P450, la surexpression concomitante de la P450 réductase ainsi que l'enzyme P450 permet d'étudier l'activité de celle-ci dans la fraction microsomale.

I.3.2 S. cerevisiae WAT21

La souche *S. cerevisiae* WAT21 est différente de *S. cerevisiae* WAT11 par le remplacement de la P450 réductase l'AthR1 par l'AthR2 d'*Arabidopsis thaliana*.

I.4 Vecteurs et plasmides recombinants

I.4.1 pCR8/GW/TOPO® (Invitrogen)

I.4.1.1 Description

Le pCR8/GW/TOPO® (Invitrogen) est un vecteur d'entrée de type Gateway utilisé pour cloner rapidement des produits de PCR. Une topoisomerase de type I (Figure 50) est fixée de façon covalente sur un site spécifique permettant la ligation de plus de 95 % des produits de PCR. Ce vecteur possède deux extrémités sortantes 3'T, qui permet d'offrir une extrémité complémentaire aux produits de PCR, dont les extrémités 3' ont été aux préalables additionnées d'une adénine.

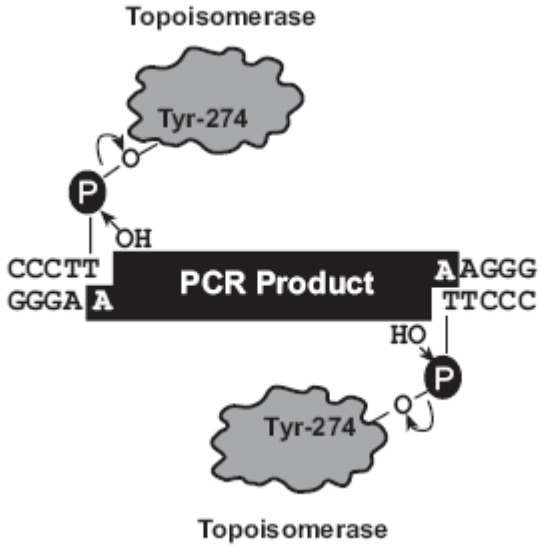


Figure 50 : Système de ligation dans le PCR8 (Invitrogen)

De plus, le vecteur pCR8/GW/TOPO® est doté : de sites spécifiques att (Figure 51) permettant une LR recombinaison rapide vers un vecteur de destination Gateway®, de deux sites *Eco*RI flanquant le site d'insertion de l'amplicon pour une excision de l'insert, et d'un

gène conférant la résistance à la spectinomycine. Les séquences complémentaires aux amorces universelles M13 et T7 permettent de vérifier par PCR le sens d'insertion du gène d'intérêt dans le plasmide.

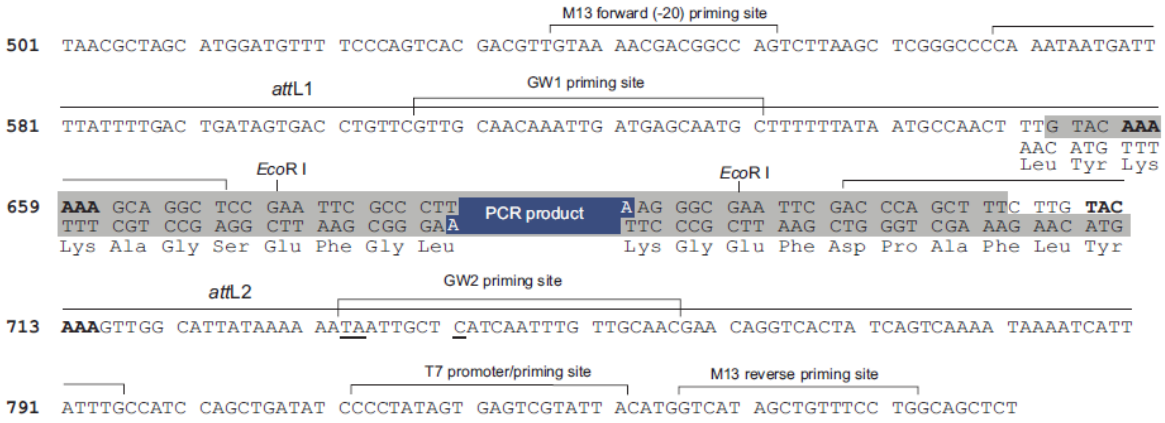


Figure 51 : Polylinker du plasmide pCR8/GW/TOPO®

I.4.1.2 Vérification du sens d'insertion du gène dans le vecteur

Le gène d'intérêt (X) peut s'insérer dans le vecteur pCR8/GW/TOPO dans le sens direct ou reverse. On suppose que ce gène possède un couple d'amorces Xdir et Xrev. Afin de vérifier le sens d'insertion du gène et de pouvoir sélectionner les plasmides ayant un insert dans le sens direct, on réalise une PCR à l'aide de l'un des couples d'amorces du gène d'intérêt (Xdir ou Xrev) et de l'amorce M13f présents sur le plasmide PCR8/GW/TOPO. La séquence M13f se situe à 145 pdb en amont de l'insert (Figure 51). Il est alors possible de déterminer le sens d'insertion du gène en fonction du couple d'amorces avec lequel on obtient un amplicon de taille supérieur à la séquence du gène (X + 145 pb). Dans le cas d'une insertion dans le sens direct, on s'attend à obtenir un amplicon de X+145 pb avec le couple d'amorces M13f/Xrev. Dans le cas d'une insertion dans le mauvais sens, on obtiendra un amplicon de (X +145 pb) avec le couple d'amorces M13f/Xdir.

I.4.2 pET-28a

Le plasmide d'expression pET-28a est équipé d'un promoteur T7 contrôlant l'expression du gène d'intérêt et nécessitant la présence de l'ARN polymérase T7 (dérive du bactériophage T7). Cette polymérase est transcrite *via* le génome de bactéries hôtes spécifiques (*E.coli*

BL21(DE3) sous le contrôle d'un promoteur *lac* inductible par l'IPTG (Figure 52). L'addition d'IPTG dans le milieu de culture, entraine l'expression d'ARN polymérase T7 *via* le génome de la bactérie, qui consécutivement, aboutit à l'expression du gène cible. Le plasmide pET est flanqué, aux extrémités N et C terminales du gène d'intérêt, de deux séquences His-Tag permettant la purification de la protéine synthétisée sur colonne chromatographique d'affinité. Un gène conférant la résistance à la kanamycine permet de sélectionner les bactéries transformées.

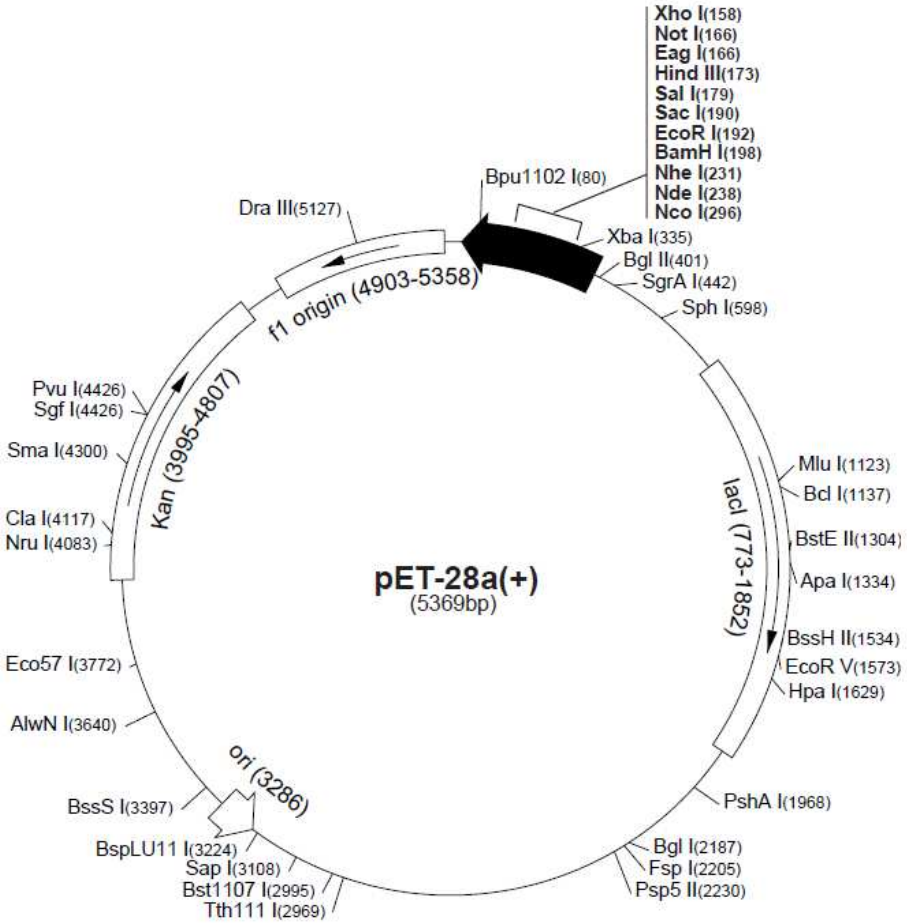


Figure 52 : Carte physique du plasmide pET-28a (Novagen).

I.4.3 pBin-GW

I.4.3.1 Présentation

La construction du pBin-GW a été réalisée par Sébastien Doerper durant sa thèse au laboratoire. Il dérive du plasmide pBin m-gfp5-ER fourni par le professeur Haselhoff. C'est un plasmide de destination Gateway® à faible nombre de copies. Entre les bordures droite et gauche délimitant l'ADN-T de ce plasmide se situent le gène de sélection npt II et la cassette de recombinaison RfB flanquée par les sites attR1 et attR2 (Figure 53). Le gène nptII est sous le contrôle du promoteur et du terminateur de la nopaline synthase (Pnos, et 3'nos), et confère la résistance à la kanamycine. La cassette de recombinaison RfB comporte le gène de résistance au chloramphénicol et le gène ccdB. Elle est encadrée en amont par le promoteur 35S et en aval par le terminateur 3'nos. L'amplification de ce plasmide est réalisée par transformation d'*E. coli* ccdB Survival. Cette souche a la capacité de survivre aux effets toxiques engendrés par l'expression du gène ccdb, et permet ainsi la propagation du plasmide. Un autre gène de résistance à la kanamycine se trouve en dehors de l'ADN-T afin de sélectionner les bactéries d'*E. coli* recombinantes.

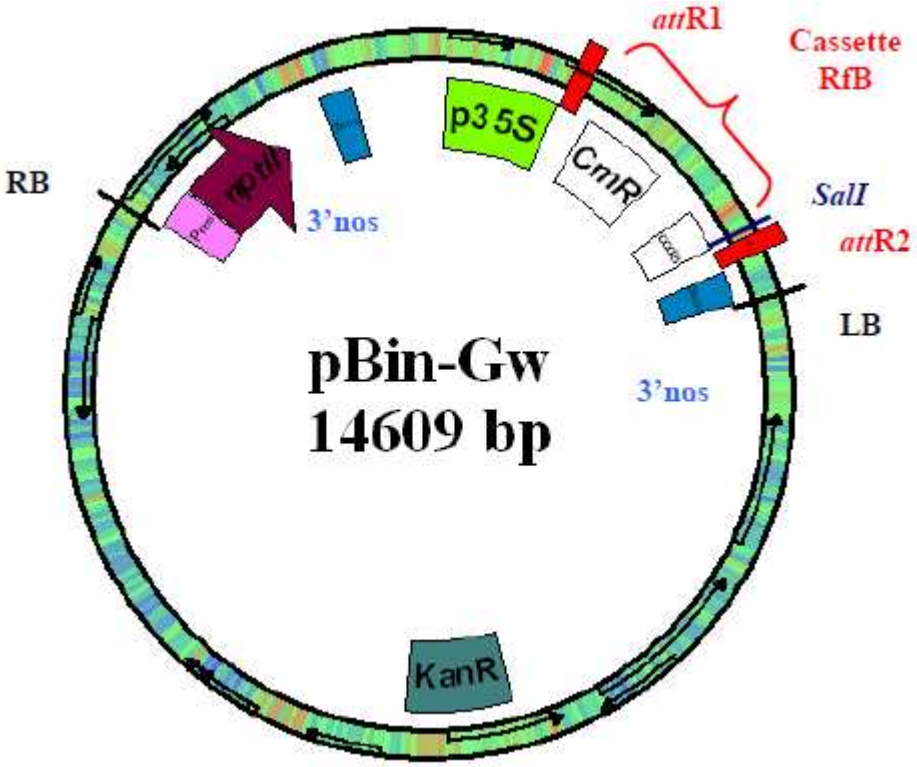


Figure 53 : Carte physique du plasmide pBin-GW (Biteau, 2009).

I.4.3.2 Recombinaison dans un vecteur de destination pBin (Technologie Gateway)

La technologie Gateway se réalise en deux étapes. La première consiste à introduire le gène d'intérêt dans un vecteur d'entrée pCR8/GW/TOPO. La deuxième étape consiste ensuite à transférer cette séquence d'intérêt dans une multitude de vecteurs de destination, tel que pBin-GW par LR recombinaison (Figure 54). Le transfert d'une séquence (flanquée par les sites att) d'ADN hétérologue, entre le vecteur d'entrée et le vecteur de destination, se fait par le système de recombinaison λ.

La séquence située entre les sites attL1 et attL2 du plasmide pCR8/GW/X (Figure 51) est recombinée avec la séquence se trouvant entre les sites attR1 et attR2 du plasmide de destination Gateway (pBin/GW). Cette recombinaison homologue se réalise en une étape.

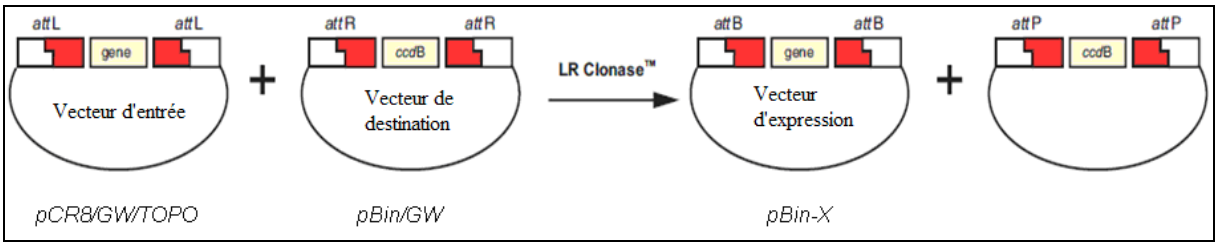


Figure 54 : Schéma représentant la LR recombinaison.

I.4.4 pYeDP6O

Ce plasmide provient du Laboratoire d'Ingénierie des Protéines Membranaires (LIPM) du CNRS à Gif-sur-Yvette. C'est un vecteur d'expression réplicatif équipé d'une origine de réplication de levure (Figure 55), d'une cassette de clonage permettant l'insertion d'un gène d'intérêt. Celui-ci est encadré en amont par un promoteur GAL10-CYC1, inductible par le galactose et réprimé par le glucose, et en aval par un terminateur PGK. Une origine de réplication *E. coli* et un gène de résistance à l'ampicilline permettent la réalisation de sous clonage. La complémentation des levures pour l'auxotrophie à l'uracile et à l'adénine est assurée par la présence dans ce plasmide de deux gènes URA3 et ADE2. Seules les levures

transformées peuvent ainsi se développer dans un milieu dépourvu en uracile ou adénine. Ce plasmide a été construit spécialement pour l'expression de P450 dans une série de souches de *Saccharomyces cerevisiae* (WAT 11, WAT 21).

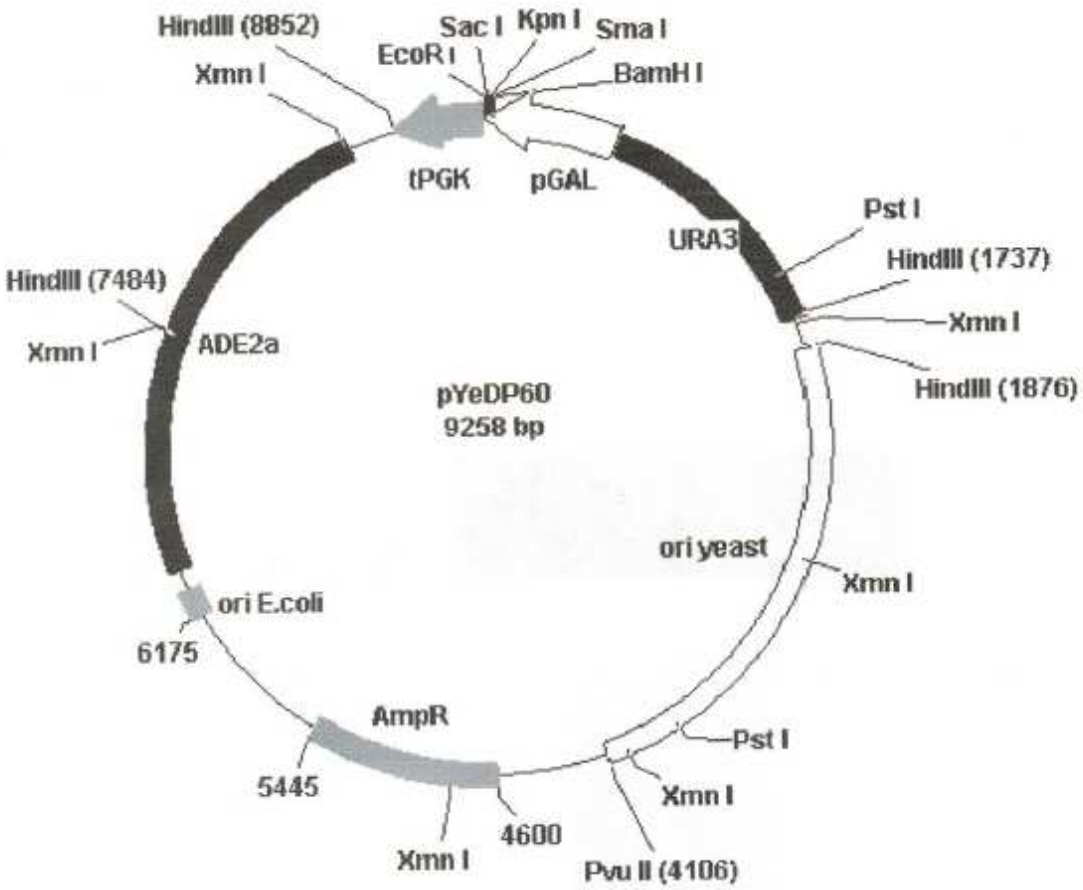


Figure 55 : Carte physique du plasmide pYedP60.

I.4.5 pTriplEx2 (Clontech)

Le pTriplEx2 (Figure 56) est un plasmide issu de l'excision et de la recircularisation du phagemide λTriplEx2. Il est constitué du promoteur et du régulateur *lac* d'*E.coli*, qui permet de réguler un gène d'intérêt dans les souches possédant le répresseur *lac*Iq.

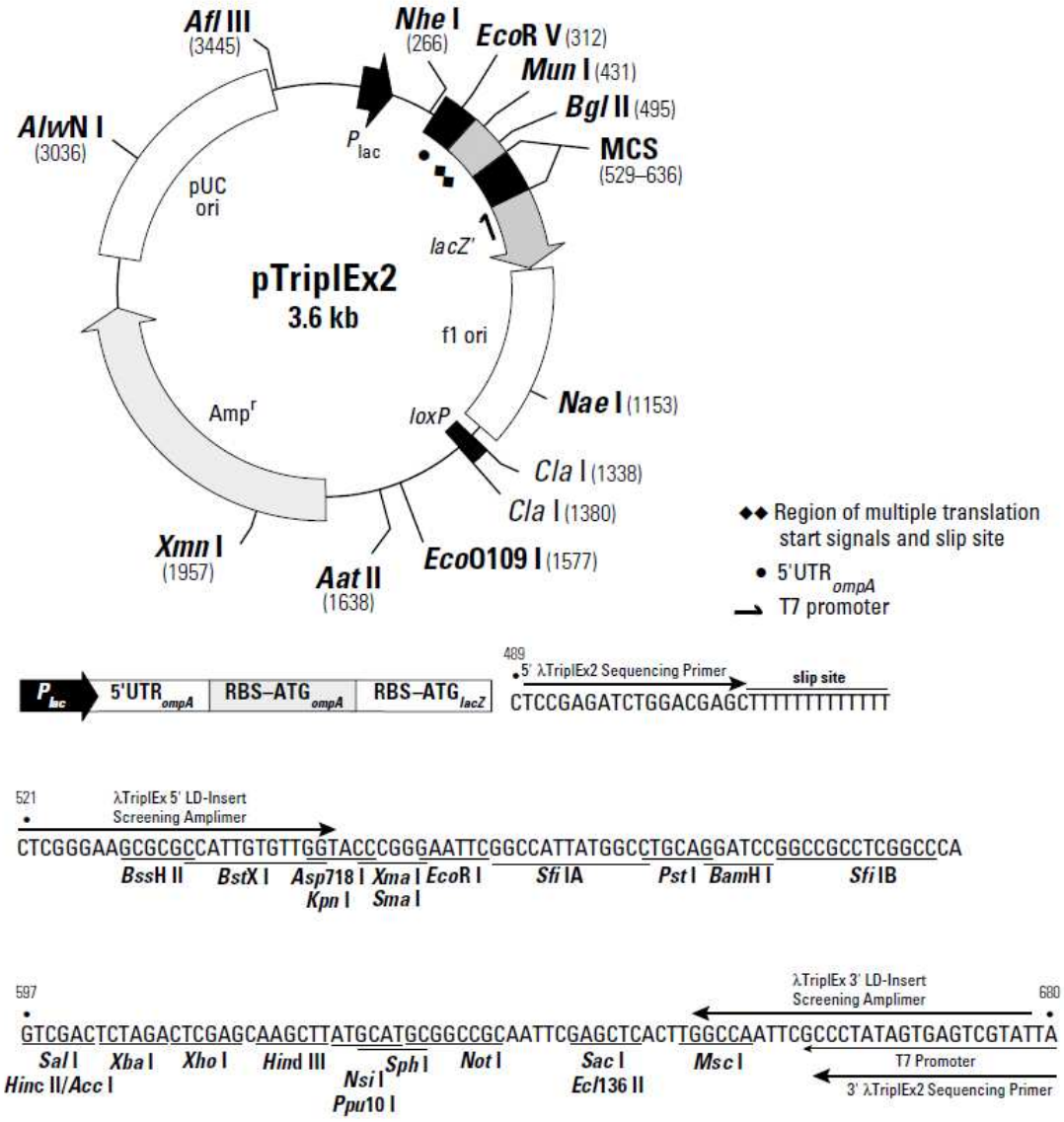


Figure 56 : Carte physique du pTriplEx2 (Clontech).

I.4.6 λTriplEx2 (Clontech)

Ce phagemide, utilisé lors de la construction de banque d'ADNc avec le « smart cDNA library construction kit » de Clontech, permet la création de banque à forte concentration. Il permet la sélection des recombinants par coloration blanche/bleu, et régule l'expression du gène d'intérêt. De plus, la transformation de ce phagemide en plasmide est facilitée par le système d'excision/recircularisation Cre-lox.

I.4.7 pQE30 (Qiagen)

Le plasmide pQE30 est un vecteur de surexpression possédant une origine de réplication de type ColE1 et le gène *bla* conférant la résistance à l'ampicilline. La transcription de l'ORF insérée dans le site multiple de clonage est placée sous le contrôle transcriptionnel du promoteur fort T5 et d'un double opérateur lacO. En absence d'inducteur (isopropyl-ß-Dthiogalactopyranoside (IPTG)), des tétramères du répresseur LacI se fixent sur les régions opératrices lacO et répriment l'expression du gène à surexprimer. Grâce à une induction par l'IPTG, ce plasmide permet la surproduction de protéines recombinantes portant une étiquette poly-histidines fusionnées à leur extrémité N-terminale (Figure 57).

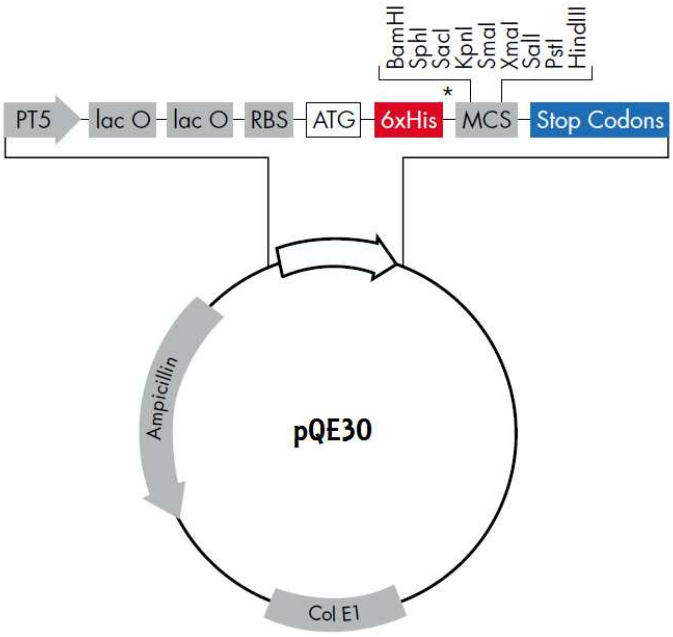


Figure 57 : Carte physique du plasmide pQE30 (Clontech).

II Milieu de culture

II.1 Milieu de culture pour bactéries

II.1.1 LB liquide

Le milieu LB (Luria-Bertani) est utilisé pour la croissance de toutes nos souches de bactéries. Ce milieu (Tableau 3) est stérilisé par autoclave à chaleur humide, puis stocké à température ambiante. La composition de ce milieu est la suivante :

Composition	NaCl	Bactotryptone	Extraits de levure	H2O ultrapure
Concentration en g/L	10	10	5	QSP 1 litre

Tableau 3 : Composition du milieu LB

II.1.2 LB solide

De l'agar à 16 g/l est ajouté au milieu LB liquide avant de réaliser l'autoclave. Ce milieu est ensuite coulé dans des boites de pétries avant de se solidifier par refroidissement. Les boites sont alors conservées à 4 °C.

II.1.3 LB $MgSO_4$ solide

Ce milieu est composé de LB solide classique additionné de $MgSO_4$ à 10 mM. Il est utilisé lors de l'étalement des banques d'ADNc et permet une meilleure adhérence des phages sur la paroi bactérienne.

II.1.4 Gélose molle

Ce milieu est composé de LB liquide additionné d'agar à 8 g/l. Cette gélose a la particularité d'être liquide au-delà de 45 °C et solide en dessous de 40 °C. Celui-ci est utilisé lors de l'étalement des banques d'ADNc.

II.1.5 YEB

Ce milieu (Tableau 4) liquide est utilisé pour la croissance des souches d'*Agrobacterium tumefaciens*. La composition de ce milieu est la suivante :

Composition	Bacto beef extract	Bacto yeast extract	Peptone	Saccharose	H_2O ultrapure
Concentration en g/L	5	1	5	5	QSP 1 litre

Tableau 4 : Composition du milieu YEB.

Le milieu YEB solide est obtenu en additionnant au milieu YEB 16 g/L d'agar.
Le milieu YEB Mg^{2+} solide est obtenu en additionnant au milieu YEB 16 g/L d'agar et 0,25 g/L de $MgSO_4$ avant autoclave.

II.2 Milieu de culture pour levures

II.2.1 YPGA

Ce milieu est un milieu riche (Tableau 5), complémenté en adénine, utilisé pour la croissance des levures WAT 11 et WAT 21 non transformées.

II.2.2 SGI

C'est un milieu sélectif dépourvu en adénine (Tableau 5) qui permet de sélectionner les levures transformées. Le plasmide pYeDP60 contenant le gène d'intérêt introduit dans les levures confère à celui-ci l'auxotrophie pour l'adénine. Seules les levures ayant intégré ce plasmide pourront se développer sur ce milieu minimum.
Des modifications ont été apportées à ce milieu présenté dans la partie résultat.

II.2.3 Milieu glucosé (YPGE)

Ce milieu permet la croissance des levures transformées tout en inhibant la synthèse de la P450 réductase d'*Arabidopis thaliana* et de la protéine d'intérêt. L'absence de pression de sélection de ce milieu permet aux levures de se régénérer en source carbonée *via* le glucose (Tableau 5).

Des modifications de la composition de ce milieu initiale ont été entreprises et sont présentées dans la partie résultat.

II.2.4 Milieu galactosé (YPL)

Ce milieu est un inducteur de synthèse de la P450 réductase ainsi que de la protéine P450 d'intérêt. Le galactose (Tableau 5) active le promoteur GAL10-CYC1 qui contrôle la synthèse concomitante de ces 2 protéines. Il est ajouté au milieu YPGE lorsque celui le glucose est entièrement consommé par les levures.

Des modifications de la composition de ce milieu initial ont été entreprises et sont présentées dans la partie résultat.

	YPGA	**SGI**	**YPGE**	**YPL**
Yeast extract (g/l)	10		10	10
Bactopeptone (g/l)	10		10	10
Hydolysat de Caséine (g/l)		1		
Yeast nitrogen base W.O aminoacids (g/l)		6,7		
Tryptophane (mg/l)		40		
Adénine (mg/l)	30			
Glucose (g/l)	20	20	20	
Galactose (200 mg/ml)				20 ml

Tableau 5: Composition des milieux de culture pour levures

II.3 Conditions et milieux de culture pour le végétal

II.3.1 Culture de *Ruta graveolens* en condition *in vitro*

La culture *in vitro* des plantes se déroule en chambre de culture dans des conditions bien spécifiques : 23 °C le jour et 20 °C la nuit, 40 % d'humidité relative, 160 µmol.m^2.s-^{1} d'intensité lumineuse avec une photopériode de 16 H. La multiplication par micropropagation des plantes est effectuée dans des pots en verre de 200 ml pour *Ruta graveolens*.

Les étapes de manipulation des graines ou de plantes *in vitro* ont lieu sous hotte à flux laminaire vertical. Les pinces et scalpels utilisés pour prélever les explants sont stérilisés au préalable au four pasteur à 180 °C. Le matériel utilisé pour la culture est stérilisé à l'autoclave.

II.3.2 Murashige et Skoog (MS)

La culture *in vitro* de *Ruta graveolens* se fait sur un milieu solide et stérile Murashige et Shoog (Tableau 6) auquel est ajouté 16 g/L d'agar et 30 g/L de saccharose.

Macroéléments mg/L		**Vitamines mg/L**		**Microéléments mg/L**		$CuSO_4$, 5 H_2O	0,025
NH_4NO_3	1650	Inositol	100	KI	0,83	$CoCL_2$, 6 H_2O	0,025
KNO_3	1900	Ac. nicotinique	0,5	H_3BO_3	6,2	**Fer mg/L**	
$CACL_2$, 2 H_2O	440	Pyridoxine HCL	0,5	$MnSO_4$, H_2O	16,9	$FeSO_4$, 7 H_2O	27,8
$MGSO_4$, 7H_2O	370	Thiamine HCL	0,1	$ZNSO_4$, 7 H_2O	8,6	Na_2EDTA, 2 H_20	37,8
KH_2PO_4	170	Glycine	2	$NaMoO_4$, 2 H_2O	0,25	**PH**	5,8

Tableau 6 : Composition du milieu MS.

Ce milieu MS est supplémenté en benzylaminopurine (BAP) à 0,1 mg/l et en acetosyringone à 50 mg/l pour la coculture des explants de Rue, après inoculation avec les agrobactéries.

Après lavage des explants, il est utilisé en tant que milieu de régénération et supplémenté en BAP (0,1 mg/l) et céfotaxime (300mg/l).

Enfin, il est utilisé comme milieu de sélection pour détecter les cellules de Rue transformées. Il est alors additionné de BAP (0,1 mg/l) et de kanamycine (75mg/l).

II.3.3 Culture des plantes en sol

Les plantes sont mises en culture sous serre (22 °C) dans des pots individualisés contenant du terreau avec un arrosage tri hebdomadaire.

II.3.4 Terreau

Le semis ainsi que le rempotage des différentes plantes ont été réalisés dans un terreau universel correspondant à la norme NF U 44-551. Celui-ci contient un assemblage de tourbe blonde et noire à hauteur de 60 %, des écorces de bois et fibres de pin compostées allégeant le terreau, de la fibre de coco afin d'assurer une bonne porosité du milieu, de la vermiculite et de la perlite pour favoriser le drainage.

II.4 Antibiotiques

L'addition d'antibiotique dans un milieu de culture entraine une pression de sélection qui permet de sélectionner les souches de bactéries et de plantes recombinantes. Les antibiotiques sont préparés à partir de poudre et sont dilués dans de l'eau, ou de l'éthanol pour le chloramphénicol. Ils sont ajoutés dans le milieu de culture après autoclave selon une concentration spécifique pour chacun (Tableau 7).

Antibiotiques	Concentrations finales
Ampicilline	50 µg/ml
Carbénicilline	100 µg/ml
Chloramphénicol	30 µg/ml
Kanamycine	50 µg/ml
Rifampicine	100 µg/ml
Spectinomycine	50 µg/ml

Tableau 7 : Concentration finale des antibiotiques utilisés.

III Outils de biologie moléculaire

III.1 Extraction d'ADN génomique de plante

Les extractions d'ADN génomique réalisées sur les différentes plantes telles que *Ruta graveolens*, *Pastinaca sativa*, *Arabidopsis thaliana* sont effectuées après un broyage des feuilles, à l'aide d'un mortier contenant de l'azote liquide. La poudre obtenue est ensuite utilisée pour effectuer l'extraction d'ADN génomique, suivant le protocole décrit par le kit « DNeasy plant minik de Qiagen » (Annexe 1). L'ADN extrait issu des différentes plantes est analysé sur gel et dosé avant d'être congelé à -20 °C.

III.2 Extraction D'ARN de plante

L'extraction totale d'ARN est une opération délicate du fait de la présence importante d' ARNase dans le milieu environnant. De plus, les ARN sont thermolabiles et se dégradent rapidement à température ambiante. Des précautions particulières sont donc prises pour pallier ces contraintes. Une solution de décontamination et de neutralisation (RNaseZap® de chez Ambion) des ARNase est appliquée sur la paillasse, ainsi que sur les gants de l'utilisateur. Après l'extraction des ARN suivant le protocole du kit « RNeasy plant minikit de Qiagen » (Annexe 2), ceux-ci sont immédiatement dosés, aliquotés et congelés à -80 °C. Un test de qualité est ensuite réalisé en effectuant une migration sur gel d'agarose. La visualisation des bandes majoritaires correspondant aux ARN ribosomiques cytoplasmiques et chloroplastiques indique l'absence de dégradation des ARN.

III.3 Amplification par PCR d'un fragment d'ADN.

III.3.1 Amplification classique

La réaction en chaine par polymérase ou PCR est une méthode d'amplification d'ADN *in vitro*. Une étape initiatrice de chauffage à 95 °C pendant 10 min est réalisée avant d'effectuer les cycles de PCR proprement dits (Tableau 8). Cette étape permet de deshybrider les ADN doubles brins, d'homogénéiser le milieu réactionnel par agitation thermique, de casser les structures secondaires présentes ou d'activer l'enzyme polymérase dans certains cas (Platinium technology). La première phase du cycle de PCR débute également par une phase de dénaturation de l'ADN à 95 °C permettant de séparer les ADN double brin.

Puis, une étape d'hybridation des amorces oligonucléotidiques à l'ADN matricielle est effectuée à une température qui leur est thermodynamiquement favorable comprise entre 50 °C et 65 °C. Cette température détermine le degré de spécificité de l'hybridation, elle est spécifique à chaque amorce et constitue la température optimale d'appariement spécifique de l'amorce au brin complémentaire. Cette température d'hybridation est calculée suivant la température de fusion (Tm) des amorces utilisées (c'est la température à laquelle la moitié de l'ADN est sous forme monobrin et l'autre moitié sous forme double brin.). La température de fusion des amorces est dépendante de leur longueur et de la quantité en base purique. Une formule mathématique permet de calculer une approximation de cette température de fusion suivant l'équation :

Tm = 69,3 + 0,41 x (% GC) – 650/N, avec N correspondant au nombre de paires de bases de l'amorce oligonucléotidique.

Afin de pouvoir effectuer une hybridation optimale simultanée des amorces sens et antisens, les températures de fusion de celles-ci doivent être proches.

La température d'hybridation correspond à Tm - 5 °C.

Enfin, la dernière étape consiste à réaliser la synthèse du brin complémentaire à l'aide de la *Taq* DNA polymérase qui effectue cette élongation à 72 °C. Ce brin est fabriqué à partir des dNTPs libres présents dans le milieu réactionnel. La durée de cette étape dépend de la longueur de l'amplicon (en général 1 min/1000 pb).

La succession de ces trois étapes répétées un certain nombre de fois, constitue le nombre de cycles nécessaires pour amplifier une séquence d'ADN.

Étapes	**Durée**	**Température**	**Cycles**
Dénaturation initiale	10 min	95 °C	1
Étape de dénaturation	30 s à 60 s	95 °C	25 à 35
Étape hybridation	30 s à 90 s	50 °C à 65 °C	
Étape élongation	30 s à 90 s	72 °C	

Tableau 8 : Les différentes étapes d'une PCR classique.

Toutes ces différentes étapes sont réalisées à l'aide d'un thermocycler Icycler (Bio Rad) ou Gradient Cycler PTC-200 (MJ Research). D'une manière générale, les réactions sont réalisées dans un volume final de 20 µl comprenant : 0,4 µM de chaque amorce oligonucléotidique

sens et antisens, 0,4 mM de dNTP, 2 µl de tampon d'activité polymérasique 10X, 2 à 5 U de *Taq* DNA polymérase ainsi que 10 à 100 ng d'ADN matriciel.

III.3.2 PCR par essais

La PCR par essais (Touchdown PCR) est un protocole utilisé pour amplifier de l'ADN faiblement représenté ou subissant une compétition sur leurs amorces par des produits de pseudogène. Les températures d'hybridations lors des premiers cycles sont très élevées permettant d'obtenir une forte stringence et d'augmenter la spécificité d'amplification. La séquence d'intérêt devient alors progressivement majoritaire vis-à-vis des ses compétiteurs, la température d'hybridation est au fur à mesure diminuée afin d'assurer une meilleure efficacité de la PCR (Tableau 9).

Étapes	Durée	Température	Cycles
Dénaturation initiale	10 min	95 °C	1
Étape de dénaturation	30 s à 60 s	95 °C	15
Étape hybridation	30 s à 90 s	65 °C – 1 °C/cycle	
Étape élongation	30 s à 90 s	72 °C	
Étape de dénaturation	30 s à 60 s	95 °C	20
Étape hybridation	30 s à 90 s	50 °C	
Étape élongation	30 s à 90 s	72 °C	

Tableau 9 : Les différentes étapes d'une PCR par essais.

III.4 Reverse Transcriptase PCR

La RT-PCR (Reverse Transcriptase PCR) est une technique qui associe une transcription inverse (RT) suivie d'une PCR. Elle permet de former par complémentarité des ADNc à partir d'un brin ARN. L'ADN polymérase ARN dépendante (ou transcriptase inverse) utilise les desoxyribonucléotides afin de former des ADNc. . Les réactions s'effectuent suivant le protocole décrit dans le kit « high capacity RNA-to-cDNA master mix » d'Applied Biosystems (Annexe 3). Les ADNc peuvent ensuite être destinés à être amplifiés par PCR.

La RT-PCR en une seule étape (single step RT-PCR) permet d'effectuer les étapes de rétrotranscription et d'amplification dans le même tube simultanément. En effet, le protocole décrit dans le kit «SuperScript™ One-Step RT-PCR with Platinum® *Taq* » mélange les réactifs de RT et de PCR. La *Taq* ne pouvant travailler sur de l'ARN, la première étape consiste a former les ADNc *via* la RT, puis la *Taq* est activé à 95 °C est commence l'amplification. Cette technique permet de réduire le risque d'inversion des tubes ou de contamination des échantillons (Annexe 4).

III.5 Digestion par des enzymes de restrictions

La majeure partie des enzymes de restrictions que nous utilisons proviennent du système « fast digest » de Fermentas qui permet de diminuer les temps d'incubation d'un facteur 10 par rapport aux systèmes standards. De plus, ce système permet d'utiliser plusieurs enzymes de restriction dans une même réaction, car les tampons d'activités sont tous compatibles pour chacune des enzymes. La réaction de digestion est généralement réalisée dans un volume final de 20 µl contenant 2 µl de tampon d'activité 10X, 10 U d'enzyme de restriction et 1 µg d'ADN à digérer. La réaction est mise à incuber pendant 5 min à 37 °C.

III.6 Ligations

III.6.1 PCR8/TOPO/TA/GW®

Les ligations dans le plasmide pCR8/GW se font à l'aide d'un kit « pCR®8/GW/TOPO® TA Cloning® ». L'avantage de ce kit réside dans le fait que la ligation du gène d'intérêt dans le plasmide PCR8/GW se fait sans utilisation au préalable d'enzyme de restriction ni de ligase. La présence de la topoisomérase (Figure 50) fixée de manière covalente sur des sites spécifiques du plasmide permet de liguer un produit de PCR sur lequel a été rajoutée au préalable une adénine à l'extrémité 3'. Cette addition du nucléotide adénine supplémentaire est réalisée lors de l'étape d'amplification du gène d'intérêt. A la fin de la PCR, une dernière étape d'extension de 10 min est réalisée à 72 °C qui permet à la Taq polymérase de rajouter cette adénine de manière non spécifique à l'extrémité 3'. Ce produit de PCR est ensuite ligué suivant le protocole décrit dans le kit pCR®8/GW/TOPO® TA Cloning®. Le volume de la réaction finale se fait dans 6 µl avec 10 ng de vecteur PCR8 linéarisé, 1,2 µmol de NaCl et

0,06 µmol de MgCl2 et 50 à 100 ng/µl de produits de PCR. Le mix réactionnel est mis à incuber pendant 1 heure à température ambiante afin d'assurer 95 % de ligation.

III.6.2 Dans les autres vecteurs

La ligation dans des vecteurs standards se fait à l'aide de l'enzyme « T4 DNA ligase » d'Invitrogen qui permet de réaliser la ligation des gènes d'intérêts avec des extrémités cohésives ou franches dans les vecteurs d'expression. La réaction s'effectue dans 20 µl de volume final contenant 4 µl de tampon d'activité 5X, 0,1 U de T4 DNA ligase, 3 à 30 fmol de vecteur et 9 à 90 fmol d'insert. La réaction est mise à incuber à 16 °C pendant 4 heures.

III.7 Électrophorèse sur gel d'agarose

L'électrophorèse sur gel d'agarose est utilisée pour séparer l'ADN ou l'ARN en fonction de leur taille et de leur charge électrique. Plus le gel d'agarose est concentré et réticulé, et plus la taille des pores du gel est petite. On peut ainsi ajuster les propriétés du gel à la taille des molécules à analyser. Généralement, un gel constitué à 1 % d'agarose permet de séparer un large éventail de molécules. Celui-ci est mis à tremper dans du tampon TAE (Tris acétate 40 mM, EDTA 1mM). Avant le dépôt dans les puits du gel, les échantillons à séparer d'ADN ou d'ARN (20 µl) sont additionnés de 4 µl de tampon de charge (Tris 50 mM pH 7,5 ; EDTA 100 mM pH 7,5 ; 10 % glycérol ; 0,05 % bleu de bromophénol). Le gel est ensuite soumis à une tension de 100 V (10 V/cm) pendant 45 min. Un marqueur de taille coloré (Fermentas massruler ready-to-use #SM0403) migre en même temps que les autres échantillons sur le gel d'agarose et permet d'identifier la taille approximative des différentes bandes.

La révélation se fait à l'aide de bromure d'éthidium ou BET. Il s'agit d'un agent intercalant communément employé comme marqueur d'acide nucléique en biologie moléculaire. Il permet de visualiser l'ADN ou l'ARN par fluorescence en présence d'un rayonnement ultra-violet (λ=254 nm). Le gel d'agarose est baigné pendant 10 min dans un bain de bromure d'éthidium à 1 mg/l et placé sous rayonnement UV.

III.8 Électrophorèse sur gel polyacrylamide en condition dénaturante ou SDS PAGE

L'électrophorèse en gel de polyacrylamide contenant du DodecylSulfate de Sodium ou SDS est une technique de biologie moléculaire couramment utilisée pour séparer des protéines suivant leur poids moléculaire apparent. La concentration en acrylamide détermine le niveau de porosité de la matrice et influe sur la vitesse de migration des protéines. Les protéines sont tout d'abord condensées dans un gel de concentration à 5 %, puis séparées dans un gel de séparation entre 7 et 15 % d'acrylamide en fonction de la taille des protéines à séparer. Un gel de séparation à 10 % d'acrylamide permet de séparer un large spectre de protéine. Le dodecylsulfate de sodium est l'agent dénaturant qui se lie aux protéines, les empêche de se replier, et leur confère une charge nette négative. 10 µl d'échantillons protéiques sont au préalable additionnés de 10 µl de tampon de charge x2 (Tris HCl pH 6,8 125 mM, SDS 5 %,Glycérol 20 %, β-mercaptoéthanol 5 %, bleu de bromophénol 0,05 %), puis chauffés 3 min à 98 °C. Les échantillons sont ensuite chargés dans les différents puits du gel de polyacrylamide. Celui-ci est immergé dans une cuve d'électrophorèse contenant un tampon de migration (Tris/Glycine x1 (à partir de tampon x10 : Tris 30,4 g, Glycine 144g pour 1L pH 8,3), SDS 0,1 %). Un champ électrique de 100 V est appliqué afin que les protéines migrent dans la matrice suivant leurs poids moléculaires. Lorsque les protéines sont séparées, elles sont visualisées par coloration au bleu de Coomassie. Le gel est ensuite décoloré par un mélange éthanol/acide acétique/eau (25/40/35) afin de ne visualiser que les bandes protéiques.

III.9 Extraction d'ADN à partir d'un gel d'agarose

Après migration et révélation aux UV, les bandes de gel correspondant aux ADN ou ARN identifiées peuvent être prélevées au vu d'un clonage à l'aide d'une lame de scalpel propre, et sont introduites dans des tubes de 2 ml. L'utilisation du kit d'extraction Qiagen « QIAquick Gel Extaraction Kit » permet de purifier, et d'extraire à partir d'un gel, les bandes d'ADN suivant le protocole décrit par le kit (Annexe 5).

III.10 Préparation de bactéries compétentes

III.10.1 Bactéries électrocompétentes

Les bactéries utilisées au laboratoire pour la transformation par chocs électriques sont des *E. coli* genhogs. Une préculture est effectuée en ensemençant 10 ml de milieu LB. Après une nuit d'incubation à 37 °C sous agitation, 1 ml de cette préculture est prélevé afin d'ensemencer 100 ml de milieu LB. La culture est incubée à 37 °C jusqu'à l'obtention d'une DO_{700} de 0,7. Une centrifugation à 5000 g sans frein à 4 °C pendant 15 min est effectuée. 2 lavages successifs avec de l'eau stérile, puis 2 lavages dans 20 ml de glycérol 10 % sont réalisés. Les bactéries compétentes sont resuspendues dans 1 ml de glycérol 10 %, aliquotées (40 µl), et congelées dans de l'azote liquide avant d'être stocké à -80 °C. La conservation peut durer 6 mois.

III.10.2 Bactéries chimiocompétentes

Différentes souches de bactéries peuvent être rendues chimiocompétentes. Une préculture d'une colonie d'E. *coli* est effectuée dans du LB additionné de l'antibiotique de sélection. 1 ml de cette préculture est utilisé pour ensemencer 50 ml de milieu LB additionné du même antibiotique. Lorsque la $D0_{600}$ atteint 0,5, les bactéries sont sédimentées par centrifugation à 4000 g pendant 10 min. Le culot est resuspendu dans 10 ml de $CaCl_2$ stérile à 100 mM. Une dernière centrifugation à 4000 g pendant 10 min permet de culotter les bactéries qui sont resuspendues dans 1 ml de $CaCl_2$. Les bactéries sont alors incubées sur de la glace pendant 1 heure, avant d'être utilisables pour une transformation par choc thermique.

III.11 Transformation de bactéries compétentes

III.11.1 Électroporation

Cette technique de transformation consiste à insérer une construction plasmidique dans des bactéries rendues électrocompétentes. L'application d'un choc électrique de courte durée et de haut voltage entraine une déstabilisation des membranes bactériennes, et rend celles-ci ponctuellement poreuses et perméables à une construction plasmidique. Une aliquote de 40 µl de bactéries électrocompétentes est mélangée avec 1 à10 ng de plasmide ou produit de ligation. Le mélange est introduit dans une cuve à électroporation (2 mm de distance entre les

2 électrodes) placée dans un pulseur (BioRad micropulser ™). L'échantillon est soumis à un choc thermique (V = 2,5 kV ; R = 200 Ω ; C = 25 µF) puis suspendu dans 500 µl de milieu LB pendant 30 min à 37 °C. Ce temps de régénération bactérien permet aux bactéries transformées d'acquérir la résistance à l'antibiotique apporté par le plasmide. La sélection des colonies transformées se fait par étalement de la culture sur boite LB sélective à 37 °C pendant une nuit.

III.11.2 Choc thermique

Cette technique est utilisée pour transformer les souches bactériennes chimiocompétentes. 1 µl de construction plasmidique est ajouté à une aliquote de 100 µl de bactéries compétentes. L'ensemble est incubé sur de la glace pendant 30 min. Un choc thermique est ensuite réalisé durant 45 secondes à 42 °C à l'aide d'un bain-marie, puis les bactéries sont immédiatement transférées sur de la glace pendant 5 min. 500 µl de milieu LB sont additionnés, ensuite le mélange est incubé pendant 30 min à 37 °C. La sélection des bactéries transformées se fait sur boite LB contenant l'antibiotique de sélection à 37 °C pendant une nuit.

III.11.3 Extraction d'ADN plasmidique

Une amplification plasmidique est réalisée en cultivant une colonie bactérienne contenant le plasmide d'intérêt à 37 °C pendant une nuit. Les bactéries sont ensuite sédimentées par centrifugation à 10 000 g pendant 3 min. L'ensemble est resuspendu dans une solution avant de lyser les bactéries par réaction alcaline. Les plasmides sont extraits sur colonne suivant le protocole décrit dans le kit Genelute™ plasmide miniprep (Sigma-Aldrich) (Annexe 6). Les 50 µl de plasmides sont stockés à -20 °C

III.12 Construction et criblage d'une banque d'ADNc

III.12.1 Construction de la banque d'ADNc

La construction de la banque d'ADNc de feuille de Rue traitée aux UV a été réalisée en 2006 dans notre laboratoire par Romain LARBAT lors de sa thèse (cf. thèse Romain LARBAT, INPL pour détails de la construction). Différentes aliquotes de 1 ml ont été conservés à -80 °C et ont été utilisés pour réaliser un criblage.

III.12.2 Titration de la banque

L'étape préliminaire consiste à évaluer la titration de la banque d'ADNc afin de connaître la quantité de plage de lyse générée par unité de volume. Pour déterminer cette titration, la banque est diluée en cascade 9 fois d'un facteur 10. Les bactéries *E.Coli* XL1 blue sont utilisées comme hôte pour la multiplication des phages. Une colonie est mise à incuber à 37 °C dans 10 ml de milieu LB additionné de $MgSO_4$ 10mM et de maltose 0,2 % jusque-là DO_{600} 0,7. Le MgSO4 et le maltose favorisent l'adsorption des phages sur les membranes des bactéries. Les bactéries sont ensuite sédimentées par centrifugation à 5000 g pendant 5 min et resuspendues dans 10 ml de MgSO4 10 mM. 300 µl de cette suspension sont mélangés avec 10 µl des différentes dilutions de la banque. L'adsorption des phages sur les bactéries se fait pendant 15 min à 37 °C. 7 ml de gélose molle maintenue à 45 °C est additionné à ce mélange. Un étalement est ensuite réalisé sur boite de pétrie (90 mm de diamètre) contenant du LB agarose $MgSO_4$ à 10 mM. L'incubation se fait à 37 °C pendant 15h. Les plages de lyses sont ensuite comptabilisées. Le titre de la banque est exprimé en Plaque Forming Unit par ml de banque (pfu/ml). La banque de feuille de Rue traitée aux UV titre à 5.1 pfu/ml.

III.12.3 Étalement de la banque d'ADNc

L'étalement de la banque se fait de la même manière que la titration de la banque. 200 000 plages de lyse réparties sur 4 grandes boites (150 mm de diamètre), soit 50 000 plages par boite, doivent être criblées. La quantité adéquate de phages est additionnée à 500 µl de bactéries *E.Coli* XL1 blue. Le mélange est incubé pendant 15 min à 37 °C. 12 ml de gélose molle sont additionnés puis un étalement est réalisé sur les grandes boites.

III.12.4 Transfert et fixation sur membrane du nylon

Les plages de lyses sont transférées sur une membrane de nylon chargée Amersham Hybond™-N+ par simple application pendant 1 min. 2 membranes sont utilisées successivement par boite de Pétrie. Les membranes sont récupérées afin de leur appliquer 3 différents lavages successifs pendant 3 min sans agitation. Un premier tampon basique de dénaturation (Tableau 10) permet de libérer l'ADN des phages, puis un deuxième tampon neutralise la dénaturation, et enfin un tampon de rinçage SSC 2X est appliqué sur la

membrane. La fixation des ADN sur la membrane se fait par formation de liaisons covalentes sous l'effet d'un rayonnement UV (254 nm, 0,12 J pendant 30 s).

Tampon de dénaturation	NaOH 0,5M	NaCl 1,5M
Tampon de neutralisation	Tris HCl pH 7	NaCl 1,5M
Tampon de rinçage SSC 20X	Citrate de Na, ajuster pH 7	NaCl 3 M

Tableau 10 : Compositions des tampons utilisés pour le criblage de banque.

III.12.5 Synthèse de la sonde

La synthèse de la sonde d'hybridation se fait suivant le protocole décrit dans le kit «Amersham Gene Images AlkPhos Direct Labelling and Detection System » (Annexe 7). Dans un premier temps, la sonde est diluée pour obtenir une concentration finale de 10 ng/µl d'ADN. 10 µl de cette sonde sont ensuite dénaturés dans un bain-marie à 95 °C pendant 5 min puis refroidi sur de la glace. 10 µl de tampon réactionnel X2 sont additionnés. 2 µl de produit de ligation contenant de la phosphatase alcaline sont introduits au mélange. L'ajout de 10 µl d'une solution de cross-linker dilué au 1/5 permet de fixer la phosphatase alcaline à la sonde de manière covalente. Le mix réactionnel est incubé pendant 30 min à 37 °C. La sonde, utilisable pendant les deux heures après sa synthèse, est stockée sur de la glace.

III.12.6 Hybridation

Le tampon d'hybridation fourni dans le kit est additionné de 0,5M de NaCl et 4 % (V/W) de Blocking reagent. Le volume de tampon d'hybridation à utiliser correspond à 0,25 ml/cm2 de membrane à hybrider. Celui-ci est préchauffé à 55 °C et les membranes y sont immergées pendant 15 min sous agitation. La sonde est ensuite ajoutée à raison de 10 ng/ml de tampon d'hybridation. L'hybridation se prolonge à 55 °C durant toute la nuit. La température d'hybridation détermine la spécificité d'hybridation de la sonde.

III.12.7 Lavage des membranes

Cette étape permet d'affiner la stringence du criblage et de se débarrasser des hybridations non spécifiques de la sonde. Les membranes sont introduites dans un premier tampon de lavage en excès (5 ml/cm2 de membranes à rincer) (Tableau 11) pendant 10 min à 55 °C avec agitations orbitales. Ce lavage est réitéré une seconde fois. Puis les membranes sont transférées 2 fois 5 min dans un tampon de lavage secondaire 1X en excès additionné de 2mM de MgCl2.

Tampon de lavage primaire	**Tampon de lavage secondaire 20X**
Urée 2M	Tris/HCl 1M
SDS 0,1 % (W/V)	NaCl 2M
NaCl 150 mM	Ajuster pH 10
Na3PO4 50 mM pH 7	Utilisable après dilution au 1/20 et addition de MgCl2 2mM
MgCl2 1 mM	
Blocking reagent 0,2 % (W/V)	

Tableau 11 : Composition des tampons de lavages.

III.12.8 Révélation

L'utilisation de CDP-Star (1,2-dioxetane) Amerscham permet de révéler la sonde fixée sur les acides nucléiques homologues. Ce substrat est métabolisé par la phosphatase alcaline en un produit chimioluminescent (Figure 58) détectable sur un film photoradiographique. Les membranes sont mises à tremper dans 10 ml de CDP-star à 0,25 mM pendant 1 min. Pour éviter d'assécher les membranes, un film plastique transparent est appliqué dessus. Les membranes sont ensuite introduites dans une cassette sous obscurité avec un film photoradiographique. Le film et les membranes sont en contact pendant 3H. Le film est ensuite trempé quelques secondes dans une solution de révélation, puis de fixation (Kodac revelator et fixator) afin de révéler les spots luminescents.

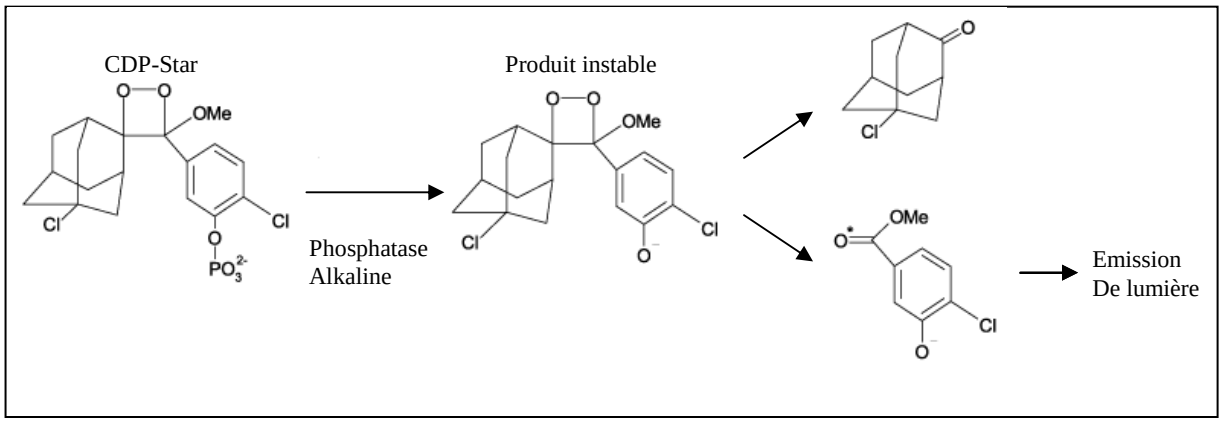


Figure 58 : Métabolisation du CDP-Star.

III.12.9 Récupération des plages de lyse

Les spots visualisés sur le film radiographique sont identifiés sur la boite de Petri correspondante. Un carottage de la plage de lyse correspondant au spot luminescent est effectué à l'aide de la pointe d'un cône de micropipette. Les carottes de géloses sont introduites dans des tubes de 2 ml contenant 500 µl de tampon SM (NaCl 100 mM ; MgSO4,7H20 10 mM, Tri HCL pH 7,5 35 mM) additionné de 100 µl de chloroforme. Les tubes sont mis à incuber pendant une nuit à 4 °C. Le lendemain, 10 µl de surnageant sont dilués 100 fois dans du $MgSO_4$ 10 mM. Cette dilution est utilisée pour réaliser un criblage secondaire.

III.12.10 Criblage secondaire

Le second criblage permet d'affiner la sélection des phages contenant un gène hétérologue à la sonde. Toutes les étapes précédemment décrites sont effectuées une seconde fois. L'étalement se fait sur petite boite de pétrie afin de cribler 50 à 100 plages de lyse par boite. Le tampon d'hybridation contenant la sonde est réutilisé pour ce deuxième criblage. Après carottage des plages de lyse, le plasmide du phage contenant le gène d'intérêt est amplifié par transformation des bactéries *E.Coli* BM 25.8, rendues au préalable chimiocompétentes.

III.12.11 Excision des plasmides

Une colonie de *E.Coli* BM 25.8 est cultivée toute la nuit à 31 °C dans 10 ml de LB additionné de $MgSO_4$ 10 mM. 200 µl de cette culture sont mélangés à 100 µl de la suspension de phage contenue dans le tampon SM. Une incubation de 30 min à 31 °C sans agitation est effectuée. 500 µl de LB sont ajoutés et incubés à 31 °C pendant 1 heure. Les phagemides λTriplEx2 contenant les gènes d'intérêt sont alors excisés et recircularisés par le système Cre/Lox en plasmide pTriplEx2 dans la bactérie BM 25.8. Un étalement sur boite de Pétri avec ampicilline permet de sélectionner les clones transformés. Une amplification plasmidique contenant le gène d'intérêt est réalisée puis envoyée à séquencer.

III.13 Banque de cDNA normalisée

Deux banques ont été commandées : l'une à partir de *Pastinaca sativa* et l'autre à partir de *Ruta graveolens* au préalable induites aux UV. Les ARN extraits (400 µg) de ces 2 plantes sont envoyés à MWG Biotech en Allemagne. Le type de séquençage est GS FLX (type 454). La banque normalisée nous est renvoyée sous forme de fichier informatique utilisable avec le logiciel Bioedit. Un blast est réalisé sur ces 2 banques avec les différentes dioxygénases isolées chez *Ruta graveolens* afin d'identifier d'autres potentielles dioxygénases chez la Rue et le panais.

III.14 Séquençage

Le séquençage des ADN est réalisé par l'entreprise MWG Biotech en Allemagne. 1 µg de plasmide contenant le gène à séquencer est envoyé à l'entreprise. Le résultat du séquençage en 3' et 5' nous est parvenu en quelques jours sous format informatique.

III.15 Induction des plantes aux UV

Les plantes de *Ruta graveolens* âgées de 2 à 3 mois sont élicitées dans une chambre à l'aide d'une lampe à UVB (λ=312 nm) placée à 20 cm des parties aériennes pendant 24 h. Les plantes, individualisées dans des pots de terre de 15 cm de diamètres, sont exposées aux UV sous différents angles afin d'assurer une homogénéisation de l'élicitions (Photo 8).

Photo 8 : Induction des plantes aux UV.

III.16 Inhibition de la synthèse de furocoumarines par le Regalis

Le prohéxadium de calcium (ou Regalis) est un inhibiteur spécifique des dioxygénases. Il nous a été gracieusement fourni par BASF France.

Le prohexadium de calcium est dilué dans de l'eau ultra pure à une concentration de 500 ppm. 1 lot de 6 plantes contrôle et 1 lot de 6 plantes est traité avec cet inhibiteur de dioxygénase. Le Regalis est administré sur les parties aériennes de la plante par pulvérisation (Photo 9). Deux traitements sont effectués à 6 heures d'intervalle. Les plantes contrôles reçoivent uniquement de l'eau ultra pure. 4 jours après, une extraction des phénylpropanoïdes est réalisée sur les 2 lots. Les résultats sont comparés et analysés en HPLC et SM.

Photo 9 : Vaporisation d'un inhibiteur spécifique des dioxygénases (le Regalis®) sur les parties aériennes des plantes de *Ruta graveolens*.

III.17 Modélisation tridimensionnelle des dioxygénases

Les modèles tridimensionnels des différentes dioxygénases ont été modélisés automatiquement par un logiciel de Swissprot (**http://swissmodel.expasy.org/**). Les modèles générés ont été établis par homologie avec les données structurales de la leucoanthocyanidin dioxygénase (1gp6) disponible sur le site Protéine Data Bank (**http://www.rcsb.org/pdb/**). La visualisation des modèles tridimensionnels s'est effectuée à l'aide du logiciel Molecular Operating Environnement (MOE) (Chemical Computing Group, Montreal).

IV Système d'expression hétérologue

Un système d'expression hétérologue permet de produire des protéines qui ne sont naturellement pas synthétisées par l'organisme qui les produit. L'organisme recombinant est obtenu par génie génétique. Le principe est de sous cloner le gène complet codant pour la protéine en aval d'un promoteur permettant une transcription importante de celui-ci. Le plasmide recombinant est ensuite placé dans une souche compétente adaptée à la production.

IV.1 Système d'expression dans les procaryotes

IV.1.1 La bactérie, un système inductible

Un système d'expression inductible permet de contrôler le niveau de production de la protéine recombinante. Le plus souvent, une première étape de culture est destinée à la croissance de la biomasse de l'organisme recombinant utilisé. Puis une deuxième étape de culture est dédiée à la production de protéines d'intérêts. L'avantage de ce système réside dans la possibilité de protéger l'organisme, contre la toxicité fréquente des protéines hétérologues accumulées. On tente donc de limiter le plus possible leurs expressions pendant la première phase de croissance.

IV.1.1.1 Expression *in vitro*

IV.1.1.1.1 Culture

Une colonie de bactéries *E.coli* BL21(DE3) est incubée dans 100 ml de milieu LB additionné de l'antibiotique de sélection. Lorsque la DO_{600} atteint 0,7, la synthèse de la protéine est induite par l'ajout dans le milieu de culture d'isopropyl-Thio-β-D-galactopyranoside (IPTG) à 1 mM. La culture induite est incubée pendant 20 h à 16 °C. Une centrifugation est ensuite réalisée à 5000g pendant 10 min à 4 °C. Le culot est lavé 3 fois dans 5 ml de PBS froid (8,1 mM Na2HPO4, 1,76 mM KH2PO4, 2,7 mM KCl, 137 mM NaCl, pH7,4) avant d'être finalement repris dans 1 ml de PBS. L'échantillon est alors soniqué (Bandelin Sonoplus HD 2070 avec une sonde MS73 à une intensité de 200W/cm2) 3 fois pendant 30 secondes dans de la glace, puis centrifugé à 16 000 g pendant 10 min à 4 °C.

IV.1.1.1.2 Purification de la protéine

La protéine d'intérêt est additionnée d'une queue Histidine lors de sa construction dans le plasmide pET-28a qui permet une purification à l'aide du kit « Ni-NTA purification system » de Qiagen. Cette protéine de fusion est purifiée sur colonne d'affinité suivant le protocole décrit dans ce kit d'Invitrogen (Annexe 8). Après 3 lavages, la protéine adsorbée sur la colonne est diluée dans 300 µl de solution contenant 500 mM d'imidazole, 300 mM Nacl et 50 mM NaH2PO4 pH8. La protéine est ensuite directement utilisée pour réaliser des tests d'activités ou aliquotée afin d'être stockée à -80 °C.

IV.2 Système d'expression dans les eucaryotes type levure

Ce système d'expression est aussi un système inductible permettant de contrôler le niveau d'expression de la protéine d'intérêt.

IV.2.1.1 Préparation de levures compétentes

Une préculture de colonie de levures WAT 11 est incubée dans 10 ml de YPGA pour la nuit à 30 °C. Cette préculture est diluée afin d'obtenir une $D0_{700}$ de 0,2. Dès que la $D0_{700}$ atteint 0,9, une centrifugation à 4000 g pendant 10 min est effectuée. Le culot est lavé dans 1 ml d'eau ultra pure stérile, avant d'être centrifugé à nouveau pendant 2 min à 4000 g. Le culot est resuspendu dans 1 ml d'acétate de lithium 100mM 1X. Une dernière centrifugation permet de concentrer le culot avec 200 µl d'acétate de lithium.

IV.2.1.2 Transformation de levures

La transformation des levures nécessite d'incuber pendant 30 min à 30 °C : 50 µl de levures compétentes, 50 µl d'acétate de lithium 100 mM/TE 1X dilué dans du PEG 4000 (50 %), 10 µl d'ADN de sperme de saumon (10 mg/ml) dénaturé pendant 20 min à 100 °C, et 100 ng de vecteurs à insérer. Un choc thermique est effectué pour réaliser la transformation, en incubant pendant 15 min le mélange à 42 °C, puis dans de la glace pendant 2 min. Après centrifugation à 3000 g pendant 1 min, les levures sont lavées dans 1 ml d'eau stérile ultra pure. Elles sont resuspendues dans 300 µl de milieu YPGA et mise à régénérer pendant 2 h à 28 °C. Après une ultime centrifugation, les levures sont reprises dans 500 µl du milieu sélectif SGI et sont incubées sur boite SGI pendant 3 nuits à 28 °C.

IV.2.1.3 Expression *in vitro*

En fonction du P450 exprimé, de nombreuses modifications ont été apportées aux méthodes mises au point par Pompon *et al* (1996) et Dieperger (1974) afin d'essayer d'obtenir une protéine fonctionnelle. La caractérisation de la fonctionnalité d'un P450 se fait par analyse du spectre d'absorbance à 450 nm après ajout de CO, et par criblage métabolique.

IV.2.1.3.1 Culture

La première étape consiste à augmenter la biomasse des levures transformées sans induire la synthèse de P450. La synthèse de P450 est réprimée par la présence de glucose dans le milieu de culture. Une préculture est réalisée en ensemençant 10 ml de milieu SGI (Tableau 5) avec une colonie de levures transformée pendant 24 h à 28 °C. 200 ml de milieu YPGE (Tableau 5) sont ensuite ensemencés avec 5 ml de cette préculture pendant 24 h à 28 °C avec une agitation orbitale. La deuxième étape consiste à induire la synthèse du P450 en substituant le glucose consommé par du galactose dans le milieu de culture. Le milieu YPGE est additionné de 20 ml de galactose (200 mg/l). Les levures sont incubées 16 h à 28 °C.

De nombreuses modifications de ce protocole initiales ont été apportées afin d'obtenir une meilleure expression. Les modifications et les résultats d'expression sont présentés dans la partie résultat.

IV.2.1.3.2 Préparation de microsomes

Les protéines P450 sont le plus souvent thermolabiles. Elles nécessitent d'être manipulées à de basses températures (4 °C) pour ne pas altérer la fonctionnalité de la protéine. Les levures, dont la synthèse de P450 est induite par l'ajout de galactose, sont sédimentées dans une centrifugeuse réfrigérée à 4 °C pendant 7 min à 5000g. Elles sont lavées dans 20 ml de TEK à 4 °C (Tableau 12). Après une centrifugation à 5000 g pendant 7 min, le TEK est remplacé par 2 ml de tampon d'extraction froid (Tableau 12). Le mélange est transféré dans un tube scintillation de 20 ml contenant pour moitié des billes de verres (Glass beads 400-625 µm, Sigma). les levures sont broyées par action mécanique, soit aprés agitation manuel, soit dans un Vibro Broyeur Retsch MM40. 3 agitations de 1 min à une fréquence de 1/30 sont effectuées. Entre chaque agitation, le tube est placé dans la glace pendant 2 min. À l'issue du broyage, les billes sont lavées 4 fois avec 5 ml de tampon d'extraction froid. Ces fractions combinées dans un tube falcon sont centrifugées à 10 000 g pendant 15 min. La précipitation des microsomes se fait soit de manière chimique (ajout de NaCl ; concentration finale 0,15M) soit par unltracentrifugation. Dans ce dernier cas, le surnageant est transféré dans un tube en plastique non rigide (Rotor Surespin 360) de 36 ml pour ultracentrifugeuse. Une ultracentrifugation est réalisée à l'aide d'un rotor thermo scientific T-1250 pendant 1 heure à 100 000 g à 4 °C (Thermo Scientific Sorvall WX ultra 80). Les microsomes culottés au fond du tube sont repris dans 2 ml de tampon TEG (Tableau 12) afin d'être homogénéisés à l'aide d'un potter. Les microsomes sont ensuite aliquotés et stockés à -80 °C.

TEK		**Tampon d'extraction**	
Tris HCL pH 7,5	50 mM	Tris HCL pH 7,5	50 mM
EDTA	1 mM	EDTA	1 mM
KCL	100mM	Sorbitol	0,6 M
TEG		BSA*	1 % (p/v)
Tris HCL pH 7,5	50 mM	B mercapto éthanol*	20mM
EDTA	1 mM	* produits ajoutés extemporanément	
Glycérol	30 % (v/v)		

Tableau 12 : Composition des tampons pour la préparation de microsomes de levures.

IV.3 Système d'expression hétérologue dans les plantes

IV.3.1 Expression constitutive

IV.3.1.1 Obtention des plasmides recombinants pBin-X

La recombinaison se réalise entre un plasmide d'entrée de type Gateway® (pCR8/GW/TOPO) et un vecteur de destination Gateway® (pBIN/GW). Des cellules électrocompétentes (E. coli GeneHogs) sont ensuite utilisées pour amplifier les plasmides recombinants. Seules les bactéries ayant incorporé les vecteurs recombinants (qui ne possèdent donc plus le gène ccdB) pourront se développer sur milieu sélectif.

La réaction de LR recombinaison se réalise dans un tube de 1,5 ml à température ambiante suivant les recommandations du kit « LR clonase ^{TM} » d'Invitrogène.

Les plasmides recombinants sont vérifiés par digestion ou par PCR (Amorce spécifique) avant d'être intégrés dans les agrobactéries par conjugaison triparentale.

IV.3.1.2 Conjugaison triparentale

La conjugaison triparentale consiste à transférer un vecteur portant l'ADN-T provenant d'une souche *E. Coli* MC1061, dans une autre souche *Agrobacterium tumefaciens.* Ce transfert est

facilité par des protéines synthétisées par un autre plasmide, le plasmide helper pRK2013 exprimé dans *E. Coli* HB101. Ce type de conjugaison est nommé la conjugaison triparentale.
Les *E.coli* MC1061 sont rendues compétentes et sont transformées avec le plasmide pBin-GW-X contenant le gène d'intérêt. Une colonie transformée est cultivée à 37 °C pendant une nuit dans 2 ml de milieu LB additionnés de kanamycine. Une colonie E. *coli* HB101 contenant le plasmide pRK2013 est cultivée à 37 °C pendant une nuit dans 2 ml de milieu LB additionnés de kanamycine. Enfin, une colonie d'*Agrobacterium tumefaciens* C58C1 Rif ^{R} (pGV2260) est cultivée à 28 °C pendant une nuit dans 2 ml de milieu YEB additionnés d'ampicilline et de rifampicine. Toutes les souches sont incubées sous agitation orbitale durant la nuit. Un mélange de 100 µl de chacune des cultures est effectué le lendemain. Le mélange est ensuite étalé sur un milieu YEB Mg^{2+}. L'incubation se prolonge une nuit à 28 °C afin que la conjugaison triparentale se réalise. Les bactéries transformées sont sélectionnées sur milieu YEB additionnées d'antibiotiques appropriés. Seules les souches recombinantes d'*Agrobacterium tumefaciens*, ayant incorporé le plasmide pGV2260 et le plasmide pBin-GW-D4, sont sélectionnées. L'incubation se fait pendant 48H à 28 °C, puis les colonies transformées sont repiquées sur milieu nutritif.

IV.3.1.3 Inoculum d'agrobactérie utilisé pour la transformation génétique de plantes

Les agrobactéries C58C1Rif^{R} transformées (pGV2260, pBin-GW-D) sont cultivées pendant 3 jours à 28 °C dans du milieu YEB solide supplémenté des antibiotiques de sélections. Une colonie est ensuite prélevée et cultivée dans 200 ml de YEB liquide sélectif pendant 36 heures sous une agitation à 200 rpm. De l'acétosyringone (concentration finale 125 µM) est rajouté au milieu. Ceci permet d'activer les gènes de virulence de l'agrobactérie. Lorsque la DO_{500} du milieu est comprise entre 0,6 et 0,8, les agrobactéries sont centrifugées à 5000 g pendant 20 min. Le culot est ensuite resuspendu dans 20 ml de milieu liquide YEB. Avant d'être utilisables pour transformer les explants végétaux, 3 lavages des agrobactéries sont nécessaires avec de l'eau stérile ultrapure afin d'éliminer les antibiotiques.

IV.3.1.4 Transformation génétique de *Ruta graveolens*

IV.3.1.4.1 Germination des graines de *Ruta graveolens L*

La scarification des graines de Rue se fait pendant 10 min dans de l'acide sulfurique à 95 %. Puis une immersion des graines dans de l'hypochlorite de calcium à 7 % (m/v) pendant 5 min permet de les stériliser. Les graines sont ensuite lavées 3 fois dans de l'eau ultrapure stérile avant d'être déposées de manière stérile sur milieu MS solide (Photo 10). Une trentaine de graines par boite sont mises à germer.

Photo 10 : Germination de graine de *Ruta graveolens.*

IV.3.1.4.2 Transformation d'hypocotyles de *Ruta graveolens*

Les hypocotyles de Rue utilisés pour la transformation génétique sont âgés de 2 à 3 semaines. Le taux de transformation des plantes étant extrêmement faible (1 %), il est nécessaire de réaliser la transformation sur une centaine d'hypocotyles pour une construction génétique. Les explants sont prélevés de manière stérile et sont incubés avec l'inoculum bactérien pendant 30 min. L'excès de bactéries est éliminé par adsorption sur papier filtre. Les hypocotyles sont transférés sur un milieu de co-culture [MS(30), BAP(0,1), acétosyringone (50mg/L)]. 3 jours de co-culture à l'obscurité sont nécessaires avant d'éliminer les agrobactéries par 2 lavages successifs. Les explants sont immergés une première fois dans de l'eau ultrapure stérile pendant 15 min, puis une seconde fois dans une solution de MS(30) BAP(0,1) additionnée de

céfotaxime à 500 mg/l. Les explants sont séchés sur du papier filtre stérile et placés sur un milieu de régénération [MS(30), BAP(0,1), céfotaxime (300mg/l)] pendant une semaine. Puis, ils sont transférés sur un milieu sélectif contenant un antibiotique [MS(30), BAP(0,1), kanamycine (75mg/l) céfotaxime (300mg/l)] afin de sélectionner les cellules transformées (Photo 11).

Photo 11 : Transformation d'hypocotyles de *Ruta graveolens.*

Après un mois de culture sur milieu sélectif, certaines cellules transformées donnent naissance à des tiges feuillées de quelques centimètres. Elles sont alors repiquées dans des pots de 200 ml contenant le même milieu sélectif MS(30), BAP(0,1), kanamycine (75mg/l) pour maintenir cette pression de sélection (Photo 12).

Photo 12 : Repiquage des plantes transformées.

IV.3.1.5 Transformation d'*Arabidopsis thaliana*

IV.3.1.5.1 Germination des graines

Les graines germent directement dans du terreau sous serre. Les plantules sont repiquées et individualisées lorsqu'elles atteignent 2 cm dans des pots de 15 cm de diamètre. Les plantes sont cultivées pendant 2 à 3 semaines jusqu'à l'apparition de jeunes fleurs.

IV.3.1.5.2 Transformation des inflorescences d'*Arabidopsis thaliana*

La méthode « floral dip » (Clough and Bent, 1998) est utilisée pour transformer *Arabidopsis thaliana.* C'est une méthode simplifiée qui permet de transformer rapidement les inflorescences d'*Arabidopsis thaliana*. La plante entière est immergée dans une solution contenant les agrobactéries transformées. L'inoculum d'*Agrobacterium tumefaciens* utilisé pour la transformation est dilué dans une solution de sucrose à 5 % (w/v) et de silwet L-77 à 0,05 % afin d'obtenir une suspension bactérienne dont la DO_{600} 0,6. Les plantes sont plongées intégralement, tête en bas, dans cette solution pendant 10 secondes, puis l'excédent est séché

sur papier absorbant. Ensuite, les plantes sont cultivées dans les phytotrons S2 jusqu'à l'apparition de fruit (les silics) contenant les graines. Quand celle-ci apparaissent, l'arrosage des plantes est limité afin de faciliter la maturation des graines. Les graines sont collectées dans un sac en papier lorsque les siliques deviennent brunes et bien sèches. Elles sont stockées dans un tube à 4 °C pendant quelques jours pour la stratification. Les graines récoltées sont mises à germer et sélectioner à l'aide d'une solution contenant de la kanamycine qui est vaporisé sur les jeunes plantules. Après quelques jours, seules les graines transformées se développent sur le terreau (Photo 13).

Photo 13 : ***Arabidopsis thaliana.***

IV.3.2 Expression transitoire

IV.3.2.1 Principe de l'expression transitoire

L'expression transitoire consiste à exprimer un gène d'intérêt dans le noyau de la cellule, sans insérer celui-ci à l'intérieur du génome de la cellule. Ce gène, non intégré dans le génome, ne sera pas transmis comme un caractère mendélien à la descendance. La protéine d'intérêt sera exprimée seulement quelques jours.

IV.3.2.2 Préparation de l'inoculum bactérien

Une colonie d'agrobactéries transformées avec le plasmide pBin-GW-X contenant le gène d'intérêt et une colonie transformée avec le plasmide pBin-GW-P19 contenant le gène suppresseur viral P19 sont repiquées dans 2 tubes contenant chacun 15 ml de milieu YEB. La

culture est incubée à 30 °C pendant 36H. 3 lavages successifs sont ensuite réalisés après centrifugation à 4000 g pendant 10 min. Les culots sont mélangés de manière à obtenir une DO_{600} 0,2 pour les agrobactéries contenant le gène d'intérêt, et une DO_{600} 0,4 pour les agrobactéries contenant le gène codant pour la protéine P19. La solution est homogénéisée et servira à l'infiltration des feuilles de tabac. De même, une solution d'agrobactérie contenant le plasmide pBin-GW vide est préparé afin de constituer un contrôle négatif.

IV.3.2.3 Agroinfiltration de *Nicotiana benthamiana*

Les feuilles de *Nicotiana benthamiana* (Photo 14) âgées de 1 à 2 mois sont co-infiltrées avec un mélange d'agrobactéries contenant soit le gène d'intérêt, soit le gène codant pour la protéine P19. Le tissu est lésé à l'aide d'une aiguille sur la face inférieure, et la solution d'agrobactérie est inoculée à l'aide d'une seringue directement dans le tissu (Photo 15). 0,2 ml de la solution d'agrobactérie est infiltré dans toute la feuille, puis les plantes sont placées dans le phytotron pendant 3 jours.

Photo 14 : ***Nicotiana benthamiana.***

Photo 15 : Expression transitoire chez *Nicotiana benthamiana.*

V Méthode d'analyse biochimique

V.1 Dosage de protéines

Les protéines purifiées surexprimées par *E.coli* BL21(DE3) sont dosées par fluorométrie en utilisant le Qubit TM fluorometer (Invitrogen). Cette méthode d'analyse repose sur l'intervention de fluorophores qui deviennent fluorescents lorsqu'ils se lient aux protéines. Cette technique d'analyse est 1000 fois plus sensible que le dosage de Bradford. De très faibles quantités de protéine sont suffisantes pour effectuer un dosage précis et spécifique. Une calibration de l'appareil est effectuée au préalable avec les solutions standards de BSA fournies dans le kit Qubit. Le volume d'échantillon à doser (1 à 20 µl) est complété à 200 µl avec la solution de réaction (Tampon + fluorophore). La lecture se fait à l'aide d'un fluoromètre après 15 min d'incubation à température ambiante.

V.2 Quantification des P450 par spectre CO

Cette méthode a été mise au point par Omura et Satoh (a et b) (1964). La mesure est réalisée sur les microsomes de levures que l'on dilue au préalable 5 fois dans du TEG (Tableau 12). L'addition de dithionite de sodium permet de réduire l'hème des microsomes. L'échantillon est alors séparé dans deux cuves de spectrophotométrie dont l'une sera saturée en CO par bullage

de 30 secondes. La mesure d'absorbance est alors réalisée pour la cuve témoin de même que celle saturée en CO. Le CO réagit avec le fer de l'hème des P450 en provoquant un changement d'état de spin du fer. Ceci entraîne une modification du spectre d'absorption du P450. Il apparaît un pic d'absorption caractéristique à 450 nm. L'apparition d'un pic à 450 nm témoigne de la présence de P450 fonctionnel dans les microsomes.

V.3 Mesure d'activités enzymatiques

V.3.1 Détermination des constantes cinétiques

V.3.1.1 Dioxygénase

Afin de déterminer les paramètres cinétiques, les réactions sont réalisées dans un volume final de 100 µl contenant 1 µg de protéine purifiée, 0.1 mM Tris/HCl pH7.2, 0.2 mM $FeSO_4$, 5 mM αcétoglutarate, 10 mM ascorbate de sodium et le substrat à des concentrations croissantes de 5 à 200 µM. La réaction est conduite à 30 °C avec une faible agitation et stoppée après 3 min d'incubation par l'ajout de 20 µl de NaOH 3M. Le pH est ensuite ajusté à la neutralité par ajout de 10 µl d'acide acétique glacial. La solution est alors centrifugée pendant 15 min à 10 000 g, et le surnageant est prélevé pour être analysé en HPLC. Chaque analyse est répétée 3 fois. Le temps d'incubation et la quantité de protéine purifiée utilisée pour chaque réaction sont adaptés pour que la conversion du substrat n'excède pas 50 %. Les constantes cinétiques sont déterminées à l'aide du logiciel Sigmaplot.

V.3.1.2 P450

Les réactions sont conduites dans un volume total de 200 µl de tampon d'activité phosphate de sodium (NaPi) 0,1 M pH 7 contenant 20 µl de microsomes, 1 mM de NADPH et 1 mM de substrat. L'incubation est menée à 27 °C sous faible agitation et stoppée après 10 min d'incubation par ajout d'acétonitrile/HCl (1 %). Après centrifugation, le surnageant est prélevé pour être analysé en HPLC.

V.3.2 Test et mesure de l'inhibition par les furocoumarines

Les réactions enzymatiques effectuées pour les tests d'inhibitions des dioxygénases sont réalisées avec les mêmes cofacteurs et conduites de la même manière que pour la

détermination des constantes cinétiques décrite ci-dessus. Seul l'inhibiteur pour des concentrations de 50, 100 et 200 µM est rajouté aux milieux réactionnels. Chaque réaction est répétée 3 fois. Les constantes cinétiques apparentes sont déterminées par le logiciel Sigmaplot. La détermination du Ki est faite graphiquement par la représentation dite de Kitz-Wilson.

V.4 Analyse en HPLC

V.4.1 Esters de CoA

Les échantillons sont séparés sur une colonne en phase inverse Cosmosil 5C18-AR-II (4.6 • 150 mm; Nacalai Tesque Inc.). La séparation se fait à l'aide de deux solvants: un solvant A (H2O, acide formique 0,1 %) et un solvant B (méthanol, acide formique 0,1 %). La méthode utilise un gradient linéaire dont la concentration en solvant B passe de 10 à 70 % en 35 min avec un débit de 1ml.min-1. La détection des molécules se fait à 346 nm pour la scopolétine et 330 nm pour l'umbelliférone.

V.4.2 Furocoumarines

Les échantillons sont séparés sur une colonne en phase inverse Purospher 125-4 (Merck, Darmstadt, Allemagne). La méthode utilise un gradient linéaire entre un solvant A (H2O, acide acétique 0,1 %) et un solvant B (méthanol 50 %, acétonitrile 50 %). La concentration du solvant B passe de 0 à 67 % en 25 min. La détection des molécules se fait de manière générale à 300nm.

V.4.3 Extraits coumariniques

Le dosage des composés coumariniques se fait sur une colonne en phase inverse Cosmosil 5C18-AR-II (4.6 • 150 mm; Nacalai Tesque Inc.). La séparation se fait à l'aide de deux solvants: Un solvant A (H2O, acide formique 0,1 %) et un solvant B (méthanol, acide formique 0,1 %). La méthode utilise un gradient de 0-5 min jusqu'a 10 % du solvant B, de 5 à 45 min de 10 % jusqu'à 60 % de solvant B, de 45 à 75 min de 60 % jusqu'à 75 % de solvant B, de 75 à 76 min de 75 % jusqu'à 100 % de solvant B avec un débit de 0,7 $ml.min^{-1}$. Les

détections des molécules se font à 346 nm pour la scopolétine et 330 nm pour les composés furocoumariniques.

V.5 Analyse en spectrométrie de masse

La spectrométrie de masse (MS) est une technique physique d'analyse qui permet avec une grande précision de détecter et d'identifier des molécules d'intérêts par mesure de leur masse. Son principe réside dans la séparation en phase gazeuse de molécules chargées (ions) en fonction de leur rapport masse/charge (m/z).

Notre système HPLC-MS est équipé d'une pompe délivrant un solvant binaire et d'un spectromètre de masse à trappe ionique linéaire (LTQ-MS, Thermo Scientific, San Jose, CA, USA). L'appareil est équipé d'une source d'ionisation par électronébulisation (ESI) en mode positif. L'analyse des données a été effectuée en utilisant le logiciel Xcalibure (version 2.1). Les paramètres de mesure utilisés sont décrits ci-dessous. Le champ électrique de nébulisation est de 5kV et la température du capillaire est de 300 °C. Les paramètres d'analyses sont optimisés à partir d'une solution standard d'isopimpinelline (0,1 g/L) avec une phase mobile [eau + acide acétique 0.1 % / méthanol + acide acétique 0.1 % (90/10)] et un débit de 5 ml/min. Les coumarines et furocoumarines sont analysées en MS/MS. MS^2(163) pour l'umbelliférone et MS^2(193) pour la scopolétine.

V.6 Synthèses des substrats

V.6.1 Formation des esters de CoA

V.6.1.1 Production de la 4- coumarate CoA ligase

La synthèse des esters de CoA se fait de manière enzymatique. Un groupement coenzyme A est fixé en position 4 des différents dérivés d'acides cinnamiques par la 4-coumarate CoA ligase (4CL). La 4CL d'*Arabidopsis thaliana* liguée dans le plasmide pQE30 (fournie par le professeur Douglas) est produite dans des bactéries *E.coli* M15. Une colonie est cultivée dans 100 ml de milieu LB additionné d'ampicilline à 50 µg/ml. La culture est incubée à 37 °C jusqu'à ce que la DO_{600} atteigne 0,7. L'induction de la synthèse de protéines 4CL se fait après addition d'isopropyl-β-D-thiogalactopyranoside (IPTG) à 1 mM pendant 20 h à 30 °C. La

purification de la protéine se fait suivant le protocole décrit dans le kit « Ni-NTA Purification System » de Qiagen (Annexe 8).

V.6.1.2 Réaction de synthèse

La synthèse des différents esters CoA est réalisée dans un volume final de 1 ml à température ambiante à partir des acides suivants : ferulate, *p*-coumarate, cafeate et cinnamate à 0,2 mM. Le milieu réactionnel contient 0,2 mM de coenzyme A, 2,5 mM ATP, 1 mM DTT, 2,5 mM MgCL2 et l'enzyme 4CL purifié (20 mg protein.ml-1).

V.6.2 Suivi cinétique de la réaction chimique

Le suivi de la synthèse enzymatique des différents esters CoA est effectué par spectrophotométrie à 346 nm pour le féruloyle CoA et le caffeoyl CoA, à 333 nm pour le p-coumaroyle CoA et à 311 nm pour le cinnamoyle CoA (Figure 59).

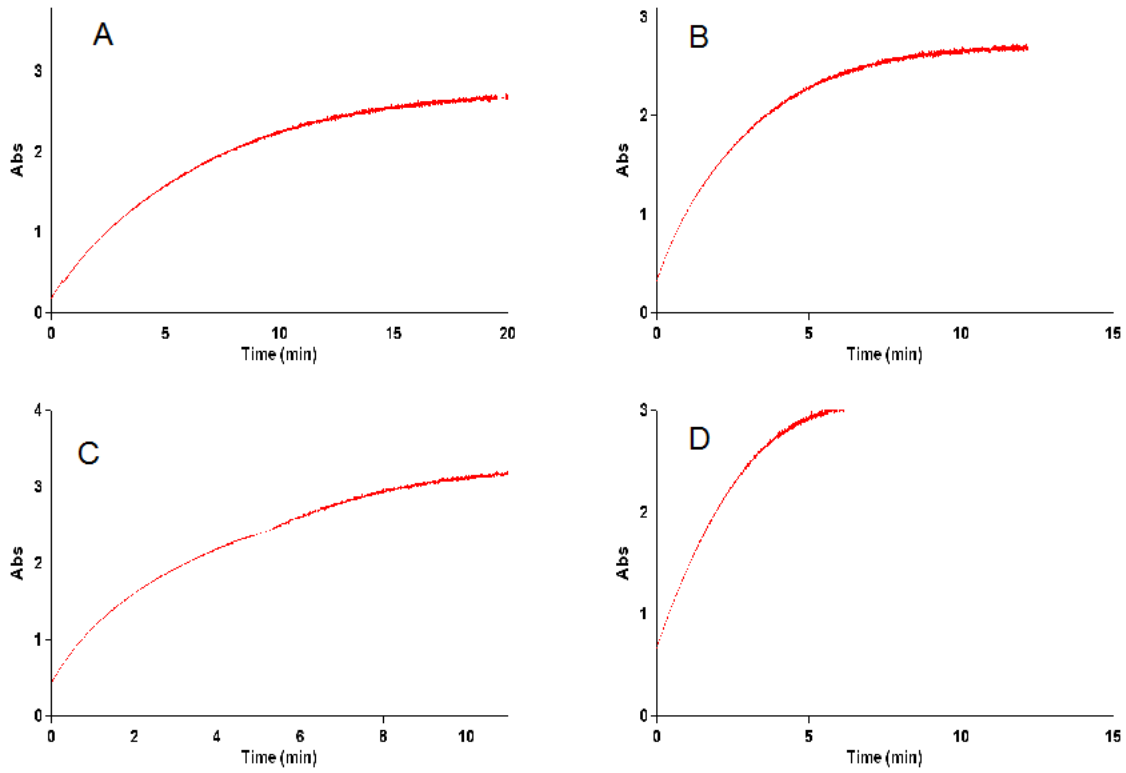


Figure 59 : suivi de la synthèse cinétique du p-coumaroyle CoA (A), féruloyle CoA (B), cinnamoyle CoA (C) et caféoylee CoA (D).

VI Méthode analytique

VI.1 Extraction des composés coumariniques

L'extraction des coumarines est réalisée sur différents tissus de *Ruta graveolens* : les feuilles, les pétioles, les tiges et les racines. L'extraction de l'umbelliférone et de la scopolétine dans les feuilles de *Nicotiana benthamiana* se fait de la même manière que pour la Rue.

200 mg de tissus frais sont prélevés et immédiatement congelés dans de l'azote liquide. Un broyage en fine poudre est effectué dans de l'azote liquide à l'aide d'un pilon et d'un mortier. 2 ml d'éthanol à 80 % sont additionnés et homogénéisés pendant 1 minute à l'aide d'un ultraturax. L'ajout de 50 µl simultanément d'un étalon interne de taxifoline à 2 mg/ml dans le milieu d'extraction est effectué avant le traitement des échantillons. Cet étalon interne de dosage permet de s'affranchir d'une part des pertes inhérentes à la préparation d'un échantillon, et d'une autre part des volumes injectés en HPLC. Une centrifugation de 15 min à 10 000 g permet de sédimenter les débris cellulaires. Le surnageant est entièrement prélevé, et évaporé à l'aide d'un « speedvac », afin de concentrer les extraits coumariniques. Le culot est ensuite resuspendu dans 200 µl de méthanol, et analysé par HPLC et SM.

VI.2 Quantification du niveau d'expression par PCR en temps réel

VI.2.1 Quantification relative

Afin de quantifier le niveau d'expression des dioxygénases chez *Ruta graveolens*, un gène de référence à été au préalable déterminé (ARN 16S) afin de réaliser une quantification relative. La concentration du gène cible est exprimée par rapport à un gène référence, obtenu à partir du même échantillon biologique. Nous avons effectué une quantification relative s'appuyant sur les efficacités respectives des gènes référence et cible, et les Cp obtenus pour les échantillons et le calibrateur, c'est la méthode des ΔΔCt.

La quantification relative fait intervenir un gène de ménage (ARN 16S) qui va servir de normalisateur. Dans notre plante modèle, ce gène est considéré comme stable, et son niveau d'expression ne varie pas en fonction des conditions analysées. Cette technique d'analyse qui met en jeu un gène de contrôle permet de compenser d'éventuels biais de PCR, provenant de variations dans la quantité et la qualité des échantillons, ou de variations liées à de possibles erreurs de pipetage. De plus, cette quantification relative permet de soustraire des variations

liées à des rendements d'extraction différents entre échantillons, ou des variations provenant de l'efficacité de la rétrotranscription des ARN en ADNc.

Les résultats sont exprimés sous forme d'un rapport gène cible/référence de chaque échantillon, normalisé par le rapport cible/référence d'un échantillon appelé calibrateur (ARN 16S). Les résultats sont calculés à l'aide du logiciel « Stepone software».

VI.2.2 Préparation du matériel

Trois extractions totales d'ARNs sont réalisées successivement sur les feuilles, les pétioles, les tiges et les racines de *Ruta graveolens,* suivant le protocole décrit dans le kit « Plant RNeasy mini kit » de Quiagen (Annexe 2).

Les ARNs sont ensuite convertis en ADNc, en utilisant une amorce oligo(dT), afin d'obtenir 1 µg d'ADNc pour chaque organe suivant les instructions du kit « high capacity RNA-to-cDNA master mix » d'Applied Biosystems (Annexe 3). Les ADNc sont ensuite dilués au 1/100ème avant d'être utilisés pour la PCR en temps réel.

VI.2.3 Mélange réactionnel

Des amorces spécifiques de notre gène de référence (ARN 16S) de *Ruta graveolens* sont commandées afin de réaliser une quantification relative (5'-CATTCGGCCCGTCTTGAA-3' forward et 5'-CCGTTGACTCGCACACATGT-3' reverse). Les amorces des gènes cibles diox 1, 2, 4 et 8 sont dessinées à l'aide du logiciel Primer Express 3.0 (Tableau 13). La spécificité de chaque couple d'amorces et les tests de validations de la méthode des ΔΔCt sont vérifiés. Les PCR sur les gènes cibles et références ont toutes deux une efficacité proche. Ceci se traduit expérimentalement par une valeur absolue de la pente pour chaque échantillon ΔCt=f(dilution successive de l'échantillon) inférieur à 0,1.

Le mélange réactionnel comprend 12,5 µl de Power SYBR Green PCR Master mix X2, 0,5 µl de chaque couple d'amorces à une concentration finale de 0,4 µM, 5 µl d'ADNc dilués au 1/100ème et 6,5 µl d'eau ultra pure stérile. La thermocycler utlisé est le Real-Time PCR StepOneplus d'Applied Biosystems.

Les conditions de PCR sont les suivantes :

95 °C pendant 10 min, puis 40 cycles qui débutent par une dénaturation à 95 °C pendant 30 s suivi d'une hybridation à 60 °C pendant 15 s et une dernière étape à 95 °C pendant 15 s. A chaque cycle la température d'hybridation initialement de 60 °C augmente de 0,3 °C.

Le traitement des données est effectué à l'aide du logiciel Stepone software

	Direct	Reverse
Diox 1	CATAAATCCTTGTGTTGGCCATAA	TCGGTGTACAAATCAGACCAGTTT
Diox 2	GGCCAAAAATAAATTGCTTCA	GAAATTGATTCGGTGTCCAGA
Diox 4	CCACCTCTCAGAGCGACTATCA	CCCATGTGAATGACGGAGTCT
Diox 8	TCCAAGGTCATGAAGTTTCCA	TTTCAATACGATCGATTTCAGACT
16 S	CATTCGGCCCGTCTTGAA	CCGTTGACTCGCACACATGT

Tableau 13 : Amorces utilisées pour évaluer l'expression des différentes dioxygénases chez la Rue.

VII Traitement statistique des données

L'analyse statistique des données est réalisée à l'aide du logiciel Excel

Chapitre III : Résultats

I Identification des gènes codant pour les dioxygénases chez les différentes plantes modèles.

La voie de biosynthèse des furocoumarines a été étudiée, au niveau biochimique, dans les années 70, par des travaux de marquage isotopique, réalisés chez différentes plantes telles que l'angélique, le figuier, la Rue ou *Ammi majus* (Brown et al., 1970; Steck and Brown, 1970; Dall'Acqua et al., 1972; Innocenti et al., 1978). Ces travaux ont permis de mettre en évidence l'existence parallèle de deux voies, aboutissant à la synthèse de furocoumarines linéaires et angulaires. Ce sont des molécules proches synthétisées par des réactions similaires. C'est l'étape de prénylation de l'umbelliférone, qui oriente la synthèse vers les formes linéaires ou angulaires. Certaines plantes ne synthétisent que des furocoumarines linéaires telles que *Ficus carica*, *Ammi majus* ou *Ruta graveolens* (Caporale et al., 1981; Milesi et al., 2001; Arab et al., 2008; Orlita et al., 2008; Oliveira et al., 2009). D'autres synthétisent des furocoumarines linéaires et des furocoumarines angulaires, telles que *Pastinaca sativa*, *Psoralea cinerea* ou *Apium graveolens* (Bourgaud et al., 1994; Peroutka et al., 2007). Il est donc primordial, de travailler avec les espèces de plantes dont le profil métabolique est en relation avec les gènes que l'on cherche à identifier, et à caractériser sur le plan fonctionnel. Différentes approches permettent de trouver la séquence codant pour une protéine, et les stratégies employées sont présentées dans le chapitre qui suit.

I.1 Recherche de gènes candidats

Les approches bibliographiques et bio-informatiques sont deux méthodes de recherche complémentaires, qui permettent de sélectionner « *in silico* » des gènes candidats. Une approche expérimentale permet ensuite de confirmer les identifications théoriques réalisées « *in silico* ».

I.1.1 Approche bibliographique

L'analyse bibliographique m'a permis d'émettre des hypothèses quant à la nature des enzymes impliquées dans la voie de biosynthèse des furocoumarines. De nombreuses études par marquages isotopiques, réalisés dans les années 60, ont démontré l'implication de P450 dans la voie de biosynthèse des furocoumarines linéaires et angulaires. La première monooxygénase caractérisée dans la voie de synthèse des furocoumarines est la psoralène synthase (CYP71AJ1) (Larbat et al., 2007), suivie par l'angélicine synthase (CYP71AJ4) (Larbat et al., 2009).

La voie de biosynthèse spécifique des furocoumarines commence à partir de la formation d'umbelliférone (7-hydroxycoumarine). Des premières expériences basées sur le suivi et l'évolution de précurseurs isotopiques, dans différents modèles de plantes, ont fourni des informations concernant la formation d'umbelliférone. Des premiers résultats ont suggéré que la conversion de la coumarine et de l'umbelliférone provenait de l'acide cinnamique chez *Hierochloe odorata* et *Melilotus officinalis* (Brown et al., 1960; Ellis and Amrhein, 1971). De plus, d'autres analyses réalisées sur *Lavandula officinalis,* ont mis en évidence la nécessité du précurseur *para*-hydroxylé de l'acide cinnamique avant d'aboutir à l'*ortho*-hydroxylation de celui-ci (Brown, 1962). L'acide 2-4-dihydroxycinnamic, provenant de l'*ortho*-hydroxylation du *cis*-*p*-coumarate, se transforme ensuite en umbelliférone par une réaction spontanée non enzymatique dépendante. Cette réaction se réalise dans des conditions de pH acide ou neutre entrainant la fermeture du cycle. Le mécanisme réationel est très similaire à celui rencontré dans la synthèse d'un composé coumarinique, la scopolétine. Dans les années 1970, cette étape d'hydroxylation de l'acide *p*-coumarique était supposée être catalysée par un cytochrome P450 (Ellis and Amrhein, 1971; Gestetner and Conn, 1974).

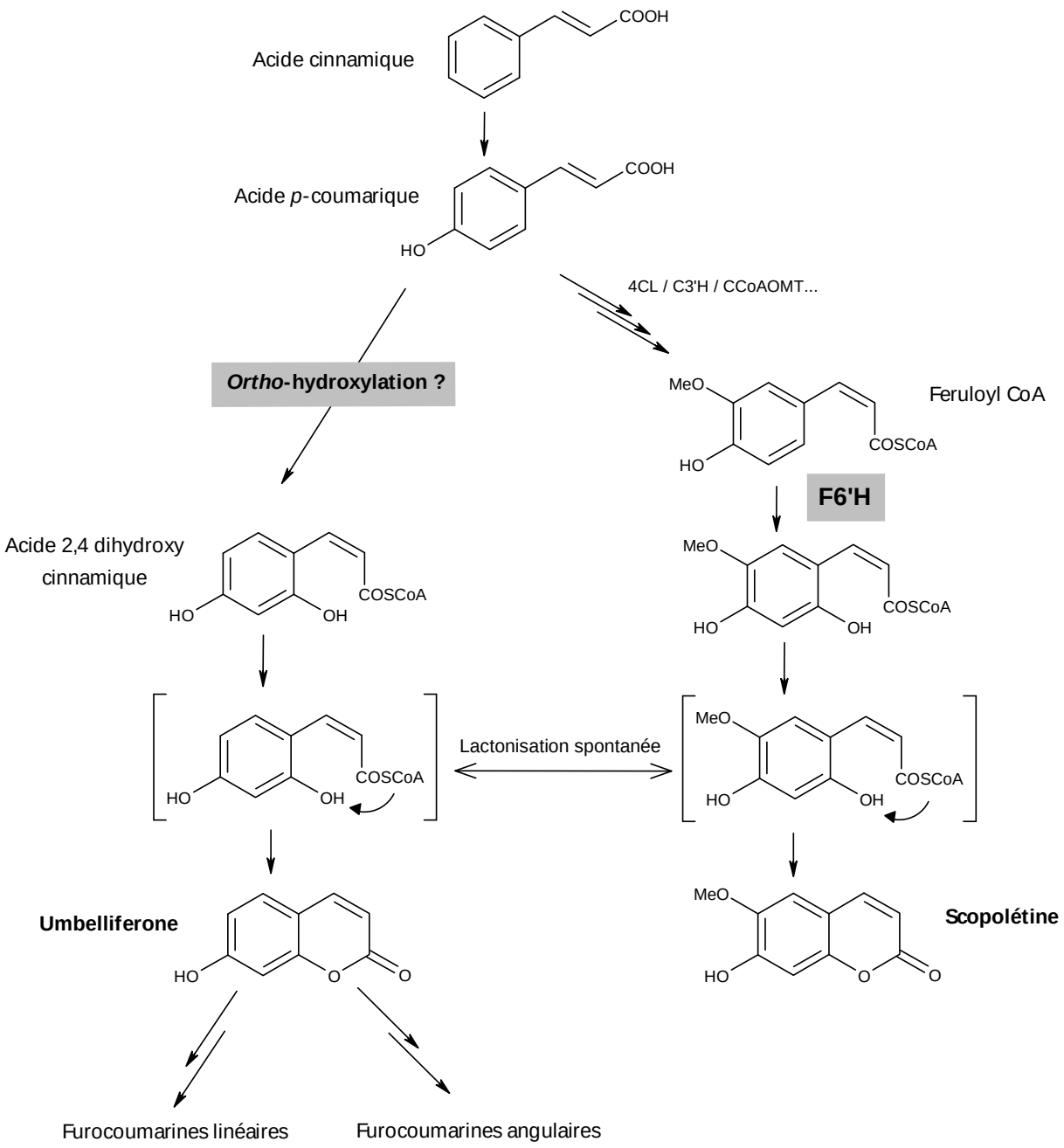


Figure 60 : Comparaison des mécanismes réactionnels mis en jeu entre la synthèse d'umbelliférone et de scopolétine.

L'équipe du professeur Mizutani, qui s'intéresse à la voie de biosynthèse des coumarines, a publié en 2008 deux séquences de gènes qui codent pour une dioxygénase chez *Arabidopsis thaliana*, la féruloyle CoA 6-Hydroxylase (F6'H). Ils ont démontré le rôle de ces deux dioxygénases Fe/α-cétoglutarates-dépendantes (At3g13610 noté F6'H1 et At1g55290 noté

F6'H2), qui interviennent dans l'hydroxylation en position 6 du féruloyle CoA pour former la scopolétine (Kai et al., 2008). Leurs résultats montrent que la F6'H1 est la dioxygénase la plus exprimée chez *Arabidopsis thaliana,* celle-ci sera notée arbitrairement Diox Ara par la suite. Par comparaison des mécanismes réactionnels mis en jeu pour former la scopolétine ou l'umbelliférone (Figure 60), je suggère que la synthèse de l'umbelliférone fasse intervenir une enzyme autre qu'un cytochrome P450 comme décrit auparavant, mais plutôt une enzyme appartenant à la même famille que celle décrite par l'équipe du professeur Mizzutani ; c'est à dire une dioxygénase Fe/α-cétoglutarate-dépendante. Cette découverte fondamentale, dans la voie de biosynthèse de la scopolétine, est à l'origine de l'orientation de mes recherches de gènes candidats vers des enzymes appartenant à la famille des dioxygénases Fe/α-cétoglutarates- dépendantes. Je vais donc dans un premier temps rechercher des gènes appartenant à cette famille d'enzyme dans les différentes banques de données informatiques disponibles.

I.1.2 Approche bio-informatique et recherche d'orthologues des F6'H sur les banques informatiques publiques et privées par tBlastn.

La recherche *in silico* de gènes orthologues, appartenant à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes chez *Ruta graveolens,* s'est révélée être infructueuse en utilisant pour sonde les séquences peptidiques de la F6'H (At3g13610 (F6'H1) et At1g55290 (F6'H2)) d'*Arabidopsis thaliana*. En effet, la recherche de gène effectuée sur la base de données EST de GenBank *via* un tBlastn, n'a pas permis de trouver de séquence fortement homologue chez la Rue, car peu de données nucléotidiques sont à ce jour disponibles concernant cette plante.

Néanmoins, la même analyse *in silico* effectuée sur GenBank a révélé différents gènes candidats appartenant à la famille des Rutacées. Deux fragments de gène (GenBank CX667249 and CX667579) ont été identifiés chez un hybride *Citrus paradisi x Poncirus trifoliata.* Ces deux fragments correspondent aux extrémités d'une même séquence, et codent vraisemblablement pour une enzyme appartenant à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes (Figure 61). Les Rutacées sont une famille de plantes dont de nombreuses espèces sont reconnues pour la synthèse des furocoumarines linéaires (Oliveira et

al., 1996; Girard et al., 2005; Cuca Suàrez and Coy Barrera, 2007). Les citrus regroupent de nombreuses espèces synthétisant des furocoumarines (Barreca et al., 2011; Scordino et al., 2011) et sont des plantes, relativement bien étudiées sur le plan phytochimique et moléculaire, pour lesquelles sont disponibles de nombreuses données bio-informatiques. Les séquences de nombreuses enzymes identifiées chez citrus sont disponibles et utilisables dans les banques de données « Expressed Sequence Tag » (EST) publiques comme Genbank. Ce candidat est donc potentiellement très intéressants.

```
AT3G13610  MAPTLLTTQFSNPAEVTDFVVYKGNGVKGLSETGIKALPEQYIQPLEERLINKFVNETDE 60
cx667249   MAPTNSSTVVTESWDIADFVIHKGHGVKGLSEKGLKSLPKQFHQPLEERFSEKRILD-HV 59
cx667579   ------------------------------------------------------------ 1

AT3G13610  AIPVIDMSNPDEDRVAEAVCDAAEKWGFFQVINHGVPLEVLDDVKAATHKFFNLPVEEKR 120
cx667249   SIPLIDMSKWESPEVAKSICDAAENWGFFQIVNHGVPLEVLERVKEATHRFFALPAEEKR 119
cx667579   ------------------------------------------------------------ 1

AT3G13610  KFTKENSLSTTVRFGTSFSPLAEQALEWKDYLSLFFVSEAEAEQFWPDICRNETLEYINK 180
cx667249   NYSKENSPINNVRYGSSFVPHVERALEWKDFLSLFYVSEEETSAFWPPVCKDEMLEYMRS 179
cx667579   ----------------FVPHVERALEWKDFLSLFYVSEEETSAFWPPVCKDEMLEYMRS 43

AT3G13610  SKKMVRRLLEYLGKNLNVKELDETKESLFMGSIRVNLNYYPICPNPDLTVGVGRHSDVSS 240
cx667249   SEVLIKRLMDVLVKGLNVKRIDEIREPMLLGSRRVNLNYYPMCPNPELTVGVGRHSDIST 239
cx667579   SEVLIKRLMDVLVKGLNVKRIDEIREPMLLGSRRVNLNYYPMCPNPELTVGVGRHSDIST 103

AT3G13610  LTILLQDQIGGLHVRSLASGNWVHVPPVAGSFVINIGDAMQIMSNGLYKSVEHRVLANGY 300
cx667249   FTILLQDDIGGLHVRKDNGTDWIHVSPISGSLIINIGDALXIMSNGRYKSIEHCVI---- 295
cx667579   FTILLQDDIGGLHVRKDNGTDWIHVAPISGSLIINIGDALQIMSNGRYKSIEHCVIANGS 163

AT3G13610  NNRISVPIFVNPKPESVIGPLPEVIANGEEPIYRDVLYSDYVKYFFRKAHDGKKTVDYAKI 361
cx667249   ------------------------------------------------------------ 295
cx667579   QNRISVPLFVNPKPEAILCPFPEVLANGEKPLYKPVLCADYSRHFYTKAHDGKKTIDFAKI 223
```

Figure 61 : Alignement de la F6'H1 d'Arabidopsis thaliana (AT3G13610) et des deux fragments de gène (GenBank CX667249 and CX667579) identifiés chez un hybride *Citrus paradisi x Poncirus trifoliata.*
Les séquences peptidiques de ces deux fragments représentés en rouge et en bleu correspondent aux extrémités d'une même séquence, et codent vraisemblablement pour une enzyme appartenant à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes.

Parallèlement, trois autres candidats appartenant à la famille des Apiacées ont été identifiés dans notre banque générée au laboratoire à partir de *Pastinaca sativa.* Cette plante synthétise des furocoumarines linéaires, telles que le bergaptène ou l'isopimpinelline, mais aussi des furocoumarines angulaires, telles que l'angélicine ou la sphondine (Stein and Posocco, 1984; Lombaert et al., 2001). Cette banque de séquence a été élaborée à partir d'un pool de cDNA normalisé, et ce, grâce à l'extraction d'ARNm dont la qualité a été au préalable contrôlé.

L'extraction d'ARNm de cette plante s'est réalisée aisément à l'aide des kits d'extractions commerciaux (Qiagen RNeasy). L'intérêt de cette plante vient du fait qu'elle exprime les gènes de la voie de biosynthèse des furocoumarines linéaires, mais aussi angulaires. Les trois candidats ont été identifiés dans cette banque de données en utilisant pour sonde les séquences peptidiques de la F6'H (F6'H1 et F6'H2).

```
F6'H1          -----------------------MAPTLLTTQFSNP--------------AEVTDFVVYK 23
F6'H2          -----------------------.NQ.-.AA..LTR--------------DQ..N...HE 22
Diox Panais    ----------------YSPLSLHLLKY..ASETMA.SPT-----------DDYRN.A.T. 33
Diox Panais2   GKRINXTMAFLXLPPLRIHNTQIPHHK.ISPHH.Y.SISNMSIADQTNSYDRLSEIKQFD 60
Diox Panais3   --RYQKMVVTS--------STMEA.VP.I.---.Y.--------------DRRRELQAFD 33

F6'H1          GN--GVKGLSETGIKALPEQYIQPLEERLINK-FVN-ETDEAIPVIDMSNP-----DEDR 74
F6'H2          ..--...........V..D.....F......F-H.KEDS..S.....I..L-----..KS 74
Diox Panais    .H--......DLKLD.........V....D---MTKVLKK.S........L-----.DPK 83
Diox Panais2   ESKL.....VDS.LTTI.RFFHHSPDTLPAP.PDLKTQPGYK..I..L.------G.RAV 114
Diox Panais3   DTKA.....VDA.VQKV.NIFVH.PDASFDKNSSDKCK-ELS..I..LKDLDKDLASHQE 92

F6'H1          VAEAVCDAAEKWGFFQVINHGVPLEVLDDVKAATHKFFNLPVEEKRKFTKENSLSTTVRF 134
F6'H2          .SK.......E...........SM...ENM.T...R..G.........SR.K....N... 134
Diox Panais    ..DQIAA.........I......I...EN..E..RR..A......I.Y.Q.Q.PTNS..L 143
Diox Panais2   .I.Q.RR.SSTL....I......VSKINGLIS.IKA.NEQ.M.I.SQYYHR-EIGLGAS. 173
Diox Panais3   IVDK.RC.S.TF.....V...I.VN..EEMLHGV.R..DQDTQV.KQ.YTR-DHANK.VY 151

F6'H1          GTSFSPLAEQALEWKDYLSLFFVSEAEAEQFWPDICRNETLEYINKSKKMVRRLLEYLGK 194
F6'H2          ......H..K...................S.L...S..S.....M.ET.PL.KK..RF..E 194
Diox Panais    T...L.KVDKV........-IL..DKKSSE...ST.K.DVK..VE..EFVMKW..KA.M. 202
Diox Panais2   Q.N.DLFQSK.AS.R.T.QVRLAPMPPDLNYV.EV....VVAWGEEV..VGDE..GL.SE 233
Diox Panais3   NSN.DLFSSP.AN.R.TFFC.MAPTPPRLEEL.AP..DIQM..SKQVM.LGVT.F.L.SE 211

F6'H1          NLNVKELDETKESLFMGSIRVNLNYYPICPNPDLTVGVGRHSDVSSLTILLQDQIGGLHV 254
F6'H2          ........K....F....T.I...........E.....................E...... 254
Diox Panais    G...-DM.-S...IL...T.I......V....E.AI.........T..F....NV....I 260
Diox Panais2   G.GLEKNKLKDLACLD.-RVMVNH...Y..Q......ITS.T.PGVV.V.V.NEV...Q. 292
Diox Panais3   A.GLEKNHLRDMECDK.-LVFACHC..P..Q.E..L.ATK.T.DGF..V.V.......QI 270

F6'H1          RSLASGNWVHVPPVAGSFVINIGDAMQIMSNGLYKSVEHRVLAN-GYNNRISVPIFVNP- 312
F6'H2          ...TT.R......IS..L..............R...........-.SY.........S.- 312
Diox Panais    .KMDTDT.IF....E.AI.......L..L...K...A..C.A..-.SKD.......T..- 318
Diox Panais2   KYG-EDD..D.E..P.AI......IL.....DE...........PLQEA.V..AV.F..- 350
Diox Panais3   LH--QDH.ID.L.SP.ALIV....LL.LIT.DRF.......V..-HVGP.V..AC.FSTS 327

F6'H1          --KPESVIGPLPEVIANGEEPIYRDVLYSDYVKYFFRKAHDGKKTVDYAKI 361
F6'H2          --.........L...E...K.V.K.I..T....H...........I.F.N. 361
Diox Panais    --S.DDI.......LK...K...KH........H.Y..S....H.L.F.QM 367
Diox Panais2   -GDR.NIY..I..IVSSEKP..F.NFPLT..L.R..T.EL...SLIN.YRL 400
Diox Panais3   LMPSTKLY..MK.LVSEDNP.K..EITVQ.F.S.SIS.GL..ESRLLHFRL 378
```

Figure 62 : Alignement des trois gènes candidats (Diox Panais, Diox Panais2 et Diox Panais3) identifiés dans notre banque d'EST privée de *Pastinaca sativa* avec les F6'H isolée chez *Arabidposis.*

Parmi ces trois candidats (Diox Panais, Diox Panais2 et Diox Panais3), la Diox Panais présente le plus de similarité avec les Diox isolées chez Arabidopsis (F6'H1 et F6'H2) : près de 60 % d'identité pour la Diox Panais, et moins de 30 % d'identité pour les Diox Panais2 et Diox Panais3 (Figure 62). Cette plante a déjà permis d'isoler des P450 impliqués dans la biosynthèse de furocoumarines linéaires et angulaires, et constitue donc un matériel végétal de choix pour l'identification d'autres gènes candidats.

Parmi toutes les séquences ainsi identifiées, j'ai choisi de retenir la Diox Citrus et la Diox Panais. En effet, la comparaison des séquences peptidiques de ces deux gènes candidats révèle prés de 60 % d'identité et 75 % d'homologie avec les deux dioxygénases isolées par Kai et ses collaborateurs (Tableau 14). Ce pourcentage d'identité et d'homologie relativement élevé laisse supposer que ces deux candidats appartiennent tous les deux à la famille des dioxygénases. Les fortes homologies peptidiques (> 75 %), entre les F6'H et les deux gènes candidats issus de *Citrus paradisi x Poncirus trifoliata* et de *Pastinaca sativa*, peuvent indiquer que ces enzymes réalisent une réaction similaire, ou qu'elles soient impliquées dans la synthèse de dérivés métaboliques proches, tels que les dérivées cinnamiques. À titre d'exemple, j'ai réalisé un alignement de séquence entre trois dioxygénases toutes impliquées dans la voie de biosynthèses des dérivées de flavonoïdes (la flavone synthase I, la flavavone 3-hydroxylase et la flavonol 6-hydroxylase), et une autre dioxygénase impliquée dans la voie de biosynthèse des gibbérellines (la gibbéréline 2-oxidase). J'ai pu constater que les trois dioxygénases qui ont un rôle dans la synthèse des flavonoïdes présentent entre elles un pourcentage d'identité situé entre 45 % et 70 %, alors qu'elles ne présentent que 20 % d'identité avec la dioxygénase impliquée dans le métabolisme des gibbérellines (Tableau 15). Je fais donc l'hypothèse, de par cette observation, que les dioxygénases sont susceptibles d'être impliquées dans la même voie de biosynthèse lorsqu'elles présentent un pourcentage d'identité supérieur à 45 %. Pour cette raison, je ne m'intéresse donc pas à Diox Panais2 et à Diox Panais3, qui présente moins de 30 % d'identité avec les F6'H d'*Arabidopsis thaliana*.

L'alignement des séquences codantes pour la Diox Citrus (EST) et la Diox Panais (EST) (effectué avec le logiciel ClustalX) indique la présence des sites consensus de liaisons spécifiques des dioxygénases Fe/α-cétoglutarates-dépendantes, et des sites de reconnaissance de l'α-cétoglutarate et du Fe^{2+} (Figure 63).

	F6'H1	F6'H2	Diox Citrus	Diox Panais
F6'H1		77	60	58
F6'H2	89		59	58
Diox Citrus (EST)	79	75		57
Diox Panais (EST)	76	75	75	

Tableau 14 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différentes dioxygénases isolées chez *Arabidopsis thaliana* (F6'H1 et F6'H2) par Kai et ses collaborateurs (Kai et al., 2008), et les candidats identifiés dans les banques d'EST publiques [Diox Citrus (EST)] et privées [Diox Panais (EST)].

	FNSI	F3H	Flav6H	GA2ox
FNSI		71	44	21
F3H	83		46	21
Flav6H	57	57		15
GA2àx	40	39	29	

Tableau 15 : Comparaison des séquences peptidiques entre 4 dioxygénases Fe/α-cétoglutarates-dépendantes.

Trois d'entre elles sont impliquées dans la voie de biosynthèse des flavonoïdes (FNSI, F3H et la Flav6H) et la dernière est impliquée dans le métabolisme des gibbérellines. La FNSI est la flavone synthase identifiée chez *Petroselinum crispum* (Q7XZQ8); la F3H correspond à la flavanone 3-hydroxylase identifiée chez *Arabidopsis thaliana* (NM_114983); la Flav6H est la flavonol 6-hydroxylase identifiée chez *Chrysosplenium americanum* (AY605048) ; la GA2ox est une gibbérelline oxydase identifiée chez *Arabidopsis thaliana* (NM_106491). Le pourcentage d'identité et d'homologie entre les différentes dioxygénases est représenté respectivement en vert et en bleu.

J'ai compilé dans un alignement de séquence les différents gènes candidats potentiellement intéressante, avec des gènes codant pour des enzymes appartenant à la famille des dioxygénases alpha-cétoglutarates-dépendantes dont les fonctions ont été caractérisées. L'arbre phylogénétique résultant montre que les différentes dioxygénases se regroupent en cluster en fonction des activités qu'elles catalysent (Figure 64). On constate ainsi que les séquences Diox Panais2 (EST) et Diox Panais3 (EST) ne sont pas localisées au même endroit que les séquences de Diox citrus (EST) et Diox Panais (EST). De même, on peut constater que la Diox citrus (EST) et la Diox Panais (EST) se regroupent avec les F6'H caractérisées chez *Arabidopsis thaliana* (Kai et al., 2008). Ce résultat renforce l'hypothèse selon laquelle les dioxygénases, issues de *Citrus paradisi x Poncirus trifoliata* (Diox citrus EST) et de *Pastinaca sativa* (Diox Panais), identifiées par bio-informatique seraient potentiellement impliquées dans le métabolisme des dérivés d'acides cinnamiques.

```
F6'H1              MAPTLLTTQFSNPAEVTDFVVYKGNGVKGLSETGIKALPEQYIQPLEERLINKFVN-ETD 59
F6'H2              .NQ.-.AA..LTRDQ..N...HE.............V..D.....F......FH.KEDS. 59
Diox Citrus (EST)  ....NSS.VVTESWDIA...IH..H.......K.L.S..K.FH......FSE.RIL--DH 58
Diox Panais (EST)  ...S-------PTDDYRN.A.T..H......DLKLD.........V....DMTK.L--KK 51

F6'H1              EAIPVIDMSNPDEDRVAEAVCDAAEKWGFFQVINHGVPLEVLDDVKAATHKFFNLPVEEK 119
F6'H2              .S.....I..L..KS.SK.......E...........SM...ENM.T...R..G...... 119
Diox Citrus (EST)  VS..L....KWESPE..KSI.....N.....IV.........ER..E...R..A..A... 118
Diox Panais (EST)  .S........L.DPK..DQIAA.........I......I...EN..E..RR..A...... 111

F6'H1              RKFTKENSLSTTVRFGTSFSPLAEQALEWKDYLSLFFVSEAEAEQFWPDICRNETLEYIN 179
F6'H2              ...SR.K....N.........H..K...................S.L...S..S.....M. 179
Diox Citrus (EST)  .NYS....PINN..Y.S..V.HV.R......F....Y...E.TSA...PV.KD.M...MR 178
Diox Panais (EST)  I.Y.Q.Q.PTNS..LT...L.KVDKV........-IL..DKKSSE...ST.K.DVK..VE 170

F6'H1              KSKKMVRRLLEYLGKNLNVKELDETKESLFMGSIRVNLNYYPICPNPDLTVGVGRHSDVS 239
F6'H2              ET.PL.KK..RF..E........K....F....T.I...........E.............. 239
Diox Citrus (EST)  S.EVLIK..MDV.V.G....RI..IR.PMLL..R........M....E.............I. 238
Diox Panais (EST)  ..EFVMKW..KA.M.G...D--MDS...IL...T.I......V....E.AI.......... 228

F6'H1              SLTILLQDQIGGLHVRSLASGNWVHVPPVAGSFVINIGDAMQIMSNGLYKSVEHRVLANG 299
F6'H2              ........E.........TT.R......IS..L..............R............. 299
Diox Citrus (EST)  TF......D.......KDNGTD.I..A.IS..LI......L......R...I..C.I... 298
Diox Panais (EST)  T..F....NV....I.KMDTDT.IF....E.AI.......L..L...K...A..C.A... 288

F6'H1              YNNRISVPIFVNPKPESVIGPLPEVIANGEEPIYRDVLYSDYVKYFFRKAHDGKKTVDYA 359
F6'H2              SY.........S..........L...E...K.V.K.I..T....H...........I.F. 359
Diox Citrus (EST)  SQ......L.......AILC.F...L....K.L.KP..CA..SRH.YT........I.F. 358
Diox Panais (EST)  SKD.......T..S.DDI.......LK...K...KH........H.Y..S....H.L.F. 348

F6'H1              KI--- 361
F6'H2              N.--- 361
Diox Citrus (EST)  ..SDI 363
Diox Panais (EST)  QM--- 350
```

Figure 63 : Alignement des séquences peptidiques des dioxygénases isolées chez *Arabidopsis thaliana* (F6'H1 et F6'H2) avec les séquences identifiées par tBlastn sur les banques publiques et privées.

Diox citrus (EST) est issue de l'hybride *Citrus paradisi x Poncirus trifoliata* et Diox Panais est issue de *Pastinaca sativa*. Les sites consensus de liaisons spécifiques des dioxygénases Fe/α-cétoglutarates-dépendantes sont encadrés en bleu et en vert pour les sites de reconnaissance de l'α-cétoglutarate et du Fe^{2+} respectivement.

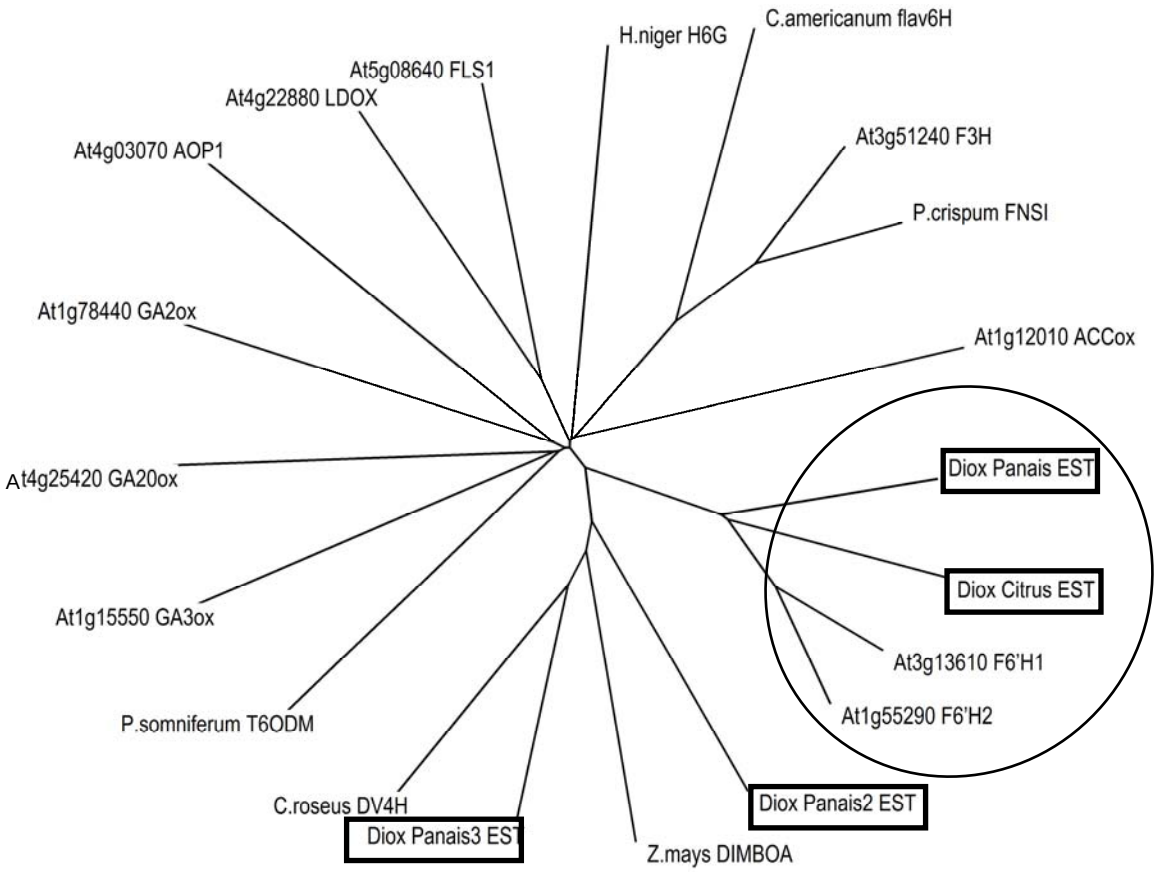


Figure 64 : Construction d'un arbre phylogénétique à partir de différentes dioxygénases cétoglutarates-dépendantes caractérisées.

La Diox Citrus EST et les Diox Panais EST, Diox Panais2 EST et Diox Panais3 EST sont encadrées dans cet arbre phylogénétique. Les numéros d'accessions sur les bases de données GenBank/EMBL/DDBJ des dioxygénases alignées sont les suivantes : At1g78440_GA2ox (NM_106491); At1g15550_GA3ox (NM_101424); At4g25420_GA20ox (NM_118674); At1g12010_ACCox (NM_101073); At3g13610_F6'H1 (NM_112207); At1g55290_F6'H2 (NM_104404); At3g51240_F3H (NM_114983); At4g22880_LDOX (NM_118417); At5g08640_FLS1 (NM_120951); At4g03070_AOP1 (AF417855); Zea mays DIMBOA 7-hydroxylase (NM_001111630) (Frey et al. 2003); P.crispum_FNSI (Q7XZQ8) ; Catharanthus roseus desacetoxyvindoline-4-hydroxylase (AF008597) (Vazquez-Flota et al. 1997); Chrysosplenium americanum flavonol 6-hydroxylase (AY605048) (Anzelotti and Ibrahim 2004); Papaver somniferum thebaine 6-O-demethylase (GQ500139) (Hagel and Facchini 2010) and Hyoscyamus niger hyoscyamine 6b-hydroxylase (HYSH6H) (Matsuda et al. 1991).

L'ensemble de ces données bio-informatiques semble indiquer d'une part, que les gènes candidats identifiés sur les banques publiques et privées (Diox Citrus (EST) et Diox Panais (EST)) appartiennent bien à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes, et

d'autre part, qu'ils pourraient jouer un rôle similaire à celui qui a été identifié chez *Arabidopsis thaliana* (F6'H1 et F6'H2) impliqué dans la synthèse de dérivés cinnamiques, tels que la scopolétine.

I.1.3 Approche expérimentale

Après avoir identifié un certain nombre de séquences candidates potentielles, la seconde étape de mon travail va consister à cloner les gènes correspondants. Pour cela, une approche de PCR a été utilisée.

I.1.3.1 Approche PCR

Cette approche expérimentale consiste à amplifier des gènes dans différentes familles de plantes, en utilisant des amorces spécifiques ou dégénérées dirigées contre un gène en particulier. L'amplification du gène recherché est généralement réalisée sur une matrice d'ARN, après transcription inverse, afin d'isoler directement la séquence codante dépourvue d'introns.

I.1.3.1.1 Diox Citrus

Les séquences nucléotidiques codant pour une dioxygénase identifiée chez l'hybride *Citrus paradisi x Poncirus trifoliata* à partir de l'analyse *in silico*, nous ont permis de générer des amorces spécifiques dirigées contre les extrémités de ce gène. Des sites de restrictions (*Bam*HI et *Sal*I) sont rajoutés aux extrémités 5' et 3', afin de faciliter le clonage dans des vecteurs d'expressions (Tableau 16).

Amorces	Sens	Séquences	Sites de restrictions
Diox Citrus	Dir	5'-CGGATCCATGGCACCAACAAATGGCTCTATCCTGGC--3'	*Bam*HI
	Rev	5'-CGTCGACTCAAAAGTCACCTATCTTTGCAAAG-3'	*Sal*I

Tableau 16 : Amorces spécifiques générées à partir de la séquence identifiée *in silico* chez l'hybride *Citrus paradisi x Poncirus trifoliata*.
Les sites de restrictions sont soulignés sur la séquence des amorces utilisées. FW = Direct, Rev = Réverse.

Le gène identifié dans la banque publique Genbank est amplifié à partir de cDNA préparé à partir d'ARN de feuille de *Citrus limetta*. L'utilisation des primers a permis d'isoler une séquence (identifiant GenBank JN245615), qui n'est pas complètement identique au gène déjà décrit dans la banque de données EST Genbank. En effet, ce candidat JN245615 présente 93 % d'identité nucléotidique et 94 % d'identité peptidique avec la séquence présente dans la banque d'EST issue de l'hybride *Citrus paradisi x Poncirus trifoliata.* Il présente 63 % d'identité nucléotidique (Tableau 18), et 61 % d'identité peptidique (Tableau 14), avec la dioxygénase fortement exprimée et isolée chez *Arabidopsis thaliana* (F6'H1 = Diox Ara).
De plus, les séquences caractéristiques (site spécifique de liaison au Fe^{2+} et à l'α-cétoglutarate), que l'on retrouve chez toutes les dioxygénases Fe/α cétoglutarates-dépendantes, sont aussi présentes sur la séquence isolée chez *Citrus limetta.* Cette séquence, isolée par cette approche expérimentale, semble donc être un bon candidat pour la caractérisation d'un gène orthologue de la F6'H chez *Citrus limetta*.

I.1.3.1.2 Pastinaca sativa

Les séquences nucléotidiques codant probablement pour une dioxygénase, identifiée chez *Pastinaca sativa* à partir de l'analyse *in silico*, m'ont permis de dessiner des amorces spécifiques dirigées contre les extrémités de ce gène, afin de le récupérer dans son intégralité et de le cloner dans un vecteur d'expression. Des sites de restrictions (*Bam*HI et *Xho*I) sont rajoutés aux extrémités 5' et 3' afin de faciliter le clonage par la suite (Tableau 17). L'extraction d'ARN, effectuée sur les feuilles de panais, m'a permis d'amplifier par PCR la séquence identifiée *in silico* (Diox Panais). Le gène récupéré par cette approche PCR est rigoureusement identique à celui identifié dans notre banque privée. Il est donc également dénommé Diox Panais. Il présente 58 % d'identité et 76 % d'homologie avec la F6'H1 (Diox Ara) isolée par kai et ses collaborateurs (Tableau 14). Il présente également les sites de liaisons au Fe^{2+} et à l'α-cétoglutarate, caractéristique de toutes les dioxygénases α-cétoglutarates-dépendantes (Figure 63). Cette séquence isolée par cette approche PCR semble donc être un très bon candidat, et nécessitera d'être caractérisé afin d'identifier un potentiel gène orthologue de la F6'H chez *Pastinaca sativa*.

	Sens	Séquences	Sites de restrictions
Diox Panais	Dir	5'-CGGATTCATGGCTCCATCTCCCACT-3'	*Bam*HI
	Rev	5'-CCTCGAGTCACATCTGTGCGAAATCAAG-3'	*Xho*I

Tableau 17 : Amorces utilisées pour amplifier la séquence codante pour une dioxygénase Fe/α-cétoglutarate-dépendante chez *Pastinaca sativa*.
Cette séquence nucléotidique a été au préalable identifiée *in silico* dans la banque d'EST privée de *Pastinaca sativa.*

I.1.3.2 Criblage de la banque de *Ruta graveolens*

Ruta graveolens est une plante étudiée dans notre laboratoire. La composition en métabolites de *Ruta graveolens* est bien documentée. Elle présente une grande variété et teneur en furocoumarines linéaires, telle que le bergaptène, le psoralène ou la xantothoxine (Kaminski et al., 2003; Orlita et al., 2008), qui sont répartis au niveau des parties aériennes (feuilles et péricarpe du fruit) (Milesi, 2001). De plus, cette plante présente la particularité d'être élicitable *in vitro* en présence de benzothiadiazole (Milesi et al., 2001), et élicitable *in vivo* sous un rayonnement UVB (Larbat, 2006). Cette plante est une bonne candidate de par ses propriétés phytochimiques, mais aussi agronomiques (plante vivace, résistance au froid) pour une production de furocoumarines au champ (Milesi et al., 2001).
Un protocole de transformation génétique a été mis au point par Karine Lièvre durant sa thèse (Lièvre., 2004), qui pourra me permettre de faire de la caractérisation par transformation génétique *in planta*. Romain Larbat a réalisé la préparation d'une banque d'ADNc, à partir de feuilles de Rue élicitées aux UVB, que nous conservons au laboratoire (Larbat, 2006). L'existence de cette banque est un atout pour l'étude des furocoumarines linéaires. En effet, peu de gènes issus de *Ruta graveolens* ont déjà été caractérisés, et les données bio-informatiques disponible dans les banques d'EST publique sont peu nombreuses (85 accessions dans Genbank).

Parallèlement aux travaux réalisés sur citrus et le panais, j'ai donc recherché des gènes orthologues chez cette plante par une approche de criblage de banque, en utilisant comme sonde soit le gène de Diox d'*Arabidopsis thaliana,* soit l'orthologue possible que j'avais déterminé chez *Citrus limetta*.

I.1.3.2.1 Choix des sondes et méthodologie

Le principe du criblage de banque étant basé sur l'appariement des bases complémentaires de la sonde, avec un gène cible présent dans la banque d'ADNc, j'ai opté pour l'utilisation de différentes sondes. Mes premiers essais ont naturellement été effectués en préparant une sonde à partir de la F6'H d'*Arabidopsis Thaliana.* En effet, la F6'H était la seule séquence dont la fonction avait été caractérisée. Toutes mes tentatives de criblage avec cette sonde ne m'ont pas permis d'isoler de nouvelles séquences orthologues chez *Ruta graveolens.* Le gène isolé chez *Citrus limetta* est quelque peu différent de celui identifié chez *Arabidopsis* (63 % d'identité nucléotidique avec la Diox Ara). *Citrus limetta* appartient à la famille des Rutacées comme *Ruta graveolens*. Le biais de codon est donc certainement moins important entre la Rue et le citrus, qu'entre la Rue et *Arabidopsis.* Or, cet usage des codons peut constituer un obstacle à utilisation de cette approche expérimentale, basée sur la complémentarité des bases nucléotidiques. J'ai donc également utilisé le gène de citrus pour réaliser un criblage de la banque de Ruta. Ce nouveau criblage m'a permis d'identifier dans cette banque de *Ruta graveolens,* quatre gènes différents qui appartiennent à la famille des dioxygénases α-cétoglutarates-dépendantes. Ces quatre gènes sont nommés Diox 1, Diox 2, Diox 4 et Diox 8.

I.1.3.2.2 Caractéristiques des gènes isolés

I.1.3.2.2.1 Comparaison avec les gènes déjà isolés

Il est intéressant de noter que les séquences nucléotidiques des quatre dioxygénases isolées chez la Rue sont plus proches du gène de *Citrus* (75 % d'identité avec la Diox Citrus) que du gène d'*Arabidopsis* (62 % d'identité avec la Diox Ara) (Tableau 18). Ceci confirme probablement l'existence d'un biais de codon plus important entre des espèces phylogénétiquements éloignées.

Bien qu'il y ait une grande différence au niveau nucléotidique, les comparaisons de séquences protéiques indiquent prés de 60 % d'identité et plus de 75 % d'homologie avec la F6'H d'*Arabidopsis thaliana* (Diox Ara) (Tableau 19). Ces gènes constituent donc de bons candidats en vue d'une future caractérisation.

Tous les clones isolés chez *Ruta graveolens* présentent les séquences caractéristiques de fixations de fer et de l'α-cétoglutarate chez les dioxygénases Fe/α-cétoglutarates-dépendantes (Figure 65).

	Diox Ara	Diox Citrus	Diox 1	Diox 2	Diox 4	Diox 8
Diox Ara		63	62	61	62	61
Diox Citrus			75	75	77	76
Diox 1				94	86	86
Diox 2					86	86
Diox 4						94
Diox 8						

Tableau 18 : Pourcentage d'identité nucléotidique entre les différentes dioxygénases candidates identifiées.
Les dioxygénases ont été isolées chez *Arabidopsis thaliana* (Diox Ara ; At3g13610), chez *Citrus limetta* (Diox citrus ; JN245615) et chez *Ruta graveolens* (Diox 1/2/4/8)

Une analyse phylogénétique incluant diverses dioxygénases caractérisées, ainsi que les dioxygénases potentielles identifiées jusque-là, montre un regroupement des différentes dioxygénases identifiées chez *Arabidopsisi thaliana*, *Citrus limetta*, *Pastinaca sativa* et *Ruta graveolens* (Figure 65). Ce groupement de gènes amène des éléments en faveur d'une fonction probablement similaire.

	Diox Ara	Diox Citrus	Diox 1	Diox 2	Diox 4	Diox 8
Diox Ara		61	59	58	60	59
Diox Citrus	78		79	79	81	80
Diox 1	76	86		95	93	91
Diox 2	75	86	97		92	91
Diox 4	77	90	96	96		96
Diox 8	77	89	96	95	99	

Tableau 19 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différents clones de dioxygénases candidates identifiées.

Les dioxygénases ont été isolées chez *Arabidopsis thaliana* (Diox ara), *Citrus limetta* (Diox Citrus) et *Ruta graveolens* (Diox 1/2/4/8).

```
Diox 1   MAPTKDFSTT TTNGAESWDD VADFVTKKGH GVKGLSERGI KTLPKPFHQP LEERFSEKKI LERASIPLID  70
Diox 2   ........-. A....D.... .......... ..N....... .......... .......... ..........  69
Diox 4   ......SV-- IHM......E ISE....... .......L.. .....Q.... .......... ..........  68
Diox 8   ......SV-- MAM......E ISE....... .......L.. R....Q.... .......... .D........  68

Diox 1   MSQWDSPEVV KSICDAAENW GFFQIVNHGV PLETLERVKE ATHRFFGLPA EEKNNYSKEN SPINNVRFGS  140
Diox 2   ..E....... .......... .......... .......... ........L. ....K..... ..........  139
Diox 4   ..K....... ........H. .......... .....Q.... ......A... ....K..... ..........  138
Diox 8   ..K....... Q.......H. .......... .........A ......A... D.N.K...D. ..........  138

Diox 1   SFVPHVEKAL EWKDFLSMFY VSEEETNTYW PPICRDEMLE YMRSSEVLIQ RLMEVLVVKG LKVKQIDEIR  210
Diox 2   .......... .......... .......... ......Q... .........K .......... ..........  209
Diox 4   .......... .......... .......... .......... .........K .......... ..........  208
Diox 8   .......... .......... .......... .......... .........K .......... ..........  208

Diox 1   EPMLVGSRRI NLNYYPKCPN PELTLGVGRH SDISTFTILL QDQIGGLHVR KLDDTGNTWV HVTPIAGSLI  280
Diox 2   .........V .......... ........... .. ....... .......... .......... ..........  279
Diox 4   .......... .......... ........... .. ....... ..E....... .......... .....S....  278
Diox 8   .........V .......... ........... .. ....... ..E....... .......... .....R....  278

Diox 1   INIGDALQIM SNGRYKSIEH MVVANGTQDR ISVPLFVNPK PQAILCPFPE VLANGEKPVY KPVFCSDYSR  350
Diox 2   .......... ........... . ......... .. ....... .......... ........L. ...L......  349
Diox 4   .......... ........... . ......... .. ....... .......... .......... ...L......  348
Diox 8   .......... ........... . ......... .. ....... .......... ........L. ...L......  348

Diox 1   HFYTKPHDGK KTVDVAMIN-  369
Diox 2   .......... ....F.L..Y  369
Diox 4   .......... ....F.LM.-  367
Diox 8   .......... ....F.LM.-  367
```

Figure 65 : Alignement des séquences peptidiques des Diox 1/2/4/8 isolées chez *Ruta graveolens*.

Les sites consensus de liaisons spécifiques des dioxygénases Fe/α-cétoglutarates-dépendantes sont encadrés en bleu et en vert pour les sites de reconnaissance du α-cétoglutarate et du Fe^{2+} respectivement.

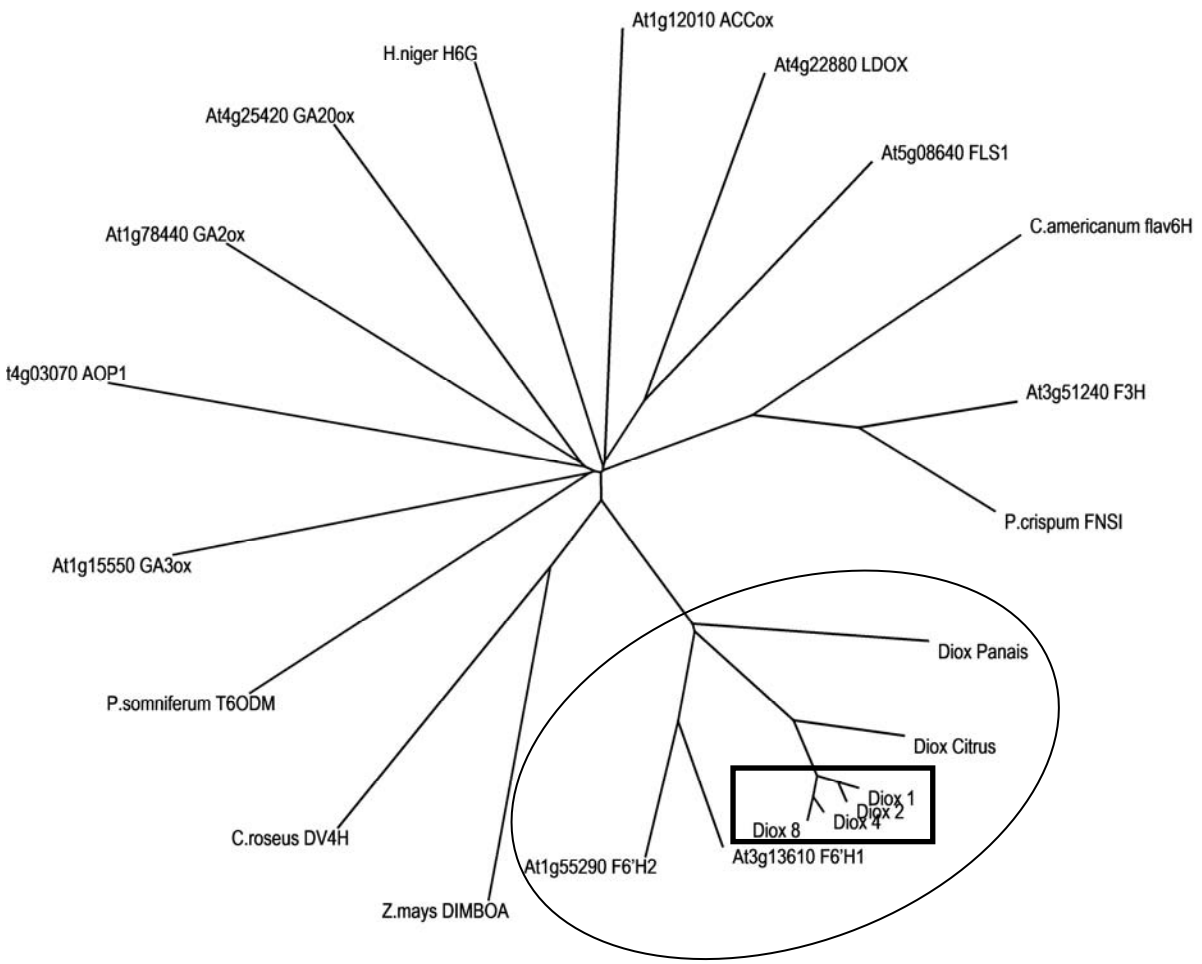


Figure 66 : Construction d'un arbre phylogénétique à partir de différents membres appartenant à la famille des dioxygénases cétoglutarates-dépendantes déjà caractérisées.
En encadré, dans cet arbre, les Diox 1/2/4/8 identifiées chez *Ruta graveolens*. Les numéros d'accessions sur les bases de données GenBank/EMBL/DDBJ des dioxygénases alignées sont décrits dans la (Figure 64).

I.1.3.2.2.2 Multiplicité de dioxygénases chez la Rue : Réalitée ou artefact ?

I.1.3.2.2.2.1 Réalisation d'un second criblage

Afin d'identifier des gènes de Ruta supplémentaires, de nouveaux criblages ont été réalisés. Pour cette nouvelle campagne, j'ai utilisé comme sonde un gène de Citrus fourni par le professeur Mizutani (référence non communiquée). Les résultats obtenus dans le cadre de ce nouveau criblage ne permettent pas d'isoler de nouveaux candidats, mais confirment bien la

présence de mes quatre dioxygénases. En effet, j'ai identifié encore une fois les Diox 1, Diox 2, Diox 4 et Diox 8 sur ce nouveau criblage. Ce résultat suggère que ces quatre dioxygénases sont présentes dans notre banque de Rue dans des proportions équivalentes.

Ces résultats ont été confirmés, à postériorité, par l'analyse d'une banque de données informatique réalisée par un séquençage haut débit (banque 454), sur des ARN extraits de feuilles de Rue traitées aux UV. Néanmoins, la comparaison des différentes séquences de dioxygénases issues du criblage expérimental, avec celles identifiées *in silico* dans notre banque EST 454, m'a permis de mettre en évidence 2 différences sur la Diox 2 (Figure 67), et une différence sur la Diox 1 (Figure 68). Ces différences, certainement liées à l'expérimentation, ne sont pas silencieuses. Or les acides aminés en question semblent être importants pour l'activité des dioxygénases, car ils sont conservés dans toutes les autres potentielles orthologues de la F6'H isolées chez *Arabidopsis thaliana* (Diox Ara), *Pastinaca sativa (*Diox Panais), *Citrus limetta* (Diox Citrus) et *Ruta graveolens* (Diox 1/2/4/8). L'extraction d'ARN de *Ruta graveolens*, et la synthèse de couples d'amorces (Tableau 20) dirigées contre chaque extrémité du gène codant pour la Diox 2, ont permis de récupérer ce gène sans les différences identifiées (Figure 67) chez *Ruta graveolens*. Celui-ci sera noté Diox 2 nvl. La même démarche est employée afin de récupérer la Diox 1 sans l'acide aminé interchangé.

```
Diox 1      ... LQDQIGGLHV RKLDDTGNTW VHVTPIAGSL IINIGDALQI MSNGRYKSIE  299 ...
Diox 2      ... LQDQIEVLHV RKLDDTGNTW VHVTPIAGSL IINIGDALQI MSNGRYKSIE  298 ...
Diox 4      ... LQDEIGGLHV RKLDDTGNTW VHVTPISGSL IINIGDALQI MSNGRYKSIE  297 ...
Diox 8      ... LQDEIGGLHV RKLDDTGNTW VHVTPIRGSL IINIGDALQI MSNGRYKSIE  297 ...
Diox Citrus ... LQDDIGGLHV RK--DNGTDW IHAAPISGSL IINIGDALQI MSNGRNKSIE  291 ...
Diox Panais ... LQDNVGGLHI RKMDTD--TW IFVPPVEGAI VINIGDALQI LSNGKYKSAE  281 ...
Diox Ara    ... LQDQIGGLHV RSLASG--NW VHVPPVAGSF VINIGDAMQI MSNGLYKSVE  292 ...
```

Figure 67 : Identification de deux acides aminés différents sur la Diox 2.

Les gènes codant pour des dioxygénases représentés dans cet alignement ont été isolés de manière expérimentale. Les différences identifiées concernent deux acides aminés situés sur la Diox 2. Cette différence n'est pas présente sur la Diox 2 identifiée dans notre banque informatique privée.

```
Diox 1      ... HMVVANGTQD RISVPLFVNP KPQAILCPFP EVLANGEKPV YKPVFCSDYS  349 ...
Diox 2      ... HMVVANGTQD RISVPLFVNP KPQAILCPFP EVLANGEKPL YKPVLCSDYS  348 ...
Diox 4      ... HMVVANGTQD RISVPLFVNP KPQAILCPFP EVLANGEKPV YKPVLCSDYS  347 ...
Diox 8      ... HMVVANGTQD RISVPLFVNP KPQAILCPFP EVLANGEKPL YKPVLCSDYS  347 ...
Diox Citrus ... HCVIANGSQN RISVPLFVNP KPEAILCPFP EVLANGEKPL YKPVLYADYS  341 ...
Diox Panais ... HCVAANGSKD RISVPIFTNP SPDDIIGPLP EVLKNGEKPI YKHVLYSDYV  331 ...
Diox Ara    ... HRVLANGYNN RISVPIFVNP KPESVIGPLP EVIANGEEPI YRDVLYSDYV  342 ...
```

Figure 68 : Identification d'un acide aminé différent sur la Diox 1

Les gènes codant pour des dioxygénases représentés dans cet alignement ont été isolés de manière expérimentale. Les différences identifiées concernent deux acides aminés situés sur la Diox 2. Cette différence n'est pas présente sur la Diox 1 identifiée dans notre banque informatique privée.

	Sens	Séquences	Sites de restrictions
Diox 1	Dir	5'-CGGATCCATGGCACCAACAAAAGAT-3'	BamH1
	Rev	5'-CCTCGAGCTAATTAATCATAGCAACGT-3'	XhoI
Diox 2	Dir	5'-CGGATCCATGGCACCCACAAAAGAT-3'	BamH1
	Rev	5'-CCTCGAGCTAGTAATTAATCAGAGCA-3'	XhoI

Tableau 20 : Amorces utilisées pour récupérer les clones de la Diox 1 et de la Diox 2 sans mutation chez *Ruta graveolens.*

I.1.3.2.2.2.2 Analyse des séquences génomiques

Afin de confirmer cette multiplicité de gènes codant pour des dioxygénases et pour mieux comprendre, le cas échéant, sur un point de vue évolutif l'émergence de quatre copies semblables d'un même gène chez *Ruta graveolens*, nous avons entrepris la recherche des séquences génomiques des Diox 1, Diox 2, Diox 4 et Diox 8. Les séquences génomiques des différentes dioxygénases ont été récupérées à partir des amorces spécifiques (Tableau 21), dirigées contre les extrémités de chacune d'entre elles, sur de l'ADN extrait de feuilles de Rue à l'aide du kit « DNeasy plant » de Qiagen.

Amorces	Sens	Séquences	Sites de restrictions
Diox 1	Dir	5'-CGGATCCATGGCACCAACAAAAGAT-3'	Bam HI
	Rev	5'-CCTCGAGCTAATTAATCATAGCAACGT-3'	Xho I
Diox 2	Dir	5'-CGGATCCATGGCACCCACAAAAGAT-3'	Bam HI
	Rev	5'-CCTCGAGCTAGTAATTAATCAGAGCA-3'	Xho I
Diox 4	Dir	5'-CGGATCCATGGCACCGACAAAAGA-3'	Bam HI
	Rev	5'-CCTCGAGTTAATTCATCAAAGCGAAGTCGACAGTTTTC-3'	Xho I
Diox 8	Dir	5'-CGGATCCATGGCACCAACAAAAGAC-3'	Bam HI
	Rev	5'-CCTCGAGCTAATTCATCAAAGCGAAG-3'	Xho I
Diox Citrus	Dir	5'-CGGATCCATGGCACCAACAAATGGCTCTATCCTGGC--3'	Bam H1
	Rev	5'-CGTCGACTCAAAAGTCACCTATCTTTGCAAAG-3'	Sal I
Diox Panais	Dir	5'-CGGATTCATGGCTCCATCTCCCACT-3'	Bam HI
	Rev	5'-CCTCGAGTCACATCTGTGCGAAATCAAG-3'	Xho I

Tableau 21 : Amorces utilisées pour amplifier les séquences génomiques des différentes dioxygénases identifiées chez *Ruta graveolens* (Diox 1/2/4/8), chez *Citrus limetta* (Diox Citrus) et chez *Pastinaca sativa* (Diox Panais).

Les séquences amplifiées à l'aide de ces couples d'amorces sont ensuite clonées dans le pCR8/TOPO/TA et sont séquencées. Cette approche nous a permis d'isoler toutes les séquences génomiques des dioxygénases identifiées chez la Rue. J'ai ainsi pu comparer les séquences codantes avec les séquences génomiques. Afin d'avoir un jeu complet, cette même approche a été réalisée sur de l'ADN extrait de feuilles de *Citrus limetta* et *Pastinaca sativa*, à partir d'amorces spécifiques de la Diox citrus et Diox Panais (Tableau 21).

L'alignement des différentes séquences codantes d'ADN génomique par ClustalX a révélé que toutes les dioxygénases identifiées chez *Ruta graveolens* (Diox 1, Diox 2, Diox 4 et Diox 8), chez *Citrus limetta* (Diox Citrus) et chez *Pastinaca sativa* (Diox Panais), présentaient un intron situé au milieu de la séquence codante. Il est intéressant de noter que l'intron est situé exactement au même endroit sur chaque dioxygénase (Figure 69). Pour la Diox Citrus issue de *Citrus limetta* l'intron est localisé entre la lysine (K) et l'acide aspartique (D), pour les Diox issues de *Ruta graveolens* entre l'arginine (R) et l'acide aspartique (D), et entre la lysine (K) et l'asparagine (N) pour la Diox issue du *Pastinaca sativa* (Tableau 22).

```
Diox Citrus ORF      SFVPHVERAL EWKDFLSLFY VSEEETSAFW PPVCK*DEMLE YMRSSEVFIK
Diox 1 ORF           SFVPHVEKAL EWKDFLSMFY VSEEETNTYW PPICR*DEMLE YMRSSEVLIQ
Diox 2 ORF           SFVPHVEKAL EWKDFLSMFY VSEEETNTYW PPICR*DQMLE YMRSSEVLIK
Diox 4 ORF           SFVPHVEKAL EWKDFLSMFY VSEEETNTYW PPICR*DEMLE YMRSSEVLIK
Diox 8 ORF           SFVPHVEKAL EWKDFLSMFY VSEEETNTYW PPICR*DEMLE YMRSSEVLIK
Diox Panais ORF      SFLPKVDKVL EWKDYLSIL- VSDKKSSEFW PSTCK*NDVKE YVEKSEFVMK
```

Figure 69 : Position de l'intron dans les différentes dioxygénases identifiées chez *Citrus limetta*, *Ruta graveolens* et *Pastinaca sativa*.

L'intron est représenté dans cette figure par une étoile. Celle-ci montre la localisation de l'intron qui est situé à chaque fois exactement au même endroit.

Dioxygénases identifiées	Position de l'intron (*)
Diox Citrus	K^{170} * D^{171}
Diox 1	R^{175} * D^{176}
Diox 2	R^{174} * D^{175}
Diox 4	R^{173} * D^{174}
Diox 8	R^{173} * D^{174}
Diox Panais	K^{162} * N^{163}

Tableau 22: Identification des acides aminés flanquant les introns dans chacune des dioxygénases identifiées.

Les alignements des séquences des quatre gènes de dioxygénases, isolées chez *Ruta graveolens*, indiquent que les séquences de ces introns présentent de nombreuses différences, et qui ne peuvent pas provenir d'erreurs liées à l'amplification. Ces résultats apportent donc la preuve que les quatre gènes codant pour les Dioxygénases de Rue sont bel et bien différents. La suite de mes expérimentations permettra de mettre en évidence si les différences d'acides aminés codant les dioxygénases engendrent une différence de fonctionnalité.

L'épissage d'un pré ARNm est dirigé par des éléments de séquence indispensable au recrutement de la machinerie d'épissage. Le dinucléotide GT est situé à la jonction exon-intron et constitue la borne 5' de l'intron. Cette partie est communément dénommée site 5' d'épissage (ou site donneur). Le site 3' d'épissage, situé à la borne 3' de l'intron, est quant à lui composé de trois éléments de séquence : le site de branchement (A), précédé d'une séquence riche en pyrimidine (Y_n) de longueur variable et d'un nucléotide terminal AG faisant partie du site accepteur. Cette conservation des extrémités des introns est appelée la

règle GT-AG. Cependant, les bases nucléotidiques situées dans l'environnement direct des ces dinucléotides jouent également un rôle important (Stephens and Schneider, 1992). Des séquences consensus ont été définies à partir de l'alignement de nombreuses séquences introniques eucaryotes. La représentation effectuée par Cartegni (Cartegni et al., 2002) de ces séquences consensus rend compte de la fréquence de chaque nucléotide aux différentes positions en fonction de la taille du nucléotide (Figure 70).

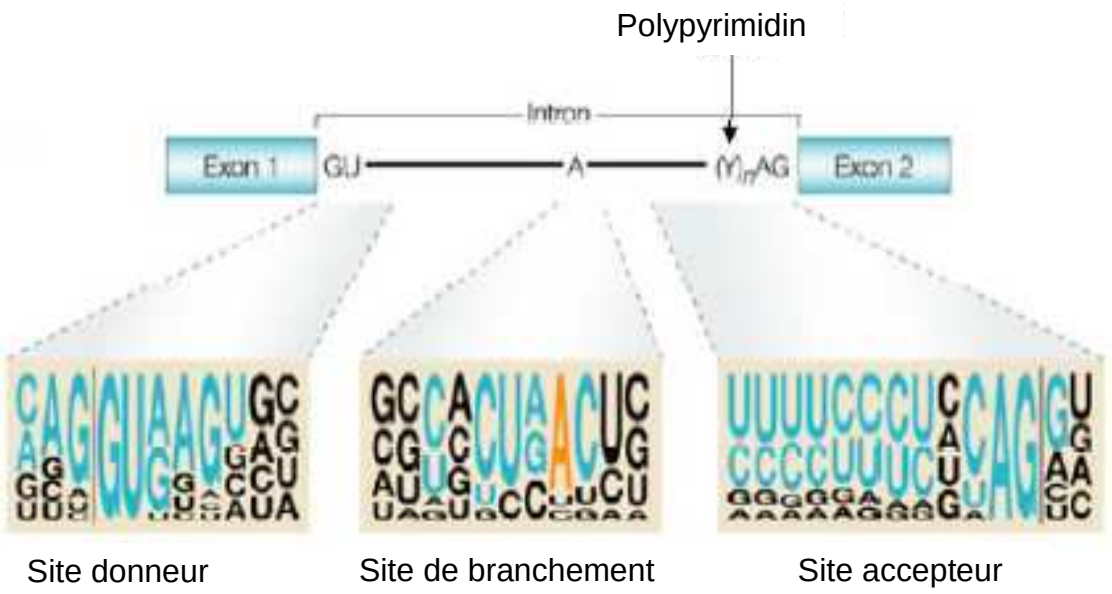


Figure 70 : Séquences consensus des sites d'épissages des introns d'eucaryotes.
Pour chacun des sites, la taille d'un nucléotide est directement proportionnelle à la fréquence de nucléotides obtenue après alignement de nombreuses séquences d'introns d'eucaryotes. Les nucléotides en bleu appartiennent aux séquences consensus classiques. La base A du site de branchement est représentée en orange. Les lignes verticales correspondent aux frontières exon-intron. Cette figure est adaptée de Cartegni (Cartegni et al., 2002).

Une observation fine de l'alignement des séquences génomiques et ORF, des différentes dioxygénases candidats isolées, permet de remarquer que la règle de jonction GT-AG est respectée et figure en rouge dans la Figure 71. Néanmoins, ces introns présentent une particularité. La très grande majorité des introns commencent en 5', juste après un triplet de nucléotide codant, et se termine en 3', juste avant le prochain triplet de nucléotide codant. En ce qui concerne les introns des différentes dioxygénases identifiées, la frontière exon-intron

du site donneur se trouve au milieu du dernier codon (représenté par un trait vertical noir sur la Figure 70), et se termine un nucléotide avant le prochain codon. Le codon situé juste en amont des introns se trouve donc tronqué du nucléotide G (Figure 71) lors de l'excision des introns. Mais du fait que l'intron se termine un nucléotide avant le prochain codon, l'excision de l'intron permet de conserver le même cadre de lecture. Qui plus est, ce nucléotide supplémentaire présent est aussi une guanine dans toutes les dioxygénases (nucléotide vert sur la Figure 70). Le triplet de nucléotide codant reste donc inchangé après excision de l'intron. Les sites donneurs et accepteurs encadrés en vert sur la Figure 70 de chaque intron ont été déterminés à partir des analogies au niveau des séquences consensus déterminées par Cartegni (Cartegni et al., 2002). Le site d'embranchement doit se situer entre ces deux sites. Après identification des différents sites consensus d'épissage, je peux en conclure que le système d'excision de ces introns se déroule de la manière classique, en faisant intervenir un complexe catalytique macromoléculaire nommé « spliceosome ». Celui-ci permet la réalisation des deux étapes de trans-esterification nécessaire au détachement de l'intron et à la liaison entre les deux exons adjacents (Black, 2003).

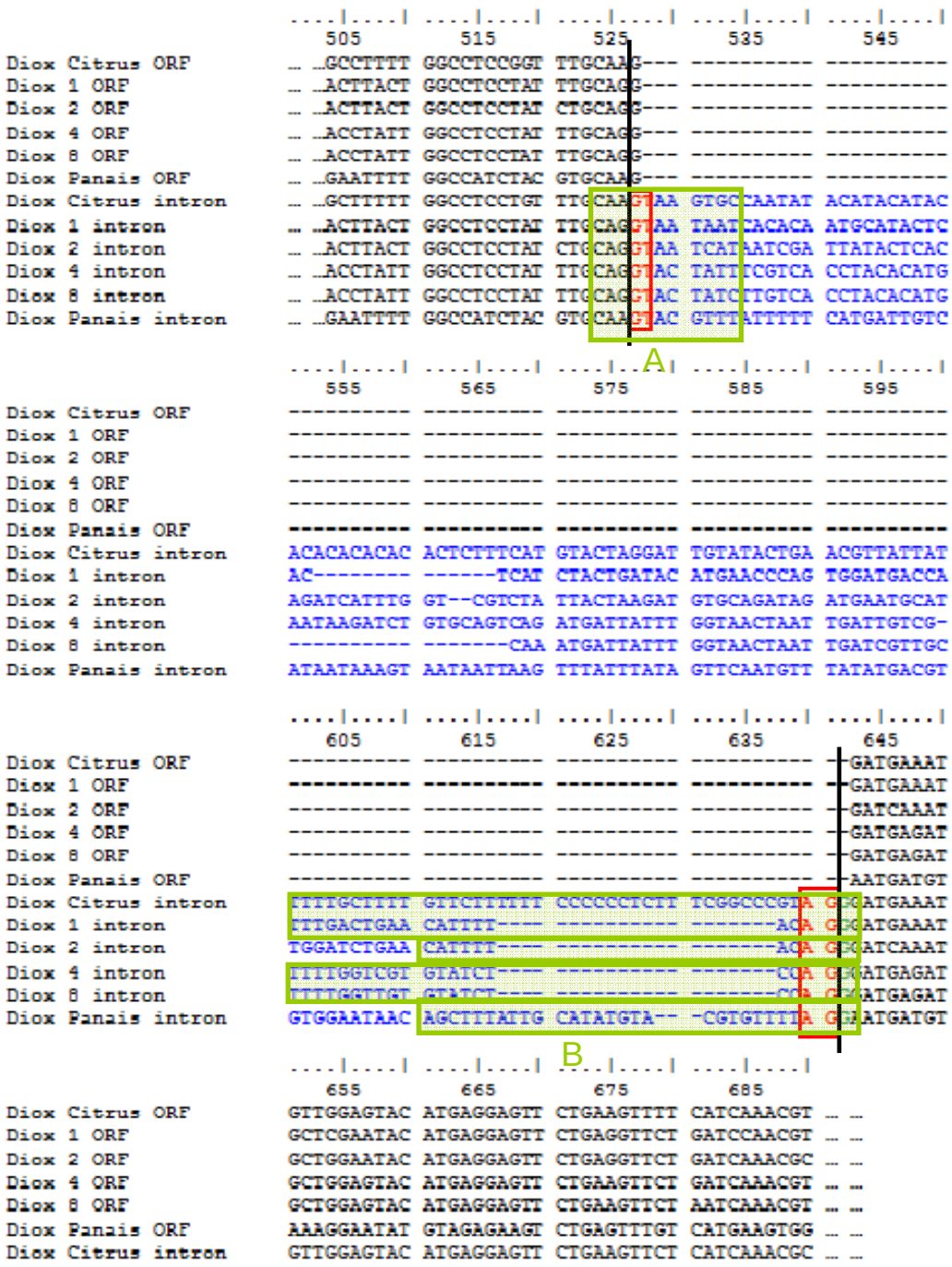


Figure 71 : Alignement des différentes séquences codantes (notées ORF) et génomiques (notées intron) sur ClustalX de toutes les dioxygénases potentielles identifiées chez *Citrus limetta*, *Ruta graveolens* et *Pastinaca sativa*.

La séquence nucléotidique en bleu correspond à l'intron. Les dinucléotides en rouge correspondent aux dinucléotides GT et AG caractéristiques des introns d'eucaryotes faisant respectivement partie du site donneur A et du site accepteur B des introns. Les sites donneurs A et accepteur B encadrés en vert ont été déterminés par analogie avec les sites consensus définis par Cartegni (Cartegni et al., 2002). Les lignes verticales représentent les frontières exon-introns.

I.2 Discussion

L'étape de prospection bibliographique fut très importante afin d'orienter ma recherche de candidats vers des gènes codant pour des dioxygénases Fe/α-cétoglutarates-dépendantes. En effet, il était supposé auparavant que l'umbelliférone résultait d'une hydroxylation en position 6 du *p*-coumarate par une enzyme appartenant à la famille des cytochromes P450, suivie immédiatement d'une lactonisation et d'une isomérisation spontanées [50, 51]. La compréhension de la voie de biosynthèse des furocoumarines a pris une nouvelle dimension avec les travaux réalisés par Kai et ses collaborateurs (Kai et al., 2008). La mise en évidence d'une enzyme impliquée dans la synthèse de coumarines ouvre la perspective de trouver des enzymes ayant des caractéristiques catalytiques similaires, permettant la synthèse d'umbelliférone. En effet, l'analogie du mécanisme réactionnel mis en jeu entre la synthèse de la scopolétine et de l'umbelliférone, ainsi que la nature des substrats et des produits formés dans cette réaction (Figure 60), m'ont conduit à rechercher des gènes orthologues de la F6'H chez des plantes productrices de furocoumarines.

Arabidopsis thaliana ne synthétise pas de furocoumarines. Dans le cadre du projet de recherche développé au laboratoire concernant l'identification de la voie de biosynthèse des furocoumarines, j'ai souhaité trouver des gènes orthologues de la F6'H. Mon premier candidat a été identifié dans les bases de données informatiques. Les séquences disponibles sur Genbank m'ont permis d'identifier deux fragments de gène (GenBank CX667249 and CX667579) provenant d'un hybride *Citrus paradisi x Poncirus trifoliata* correspondant aux extrémités d'une même séquence (Figure 61). Au vu des caractéristiques de la séquence identifiée « in silico » (62 % d'identité avec les F6'H (Tableau 14), présence des sites caractéristiques de liaisons au Fe^{2+} et à l'α-cétoglutarate (Figure 63)), il semblerait que ce gène (Diox Citrus EST) code pour une enzyme appartenant à la famille des dioxygénases Fe/α-cétoglutarate-dépendantes. Les plantes produisant des furocoumarines ne sont malheureusement pas des plantes modèles pour la communauté scientifique, et je n'ai donc pu identifier d'autres séquences intéressantes dans les banques de données publiques.

Le panais est une plante d'intérêt dans notre laboratoire. Cette plante a permis auparavant d'identifier et de caractériser la psoralène synthase (CYP71AJ3) et l'angélicine synthase (CYP71AJ4) (Larbat et al., 2009). Suite à ces travaux, le laboratoire a fait réaliser une banque de cDNA issue de *Pastinaca sativa* et les données informatiques m'ont été mises à disposition. J'ai ainsi pu identifier par tBlastn, trois gènes candidats présents dans cette

banque. Des alignements de séquence (Figure 62) permettront de me focaliser sur l'une d'entre elles. Ce gène (Diox Panais EST) présente aussi toutes les caractéristiques spécifiques des enzymes appartenant à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes (Tableau 14 et Figure 63).

La réalisation d'un arbre phylogénétique (Figure 64), à partir de différentes dioxygénases caractérisées à ce jour et comprenant les deux candidats (Diox Citrus EST et Diox Panais EST) identifiés sur les banques informatiques publiques et privées, indique qu'ils se regroupent avec les F6'H d'*Arabidopsis thaliana* caractérisées par kai (Kai et al., 2008). Ce groupe de gènes amène des éléments en faveur d'une fonction probablement similaire, et potentiellement impliquée dans l'hydroxylation des acides cinnamiques. Ce travail bio-informatique effectué en amont m'a permis, ensuite, de récupérer ces deux gènes candidats par des approches expérimentales de PCR à partir d'ARN extrait de feuilles de *Citrus limetta* et de *Pastinaca sativa*.

Ruta graveolens est une plante étudiée au laboratoire pour laquelle très peu de séquences sont disponibles sur les banques de données informatiques. Afin d'identifier d'autres gènes candidats chez *Ruta graveolens* potentiellement orthologue des F6'H d'*Arabidopsis thaliana,* j'ai réalisé un criblage expérimental d'une banque de cNDA de Rue mise a disposition au laboratoire. J'ai réalisé ce travail en utilisant une sonde codant pour la F6'H *d'Arabidopsis thalaiana* et une autre sonde issue de Citrus. L'intérêt d'identifier au préalable une potentielle dioxygénase chez le citrus s'est révélé primordial. En effet, les différentes tentatives de criblages de banque de Rue, effectuées à partir de la F6'H d'*Arabidopsis thaliana* se sont révélées être des échecs, certainement liés à l'usage de codon différent entre ces deux espèces. L'utilisation d'une sonde issue de *Citrus limetta* (Diox Citrus) appartenant à la même famille que *Ruta graveolens* pour réaliser mon criblage de banque, m'a permis d'identifier quatre autres candidats (Diox 1, Diox 2, Diox 4 et Diox 8) potentiellement orthologues de la F6'H.

Les candidats (Diox 1/2/4/8) identifiés chez *Ruta graveolens* présentent prés de 60 % d'identité avec la F6'H d'*Arabidopsis thaliana* (Tableau 19), et comportent toutes les séquences spécifiques de liaison au Fe^{2+} et à l'α-cétoglutarate (Figure 65). De plus, leur positionnement dans l'arbre phylogénétique, regroupant différents membres de la famille des dioxygénases, montre un rassemblement de ces quatre nouveaux candidats (Diox 1/2/4/8) avec les F6'H d'*Arabidopsis thaliana* ainsi que les Diox citrus et Diox Panais. Ce sous-ensemble pourrait être un groupe de gènes impliqués dans la métabolisation des dérivés de l'acide cinnamique, à l'image des dioxygénases qui se regroupent pour former un ensemble impliqué dans le métabolisme des flavonoïdes (Figure 66).

Je me suis interrogé quant à la profusion de gènes chez *Ruta graveolens* codant pour des dioxygénases dont les séquences peptidiques sont très proches. Je me suis donc demandé si ces quatre clones ne proviendraient pas de problèmes liés à l'expérimentation. Dans un premier temps, j'ai réalisé un second criblage avec une autre sonde de citrus (fourni par le professeur Mizutani) qui m'a permis d'identifier exactement les mêmes séquences. A posteriori, le laboratoire a commandé la réalisation d'une banque (454) de cDNA issue de *Ruta graveolens*. J'ai ainsi pu retrouver dans un second temps, ces quatre séquences dans la banque informatique. Néanmoins, la Diox 1 et la Diox 2 présentent de très faibles différences (un ou deux acides aminés différents). Ces différences peuvent provenir lors de la construction de la banque de cDNA par Romain Larbat (Larbat, 2006). L'étape d'expression et de caractérisation des candidats permettra de déterminer si ces différences ont un impact sur l'activité ou la fonction du gène. Enfin, la recherche des séquences génomiques des Diox 1, Diox 2, Diox 4 et Diox 8 montre la présence d'un seul et unique intron situé exactement au même endroit (Figure 69), mais qui est complètement différent d'un gène à l'autre. L'analyse des séquences des différents introns m'a permis d'identifier les sites de jonction exon-intron respectant la règle de GT-AG, ainsi que les différents sites donneurs et accepteurs caractéristiques des introns d'eucaryotes et défini par Cartegni (Cartegni et al., 2002). Ce résultat me permet de conclure qu'il existe bien quatre gènes exprimés chez la Rue (Figure 71).

Il est surprenant que *Ruta graveolens* possède et exprime quatre gènes ayant une si forte homologie sur le plan peptidique. En effet, les Diox 1/2/4/8 ont un pourcentage d'identité peptidique entre elles supérieur à 91 % (Tableau 19). La Diox 1 présente 95 % d'identité avec la Diox 2 (Tableau 19), soit 15 acides aminés différents sur les 369 acides aminés composants les deux protéines. De même, la Diox 4 révèle 96 % d'identité avec la Diox 8, soit 12 acides aminés différents sur les 368 acides aminés. Le maintien de l'expression de ces quatre gènes différents au cours de l'évolution doit certainement apporter un avantage adaptatif à *Ruta graveolens*. Les introns sont des séquences génomiques qui subissent peu de pression de sélection liée à l'environnement. Ces séquences ont donc la possibilité de muter de manière plus rapide et pourraient expliquer la différence remarquable entre tous les introns présents au sein des différentes dioxygénases identifiées (Figure 71). La pression de sélection a certainement eu un rôle important dans l'apparition et le maintien des différentes copies chez *Ruta graveolens*. Le phénomène de duplication génique est un mécanisme central dans l'évolution des gènes et du génome. Il semble être le mécanisme majeur mis en place dans l'acquisition de nouveaux gènes et contribue de ce fait énormément à l'évolution des espèces

végétales (Ober, 2005; Roth et al., 2007). Différents phénomènes peuvent être à l'origine de cette duplication génique. Cependant, la présence d'introns dans les séquences génomiques exclut néanmoins un phénomène de duplication génique par rétroposition. L'apparition des quatre clones Diox1/2/4/8 chez *Ruta graveolens* peut découler du phénomène de diploïdisation, largement rencontré chez les angiospermes, ou du phénomène d'enjambement (ou crossing-over) inégal. Ces deux phénomènes ont pour conséquence l'apparition de copies du même gène qui peuvent alors évoluer de différentes manières. La caractérisation fonctionnelle de ces différents gènes me permettra alors de déterminer si le phénomène de duplication génique aboutit à l'acquisition de nouvelles fonctions (Ohno, 1970) (néo-fonctionnalisation ou sub-fonctionnalisation (Ober, 2010)), ou si cette duplication génique engendre une redondance de l'activité du gène ancestral entrainant une meilleure activité générale (Graur and Li, 1999). La plupart du temps, la duplication génique aboutit à la formation de pseudogènes qui sont éliminés au cours du temps, car ils n'apportent pas d'avantages évolutifs (Zhang, 2003). Néanmoins, j'ai retrouvé par deux criblages différents les quatre gènes (Diox 1/2/4/8) dans la banque de Rue, ce qui me fait penser, que ces quatre gènes sont exprimés de manière équivalente et qu'ils présentent donc un avantage adaptatif pour avoir été conservés au cours du temps. L'étape de caractérisation fonctionnelle me permettra d'une part, de déterminer le rôle de la duplication génique des dioxygénases de *Ruta graveolens*, et d'une autre part, de valider mes hypothèses quant à la formation d'un « cluster » de dioxygénases métabolisant les dérivées cinnamiques.

I.3 Conclusion

Différentes stratégies ont été mises en place afin de pouvoir identifier plusieurs gènes candidats potentiels, susceptibles d'intervenir dans la voie de biosynthèse des furocoumarines. L'analyse bibliographique s'est révélée être capitale dans la décision de rechercher des enzymes appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes. En effet, la découverte par Kai et ses collaborateurs en 2008 (Kai et al., 2008), de deux dioxygénases capables de réaliser une *ortho*-hydroxyaltion du féruloyle CoA (F6'H) pour former la scopolétine chez *Arabidopsis thaliana*, a joué un rôle important dans ma stratégie d'identification de gènes candidats au cours de ma thèse. Au vu des similitudes existant au niveau réactionnel et métabolique entre la synthèse de scopolétine et d'umbelliférone, j'ai tenté d'identifier des séquences ortologues des F6'H dans des plantes productrices de furocoumarines. La possibilité de coupler des données bio-informatiques (base de données

publique et privée) avec des essaies expérimentaux m'a permis d'isoler des gènes candidats chez *Citrus limetta*, *Pastinaca sativa* et *Ruta graveolens*. L'analyse bio-informatique de l'ensemble de ces candidats m'indique que ces gènes appartiennent à la famille des dioxygénases α-cétoglutarates et qu'ils se regroupent avec les F6'H d'*Arabidopsis*. Ce sont de bons arguments pour essayer de poursuivre cette étape d'identification de gènes candidats par une étape de caractérisation fonctionnelle afin de vérifier s'ils sont réellement impliqués dans la voie de biosynthèse des furocoumarines.

II Caractérisation de gènes codant pour des dioxygénases dans des systèmes d'expressions hétérologues.

On entend par protéine recombinante une protéine produite par une cellule dont le matériel génétique a été modifié par recombinaison génétique. Cette protéine synthétisée par un organisme autre que celui qui la produit naturellement est ainsi dite hétérologue ou recombinante. Le système de production de protéines recombinantes est caractérisé par l'utilisation d'un vecteur d'expression, jouant le rôle de transporteur génétique du gène candidat codant pour la protéine recherchée, et d'une cellule ou organisme hôte qui a pour objectif de synthétiser la protéine.

Il existe divers systèmes d'expression hétérologue consacrés à la production de protéines recombinantes. Les systèmes les plus couramment utilisés sont la bactérie *E. coli*, la levure *Saccharomyces cerevisiae* et les cellules *Chinese Hamster Ovary* (CHO). D'autres systèmes d'expressions sont de plus en plus présents dans les domaines de la recherche et des biotechnologies, tels que les cellules d'insectes, les animaux ou les plantes transgéniques.

La nature de la protéine (membranaire ou soluble) à exprimer, ainsi que les objectifs de caractérisation (*in vitro* ou *in planta*) déterminent les choix de l'hôte d'expression utilisé.

Dans cette second partie du chapitre, je présente les différents résultats obtenus après les tentatives de caractérisation fonctionnelle de mes différentes dioxygénases $Fe^{2+/}$α-cétoglutarates-dépendantes isolées chez *Citrus limetta*, *Pastinaca sativa* et *Ruta graveolens.* Puis dans un second temps, je développe de manière plus fine la caractérisation *in vitro* et *in planta* de la Diox 4 (C2'H), une dioxygénase qui présente une activité encore jamais décrite.

II.1 Expression hétérologue dans un système procaryote : *E.coli* BL21(DE3)

Les protéines appartenant à la famille des dioxygénases Fe/α-cétoglutarates-dépendantes sont des protéines solubles. Le système d'expression hétérologue dans les bactéries *E. coli* est particulièrement bien adapté pour ce type de protéines. En effet, *E. coli* peut accumuler dans le cytoplasme ou le périplasme une grande quantité de protéines solubles, jusqu'à 20 % des protéines cellulaires totales. De plus, il existe un grand nombre de vecteurs plasmidiques

disponibles afin d'insérer et d'exprimer un gène étranger. Ces vecteurs permettent pour la majorité d'entre eux de sélectionner ensuite les clones recombinants.

II.1.1 Clonage des dioxygénases dans le plasmide pET28a

Dans un premier temps, l'ORF d'intérêt est amplifié par PCR et cloné dans un plasmide transitoire de petite taille, le pCR8/TOPO/TA. Ce plasmide « high copy» permettra d'obtenir une grande quantité de plasmides qui sera utilisée pour vérifier, d'une part, la fidélité de l'amplification par PCR, et pour faciliter, d'une autre part, le sous-clonage dans différents vecteurs d'expression. Le clonage des ORF dans le pCR8 se fait après amplification du gène à l'aide d'un couple d'amorces spécifiques, sur lequel j'ai rajouté des sites de restriction aux extrémités (Tableau 23). Le plasmide pCR8 recombinant ainsi que le plasmide receveur pET28a sont ensuite digérés par les enzymes de restrictions *Bam*HI et *Xho*I. Les fragments obtenus sont purifiés et ligaturés ensemble. Ils sont ensuite utilisés pour transformer des *E.coli* BL21(DE3) compétentes en faisant une sélection sur un milieu contenant de la kanamycine.

Amorces	Sens	Séquences	Sites de restrictions
Diox 1	Dir	5'-CGGATCCATGGCACCAACAAAAGAT-3'	*Bam*HI
	Rev	5'-CCTCGAGCTAATTAATCATAGCAACGT-3'	*Xho*I
Diox 2	Dir	5'-CGGATCCATGGCACCCACAAAAGAT-3'	*Bam*HI
	Rev	5'-CCTCGAGCTAGTAATTAATCAGAGCA-3'	*Xho*I
Diox 4	Dir	5'-CGGATCCATGGCACCGACAAAAGA-3'	*Bam*HI
	Rev	5'-CCTCGAGTTAATTCATCAAAGCGAAGTCGACAGTTTTC-3'	*Xho*I
Diox 8	Dir	5'-CGGATCCATGGCACCAACAAAAGAC-3'	*Bam*HI
	Rev	5'-CCTCGAGCTAATTCATCAAAGCGAAG-3'	*Xho*I
Diox Citrus	Dir	5'-CGGATCCATGGCACCAACAAATGGCTCTATCCTGGC--3'	*Bam*HI
	Rev	5'-CGTCGACTCAAAAGTCACCTATCTTTGCAAAG-3'	*Sal*I
Diox Panais	Dir	5'-CGGATTCATGGCTCCATCTCCCACT-3'	*Bam*HI
	Rev	5'-CCTCGAGTCACATCTGTGCGAAATCAAG-3'	*Xho*I

Tableau 23 : Amorces utilisées pour amplifier les séquences codantes des dioxygénases Fe/α-cétoglutarates-dépendantes identifiées chez *Citrus limetta* (Diox Citrus), *Ruta graveolens* (Diox Ruta) et *Pastina sativa* (Diox Panais)

II.1.2 Production et purification des protéines recombinantes

Les plasmides pET28a contenant les différentes dioxygénases d'intérêts (Diox Citrus, Diox Panais, Diox 1/2/2nvl/4/8) sont introduits par choc thermique dans les bactéries *E.coli* BL21(DE3). Parallèlement, afin d'avoir un contrôle positif, j'ai également transformé ces bactéries avec un plasmide pET28a contenant le gène codant pour la F6'H1 d'*Arabidopsis*. Ce plasmide recombinant nous a été fourni par l'équipe du professeur Mizutani. Ces bactéries recombinantes sont ensuite mises en culture dans un milieu sélectif (kanamycine) afin d'obtenir une quantité de bactéries recombinantes suffisante (DO_{600} =0,6) pour réaliser l'expression hétérologue. L'induction des protéines d'intérêt est réalisée par ajout d'IPTG dans le milieu de culture, suivant le protocole décrit dans le chapitre « Matériel et méthodes ». Après leur extraction par lyse bactérienne à l'aide d'un sonicateur, les protéines sont purifiées sur colonne, par chromatographie d'affinité sur des ions métalliques immobilisés. Les protéines sont ensuite visualisées par SDS PAGE suivi d'une coloration au bleu de Coomassie. Un récapitulatif des protéines visualisées est présenté dans le Tableau 24.

	F6'H1	Diox Citrus	Diox Panais	Diox 8	Diox 4	Diox 2	Diox 2 nvl	Diox 1
SDS PAGE	+	+	-	+	+	+	+	-

Tableau 24 : Récapitulatif des différentes protéines produites par le système bactérien *E. coli* BL21(DE3). Ces protéines ont été visualisées après lyse par sonication des bactéries dans la fraction du surnageant. La migration des protéines est réalisée sur SDS PAGE à 10 % et visualisée après coloration au bleu de Coomassie. (+) symbolise la présence de la protéine ; (-) symbolise l'absence de protéine.

Les résultats que j'ai obtenus indiquent que les Diox Panais et Diox 1 ne sont pas produites dans le système d'expression bactérien alors que les Diox Citrus, la Diox 8, la Diox 4, la Diox 2 et la Diox 2 nvl sont visibles sur SDS PAGE coloré au bleu de coomassie (Tableau 24). Afin de standardiser mes tests enzymatiques, j'ai quantifié les protéines purifiées en utilisant une approche de fluorométrie (Qubit TM fluorometer d'Invitrogen). Cette méthode, a priori

plus fiable et reproductible qu'une coloration de Bradford classique m'a permis d'utiliser des solutions de protéines standardisées à 300 µg/ml.

II.1.3 Tests d'incubations

Dès que les protéines sont purifiées, j'ai entrepris les incubations en présence de différents substrats. En effet, celles-ci ne peuvent être stockées au congélateur, car leur activité décroit drastiquement. De même, la qualité et l'ordre des substrats et cofacteurs utilisés se sont révélés être très importants (cf « Matériel et méthodes »).

La caractérisation enzymatique des différentes dioxygénases, dont l'expression a été vérifiée sur SDS PAGE, est réalisée à partir de protéines purifiées. Toutes ces protéines ont été incubées en présence des cofacteurs indispensables à l'activité (α-cétoglutarate, Fe (III) et acide ascorbique) ainsi que les différents dérivés d'acides cinnamiques et de leurs formes estérifiées, suivant le protocole décrit dans le chapitre « Matériel et méthodes » (Tableau 25).

Substrats	F6'H1	Diox Citrus	Diox 8	Diox 4	Diox 2	Diox 2 nvl
Acide cinnamique	-	-	-	-	-	-
Acide *p*-coumarique	-	-	-	-	-	-
Acide férulique	-	-	-	-	-	-
Acide caféique	-	-	-	-	-	-
Cinnamoyle CoA	-	-	-	-	-	-
p-Coumaroyle CoA	-	-	-	+	-	-
Féruloyle CoA	+	-	+	+	-	+
Caffeoyl CoA	-	-	-	-	-	-

Tableau 25 : Résultats des incubations des différentes dioxygénases purifiées (Diox Citrus, Diox 8/4/2) avec les différents acides cinnamiques et leurs formes estérifiées.
Un témoin positif a été réalisé avec la F6'H1 d'*Arabidopsis thaliana* caractérisée par Kai (Kai et al., 2008). (-) signifie qu'aucune métabolisation n'a été mise en évidence. (+) signifie que le substrat à été métabolisé. L'analyse a été faite par HPLC-DAD.

Les résultats obtenus avec l'incubation de la F6'H1 d'*Arabidopsis thaliana* vont dans le sens des observations faites par Kai et ses collaborateurs. On constate une métabolisation limitée au féruloyle CoA confirmant ainsi la validité de mon témoin positif. Les autres résultats obtenus montrent que la Diox Citrus ainsi que la Diox 2 ne présentent pas d'activité avec les substrats utilisés. Les Diox 8 et Diox 2 nvl présentent une activité uniquement avec le féruloyle CoA (Tableau 25). Les produits P formés par ces deux protéines semblent identiques (Figure 72). Ils ont le même temps de rétention ainsi que le même spectre UV. Par comparaison avec les temps de rétentions et les spectres UV des molécules standards que nous possédons au laboratoire, j'ai pu identifier le produit P formé par la Diox 8 et la Diox 2 nvl comme étant de la scopolétine. Ces résultats ont été confirmés par spectrométrie de masse. Ces deux dioxygénases sont donc chez *Ruta graveolens* des orthologues de la F6'H isolées chez *Arabidopsis thaliana* par Kai et ses collaborateurs. Elles possèdent la même spécificité de substrat et métabolisent le féruloyle CoA pour former de la scopolétine. Il est intéressant de noter que les deux acides aminés interchangés (EV à la place de GG) dans la séquence codante pour la Diox 2 (Figure 67) sont certainement responsables de l'absence d'activité de la Diox 2. Ces acides aminés ont donc un rôle très important concernant les relations structures-activités des dioxygénases.

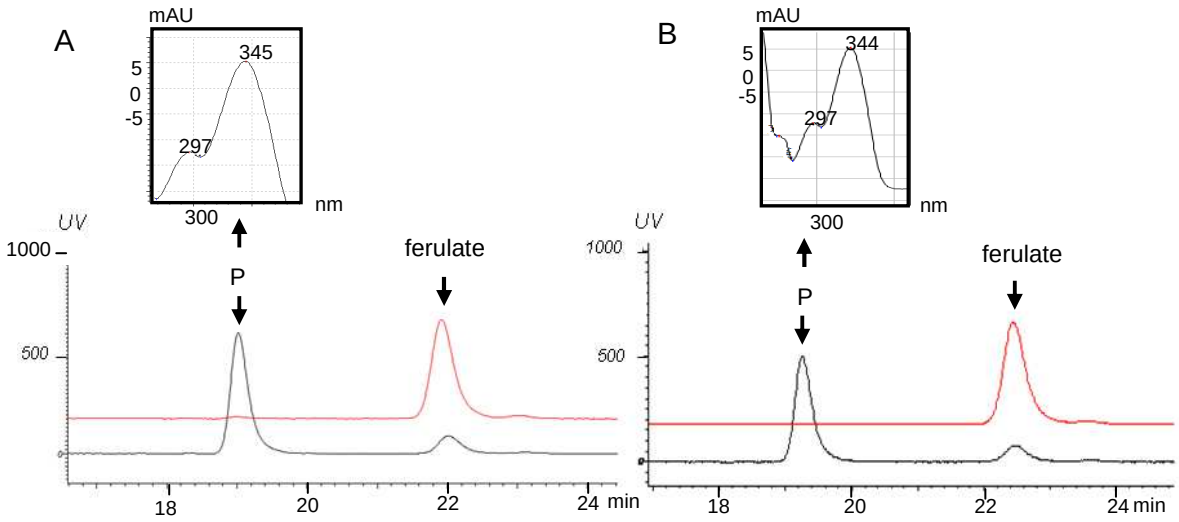


Figure 72 : Caractérisation cinétique de la Diox 2 nvl (A) et de la Diox 8 (B) isolées chez *Ruta graveolens*. L'analyse HPLC et les spectres UV montrent que le féruloyle CoA est métabolisé en un produit P avec la Diox 2 nvl (A) et la Diox 8. Ce produit a été identifié comme étant la scopolétine.

Les résultats obtenus pour la Diox 4 mettent en évidence une bi fonctionnalité. En effet, elle métabolise à la fois le *p*-coumaroyle CoA et le féruloyle CoA (Tableau 25) pour former respectivement deux produits P1 et P2.

Le temps de rétention, le spectre UV de ces produits, et l'analyse en spectrométrie de masse ont permis d'identifier le produit formé P1 comme étant l'umbelliférone, et le produit P2 de la scopolétine (Figure 73). Cette Diox 4 présente une fonction nouvelle dans la mesure où elle métabolise le *p*-coumaroyle CoA *in vitro* et contribue à l'accumulation d'umbelliférone, un précurseur de la voie de biosynthèse des furocoumarines.

Ces résultats préliminaires mettent donc en évidence l'existence de variantes de dioxygénases ayant des activités non décrites jusque-là, mais qui répondent à nos objectifs de recherche. Il semblerait que, conformément à mes hypothèses, la première étape d'entrée dans la voie de biosynthèse des furocoumarines puisse être catalysée par une enzyme de la famille des dioxygénases. Cette Diox 4 sera rebaptisée par la suite C2'H de par cette nouvelle activité vis-à-vis du *p*-coumaroyle CoA. Dans la suite de ce travail, je me suis focalisé sur la caractérisation fonctionnelle « *in vitro* », mais aussi « *in planta* » de cette enzyme.

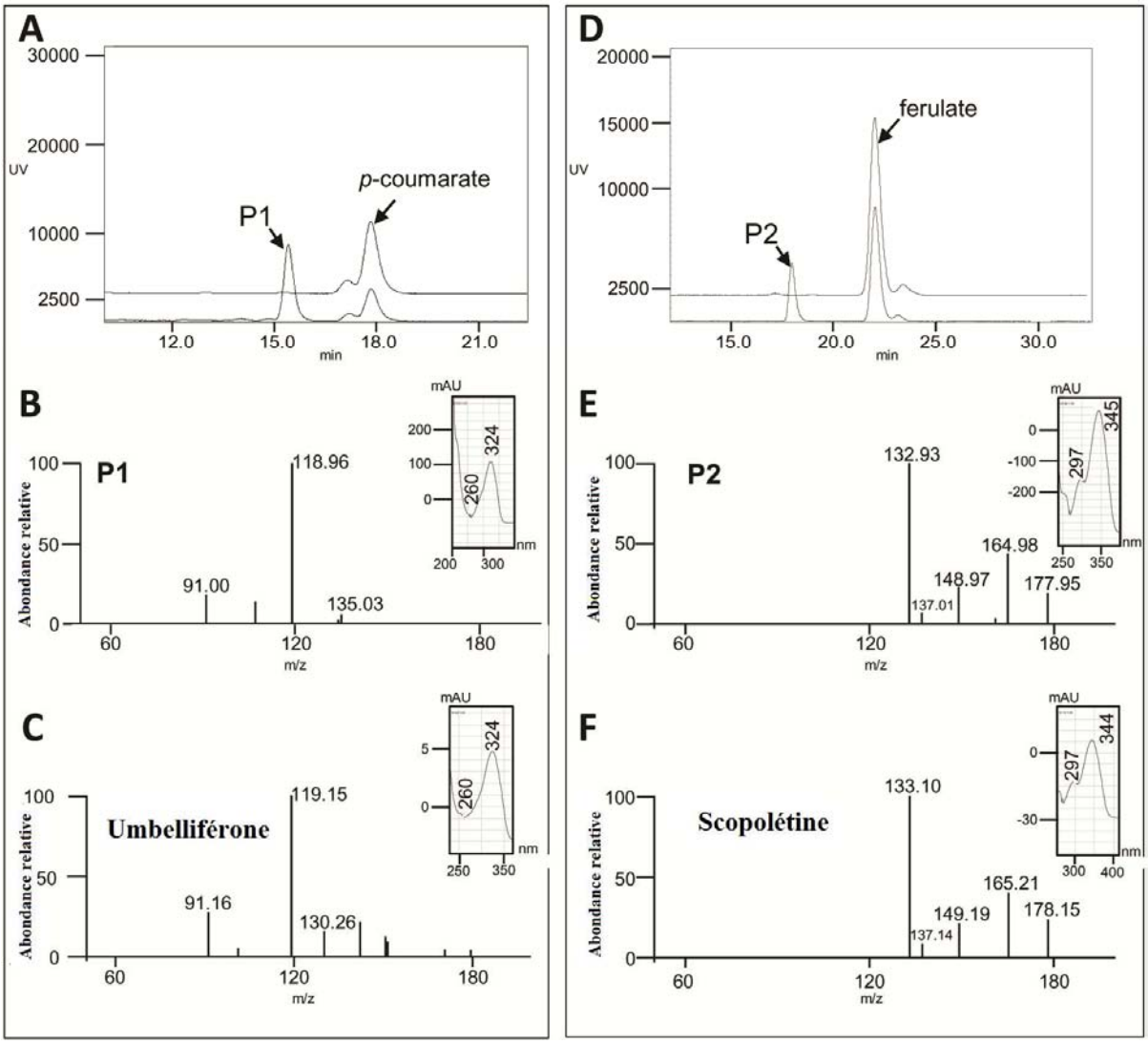


Figure 73 : Caractérisation fontionnelle de la Diox 4 isolée chez *Ruta graveolens*.
L'analyse HPLC-DAD montre que le *p*-coumaroyle CoA (A) et le féruloyle CoA (B) sont métabolisés respectivement en produit P1 et P2. Les figures B et E montrent les spectres UV et l'analyse en spectrométrie de masse des produits P1 et P2 respectivement. Les figures C et F montrent les spectres UV et l'analyse en spectrométrie de masse des molécules standards de l'umbelliférone et de la scopolétine respectivement.

II.2 Caractérisation de la Diox 4

II.2.1 Constantes cinétiques de la D4 (ou C2'H)

II.2.1.1 Paramètres cinétiques

Afin de déterminer les paramètres cinétiques de la Diox 4, j'ai réalisé des incubations avec les deux substrats préférentiels de l'enzyme, le féruloyle CoA et le *p*-coumaroyle CoA, suivant le

protocole décrit dans le chapitre « Matériel et méthodes ». Les résultats de mes cinétiques d'incubations révèlent des courbes aux allures d'hyperboles caractéristiques des enzymes de type michaeliennes (Figure 74). Les constantes cinétiques ont ensuite été calculées à partir de ces courbes à l'aide du logiciel Sigmaplot (Tableau 26).

Les résultats indiquent une affinité relativement proche pour le féruloyle CoA (K_m = 37 +/- 13 µM, $k_{cat\ fer}$ /K_m = 0.75 µmol min^{-1}) et le *p*-coumaroyle CoA (K_m= 50 +/-10 µM, $k_{cat\ coum}$ /K_m = 0.85 µmol min^{-1}).

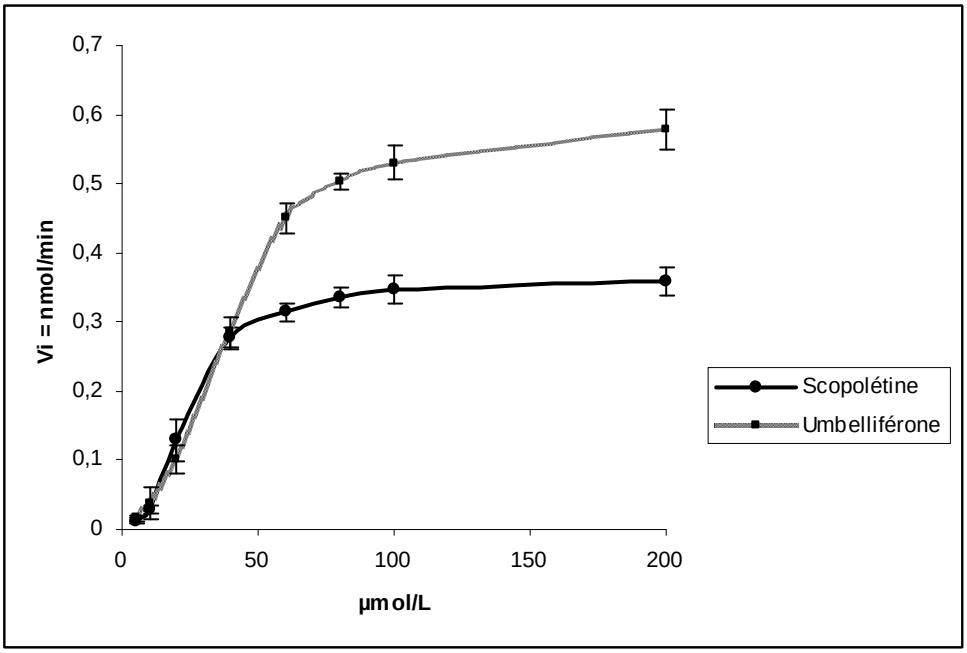


Figure 74 : Représentation des courbes cinétiques de type michaeliennes de la Diox 4 en fonction de la concentration (5 à 200 µM) en féruloyle CoA (trait plein) et en *p*-coumaroyle CoA (pointillé).

Constantes cinétiques	Féruloyle CoA	*p*-Coumaroyle CoA
V_{max} (s^{-1})	0.46 +/- 0.05	0.71 +/- 0,06
K_m (µM)	37 +/- 13	50 +/-10
K_{cat}/K_m (µmol.min^{-1})	0.75	0.85

Tableau 26 : Constantes cinétiques de la Diox 4 calculées à l'aide du logiciel Sigmaplot.
Les paramètres cinétiques ont été calculés à partir des données d'incubation de la Diox 4 avec le féruloyle CoA et le p-poumaroyl CoA.

II.2.1.2 Inhibition compétitive de la D4

Il est fréquent d'observer une inhibition des enzymes par des produits se trouvant en fin de chaine de la voie de biosynthèse qu'elles catalysent. Pour cette raison, j'ai réalisé des tests d'inhibition avec différents composés appartenant à la voie de biosynthèse des furocoumarines (umbelliférone, psoralène, xanthotoxine, bergaptène et isopimpinelline). J'ai réalisé l'incubation de la Diox 4 avec le *p*-coumaroyle CoA et le féruloyle CoA en présence de différentes concentrations (50 µM, 100 µM, et 200 µM) de ces différents composés furocoumariniques. Parmi eux, seul le psoralène permet de mettre en évidence une diminution significative de la quantité de produits formés révélant ainsi un phénomène d'inhibition de la Diox 4. Les constantes cinétiques de la réaction obtenues en présence de psoralène ont été calculées à l'aide du logiciel Sigmaplot.

Féruloyle CoA			
Constantes cinétiques	**Inhibition avec le psoralène**		
	50 µM	100 µM	200 µM
V_{max} (s^{-1})	**$V_{max\ app}$ (s^{-1})**	**$V_{max\ app}$ (s^{-1})**	**$V_{max\ app}$ (s^{-1})**
0.46 +/- 0.05	0.47 +/- 0.05	0.45 +/- 0.05	0.46 +/- 0.06
K_m (µM)	**$K_{m\ app}$ (µM)**	**$K_{m\ app}$ (µM)**	**$K_{m\ app}$ (µM)**
37 +/- 13	64 +/- 14	73 +/- 14	107 +/- 17

***p*-Coumaroyle CoA**			
Constantes cinétiques	Inhibition avec le psoralène		
	50 µM	100 µM	200 µM
V_{max} (s^{-1})	**$V_{max\ app}$ (s^{-1})**	**$V_{max\ app}$ (s^{-1})**	**$V_{max\ app}$ (s^{-1})**
0.71 +/- 0,06	0.69 +/- 0.08	0.69 +/- 0.08	0.65 +/- 0.08
K_m (µM)	**$K_{m\ app}$ (µM)**	**$K_{m\ app}$ (µM)**	**$K_{m\ app}$ (µM)**
50 +/-10	77 +/- 9	96 +/- 9	118 +/- 9

Tableau 27 : Constantes cinétiques de la Diox 4 calculées à l'aide du logiciel Sigmaplot.
Ces paramètres cinétiques ont été calculés à partir des données d'incubation de la Diox 4 avec le féruloyle CoA et le *p-coumaroyle CoA* en présence de concentrations croissantes de psoralène (50, 100, 200 µM**).**

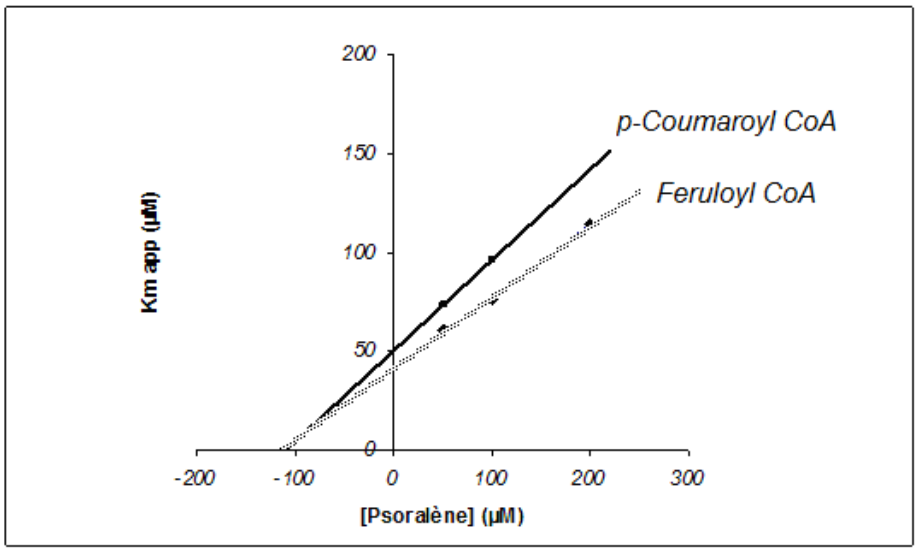


Figure 75 : Estimation du Ki du psoralène par la représentation de Kitz Wilson.
L'inhibition de la métabolisation du féruloyle CoA par le psoralène (Ki = 122 µM) est représentée en pointillé ; l'inhibition de la métabolisation du *p*-coumaroyle CoA par le psoralène (Ki = 111 µM) est symbolisée en trait plein.

On constate ainsi que la vitesse de réaction maximale n'est pas modifiée ($V_{max\ app} = V_{max}$), mais que l'affinité de l'enzyme pour le substrat est diminuée ($K_{m\ app} > K_m$) (Tableau 27). Ce type d'inhibition est caractéristique des inhibitions compétitives de type K. La représentation dite de Kitz–Wilson permet d'obtenir la valeur de Ki par une lecture graphique (Figure 75). L'estimation du Ki pour le *p*-coumaroyle CoA est 111 µM et pour le féruloyle CoA de 122 µM.

II.2.2 Relation entre la teneur en coumarines et furocoumarines et la Diox 4 (C2'H).

Les premières étapes de la caractérisation fonctionnelle ont été réalisées *in vitro* sur des extraits acellulaires de protéines. Il a déjà été démontré, par ailleurs, que ce type d'approche n'était pas toujours fidèle à la réalité et pouvait donner des résultats artéfactuels. Pour cette raison, j'ai choisi de compléter mon travail par une analyse du comportement de cette enzyme au sein de la plante, et plus précisément, au sein de la plante dont elle est originaire.

Afin d'observer s'il existe une corrélation entre le niveau d'expression du gène codant pour la Diox 4 (ou C2'H) et la concentration en coumarine et furocoumarine chez *Ruta graveolens*, j'ai analysé dans un premier temps la variation des concentrations en coumarines et furocoumarines entre des plantes de Rue ayant subi une élicitions de 24 h aux UV et des plantes témoins. Larbat avait déjà démontré durant sa thèse une faible augmentation des furocoumarines chez les plantules de *Ruta graveolens* soumises à un stress UV. Je me suis donc inspiré des ses résultats préliminaires pour mettre au point cette expérience d'élicitation. Dans un second temps, j'ai réalisé une extraction d'ARN total sur le même matériel végétal élicité et témoin, afin de quantifier le niveau d'expression de la Diox 4 par des approches de PCR en temps réel. Enfin, j'ai évalué l'impact d'un inhibiteur spécifique de dioxygénase (le Regalis®), vaporisé sur les feuilles de *Ruta graveolens* élicitées et témoins, au niveau des concentrations en coumarines et furocoumarines.

II.2.2.1 Analyse de la composition en furocoumarines chez des plantes de Rue

L'analyse quantitative de la composition en furocoumarines des plantes de *Ruta graveolens* est réalisée par HPLC. Un lot de trois plantes ayant subi un traitement aux UV pendant 24 h, et un autre lot de trois plantes contrôles sans traitement sont utilisés pour réaliser une extraction de phénylpropanoïdes au niveau des feuilles, des pétioles, des tiges et des racines. Le matériel végétal prélevé, qui n'est pas utilisé pour l'extraction, est conservé à -80 °C afin de réaliser par la suite une quantification du niveau d'expression de la Diox4. Les résultats sont exprimés en $mg.g^{-1}$ de matière fraiche.

Une comparaison est effectuée entre les plantes ayant subi une élicitation aux UV et les plantes contrôles. L'analyse statistique révèle une augmentation significative ($p<0,5$) de la teneur en umbelliférone uniquement dans les feuilles de *Ruta graveolens* après une exposition de 24 h aux UV (Figure 76). Les pétioles, tiges et racines ne sont pas affecté par le traitement aux UV. Par ailleurs, les concentrations des autres composés furocoumariniques (xanthotoxine, bergaptène, marmésine, psoralène) ne varient pas de manière significative lors de l'induction aux UV.

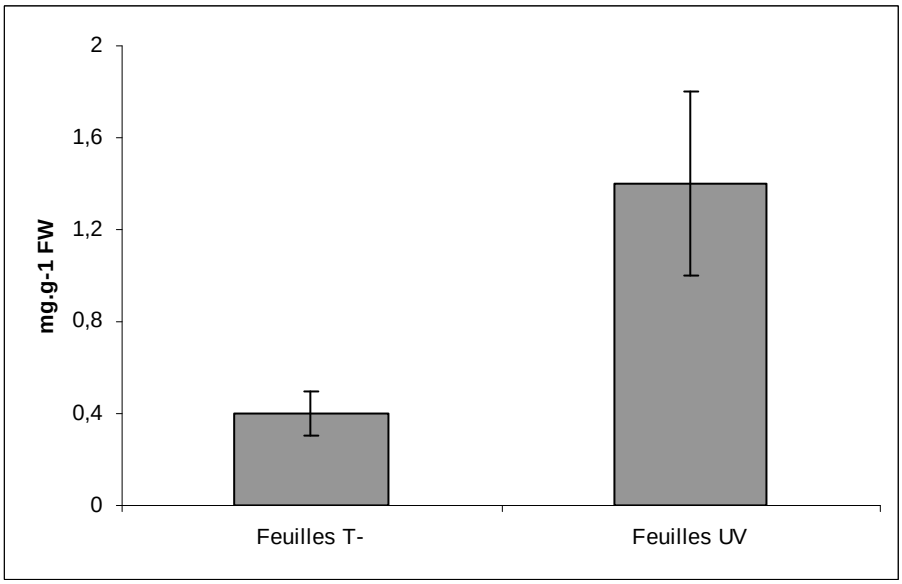


Figure 76 : Analyse quantitative de la teneur en umbelliférone dans les feuilles de Ruta graveolens.
L'analyse quantitative de l'umbelliférone, la xanthotoxine, le bergaptène, la marmésine et le psoralène a été réalisée sur les feuilles, les pétioles, les tiges et les racines d'un lot de trois plantes de *Ruta graveolens*. La même analyse est effectuée sur des plantes de Rue ayant subi un traitement de 24 h aux UV. Seule la teneur en umbelliférone révèle une augmentation significative ($p<0,5$) dans les feuilles lorsque l'on soumet les plantes de *Ruta graveolens* aux UV.

II.2.2.2 PCR en temps réel sur des ARN extraits de Rue

Afin d'établir s'il y a une corrélation entre la quantité de furocoumarines et l'expression du gène codant pour la Diox 4, j'ai inité une analyse par PCR en temps réel.

Les ARN totaux sont extraits à partir des feuilles, des pétioles, des tiges et des racines de *Ruta graveolens* ayant subi une induction aux UV pendant 24 h. Cette même extraction d'ARN totaux est réalisée sur des plantes de Rue sans traitement aux UV. Un couple d'amorces spécifiques pour la Diox 4 est généré avec le logiciel Primer Express 3.0 (Diox 4 FW : CCACCTCTCAGAGCGACTATCA ; Diox 4 Rev : CCCATGTGAATGACGGAGTCT). Les conditions de PCR sont décrites dans le chapitre « Matériel et méthodes ».

Les résultats obtenus après traitement des données par le logiciel Stepone software indiquent que la Diox 4 est exprimée dans tous les organes collectés (feuilles, pétioles, tiges, racines) (Figure 77). Le niveau de cette expression est comparable dans tous les organes lorsque les plantes ne sont pas traitées aux UV. Par contre lors d'une exposition de 24 h aux UV, le niveau d'expression de la Diox 4 est beaucoup plus important dans les feuilles et les pétioles de *Ruta graveolens* (Figure 77).

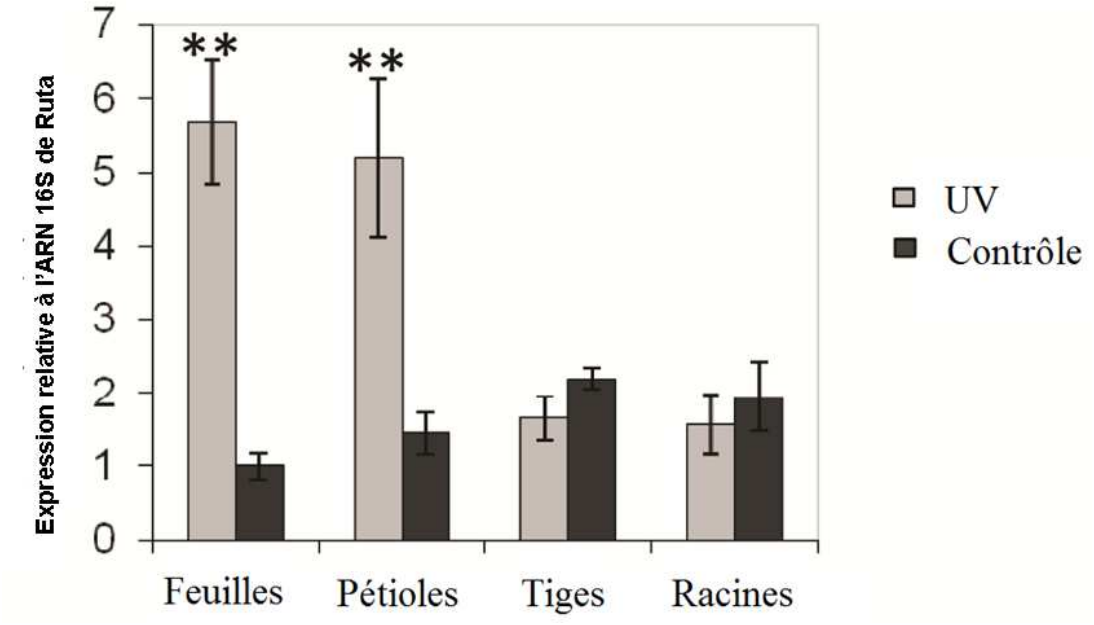


Figure 77 : Quantification du niveau d'expression de la Diox 4 réalisée sur les feuilles, les pétioles, les tiges et les racines de *Ruta graveolens* ayant subit ou non un traitement de 24 h aux UV.
(**) Indique une différence hautement significative.

Les résultats d'élicitation des composés coumariniques par les UV pendant 24 h montrent donc qu'il existe une corrélation positive, entre le niveau d'expression de la Diox 4 et la teneur en umbelliférone chez *Ruta graveolens*. En effet, l'augmentation de la synthèse d'umbelliférone (Figure 76) dans les feuilles est accompagnée d'une augmentation du niveau d'expression de la Diox 4 (Figure 77) dans ces mêmes feuilles.

II.2.2.3 Inhibition par le Regalis®

Les résultats décrits ci-dessus semblent donc indiquer une corrélation entre la concentration en umbelliférone et le niveau d'expression de la Diox4 dans les feuilles. Pour renforcer cette observation, j'ai tenté d'inhiber cette enzyme *in planta* afin d'observer une éventuelle variation de la teneur en umbelliférone. Pour cela, j'ai vaporisé une solution contenant un inhibiteur spécifique de dioxygénases à 500 ppm (le prohexadium de calcium ou Regalis®) sur les feuilles de *Ruta graveolens* élicitées aux UV et les témoins, suivant le protocole décrit dans le chapitre « Matériel et méthodes ». J'ai quantifié ensuite la teneur en umbelliférone et furocoumarines chez les plantes traitées et non traitées au Regalis® 4 jours après le

traitement. Les résultats montrent une diminution significative ($p<0,5$) de toutes les furocoumarines chez les plantes ayant reçu un traitement par le Regalis® avec ou sans élicitation au préalable par des UV (Tableau 28). Cette observation indique que la voie de biosynthèse des furocoumarines fait bien intervenir « *in planta* » une enzyme appartenant à la famille des dioxygénases, même si, à ce stade on ne peut pas conclure que cet effet soit spécifiquement lié à la Diox 4.

$mg.g^{-1}$ FW	**Témoins**	**Témoins +Regalis**	**UV**	**UV + Regalis**
Umbelliférone	0,05 ± 0,02	0,03 ± 0,01	0,08 ± 0,01	0,04 ± 0,02
Xanthotoxine	6,7 ± 0,5	5,9 ± 0,6	6,8 ± 0,4	6,1 ± 0,7
Bergaptène	22,6 ± 2,6	15,1 ± 7,1	24,4 ± 1,5	16,8 ± 5,9
Marmésine	1,8 ± 0,45	1,1 ± 0,3	1,8 ± 0,4	1,2 ± 0,3
Psoralène	1,0 ± 0,2	0,6 ± 0,3	1,1 ± 0,2	0,7 ± 0,3

Tableau 28 : Analyse quantitative en $mg.g^{-1}$ de la teneur en composés furocoumariniques chez *Ruta graveolens* après traitement par le Regalis®.
La quantification de l'umbelliférone et des composés furocoumariniques est réalisée sur quatre lots de six plantes de *Ruta graveolens* avec ou sans éliciation au préalable par les UV et avec ou sans traitement par le Regalis®. Une diminution significative de tous ces composés est constatée après 4 jours de traitement par le Regalis®.

L'ensemble des résultats qui sont présentés dans cette partie permet, d'une part, de corréler la teneur en umbelliférone des feuilles de *Ruta graveolens* au niveau d'expression dans les feuilles de la Diox 4 grâce aux expériences d'élicitation et de quantification par PCR en temps réel. D'une autre part, l'utilisation d'un inhibiteur de dioxygénase (le Regalis®) montre la diminution de la teneur en umbelliférone et furocoumarines chez les plantes élicitées ou non. Ces résultats indiquent donc que l'umbelliférone et les furocoumarines sont dépendantes « *in vivo* » de l'activité d'une enzyme appartenant à la famille des dioxygénases, la Diox 4 chez *Ruta graveolens*.

II.3 Caractérisation de la Diox 4 dans un système eucaryote *in planta*

Les étapes décrites ci-dessus indiquent que la Diox 4 est impliquée dans la synthèse de scopolétine et d'umbelliférone «*in vitro*». J'ai également montré une corrélation positive entre la synthèse de furocoumarines et le niveau d'expression de la Diox 4. Il me faut maintenant montrer sans ambiguïté que ces fonctions ne sont pas artéfactuelles, et que cette activité peut être mise en évidence chez une plante qui n'est pas capable de synthétiser de l'umbelliférone à l'état naturel. J'ai donc choisi d'exprimer la protéine chez le tabac.

II.3.1 Expression transitoire

II.3.1.1 Transformation de Nicotiana benthamiana

L'utilisation de ce système d'expression, déjà décrit dans la littérature, a été développée au sein du laboratoire. L'acquisition de la méthode a nécessité une phase de mise au point.

II.3.1.1.1 Test préliminaire avec la Green Fluorescent Protein ou GFP

II.3.1.1.1.1 Clonage de la GFP dans le pBin

Alexandre Olry et Fazeelat Karamat ont réalisé les premiers essais à l'aide d'une protéine dont le niveau d'expression est visible et quantifiable en loupe binoculaire, la GFP. Pour réaliser l'expression transitoire de la GFP, il faut au préalable insérer le gène codant pour la GFP dans un plasmide d'expression de plantes, le plasmide binaire pBin /GW comportant un promoteur constitutif fort (le promoteur 35S du CaMV). Pour cela, la première étape consiste à cloner le gène codant pour la GFP dans le sens direct du plasmide PCR8/TOPO/TA/GW (voir chapitre « Matériel et méthodes »). La technologie Gateway permet ensuite de transférer le gène d'intérêt vers un vecteur de destination (pBin/GW) par LR recombinaison.

II.3.1.1.1.2 Préparation de l'inoculum d'agrobactérie

Les plasmides recombinants sont intégrés dans les agrobactéries (*Agrobacterium tumefaciens*) par conjugaison triparentale suivant le protocole décrit dans le chapitre « Matériel et méthodes ». La sélection des souches recombinantes se fait sur un milieu sélectif contenant de

la kanamycine et une culture d'agrobactéries recombinantes est effectuée pendant 36 heures. L'acétosyringone additionnée au milieu permet d'activer les gènes de virulence d'*Agrobacterium tumefaciens*. Cet inoculum est utilisé pour infiltrer les feuilles de *Nicotiana benthamiana.*

II.3.1.1.1.3 Infiltration et visualisation en binoculaire

La face inférieure des feuilles de *Nicotiana benthamiana* est infiltrée avec l'inoculum d'agrobactérie recombinant, contenant le gène codant pour la GFP (AgrGFP), à l'aide d'une seringue suivant le protocole décrit dans le chapitre « Matériel et méthodes ». Quatre jours après inoculation, les tissus infiltrés sont visualisés sous binoculaire (Figure 78). Un témoin négatif est réalisé en infiltrant les feuilles de tabac avec de l'eau. Les résultats obtenus en binoculaire montrent que la GFP est bien exprimée suivant le protocole d'expression transitoire dans les feuilles de tabac. L'absence de fluorescence du témoin négatif confirme que la fluorescence provient bien de l'expression de la GFP. Néanmoins, l'intensité de la fluorescence n'est pas très importante. Pour l'améliorer de manière significative, l'expérience a été répétée en co-infiltrant les AgrGFP avec des agrobactéries contenant un gène codant pour une protéine suppresseur de l'activité RNA silencing, la protéine P19 du thombusvirus.

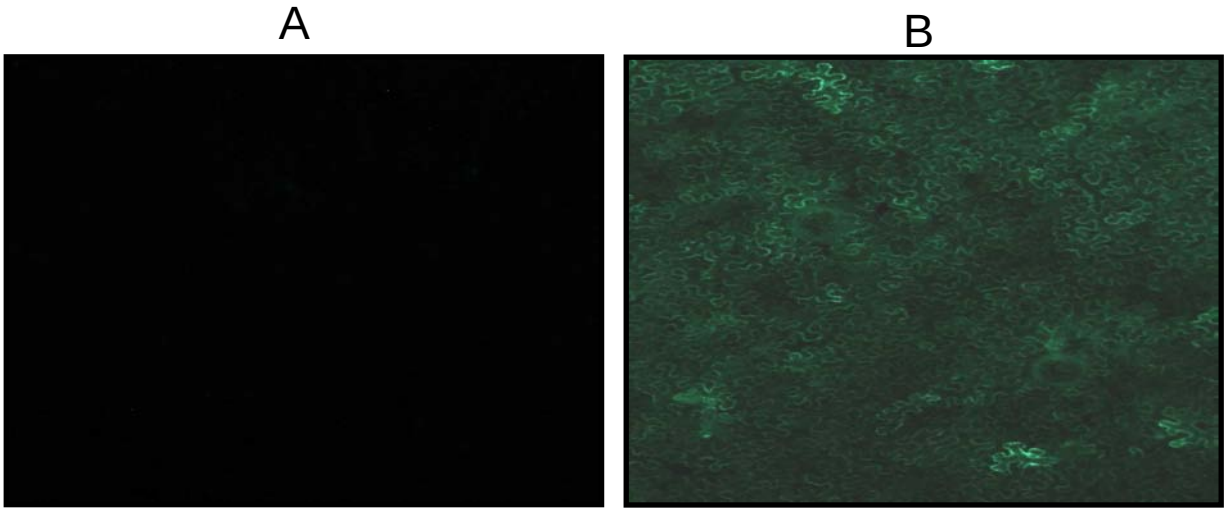


Figure 78 : Expression transitoire de la GFP dans les feuilles de *Nicothiana benthamiana*.
Visualisation à la binoculaire de l'expression transitoire de la GFP dans les feuilles de *Nicotiana benthamiana* après 4 jours d'incubation (B). Un témoin négatif est réalisé en infiltrant les feuilles de tabac avec de l'eau (A). L'intensité de la coloration verte est proportionnelle au niveau d'expression de la GFP.

II.3.1.1.1.4 Co-infiltration de la GFP avec la protéine P19

Le RNA silencing est un mécanisme utilisé par les plantes comme moyen de lutte antivirale. Pour contrecarrer ce système de défense, les virus ont développé des protéines dites « suppresseurs de RNA silencing » capables d'inhiber les étapes de la cascade réactionnelle du RNA silencing (Chapman and Carrington, 2007). La protéine p19 provient des thombusvirus. Elle est capable de séquestrer les siRNA impliqués dans le processus de RNA silencing. L'utilisation de cette protéine en co-expression avec la protéine GFP permet de contourner le système de défense naturelle de *Nicotiana benthamiana*, et ainsi de surexprimer la GFP. Le protocole d'expression transitoire du gène codant pour la protéine p19 (AgrP19) dans les feuilles de tabac se fait de la même manière que celui décrit pour la GFP. Les deux solutions d'*Agrobacterium tumefaciens* (AgrGFP et AgrP19) sont infiltrées simultanément dans les feuilles de *Nicotiana benthamiana*. Celles-ci sont ensuite visualisées par observation avec une loupe binoculaire. Un contrôle négatif est réalisé avec la co-infiltration d'eau et d'Agrobacterium recombinant contenant le gène codant pour la protéine p19 (AgrP19). Un autre contrôle est effectué avec l'infiltration des agrobactéries recombinante contenant le gène codant pour la GFP et de l'eau. Le niveau d'expression de la GFP dans les feuilles de tabac est ensuite visualisé à la loupe binoculaire (Figure 79). Les résultats de fluorescence indiquent que le niveau d'expression de la GFP est nettement augmenté lorsqu'elle est co-infiltrée avec le gène codant pour la protéine P19.

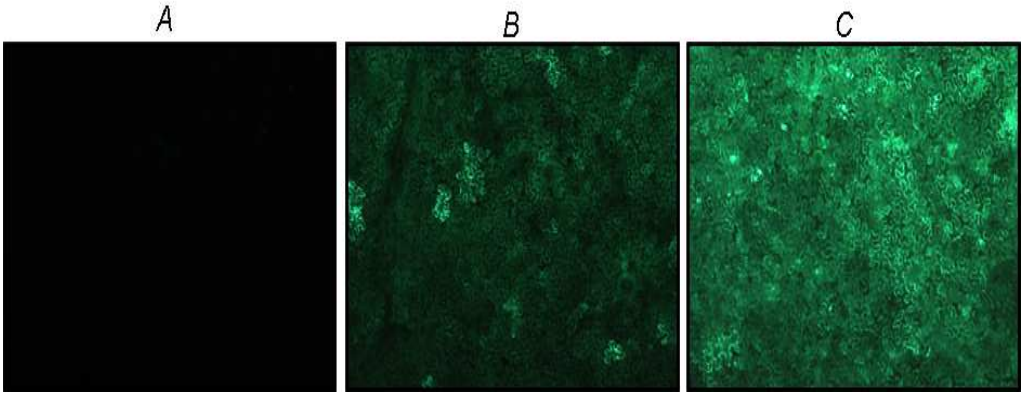


Figure 79 : Co-expression transitoire de la GFP et de la protéine P19 dans les feuilles de *Nicothiana benthamiana.*

Visualisation à la loupe binoculaire du niveau d'expression de la GFP dans les feuilles de *Nicotiana benthamiana* par transformation transitoire après 5 jours d'incubation. (A) Un témoin négatif est réalisé en infiltrant les feuilles de tabac avec de l'eau et les agrobactéries recombinantes contenant le gène codant la protéine p19 (AgrP19). (B) Un second contrôle est réalisé en co-infiltrant les agrobactéries recombinantes contenant le gène codant pour la GFP (AgrGFP) avec de l'eau. (C) co-infiltration de AgrGFP et AgrP19. Expérience réalisée par Alexandre Olry et Fazeelat Karamat.

II.3.1.1.2 Infiltration de la Diox 4

Suivant le protocole décrit pour transformer les AgrGFP, j'ai introduit le gène codant pour la Diox 4 dans le plasmide pBin/GW avant d'utiliser le plasmide recombinant pour transformer les *Agrobacterium tumefaciens* (AgrDiox4). J'ai également transformé des agrobactéries avec le plasmide pBin/GW vide afin de constituer un contrôle négatif (AgrT-). J'ai donc choisi de co-exprimer le gène Diox 4 avec celui codant pour la protéine p19. Pour cela, j'ai préparé trois inoculums d'Agrobactéries recombinantes différentes : un inoculum d'AgrDiox4, un second inoculum d'AgrP19 et un troisième d'AgrT-. Ensuite, j'ai réalisé la co-infiltration de deux lots de trois plantes de *Nicotiana benthamiana* suivant le protocole décrit dans le chapitre « Matériel et méthodes ». Le premier lot de plantes a été infiltré avec les AgrDiox4 et les AgrP19, et le second lot a été infiltré avec les AgrT- et les AgrP19 afin de constituer un lot de contrôle négatif. Après trois jours d'incubation, les feuilles de *Nicotiana benthamiana* co-infiltrées ont été prélevées afin de réaliser une extraction des composés phénylpropanoïdes.

II.3.1.2 Extraction coumarinique et identification par SM

Trois jours après agro-infiltration les composés coumariniques sont extraits des feuilles de tabac et sont analysés par spectrométrie de masse. Les analyses de spectrométrie de masse sont très sensibles et très spécifiques. Deux molécules ont été spécifiquement recherchées grâce à un filtre de masse : l'umbelliférone et la scopolétine. Ces deux composés ont été identifiés uniquement dans le lot de plantes co-infiltré avec les AgrDiox4 et AgrP19 (Figure 80 A et C). La comparaison de ces deux composés avec les spectres de masse des molécules standards de l'umbelliférone et de la scopolétine standards (Figure 80 B et D) ont certifié la production de ces deux molécules dans le lot de plantes co-infiltré avec les AgrDiox4 et AgrP19. Ces deux composés n'ont pas été détectés dans le lot de plantes co-infiltré avec les AgrT- et AgrP19. De même, la recherche de forme glycosylée de la scopolétine (scopoline) et de l'umbelliférone (skimmine) a montré l'absence de ces composés dans les plantes co-infiltrées et contrôles.

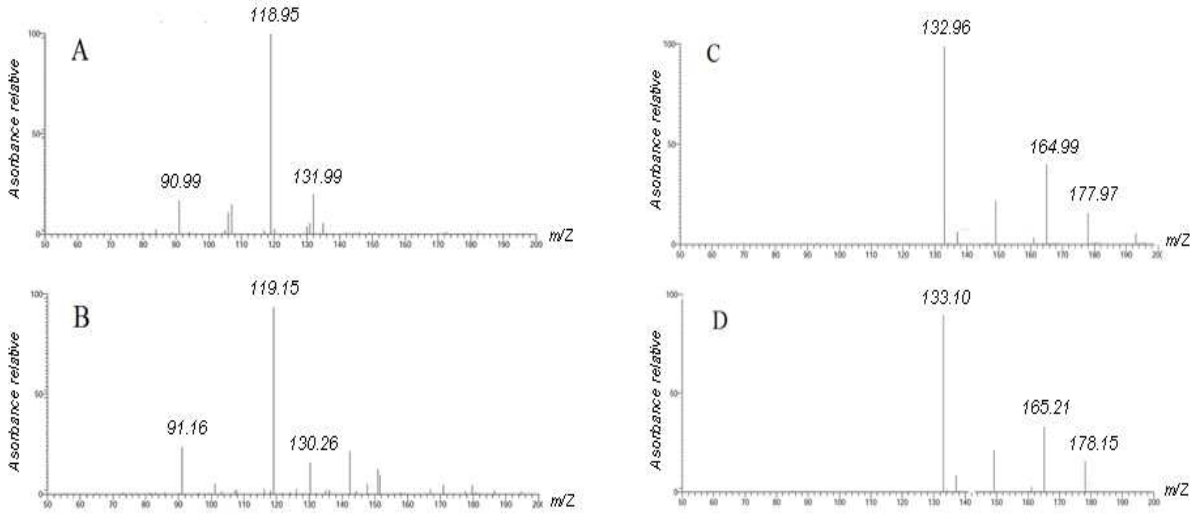


Figure 80 : Identification de l'umbelliférone et de la scopolétine en spectrométrie de masse chez *Nicotiana benthamiana* surexprimant la Diox 4.

L'analyse des composés présents chez les feuilles de *Nicotiana benthamiana* après expression transitoire avec les AgrDiox4 et AgrP19 (A et C) est réalisée à l'aide d'un spectromètre de masse. Ces composés sont absents dans les plantes témoins. La comparaison avec les spectres de masse obtenue pour des molécules standards (B : umbelliférone et D : scopolétine) ont permis d'identifier les composées synthétisés comme étant de l'umbelliférone (A) et de la scopolétine (C).

II.3.1.3 Dosage par HPLC

Les approches quantitatives par spectrométrie de masse nécessitent beaucoup de précalibration en comparaison aux approches par UV. Cela vient du fait que la réponse du détecteur n'est pas parfaitement proportionnelle à la concentration de l'analyte cible. J'ai donc choisi d'effectuer la quantification par des approches UV en HPLC-DAD.

Afin de tester le niveau de production des molécules et de spécificité de l'enzyme, l'umbelliférone et la scopolétine ont été quantifiés. Pour cela, j'ai additionné lors de mon extraction de composés phénylpropanoïdes un étalon interne de dosage (la taxifoline) me permettant de m'affranchir des variations potentielles liées à l'expérimentation (perte de produit pendant l'extraction par exemple). Après concentration de mes solutions d'extraction suivant le protocole décrit dans le chapitre « Matériel et méthodes », je les ai analysées et quantifiées par HPLC-DAD. Le dosage de l'umbelliférone et de la scopolétine chez les feuilles de *Nicotiana benthamiana* co-transformées avec les AgrDiox4 et les AgrP19, révèle une quantité équivalente en umbelliférone et en scopolétine (respectivement 0,3 $mg.g^{-1}$ et 0,25

mg.g^{-1} de matière fraîche) (Figure 81). Ces deux composés ne sont toujours pas détectés dans les feuilles de *Nicotiana benthamiana* contrôles, co-transformées avec les AgrT- vide et les AgrP19.

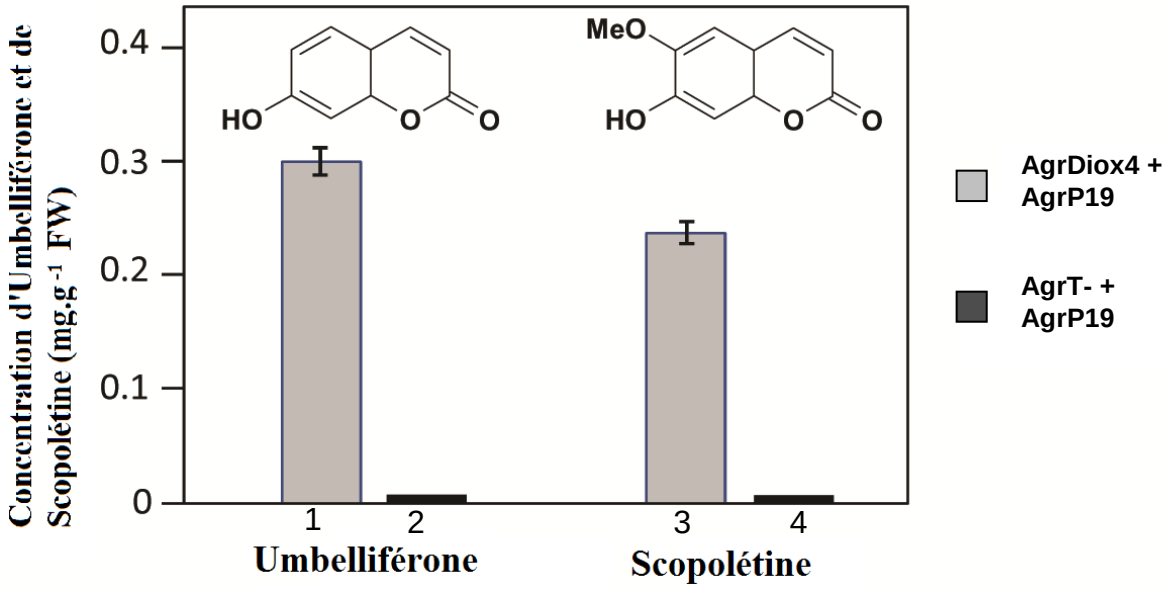


Figure 81 : Dosage de la quantité d'umbelliférone et de scopolétine exprimé en mg.g^{-1} de matière fraiche chez les feuilles de *Nicotiana benthamiana* issues de l'expression transitoire.
Un lot de trois plantes a été co-infiltré avec AgrDiox 4 et AgrP19 (1 et 3), et un lot de trois plantes contrôles a été co-infiltré avec AgrT- et AgrP19 (2 et 4).

Ces résultats d'expressions transitoires des feuilles de *Nicotiana benthamiana* réalisées avec la co-infiltration simultanée de deux inoculums agrobactéries recombinantes transformées soit avec le plasmide pBin/GW contenant le gène Diox 4 (AgrDiox4), soit avec le plasmide contenant le gène codant pour la protéine P19 (AgrP19), ont montré que la dioxygénase α-cétoglutarate-dépendante issue de *Ruta graveolens* (Diox4) est impliqué dans la synthèse d'umbelliférone et de scopolétine « *in planta* ».

II.3.2 Expression stable

À ce stade il me manque encore un volet concernant l'étude de la fonction de la Diox 4, à savoir, s'il est possible de moduler la synthèse de furocoumarines en surexprimant la C2'H dans *Ruta graveolens*.

II.3.2.1 Transformation de *Ruta graveolens*

Afin de poursuivre cette caractérisation « *in planta* » de la Diox 4, j'ai réalisé la transformation génétique des hypocotyles de *Ruta graveolens*. Le protocole de transformation génétique de Rue a été mis au point par Karine Lièvre durant sa thèse (Lièvre., 2004) qui a développé une méthodologie permettant la régénération *in vitro* d'hypocotyles de *Ruta graveolens,* et l'insertion stable de gène dans le génome de cette plante. Selon ce protocole de transformation génétique décrit plus en détail dans le chapitre « Matériel et méthodes », j'ai réalisé la transformation génétique avec l'inoculum d'agrobactéries recombinantes transformées avec le plasmide pBin/GW/Diox4 (AgrDiox4). Un contrôle négatif est réalisé en incubant des hypocotyles de Rue dans un inoculum d'agrobactéries recombinantes avec le plasmide pBin/GW vide (AgrT-). Le protocole de préparation de ces inoculums se fait de la même manière que ceux utilisés pour l'expression transitoire dans les feuilles de tabac. Après immersion des explants d'hypocotyles de Rue dans l'inoculum d'agrobactérie AgrDiox4 ou AgrT-, ils sont incubés sur un milieu de régénération avant d'être transférés sur un milieu sélectif contenant de la kanamycine pour sélectionner les cellules transformées (Photo 16).

Photo 16 : Sélection des explants d'hypocotyles de *Ruta graveolens* sur milieu sélectif contenant de la kanamycine pendant un à deux mois.

Certaines cellules transformées donnent naissance à des tiges feuillées de quelques centimètres qui sont alors repiquées dans des pots contenant le même milieu sélectif pour maintenir la pression de sélection. Ces plantules sont actuellement en cours de régénération (Photo 17) et il reste donc à faire la caractérisation moléculaire puis enfin les analyses du profil métaboliques des plantes transformées.

Photo 17 : Plantule de *Ruta graveolens in vitro* potentiellement transformée.

II.3.2.2 Transformation Arabidopsis thaliana

Le même inoculum d'agrobactérie AgrDiox4 est utilisé pour réaliser la transformation d'*Arabidopsis thaliana*. Un inoculum d'agrobactérie AgrT- est réalisé afin de constituer un lot de plantes contrôles. Les inoculums d'agrobactéries sont diluées dans une solution de saccharose à 5 % et de silwet L-77 à 0,05 % avant d'être utilisés. La méthode dite « floral dip » (Clough and Bent, 1998) est utilisé pour réaliser la transformation des plantes d'Arabidopsis. Après immersion des plantes (tête en bas) dans l'inoculum d'agrobactéries, celles-ci sont cultivées en phytotrons jusqu'à l'apparition de graines. Les siliques des plantes

transformées sont récupérées et les graines sont collectées dans un sac en papier avant d'être mises à germer sur un milieu MS sélectif contenant de la kanamycine. La caractérisation moléculaire ainsi que la comparaison des profils métaboliques des plantes transformés avec les AgrDiox4 avec les plantes contrôle reste à faire.

II.4 Discussion

L'umbelliférone est une coumarine que l'on retrouve de manière ubiquiste dans le règne végétal. Les premières études indiquaient que l'umbelliférone provenait de l'*ortho*-hydroxylation de l'acide *p*-coumarique (Ellis and Amrhein, 1971; Gestetner and Conn, 1974). Cependant, les gènes impliqués dans sa synthèse n'avaient jamais été identifiés et caractérisés jusqu'à nos jours. Une étude récente publiée en 2008 par Kai et ses collaborateurs (Kai et al., 2008) a contribué à identifier et caractériser cette étape d'*ortho*-hydroxylation. En effet, ces auteurs ont montré qu'une dioxygénase α-cétoglutarate-dépendante hydroxyle le féruloyle CoA (F6'H) pour former la scopolétine chez *Arabidopsis thaliana*. Cependant, leurs expériences ont révélé que la F6'H (F6'H1 et F6'H2) d'*Arabidopsis thaliana* n'était pas impliquée dans la synthèse d'umbelliférone (Kai et al., 2008). Les gènes impliqués dans la voie de biosynthèse de l'umbelliférone n'étaient donc toujours pas identifiés.

Du fait de l'analogie structurale et réactionnelle mise en jeu dans la synthèse de scopolétine et d'umbelliférone, j'ai essayé de caractériser les différents candidats potentiellement orthologues des F6'H, issus de *Ruta graveolens* (Diox 1, Diox 2, Diox 2 nvl, Diox 4 et Diox8), de *Citrus limetta* (Diox Citrus) et de *Pastinaca sativa* (Diox Panais). Tous ces candidats appartiennent à la famille des dioxygénases α-cétoglutarates-dépendantes et ont été identifiés par les méthodes citées dans la partie précédent du chapitre. L'expression hétérologue dans les *E.coli* BL21(DE3) m'a permis de visualiser les protéines sur SDS PAGE résultant de l'expression des gènes candidats Diox 2, Diox 2 nvl, Diox 4, Diox 8 et Diox Citrus (Tableau 24). Après purification sur colonne chromatographique d'affinité, j'ai réalisé des tests de métabolisation de différents dérivés cinnamiques avec toutes les protéines isolées (Tableau 25).

La caractérisation fonctionnelle de la Diox 2 nvl et de la Diox 8 dans les *E.coli* BL21(DE3) a permis de mettre en évidence *in vitro* la métabolisation spécifique du féruloyle CoA pour ces deux enzymes. Ce sont donc deux gènes orthologues des F6'H d'*Arabidopsis thaliana*

caractérisées par Kai (Kai et al., 2008) qui présentent la même spécificité de substrat vis-à-vis du féruloyle CoA.

La Diox 4 montre *in vitro* une double spécificité de substrats. Elle est capable de métaboliser le féruloyle CoA et le *p*-coumaroyle CoA pour former respectivement la scopolétine et l'umbelliférone (Figure 73) avec la même efficacité pour les deux substrats (Tableau 26).

Cette bifonctionnalité a été confirmée *in vivo* par l'expression transitoire de la Diox 4 dans les feuilles de *Nicotiana benthamiana,* sous le contrôle du promoteur 35S. Le tabac est une plante qui contient une très faible teneur en scopolétine et scopoline (forme glycosylée de la scopolétine), mais pas d'umbelliférone (Chong et al., 2002; Costet et al., 2002). Les plantes transformées par agro-infiltration, récoltées trois jours après infection, montrent un niveau élevé d'umbelliférone et de scopolétine (Figure 81). Or ces composés n'ont pas été détectés, ou justes à l'état de traces, chez les plantes contrôles. L'accumulation de ces deux coumarines dans les feuilles de *N. benthamiana* confirme le caractère bifonctionnel de cette enzyme chez les plantes. La forte teneur en umbelliférone qui a été retrouvée dans les plantes de tabac exprimant la Diox 4 est particulièrement intéressante afin de démontrer le rôle de cette enzyme dans la synthèse d'umbelliférone, car ces plantes n'accumulent naturellement pas d'umbelliférone. De plus, le niveau d'expression de la Diox 4 chez *Ruta graveolens* indique une corrélation positive avec l'accumulation d'umbelliférone chez les feuilles et les pétioles traités aux UV (Figure 76 et 77). Les résultats de l'expression « *in planta* » corroborent donc le caractère multifonctionnel de cette enzyme.

L'ensemble de ces résultats me permet de conclure que la Diox 4 présente à la fois une activité F6'H et *p*-coumaroyle CoA 2'-hydroxylase (C2'H) chez *Ruta graveolens* avec une efficacité similaire. D'après mes incubations *in vitro*, l'umbelliférone est issu de l'*ortho*-hydroxylation du *p*-coumaroyle CoA et non de l'acide *p*-coumarique (Tableau 25). Ce résultat suggère que la Diox 4 travaille en collaboration avec une 4-coumarate CoA ligase (4CL) effectuant la synthèse au préalable de l'ester de *p*-coumaroyle CoA. Cette théorie est en accord avec d'autres résultats observés chez *Ruta graveolens*, qui décrivent la présence de deux 4CL différentes (Endler et al., 2008) : l'une d'entre elles est supposée être impliquée spécifiquement dans la synthèse de coumarines (Endler et al., 2008) et pourrait agir en tandem avec la Diox 4 (C2'H) décrite précédemment.

Parallèlement à ce travail, Shimidzu et ses collaborateurs ont isolé un autre gène candidat issu de *Ipomea batatas* codant pour une autre dioxygénase Fe/α-cétoglutarate-dépendante (Matsumoto et al., 2011). La caractérisation fonctionnelle qu'ils ont effectué *in vitro* par expression hétérologue dans les *E.coli* va dans le même sens, et indique qu'il existe aussi chez

Ipomea des enzymes bifonctionnelles. En effet, la dioxygénase identifiée chez ipomée présente aussi une double activité avec le féruloyle CoA et le *p*-coumaroyle CoA.

Une autre dioxygénase α-cétoglutarate-dépendante bifonctionnelle a déjà été décrite dans la voie de biosynthèse des phénylpropanoïdes, et plus précisément dans le métabolisme des flavonoïdes chez Citrus (LukaÄina et al., 2003). Le caractère multifonctionnel de certaines enzymes est un élément essentiel pour expliquer la complexité et la diversité des métabolites identifiés chez les plantes (Schwab, 2003). En plus de l'activité F6'H, l'acquisition d'une activité supplémentaire C2'H par la Diox 4 chez *Ruta graveolens* ouvre la voie de biosynthèse de l'umbelliférone et de nombreux autres dérivés du 7- hydroxycoumarine, tels que les furocoumarines, qui jouent un rôle primordial dans la défense des plantes (McCloud et al., 1992; Neal, 1994). Ces furocoumarines ne sont pas présentes chez *Arabidopsis thaliana* qui ne contient qu'une F6'H monofonctionelle (Kai et al., 2008). La scopolétine, qui résulte de l'activité F6'H, a été recherchée chez *Ruta graveolens*, mais n'a pas été détectée de manière significative dans les tissus de la plante. Cette absence de scopolétine chez *Ruta graveolens* est surprenante dans la mesure où, la Diox 4 possède une activité F6'H équivalente à l'activité C2'H *in vitro* (Tableau 26) et *in planta* (Figure 81). Il aurait été concevable de trouver la présence concomitante de scopolétine et d'umbelliférone chez *Ruta graveolens*. La formation de produits dérivés, en aval de la voie de biosynthèse de scopolétine, pourrait expliquer cette absence de scopolétine chez la Rue. Cependant, la recherche de scopoline (forme glycosylée de la scopolétine) à laquelle nous avons procédé par des approches de spectrométrie de masse n'indique pas non plus de teneur significative dans les tissus de *Ruta graveolens*. La scopoline est d'ailleurs considérée comme une coumarine mineure présente chez la Rue (Steck et al., 1971; Paszkiewicz et al., 2008). Une autre étude récente confirme que les quantités de scopolétine et scopoline sont inférieures chez *Ruta graveolens* à la limite de détection de l'HPLC-UV (<2,5 $\mu g.g^{-1}$ FW) (Adam et al., 2009).
De plus, l'identification et la caractérisation de la Diox 2 nvl et de la Diox 8 chez *Ruta graveolens* possédant uniquement une activité F6'H *in vitro* est en contradiction avec l'absence de scopolétine chez la Rue. D'autres études expérimentales devront être menées afin de mieux comprendre cette ambiguïté et d'expliquer cette absence de scopolétine et de scopoline chez *Ruta graveolens*.

Les expériences cinétiques d'inhibition de la Diox 4 avec les différents composés coumariniques et furocoumariniques ont montré que le psoralène était un inhibiteur compétitif

de l'activité F6'H et C2'H de la Diox 4 (Tableau 27). Les furocoumarines sont des composés naturels possédant des fonctions allélopathiques de par leur importante activité phytotoxique. En effet, les furocoumarines linéaires sont connues pour leurs capacités d'interagir avec l'ADN. Ce sont des molécules intercalantes qui sont capables d'établir des liaisons covalentes entre deux résidus de thymine (Dall'Acqua et al., 1978b). Pour cette raison les furocoumarines linéaires présentent une toxicité pour tous les organismes, et potentiellement pour ceux qui les produisent en grande quantité tels que *Ruta graveolens*. Le Ki calculé expérimentalement concernant l'inhibition de la Diox 4 par le psoralène correspond au double de la valeur du *Km* pour le féruloyle CoA et *p*-coumaroyle CoA (Figure 75). Bien que l'affinité de la Diox 4 soit moins importante pour le psoralène que pour ses deux substrats (féruloyle CoA et *p*-coumaroyle CoA), le psoralène pourrait agir par un rétrocontrôle inhibiteur de l'activité C2'H lorsque la concentration en psoralène deviendrait excessive et nocive pour l'organisme. Ce *retro*-control négatif permettrait de réguler la voie de biosynthèse des furocoumarines et d'ajuster les teneurs en furocoumarines dans la plante. Chez *Ruta graveolens*, cette voie de régulation est cohérente avec la synthèse de furocoumarines qui se fait de manière constitutive dans tous les tissus de la plante (Eiler, 1989; Milesi et al., 2001), et avec les résultats obtenus lors de l'élicitation des Rue aux UV ne montrant pas d'augmentation significative de la concentration en furocoumarines dans les différents tissus. Contrairement à d'autres plantes qui accumulent des furocoumarines et qui présentent une forte sensibilité à l'élicitation (Hamerski and Matern, 1988a), par régulation du niveau d'expression des gènes (Larbat et al., 2007), *Ruta graveolens* pourrait avoir développé un système alternatif de régulation, tel que la rétro-inhibition enzymatique.

L'analyse bio-informatique des séquences peptidiques de la Diox 4 et Diox 8 identifiées chez *Ruta graveolens* révèle une très forte identité (96 %) entre ces deux dioxygénases. Seulement 12 différences sur les 368 acides aminés composant les deux dioxygénases ont été identifiées. Malgré cette forte ressemblance peptidique, la caractérisation fonctionnelle des deux protéines a montré une spécificité de substrat différente. La Diox 8 métabolise uniquement le féruloyle CoA alors que la Diox 4 métabolise le féruloyle CoA et le *p*-coumaroyle CoA. Afin de mieux comprendre les relations structures fonctions qui existent dans cette famille d'enzyme, des travaux de mutagenèses dirigés ont été entrepris pour identifier les acides aminés impliqués dans la spécificité de substrat. La modélisation tridimensionnelle de ces deux dioxygénases à l'aide du logiciel MOE, m'a permis avec l'aide de Romain Larbat, d'identifier (Figure 82) un acide aminé différent sur la Diox 8 situé proche de la poche catalytique de l'enzyme : la

valine (Val^{218}) située sur la Diox 8 et remplacée sur la Diox 4 par une isoleucine (Ile). Un second acide aminé différent entre la Diox 8 et la Diox 4 a été sélectionné (Figure 82), car ils présentent le plus de divergence. En effet, l'acide glutamique (Glu^{104}) présent sur la Diox 8 est remplacé par la glutamine (Gln^{104}) sur la Diox 4. La fonction acide carboxylique de la Diox 8 est remplacée par une fonction amine sur l'acide aminé 104 dans la Diox 4. Cette divergence de fonction peut entrainer des modifications d'affinités du substrat, et nous avons donc décidé d'interchanger aussi celui-ci. Les expériences de mutagenèse dirigée vont donc permettre, par des approches de PCR, de remplacer sur la Diox 8 la Val^{218} par une Ile^{218} et/ou la Glu^{104} par une Gln^{104}. Pour cela, j'ai commandé des couples d'amorces spécifiques qui comprennent l'acide aminé interchangé (Tableau 29). Les expériences sont en cours d'analyse.

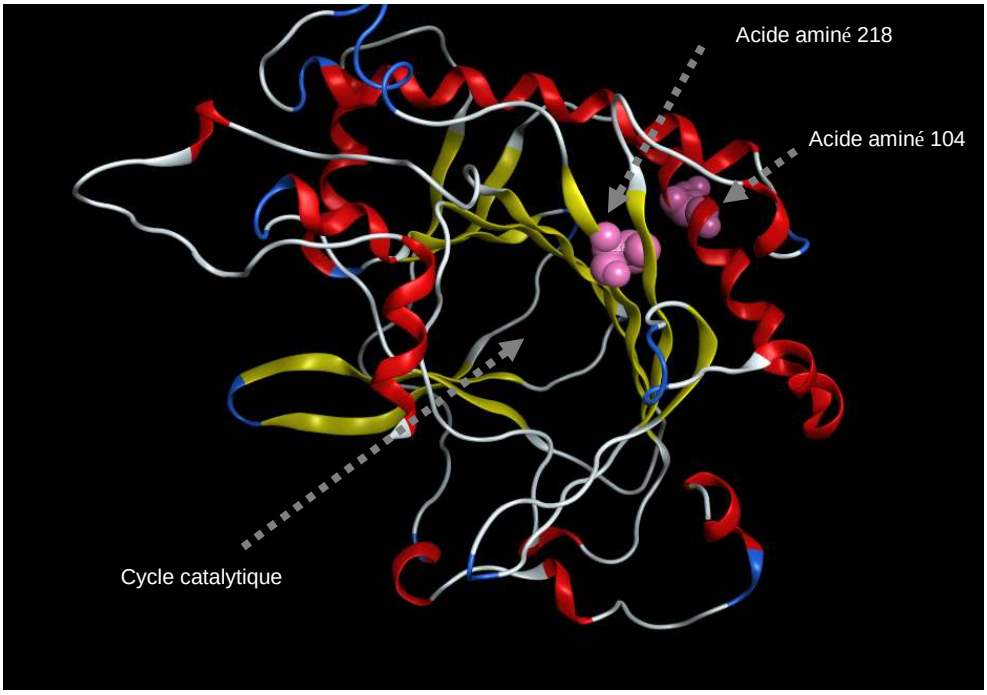


Figure 82 : Modélisation tridimensionelle de la Diox 8 à l'aide du logiciel MOE.
L'acide aminée 218 ainsi que l'acide aminé 104 ont été sélectionnés pour réaliser de la mutagénèse dirigée. Remplacement de la Val^{218} par une Ile^{218} et/ou de la Glu^{104} par la Gln^{104}.

Diox 8	Amorce avec la mutation
$Val^{218} \rightarrow Ile^{218}$	5'-GTA GTT GAG GTT GAT TCT CCT GGA ACC-3'
	3'-GGT TCC AGG AGA ATC AAC CTC AAC TAC-5'
$Glu^{104} \rightarrow Gln^{104}$	5'- CTT CAC CCT CTG CAG AGT TTC C -3'
	3'- GGA AAC TCT GCA GAG GGT GAA G-5'

Tableau 29 : Couple d'amorces utilisé pour la mutagenèse dirigée contre l'acide aminé valine (Val^{218}) et leucine (Glu^{104}) de la Diox 8.
Ces acides aminés seront remplacés respectivement par l'isoleucine et la glutamine (souligné sur les amorces).

De plus, les résultats de caractérisation fonctionnelle de la Diox 2 et de la Diox 2 nvl permettent de conclure quant à la fonction des acides aminés différents identifiés (Figure 67). En effet, les deux protéines correspondantes sont exprimées et visualisées par SDS PAGE après coloration au bleu de Coomassie (Tableau 24), cependant seule la protéine correspondante au gène Diox 2 nvl révèle une activité. J'en conclus que les deux acides aminés glycines (GG) (Figure 67) identifiées chez tous les gènes candidats, hormis sur la Diox 2, sont des acides aminés indispensables à l'activité des dioxygénases Fe^{2+}/α-cétoglutarates-dépendantes.

De même l'absence d'expression de la Diox 1, peut provenir de l'acide aminé différent identifié (Figure 68). Sur cette Diox 1, la leucine présente chez toutes les autres dioxygénases identifiées, est remplacée par une phénylalanine. Cet acide aminé pourrait être impliqué dans le bon repliement de la protéine. L'identification et la caractérisation de cette Diox 1 non muté, permettront de valider cette hypothèse. Une autre hypothèse de cette « non-expression » pourrait être en lien avec la formation de corps d'inclusion dans les bactéries. En effet, les cellules bactériennes sont susceptibles de renfermer dans des corps d'inclusion, de grandes quantités de protéines inutiles pour la bactérie sous forme insolubles et dénaturées. Il est possible de purifier ces corps d'inclusion et d'en extraire les protéines, mais il faut ensuite renaturer les protéines. Or cette étape possède un rendement très faible ; c'est pourquoi il est préférable de changer de vecteur d'expression. À titre d'exemple, je pourrais essayer d'exprimer la Diox 1 et/ou la Diox Panais ; qui ne s'expriment pas non plus, en utilisant le plasmide pGEX produisant des protéines de fusion GST régulées par le promoteur *tac*. En ce

qui concerne la Diox Citrus, l'équipe du professeur Mizutani à réussi à identifier des mutations dans ce clone et ils sont finalement parvenu, après modification de la séquence, à caractériser ce gène qui montre aussi une activité pour le féruloyle CoA et le *p*-coumaroyle CoA (communication personnelle).

La présence de quatre dioxygénases différentes (présence d'introns différents) chez *Ruta graveolens* est certainement apparue suite à des phénomènes de duplication génique. Dans le précédent chapitre, je suppose que la conservation de ces gènes au sein de l'espèce pourrait présenter un avantage adaptatif, aboutissant à une redondance de l'activité nécessaire pour la plante, ou au contraire, à une différenciation de l'activité et la formation de nouvelle fonction enzymatique. Après avoir caractérisé les différents candidats codant pour les dioxygénases chez la Rue, je pencherai plutôt pour cette deuxième hypothèse de néo-fonctionalisation ou du subfonctionalisation. Dans le cas de néo-fonctionnalisation, la fonction ancestrale pourrait être la féruloyle CoA hydroxylase (F6'H), à partir duquel aurait émergé la fonction *p*-coumaroyle CoA hydroxylase (C2'H), aboutissant à des enzymes bifonctionnelles telle que la Diox 4. Dans le cas de sub-fonctionalisation, la fonction ancestrale pourrait être la double activité F6'H/C2'H caractérisée chez la Diox 4, qui aurait émergé pour former des enzymes plus spécifiques mono-fonctionnelle telles que la Diox 8 ou la Diox 2 présentant une activée F6'H uniquement. Dans ce cas, je devrai pouvoir identifier un clone chez la Rue qui possède uniquement l'activité C2'H. Ce clone pourrait bien être la Diox 1 par exemple.

Enfin, deux études indépendantes réalisées dans les années 1970 (Ellis and Amrhein, 1971; Gestetner and Conn, 1974) suggéraient que l'umbelliférone était métabolisé par une enzyme, appartenant à la famille des cytochromes P450, réalisant une *ortho*-hydroxylation de l'acide *p*-coumarique. Mon travail démontre que la synthèse d'umbelliférone chez *Ruta graveolens*, est réalisée par une *ortho*-hydroxylation du *p*-coumaroyle CoA par l'intermédiaire d'une enzyme appartenant à la famille des dioxygénases *α*-cétoglutarates-dépendantes. Cependant, il est noté dans la littérature que certaines fonctions enzymatiques sont assurées alternativement, soit par des P450, soit par des dioxygénases en fonction de l'espèce étudiée. Pour ne citer qu'un exemple, chez les Apiacées et *Equisetum arvense*, la [2S]-naringénine est convertie en apigénine par une enzyme appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes, alors que chez les Fabacées tels que *Glycyrrhiza echinita* (Akashi et al., 1998) ou dans le soja (Fliegmann et al., 2010), cette même conversion est assurée par le CYP93B, une enzyme appartenant à la famille des P450. L'utilisation d'un inhibiteur spécifique de

dioxygénase (le Regalis®) appliqué sur les parties aériennes de *Ruta graveolens* m'a permis de mettre en évidence une diminution significative de l'umbelliférone et de tous les composés furocoumariniques chez la Rue (Tableau 28). Cette observation est un bon indicateur pour révéler l'implication d'une dioxygénase dans le métabolisme des furocoumarines chez *Ruta graveolens*. Cependant d'autres analyses doivent être menées afin de montrer si cette activité C2'H est assurée uniquement par une enzyme appartenant à la famille des dioxygénases α-cétoglutarates, ou si d'autres enzymes mono-oxygénases (telles que les P450) sont aussi capables d'assurer cette même fonction. On ne peut donc exclure, que la fonction C2'H soit assurée par un P450 dans d'autres espèces de plante.

Les furocoumarines sont des molécules utilisées en thérapeutique : elles présentent de nombreuses applications en dermatologie (psoriasis et vitiligo) (Llano et al., 2003) ; mais aussi dans le traitement de certains lymphomes cutanés à cellules T (Van Iperen and Beijersbergen Van Henegouwen, 1997). Ces molécules semblent promises à de nouvelles applications thérapeutiques dans le traitement de la dépression (Chen et al., 2007) ou de la β-thalassémie (Viola et al., 2008; AdiÅŸen and GÃ¼rer, 2010). Or la synthèse chimique de ces composés et un processus couteux et la production actuelle de furocoumarines provient essentiellement de l'huile de bergamote (*Citrus bergamia*), cultivée principalement en Italie (Calabre) et en Côte-d'Ivoire. La culture de *Citrus bergamia* étant en pleine décroissante, Milesi *et al* ont recherché de nouvelles espèces de plantes pour la bioproduction de furocoumarines (Milesi et al., 2001). D'après leurs conclusions, *Ruta* pourrait être une plante adaptée à la production importante de ces métabolites (Milesi et al., 2001), du fait de sa capacité de production en furocoumarines nettement supérieure aux autres familles de plantes synthétisant ce type de composé (Apiacée, Fabacée, Moracée).
Il serait intéressant de sur-exprimer de manière stable la Diox 4 (ou C2'H) chez *Ruta graveolens* afin d'évaluer l'augmentation de la teneur en umbelliférone et furocoumarine chez les plantes transformées génétiquement. J'ai lancé des analyses de transformation génétique de Rue qui devront être analysées quand le matériel végétal transformé sera suffisamment développé pour réaliser une extraction des phénylpropanoïdes (Photo 17). L'approche d'ingénierie métabolique pourrait certainement modifier de manière significative les teneurs en furocoumarines, et ainsi augmenter le potentiel de production des furocoumarines par *Ruta graveolens*.

De même, j'ai réalisé la transformation génétique d'*Arabidopsis thaliana* dont la teneur en umbelliférone et en scopolétine devra être analysée lorsque la croissance du matériel végétal permettra de réaliser une extraction des phénylpropanoïdes. *A. thaliana* ne synthétise naturellement pas de furocoumarines, mais il serait intéressant à long terme de reconstituer la voie de biosynthèse des furocoumarines par génie métabolique chez cette plante modèle, lorsque toutes les enzymes de cette voie seront identifiées et caractérisées. De même, il serait alors possible de permettre la production des furocoumarines dans n'importe quel type de plante. De nombreuses applications agronomiques (liées au potentiel allélopathique des furocoumarines), sanitaires (liées a la toxicité des furocoumarines) et industrielles (par la production de ces molécules thérapeutique) pourraient en découler.

II.5 Conclusion

Malgré leur importance dans les processus de défense vis-à-vis des prédateurs, les voies de biosynthèse des furocoumarines linéaires et angulaires restent encore très peu documentées au niveau moléculaire. Les résultats, présentés dans cette seconde partie du chapitre III, décrivent différents systèmes d'expression hétérologue « *in vitro* » et « *in planta* » que j'ai mis en place afin de caractériser des gènes candidats potentiellement impliqués dans la voie de biosynthèse des furocoumarines. La caractérisation dans le système d'expression hétérologue bactérien des dioxygénases α-cétoglutarates identifiées chez *Ruta graveolens,* m'a permis de déterminer l'activité de la Diox 2 (F6'H), de la Diox 4 (F6'H et C2'H) et de la Diox 8 (F6'H). Ce système d'expression est relativement facile et rapide à mettre en place, car il existe de nombreuses souches bactériennes améliorées (ex : *E.Coli* BL21DE3) et de nombreux vecteurs d'expression permettant de moduler le niveau d'expression des protéines (ex : pET28a). Les bactéries sont particulièrement bien adaptées pour la production de protéines solubles telles que les dioxygénases. Elles présentent l'intérêt d'être un système inductible afin de pallier à d'éventuels problèmes de toxicité envers l'hôte de production.

Mon travail s'est essentiellement axé sur la caractérisation fonctionelle de la Diox 4 présentant une double activité (*ortho*-hydroxylation du féruloyle CoA (F6'H) et du *p*-coumaroyle CoA (C2'H)) encore jamais décrite dans la littérature. J'ai montré que cette dioxygénase était sensible à l'inhibition compétitive du psoralène, une furocoumarine synthétisée par la Rue, hautement toxique de fait de ses propriétés intercalentes avec l'ADN, ce qui constitue probablement un système de régulation par rétro-inhibition enzymatique.

Ce travail *in vitro* a été poursuivi *in planta* par comparaison du profil métabolique de la Rue avec le niveau d'expression de la Diox 4 en fonction de l'élicitation aux UV. Une corrélation positive a été montrée entre l'accumulation d'umbelliférone dans les feuilles de *Ruta graveolens* et le niveau d'expression de la Diox 4 dans ces mêmes feuilles. La vaporisation d'un inihibteur spécifique de dioxygénase sur les feuilles de Rue a montré une diminution du profil d'accumulation de toutes les furocoumarines identifiées et confirme donc l'intervention d'une dioxygénase *in planta* situé an amont de la voie de biosynthèse des furocoumarines. Pour compléter cette étude *in vivo*, j'ai réalisé l'expression transitoire de la Diox 4 dans les feuilles de *Nicotiana benthamiana*. L'analyse du profil métabolique des feuilles de tabac surexprimant la Diox 4 a confirmé la double activité F6'H et C2'H de la dioxygénase. Ce système d'expression hétérologue est très intéressant, car il permet d'étudier les protéines candidats avec une certaine réalité biologique liée à l'expression dans un « système plante ». Contrairement à l'expression *in vitro*, l'expression transitoire *in planta* présente l'avantage de caractériser la protéine avec une vision holistique, et prend en compte d'éventuelles interactions biologiques qui ne sont pas contrôlables *in vitro*. C'est une approche systémique qui, me semble t-il, sera de plus en plus développée pour la caractérisation des gènes candidats de la voie de biosynthèse des furocoumarines. L'étape qui oriente la synthèse de furocoumarines à partir de dérivées cinnamiques (*p*-coumaroylec CoA) est donc métabolisée par une dioxygénase α-cétoglutarate-dépendante présentant une activité C2'H/F6'H (Diox4). La caractérisation moléculaire de cette étape constitue donc une avancée majeure dans le projet d'identification des gènes impliqués dans le métabolisme des furocoumarines

III Identification et tentatives de caractérisation des cytochromes P450.

III.1 Recherche de P450 candidats

III.1.1 Les collaborations

III.1.1.1 Collaboration avec l'équipe du professeur Matern

III.1.1.1.1 Présentation

L'équipe du Prof. Matern de l'université de Marburg avait effectué des travaux de recherche sur la culture et l'élicitation des cellules d'*Ammi majus* depuis les années 1980 (Hamerski and Matern, 1988a, , 1988b; Hamerski et al., 1990 a; Hamerski et al., 1990 b). Une collaboration a été entreprise entre nos deux laboratoires en 2005 afin d'identifier et de caractériser de nouveaux candidats dans la voie de biosynthèse des furocoumarines. La première partie de ce projet, réalisée à l'université de Marburg, était destinée à isoler des séquences codantes de P450 exprimées spécifiquement dans les cultures cellulaires d'*Ammi majus* stimulées pour la production de furocoumarines. Les travaux de caractérisation fonctionnelle des différents P450 identifiés et isolés sont ensuite effectués dans notre laboratoire. L'ensemble des travaux d'isolement de séquences de P450 par differential display a été réalisé par les deux doctorantes Specker S. et Kellner S. au sein de l'équipe du Pr. Matern. Une partie de ces gènes a été caractérisée dans notre laboratoire par Romain Larbat durant sa thèse. Les autres gènes m'ont ensuite été confiés afin de déterminer leurs fonctions.

III.1.1.1.2 Méthode d'identification des P450 candidats

Les premières expériences d'élicitation des cultures cellulaires d'*Ammi majus* par les extraits de champignons *Phytophtora megasperma* (*Pmg*) ont mis en évidence l'accumulation de coumarines et de furocoumarines dans le milieu après un délai de 3 à 5H (Hamerski et al., 1990 b). L'équipe du professeur Matern a pour objectif d'isoler les gènes qui entrainent l'accumulation des furocoumarines par une technique de tri d'ARNm (ou differential display) exprimés dans le milieu. Elle est fondée sur trois étapes majeures : la transcription inverse à partir des ARNm aboutissant à l'obtention d'ADNc monocaténaire ; l'amplification par PCR

de ces ADNc ; et la séparation des fragments d'ADNc bicaténaires par électrophorèse. Afin d'isoler les gènes impliqués dans le métabolisme spécifique des furocoumarines, les ARN totaux sont extraits des milieux de culture cellulaire d'*Ammi majus* avec ou sans traitement par les extraits de champignons *Pmg*. Ces ARNm sont utilisés comme matrice pour la Differential-Diplay Reverse Transcription PCR (DD-RT-PCR). Les séquences de P450 sont amplifiées sélectivement en utilisant des amorces dégénérées dirigées contre les régions peptidiques consensus PERF et PFG de P450 définies par Akashi et Schöpfer (Akashi et al., 1997; Schopfer and Ebel, 1998). Enfin, la comparaison des gels électrophorétiques permet d'identifier des bandes d'ADNc différentiellement accumulées dans les échantillons. Cette technique de tri d'ARNm a permis, à l'équipe du professeur Matern, d'isoler sept fragments d'ADNc codant pour des cytochromes P450 spécifiquement exprimés dans les cellules d'*Ammi majus* élicitées. Les séquences complètes de ces P450 ont ensuite été obtenues par des techniques de RACE PCR. Deux de ces P450 appartiennent aux familles des CYP73 et CYP98 (CYP73A43 et CYP98A21) qui codent respectivement une C4H et une C3'H dans le métabolisme des phénylpropanoïdes (Hubner et al., 2003). Un troisième P450 appartenant à la famille des CYP71AJ (CYP71AJ1) a été caractérisé dans notre laboratoire par Romain Larbat (Larbat et al., 2007). Ce P450 code pour la psoralène synthase impliquée dans la voie de biosynthèse spécifique des furocoumarines linéaires. Les cinq autres séquences ont été classées et nommées d'après la nomenclature internationale des P450 basée sur l'homologie de séquences protéiques. Trois d'entre eux appartiennent à la famille des CYP71 largement représentée chez les végétaux supérieurs, avec CYP71D97, CYP71D98 et CYP71AZ1. Un autre gène appartenant à la famille des CYP76, avec CYP76B8, et un dernier appartenant à la famille des CYP82, avec CYP82H1, ont été identifiés. Tous ces gènes sont potentiellement impliqués dans la voie de biosynthèse des furocoumarines étant donné que les 5 gènes sont fortement exprimés suite à une élication.

III.1.1.1.3 Modification de l'extrémité N-terminale de CYP71D98 et CYP71AZ1.

Suite à de nombreux échec de caractérisation des P450 candidats identifiés chez *Ammi majus*, des modifications au niveau de l'extrémité N-terminale ont été entrepris par le laboratoire du Prof. Matern. En effet, les travaux réalisés par Batard et al (Batard et al., 2000) ont montré que le remplacement de l'ancre transmembranaire d'une C4H de blé (CYP73A17) par celle de topinambour (CYP73A1) engendrait une meilleure expression de l'enzyme dans les levures

WAT 11. Cette stratégie a d'ailleurs été reprise avec succès par Larbat et al (Larbat et al., 2007), afin d'obtenir un niveau d'expression détectable de CYP71AJ1 dans les levures WAT 11. Les 30 premiers résidus de la protéine sont destinés à ancrer l'enzyme dans la membrane du réticulum endoplasmique, et n'interviennent pas dans la reconnaissance du substrat (Sagara et al., 1993; Sueyoshi et al., 1995). La protéine chimère qui résulte de cette modification présente donc les mêmes spécificités de substrat. Le succès de cette procédure a conduit Sandra Kellner, du laboratoire de Malburg, à remplacer la séquence N-terminale de CYP71D98 et CYP71AZ1 par celle de topinambour (CYP73A1). Ces séquences seront annotées CYP71D98-nt et CYP71AZ1-nt. Ces produits de PCR nous ont été envoyés et m'ont par la suite été confiés.

III.1.1.2 Collaboration avec l'équipe du professeur Simonsen

Comme avec le laboratoire de l'université de Marburg, une collaboration s'est créée en 2009 avec l'université de Copenhague et l'équipe de du Prof. Henrik Simonsen. Un de leurs axes de recherche concerne la compréhension moléculaire de la voie de biosynthèse de la thapsigargine, un principe actif en cours de validation dans le traitement du cancer de la prostate. Ils disposent de plusieurs banque d'EST privée provenant de *Thapsia garganica*, une plante appartenant à la famille des Apiacées qui synthétise de la thapsigargine (Liu et al., 2006), mais aussi des coumarines, tels que la scopolétine et le 6-methoxy-7-geranyloxycoumarine (Rasmussen et al., 1981; Avato, 1991). Le criblage de cette banque avec les différents CYP71AJ identifiés par RomainLarbat a permis d'identifier un autre membre de cette famille, CYP71AJ5. L'extraction d'ARN de *Thapsia garganica* et l'utilisation d'amorces spécifiques dirigées contre les extrémités du gène CYP71AJ5, leur ont permis d'isoler ce nouveau membre de la famille des CYP71AJ. Ce nouveau gène m'a ensuite été confié pour déterminer sa fonction.

III.1.2 Les bonnes raisons pour étudier ces différents P450 candidats

III.1.2.1.1 CYP71AJ5, le meilleur candidat.

Cette enzyme me semble être le candidat le plus probablement impliqué dans la voie de biosynthèse des furocoumarines. En effet, la première monooxygénase caractérisée dans la

voie de synthèse des furocoumarines est la psoralène synthase (CYP71AJ1) (Larbat et al., 2007), suivie par l'angélicine synthase (CYP71AJ4) (Larbat et al., 2009). Deux étapes similaires de la voie de biosynthèse des furocoumarines sont donc métabolisées par des P450 appartenant à la même sous-famille, les CYP71AJ. Tel qui l'a été décrit dans la voie de biosynthèse des benxazinoides, il est possible que la voie de biosynthèse des furocoumarines dérive intégralement d'un même gène ancestral qui aurait acquis de nouvelles fonctions par des phénomènes de duplications géniques (Frey et al., 2009). Dans ce cas, il est possible que les autres P450 intervenant dans cette voie de biosynthèse appartiennent pour la plupart à la même sous famille de P450, les CYP71AJ.

L'analyse bio-informatique de ce candidat et de cette sous-famille de P450 (CYP71AJ) révèle des pistes intéressantes. La comparaison des pourcentages d'identités et d'homologies peptidiques entre les différents membres de la sous-famille des CYP71AJ indique que les CYP71AJ1/2/3 sont plus éloignés de CYP71AJ4 et CYP71AJ5. En effet, CYP71AJ1/2/3 présente 81-88 % d'identité entre eux et seulement 59-71 % avec CYP71AJ4 et CYP71AJ5 (Tableau 30). CYP71AJ5 est la séquence peptidique la plus éloignée de cette sous-famille de CYP71AJ avec seulement 55-61 % d'identité avec les autres membres. L'arbre phylogénétique généré à partir de ClustalX (Figure 83) montre que CYP71AJ1/2/3 forment un sous-groupe phylogénétiquement à l'intérieur de la famille des CYP71AJ alors que CYP71AJ4 et CYP71AJ5 sont des membres plus éloignés. Un alignement de CYP71AJ5 et des différentes séquences peptidiques des CYP71AJ1/2/3/4 (Figure 84) montrent les sites de reconnaissances du substrat (SRS) défini par Gotoh (Gotoh, 1992) et Larbat (Larbat et al., 2009). Quand on compare les séquences spécifiques SRS de CYP71AJ1 avec les autres membres de la sous-famille, il parait évident que les CYP71AJ1/2/3 sont similaires. Le pourcentage d'identité des différentes séquences SRS varie seulement de 73 à 100 % (Tableau 31). CYP71AJ4 et CYP71AJ5 ont des séquences SRS qui présentent un pourcentage d'identité avec CYP71AJ1 nettement plus bas (30 à 67 %).
Cependant, la comparaison des séquences SRS entre CYP71AJ4 et CYP71AJ5 révèle aussi des différences très importantes (10 à 55 %). D'après les travaux de Larbat *et al*, les SRS2 et SRS3 constitueraient un canal permettant l'entrée du substrat dans l'enzyme. Au niveau du SRS2, il y a une très grande différence de CYP71AJ5 vis-à-vis des autres membres de la sous-famille (30 % d'identité avec CYP71A1 et 10 % avec CYP71AJ4). Ceci suggère que ce canal permet l'entrée d'un substrat différent des autres membres de la sous-famille des CYP71AJ.

L'acide aminé Thr[301] présent dans le SRS4, chez CYP71AJ1, est bien retrouvé dans tous les CYP71AJ, y compris CYP71AJ5 (Figure 84). Cet acide aminé situé dans la poche catalytique de l'enzyme semble être impliqué dans l'activité de la protéine d'après les travaux de modélisation de Romain Larbat. Cependant, l'ensemble de la séquence SRS4 de CYPA71AJ5 est très différent des autres membres de la sous-famille (42-46%) (Tableau 31).

Dans la séquence SRS5, le triplet YFT est très important dans la reconnaissance du substrat. L'acide aminé Thr est important dans ce triplet pour l'activité catalytique de la protéine. Le motif YFT est bien présent chez les CYP71AJ1/2/3 qui métabolisent la marmésine en psoralène. Celui-ci est remplacé par le triplet YIT chez CYP71AJ4 qui métabolise la columbianétine en angélicine. L'Iso[367] de CYP71AJ4 remplace donc la Phe de CYP71AJ1/2/3 dans le triplet YFT. Or ces deux acides aminés sont apolaires et ne doivent certainement pas trop modifier la conformation du site actif.

Le triplet YFT présent chez les CYP71AJ1/2/3 est remplacé par YTP chez CYP71AJ5. La Pro[367] est donc située au même emplacement que la Thr de CYP71AJ1/2/3/4. Cependant, on retrouve bien la Thr dans ce triplet, mais positionnée différemment dans le site actif. Ce changement de position de la Thr, et l'apparition d'un résidu proline doit certainement modifier la conformation du site actif et pourrait indiquer une spécificité de substrat différent de la mamésine ou de la columbianétine.

L'analyse de l'ensemble des données bio-informatiques semble montrer une divergence de spécificité de substrats de CYP71AJ5 par rapport aux autres membres de la famille des CYP71AJ déjà caractérisés.

Il est à noter que chez *Thapsia garganica*, la scopolétine et la 6-methoxy-7-geranyloxycoumarine (Rasmussen et al., 1981) ont déjà été identifiées, mais, à notre connaissance, il n'y a pas de furocoumarines qui aient été décrites chez cette plante.

	CYP71AJ1	CYP71AJ2	CYP71AJ3	CYP71AJ4	CYP71AJ5
CYP71AJ1		88	81	69	61
CYP71AJ2	92		81	71	58
CYP71AJ3	89	89		70	59
CYP71AJ4	80	83	83		55
CYP71AJ5	74	72	73	72	

Tableau 30 : Pourcentage d'identité (en vert) et d'homologie (en bleue) peptidique entre les différents clones isolés de la famille des CYP71AJ.

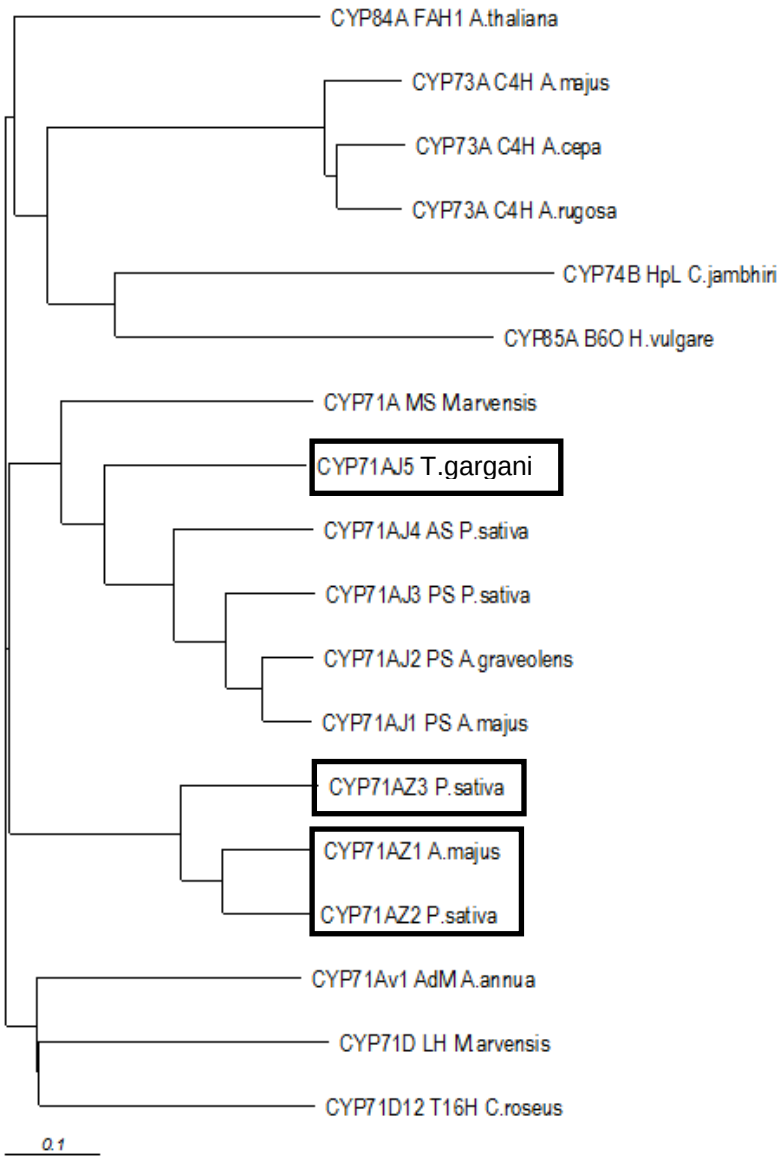


Figure 83 : Arbre phylogénétique de différents P450 caractérisés incluant les gènes candidats appartenant à la famille CYP71AJ et CYP71AZ.

L'arbre est généré à partir de ClustalX et visualisé grâce au logiciel TreeView. J'ai intégré dans cet arbre différents P450 de plantes dont la fonction a déjà été caractérisée. Les gènes encadrés correspondent à CYP71AJ5 et 3 autres P450 candidats identifiés CYP71AZ1, CYP71AZ2 et CYP71AZ3. Les numéros d'accession sur GenBank/EMBL/DDBJ des gènes caractérisés sont les suivants : CYP73A_C4H_A.majus (AAO62904) ; CYP73A C4H A.cepa(AAS48416) ; CYP73A C4H A.rugosa(AAU09021) ; CYP71A MS M.arvensis(ABW86890) ; CYP71D LH M.arvensis(ABQ24001) ; CYP71D12 T16H C.roseus(ACM92061) ; CYP74B HpL C.jambhiri(BAC55161) ; CYP71Av1 AdM A.annua(ABC41927) ; CYP84A FAH1 A.thaliana(CAA18128) ; CYP85A B6O H.vulgare(ABH01181) ; CYP71AJ4 AS P.sativa(C0SJS3) ; CYP71AJ3 PS P.sativa(C0SJS2) ; CYP71AJ2 PS A.graveolens(C0SJS4) ; CYP71AJ1 PS A.majus (Q6QNI4).

```
CYP71AJ1   MKMLEQNPQY LYFFSLFLVT IFLYKWLTLK KTPLKNLPPS PPQYPIIGNL HQIGPDPQAS LRDLAQKYGP   70
CYP71AJ2   .......... .......... V.V..L.... ...S...... ..R....... ........H. .....L....   70
CYP71AJ3   .......... ....P..... .......-V. ...S...... ..RL...... ........I. .....RE...   69
CYP71AJ4   .......... .......... .......-A. ...S...... ..RL...... ......LHI. .....R....   69
CYP71AJ5   -M.MD.QT-- -L.L..CF.L V...I..STS ..TG...... .QKL...... ..VNQ..HI. ..S..K....   66
                                                      SRS1
CYP71AJ1   LMFLKFGTVP VLVVSSADAA REALKTHDLV FADRPYSSVA NKIFYNGKDM VFARYTEYWR QVKSICVTQL  140
CYP71AJ2   ..S....... .......... ..V......I .......... ..V....... .......... .......... 140
CYP71AJ3   V.H....S.. ........G. ..IF...... .......... .R.....R.. .......... ....T..... 139
CYP71AJ4   ..Q.QL.RI. .......E.T ..V.....V. .SQ..IT.AI D.LC.K.R.V A.S..S.... ..R.T..... 139
CYP71AJ5   V.Q.H..SI. .......... M.IM.....A ..N..N..IW D........V ..P.S.... .......L.. 136

CYP71AJ1   LSNKRVNSFH YVREEEVDLL VQNLENSHSK VANLTELLIE VTGNVVCRVS VG-------S GDKVDSYKIL  203
CYP71AJ2   .........Q N......... ...I...C.. .I........ .......K.. ..-------. .......... 203
CYP71AJ3   ..V...Q... N......A.. LD.I...K.. .I..S.M... L.......AA L.-------. .YN.....S.  202
CYP71AJ4   ...S..H... NI.....A.. I..I...A.E .I..G.Q..Q L.R....... ..SE---YL. .H.GKL.QK.  206
CYP71AJ5   .....R..Q T......A.. .E.VRE.G.. TV..S..FYT LLS...S.IA L.RKYAITTE .G.EN.F.E.  206
           SRS2                                       SRS3
CYP71AJ1   ILEIMDMLGY SRSIEDFFPL LGWVDWLTGL RGKVAEAAKG VDTFLEGVLK EHLSTTG--- -SKYNDFVSI  269
CYP71AJ2   .....E.... .........M F.......... .....K.... ..D....... ...TARASNN A.AD......  273
CYP71AJ3   L.Q....... .........S ......I... K...EK..N. ..A....... N.TNPSTS-- -.ANK.....  269
CYP71AJ4   LA.VTE..A. TY..G..I.. .......S.S KA..EKT..E ..A....A.R D.IK.MASNK G.AND..L..  276
CYP71AJ5   FQN.AQLI.. -F.VG.YI.W .F.I.SVN.. K.R.EK..NE A.L.....I. D.SIALDN-- GVSTD.LLYN  273
                      SRS4
CYP71AJ1   LLEIQEADAG S--SMDNECI KSLIWDMLGA GTETISTALE WTLAALIKNP DAMFKLQNEV REIGKGKSKI  337
CYP71AJ2   .......... .--T...... .......... .......... ..........  ...L...... .......... 341
CYP71AJ3   .......... .--...K... .......... .....A.... ..IG...S.. ...S...K.. ........R.  337
CYP71AJ4   ....R..... .--TL.E... .AIV...ILG ....T..T.. .IV..I.... .V.....K.. .......... 344
CYP71AJ5   .....KQ.TN .AF.I.KDS. .GV.LN.FFD ..D.T.AV.. ..M.....H. .I.C..K... ....R..... 343
           SRS5
CYP71AJ1   SEADLVKMNY LQAVMKESMR LYFTAPLLVP REARQDIKFM GYDISSGTQV LINAWAIARD PLLWDKPEEF  407
CYP71AJ2   .....G..T. .......... .......... ..S...V... .....A.... ...V...... .S..E.....  411
CYP71AJ3   E.G....D. .K........ .......... ....V..... ....K..... .......... .SS..N....  407
CYP71AJ4   E.V....... .K........ ..I..-F.L. ...K..V.L. .......... ...T.ET... .S...N....  413
CYP71AJ5   .GD..E..H. .K..V..... V.TPV....A ...M..V.V. ...VKA.... ........T. .A...N..K.  413
                                                                                SRS6
CYP71AJ1   RPERFLNSPI DYKGFHYEFL PFGAGRRGCP GIQFAMCINE LVVANLVHKF NFELPDGKRL EDLDMTAASG  477
CYP71AJ2   ........H. .....N..Y. .......... ......AV.. .....VI... .......E.. ....V..... 481
CYP71AJ3   .......... .....N..YI .......... .....ISV.. .....V.N.. .......... .EM...ST.  477
CYP71AJ4   .......... ....L...Y. ...G...... ......AV.. .A...V.Y.. D.KM...E.F ....SGVP.  483
CYP71AJ5   I......N.. ....L.F..I .......... ...Y..A... .AL.....I. D.A....R.. .....SET.  483

CYP71AJ1   ITLRKKSPLL VVARPHV 494
CYP71AJ2   ......---- ------- 487
CYP71AJ3   ..FH...--- ------- 484
CYP71AJ4   .S.YR.---- ------- 489
CYP71AJ5   M..H...... .I.TSR. 500
```

Figure 84 : Alignement des séquences peptidiques des différents membres de la famille des CYP71AJ1-4 et du nouveau CYP71AJ5 isolé chez *Thapsia garganica*.

En rouge sont encadrées les six séquences de reconnaissance du Substrat (SRS) comme décrit par Larbat R. (Larbat et al., 2009).

A	SRS1	SRS2	SRS3	SRS4	SRS5	SRS6
CYP71AJ2	95	90	91	100	100	85
CYP71AJ3	89	80	73	92	100	69
CYP71AJ4	37	40	36	61	67	38
CYP71AJ5	68	30	55	46	58	62

B	SRS1	SRS2	SRS3	SRS4	SRS5	SRS6
CYP71AJ5	47	10	55	42	33	31

Tableau 31 : Pourcentage d'identité des séquences peptidiques codant pour les SRS des différents membres de la famille des CYP71AJ.

Les séquences SRS ont été identifiées sur la Figure 84. Ce tableau présente les pourcentages d'identités entre la psoralène synthase (CYP71AJ1) et les autres membres de la famille des CYP71AJ (Tableau A). Le tableau B présente de même le pourcentage d'identité entre l'angélicine synthase (CYP71AJ4) et CYP71AJ5.

III.1.2.1.2 CYP71AZ1, un grand potentiel.

CYP71AZ1 définit une nouvelle sous-famille de cytochrome chez les CYP71. Cette famille est la plus grande famille de P450 décrite chez les végétaux (plus de 440 séquences identifiées d'après http://drnelson.uthsc.edu/CytochromeP450.html), mais relativement peu de CYP71 ont à ce jour des fonctions caractérisées. À titre d'exemple, on trouve dans cette famille de P450 la menthofurane synthase avec CYP71A, la tabersonine 16-hydroxylase, la limonène 3-hydroxylase dans la sous-famille des CYP71D, et bien entendu la psoralène synthase ou l'angélicine synthase dans la sous-famille des CYP71AJ.

Cette famille de P450 appartenant à CYP71 catalyse donc deux étapes clefs de la synthèse de psoralène et d'angélicine. De plus, la comparaison du mécanisme réactionnel de cyclisation de pulegone en menthofurane chez *Mentha pulegium* (Bertea et al., 2001) avec la réaction catalysée par la marmésine synthase présente certaines analogies (Figure 85). Cette réaction est catalysée chez la menthe par une enzyme appartenant à la famille des P450, la menthofurane synthase (CYP71A) (Bertea et al., 2001). Cette analogie réactionnelle m'amène à faire l'hypothèse que la marmésine synthase pourrait aussi être un P450 appartenant à la famille des CYP71 telle que CYP71AZ1.

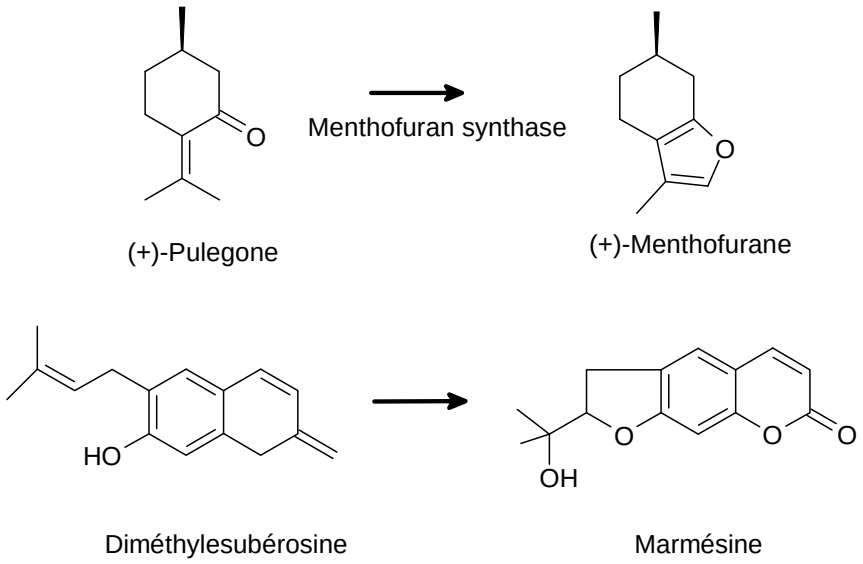


Figure 85 : Comparaison du mécanisme de cyclisation lors de la synthèse de menthofurane et de la marmésine.

De plus, ce gène est particulièrement intéressant de par son profil d'expression. En effet, des analyses cinétiques et quantitatives du niveau d'expression de CYP71AZ1 montrent que son profil d'expression ressemble au profil d'accumulation des furocoumarines après addition de l'éliciteur dans le milieu (communication personnelle). L'expression de ce gène précéderait l'accumulation de psoralène, tout comme cela a été démontré pour CYP71AJ1 (Figure 86) (Larbat et al., 2007).

En complément de CYP71AZ1, j'ai identifié deux autres clones de CYP71AZ1 chez *Patinaca sativa* (CYP71AZ2 et CYP71AZ3). À l'aide de CYP71AZ1 isolés chez *Ammi majus* par l'équipe du Prof. Mattern j'ai réalisé une recherche par tBlastn dans la banque EST-454 de Panais que nous avons fait réaliser Les comparaisons des identités et homologies peptidiques indiquent que CYP71AZ1 est plus proche de CYP71AZ2 (81 % d'identité) que de CYP71AZ3 (71 %) (Tableau 32). L'arbre phylogénétique construit à partir de différents P450 caractérisés à ce jour, montre que les deux gènes CYP71AZ1 et CYP71AZ2 se regroupent ensemble alors que CYP71AZ3 semble diverger (Figure 83). L'analogie est flagrante avec la famille des CYP71AJ représenté aussi dans ce même arbre phylogénétique. Les

CYP71AJ1/2/3 codant pour la psoralène synthase se regroupent tous ensemble alors que CYP71AJ4 codant pour l'angélicine synthase diverge légèrement de ce cluster (Figure 83).

De plus, quand on analyse les analogies au niveau des SRS (SRS1/4/5/6), CYP71AZ1 présente 73 % d'identité avec CYP71AZ2 et seulement 61 % d'identité avec CYP71AZ3. De même, les comparaisons de séquences entre CYP71AZ2 et CYP71AZ3 montrent plus de 73 % d'identité alors que leurs séquences SRS ne montrent que 63 % d'identité. Il est intéressant de noter que CYP71AZ1 a été identifié chez *Ammi majus* qui ne synthétise qu'un seul type de furocoumarines : les furocoumarines linéaires. Alors que CYP71AZ2 et CYP71AZ3 ont été identifiés chez *Pastinaca sativa* qui produit des furocoumarines linéaires et angulaires. La présence de deux copies du même gène, présentant des différences au niveau des SRS, chez *Pastinaca sativa* peut laisser supposer que l'une des copies soit impliquée dans la voie de biosynthèse des furocoumarines linéaires et l'autre copie dans la voie de biosynthèse des furocoumarines angulaires.

Une extraction d'ARN de *Pastinaca sativa*, ainsi que des couples d'amorces ont été commandés afin de récupérer ces nouveaux clones codant pour des cytochromes P450 de la famille des CYP71AZ.

	CYP71AZ1	**CYP71AZ2**	**CYP71AZ3**
CYP71AZ1		81	71
CYP71AZ2	96		73
CYP71AZ3	84	85	

Tableau 32 : Comparaison des identités et homologies peptidiques des clones CYP71AZ2 et CYP71AZ3 identifiés dans la banque d'EST de *Pastinaca sativa* et du clone CYP71AZ1 identifié chez *Ammi majus*.

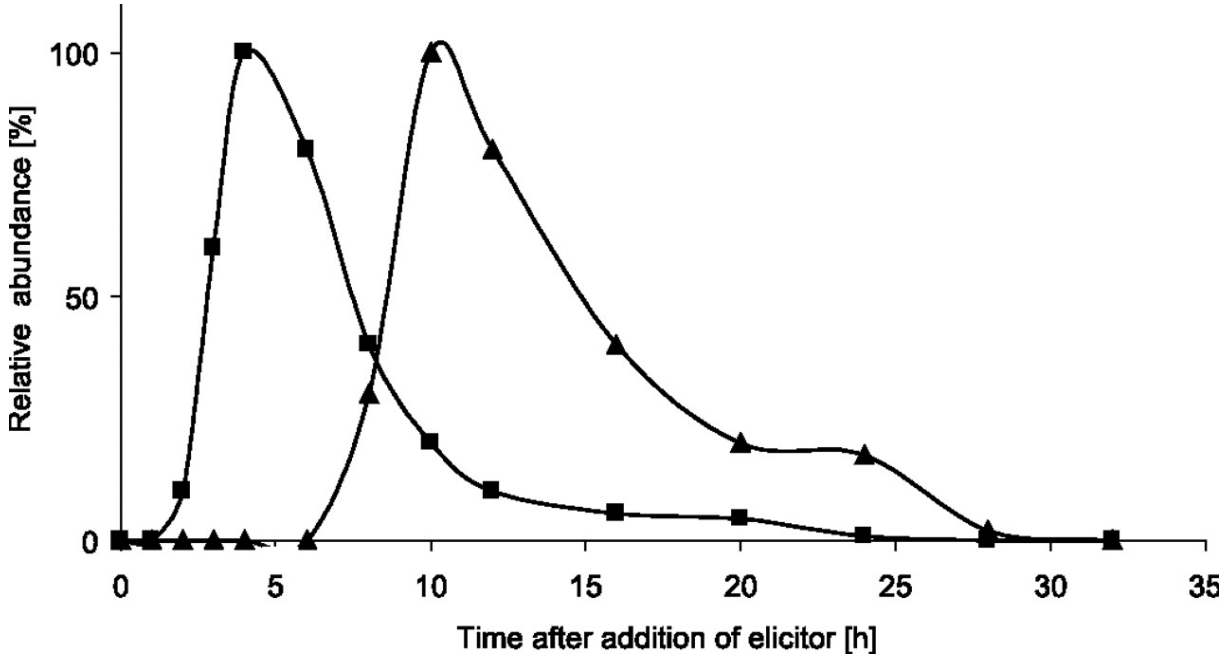


Figure 86 : Comparaison du niveau d'expression relatif de *CYP71AJ1* par rapport au profil d'accumulation du psoralène après induction du milieu de culture cellulaire d'*Ammi majus* par *Phytophtora megasperma (Larbat et al., 2007).*

III.1.2.1.3 CYP71D97 et CYP71D98, les outsiders

CYP71D97 et CYP71D98 sont des candidats intéressants dans la mesure où ils appartiennent eux aussi à la grande famille des CYP71. Afin d'apprécier une potentielle implication de ces gènes candidats dans la synthèse des coumarines et furocoumarines, je réalise une analyse de comparaison de séquence pour définir si ces gènes candidats présentent des homologies avec des gènes identifiés et/ou caractérisés. La comparaison est réalisée avec les séquences disponibles dans les banques de données EST publiques (Genbank). CYP71D97 présente 54 % d'identité avec CYP71D55, une premnaspirodienne oxygénase qui catalyse l'hydroxylation de nombreux sesquiterpènes (Takahashi et al., 2007) et 45 % d'identité avec CYP71D9 une flavonoïde 6-hydroxyalse. CYP71D9 métabolise avec une spécificité et une activité moins importante la liquiritigénine, la lutéoline, la naringénine, l'apigénine et l'ériodictyole (Latunde-Dada et al., 2001). Cette enzyme présente un large spectre d'activité et les flavonoïdes dérivent aussi de la voie de biosynthèses des phénylpropanoïdes. Ce sont des molécules relativement proches structurellement des furocoumarines. La sous-famille des CYP71D semble être une famille de gène intéressante dans le cadre de mes recherches.

CYP71D98 présente 68 % d'identité avec CYP71D97 et 55 % d'identité avec CYP71D55.

III.1.2.1.4 CYP73A43 (C4H) et CYP98A21 (C3'H), une alternative à la C2'H ?

Dans la voie de biosynthèse des phénylpropanoïdes, les acides cinnamiques subissent des hydroxylations en différentes positions du noyau phénol. Ce noyau phénol est dans de nombreux cas hydroxylés par des enzymes appartenant à la famille des P450. Tel est le cas pour la cinnamate 4-hydroxylase (C4H appartenant à la famille des CYP73A), pour la p-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H appartenant à la famille des CYP98) ou pour la féruloyle 5 hydroxylase (F5H appartenant à la famille des CYP84) (Meyer et al., 1996; Schoch et al., 2001b) (Figure 87).

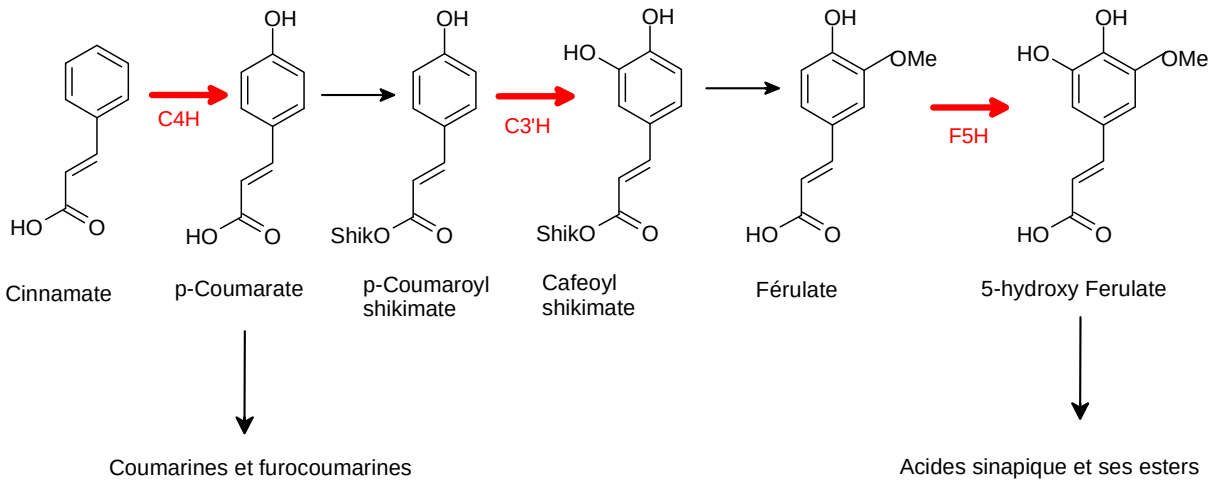


Figure 87 : Représentation des hydroxylations des acides cinnamiques catalysées par la cinnamate 4-hydroxylase (C4H), la p-coumaroyleshikimate/quinate 3'-hydrolxylase (C3'H) et la ferulate 5-hydroxylase (F5H).

D'après mes travaux effectués sur les dioxygénases et présentés dans les précédents chapitres, l'hydroxylation en position 2 du noyau phénol des acides cinnamiques est effectuée par une enzyme appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes tout du moins chez *Ruta graveolens*, mais aussi chez *Arabidopsisi thaliana* d'après les travaux de Kai (Kai et al., 2008). Néanmoins, cette hydroxylation pourrait aussi être effectuée par une enzyme appartenant à la famille des P450 chez d'autres espèces de plantes ou en collaboration avec la dioxygénase (C2'H). Cette dualité entre P450 et dioxygénase a souvent été décrite au sein

d'une même vois de biosynthèse. C'est notamment le cas dans la voie de biosynthèse des flavonoïdes ou la synthèse de l'apigénine est réalisé par une dioxygénase chez la Apiacées (Bredebach et al., 2011) et par un P450 chez les Fabacées (Fliegmann et al., 2010). D'autres exemples ont déjà été rapportés notamment dans la voie de biosynthèse des gibbérellines avec la gibbérelline 7-oxydase (Hedden, 1999). Au regard, des hydroxylations effectué par la C4H, la C3'H ou la F5H, on peut supposer que cette dualité entre dioxygénase et P450 existe aussi dans la voie de biosynthèse des coumarines et furocoumarines au niveau de la C2'H. CYP73A43(C4H) et CYP98A21 (C3'H) pourraient être des enzymes bifonctionnelle capables d'assurer aussi cette fonction C2'H à l'image de la bifonctionnalité décrite pour la Diox 4 dans le chapitre précédent.

III.1.2.1.5 CYP82H1 et CYP76B8, les peu probables

CYP82H1 définit comme CYP71AZ1, une nouvelle sous-famille de cytochromes chez les CYP82. D'autres membres appartenant à la sous-famille CYP82H ont été identifiés depuis chez *Vitis vinifiera*. Récemment, la première fonction de la famille des CYP82 a été caractérisée comme étant impliquée dans la voie de biosynthèse des homoterpènes (CYP82G1) chez *Arabidopsis thaliana* lorsque celle-ci est prédatée par des herbivores (Lee et al., 2010; Tholl et al., 2011). Peu d'éléments sont donc disponibles concernant cette famille de P450.

L'analyse des banques de données de NCBI indique que CYP76B8 présente 57 % d'identité avec CYP76B6 (géraniol 10-hydroxylase) de *Catharantus roseus* impliqué dans le métabolisme des terpènes (Collu et al., 2001) et 53 % d'identité avec CYP76B1 impliqué dans le métabolisme des herbicides de la famille des phénylurés (Batard et al., 1998; Robineau et al., 1998).

Néanmoins, ces deux gènes candidats ont été exprimés suite à l'élicitation des cultures cellulaire d'*Ammi majus* et par ce fait, ils constituent des candidats à prendre en compte que je tenterai de caractériser.

III.2 La levure : un système plébiscité, mais pas optimisé pour des P450 de plantes « exotiques »

Les cytochromes P450 microsomaux isolés à partir de culture cellulaire d'*Ammi majus* par l'équipe du professeur Matern de l'université de Marburg ont été confiés à notre laboratoire afin de réaliser la caractérisation de ces gènes. Parmi les huit gènes candidats identifiés par la technique du différentiel display, trois d'entre eux (CYP73A43, CYP98A21 et CYP71AJ1) ont déjà été caractérisés *via* un système d'expression hétérologue utilisé depuis longtemps au laboratoire, la levure. Ce système d'expression a de nombreuses fois fait ses preuves dans la littérature, également dans notre laboratoire avec l'expression de la C4H (CYP73A1) de topinambour. Les cinq autres gènes identifiés chez *Ammi majus* m'ont ensuite été confiés. Il en a été de même avec le CYP71AJ5 isolé chez *Thapsia garganica* par l'équipe du professeur Simonsen de l'université de Copenhague. Auparavant, certains P450 (CYP71AZ1-nt et CYP71D98-nt) ont subi un remplacement de l'ancre transmembranaire en vue d'améliorer le niveau d'expression. J'ai donc tenté d'exprimer ces 6 gènes dans la levure en faisant de nombreuses mises au point et j'ai réalisé un criblage biochimique large pour l'ensemble de ces gènes candidats.

III.2.1 Clonage dans le plasmide pYeDP60

Le plasmide pYeDP60, qui provient du Laboratoire d'Ingénierie des Protéines Membranaires du CNRS (Gif/Yvette), a été spécialement construit pour l'expression de P450 dans une série de souches de *Saccharomyces cerevisiae* (WAT 11, WAT 21, WAT(R)). C'est un plasmide non intégratif qui nécessite une pression de sélection pour subsister. Les séquences CYP76B8 et CYP82H1 ont été clonées dans le vecteur pYeDP60 par A. Hehn et R. Larbat durant la thèse de ce dernier en 2006. De même, CYP71AJ5 m'a été confié après qu'il soit cloné dans le pYeDP60. En ce qui me concerne, j'ai cloné les produits de PCR de CYP71D97, CYP71D98-nt et CYP71AZ1-nt dans le pYeDP60 après les avoir amplifiés par PCR en ajoutant aux extrémités des sites de restriction (Tableau 33). J'obtiens donc des produits de PCR avec des sites de restrictions flanqués de part et d'autre de l'ORF du P450 candidats qui sont compatibles avec la cassette de clonage présente sur le plasmide d'expression pYeDP60.

Amorces	Séquences	Sites de restrictions
CYP71D97-F	5'-CCGGATCCATGGCTCTCCAGTTTGTGCC-3'	*Bam*HI
CYP71D97-R	5'-CCGGATCCATGGCTCTCCAGTTTGTGCC-3'	*Eco*RI
CYP73A1-D	5'-GGATCCATGGACCTCCTCCTCATAGA-3'	*Kpn*I
CYP71D98-R	5'-CCGAATTCAATCTTCAGCTACGATTCGCCCTT-3'	EcoRI
CYP73A1-D	5'-GGATCCATGGACCTCCTCCTCATAGA-3'	KpnI
CYP71AZ1-R	5'- GGCCATGGAATACTTTTGTATATACAAAATC -3'	KpnI

Tableau 33 : Amorces utilisées pour amplifier les séquences codantes pour les cytochromes P450 ; CYP71D97, CYP71D98-nt et CYP71AZ1-nt.

Les différents gènes, après avoir été amplifiés, clonés dans le pCR8/TOPO et séquencés ont été introduits dans le plasmide pYeDP60. Pour cela, le plasmide pYeDP60 et les plasmides pCR8 recombinants sont digérés au préalable par un couple d'enzyme comme décrit dans le Tableau 33. La ligation des gènes d'intérêts CYP71D97, CYP71D98-nt et CYP71AZ1-nt dans le plasmide pYeDP60 se fait suivant le protocole décrit par le kit T4 DNA ligase d'Invitrogen. Le produit de ligation est ensuite utilisé pour transformer les bactéries GeneHogs d'Invitrogène. Puis, le plasmide pYeDP60 recombinant est amplifié par culture bactérienne avant d'être extrait et vérifié par digestion enzymatique. Les plasmides sont ensuite envoyés à séquencer.

III.2.2 Transformations des levures

Les six constructions plasmidiques pYeDP60 recombinants, contenant les gènes d'intérêts CYP71AJ5, CYP71D97, CYP71D98-nt, CYP71AZ1-nt, CYP76B8 ou CYP82H1 sont introduites indépendamment dans la souche de levure WAT 11 suivant la méthode de transformation décrite dans le chapitre « Matériel et méthodes ». Un étalement sur milieu sélectif SGI est ensuite effectué pour chaque transformation afin d'isoler une colonie de levures transformées, qui sera cultivée pour exprimer le cytochrome P450 candidat.

III.2.3 Amélioration de l'expression des P450 dans la levure

L'expression hétérologue de P450 végétaux en système levure est une technique largement documentée dans la littérature. Le protocole d'expression et de préparation des microsomes a

été mis au point par Pompon *et al* (Pompon et al., 1996a) et Diesperger (Diesperger et al., 1974). Le protocole initial décrit selon la méthode de Pompon a été optimisé par Jiang et Morgan en comparant les niveaux d'activités de la cinnamate 4-hydroxylase (C4H) et de la ferulate 5-hydroxylase (F5H) (Jiang and Morgan, 2004). Ils ont apporté des modifications au niveau de la composition des milieux et des protocoles de culture des levures. En effet, de nombreux facteurs influencent l'activité des P450 *in vivo* tels que le pH intracellulaire, la concentration en enzymes ou en NADPH réductases, ainsi que la teneur en oxygène (Jiang and Morgan, 2004). De nombreux paramètres peuvent donc être modifiés afin d'améliorer l'expression de chaque P450 dans la levure. Après de nombreuses tentatives infructueuses d'expression suivant le protocole décrit par Pompon, j'ai modulé différents paramètres de culture pour toutes les levures transformées comportant les constructions plasmidiques qui m'ont été confiées. Mes travaux ont porté d'une part sur l'amélioration du niveau d'expression des P450 candidats dans la levure, et d'une autre part sur l'amélioration de l'efficacité de la préparation de microsome.

III.2.3.1 Amélioration du niveau d'expression

Certaines protéines hétérologues présentent des propriétés thermolabiles. Elles sont sensibles à la chaleur et présentent un risque de dégradation lorsqu'elles se trouvent dans un milieu à température ambiante. Pour palier à ce risque, j'ai opté de jouer sur les températures de culture des levures. D'autres protéines présentant quand à elle des propriétés de cytotoxicités entrainent une altération des cellules. Elles sont donc néfastes pour l'organisme les produisant, la levure, lorsqu'elles sont exprimées en trop grandes quantités. Pour moduler la production de protéine produite par la levure il est possible de varier la durée d'induction. Enfin, certaines protéines présentent ces deux propriétés et sont à la fois thermolabiles et cytotoxiques. Il faut donc trouver une adéquation entre la température et la durée d'induction afin d'augmenter la synthèse des protéines dans le système levure.

III.2.3.1.1 T°C induction

Le métabolisme des levures est optimal pour une température avoisinant les 30 °C. C'est à cette température que la croissance des levures est la plus importante. Le risque de dégradation des protéines est cependant envisageable à cette température. Pour palier à ce

risque des essais d'inductions avec l'ensemble des gènes P450 candidats ont été réalisés à 18 °C, 22 °C et 28 °C.

III.2.3.1.2 Durée induction

La durée d'induction est un paramètre qui influence directement la quantité de protéines produites par la levure. Cette durée d'induction doit être corrélée à la température d'induction. En effet, le métabolisme des levures est ralenti lorsqu'on diminue la température de culture.

Afin de limiter le risque de dégradation des P450, j'ai fait varier les durées d'induction pour une température constante soit de 28 °C soit de 18 °C. Il a fallu trouver un juste milieu entre une durée d'induction trop courte ne permettant pas l'expression suffisante de protéines, et une durée d'induction trop longue. En effet, le milieu de culture lors de l'induction n'est pas un milieu sélectif et ne permet pas le maintien du plasmide non intégratif au sein des levures (pression de sélection inexistante). Jiang et Morgan ont montré qu'une perte d'activité du P450 exprimé était notable dès que l'induction dépassait 6 h (Jiang and Morgan, 2004). Le risque de toxicité de la protéine pour les levures est aussi à prendre en compte si le niveau de protéine exprimée est trop important. Pour les différents gènes d'intérêt, j'ai testé des durées d'induction de 3 h, 6 h, et 15 h à une température de 28 °C ou de 18 °C.

III.2.3.2 Amélioration de la préparation de microsomes

Si les conditions d'expressions sont importantes, la purification (même partielle) des protéines membranaire est une étape essentielle pour récolter des protéines fonctionnelles.

III.2.3.2.1 Broyage mécanique automatisé des levures

Pompon recommande de réaliser un broyage des levures à l'aide de billes que l'on incorpore dans la solution de levures. Puis une agitation manuelle de 1 min est effectuée par l'opérateur. Cette étape est renouvelée cinq fois, entrecoupée par une pause de 1 min où le prélèvement est placé dans la glace.

Pour des raisons de reproductibilité et d'efficacité, j'ai opté pour un broyage mécanique automatisé à l'aide du Vibro Broyeur Retsch MM40. La fréquence de broyage est effectuée au 1/30eme. J'ai répété 3 fois ce broyage d'une durée de 1 min.

III.2.3.2.2 Précipitation par ultracentrifugation

La précipitation des microsomes n'est plus effectuée de manière chimique par l'ajout de NaCl (concentration finale 0,15M) et de PEG 4000 (concentration finale 0,1 g/ml) dans le milieu, mais elle a été réalisée par une ultracentrifugation à 100 000 g effectuée à l'aide d'une ultracentrifugeuse (Scientific Sorvall WX ultra 80). L'acquisition de cette ultrcentrifugeuse au laboratoire au cours de ma thèse ma donc permis de précipiter mes protéines différemment.

III.2.3.3 Composition des milieux

Les milieux et procédures de cultures habituellement utilisés depuis l'optimisation effectuée par Pompon (Pompon et al., 1996a) sont les suivants (Tableau 34) :

	SGI (g/l)	YPGE (g/l)
Yeast extract		10
Bactopeptone		10
Hydrolysat de Caséine	1	
Bactocasaminoacids w.o aminoacids	6,7	
Tryptophane	40	
Glucose	20	5
Éthanol		3 %(V/V)

Tableau 34 : Composition des milieux de culture selon le protocole décrit par Pompon pour les levures transformées avec le plasmide pYeDP60.

Pompon recommande de cultiver les levures dans un milieu sélectif SGI, puis de les transférer dans un milieu non sélectif contenant du glucose et de l'éthanol (3 %) afin d'augmenter la quantité de levures tout en réprimant la synthèse des protéines d'intérêts grâce au glucose. Enfin, il conseille d'induire la production de protéine par un ajout de 20 ml de galactose à 200 mg/L. L'ajout de galactose se fait en théorie après épuisement complet du glucose dans le

milieu YPGE et permet ainsi l'expression concomitante de la réductase AtHR1 et du P450 d'intérêt. Ce protocole a été utilisé pour tenter d'exprimer tous les P450 qui m'avaient été confiés, mais sans succès.

D'après les analyses effectuées par Jiang et Morgan, le transfert des levures dans le milieu de culture non sélectif YPGE provoque une perte très importante du plasmide pYeDP60 entrainant une diminution de l'expression des protéines d'intérêt. J'ai donc modifié la procédure de culture ainsi que la composition des milieux suivant certaines recommandations apportées par Jiang et Morgan (Jiang and Morgan, 2004) (Tableau 35).

	SGI	SGI modifié	YPL modifié
Yeast extract (g/l)			10
Bactopeptone (g/l)			10
Hydrolysat de Caséine (g/l)	1	5	
Yeast nitrogen base W.O aminoacids (g/l)	6,7	6,7	
Tryptophane (mg/l)	40	40	
Fe (III) (µM)			40
Acide aminolevulinique (mM)			1
Glucose (g/l)	20	20	20

Tableau 35 : Composition des milieux de culture du SGI, du SGI modifié et du YPL modifié suivant les recommandations de Jiang et Morgan (Jiang and Morgan, 2004).

Basé sur l'hypothèse que la source d'azote est certainement un facteur limitant du rendement de croissance des levures, le milieu sélectif SGI modifié contient une quantité supérieure en hydrolysats de caséine. De plus, j'ai rajouté dans le milieu d'induction YPL des précurseurs du groupement héminique (acide aminolevulinique et Fe(III)). Le nouveau protocole de

culture consiste à cultiver les levures pendant 24 h dans la solution de SGI modifiée à 28 °C, puis à renouveler ce milieu une seconde fois pour une incubation de 24 h à 28 °C. Ce renouvellement du milieu permet de combler la consommation des nutriments essentiels à la croissance des levures. Au lieu d'introduire le galactose dans la solution de SGI modifiée, les levures sont sédimentées par une centrifugation à 7000g et sont suspendues de nouveau dans le milieu d'induction YPL modifié, puis incubées pendant 6 h à 18 °C. Cette étape de centrifugation permet de se débarrasser complètement du glucose initialement présent dans la solution de SGI modifié.

III.2.3.4 Analyse de l'expression fonctionnelle des P450

Avant de réaliser un criblage fonctionnel, il est souhaitable de vérifier le niveau d'expression des P450. Les P450 ont la particularité de présenter une absorption intense dans le spectre visible à 450 nm lorsque le fer de l'hème, à l'état réduit, est complexé au monoxyde de carbone. La quantification des P450 exprimés se fait donc par le biais d'une analyse différentielle du spectre d'absorbance à 450 nm en présence de monoxyde de carbone. Cette méthode mise au point par Omura et Sato (Omura and Sato, 1964) permet de déterminer la quantité de P450 fonctionnelles exprimées dans la culture de levures. J'ai obtenu des pics caractéristiques d'absorbances à 450 nm avec CYP71D97 (Figure 88 A et B) et CYP71AZ1-nt (Figure 88 C). Les différentes conditions de culture me permettant d'obtenir l'expression de CYP71D97 et CYP71AZ1-nt sont récapitulées dans le Tableau 36). Pour CYP71D97, il semblerait que l'utilisation du milieu modifié améliore légèrement l'expression de celui-ci (1,4 µM contre 1,1 µM). Le spectre d'absorbance de CYP71AZ1-nt a seulement été visible lors de mes premiers essais d'optimisation du niveau d'expression, c'est-à-dire lorsque j'avais réalisé une induction à 18 °C pendant 15 h dans le milieu de culture décrit par Pompon.

Les cytochromes P450 de plantes sont des enzymes membranaires dont l'expression dans le système levure est extrêmement variable. Il ne semble pas exister de protocole universel qui permet à coup sûr de produire des protéines fonctionnelles. En outre, l'absence d'expression détectable pour les autres P450 peut résulter d'une synthèse inefficace (problème au niveau transcriptionnel ou traductionnel), d'une dégradation importante de l'enzyme ou encore de la limite de détection du spectrophotomètre. Une expression même faible peut permettre une métabolisation de molécules. Les travaux expérimentaux de Larbat l'ont démontré avec CYP71AJ1. C'est pour cette dernière raison que même en l'absence de spectre d'absorbance à

450 nm, j'ai testé l'activité de mes microsomes en réalisant un criblage des substrats de la voie de biosynthèse des furocoumarines.

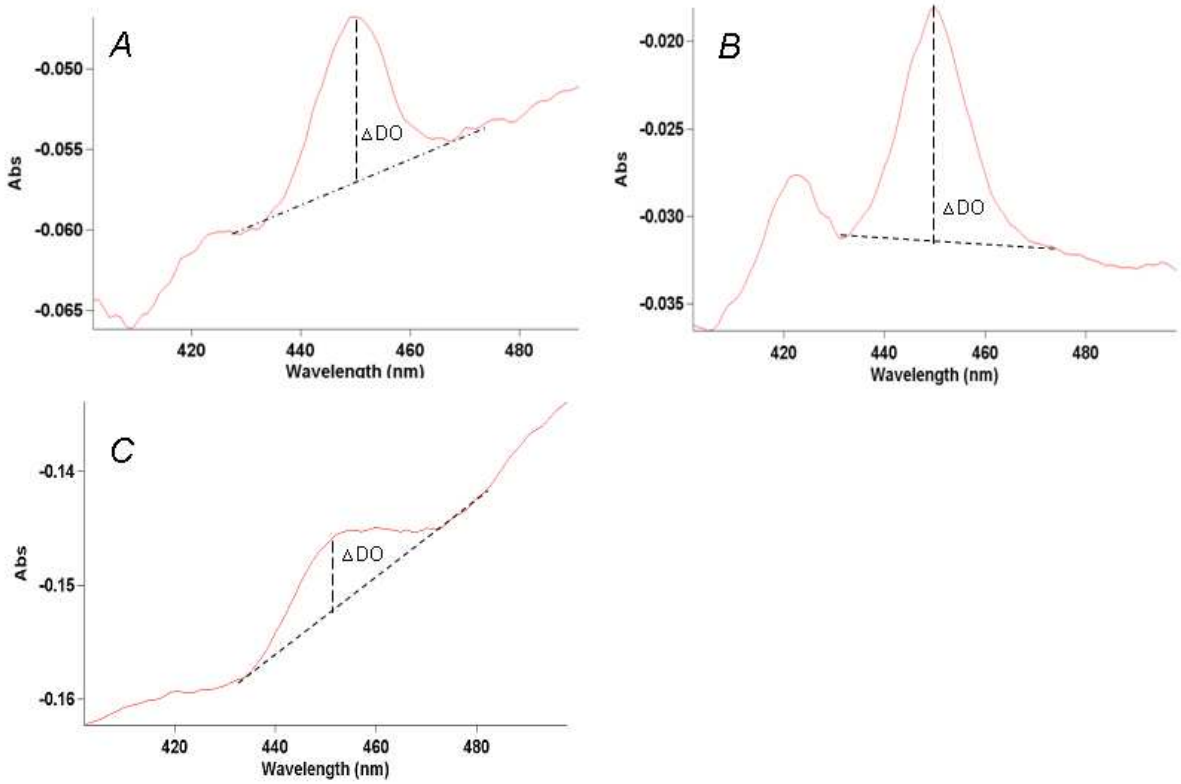


Figure 88 : Spectres différentiels au CO enregistrés avec les microsomes de levures transformés avec pYe71D97 (A) et (B) et pYeAZ1-nt (C).

Les conditions de cultures sont décrites dans le Tableau 36**.**

Conditions de culture	CYP71D97	CYP71D97	CYP71AZ1-nt
Température d'induction	18 ℃	18 ℃	18 ℃
Durée d'induction	6 h	6 h	15 h
Broyage des levures	Automatisé	Automatisé	Manuel
Précipitation des microsomes	Ultracentrifugation	Ultracentrifugation	Précipitation chimique
Compositions du milieu	SGI et YPGE	SGI modifié et YPL modifié	SGI et YPGE
[P450] en µM	1,1	1,4	0,5

Tableau 36 : Conditions de culture permettant l'expression de CYP71D97 et CYP71AZ1-nt.

III.2.3.5 Criblage métabolique

La préparation des différents microsomes a toujours été immédiatement suivie de tests de métabolisation avec les différents substrats appartenant à la voie de synthèse des furocoumarines pour tous les gènes candidats. En effet, une faible activité peut être décelable sur les chromatogrammes d'HPLC même en l'absence de spectre d'absorption à 450 nm.

III.2.3.5.1 CYP71D97

CYP71D97 est un P450 dont l'expression dans les levures a pu être vérifiée par la présence d'un spectre d'absorbance à 450 nm (Figure 88). L'ensemble de mes tests d'activité avec les dérivés coumariniques et furocoumariniques ne m'ayant pas permit de mettre en évidence une activité (Annexe 9), j'ai élargi mon domaine de recherche d'activité à la famille des flavonoïdes. CYP71D9, une flavonoïde 6-hydroxylase, est un P450 présentant 46 % d'identité peptidique avec CYP71D97. Cette enzyme présente une spécificité de substrats relativement larges en métabolisant la liquiritigénine, la lutéoline, la naringénine, l'apigénine et l'ériodictyole.

Un criblage de molécules appartenant à la famille des flavanones et des flavones (Annexe 9) avec CYP71D97 est réalisé. Ce premier criblage a été effectué par une technique de spectrofluorométrie (Olry et al., 2007) mise au point par Olry *et al*, au laboratoire de l' IBMP-CNRS de Strasbourg (groupe de D. Werck-Reichhart) . Ce criblage semble montrer que CYP71D97 métabolise la 5-OH flavone et dans une moindre mesure la 7-OH flavone (données non présentées). Cette indication m'a encouragé à élargir le criblage à des molécules appartenant à la voie de biosynthèse des flavonoïdes et à réaliser une analyse par HPLC. Contrairement à CYP71D9 qui présente une activité avec différents flavonoïdes, seule la 5OH-flavone a été métabolisée parmi tous les flavonoïdes testés et analysés. Les résultats obtenus (Figure 89) indiquent que la 5-OH flavone est métabolisée en un produit P1 plus hydrophile que le substrat. Celui-ci n'a pas été identifié, mais pourrait vraisemblablement être une molécule hydroxylée.

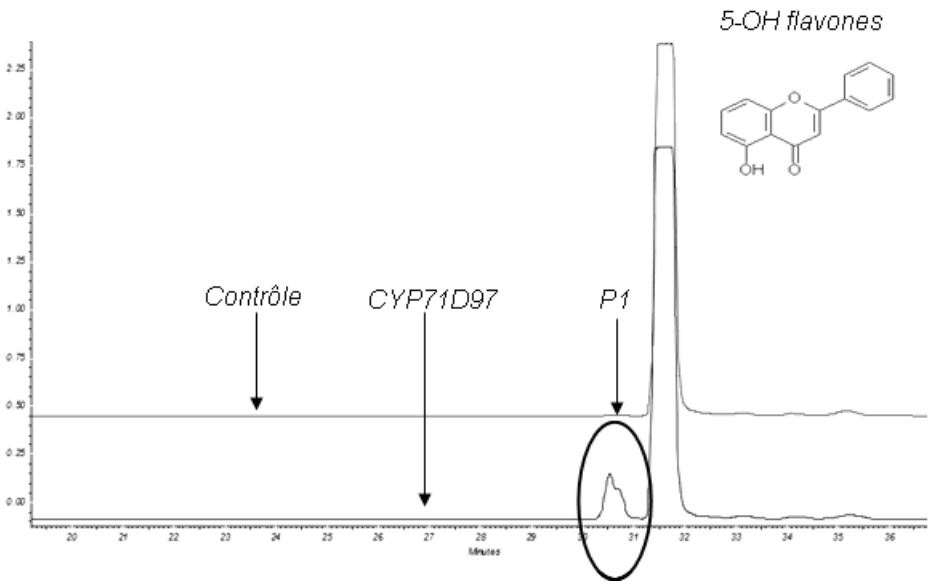


Figure 89 : Chromatogramme obtenu après incubation de CYP71D97 avec la 5-0Hflavone.
Cette figure représente une comparaison des chromatogrammes issus du test de métabolisation de CYP71D97 avec le 5-OH flavone en présence et en absence de NADPH (contrôle). Le produit P1 apparait uniquement lors de l'incubation de CYP71D97 en présence de NADPH.

III.2.3.5.2 CYP71AJ5

CYP71AJ5 est proche phylogénétiquement de CYP71AJ1 et CYP71AJ4. Il constitue donc un bon candidat pour catalyser une réaction appartenant à la voie de biosynthèse des furocomuairnes. À cet effet, des incubations ont été réalisées avec les différents intermédiaires de la voie de biosynthèse malgré l'absence d'absorbance caractéristique des P450 fonctionnelles à 450 nm. Des résultats préliminaires qui nécessitent d'être approfondis ont mis en évidence l'apparition d'un produit P1 avec le bergaptol et d'un produit P2 avec le bergaptène (Figure 90). Malgré des tentatives pour augmenter l'expression de CYP71AJ5 dans le système levure (voir modification du protocole d'expression de Pompon), les autres essais de métabolisation du bergaptol et du bergaptène n'ont pas permis d'obtenir à nouveau ces deux produits. Atréfact ou réalité ? Des expérimentations complémentaires doivent être effectuées.

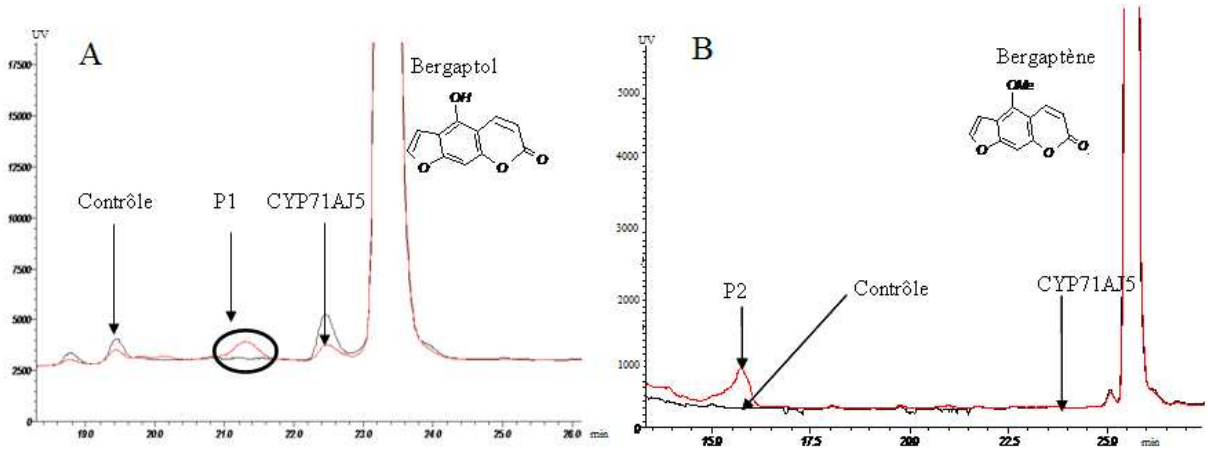


Figure 90 : Chromatogramme obtenu après incubation de CYP71AJ5 avec le bergapol et le bergaptène.
A) Comparaison des chromatogrammes issus du test de métabolisation de CYP71AJ5 avec le bergaptol en présence (rouge) et en absence de NADPH (noire). Le produit P1 apparait uniquement lors de l'incubation de CYP71AJ5 en présence de NADPH. B) Comparaison des chromatogrammes issus du test de métabolisation de CYP71AJ5 avec le bergaptène en présence (rouge) et en absence de NADPH (noire). Le produit P2 apparait uniquement lors de l'incubation de CYP71AJ5 en présence de NADPH.

III.3 Discussion

Les différentes étapes de la voie de biosynthèse des furocoumarines ont été identifiées depuis le début des années 1960 sur le plan des réactions chimiques. À ce jour, seulement quatre étapes ont été caractérisées d'un point de vue moléculaire. Parmi toutes les réactions catalysées, près de 60 % le sont par des enzymes de la famille des cytochromes P450. La caractérisation moléculaire de la voie de biosynthèse consiste donc à relever un challenge de taille. Il faut d'une part identifier des gènes codant pour des P450 à partir de plantes qui ne sont pas des plantes modèles pour la communauté scientifique et pour lesquelles il n'y a que très peu de ressources informatiques. D'autre part, il faut disposer d'un moyen d'expression des protéines correspondantes. Si cette la levure est très largement plébiscitée, cette méthode n'est pas forcément toujours la mieux adaptée en fonction du P450 étudié. Enfin, la dernière étape de criblage métabolique est sans doute la plus aléatoire du fait de la grande diversité de molécules existantes au sein des plantes (molécules glycosylées, prénylées, estérifiées).

Pour répondre à la première question, les collaborations avec deux équipes européennes nous ont permis d'accéder à un certain nombre de gènes de fonction inconnue, mais potentiellement impliqués dans la synthèse de furocoumarines. CYP71AZ1 semble être un candidat intéressant. Les données, communiquées par Kellner S., indiquent que le profil

d'expression de CYP71AZ1 semble être corrélé à celui de CYP71AJ1. De plus, le criblage de la banque d'EST de *Pastinaca sativa* a révélé la présence de deux copies de CYP71AZ1 chez cette plante alors qu'une seule copie n'est identifiée chez *Ammi majus.* Ce résultat reste cependant à vérifier par un criblage expérimental. Cette distribution de CYP71AZ1 d'après le criblage informatique ressemble beaucoup à celle des gènes codant pour la famille des CYP71AJ comme peut l'attester l'arbre phylogénétique généré à partir de différents P450 dont la fonction a été caractérisée à ce jour. En effet, un seul clone de CYP71AJ a été identifié et caractérisé chez les plantes ne produisant que des furocoumarines linéaires comme *Ammi majus* (CYP71AJ1) (Larbat et al., 2007) ou *Apium graveolens* (CYP71AJ2) , alors que ce gène se trouve sous forme de deux copies chez une plante productrice à la fois de furocoumarines linéaires et angulaires comme *Pastinaca sativa* (Larbat et al., 2009). De plus, ces deux copies CYP71AZ2 et CYP71AZ3 présentent 73 % d'identité entre elles, mais ne présente que 63 % d'identité au niveau de leurs séquences SRS. Le même constat a été fait auparavant pour la psoralène synthase (CYP71AJ3) et l'angélicine synthase (CYP71AJ4), qui présentent 70 % d'identité peptidique au niveau de la séquence intégrale, mais seulement 40 % d'identité au niveau des séquences SRS (Larbat et al., 2009).

CYP71AJ5 me semble aussi être un bon candidat. En effet, il est fortement probable que les différentes étapes de la voie de biosynthèse des furocoumarines soient métabolisées par des gènes issus d'un même gène ancestral. Cette dérive génétique apparue par des phénomènes de duplication génique a déjà été décrite dans la voie de biosynthèse des benzoxazinoïdes (Frey et al., 2009). Les benzoxazinoides sont des métabolites secondaires synthétisés par certaines Poacées, dont la synthèse fait intervenir de nombreuses étapes successives catalysées par des cytochromes P450 appartenant tous à la même sous-famille et qui sont apparus suite à des duplications en tandem d'un même gène ancestral (Frey et al., 2009). Le même scénario d'évolution est envisageable pour la synthèse des furocoumarines d'autant plus, que la psoralène synthase (CYP71AJ1) et l'angélicine synthase (CYP71AJ4), semblent toutes les deux dériver d'un même gêne ancestral (Larbat et al., 2009). D'après, l'arbre phylogénétique construit à partir de P450 déjà caractérisés et des séquences peptidiques des différents membres de la sous-famille des CYP71AJ (Figure 83), le clone CYP71AJ5 se détache de l'angélicine synthase et du cluster de gène codant pour la psorène synthase. De plus, la comparaison des acides aminés des SRS des différents membres de la sous-famille des CYP71AJ semble indiquer une divergence d'affinité de substrat de CYP71AJ5. Les séquences SRS de la psoralène synthase sont différentes de l'angélicine synthase, elle-même

différente du nouveau clone CYP71AJ5. L'ensemble de ces données permet de supposer que ce nouveau clone dérive d'un même gène ancestral que la psoralène synthase et l'angélicine synthase, mais qui aurait acquis une autre fonction enzymatique, par des phénomènes de duplication génique.

Les autres candidats CYP71D97, CYP71D98, CYP82H1 et CYP76B8 me semblent être intéressants dans la mesure où ils sont exprimés suite à une élicitation des cultures cellulaire d'*Ammi majus*. Toutefois, l'induction des cultures cellulaires d'*Ammi majus* par un champignon peut entrainer la synthèse de nombreux autres métabolites secondaires impliqués dans des processus allélopathiques tels que les flavonoïdes ou les terpènes. D'ailleurs, l'élicitation des cellules d'*Ammi majus* par *Phytophtora* s'accompagne, en plus des furocoumarines, d'une forte production de phtalide et d'une dizaine de molécules encore non identifiées (Kellner S., communication personnelle). C'est une des limites me, semble-t-il, du differential display. Ces gènes peuvent intervenir dans d'autres voies de biosynthèses que celle des furocoumarines.

De nouvelles pistes pourront s'ouvrir dans les mois (ou année à venir), car les programmes de séquençage deviennent de plus en plus routiniers. Ainsi, au laboratoire nous avons généré deux banques de cDNA issues de plantes qui produisent des furocoumarines (*Pastinaca sativa* et *Ruta graveolens*). L'analyse de la banque de panais, comme cela été développé plus tôt a permis d'identifier des orthologues de CYP71AZ1. D'autres pistes sont explorées dans le cadre de la thèse de Celia Krieger pour cloner des orthologues de CYP71AJ. La seconde banque provenant de la Rue est en cours d'analyse. À côté de ces banques, il existe des programmes de séquençage extensif qui génèrent des données nouvelles quasi quotidiennement. Parmi ces programmes, on trouve le consortium Phytometasyn canadien qui a pour objectif le séquençage de 75 plantes médicinales.

Le second point crucial dans le challenge, consiste à améliorer un système d'expression déjà existant et largement utilisé dans la communauté scientifique des « P450istes » : la levure. Les premières tentatives d'expression dans la levure des P450 isolés chez *Ammi majus* (CYP71D97, CYP71D98, CYP71AZ1, CYP82H1 et CYP76B8) ainsi que chez les P450 (CYP71AJ5) isolés chez *Thapsia garganica* n'ont pas permis de mettre en évidence une expression des P450 d'intérêts. Cette absence, ou très faible expression de l'enzyme est caractérisée par l'absence de spectre d'absorbance à 450 nm en présence de monoxyde de

carbone. Une première tentative de modification de l'extrémité N-terminale a été entreprise pour CYP71D98-nt et CYP71AZ1-nt afin d'améliorer le niveau d'expression de ces enzymes dans les levures comme l'a déjà montré Batard *et al* avec la C4H (Batard et al., 2000) et Larbat *et al* avec CYP71AJ1 (Larbat et al., 2007). L'expression de ces gènes chimères n'a cependant pas été concluante, et j'ai du entreprendre des modifications du protocole de culture initialement décrit par Pompon (Pompon et al., 1996a). J'ai modifié différents paramètres tels que la température d'induction, la durée d'induction, le mode de broyage et de précipitation des levures ainsi que la composition des milieux de culture. Ces modifications m'ont permis d'exprimer de manière quantifiable les CYP71D97 (1,4 µM) et CYP71AZ1-nt (0,5 µM). Néanmoins, je ne suis pas parvenu à exprimer de manière notable les CYP71D98-nt, CYP82H1, CYP76B8 et CYP71AJ5.

Des hypothèses peuvent être envisageables pour expliquer cet échec : une dégradation excessive de l'enzyme, un défaut de synthèse au niveau de la transcription ou de la traduction, ou encore un niveau d'expression trop faible pour être détecté par spectrométrie d'absorption. J'ai présenté, dans ce travail d'expression hétérologue chez la levure, différents leviers d'action permettant de jouer sur l'expression des enzymes. Mon travail ne consistait pas à optimiser l'expression d'un P450 déjà caractérisé comme l'a fait Jiang et Morgan avec la C4H et la F5H (Jiang and Morgan, 2004), mais d'obtenir une expression suffisante des gènes candidats qui m'avaient été confiés (CYP71D97, CYP71D98-nt, CYP71AZ1-nt, CYP82H1, CYP76B8 et CYP71AJ5) afin de déterminer leur(s) activité(s). Cependant, il semblerait qu'il n'existe pas de protocole universel d'expression de P450 chez la levure qui puisse avec certitude produire des protéines fonctionnelles.

D'autres pistes sont à présent envisagées. Les alternatives consistent par exemple à recoder intégralement le gène en vue de produire la protéine chez la levure avec les acides aminés préférentiels utilisés par cet organisme. Une autre alternative plus simple à mettre en place consiste à rajouter au niveau de l'extrémité du gène candidat une queue poly histidine. Cette étiquette permettra de tester le niveau d'expression en réalisant un western blot. Il existe cependant un risque avec cette méthode d'altérer l'activité de la protéine.

Une autre option consiste à changer de système d'expression comme le tabac qui a fonctionné avec la Diox 4 (chapitre précédent). Des travaux de mise au point sont actuellement menés au laboratoire par Alexandre Olry. J'ai également initié un travail d'expression de ces P450 dans des cellules d'insectes qui sera repris lors d'une collaboration avec le Japon.

La troisième étape du challenge n'est pas la plus facile. Elle consiste à identifier la fonction du P450 candidat. Cette étape est laborieuse, car il faut réaliser un criblage métabolique très large. À ce jour, seul 20 %, des P450 d'*Arabidopsis thaliana* se sont vus assigner une fonction. Sur les sept P450 testés dans le cadre de ce travail des pistes ont été trouvées pour deux d'entre eux. Mes tests d'incubation avec CYP71D97 dont le niveau d'expression était détectable en spectrométrie d'absorption ne me permettent pas de statuer de manière définitive sur l'activité de ce P450. En effet, j'ai obtenu une métabolisation du 5-OH flavone à deux reprises avec des techniques d'analyse différentes (HPLC et fluorométrie), mais qui n'ont pu être observées de manière répétée. L'expression de ce P450 est fonctionnelle d'après le spectre d'absorption à 450 nm, ce qui me fait supposer que les difficultés pour confirmer à nouveau cette métabolisation ne proviennent non pas de l'enzyme, mais plutôt du produit formé après métabolisation (volatile, instable ?). Je ne peux cependant pas exclure que la 5-OH flavone ne soit pas le substrat préférentiel de CYP71D97, ce qui expliquerait aussi certaines difficultés à reproduire cette métabolisation. Dans ce cas, il faudrait cribler d'autres flavonoïdes, plus ou moins proches de la 5-OH flavone. Le résultat obtenu avec le CYP71D97 est sans doute une indication sur la famille ou le type de molécule que ce P450 est capable de métaboliser. En effet, les plantes appartenant à la famille des Apiacées sont reconnues pour accumuler des flavonoïdes de type flavone et flavonol (Gebhardt et al., 2005). D'ailleurs, les formes glucosides de l'isorhamnétine et du kaempférol (isorhamnetin-3-O-rutinoside, kaempferol-3-O- glucoside et l'isorhamnetin-3-O-glucoside) ont déjà été identifiées dans les parties aériennes de *Ammi majus* (Singab, 1998). Le kaempférol (ou 3,4',5,7-tétrahydroxyflavone) est un flavonoïde proche structurellement de la 5-OH flavone qui devra être testé dans les prochains tests de métabolisation.

Pour CYP71AJ5, le scénario est quelque peu différent dans la mesure où je ne peux quantifier le niveau d'expression de ce P450 dans la levure. Il semblerait que les différentes modifications du protocole de cultures de Pompon n'aient pas suffi à améliorer de façon significative la production de P450 fonctionnel. L'absence de spectre d'absorption à 450 nm ne me permet pas de confirmer la présence de protéine fonctionnelle dans mes extraits de microsomes. Cependant, il est possible qu'une très faible quantité de protéines soit synthétisée et qu'elle ne soit pas détectée en spectrophotométrie d'absorption. Cette deuxième hypothèse semble être le cas pour ce P450 dans la mesure où j'ai obtenu, lors d'un test d'activité, une faible métabolisation du bergaptol et du bergaptène. Néanmoins, les difficultés à produire cette enzyme *via* la levure en quantité notable ne m'ont actuellement pas permis de réitérer ce

résultat : soit l'enzyme n'est pas exprimée dans mes levures (ce qui expliquerait que je n'obtienne plus de métabolisation) ; soit ce ne sont pas les substrats préférentiels du CYP71AJ5. Chez *Thapsia garganica,* la scopolétine et la 6-methoxy-7-geranyloxycoumarine (Rasmussen et al., 1981) ont déjà été identifiées, mais, à notre connaissance, il n'y a pas de furocoumarines qui aient été décrites chez cette plante. Dans l'hypothèse où le bergaptol et le bergaptène seraient les véritables substrats du CY71AJ5, il serait concevable que l'évolution du patrimoine génétique de *Thapsia garganica* ait permis de supprimer une étape préliminaire de la voie de biosynthèse des furocoumarines. Le bénéfice pour la plante de produire ces composés toxiques n'a peut-être pas été suffisant au regard du cout métabolique qu'ils entrainent. Il en résulterait une suppression de synthèse des furocoumarines, mais dont l'expression de certains gènes serait encore visible comme pour le CYP71AJ5. En effet, il est intéressant de noter que *Thapsia garganica* appartient à la famille des Apiacées qui comprend un grand nombre d'espèces synthétisant des furocoumarines linéaires et angulaires telles que la carotte (*D. carota*), le céleri (*A. graveolens*) ou le panais (*Pastinaca sativa*) (Peroutka et al., 2007; Schulzova et al., 2007). Cette hypothèse est cependant très spéculative et n'a pas été, à ma connaissance, encore démontrée chez les végétaux supérieurs.

III.4 Conclusion

Dans le cadre de ce troisième chapitre, j'ai abordé l'identification et la caractérisation de gènes codant pour des cytochromes P450, enzyme majoritaire de la voie de biosynthèse des furocoumarines. Les cytochromes P450 ont évolué afin de former la plus grande famille d'enzymes qui participent au métabolisme végétal. Ils sont considérés comme étant des acteurs majoritaires dans les voies du métabolisme secondaire. On les retrouve notamment au niveau des étapes limitantes et irréversibles dans toutes les voies métaboliques, le plus souvent en réalisant des hydroxylations ou des oxygénations. Plusieurs centaines de gènes codant des P450 ont été déterminés chez chacune des espèces séquencées. À titre d'exemple, plus de 250 P450 ont été recensés chez *Arabidopsis* et plus de 330 chez le riz (*Oryza sativa*). Mon but a été de déterminer la fonction de P450 orphelins susceptibles d'intervenir dans la synthèse de furocoumarines. Ces P450 candidats m'ont été confiés par l'équipe du Professeur Matern de l'université de Malburg et par l'équipe du Prof. Simonsen de l'université de Copenhague.

Une première étape de recherche bibliographique et bio-informatique m'a permis de déterminer lesquels de ces candidats sont les plus pertinents pour la synthèse de

furocoumarines et lesquels le sont moins. Il semblerait que CYP71AJ5 soit un très bon candidat dans la mesure où il appartient à la même sous-famille que les deux autres gènes déjà caractérisés dans la voie de synthèse des furocoumarines (CYP71AJ1 et CYP71AJ4). De plus, l'analyse de la séquence peptidique semble montrer quelques divergences avec la psoralène synthase (CYP71AJ1) et l'angélicine synthase (CYP71AJ4), notamment au niveau des séquences SRS. Néanmoins, la tentative d'expression de ce P450 ne m'a pas permis d'identifier une fonction précise de CYP71AJ5 à ce jour. J'ai obtenu lors d'un criblage métabolique une faible métabolisation du bergaptol et du bergaptène, mais qui n'ont jamais pu être reconduites. Malgré d'importantes modifications du protocole d'expression de ce gène dans la levure et du protocole de préparation des microsomes, je ne suis pas parvenu à statuer de façon définitive sur la fonction de CYP71AJ5.

CYP71AZ1 semble aussi être un bon gène candidat d'après l'analyse bio-informatique effectuées. La distribution du nombre de copies de ce gène au niveau d'*Ammi majus* (furocoumarines linaire) et de *Pastinaca sativa* (furocoumarines linéaires et angulaires) ressemble beaucoup à celle rencontrée dans la famille des CYP71AJ. De plus, les études expérimentales ont montré que le niveau d'expression de CYP71AZ1 était corrélé au profil d'accumulation des furocoumarines comme cela a déjà été démontré pour CYP71AJ1. Les modifications apportées au niveau de l'ancre membranaire ainsi que celles apportées au niveau du protocole d'expression de ce gène ont permis d'améliorer la production de CYP71AZ1-nt, mais le criblage métabolique effectué n'a pas révélé de substrats métabolisés à ce jour.

CYP71D97 et CYP71D98 appartiennent tous les deux à la famille des CYP71 et par conséquent, constituent d'autres candidats potentiels. Seul CYP71D97 a révélé une expression quantifiable en spectrométrie d'absorption et le criblage métabolique a permis d'obtenir par deux reprises et par deux techniques d'analyse différentes (fluorométrie et HPLC) la métabolisation du 5-0H flavone. Cependant, de possibles problèmes de stabilité du produit formé ne m'ont pas permis de reproduire ces résultats. Les autres P450 candidats n'ont à ce jour pas pu être correctement exprimé ni même caractérisé.

L'acquisition de connaissances sur les enzymes appartenant à la famille des P450 est primordiale quand on sait que les réactions qu'elles catalysent sont généralement difficiles ou impossibles à réaliser par chimie de synthèse. La grande diversité de P450 présentes chez toutes les espèces végétales et de réactions, souvent complexes, qui peuvent être ainsi catalysées offrent une ressource considérable pour l'industrie et pour la bioremédiation.

Chapitre IV : Conclusion générale et perspectives

I Conclusion générale

Ce travail de thèse avait pour objectif d'acquérir des connaissances sur une composante moléculaire de la voie de biosynthèse des furocoumarines. Malgré la caractérisation récente de deux gènes (la psoralène synthase et l'angélicine synthase) par Romain Larbat qui sont directement impliqués dans cette voie de biosynthèse, relativement peu d'étapes sont à ce jour caractérisées concernant ces molécules, dont l'intérêt pharmaceutique est indéniable. Mon travail de thèse s'est donc partagé en deux grands axes : d'une part, identifier de nouveaux gènes candidats susceptibles d'intervenir au sein même de cette voie de biosynthèse, et d'autre part de tenter de caractériser des gènes identifiés quelque temps, mais qui n'avaient encore jamais été exprimés.

Identification et caractérisation de nouveaux gènes candidats appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes.

La première partie de mes travaux présentée dans cette thèse avait pour objectif d'expliquer les différentes approches que j'ai mises en place, afin d'identifier des gènes candidats intervenant au niveau de la voie de biosynthèse des furocoumarines. La recherche bibliographique était mon premier outil d'analyse, ce qui fut un véritable tournant dans l'orientation de mes recherches concernant les familles de gènes susceptibles d'intervenir dans la synthèse des furocoumarines. En effet, la voie de biosynthèse spécifique des furocoumarines est initiée par la conversion d'umbelliférone (7-hydroxycoumarine) qui initie le début des cascades de réactions chimiques aboutissant à la formation de furocoumarines. Des premières expériences basées sur le suivi et l'évolution de précurseurs isotopiques ont supposé que la majorité des étapes de cette voie de biosynthèse était métabolisées par des enzymes appartenant à la famille des cytochromes P450. D'après des études préliminaires effectuées sur *Melilotus officinalis* (Brown et al., 1960) et sur *Lavandula officinalis* (Brown, 1962), la conversion de l'umbelliférone était supposée provenir de l'*ortho*-hydroxylation de

l'acide *cis-p*-coumarique par une enzyme appartenant à la famille des cytochromes P450. L'identification et la caractérisation, en 2008 par l'équipe du professeur Mizutani d'une enzyme appartenant à la famille des dioxygénases α-cétoglutarates-dépendantes (F6'H) chez *Arabidopsis thaliana*, est un préalable essentiel à mon travail de thèse. Celle-ci étant impliquée dans la conversion de la scopolétine à partir du féruloyle CoA (Kai et al., 2008), j'ai supposé, aux vues des analogies réactionnelles mises en jeu avec la synthèse d'umbelliférone, que l'étape de conversion du *p*-coumarate en umbelliférone faisait intervenir un intermédiaire (le *p*-coumaroyle CoA) et que l'enzyme appartenait aussi à une dioxygénase α-cétoglutarate-dépendante. Basée sur cette hypothèse, j'ai initié un travail d'identification de nouveaux gènes appartenant à la famille des dioxygénases par des approches bio-informatiques dans un premier temps, puis expérimentales dans un second temps. L'association de ces deux approches m'a permis d'identifier chez des espèces productrices de furocoumarines (*Citrus limetta*, *Pastinaca sativa*, *Ruta graveolens*) des gènes potentiellement orthologues des F6'H isolés chez *Arabidopsis thaliana.* Les bases de données informatiques publiques (Genbank) et privées que nous possédons au laboratoire depuis peu, m'ont permis d'identifier la séquence complète de deux dioxygénases possédant de fortes homologies (60 %) avec les F6'H chez le citrus (Diox Citrus) et chez le panais (Diox Panais) ; par ailleurs, elles présentent toutes les séquences caractéristiques de fixations au Fe^{2+} et à l'α-cétoglutarate comme les F6'H. Des approches expérimentales par PCR m'ont ensuite permis de récupérer ces clones à partir d'extraits d'ARN de plantes, et de les sous-cloner dans un vecteur d'expression pour effectuer leurs caractérisations. Quatre autres gènes candidats (Diox 1, Diox 2, Diox 4 et Diox 8) ont aussi été identifiés chez *Ruta graveolens* par une approche expérimentale de criblage de banque de cDNA construite à partir d'extraits d'ARN de Rue par Larbat R. au cours de sa thèse. L'ensemble de ces candidats révèle une forte identité (60 %) peptidique avec les F6'H d'*Arabidopsis* et présente aussi les séquences consensus de fixation au Fe^{2+} et l'α-cétoglutarate. La recherche des séquences génomiques des six nouveaux gènes identifiés a révélé la présence d'un intron situé exactement au même endroit, mais présentant des séquences nucléotidiques totalement différentes. Ce constat confirme l'existence biologique de quatre copies très fortement semblables (>91 % identité peptidique) chez *Ruta graveolens* qui seraient, selon mes hypothèses, apparues suite à des phénomènes de duplication génique aboutissant à la création de nouvelles fonctions enzymatiques. L'ensemble des gènes identifiés chez ces trois espèces végétales fera l'objet de tentatives de caractérisation fonctionnelle dans la seconde partie de ce premier axe.

Caractérisation *in vitro* et *in planta* des gènes candidats codant pour des dioxygénases α-cétoglutarates-dépendantes.

L'ensemble des gènes identifiés chez *Citrus limetta*, *Pastinaca sativa* et *Ruta graveolens* a tenté d'être exprimé *in vitro* dans un système d'expression hétérologue bactérien (*E.coli* BL21(DE3). Les résultats de ces expérimentations ont montré que la Diox 2 et la Diox 8 isolées chez *Ruta graveolens* étaient des orthologues de la F6'H d'*Arabidopsis thaliana* capable de métaboliser le féruloyle CoA en scopolétine. Quant à la Diox 4 isolée aussi chez la Rue, elle présente une bi fonctionnalité novatrice d'*ortho*-hydroxylation du *p*-coumaroyle CoA et du féruloyle CoA. Les paramètres cinétiques calculés pour cette enzyme indiquent qu'elle métabolise avec la même affinité et la même efficacité ces deux substrats *in vitro*. De plus, les expériences cinétiques d'inhibition de la Diox 4 avec les différents composés coumariniques et furocoumariniques ont montré que le psoralène était un inhibiteur compétitif de l'activité F6'H et C2'H. Le Ki calculé expérimentalement correspond approximativement au double de la valeur du Km pour le féruloyle CoA et p-coumaroyle CoA. Bien que l'affinité de la Diox 4 soit moins importante pour le psoralène que pour ses deux substrats, le psoralène peut agir tel un rétrocontrôle inhibiteur de l'activité C2'H, lorsque la concentration en psoralène devient excessive et nocive pour la plante. Ce feedback négatif permettrait de réguler la voie de biosynthèse des furocoumarines et d'ajuster les teneurs en furocoumarines chez *Ruta graveolens*. J'ai poursuivi ce travail de caractérisation *in vitro* par la caractérisation fonctionnelle de la Diox 4 *in planta*. Une corrélation positive a été montrée entre la variation de la concentration en umbelliférone dans les feuilles de *Ruta graveolens* et le niveau d'expression de la Diox 4. De même, la vaporisation d'un inhibiteur spécifique de dioxygénase (le Regalis®), sur les parties aériennes de la Rue, a montré une diminution en umbelliférone et en furocoumarines situées en aval de la Diox 4. L'intervention d'une dioxygénase dans le métabolisme des furocoumarines chez *Ruta graveolens* est par conséquent une évidence. Un autre travail d'expression transitoire de la Diox 4 réalisé dans des feuilles de *Nicotiana benthamiana*, a montré que la surexpression du gène codant pour cette dioxygénase entraînait l'accumulation d'umbelliférone et de scopolétine chez cette plante qui ne produit habituellement pas de coumarines. L'identification de ces deux composés chez *Nicotiana benthamiana* qui surexpriment la Diox 4, dans des proportions équivalentes, confirme le caractère bi fonctionnel de cette dioxygénase. L'ensemble des

résultats qui fait suite à la caractérisation *in vitro* d'une part, et *in planta* d'autre part, montre que l'étape initiatrice de la voie de biosynthèse des furocoumarines n'est pas métabolisée par une enzyme appartenant à la famille des cytochromes P450 chez *Ruta graveolens*, mais, par une dioxygénase α-cétoglutarate-dépendante (Diox 4).

Tentative d'expression des cytochromes P450 candidats dans la levure

Le deuxième grand axe de mes travaux de recherche consistait à exprimer d'une part différents gènes candidats appartenant à la famille des P450 et d'autre part, à caractériser ces candidats. L'ensemble des gènes candidats qui m'a été confié a été identifié par l'équipe du Prof. Matern de l'université de Malburg et l'équipe du Prof. Simonsen de l'université de Copenhague, après la mise en place d'une collaboration avec notre laboratoire. Mon objectif consistait donc à améliorer le niveau d'expression de ces gènes qui ont, pour la plupart, essayé d'être exprimés sans véritable succès dans la levure, par mes prédécesseurs. L'analyse bio-informatique des différents candidats a montré que CYP71AJ5, CYP71AZ1, CYP71D97 et CYP71D98 étaient de sérieux candidats susceptibles d'intervenir dans la voie de biosynthèse des furocoumarines. Les différentes mises au point que j'ai pu apporter au protocole initial d'expression hétérologue dans la levure n'ont pas permis de statuer de façon définitive sur la fonction de ces gènes candidats. Néanmoins, j'ai obtenu, à deux reprises, la métabolisation du 5-0H-flavone par CYP71D97 et une légère métabolisation du bergaptol et du bergaptène par le CYP71AJ5. L'impossibilité de reproduire ces résultats ne me permet cependant pas de conclure quant à la fonction de ces gènes, mais peut néanmoins donner certaines pistes de recherche. Je peux cependant conclure qu'il n'existe pas de protocole standard permettant d'exprimer avec certitude n'importe quel P450 dans la levure. Aux vues de mes résultats d'expression, les mises au point du protocole d'expression doivent être apportées de manière individuelle pour chaque P450 étudié.

II Perspectives

L'ensemble des gènes candidats que j'ai identifié chez *Ruta graveolens*, *Citrus limetta* et *Pastina sativa* n'a pas permis à ce jour de leur attribuer à tous une fonction. Selon mes hypothèses évolutives (duplication génique et sub-fonctionalisation) concernant l'apparition

des différentes copies de gènes codant pour des dioxygénases chez la Rue, il serait étonnant de ne pas réussir à trouver un clone capable de métaboliser uniquement le p-coumaroyle CoA chez la Rue comme il a été trouvé pour le féruloyle CoA (Diox 8 et Diox 2). La Diox 1, que je n'ai pas réussi à exprimer dans les bactéries, a révélé, après analyse bio-informatique, une mutation sur des acides aminés conservés habituellement chez les autres dioxygénases. La récupération de ce clone « non muté » et la caractérisation de celui-ci, est peut être la clef du mystère, d'autant plus qu'une très faible variation de la séquence peptidique entraîne une spécificité différente des dioxygénases, à l'image de la Diox 4 (C2'H et F6'H) et de la Diox 8 (F6'H). Cette très forte analogie peptidique entre ces deux dioxygénases qui possèdent des fonctions différentes, m'a conduit à entreprendre un travail de relation structure-fonction au niveau des dioxygénases de *Ruta graveolens*. En effet, la modélisation tridimensionnelle de mes différentes dioxygénases, nous a permis d'identifier avec Larbat R. chez la Diox 4 et la Diox 8, deux acides aminés différents situés au niveau de la poche catalytique. De ce fait, nous avons entrepris au laboratoire des travaux de mutagenèse dirigée, actuellement en cours d'analyse, qui permettront de statuer sur la fonction de ces deux acides aminés.

Les dioxygénases α-cétoglutarates-dépendantes semblent avoir un rôle primordial dans la synthèse de composés coumariniques. A ce jour, seuls des clones capables de métaboliser le féruloyle CoA, le *p*-coumaroyle CoA et le caféoyle CoA ont pu être identifiés (communication personnelle du professeur Mizutani). Il ne serait donc pas surprenant de trouver aussi une dioxygénase capable de métaboliser le cinnamoyle CoA chez des plantes dont la composition est riche en coumarines, à moins que cette réaction ne soit réalisée par une enzyme appartenant à la famille des P450. La dualité entre dioxygénases et P450 pour une même fonction chez des espèces différentes a déjà été décrite à de nombreuses reprises (flavonoides, giberreline), et il serait intéressant de continuer à rechercher des P450 potentiellement impliqués dans la conversion de l'acide *p*-coumarique en umbelliférone afin de vérifier s'il existe aussi une dualité entre ces deux familles d'enzymes pour catalyser la formation d'umbelliférone. La difficulté rencontrée pour identifier des dioxygénases impliquées dans la conversion du cinnamoyle CoA en coumarine me fait penser que cette réaction est peut-être réalisée par un P450 chez les plantes productrices de coumarines.

En ce qui concerne les gènes candidats appartenant à la famille des P450 qui m'ont été confiés, j'essaierai de les exprimer dans d'autres systèmes d'expressions hétérologues, tels que les cellules d'insectes. L'équipe du professeur Mizutani utilise couramment ce système d'expression hétérologue et m'a permis d'initier ce travail de caractérisation lors de mon passage dans leur laboratoire japonais. L'expression transitoire dans les feuilles de *Nicotiana*

benthamiana s'est révélée être un succès avec ma Diox 4. Il serait peut être intéressant de tenter aussi l'expression transitoire de P450 dans ce système *in planta*.

La découverte de l'activité C2'H réalisée par une dioxygénase, entrouvre de nombreuses possibilités d'ingénierie métabolique pour la suite, car il est maintenant possible d'initier la voie de biosynthèse de furocoumarines par accumulation d'umbelliférone chez des plantes non productrices de furocoumarines, telles que le tabac. L'identification des étapes situées en aval permettra certainement de reconstruire un jour l'intégralité de cette voie de biosynthèse dans des plantes modèles, telles qu'*Aribodopsis thaliana*. De nombreuses applications pharmaceutiques (production à grande échelle de furocoumarines), agronomiques (phytoallexines) et sanitaires (phototoxicité des furocoumarines) pourraient en découler.

Chapitre V : Références bibliographiques

Aas, P.A., Otterlei, M., Falnes, P.O., Vagbo, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E., and Krokan, H.E. (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature **421,** 859-863.

Abdalla, D.S.P., Monteiro, H.P., Oliveira, J.A.C., and Bechara, E.J.H. (1986). Activities of superoxide dismutase and glutathione peroxidase in schizophrenic and manic-depressive patients. Clinical Chemistry **32,** 805-807.

Abreu, M.T., von Tirpitz, C., Hardi, R., Kaatz, M., Van Assche, G., Rutgeerts, P., Bisaccia, E., Goerdt, S., Hanauer, S., Knobler, R., Mannon, P., Mayer, L., Ochsenkuhn, T., Sandborn, W.J., Parenti, D., Lee, K., and Reinisch, W. (2009). Extracorporeal photopheresis for the treatment of refractory Crohn's disease: Results of an open-label pilot study. Inflammatory Bowel Diseases **15,** 829-836.

Abyshev, A.Z., Gindin, V.A., Kerimov, Y.B., Ismailov, E.S., Agaev, E.M., and Isaev, N.Y. (1992). Furocoumarins of Ruta graveolens. Chemistry of Natural Compounds **28,** 382-383.

Achaz, G., Netter, P., and Coissac, E. (2001). Study of intrachromosomal duplications among the eukaryote genomes. Molecular Biology and Evolution **18,** 2280-2288.

Adam, M., Dobias, P., Eisner, A., and Ventura, K. (2009). Extraction of antioxidants fom plants using ultrasonic methods and their antioxidant capacity. Journal of Separation Science **32,** 288-294.

Adams, M., Ettl, S., Kunert, O., Wube, A.A., Haslinger, E., Bucar, F., and Bauer, R. (2006). Antimycobacterial activity of geranylated furocoumarins from Tetradium daniellii. Planta Medica **72,** 1132-1135.

Adisen, E., and Gürer, M.A. (2010). Photo(chemo)therapy for cutaneous T cell lymphoma. **44,** 105-111.

Adişen, E., Karaca, F., Oztas, M., and Gürer, M.A. (2008). Efficacy of local psoralen ultraviolet A treatments in psoriasis, vitiligo and eczema. Clinical and Experimental Dermatology **33,** 344-345.

Akashi, T., Aoki, T., and Ayabe, S.I. (1998). Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (Glycyrrhiza echinata L.; Fabaceae) which represents licodione synthase and flavone synthase II. FEBS Letters **431,** 287-290.

Akashi, T., Aoki, T., Takahashi, T., Kameya, N., Nakamura, I., and Ayabe, S.I. (1997). Cloning of cytochrome P450 cDNAs from cultured Glycyrrhiza echinata L. cells and their transcriptional activation by elicitor-treatment. Plant Science **126,** 39-47.

Anterola, A.M., Jeon, J.H., Davin, L.B., and Lewis, N.G. (2002). Transcriptional control of monolignol biosynthesis in Pinus taeda: Factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. Journal of Biological Chemistry **277,** 18272-18280.

Anzellotti, D., and Ibrahim, R.K. (2004). Molecular characterization and functional expression of flavonol 6-hydroxylase. BMC Plant Biology **4**.

Arab, S.T., Al-Nami, S.Y., and Abu-Mustafa, E.A. (2008). Chemical composition of ammi majus l. and its inhibition activity against corrosion. Egyptian Journal of Chemistry **51,** 115-128.

Avato, P. (1991). Essential oil of Thapsia garganica. Planta Medica **57,** 585-586.

Bak, S., Kahn, R.A., Olsen, C.E., and Halkier, B.A. (1997). Cloning and expression in Escherichia coli of the obtusifoliol 14α-demethylase of Sorghum bicolor (L.) moench, a cytochrome p450 orthologous to the sterol 14α-demethylases (CYP51) from fungi and mammals. Plant Journal **11,** 191-201.

Bak, S., Paquette, S.M., Morant, M., Morant, A.V., Saito, S., Bjarnholt, N., Zagrobelny, M., JÃ¸rgensen, K., Osmani, S., Simonsen, H.T., PÃ©rez, R.S., Van Heeswijck, T.B., JÃ¸rgensen, B., and MÃ¸ller, B.L. (2006). Cyanogenic glycosides: A case study for evolution and application of cytochromes P450. Phytochemistry Reviews **5,** 309-329.

Baldwin, J.E., Adlington, R.M., Coates, J.B., Crabbe, M.J., Crouch, N.P., Keeping, J.W., Knight, G.C., Schofield, C.J., Ting, H.H., and Vallejo, C.A. (1987). Purification and initial characterization of an enzyme with deacetoxycephalosporin C synthetase and hydroxylase activities. Biochemical Journal **245,** 831-841.

Bao, A.M., Hestiantoro, A., Van Someren, E.J.W., Swaab, D.F., and Zhou, J.N. (2005). Colocalization of corticotropin-releasing hormone and oestrogen receptor-α in the paraventricular nucleus of the hypothalamus in mood disorders. Brain **128,** 1301-1313.

Barouki, R., and Morel, Y. (2001). Repression of cytochrome P450 1A1 gene expression by oxidative stress: Mechanisms and biological implications. Biochemical Pharmacology **61,** 511-516.

Barreca, D., Bellocco, E., Caristi, C., Leuzzi, U., and Gattuso, G. (2011). Distribution of C- and O-glycosyl flavonoids, (3-hydroxy-3-methylglutaryl)glycosyl flavanones and furocoumarins in Citrus aurantium L. juice. Food Chemistry **124,** 576-582.

Baskin, J.M., Ludlow, C.J., Harris, T.M., and Wolf, F.T. (1967). Psoralen, an inhibitor in the seeds of Psoralea subacaulis (Leguminosae). Phytochemistry **6,** 1209-1213.

Basson, A.E., and Dubery, I.A. (2007). Identification of a cytochrome P450 cDNA (CYP98A5) from Phaseolus vulgaris, inducible by 3,5-dichlorosalicylic acid and 2,6-dichloro isonicotinic acid. Journal of Plant Physiology **164,** 421-428.

Batard, Y., LeRet, M., Schalk, M., Robineau, T., Durst, F., and Werck-Reichhart, D. (1998). Molecular cloning and functional expression in yeast of CYP76B1, a xenobiotic-inducible 7-ethoxycoumarin O-de-ethylase from Helianthus tuberosus. Plant Journal **14,** 111-120.

Batard, Y., Hehn, A., Nedelkina, S., Schalk, M., Pallett, K., Schaller, H., and Werck-Reichhart, D. (2000). Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. Archives of Biochemistry and Biophysics **379,** 161-169.

Baudry, J., Li, W., Pan, L., Berenbaum, M.R., and Schuler, M.A. (2003). Molecular docking of substrates and inhibitors in the catalytic site of CYP6B1, an insect cytochrome P450 monooxygenase. Protein Engineering **16,** 577-587.

Bauer, N., Fulgosi, H., and Jelaska, S. (2011). Overexpression of phenylalanine ammonia-lyase in transgenic roots of Coleus blumei alters growth and rosmarinic acid synthesis. Food Technology and Biotechnology **49,** 24-31.

Bazarbachi, A., Suarez, F., Fields, P., and Hermine, O. (2011). How I treat adult T-cell leukemia/lymphoma. Blood **118,** 1736-1745.

Becerra, J.X. (2001). Interactions between chemical and mechanical defenses in the plant genus bursera and their implications for herbivores. American Zoologist **41,** 865-876.

Becerra, J.X. (2007). The impact of herbivore-plant coevolution on plant community structure. Proceedings of the National Academy of Sciences of the United States of America **104,** 7483-7488.

Becerra, J.X., Noge, K., and Venable, D.L. (2009). Macroevolutionary chemical escalation in an ancient plant-herbivore arms race. Proceedings of the National Academy of Sciences of the United States of America **106,** 18062-18066.

Beier, R.C., and Oertli, E.H. (1983). Psoralen and other linear furocoumarins as phytoalexins in celery. Phytochemistry **22,** 2595-2597.

Benderoth, M., Textor, S., Windsor, A.J., Mitchell-Olds, T., Gershenzon, J., and Kroymann, J. (2006). Positive selection driving diversification in plant secondary metabolism. Proceedings of the National Academy of Sciences of the United States of America **103,** 9118-9123.

Benedetto, A.V. (1977). The psoralens. An historical perspective. Cutis **20,** 469-471.

Bennett, M.D., Leitch, I.J., Price, H.J., and Johnston, J.S. (2003). Comparisons with Caenorhabditis (100 Mb) and Drosophila (175 Mb) using flow cytometry show genome size in Arabidopsis to be 157 Mb and thus 25% larger than the Arabidopsis genome initiative estimate of 125 Mb. Annals of Botany **91,** 547-557.

Berenbaum, M., and Feeny, P. (1981). Toxicity of angular furanocoumarins to swallowtail butterflies: Escalation in a coevolutionary arms race? Science **212,** 927-929.

Berenbaum, M.R. (1990). Evolution of specialization in insect-umbellifer associations. Annual Review of Entomology **35,** 319-343.

Berenbaum, M.R. (2002). Postgenomic chemical ecology: From genetic code to ecological interactions. Journal of Chemical Ecology **28,** 873-896.

Bernhardt, R. (1995). Cytochrome P450: Structure, function, and generation of reactive oxygen species. Reviews of Physiology Biochemistry and Pharmacology **127,** 137-221.

Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. Journal of Biotechnology **124,** 128-145.

Bertea, C.M., Schalk, M., Karp, F., Maffei, M., and Croteau, R. (2001). Demonstration that menthofuran synthase of mint (Mentha) is a cytochrome P450 monooxygenase: Cloning, functional expression, and characterization of the responsible gene. Archives of Biochemistry and Biophysics **390,** 279-286.

Bhushan, B., Trott, S., Spain, J.C., Halasz, A., Paquet, L., and Hawari, J. (2003). Biotransformation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by a rabbit liver cytochrome P450: Insight into the mechanism of RDX biodegradation by

Rhodococcus sp. strain DN22. Applied and Environmental Microbiology **69,** 1347-1351.

Bienz, S., Bisegger, P., Guggisberg, A., and Hesse, M. (2005). Polyamine alkaloids. Natural Product Reports **22,** 647-658.

Bilici, M., Efe, H., Köroglu, M.A., Uydu, H.A., Bekaroglu, M., and Deger, O. (2001). Antioxidative enzyme activities and lipid peroxidation in major depression: Alterations by antidepressant treatments. Journal of Affective Disorders **64,** 43-51.

Biteau, F. (2009). Production de protéines recombinantes par des plantes carnivores génétiquement transformées : Application à Drosera rotundifolia et transfert de la technologie à Nepenthes alata. In Laboratoire Agronomie Environnement ENSAIA/INRA (Vandoeuvre-les-Nancy: Institue National Polytechnique de Lorraine-ENSAIA).

Björkholm, M., Svedmyr, E., and Sjöberg, J. (2011). How we treat elderly patients with Hodgkin lymphoma. Current Opinion in Oncology **23,** 421-428.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. In Annual Review of Biochemistry, pp. 291-336.

Blanco, E., Pignatelli, M., Beltran, S., Punset, A., Pérez-Lluch, S., Serras, F., Guigo, R., and Corominas, M. (2008). Conserved chromosomal clustering of genes governed by chromatin regulators in Drosophila. Genome Biology **9**.

Blumenthal, T., and Gleason, K.S. (2003). Caenorhabditis elegans operons: Form and function. Nature Reviews Genetics **4,** 112-120.

Bono, J.M., Matzkin, L.M., Castrezana, S., and Markow, T.A. (2008). Molecular evolution and population genetics of two Drosophila mettleri cytochrome P450 genes involved in host plant utilization. Molecular Ecology **17,** 3211-3221.

Bordin, F., Carlassare, F., Busulini, L., and Baccichetti, F. (1993). Furocoumarin sensitization induces DNA-protein cross-links. Photochemistry and Photobiology **58,** 133-136.

Boudet, A.M. (2007). Evolution and current status of research in phenolic compounds. Phytochemistry **68,** 2722-2735.

Bourgaud, F., Poutaraud, A., and Guckert, A. (1994). Extraction of coumarins from plant material (Leguminosae). Phytochemical Analysis **5,** 127-132.

Bourgaud, F., Hehn, A., Larbat, R., Doerper, S., Gontier, E., Kellner, S., and Matern, U. (2006). Biosynthesis of coumarins in plants: A major pathway still to be unravelled for cytochrome P450 enzymes. Phytochemistry Reviews **5,** 293-308.

Bourgaud F., A.N., Guckert A., Forlot P. (1989). Natural source of furocoumarins (psoralens). In Psoralens: Past, Present and Future of Photochemoprotection and other biological activities. (paris: John Libbey Eurotext.

Bredebach, M., Matern, U., and Martens, S. (2011). Three 2-oxoglutarate-dependent dioxygenase activities of Equisetum arvense L. forming flavone and flavonol from (2S)-naringenin. Phytochemistry **72,** 557-563.

Brosius, J. (1991). Retroposons - Seeds of evolution. Science **251,** 753.

Brosius, J. (2003). The contribution of RNAs and retroposition to evolutionary novelties. Genetica **118,** 99-116.

Brown, S.A. (1962). Biosynthesis of coumarin and herniarin in lavender. Science **137,** 977-978.

Brown, S.A. (1970). Biosynthesis of furanocoumarins in parsnips. Phytochemistry **9,** 2471-2475.

Brown, S.A., Towers, G.H., and Wright, D. (1960). Biosynthesis of the coumarins. Tracer studies on coumarin formation in Hierochloe odorata and Melilotus officinalis. Canadian journal of biochemistry and physiology **38,** 143-156.

Brown, S.A., el-Dakhakhny, M., and Steck, W. (1970). Biosynthesis of linear furanocoumarins. Canadian journal of biochemistry **48,** 863-871.

Bugg, T.D.H. (2003). Dioxygenase enzymes: Catalytic mechanisms and chemical models. Tetrahedron **59,** 7075-7101.

Bugg, T.D.H., and Winfield, C.J. (1998). Enzymatic cleavage of aromatic rings: Mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. Natural Product Reports **15,** 513-530.

Cabello-Hurtado, F., Durst, F., Jorrin, J.V., and Werck-Reichhart, D. (1998). Coumarins in Helianthus tuberosus: Characterization, induced accumulation and biosynthesis. Phytochemistry **49,** 1029-1036.

Cadet, J., Anselmino, C., Douki, T., and Voituriez, L. (1992). Photochemistry of nucleic acids in cells. Journal of Photochemistry and Photobiology B: Biology **15,** 277-298.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biology **4**.

Caporale, G., Innocenti, G., Guiotto, A., Rodighiero, P., and Dall'Acqua, F. (1981). Biogenesis of linear O-alkylfuranocoumarins: A new pathway involving 5-hydroxymarmesin. Phytochemistry **20,** 1283-1287.

Cappelletti, E.M., Innocenti, G., and Caporale, G. (1984). Furocoumarins localization in the fruit and seed of Psoralea corylifolia L. Plantes Medicinales et Phytotherapie **18,** 181-189.

Carpinella, M.C., Ferrayoli, C.G., and Palacios, S.M. (2005). Antifungal synergistic effect of scopoletin, a hydroxycoumarin isolated from Melia azedarach L. fruits. Journal of Agricultural and Food Chemistry **53,** 2922-2927.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nature Reviews Genetics **3,** 285-298.

Carzoli, F.G., Michelotti, V., Fambrini, M., Salvini, M., and Pugliesi, C. (2008). Molecular Cloning and Organ-Specific Expression of Two Gibberellin 20-Oxidase Genes of Helianthus annuus. Plant Molecular Biology Reporter**,** 1-9.

Chapman, E.J., and Carrington, J.C. (2007). Specialization and evolution of endogenous small RNA pathways. Nature Reviews Genetics **8,** 884-896.

Chapple, C. (1998). Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. In Annual Review of Plant Biology, pp. 311-343.

Charrier, B., Coronado, C., Kondorosi, A., and Ratet, P. (1995). Molecular characterization and expression of alfalfa (Medicago sativa L.) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes. Plant Molecular Biology **29,** 773-786.

Charrier, B., Leroux, C., Kondorosi, A., and Ratet, P. (1996). The expression pattern of alfalfa flavanone 3-hydroxylase promoter-gus fusion in Nicotiana benthamiana correlates with the presence of flavonoids detected in situ. Plant Molecular Biology **30,** 1153-1168.

Chaudhary, S.K., Ceska, O., Warrington, P.J., and Ashwood-Smith, M.J. (1985). Increased furocoumarin content of celery during storage. Journal of Agricultural and Food Chemistry **33,** 1153-1157.

Chen, C.D., and Kemper, B. (1996). Different structural requirements at specific proline residue positions in the conserved proline-rich region of cytochrome P450 2C2. Journal of Biological Chemistry **271,** 28607-28611.

Chen, F., Ro, D.K., Petri, J., Gershenzon, J., Bohlmann, J., Pichersky, E., and Tholl, D. (2004). Characterization of a root-specific arabidopsis terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. Plant Physiology **135,** 1956-1966.

Chen, S., Glawischnig, E., Jorgensen, K., Naur, P., Jorgensen, B., Olsen, C.E., Hansen, C.H., Rasmussen, H., Pickett, J.A., and Halkier, B.A. (2003). CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in Arabidopsis. Plant Journal **33,** 923-937.

Chen, W., Zeng, H., and Tan, H. (2000). Cloning, sequencing, and function of sanF: A gene involved in nikkomycin biosynthesis of Streptomyces ansochromogenes. Current Microbiology **41,** 312-316.

Chen, Y., Wang, H.D., Xia, X., Kung, H.F., Pan, Y., and Kong, L.D. (2007). Behavioral and biochemical studies of total furocoumarins from seeds of Psoralea corylifolia in the chronic mild stress model of depression in mice. Phytomedicine **14,** 523-529.

Chiesa-Fuxench, Z.C., and González-Chávez, J. (2010). Extracorporeal photopheresis: A review on the immunological aspects and clinical applications. Puerto Rico Health Sciences Journal **29,** 337-347.

Chong, J., Baltz, R., Schmitt, C., Beffa, R., Fritig, B., and Saindrenan, P. (2002). Downregulation of a pathogen-responsive tobacco UDP-Glc:phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. Plant Cell **14,** 1093-1107.

Cianfrogna, J.A., Zangerl, A.R., and Berenbaum, M.R. (2002a). Effects of furanocoumarins on feeding behavior of parsnip webworms Depressaria pastinacella. Journal of Chemical Ecology **28,** 1365-1375.

Cianfrogna, J.A., Zangerl, A.R., and Berenbaum, M.R. (2002b). Dietary and developmental influences on induced detoxification in an oligophage. Journal of Chemical Ecology **28,** 1349-1364.

Clark, N.L., Aagaard, J.E., and Swanson, W.J. (2006). Evolution of reproductive proteins from animals and plants. Reproduction **131,** 11-22.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant Journal **16,** 735-743.

Cohen, M.B., Schuler, M.A., and Berenbaum, M.R. (1992). A host-inducible cytochrome P-450 from a host-specific caterpillar: Molecular cloning and evolution. Proceedings of the National Academy of Sciences of the United States of America **89,** 10920-10924.

Collu, G., Unver, N., Peltenburg-Looman, A.M.G., Van der Heijden, R., Verpoorte, R., and Memelink, J. (2001). Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. FEBS Letters **508,** 215-220.

Costet, L., Fritig, B., and Kauffmann, S. (2002). Scopoletin expression in elicitor-treated and tobacco mosaic virus-infected tobacco plants. Physiologia Plantarum **115,** 228-235.

Coxon, D.T., Curtis, R.F., Price, K.R., and Levett, G. (1973). Abnormal metabolites produced by Daucus carota roots stored under conditions of stress. Phytochemistry **12,** 1881-1885.

Cuca Suàrez, L.E., and Coy Barrera, C.A. (2007). Metabolites isolated from Esenbeckia alata (Karst & Triana) (Rutaceae). Biochemical Systematics and Ecology **35,** 386-388.

Curini, M., Cravotto, G., Epifano, F., and Giannone, G. (2006). Chemistry and biological activity of natural and synthetic prenyloxycoumarins. Current Medicinal Chemistry **13,** 199-222.

Cushnie, T.P.T., and Lamb, A.J. (2011). Recent advances in understanding the antibacterial properties of flavonoids. International Journal of Antimicrobial Agents **38,** 99-107.

Dall'Acqua, F., and Martelli, P. (1991). Photosensitizing action of furocoumarins on membrane components and consequent intracellular events. Journal of Photochemistry and Photobiology, B: Biology **8,** 235-254.

Dall'Acqua, F., Vedaldi, D., and Recher, M. (1978a). The photoreaction between furocoumarins and various DNA with different base compositions. Photochem. Photobiol. **27,** 33-36.

Dall'Acqua, F., Vedaldi, D., and Recher, M. (1978b). The photoreaction between furocoumarins and various DNA with different base compositions. Photochemistry and Photobiology **27,** 33-36.

Dall'Acqua, F., Capozzi, A., Marciani, S., and Caporale, G. (1972). Biosynthesis of furocoumarins: further studies on Ruta graveolens. Zeitschrift fur Naturforschung. Teil B. Anorganische Chemie, organische Chemie, Biochemie, Biophysik, Biologie **27,** 813-817.

Damjanic, A., and Akacic, B. (1974). Furocoumarins in Ficus carica. Planta Medica **26,** 119-123.

Dardalhon, M., De Massy, B., Nicolas, A., and Averbeck, D. (1998). Mitotic recombination and localized DNA double-strand breaks are induced after 8-methoxypsoralen and UVA irradiation in Saccharomyces cerevisiae. Current Genetics **34,** 30-42.

David, J.P., Boyer, S., Mesneau, A., Ball, A., Ranson, H., and Dauphin-Villemant, C. (2006). Involvement of cytochrome P450 monooxygenases in the response of mosquito larvae to dietary plant xenobiotics. Insect Biochemistry and Molecular Biology **36,** 410-420.

De Carolis, E., Chan, F., Balsevich, J., and De Luca, V. (1990). Isolation and characterization of a 2-oxoglutarate dependent dioxygenase involved in the second-to-last step in vindoline biosynthesis. Plant Physiology **94,** 1323-1329.

De Graaf, C., Vermeulen, N.P.E., and Feenstra, K.A. (2005). Cytochrome P450 in silico: An integrative modeling approach. Journal of Medicinal Chemistry **48,** 2725-2755.

Dercks, W., Trumble, J., and Winter, C. (1990). Impact of atmospheric pollution on linear furanocoumarin content in celery. Journal of Chemical Ecology **16,** 443-454.

Diesperger, H., Mueller, C.R., and Sandermann Jr, H. (1974). Rapid isolation of a plant microsomal fraction by MG2+ precipitation. FEBS Letters **43,** 155-158.

Dixon, R.A., and Paiva, N.L. (1995). Stress-induced phenylpropanoid metabolism. Plant Cell **7,** 1085-1097.

Djokic, L., Narancic, T., Nikodinovic-Runic, J., Savic, M., and Vasiljevic, B. (2011). Isolation and characterization of four novel Gram-positive bacteria associated with the rhizosphere of two endemorelict plants capable of degrading a broad range of aromatic substrates. Applied Microbiology and Biotechnology **91,** 1227-1238.

Dotzlaf, J.E., and Yeh, W.K. (1989). Purification and properties of deacetoxycephalosporin C synthase from recombinant Escherichia coli and its comparison with the native enzyme purified from Streptomyces clavuligerus. Journal of Biological Chemistry **264,** 10219-10227.

Durbin, M.L., McCaig, B., and Clegg, M.T. (2000). Molecular evolution of the chalcone synthase multigene family in the morning glory genome. Plant Molecular Biology **42,** 79-92.

Edelson, R.L. (1988). Light-activated drugs. Scientific American **259,** 68-75.

Ehlers, D., Pfister, M., Bork, W.R., and Toffel-Nadolny, P. (1995). HPLC analysis of tonka bean extracts. Zeitschrift fÃ¼r Lebensmittel-Untersuchung und -Forschung **201,** 278-282.

Ehrlich, P.R., and Raven, P.H. (1964). Butterflies and plants : a study in coevolution. Evolution **4,** 586-608.

Eiler, U. (1989). Elicitor induction of secondary metabolism in differenciated in vitro system of Ruta graveoelens. (Berlin: Springer-Verlag).

Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J., and Schofield, C.J. (2003). Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1α. Journal of Biological Chemistry **278,** 1802-1806.

Ellis, B.E., and Amrhein, N. (1971). The 'NIH-Shift' during aromatic ortho-hydroxylation in higher plants. Phytochemistry **10,** 3069-3072.

Endler, A., Martens, S., Wellmann, F., and Matern, U. (2008). Unusually divergent 4-coumarate:CoA-ligases from Ruta graveolens L. Plant Molecular Biology **67,** 335-346.

Evans, P.H., and Becerra, J.X. (2006). Non-terpenoid essential oils from Bursera chemapodicta. Flavour and Fragrance Journal **21,** 616-618.

Evans, P.H., Becerra, J.X., Venable, D.L., and Bowers, W.S. (2000). Chemical analysis of squirt-gun defense in Bursera and counterdefense by chrysomelid beetles. Journal of Chemical Ecology **26,** 745-754.

Facchini, P.J. (2001). Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications. In Annual Review of Plant Biology, pp. 29-66.

Fahrendorf, T., and Dixon, R.A. (1993). Stress responses in alfalfa (Medicago sativa L.) XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450. Archives of Biochemistry and Biophysics **305,** 509-515.

Fellenberg, C., Milkowski, C., Hause, B., Lange, P.R., Böttcher, C., Schmidt, J., and Vogt, T. (2008). Tapetum-specific location of a cation-dependent O-methyltransferase in Arabidopsis thaliana. Plant Journal **56,** 132-145.

Fink, G.R. (1998). Anatomy of a revolution. Genetics **149,** 473-477.

Fitzpatrick, T.B., and Pathak, M.A. (1959). Historical aspects of methoxsalen and other furocoumarins. The Journal of investigative dermatology **32,** 229-231.

Fliegmann, J., FurtwÃ¤ngler, K., Malterer, G., Cantarello, C., Schüler, G., Ebel, J., and Mithöfer, A. (2010). Flavone synthase II (CYP93B16) from soybean (Glycine max L.). Phytochemistry **71,** 508-514.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151,** 1531-1545.

Forslund, K., Morant, M., JÃ¸rgensen, B., Olsen, C.E., Asamizu, E., Sato, S., Tabata, S., and Bak, S. (2004). Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in Lotus japonicus. Plant Physiology **135,** 71-84.

Fraenkel, G.S. (1959). The raison d'être of secondary plant substances. Science **129,** 1466-1470.

Fraissinet-Tachet, L., Baltz, R., Chong, J., Kauffmann, S., Fritig, B., and Saindrenan, P. (1998). Two tobacco genes induced by infection, elicitor and salicylic acid encode glucosyltransferases acting on phenylpropanoids and benzoic acid derivatives, including salicylic acid. FEBS Letters **437,** 319-323.

Franke, R., Hemm, M.R., Denault, J.W., Ruegger, M.O., Humphreys, J.M., and Chapple, C. (2002a). Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis. Plant Journal **30,** 47-59.

Franke, R., Humphreys, J.M., Hemm, M.R., Denault, J.W., Ruegger, M.O., Cusumano, J.C., and Chapple, C. (2002b). The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant Journal **30,** 33-45.

Frazee, R.W., Orville, A.M., Dolbeare, K.B., Yu, H., Ohlendorf, D.H., and Lipscomb, J.D. (1998). The axial tyrosinate Fe3+ ligand in protocatechuate 3,4-dioxygenase influences substrate binding and product release: Evidence for new reaction cycle intermediates. Biochemistry **37,** 2131-2144.

Frear, D.S. (1995). Wheat microsomal cytochrome P450 monooxygenases: Characterization and importance in the metabolic detoxification and selectivity of wheat herbicides. Drug Metabolism and Drug Interactions **12,** 329-357.

Frey, M., Schullehner, K., Dick, R., Fiesselmann, A., and Gierl, A. (2009). Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. Phytochemistry **70,** 1645-1651.

Frey, M., Stettner, C., Paré, P.W., Schmelz, E.A., Tumlinson, J.H., and Gierl, A. (2000). An herbivore elicitor activates the gene for indole emission in maize. Proceedings of the National Academy of Sciences of the United States of America **97,** 14801-14806.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., and Gierl, A. (1997). Analysis of a chemical plant defense mechanism in grasses. Science **277,** 696-699.

Fritig, B., Hirth, L., and Ourisson, G. (1970). Biosynthesis of the coumarins: Scopoletin formation in tobacco tissue cultures. Phytochemistry **9,** 1963-1975.

Frugoli, J.A., McPeek, M.A., Thomas, T.L., and Robertson McClung, C. (1998). Intron loss and gain during evolution of the catalase gene family in angiosperms. Genetics **149,** 355-365.

Gabriac, B., Werck-Reichhart, D., Teutsch, H., and Durst, F. (1991). Purification and immunocharacterization of a plant cytochrome P450: The cinnamic acid 4-hydroxylase. Archives of Biochemistry and Biophysics **288,** 302-309.

Gattuso, G., Barreca, D., Caristi, C., Gargiulli, C., and Leuzzi, U. (2007). Distribution of flavonoids and furocoumarins in juices from cultivars of Citrus bergamia risso. Journal of Agricultural and Food Chemistry **55,** 9921-9927.

Gebhardt, Y., Witte, S., Forkmann, G., Lukacin, R., Matern, U., and Martens, S. (2005). Molecular evolution of flavonoid dioxygenases in the family Apiaceae. Phytochemistry **66,** 1273-1284.

Gestetner, B., and Conn, E.E. (1974). The 2 hydroxylation of trans cinnamic acid by chloroplasts from Melilotus alba Desr. Archives of Biochemistry and Biophysics **163,** 617-624.

Gierl, A., and Frey, M. (2001). Evolution of benzoxazinone biosynthesis and indole production in maize. Planta **213,** 493-498.

Girard, C., Roux, D., Muyard, F., Colombain, M., Tillequin, F., Waterman, P.G., and Bevalot, F. (2005). 3-Substituted coumarins from the twigs of Rhadinothamnus rudis ssp. amblycarpus. Zeitschrift fur Naturforschung - Section B Journal of Chemical Sciences **60,** 561-564.

Girennavar, B., Jayaprakasha, G.K., and Patil, B.S. (2007). Potent inhibition of human cytochrome P450 3A4, 2D6, and 2C9 isoenzymes by grapefruit juice and its furocoumarins. Journal of Food Science **72,** C417-C421.

Girennavar, B., Jayaprakasha, G.K., Jifon, J.L., and Patil, B.S. (2008a). Variation of bioactive furocoumarins and flavonoids in different varieties of grapefruits and pummelo. European Food Research and Technology **226,** 1269-1275.

Girennavar, B., Cepeda, M.L., Soni, K.A., Vikram, A., Jesudhasan, P., Jayaprakasha, G.K., Pillai, S.D., and Patil, B.S. (2008b). Grapefruit juice and its furocoumarins inhibits autoinducer signaling and biofilm formation in bacteria. International Journal of Food Microbiology **125,** 204-208.

Godar, D.E. (1999). UVA1 radiation triggers two different final apoptotic pathways. Journal of Investigative Dermatology **112,** 3-12.

Godar, D.E. (2008). Can dietary furocoumarins really be responsible for the increase in melanoma? Medical Hypotheses **71,** 613-614.

Gomez-Vasquez, R., Day, R., Buschmann, H., Randles, S., Beeching, J.R., and Cooper, R.M. (2004). Phenylpropanoids, phenylalanine ammonia lyase and peroxidases in elicitor-challenged cassava (Manihot esculenta) suspension cells and leaves. Annals of Botany **94,** 87-97.

Gorgus, E., Lohr, C., Raquet, N., Guth, S., and Schrenk, D. (2010). Limettin and furocoumarins in beverages containing citrus juices or extracts. Food and Chemical Toxicology **48,** 93-98.

Gotoh, O. (1992). Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. Journal of Biological Chemistry **267,** 83-90.

Graur, D., and Li, W.-H. (1999). Fundamentals of molecular evolution. (Sunderland, MA, USA: Sinauer).

Gravot, A. (2002). Etude de P450 impliqués dans la biosynthèse des furocoumarines chez Ruta graveolens. In Laboratoire Agronomie Environnement - ENSAI/INPL (Vandoeuvre-lès-Nancy: Institut National Polytecnique de Lorraine).

Gravot, A., Larbat, R., Hehn, A., Lievre, K., Gontier, E., Goergen, J.L., and Bourgaud, F. (2004). Cinnamic acid 4-hydroxylase mechanism-based inactivation by psoralen derivatives: Cloning and characterization of a C4H from a psoralen producing plant - Ruta graveolens - Exhibiting low sensitivity to psoralen inactivation. Archives of Biochemistry and Biophysics **422,** 71-80.

Grennan, A.K. (2006). Gibberellin metabolism enzymes in rice. Plant Physiology **141,** 524-526.

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 hydroxycinnamoyle spermidines. Plant Journal **58,** 246-259.

Guyot, A.D., Farhi, D., Ingen-Housz-Oro, S., Bussel, A., Parquet, N., Rabian, C., Bachelez, H., and Francès, C. (2007). Treatment of refractory erosive oral lichen planus with extracorporeal photochemotherapy: 12 Cases. British Journal of Dermatology **156,** 553-556.

Hadacek, F., Muller, C., Werner, A., Greger, H., and Proksch, P. (1994). Analysis, isolation and insecticidal activity of linear furanocoumarins and other coumarin derivatives from Peucedanum (Apiaceae: Apioideae). Journal of Chemical Ecology **20,** 2035-2054.

Hagel, J.M., and Facchini, P.J. (2010). Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. Nature Chemical Biology **6,** 273-275.

Hale, A.L., Meepagala, K.M., Oliva, A., Aliotta, G., and Duke, S.O. (2004). Phytotoxins from the leaves of Ruta graveolens. Journal of Agricultural and Food Chemistry **52,** 3345-3349.

Halkier, B.A., and Du, L. (1997). The biosynthesis of glucosinolates. Trends in Plant Science **2,** 425-431.

Halkier, B.A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. In Annual Review of Plant Biology, pp. 303-333.

Hamerski, D., and Matern, U. (1988a). Elicitor-induced biosynthesis of psoralens in Ammi majus L. suspension cultures. Microsomal conversion of demethylsuberosin into (+)marmesin and psoralen. European Journal of Biochemistry **171,** 369-375.

Hamerski, D., and Matern, U. (1988b). Biosynthesis of psoralens Psoralen 5-monooxygenase activity from elicitor-treated Ammi majus cells. FEBS Letters **239,** 263-265.

Hamerski, D., Schmitt, D., and Matern, U. (1990 a). Induction of two prenyltransferases for the accumulation of coumarin phytoalexins in elicitor-treated Ammi majus cell suspension cultures. Phytochemistry **29,** 1131-1135.

Hamerski, D., Beier, R.C., Kneusel, R.E., Matern, U., and Himmelspacht, K. (1990 b). Accumulation of coumarins in elicitor-treated cell suspension cultures of Ammi majus. Phytochemistry **29,** 1137-1142.

Harborne, J.B. (1971). Flavonoid and phenylpropanoid patterns in the Umbelliferae. (London: Accademic press).

Hashimoto, T., and Yamada, Y. (1987). Purification and characterization of hyoscyamine 6 beta-hydroxylase from root cultures of Hyoscyamus niger L. Hydroxylase and epoxidase activities in the enzyme preparation. European Journal of Biochemistry **164,** 277-285.

Hausinger, R.P. (2004). Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. Critical Reviews in Biochemistry and Molecular Biology **39,** 21-68.

Hedden, P. (1999). Recent advances in gibberellin biosynthesis. Journal of Experimental Botany **50,** 553-563.

Hedden, P., and Phillips, A.L. (2000). Gibberellin metabolism: New insights revealed by the genes. Trends in Plant Science **5,** 523-530.

Hehmann, M., Lukacin, R., Ekiert, H., and Matern, U. (2004). Furanocoumarin biosynthesis in Ammi majus L. Cloning of bergaptol O-methyltransferase. European Journal of Biochemistry **271,** 932-940.

Helariutta, Y., Kotilainen, M., Elomaa, P., Kalkkinen, N., Bremer, K., Teeri, T.H., and Albert, V.A. (1996). Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: Evolution with substrate change and catalytic simplification. Proceedings of the National Academy of Sciences of the United States of America **93,** 9033-9038.

Hinman, V.F., and Davidson, E.H. (2007). Evolutionary plasticity of developmental gene regulatory network architecture. Proceedings of the National Academy of Sciences of the United States of America **104,** 19404-19409.

Hivelin, M., Siemionow, M., Grimbert, P., and Lantieri, L. (2009). Extracorporeal photopheresis: From solid organs to face transplantation. Transplant Immunology **21,** 117-128.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B., and Legrand, M. (2004). Silencing of hydroxycinnamoyle-coenzyme A shikimate/quinate hydroxycinnamoyletransferase affects phenylpropanoid biosynthesis. The Plant Cell **16,** 1446-1465.

Hoffmeister, D., and Keller, N.P. (2007). Natural products of filamentous fungi: Enzymes, genes, and their regulation. Natural Product Reports **24,** 393-416.

Hogge, L.R., Reed, D.W., Underhill, E.W., and Haughn, G.W. (1988). HPLC separation of glucosinolates from leaves and seeds of Arabidopsis thaliana and their identification using thermospray liquid chromatography/mass spectrometry. Journal of Chromatographic Science **26,** 551-556.

Holton, T.A., Brugliera, F., and Tanaka, Y. (1993). Cloning and expression of flavonol synthase from Petunia hybrida. Plant Journal **4,** 1003-1010.

Hubner, S., Hehmann, M., Schreiner, S., Martens, S., Lukacin, R., and Matern, U. (2003). Functional expression of cinnamate 4-hydroxylase from Ammi majus L. Phytochemistry **64,** 445-452.

Hughes, A.L. (1994). The evolution of functionally novel proteins after gene duplication. Proceedings of the Royal Society B: Biological Sciences **256,** 119-124.

Hughes, A.L. (2002). Adaptive evolution after gene duplication. Trends in Genetics **18,** 433-434.

Hughes, M.K., and Hughes, A.L. (1993). Evolution of duplicate genes in a tetraploid animal, Xenopus laevis. Molecular Biology and Evolution **10,** 1360-1369.

Hung, C.-F., Prapaipong, H., Berenbaum, M.R., and Schuler, M.A. (1995). Differential induction of cytochrome P450 transcripts in Papilio polyxenes by linear and angular furanocoumarins. Insect Biochemistry and Molecular Biology **25,** 89-99.

Hurst, L.D., Pal, C., and Lercher, M.J. (2004). The evolutionary dynamics of eukaryotic gene order. Nature Reviews Genetics **5,** 299-310.

Hutton Jr, J.J., Tappel, A.L., and Udenfriend, S. (1966). Requirements for α-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase. Biochemical and Biophysical Research Communications **24,** 179-184.

Ilg, A., Beyer, P., and Al-Babili, S. (2009). Characterization of the rice carotenoid cleavage dioxygenase 1 reveals a novel route for geranial biosynthesis. FEBS Journal **276,** 736-747.

Innocenti, G., Dall'Acqua, F., Rodighiero, P., and Caporale, G. (1978). Biosynthesis of O-alkylfurocoumarins in Angelica archangelica. PLANTA MED. **34,** 167-171.

Innocenti, G., Bourgaud, F., Piovan, A., and Favretto, D. (1997). Furocoumarins and other secondary metabolites from Psoralea canescens. Pharmaceutical Biology **35,** 232-236.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J.M., Lane, W.S., and Kaelin W.G, Jr. (2001). HIFα targeted for VHL-mediated destruction by proline hydroxylation: Implications for O2 sensing. Science **292,** 464-468.

J.Kagan. (1993). Organic Photochemistry. Principles and Applications, A. presse, ed (London.

Jacob, F., Perrin, D., Sanchez, C., and Monod, J. (2005). The operon: A group of genes with expression coordinated by an operator. L'opéron: Groupe de gènes à expression coordonnée par un opérateur [C. R. Acad. Sci. Paris 250 (1960) 1727-1729] **328,** 514-520.

Jagasia, M.H., Savani, B.N., Stricklin, G., Engelhardt, B., Kassim, A., Dixon, S., Chen, H., Chinratanalab, W., Goodman, S., Greer, J.P., and Schuening, F. (2009).

Classic and Overlap Chronic Graft-versus-Host Disease (cGVHD) Is Associated with Superior Outcome after Extracorporeal Photopheresis (ECP). Biology of Blood and Marrow Transplantation **15,** 1288-1295.

Janc, J.W., Egan, L.A., and Townsend, C.A. (1995). Purification and characterization of clavaminate synthase from Streptomyces antibioticus. A multifunctional enzyme of clavam biosynthesis. Journal of Biological Chemistry **270,** 5399-5404.

Janda, P.H., Bhambri, S., Del Rosso, J.Q., and Mobini, N. (2008). Phytophotodermatitis: Case report and review of the literature. Cosmetic Dermatology **21,** 99-102.

Jansen, G.A., Mihalik, S.J., Watkins, P.A., Moser, H.W., Jakobs, C., Denis, S., and Wanders, R.J.A. (1996). Phytanoyl-CoA hydroxylase is present in human liver, located in peroxisomes, and deficient in Zellweger syndrome: Direct, unequivocal evidence for the new, revised pathway of phytanic acid α-oxidation in humans. Biochemical and Biophysical Research Communications **229,** 205-210.

Janzen, D.H., and Muller, C.H. (1969). Coevolution. Science **165,** 415-416.

Jeong, H.K., Young, M.C., Kim, B.G., and Ahn, J.H. (2008). Analysis of flavonoids and characterization of the OsFNS gene involved in flavone biosynthesis in rice. Journal of Plant Biology **51,** 97-101.

Jia, S., VanDusen, W.J., Diehl, R.E., Kohl, N.E., Dixon, R.A.F., Elliston, K.O., Stern, A.M., and Friedman, P.A. (1992). cDNA cloning and expression of bovine aspartyl (asparaginyl) Î²- hydroxylase. Journal of Biological Chemistry **267,** 14322-14327.

Jiang, H., and Morgan, J.A. (2004). Optimization of an in Vivo Plant P450 Monooxygenase System in Saccharomyces cerevisiae. Biotechnology and Bioengineering **85,** 130-137.

Kaessmann, H., Vinckenbosch, N., and Long, M. (2009). RNA-based gene duplication: Mechanistic and evolutionary insights. Nature Reviews Genetics **10,** 19-31.

Kagan.J. (1993). Organic Photochemistry. Principles and Applications, A. presse, ed (London.

Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K., and Shimizu, B.I. (2008). Scopoletin is biosynthesized via ortho-hydroxylation of féruloyle CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. Plant Journal **55,** 989-999.

Kakkar, R.K., and Rai, V.K. (1993). Plant polyamines in flowering and fruit ripening. Phytochemistry **33,** 1281-1288.

Kaminski, M., Kartanowicz, R., Kaminski, M.M., Krolicka, A., Sidwa-Gorycka, M., Lojkowska, E., and Gorzen, W. (2003). HPLC-DAD in identification and quantification of selected coumarins in crude extracts from plant cultures of Ammi majus and Ruta graveolens. Journal of Separation Science **26,** 1287-1291.

Kandel, S., Sauveplane, V., Olry, A., Diss, L., Benveniste, I., and Pinot, F. (2006). Cytochrome P450-dependent fatty acid hydroxylases in plants. Phytochemistry Reviews **5,** 359-372.

Kang, H.K., Shin, E.J., and Shim, S.C. (1992). Transient absorption spectra and quenching of coumarin excited states by nucleic acid bases. Journal of Photochemistry and Photobiology B: Biology **13,** 19-28.

Kang, S.W., Kim, C.Y., Song, D.G., Pan, C.H., Cha, K.H., Lee, D.U., and Um, B.H. (2010). Rapid identification of furanocoumarins in Angelica dahurica using the online LC-MMR-MS and their nitric oxide inhibitory activity in RAW 264.7 cells. Phytochemical Analysis **21,** 322-327.

Kawase, M., Sakagami, H., Motohashi, N., Hauer, H., Chatterjee, S.S., Spengler, G., Vigyikanne, A.V., MolnÃ¡r, A., and MolnÃ¡r, J. (2005). Coumarin derivatives with tumor-specific cytotoxicity and multidrug resistance reversal activity. In Vivo **19,** 705-712.

Keating, G.J.a.O.K., R. (1997). The chemistry and occurrence of coumarins. In In Coumarins: Biology, Applications and Mode of Action, R.a.T. O'Kennedy, R.D., eds, ed (Chichester: Chichester), pp. pp. 23–66.

Keeling, C.I., and Bohlmann, J. (2006). Diterpene resin acids in conifers. Phytochemistry **67,** 2415-2423.

Kemper, B. (2004). Structural basis for the role in protein folding of conserved proline-rich regions in cytochromes P450. Toxicology and Applied Pharmacology **199,** 305-315.

Kindl, H. (1971). Ortho-hydroxylation of aromatic carboxylic acids in higher plants. Hoppe-Seyler's Zeitschrift fur Physiologische Chemie **352,** 78-84.

Kitamura, N., Kohtani, S., and Nakagaki, R. (2005). Molecular aspects of furocoumarin reactions: Photophysics, photochemistry, photobiology, and structural analysis. Journal of Photochemistry and Photobiology C: Photochemistry Reviews **6,** 168-185.

Kitamura, Y., Ikenaga, T., Ooe, Y., Hiraoka, N., and Mizukami, H. (1998). Induction of furanocoumarin biosynthesis in Glehnia littoralis cell suspension cultures by elicitor treatment. Phytochemistry **48,** 113-117.

Kivirikko, K.I., and Pihlajaniemi, T. (1998). Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. Advances in Enzymology and Related Areas of Molecular Biology **72,** 325-398.

Kliebenstein, D.J., Kroymann, J., and Mitchell-Olds, T. (2005). The glucosinolate-myrosinase system in an ecological and evolutionary context. Current Opinion in Plant Biology **8,** 264-271.

Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J., and Mitchell-Olds, T. (2001). Gene duplication in the diversification of secondary metabolism: Tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in arabidopsis. Plant Cell **13,** 681-693.

Klingenberg, M. (1958). Pigments of rat liver microsomes. Archives of Biochemistry and Biophysics **75,** 376-386.

Koenigs, L.L., and Trager, W.F. (1998). Mechanism-based inactivation of cytochrome P450 2B1 by 8-methoxypsoralen and several other furanocoumarins. Biochemistry **37,** 13184-13193.

Korioth, F., Gieffers, C., and Frey, J. (1994). Cloning and characterization of the human gene encoding aspartyl β-hydroxylase. Gene **150,** 395-399.

Kosuge, T. (1961). Studies on the identity of bound coumarin in sweet clover. Archives of Biochemistry and Biophysics **95,** 211-218.

Kriechbaumer, V., Weigang, L., FieÃŸelmann, A., Letzel, T., Frey, M., Gierl, A., and Glawischnig, E. (2008). Characterisation of the tryptophan synthase alpha subunit in maize. BMC Plant Biology **8**.

Krueger, J.G. (2002). The immunologic basis for the treatment of psoriasis with new biologic agents. Journal of the American Academy of Dermatology **46,** 1-23.

Kung, A.C., Stephens, M.B., and Darling, T. (2009). Phytophotodermatitis: Bulla formation and hyperpigmentation during spring break. Military Medicine **174,** 657-661.

Lange, T., Hedden, P., and Graebe, J.E. (1994a). Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. Proceedings of the National Academy of Sciences of the United States of America **91,** 8552-8556.

Lange, T., Schweimer, A., Ward, D.A., Hedden, P., and Graebe, J.E. (1994b). Separation and characterisation of three 2-oxoglutarate-dependent dioxygenases

from Cucurbita maxima L. endosperm involved in gibberellin biosynthesis. Planta **195,** 98-107.

Larbat, R. (2006). Contribution à l'étude des P450 impliqués dans la biosynthèse des furocoumarines. In Laboratoire Agronomie Environnement ENSAIA/INRA (Vandoeuvre-lès-Nancy: Institut National Poytechnique de Lorraine).

Larbat, R., Kellner, S., Specker, S., Hehn, A., Gontier, E., Hans, J., Bourgaud, F., and Matern, U. (2007). Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis. Journal of Biological Chemistry **282,** 542-554.

Larbat, R., Hehn, A., Hans, J., Schneider, S., Jugde, H., Schneider, B., Matern, U., and Bourgaud, F. (2009). Isolation and functional characterization of CYP71AJ4 encoding for the first P450 monooxygenase of angular furanocoumarin biosynthesis. Journal of Biological Chemistry **284,** 4776-4785.

Latunde-Dada, A.O., Cabello-Hurtado, F., Czittrich, N., Didierjean, L., Schopfer, C., Hertkorn, N., Werck-Reichhart, D., and Ebel, J. (2001). Flavonoid 6-hydroxylase from soybean (Glycine max L.), a novel plant P-450 monooxygenase. Journal of Biological Chemistry **276,** 1688-1695.

Lee, S., Badieyan, S., Bevan, D.R., Herde, M., Gatz, C., and Tholl, D. (2010). Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **107,** 21205-21210.

Leiss, K.A., Maltese, F., Choi, Y.H., Verpoorte, R., and Klinkhamer, P.G.L. (2009). Identification of Chlorogenic Acid as a Resistance Factor for Thrips in Chrysanthemum. Plant Physiology **150,** 1567-1575.

Lerner, A.B., Denton, C.R., and Fitzpatrick, T.B. (1953). Clinical and experimental studies with 8-methoxypsoralen in vitiligo. The Journal of investigative dermatology **20,** 299-314.

Lièvre, K., Hehn, A., Tran, T.L.M., Gravot, A., Thomasset, B., Bourgaud, F., and Gontier, E. (2005). Genetic transformation of the medicinal plant Ruta graveolens L. by an Agrobacterium tumefaciens-mediated method. Plant Science **168,** 883-888.

Lièvre. (2004). Modification de la composition en molécules pharmaceutiques (furocoumairnes) de la Rue officinale (Ruta graveoelens) par transformation génétique. In Laboratoire Agronomie et Environnement ENSAI/INRA (Vandoeuvre-lès-Nancy: Institut National Polytechnique de Lorraine).

Liu, H., Jensen, K.G., Tran, L.M., Chen, M., Zhai, L., Olsen, C.E., SÃ¸hoel, H., Denmeade, S.R., Isaacs, J.T., and BrÃ¸gger Christensen, S. (2006). Cytotoxic phenylpropanoids and an additional thapsigargin analogue isolated from Thapsia garganica. Phytochemistry **67,** 2651-2658.

Liu, S., Hu, Y., Wang, X., Han, L., Song, S., Cheng, H., and Lin, Z. (2009). Isolation and characterization of a gene encoding cinnamate 4-hydroxylase from Parthenocissus henryana. Molecular Biology Reports **36,** 1605-1610.

Llano, J., Raber, J., and Eriksson, L.A. (2003). Theoretical study of phototoxic reactions of psoralens. Journal of Photochemistry and Photobiology A: Chemistry **154,** 235-243.

Lloyd, M.D., Merritt, K.D., Lee, V., Sewell, T.J., Wha-Son, B., Baldwin, J.E., Schofield, C.J., Elson, S.W., Baggaley, K.H., and Nicholson, N.H. (1999). Product-substrate engineering by bacteria: Studies on clavaminate synthase, a trifunctional dioxygenase. Tetrahedron **55,** 10201-10220.

Logemann, E., Tavernaro, A., Schulz, W., Somssich, I.E., and Hahlbrock, K. (2000). UV light selectively coinduces supply pathways from primary metabolism and flavonoid secondary product formation in parsley. Proceedings of the National Academy of Sciences of the United States of America **97,** 1903-1907.

Lombaert, G.A., Siemens, K.H., Pellaers, P., Mankotia, M., and Ng, W. (2001). Furanocoumarins in celery and parsnips: Method and multiyear Canadian survey. Journal of AOAC International **84,** 1135-1143.

Long, M., Betran, E., Thornton, K., and Wang, W. (2003). The origin of new genes: Glimpses from the young and old. Nature Reviews Genetics **4,** 865-875.

Lukacin, R., and Britsch, L. (1997). Identification of strictly conserved histidine and arginine residues as part of the active site in Petunia hybrida flavanone 3Î²-hydroxylase. European Journal of Biochemistry **249,** 748-757.

Lukacin, R., Wellmanna, F., Britsch, L., Martens, S., and Matern, U. (2003). Flavonol synthase from Citrus unshiu is a bifunctional dioxygenase. Phytochemistry **62,** 287-292.

Lukačin, R., Gröning, I., Pieper, U., and Matern, U. (2000). Site-directed mutagenesis of the active site serine290 in flavanone 3β-hydroxylase from Petunia hybrida. European Journal of Biochemistry **267,** 853-860.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290,** 1151-1155.

Lynch, M., and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. Genetics **154,** 459-473.

Lynch, M., and Conery, J.S. (2003). The evolutionary demography of duplicate genes. Journal of Structural and Functional Genomics **3,** 35-44.

Mahesh, V., Million-Rousseau, R., Ullmann, P., Chabrillange, N., Bustamante, J., Mondolot, L., Morant, M., Noirot, M., Hamon, S., De Kochko, A., Werck-Reichhart, D., and Campa, C. (2007). Functional characterization of two p-coumaroyle ester 3â€²-hydroxylase genes from coffee tree: Evidence of a candidate for chlorogenic acid biosynthesis. Plant Molecular Biology **64,** 145-159.

Maier, W., Schmidt, J., Nimtz, M., Wray, V., and Strack, D. (2000). Secondary products in mycorrhizal roots of tobacco and tomato. Phytochemistry **54,** 473-479.

Mansuy, D. (1998). The great diversity of reactions catalyzed by cytochromes P450. Comparative Biochemistry and Physiology - C Pharmacology Toxicology and Endocrinology **121,** 5-14.

Mao, W., Schuler, M.A., and Berenbaum, M.R. (2007). Cytochrome P450s in Papilio multicaudatus and the transition from oligophagy to polyphagy in the Papilionidae. Insect Molecular Biology **16,** 481-490.

Mao, W., Berenbaum, M.R., and Schuler, M.A. (2008). Modifications in the N-terminus of an insect cytochrome P450 enhance production of catalytically active protein in baculovirus-Sf9 cell expression systems. Insect Biochemistry and Molecular Biology **38,** 66-75.

Mari, M., and Cederbaum, A.I. (2001). Induction of catalase, alpha, and microsomal glutathione S-transferase in CYP2E1 overexpressing HepG2 cells and protection against short-term oxidative stress. Hepatology **33,** 652-661.

Marzano, C., Chilin, A., Bordin, F., Baccichetti, F., and Guiotto, A. (2002). DNA damage and biological effects induced by photosensitization with new N1-unsubstituted furo[2,3-h]quinolin-2(1H)-ones. Bioorganic and Medicinal Chemistry **10,** 2835-2844.

Massot, B., Milesi, S., Gontier, E., Bourgaud, F., and Guckert, A. (2000). Optimized culture conditions for the production of furanocoumarins by micropropagated shoots of Ruta graveolens. Plant Cell, Tissue and Organ Culture **62,** 11-19.

Matern, U. (1991). Coumarins and other phenylpropanoid compounds in the defense response of plant cells. Planta Medica **57,** S15-S20.

Matsumoto, S., Mizutani, M., Sakata, K., and Shimizu, B.I. (2011). Molecular cloning and functional analysis of the ortho-hydroxylases of p-coumaroyle coenzyme A/féruloyle coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato, Ipomoea batatas (L.) Lam. Phytochemistry.

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.

McCloud, E.S., Berenbaum, M.R., and Tuveson, R.W. (1992). Furanocoumarin content and phototoxicity of rough lemon (Citrus jambhiri) foliage exposed to enhanced ultraviolet-B (UVB) irradiation. Journal of Chemical Ecology **18,** 1125-1137.

Meurer-Grimes, B. (1995). New evidence for the systematic significance of acylated spermidines and flavonoids in pollen of Higher Hamamelidae. Brittonia **47,** 130-142.

Meyer, K., Cusumano, J.C., Somerville, C., and Chapple, C.C.S. (1996). Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases. Proceedings of the National Academy of Sciences of the United States of America **93,** 6869-6874.

Michalak, P. (2008). Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. Genomics **91,** 243-248.

Middleton, E.M., and Teramura, A.H. (1993). The role of flavonol glycosides and carotenoids in protecting soybean from ultraviolet-B damage. Plant Physiology **103,** 741-752.

Miles, E.W. (2001). Tryptophan synthase: A multienzyme complex with an intramolecular tunnel. Chemical Records **1,** 140-151.

Milesi, S. (2001). Etude de la production de furocoumarines par des plantes du genre *Ruta* cultivées au champ et des suspensions cellulaires de *Ruta graveolens* In Laboratoire Agronomie Environnement - ENSAIA/INRA (Vandoeuvre-lès-Nancy: Institut National Polytechnique de Lorraine).

Milesi, S., Massot, B., Gontier, E., Bourgaud, F., and Guckert, A. (2001). Ruta graveolens L.: A promising species for the production of furanocoumarins. Plant Science **161,** 189-199.

Miller, K.G., Poole, C.F., and Pawlowski, T.M.P. (1996). Classification of the botanical origin of cinnamon by solid-phase microextraction and gas chromatography. Chromatographia **42,** 639-646.

Mizutani, M., and Ohta, D. (2010). Diversification of P450 genes during land plant evolution. In Annual Review of Plant Biology, pp. 291-315.

Mizutani, M., Ohta, D., and Sato, R. (1993a). Purification and characterization of a cytochrome p450 (trans-cinnamic acid 4-hydroxylase) from etiolated mung bean seedlings. Plant and Cell Physiology **34,** 481-488.

Mizutani, M., Ward, E., and Ohta, D. (1998). Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differential expression, and RFLP mapping of multiple cytochromes P450. Plant Mol Biol **37,** 39-52.

Mizutani, M., Ward, E., DiMaio, J., Ohta, D., Ryals, J., and Sato, R. (1993b). Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. Biochemical and Biophysical Research Communications **190,** 875-880.

Moglia, A., Comino, C., Portis, E., Acquadro, A., De Vos, R.C.H., Beekwilder, J., and Lanteri, S. (2009). Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. Plant Cell Reports **28,** 963-974.

Morant, M., Hehn, A., and Werck-Reichhart, D. (2002). Conservation and diversity of gene families explored using the CODEHOP strategy in higher plants. BMC Plant Biology **2**.

Morant, M., Schoch, G.A., Ullmann, P., Ertunc, T., Little, D., Olsen, C.E., Petersen, M., Negrel, J., and Werck-Reichhart, D. (2007). Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. Plant Molecular Biology **63,** 1-19.

Morikawa, T., Mizutani, M., Aoki, N., Watanabe, B., Saga, H., Saito, S., Oikawa, A., Suzuki, H., Sakurai, N., Shibata, D., Wadano, A., Sakata, K., and Ohta, D. (2006). Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. Plant Cell **18,** 1008-1022.

Müller, I., Kahnert, A., Pape, T., Sheldrick, G.M., Meyer-Klaucke, W., Dierks, T., Kertesz, M., and UsÃ³n, I. (2004). Crystal Structure of the Alkylsulfatase AtsK: Insights into the Catalytic Mechanism of the Fe(II) α-Ketoglutarate-Dependent Dioxygenase Superfamily. Biochemistry **43,** 3075-3088.

Murray, E.E., Lotzer, J., and Eberle, M. (1989). Codon usage in plant genes. Nucleic Acids Research **17,** 477-498.

Murray, M. (2000). Mechanisms of Inhibitory and Regulatory Effects of Methylenedioxyphenyl Compounds on Cytochrome P450-Dependent Drug Oxidation. Current Drug Metabolism **1,** 67-84.

Musajo, L., Rodighiero, G., and Caporale, G. (1954). Photodynamic activity of natural coumarin derivatives. Bulletin de la Société de chimie biologique **36,** 1213-1224.

Musajo, L., Rodighiero, G., Colombo, G., Torlone, V., and Dall'Acqua, F. (1965). Photosensitizing furocoumarins: Interaction with DNA and photo-inactivation of DNA containing viruses. Experientia **21,** 22-24.

Nafisi, M., Goregaoker, S., Botanga, C.J., Glawischnig, E., Olsen, C.E., Halkier, B.A., and Glazebrook, J. (2007). Arabidopsis cytochrome P450 monooxygenase 71A13 catalyzes the conversion of indole-3-acetaldoxime in camalexin synthesis. Plant Cell **19,** 2039-2052.

Nair, R.B., Xia, Q., Kartha, C.J., Kurylo, E., Hirji, R.N., Datla, R., and Selvaraj, G. (2002). Arabidopsis CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. Plant Physiol **130,** 210-220.

Neal, J.J. (1994). Inhibition of insect cytochromes P450 by furanocoumarins. Pesticide Biochemistry and Physiology **50,** 43-50.

Nelson, D.R. (1999). Cytochrome P450 and the individuality of species. Archives of Biochemistry and Biophysics **369,** 1-10.

Nelson, D.R., Schuler, M.A., Paquette, S.M., Werck-Reichhart, D., and Bak, S. (2004). Comparative genomics of rice and arabidopsis. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. Plant Physiology **135,** 756-772.

Nelson, D.R., Kamataki, T., Waxman, D.J., Guengerich, F.P., Estabrook, R.W., Feyereisen, R., Gonzalez, F.J., Coon, M.J., Gunsalus, I.C., Gotoh, O., Okuda, K., and Nebert, D.W. (1993). The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA and Cell Biology **12,** 1-51.

Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C., and Nebert, D.W. (1996). P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics **6,** 1-42.

Nguyen. (1992). Contribution à l'étude de la production de psoralène (Furocoumarine) par la culture in vitro de plantes du genre psoralea (Leguminosae) (Nancy: Institue National Polytechnique de Lorraine), pp. 161.

Nishiyama, Y., Yun, C.S., Matsuda, F., Sasaki, T., Saito, K., and Tozawa, Y. (2010). Expression of bacterial tyrosine ammonia-lyase creates a novel p-coumaric acid pathway in the biosynthesis of phenylpropanoids in Arabidopsis. Planta **232,** 209-218.

Nitao, J.K., and Zangerl, A.R. (1987). Floral development and chemical defense allocation in wild parsnip (Pastinaca sativa). Ecology **68,** 521-529.

Nograles, K.E., Davidovici, B., and Krueger, J.G. (2010). New Insights in the Immunologic Basis of Psoriasis. Seminars in Cutaneous Medicine and Surgery **29,** 3-9.

Nur Fariza, M.S., Pang, S.L., Choong, C.Y., and Wickneswari, R. (2008). Extensive DNA sequence variations in two lignin genes, Cinnamate 4-hydroxylase and Cinnamyl Alcohol Dehydrogenase from Acacia mangium and Acacia auriculiformis. Journal of Biological Sciences **8,** 687-690.

Oba, K., Conn, E.E., Canut, H., and Boudet, A.M. (1981). Subcellular localization of 2-(β-D-glucosyloxy)-cinnamic acids and related β-glucosidase in leaves of Melilotus alba Desr. Plant Physiology **68,** 1359-1363.

Ober, D. (2005). Seeing double: Gene duplication and diversification in plant secondary metabolism. Trends in Plant Science **10,** 444-449.

Ober, D. (2010). Gene duplications and the time thereafter - examples from plant secondary metabolism. Plant Biology **12,** 570-577.

Ogunkanmi A.B., Tucker D.J., and T.A., M. (1973). An improved bio-assay for abscisic acid and other antitranspirants. New Phytol. **72,** 277-282.

Ohno, S. (1970). Evolution by gene duplication. (Berlin, Germany: Springer).

Ohno, S. (1973). Ancient linkage groups and frozen accidents. Nature **244,** 259-262.

Ojala, T., Remes, S., Haansuu, P., Vuorela, H., Hiltunen, R., Haahtela, K., and Vuorela, P. (2000). Antimicrobial activity of some coumarin containing herbal plants growing in Finland. Journal of Ethnopharmacology **73,** 299-305.

Okada, A., Okada, K., Miyamoto, K., Koga, J., Shibuya, N., Nojiri, H., and Yamane, H. (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. Journal of Biological Chemistry **284,** 26510-26518.

Oliveira, A.P., Valentao, P., Pereira, J.A., Silva, B.M., Tavares, F., and Andrade, P.B. (2009). Ficus carica L.: Metabolic and biological screening. Food and Chemical Toxicology **47,** 2841-2846.

Oliveira, F.M., Santana, A.E.G., Conserva, L.M., Maia, J.G.S., and Guilhon, G.M.P. (1996). Alkaloids and coumarins from Esenbeckia species. Phytochemistry **41,** 647-649.

Olry, A., Schneider-Belhaddad, F., Heintz, D., and Werck-Reichhart, D. (2007). A medium-throughput screening assay to determine catalytic activities of oxygen-consuming enzymes: A new tool for functional characterization of cytochrome P450 and other oxygenases. Plant Journal **51,** 331-340.

Omura, T., and Sato, R. (1962). A new cytochrome in liver microsomes. The Journal of biological chemistry **237,** 1375-1376.

Omura, T., and Sato, R. (1964). The carbon monoxide-binding pigment of liver microsomes.I. Evidence. The Journal of biological chemistry **239,** 2370-2378.

Orlita, A., Sidwa-Gorycka, M., Kumirska, J., Malinski, E., Siedlecka, E.M., Gajdus, J., Lojkowska, E., and Stepnowski, P. (2008). Identification of Ruta graveolens L. metabolites accumulated in the presence of abiotic elicitors. Biotechnology Progress **24,** 128-133.

Osbourn, A.E., and Field, B. (2009). Operons. Cellular and Molecular Life Sciences **66,** 3755-3775.

Paine, M.F., Scrutton, N., Munro, A., Gutierrez, A., Roberts, G., and Wolf, C.R. (2005). Cytochrome P450: Structure, Mechanism and Biochemistry. (New York: plenum).

Palmer, J.M., and Keller, N.P. (2010). Secondary metabolism in fungi: Does chromosomal location matter? Current Opinion in Microbiology **13,** 431-436.

Pan, R., Gao, X.H., Li, Y., Xia, Y.F., and Dai, Y. (2010). Anti-arthritic effect of scopoletin, a coumarin compound occurring in Erycibe obtusifolia Benth stems, is associated with decreased angiogenesis in synovium. Fundamental and Clinical Pharmacology **24,** 477-490.

Paquette, S.M., Bak, S., and 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,** 307-317.

Park, N.I., Park, J.H., and Park, S.U. (2011). Overexpression of Cinnamate 4-Hydroxylase Gene Enhances Biosynthesis of Decursinol Angelate in Angelica gigas Hairy Roots. Molecular Biotechnology**,** 1-7.

Parker, H.G., VonHoldt, B.M., Quignon, P., Margulies, E.H., Shao, S., Mosher, D.S., Spady, T.C., Elkahloun, A., Cargill, M., Jones, P.G., Maslen, C.L., Acland, G.M., Sutter, N.B., Kuroki, K., Bustamante, C.D., Wayne, R.K., and Ostrander, E.A. (2009). An expressed fgf4 retrogene is associated with breed-defining chondrodysplasia in domestic dogs. Science **325,** 995-998.

Parrish, J.A., Fitzpatrick, T.B., Tanenbaum, L., and Pathak, M.A. (1974). Photochemotherapy of psoriasis with oral methoxsalen and longwave ultraviolet light. New England Journal of Medicine **291,** 1207-1211.

Passoja, K., Rautavuoma, K., Ala-Kokko, L., Kosonen, T., and Kivirikko, K.I. (1998). Cloning and characterization of a third human lysyl hydroxylase isoform. Proceedings of the National Academy of Sciences of the United States of America **95,** 10482-10486.

Paszkiewicz, M., Orlita, A., Dziabas, A., Gotebiowski, M., Lojkowska, E., Szafranek, J., Malinski, E., and Stepnowski, P. (2008). Simplex optimized LC analysis of plant coumarins and furanocoumarins. Chromatographia **67,** 653-657.

Pathak, M.A., and Fitzpatrick, T.B. (1992). The evolution of photochemotherapy with psoralens and UVA (PUVA): 2000 BC to 1992 AD. Journal of Photochemistry and Photobiology B: Biology **14,** 3-22.

Pathak, M.A., Daniels, F., Jr, and Fitzpatrick, T.B. (1962). The presently known distribution of furocoumarins (psoralens) in plants. The Journal of investigative dermatology **39,** 225-239.

Pereira, L.E., Villinger, F., Wulff, H., Sankaranarayanan, A., Raman, G., and Ansari, A.A. (2007). Pharmacokinetics, toxicity, and functional studies of the selective Kv1.3 channel blocker 5-(4-phenoxybutoxy)psoralen in rhesus macaques. Experimental Biology and Medicine **232,** 1338-1354.

Peroutka, R., Schulzova, V., Botek, P., and Hajslova, J. (2007). Analysis of furanocoumarins in vegetables (Apiaceae) and citrus fruits (Rutaceae). Journal of the Science of Food and Agriculture **87,** 2152-2163.

Petersen, R.A., Zangerl, A.R., Berenbaum, M.R., and Schuler, M.A. (2001). Expression of CYP6B1 and CYP6B3 cytochrome P450 monooxygenases and furanocoumarin

metabolism in different tissues of Papilio polyxenes (Lepidoptera: Papilionidae). Insect Biochemistry and Molecular Biology **31,** 679-690.

Piri, E., Babaeian, M., Tavassoli, A., and Esmaeilian, Y. (2011). Effects of UV irradiation on plants. African Journal of Microbiology Research **5,** 1710-1716.

Plumas, J., Drillat, P., Jacob, M.C., Richard, M.J., and Favrot, M.C. (2003). Extracorporeal photochemotherapy for treatment of clonal T cell proliferations. La photochimiothÃ©rapie extracorporelle dans le traitement de pathologies du systÃ¨me lymphocytaire T **90,** 763-770.

Pokrovskii, O.I., Markoliya, A.A., Lepeshkin, F.D., Kuvykin, I.V., Parenago, O.O., and Gonchukov, S.A. (2009). Extraction of linear furocoumarins from Ammi Majus seeds by means of supercritical fluid extraction and supercritical fluid chromatography. Russian Journal of Physical Chemistry B **3,** 1165-1171.

Polat, M., Oztas, P., Dikilitas, M.C., and Alli, N. (2008). Phytophotodermatitis due to Ficus carica. Dermatology Online Journal **14**.

Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996a). Yeast expression of animal and plant P450s in optimized redox environments. In Methods in Enzymology, pp. 51-64.

Pompon, D., Louerat, B., Bronnie, A., and Urban, P. (1996b). Yeast expression of Animal and Plant P450s in optimized redox environments. Methods in Enzymology **272,** 51-64.

Pramod, K.K., Singh, S., and Jayabaskaran, C. (2010). Biochemical and structural characterization of recombinant hyoscyamine 6β-hydroxylase from Datura metel L. Plant Physiology and Biochemistry **48,** 966-970.

Proctor, R.H., McCormick, S.P., Alexander, N.J., and Desjardins, A.E. (2009). Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus Fusarium. Molecular Microbiology **74,** 1128-1142.

Purpero, V., and Moran, G.R. (2007). The diverse and pervasive chemistries of the α-keto acid dependent enzymes. Journal of Biological Inorganic Chemistry **12,** 587-601.

Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., and Osbourn, A. (2004). A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants. Proceedings of the National Academy of Sciences of the United States of America **101,** 8233-8238.

Que Jr, L., Kolanczyk, R.C., and White, L.S. (1987). Functional models for catechol 1,2-dioxygenase. Structure, reactivity, and mechanism. Journal of the American Chemical Society **109,** 5373-5380.

Rani, A., Singh, K., Sood, P., Kumar, S., and Ahuja, P.S. (2009). p-Coumarate: CoA ligase as a key gene in the yield of catechins in tea [Camellia sinensis (L.) O. Kuntze]. Functional and Integrative Genomics **9,** 271-275.

Rasmussen, U., Christensen, S.B., and Sandberg, F. (1981). Phytochemistry of the genus Thapsia. Planta Medica **43,** 336-341.

Razavi, S.M. (2011). Plant coumarins as allelopathic agents. International Journal of Biological Chemistry **5,** 86-90.

Riveiro, M.E., De Kimpe, N., Moglioni, A., Vazquez, R., Monczor, F., Shayo, C., and Davio, C. (2010). Coumarins: Old compounds with novel promising therapeutic perspectives. Current Medicinal Chemistry **17,** 1325-1338.

Ro, D.K., Mah, N., Ellis, B.E., and Douglas, C.J. (2001). Functional characterization and subcellular localization of poplar (Populus trichocarpa x Populus deltoides) cinnamate 4-hydroxylase. Plant Physiol **126,** 317-329.

Roach, P.L., Clifton, I.J., Fulop, V., Harlos, K., Barton, G.J., Hajdu, J., Andersson, I., Schofield, C.J., and Baldwin, J.E. (1995). Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature **375,** 700-704.

Robineau, T., Batard, Y., Nedelkina, S., Cabello-Hurtado, F., LeRet, M., Sorokine, O., Didierjean, L., and Werck-Reichhart, D. (1998). The chemically inducible plant cytochrome P450 CYP76B1 actively metabolizes phenylureas and other xenobiotics. Plant Physiology **118,** 1049-1056.

Rodin, S.N., and Riggs, A.D. (2003). Epigenetic silencing may aid evolution by gene duplication. Journal of Molecular Evolution **56,** 718-729.

Roth, C., Rastogi, S., Arvestad, L., Dittmar, K., Light, S., Ekman, D., and Liberles, D.A. (2007). Evolution after gene duplication: Models, mechanisms, sequences, systems, and organisms. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution **308,** 58-73.

Roy, S.W., and Gilbert, W. (2005). The pattern of intron loss. Proceedings of the National Academy of Sciences of the United States of America **102,** 713-718.

Ruetschi, U., Nordin, I., Odelhog, B., Jornvall, H., and Lindstedt, S. (1993). β--Butyrobetaine hydroxylase structural characterization of the Pseudomonas enzyme. European Journal of Biochemistry **213,** 1075-1080.

Sagara, Y., Barnes, H.J., and Waterman, M.R. (1993). Expression in Escherichia coli of functional cytochrome P450(c17) lacking its hydrophobic amino-terminal signal anchor. Archives of Biochemistry and Biophysics **304,** 272-278.

Sah, P., Agarwal, D., and Garg, S.P. (2006). Isolation and identification of furocoumarins from the seeds of Psoralea corylifolia linn. Indian Journal of Pharmaceutical Sciences **68,** 768-771.

Sandermann Jr, H. (1994). Higher plant metabolism of xenobiotics: The 'green liver' concept. Pharmacogenetics **4,** 225-241.

Santana, L., Uriarte, E., Roleira, F., Milhazes, N., and Borges, F. (2004). Furocoumarins in medicinal chemistry. Synthesis, natural occurrence and biological activity. Current Medicinal Chemistry **11,** 3239-3261.

Sardari, S., Mori, Y., Horita, K., Micetich, R.G., Nishibe, S., and Daneshtalab, M. (1999). Synthesis and antifungal activity of coumarins and angular furanocoumarins. Bioorganic and Medicinal Chemistry **7,** 1933-1940.

Sasaki, M., Meguro, F., Kumazawa, E., Fujita, H., Kakishima, H., and Sakata, T. (1988). Evidence for uptake of 8-methoxypsoralen and 5-methoxypsoralen by cellular nuclei. Mutation Research **197,** 51-58.

Sayre, R.M., and Dowdy, J.C. (2008). The increase in melanoma: Are dietary furocoumarins responsible? Medical Hypotheses **70,** 855-859.

Schönberg, A., and Sina, A. (1950). Experiments with xanthotoxin and imperatorin obtained from the fruits of ammi majus (L.). Journal of the American Chemical Society **72,** 4826-4828.

Schalk, M., Cabello-Hurtado, F., Pierrel, M.A., Atanossova, R., Saindrenan, P., and Werck-Reichhart, D. (1998). Piperonylic acid, a selective, mechanism-based inactivator of the trans-cinnamate 4-hydroxylase: A new tool to control the flux of metabolites in the phenylpropanoid pathway. Plant Physiology **118,** 209-218.

Schmitt, I.M., Maxwell, V.M., Olack, G.A., Edelson, R.L., Chimenti, S., and Gasparro, F.P. (1994). The specific effects of 8-methoxypsoralen photoadducts on cell growth: HPLC analysis of monoadduct and crosslink formation in cells exposed to split-dose treatment. Journal of Photochemistry and Photobiology B: Biology **22,** 17-21.

Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D. (2001a). CYP98A3 from Arabidopsis thaliana is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem **276,** 36566-36574.

Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D. (2001b). CYP98A3 from Arabidopsis thaliana Is a 3β-Hydroxylase of Phenolic Esters, a Missing Link in the Phenylpropanoid Pathway. Journal of Biological Chemistry **276,** 36566-36574.

Schoch, G.A., Morant, M., Abdulrazzak, N., Asnaghi, C., Goepfert, S., Petersen, M., Ullmann, P., and Werck-Reichhart, D. (2006). The meta-hydroxylation step in the phenylpropanoid pathway: A new level of complexity in the pathway and its regulation. Environmental Chemistry Letters **4,** 127-136.

Schönberg, A. (1968). Preparative Organic Photochemistry.

Schopfer, C.R., and Ebel, J. (1998). Identification of elicitor-induced cytochrome P450s of soybean (Glycine max L.) using differential display of mRNA. Molecular and General Genetics **258,** 315-322.

Schuler, M.A. (1996a). Plant cytochrome P450 monooxygenases. Critical Reviews in Plant Sciences **15,** 235-284.

Schuler, M.A. (1996b). The role of cytochrome P450 monooxygenases in plant-insect interactions. Plant Physiology **112,** 1411-1419.

Schuler, M.A. (2011). P450s in plant-insect interactions. Biochimica et Biophysica Acta - Proteins and Proteomics **1814,** 36-45.

Schuler, M.A., and Berenbaum, M.R. (2003). Diversification of furanocoumarin-metabolizing cytochrome P450 monooxygenases in two papilionids: Specificity and substrate encounter rate. Proc Natl Acad Sci U S A **100,** 14593-14598.

Schuler, M.A., and Werck-Reichhart, D. (2003). Functional Genomics of P450s. In Annual Review of Plant Biology, pp. 629-667.

Schulzova, V., Hajslova, J., Botek, P., and Peroutka, R. (2007). Furanocoumarins in vegetables: Influence of farming system and other factors on levels of toxicants. Journal of the Science of Food and Agriculture **87,** 2763-2767.

Schwab, W. (2003). Metabolome diversity: Too few genes, too many metabolites? Phytochemistry **62,** 837-849.

Scordino, M., Sabatino, L., Belligno, A., and Gagliano, G. (2011). Flavonoids and furocoumarins distribution of unripe chinotto (Citrus* myrtifolia Rafinesque) fruit: beverage processing homogenate and juice characterization. European Food Research and Technology**,** 1-9.

Sedgwick, B., and Lindahl, T. (2002). Recent progress on the Ada response for inducible repair of DNA alkylation damage. Oncogene **21,** 8886-8894.

Serrano-Pérez, J.J., Merchan, M., and Serrano-Andrés, L. (2008). Photoreactivity of furocoumarins and DNA in PUVA therapy: Formation of psoralen-thymine adducts. Journal of Physical Chemistry B **112,** 14002-14010.

Serrano-Pérrez, J.J., Gonzalez-Luque, R., Merchan, M., and Serrano-Andrés, L. (2008). The family of furocoumarins: Looking for the best photosensitizer for phototherapy. Journal of Photochemistry and Photobiology A: Chemistry **199,** 34-41.

Shimada, N., Aoki, T., Sato, S., Nakamura, Y., Tabata, S., and Ayabe, S.I. (2003). A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in Lotus japonicus. Plant Physiology **131,** 941-951.

Shimura, K., Okada, A., Okada, K., Jikumaru, Y., Ko, K.W., Toyomasu, T., Sassa, T., Hasegawa, M., Kodama, O., Shibuya, N., Koga, J., Nojiri, H., and Yamane, H. (2007). Identification of a biosynthetic gene cluster in rice for momilactones. Journal of Biological Chemistry **282,** 34013-34018.

Siatka, T., and Kasparova, M. (2009). Effects of combined hormonal deprivation and fungal elicitation on the production of coumarins in cell suspension cultures of Angelica archangelica L. Ceska a Slovenska Farmacie **58,** 168-171.

Sicker, D., Frey, M., Schulz, M., and Gierl, A. (2000). Role of natural benzoxazinones in the survival strategy of plants. In International Review of Cytology, pp. 319-346.

Silva, W.P.K., Deraniyagala, S.A., Wijesundera, R.L.C., Karunanayake, E.H., and Priyanka, U.M.S. (2002). Isolation of scopoletin from leaves of Hevea brasiliensis and the effect of scopoletin on pathogens of H. brasiliensis. Mycopathologia **153,** 199-202.

Singab, A.N.B. (1998). Acetylated flavonol triglycosides from Ammi majus L. Phytochemistry **49,** 2177-2180.

Sinkel, C., Schwarzer, M.C., Frenking, G., Greiner, A., and Agarwal, S. (2011). Polymer-bound 4-methylcoumarin/1-heptanoyl-5-fluorouracil photodimers: NMR elucidation of dimer structure. Magnetic Resonance in Chemistry **49,** 70-75.

Soltis, P.S. (2005). Ancient and recent polyploidy in angiosperms. New Phytologist **166,** 5-8.

Song, K., Xue, X.P., Wang, W., Huyan, T., Wang, H., Yang, H., and Xie, Q. (2010). The distribution and expression pro-files of Aspartyl/Asparaginyl beta-hydroxylase

(ASPH) in some tumorous cell lines and tissues. Chinese journal of cellular and molecular immunology **26,** 141-144.

Spencer, G.F., Tjarks, L.W., and Powell, R.G. (1987). Analysis of linear and angular furanocoumarins by dual-column high-performance liquid chromatography. Journal of Agricultural and Food Chemistry **35,** 803-805.

Springob, K., Nakajima, J.I., Yamazaki, M., and Saito, K. (2003). Recent advances in the biosynthesis and accumulation of anthocyanins. Natural Product Reports **20,** 288-303.

Stanfill, S.B., Brown, C.R., Yan, X., Watson, C.H., and Ashley, D.L. (2006). Quantification of flavor-related compounds in the unburned contents of bidi and clove cigarettes. Journal of Agricultural and Food Chemistry **54,** 8580-8588.

Steck, W., and Brown, S.A. (1970). Biosynthesis of angular furanocoumarins. Canadian journal of biochemistry **48,** 872-880.

Steck, W., Bailey, B.K., Shyluk, J.P., and Gamborg, O.L. (1971). Coumarins and alkaloids from cell cultures of Ruta graveolens. Phytochemistry **10,** 191-194.

Stein, M.L., and Posocco, E. (1984). Furocoumarins of Pastinaca sativa subsp. sylvestris. Fitoterapia **55,** 119-122.

Stephens, R.M., and Schneider, T.D. (1992). Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. Journal of Molecular Biology **228,** 1124-1136.

Su, V., and Hsu, B.D. (2003). Cloning and expression of a putative cytochrome P450 gene that influences the colour of Phalaenopsis flowers. Biotechnology Letters **25,** 1933-1939.

Su, V., and Hsu, B.-D. (2010). Transient Expression of the Cytochrome p450 CYP78A2 Enhances Anthocyanin Production in Flowers. Plant Molecular Biology Reporter **28,** 302-308.

Sueyoshi, T., Park, L.J., Moore, R., Juvonen, R.O., and Negishi, M. (1995). Molecular engineering of microsomal P450 2a-4 to a stable, water-soluble enzyme. Archives of Biochemistry and Biophysics **322,** 265-271.

Sullivan, M.L., and Zarnowski, R. (2010). Red clover coumarate 3'-hydroxylase (CYP98A44) is capable of hydroxylating p-coumaroyle-shikimate but not p-coumaroyle-malate: implications for the biosynthesis of phaselic acid. Planta **231,** 319-328.

Swiegers, J.H., Vaz, F.M., Pretorius, I.S., Wanders, R.J.A., and Bauer, F.F. (2002). Carnitine biosynthesis in Neurospora crassa: Identification of a cDNA coding for □-N-trimethyllysine hydroxylase and its functional expression in Saccharomyces cerevisiae. FEMS Microbiology Letters **210,** 19-23.

Takahashi, S., Yeo, Y.S., Zhao, Y., O'Maille, P.E., Greenhagen, B.T., Noel, J.P., Coates, R.M., and Chappell, J. (2007). Functional Characterization of Premnaspirodiene Oxygenase, a Cytochrome P450 Catalyzing Regio- And Stereo-specific Hydroxylations of Diverse Sesquiterpene Substrates. Journal of Biological Chemistry **282,** 31744-31754.

Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Current Opinion in Plant Biology **9,** 297-304.

Tholl, D., Sohrabi, R., Huh, J.H., and Lee, S. (2011). The biochemistry of homoterpenes - Common constituents of floral and herbivore-induced plant volatile bouquets. Phytochemistry **72,** 1635-1646.

Tietjen, K.G., and Matern, U. (1983). Differential response of cultured parsley cells to elicitors from two non-pathogenic strains of fungi. II. Effects on enzyme activities. European Journal of Biochemistry **131,** 409-413.

Tietjen, K.G., Hunkler, D., and Matern, U. (1983). Differential response of cultured parsley cells to elicitors from two non-pathogenic strains of fungi. I. Identification of induced products as coumarin derivatives. European Journal of Biochemistry **131,** 401-407.

Tijet, N., Helvig, C., and Feyereisen, R. (2001). The cytochrome P450 gene superfamily in Drosophila melanogaster: Annotation, intron-exon organization and phylogeny. Gene **262,** 189-198.

Tognolli, M., Penel, C., Greppin, H., and Simon, P. (2002). Analysis and expression of the class III peroxidase large gene family in Arabidopsis thaliana. Gene **288,** 129-138.

Tranchimand, S., Ertel, G., Gaydou, V., Gaudin, C., Tron, T., and Iacazio, G. (2008). Biochemical and molecular characterization of a quercetinase from Penicillium olsonii. Biochimie **90,** 781-789.

Tremblay, R., Wang, D., Jevnikar, A.M., and Ma, S. Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. Biotechnology Advances **28,** 214-221.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature **419,** 174-178.

Trieu, A.T., Burleigh, S.H., Kardailsky, I.V., Maldonado-Mendoza, I.E., Versaw, W.K., Blaylock, L.A., Shin, H., Chiou, T.J., Katagi, H., Dewbre, G.R., Weigel, D., and Harrison, M.J. (2000). Transformation of Medicago truncatula via infiltration of seedlings or flowering plants with Agrobacterium. Plant Journal **22,** 531-541.

Umezawa, T. (2010). The cinnamate/monolignol pathway. Phytochemistry Reviews **9,** 1-17.

Urban, P., Werck-Reichhart, D., Teutsch, H.G., Durst, F., Regnier, S., Kazmaier, M., and Pompon, D. (1994). Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway. European Journal Of Biochemistry / FEBS **222,** 843-850.

Valegard, K., Terwisscha Van Scheltinga, A.C., Dubus, A., Ranghino, G., Ã–ster, L.M., Hajdu, J., and Andersson, I. (2004). The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. Nature Structural and Molecular Biology **11,** 95-101.

Valle, T., Lopez, J.L., Hernandez, J.M., and Corchete, P. (1997). Antifungal activity of scopoletin and its differential accumulation in Ulmus pumila and Ulmus campestris cell suspension cultures infected with Ophiostoma ulmi spores. Plant Science **125,** 97-101.

Van Henegouwen, G.M.J.B., Wijn, E.T., Schoonderwoerd, S.A., and Dall'Acqua, F. (1989). A method for the determination of PUVA-induced in vivo irreversible binding of 8-methoxypsoralen (8-MOP) to epidermal lipids, proteins and DNA/RNA. Journal of Photochemistry and Photobiology, B: Biology **3,** 631-635.

Van Iperen, H.P., and Beijersbergen Van Henegouwen, G.M.J. (1997). Clinical and mechanistic aspects of photopheresis. Journal of Photochemistry and Photobiology B: Biology **39,** 99-109.

Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., and Panopoulos, N. (2007). Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. Biotechnology Journal **2,** 1214-1234.

Vialart, G., Hehn, A., Olry, A., Ito, K., Krieger, C., Larbat, R., Paris, C., Shimizu, B., Sugimoto, Y., Mizutani, M., and Bourgaud, F. (2012). A 2-oxoglutarate-dependent dioxygenase from Ruta graveolens L. exhibits p-coumaroyle coA 2'-hydroxylase activity (C2'H): a missing step in the synthesis of umbelliferone in plants. The plant Journal **In press**.

Viola, G., Vedaldi, D., Dall'Acqua, F., Lampronti, I., Bianchi, N., Zuccato, C., Borgatti, M., and Gambari, R. (2008). Furocoumarins photolysis products induce differentiation of human erythroid cells. Journal of Photochemistry and Photobiology B: Biology **92,** 24-28.

Vlad, F., Tiainen, P., Owen, C., Spano, T., Daher, F.B., Oualid, F., Senol, N.O., Vlad, D., Myllyharju, J., and Kalaitzis, P. (2010). Characterization of two carnation petal prolyl 4 hydroxylases. Physiologia Plantarum **140,** 199-207.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant Journal **33,** 949-956.

Wanders, R.J.A., Komen, J., and Ferdinandusse, S. (2011). Phytanic acid metabolism in health and disease. Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids **1811,** 498-507.

Wang, Q., Hillwig, M.L., and Peters, R.J. (2011). CYP99A3: Functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. Plant Journal **65,** 87-95.

Wang, W., Li, S., Zhao, X., Lin, B., and Du, Y. (2007). [Determination of six secondary metabolites including chlorogenic acid in tobacco using high performance liquid chromatography with coulometric array detection]. Se pu = Chinese journal of chromatography / Zhongguo hua xue hui **25,** 848-852.

Wang, W., Zheng, H., Fan, C., Li, J., Shi, J., Cai, Z., Zhang, G., Liu, D., Zhang, J., Vang, S., Lu, Z., Wong, G.K.S., Long, M., and Wang, J. (2006). High rate of chimeric gene origination by retroposition in plant genomes. Plant Cell **18,** 1791-1802.

Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. Nature **449,** 54-61.

Watkins, P.A., Howard, A.E., and Mihalik, S.J. (1994). Phytanic acid must be activated to phytanoyl-CoA prior to its α-oxidation in rat liver peroxisomes. Biochimica et Biophysica Acta - Lipids and Lipid Metabolism **1214,** 288-294.

Weaver, L.M., and Herrmann, K.M. (1997). Dynamics of the shikimate pathway in plants. Trends in Plant Science **2,** 346-351.

Weisshaar, B., and Jenkinst, G.I. (1998). Phenylpropanoid biosynthesis and its regulation. Current Opinion in Plant Biology **1,** 251-257.

Wellesen, K., Durst, F., Pinot, F., Benveniste, I., Nettesheim, K., Wisman, E., Steiner-Lange, S., Saedler, H., and Yephremov, A. (2001). Functional analysis of the LACERATA gene of Arabidopsis provides evidence for different roles of fatty acid ω-hydroxylation in development. Proceedings of the National Academy of Sciences of the United States of America **98,** 9694-9699.

Wen, Z., Berenbaum, M.R., and Schuler, M.A. (2006a). Inhibition of CYP6B1-mediated detoxification of xanthotoxin by plant allelochemicals in the black swallowtail (Papilio polyxenes). Journal of Chemical Ecology **32,** 507-522.

Wen, Z., Pan, L., Berenbaum, M.R., and Schuler, M.A. (2003). Metabolism of linear and angular furanocoumarins by Papilio polyxenes CYP6B1 co-expressed with NADPH cytochrome P450 reductase. Insect Biochemistry and Molecular Biology **33,** 937-947.

Wen, Z., Rupasinghe, S., Niu, G., Berenbaum, M.R., and Schuler, M.A. (2006b). CYP6B1 and CYP6B3 of the black swallowtail (Papilio polyxenes): Adaptive evolution through subfunctionalization. Molecular Biology and Evolution **23,** 2434-2443.

Wen, Z., Zeng, R.S., Niu, G., Berenbaum, M.R., and Schuler, M.A. (2009). Ecological significance of induction of broad-substrate cytochrome p450s by natural and synthetic inducers in Helicoverpa Zea. Journal of Chemical Ecology **35,** 183-189.

Werck-Reichhart, D. (1995a). Cytochromes P450 in phenylpropanoid metabolism. Drug Metabolism and Drug Interactions **12,** 221-243.

Werck-Reichhart, D. (1995b). Herbicide metabolism and selectivity: role of cytochrome P450. In Proceedings British Crop Protection Conference-Weeds, pp. 813-822.

Werck-Reichhart, D., and Feyereisen, R. (2000). Cytochromes P450: a success story. Genome Biology **1**.

Wickramasinghe, R.H., and Villee, C.A. (1975). Early role during chemical evolution for cytochrome P450 in oxygen detoxification. Nature **256,** 509-511.

Williams, P.A., Cosme, J., Sridhar, V., Johnson, E.F., and McRee, D.E. (2000). Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. Molecular Cell **5,** 121-131.

Wilmouth, R.C., Turnbull, J.J., Welford, R.W.D., Clifton, I.J., Prescott, A.G., and Schofield, C.J. (2002). Structure and mechanism of anthocyanidin synthase from Arabidopsis thaliana. Structure **10,** 93-103.

Wittstock, U., and Halkier, B.A. (2000). Cytochrome P450 CYP79A2 from Arabidopsis thaliana L. catalyzes the conversion of L-phenylalanine to phenylacetaldoxime in the biosynthesis of benzylglucosinolate. Journal of Biological Chemistry **275,** 14659-14666.

Wu, H., De Graaf, B., Mariani, C., and Cheung, A.Y. (2001). Hydroxyproline-rich glycoproteins in plant reproductive tissues: Structure, functions and regulation. Cellular and Molecular Life Sciences **58,** 1418-1429.

Wu, S., and Chappell, J. (2008). Metabolic engineering of natural products in plants; tools of the trade and challenges for the future. Current Opinion in Biotechnology **19,** 145-152.

Wu, S.B., Pang, F., Wen, Y., Zhang, H.F., Zhao, Z., and Hu, J.F. (2010). Antiproliferative and apoptotic activities of linear furocoumarins from notopterygium incisum on cancer cell lines. Planta Medica **76,** 82-85.

Xu, H., Park, N.I., Li, X., Kim, Y.K., Lee, S.Y., and Park, S.U. (2010). Molecular cloning and characterization of phenylalanine ammonia-lyase, cinnamate 4-hydroxylase and genes involved in flavone biosynthesis in Scutellaria baicalensis. Bioresource Technology **101,** 9715-9722.

Xu, R.Y., Nan, P., Yang, Y., Pan, H., Zhou, T., and Chen, J. (2011). Ultraviolet irradiation induces accumulation of isoflavonoids and transcription of genes of enzymes involved in the calycosin-7-O-β-d-glucoside pathway in Astragalus membranaceus Bge. var. mongholicus (Bge.) Hsiao. Physiologia Plantarum **142,** 265-273.

Yang, H., Song, K., Xue, T., Xue, X.P., Huyan, T., Wang, W., and Wang, H. (2010). The distribution and expression profiles of human aspartyl/asparaginyl beta-hydroxylase in tumor cell lines and human tissues. Oncology Reports **24,** 1257-1264.

Yin, X., O'Hare, T., Gould, S.J., and Zabriskie, T.M. (2003). Identification and cloning of genes encoding viomycin biosynthesis from Streptomyces vinaceus and evidence for involvement of a rare oxygenase. Gene **312,** 215-224.

Yuasa, H.J., Ushigoe, A., and Ball, H.J. (2011). Molecular evolution of bacterial indoleamine 2,3-dioxygenase. Gene **485,** 22-31.

Zangar, R.C., Davydov, D.R., and Verma, S. (2004). Mechanisms that regulate production of reactive oxygen species by cytochrome P450. Toxicology and Applied Pharmacology **199,** 316-331.

Zangerl, A.R., and Berenbaum, M.R. (1987). Furanocoumarins in wild parsnip: effects of photosynthetically active radiation, ultraviolet light, and nutrients. Ecology **68,** 516-520.

Zarebska, Z., Waszkowska, E., Caffieri, S., and Dall'Acqua, F. (2000). PUVA (psoralen + UVA) photochemotherapy: Processes triggered in the cells. Farmaco **55,** 515-520.

Zeng, R.S., Wen, Z., Niu, G., Schuler, M.A., and Berenbaum, M.R. (2009). Enhanced toxicity and induction of cytochrome P450s suggest a cost of "eavesdropping" in a multitrophic interaction. Journal of Chemical Ecology **35,** 526-532.

Zeringue Jr, H.J. (1984). The accumulation of five fluorescent compounds in the cotton leaf induced by cell-free extracts of Aspergillus flavus. Phytochemistry **23,** 2501-2503.

Zhang, C.Z., Wang, S.X., Zhang, Y., Chen, J.P., and Liang, X.M. (2005a). In vitro estrogenic activities of Chinese medicinal plants traditionally used for the management of menopausal symptoms. Journal of Ethnopharmacology **98,** 295-300.

Zhang, J. (2003). Evolution by gene duplication: An update. Trends in Ecology and Evolution **18,** 292-298.

Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W., and Chua, N.H. (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nature Protocols **1,** 641-646.

Zhang, Y., Wu, Y., Liu, Y., and Han, B. (2005b). Computational identification of 69 retroposons in Arabidopsis. Plant Physiology **138,** 935-948.

Zhang, Z., Ren, J., Stammers, D.K., Baldwin, J.E., Marios, K., and Schofield, C.J. (2000). Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. Nature Structural Biology **7,** 127-133.

Zhou, Q., Wintersteen, C.L., and Cadwallader, K.R. (2002). Identification and quantification of aroma-active components that contribute to the distinct malty flavor of buckwheat honey. Journal of Agricultural and Food Chemistry **50,** 2016-2021.

Zhou, Y.T., Xiao, H.B., and Bi, M.G. (2011). ROS and endoplasmic reticulum stress. Chinese Pharmacological Bulletin **27,** 597-600.

Zhu, W. (2008). CYP2D6: A key enzyme in morphine synthesis in animals. Medical Science Monitor **14,** SC15-SC18.

Zhu, W.Y. (2010). Three cases of phytophotodermatitis resulting from Chinese herbal medicine decoction. Journal of Clinical Dermatology **39,** 495-497.

Ziegler, J., and Facchini, P.J. (2008). Alkaloid biosynthesis: Metabolism and trafficking. In Annual Review of Plant Biology, pp. 735-769.

Zimmer, E.A., Martin, S.L., and Beverley, S.M. (1980). Rapid duplication and loss of genes coding for the Î± chains of hemoglobin. Proceedings of the National Academy of Sciences of the United States of America **77,** 2158-2162.

Zobel, A.M. (1997). Coumarins in fruits and vegetables. In In Phytochemistry of Fruits and Vegetables, F.A.A.a.R. Tomas-Barberan, R.J., eds, ed (Oxford:: Clarendon Press), pp. pp. 173–203.

Zobel, A.M., and Brown, S.A. (1990a). Seasonal changes of furanocoumarin concentrations in leaves of Heracleum lanatum. Journal of Chemical Ecology **16,** 1623-1634.

Zobel, A.M., and Brown, S.A. (1990b). Dermatitis-inducing furanocoumarins on leaf surfaces of eight species of Rutaceous and Umbelliferous plants. Journal of Chemical Ecology **16,** 693-700.

Zobel, A.M., and Brown, S.A. (1993). Influence of low-intensity ultraviolet radiation on extrusion of furanocoumarins to the leaf surface. Journal of Chemical Ecology **19,** 939-952.

Zobel, A.M., Brown, S.A., and Glowniak, K. (1990). Localization of furanocoumarins in leaves, fruits, and seeds of plants causing contact photodermatitis. Planta Medica **56,** 571-572.

Zobel, A.M., Wang, J., March, R.E., and Brown, S.A. (1991). Identification of eight coumarins occurring with psoralen, xanthotoxin, and bergapten on leaf surfaces. Journal of Chemical Ecology **17,** 1859-1870.

Zulak, K.G., and Bohlmann, J. (2010). Terpenoid biosynthesis and specialized vascular cells of conifer defense. Journal of Integrative Plant Biology **52,** 86-97.

Zumwalt, J.G., and Neal, J.J. (1993). Cytochromes P450 from Papilio polyxenes: Adaptations to host plant allelochemicals. Comparative Biochemistry and Physiology - C Pharmacology Toxicology and Endocrinology **106,** 111-118.

Chapitre VI : Annexes

Annexe 1 : Extraction de l'ADN de plantes suivant le protocole de « DNeasy plant minik de Qiagen »

1. **Harvest leaves and place up to 50 mg into each tube in 2 collection microtube racks.**

   Unless a different optimal amount of starting material has been previously determined, do not use more than 50 mg (wet weight) per sample (see "Sample size", page 16).

   Most leaf material can be stored at 4°C for at least 24 h prior to processing without affecting DNA yield or quality.

   Keep the clear covers from the collection microtube racks for use in step 10.

   Use the plate register cards provided to record the position of each sample in the racks.

2. **Add one tungsten carbide bead to each collection microtube.**

3. **Combine Buffer AP1, RNase A, and Reagent DX, according to the table below, to make a working lysis solution. Pipet 400 µl of the working lysis solution into each collection microtube. Seal the microtubes with the caps provided.**

   It is important to prepare a fresh working lysis solution. To allow thorough mixing of the solution, combine the components in a tube and vortex to mix, then dispense the solution into a reagent reservoir for use with a multichannel pipet.

	Volume per sample	Volume for 2 x 96 samples*
Buffer AP1 (preheated to 65°C)	400 µl	90 ml
RNase A (100 mg/ml)	1 µl	225 µl
Reagent DX†	1 µl	225 µl

* 15% excess mixture is included in these calculations to allow for pipetting errors.

† Reagent DX is viscous.

4. **Sandwich each rack of collection microtubes between adapter plates and fix into TissueLyser clamps as described in the *TissueLyser User Manual*.**

   Note: Ensure that the microtubes are properly sealed with caps.

   IMPORTANT: Two plate sandwiches must be clamped to the TissueLyser to provide balance. To process 96 samples or less, assemble a second plate sandwich using a rack of collection microtubes containing tungsten carbide beads, but no samples or buffers, and fix it into the empty clamp.

5. **Grind the samples for 1.5 min at 30 Hz.**

   IMPORTANT: Prolonging the disruption time may result in DNA shearing.

6. **Remove and disassemble the plate sandwiches. Ensure that the collection microtubes are tightly closed. Reassemble the plate sandwiches so that the collection microtubes nearest the TissueLyser in steps 4 and 5 are now outermost. Reinsert the plate sandwiches into the TissueLyser.**

   Rotating the racks of collection microtubes in this way ensures that all samples are thoroughly disrupted (See "Disruption and homogenization using the TissueLyser System", page 17). More foam will form in the tubes that were outermost during the initial disruption step.

   IMPORTANT: Merely rotating the entire plate sandwich so that the QIAGEN logos are upside down when reinserted into the mixer mill is not sufficient, since the same samples that were outermost during the initial disruption will remain outermost in the second disruption step.

7. **Grind the samples for another 1.5 min at 30 Hz.**

   IMPORTANT: Prolonging the disruption time may result in shearing of DNA.

8. **Remove the plate sandwiches from the TissueLyser and remove the adapter plates from each rack of collection microtubes. To collect any solution from the caps, centrifuge the collection microtubes. Allow the centrifuge to reach 3000 rpm, and then stop the centrifuge.**

   Do not prolong this step.

9. **Remove and discard caps. Add 130 µl Buffer AP2 to each collection microtube.**

10. **Close the microtubes carefully with new caps (provided); ensure that the microtubes are properly sealed to avoid leakage during shaking. Place a clear cover (saved from step 1) over each rack of collection microtubes and shake the racks vigorously up and down for 15 s. To collect any solution from the caps, centrifuge the collection microtubes. Allow the centrifuge to reach 3000 rpm, and then stop the centrifuge.**

    Do not prolong this step.

    Note: To ensure optimal DNA yields, it is important to shake the racks of collection microtubes vigorously up and down with both hands for the full 15 s. The genomic DNA will not be sheared by vigorous shaking.

    The centrifugation step prevents precipitates from freezing to the caps, which would otherwise be difficult to remove after incubation at –20°C (step 11).

    Keep the clear covers from the collection microtube racks for use in step 15.

11. **Incubate the racks of collection microtubes for 10 min at –20°C.**

    This incubation aids the precipitation of proteins and inhibitors of downstream applications following addition of Buffer AP2.

12. **Centrifuge the racks of collection microtubes for 5 min at 6000 rpm.**

    Compact pellets will form, but some particles may float. Be careful not to transfer any of these particles in the following step.

13. **Remove and discard the caps. Carefully transfer 400 µl of each supernatant to new racks of collection microtubes (provided), ensuring that the new tubes are in the correct orientation.**

    Do not discard the pellets as they contain the tungsten carbide beads, which can be recovered and reused (see Appendix B, page 50).

    Do not transfer more than 400 µl of the supernatant as otherwise the capacity of the DNeasy 96 plates and the S-Blocks used in subsequent steps will be exceeded.

    If less than 400 µl supernatant is recovered, adjust the amount of Buffer AP3/E in step 14 accordingly.

    Collection microtubes are connected in strips of 8. To avoid transferring particulate matter, it is helpful to remove the strips from the rack so that the contents of the microtubes are visible, and to use a multichannel pipet on its lowest speed setting.

    Save the used collection microtubes to recover the tungsten carbide beads at a later stage (see Appendix B, page 50).

14. **Add 1.5 volumes (typically 600 µl) of Buffer AP3/E to each sample.**

    Note: Ensure that ethanol has been added to Buffer AP3/E prior to use. See "Things to do before starting", page 32.

    A white precipitate may form upon addition of Buffer AP3/E. This precipitate does not interfere with the DNeasy 96 Plant procedure or any subsequent application.

15. **Close the collection microtubes with new caps (provided); ensure that the tubes are properly sealed to prevent leakage during shaking. Place a clear cover over each rack of collection microtubes and shake the racks vigorously up and down for 15 s. To collect any solution from the caps, centrifuge the collection microtubes. Allow the centrifuge to reach 3000 rpm, and then stop the centrifuge.**

    Do not prolong this step.

    Note: To ensure optimal DNA yields, it is important to shake the racks of collection microtubes vigorously up and down with both hands for the full 15 s. The genomic DNA will not be sheared by vigorous shaking.

16. **Place two DNeasy 96 plates on top of S-Blocks (provided). Mark the DNeasy 96 plates for later sample identification.**

17. **Remove and discard the caps from the collection microtubes. Carefully transfer 1 ml of each sample to each well of the DNeasy 96 plates.**

    Take care not to wet the rims of the wells to avoid aerosols during centrifugation. Do not transfer more than 1 ml per well.

    Note: Lowering pipet tips to the bottoms of the wells may cause sample overflow and cross-contamination. Therefore, remove one set of caps at a time, and begin drawing up the samples as soon as the pipet tips contact the liquid. Repeat until all the samples have been transferred to the DNeasy 96 plates.

18. **Seal each DNeasy 96 plate with an AirPore Tape Sheet (provided). Centrifuge for 4 min at 6000 rpm.**

    AirPore Tape prevents cross-contamination between samples during centrifugation.

    After centrifugation, check that all of the lysate has passed through the membrane in each well of the DNeasy 96 plates. If lysate remains in any of the wells, centrifuge for a further 4 min.

19. **Remove the tape. Carefully add 800 µl Buffer AW to each sample.**

    Note: Ensure that ethanol has been added to Buffer AW prior to use. See "Things to do before starting", page 32.

20. **Centrifuge for 15 min at 6000 rpm to dry the DNeasy membranes.**

    For efficient drying, do not reseal the DNeasy 96 plate with AirPore Tape.

    IMPORTANT: Residual ethanol in the DNeasy membranes derived from Buffer AW may inhibit PCR and must be removed by centrifugation before elution of the DNA.

    Note: DNeasy membranes are sometimes slightly colored after this wash step. This should not affect the DNeasy 96 Plant procedure. A very dark membrane could indicate that too much starting material was used. A second wash step with 800 µl ethanol (96–100%) may improve DNA quality in these cases. Empty the flow-through from the S-Block before performing this second wash step.

21. **To elute the DNA, place each DNeasy 96 plate in the correct orientation on a new rack of Elution Microtubes RS (provided), add 100 µl Buffer AE to each sample, and seal the DNeasy 96 plates with new AirPore Tape Sheets (provided). Incubate for 1 min at room temperature (15–25°C). Centrifuge for 2 min at 6000 rpm.**

    Elution in 2 x 50 µl (instead of 2 x 100 µl) increases DNA concentration, but decreases the overall DNA yield (see "Elution", page 21).

22. **Repeat step 21 with another 100 µl Buffer AE.**

    Use new caps (provided) to seal the Elution Microtubes RS for storage.

Annexe 2 : Extraction d'ARN de plantes suivant le protocole « RNeasy plant minikit de Qiagen ».

1. **Determine the amount of plant material. Do not use more than 100 mg.**

   Weighing tissue is the most accurate way to determine the amount.

2. **Immediately place the weighed tissue in liquid nitrogen, and grind thoroughly with a mortar and pestle. Decant tissue powder and liquid nitrogen into an RNase-free, liquid-nitrogen–cooled, 2 ml microcentrifuge tube (not supplied). Allow the liquid nitrogen to evaporate, but do not allow the tissue to thaw. Proceed immediately to step 3.**

   RNA in plant tissues is not protected until the tissues are flash-frozen in liquid nitrogen. Frozen tissues should not be allowed to thaw during handling. The relevant procedures should be carried out as quickly as possible.

3. **Add 450 µl Buffer RLT or Buffer RLC (see "Important points before starting") to a maximum of 100 mg tissue powder. Vortex vigorously.**

   A short 1–3 min incubation at 56°C may help to disrupt the tissue. However, do not incubate samples with a high starch content at elevated temperatures, otherwise swelling of the sample will occur.

4. **Transfer the lysate to a QIAshredder spin column (lilac) placed in a 2 ml collection tube, and centrifuge for 2 min at full speed. Carefully transfer the supernatant of the flow-through to a new microcentrifuge tube (not supplied) without disturbing the cell-debris pellet in the collection tube. Use only this supernatant in subsequent steps.**

   It may be necessary to cut off the end of the pipet tip to facilitate pipetting of the lysate into the QIAshredder spin column. Centrifugation through the QIAshredder spin column removes cell debris and simultaneously homogenizes the lysate. While most of the cell debris is retained on the QIAshredder spin column, a very small amount of cell debris will pass through and form a pellet in the collection tube. Be careful not to disturb this pellet when transferring the lysate to the new microcentrifuge tube.

5. **Add 0.5 volume of ethanol (96–100%) to the cleared lysate, and mix immediately by pipetting. Do not centrifuge. Proceed immediately to step 6.**

   Note: The volume of lysate may be less than 450 µl due to loss during homogenization.

   Note: Precipitates may be visible after addition of ethanol. This does not affect the procedure.

6. **Transfer the sample (usually 650 µl), including any precipitate that may have formed, to an RNeasy spin column (pink) placed in a 2 ml collection tube (supplied). Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm). Discard the flow-through.***

   Reuse the collection tube in step 7.

   If the sample volume exceeds 700 µl, centrifuge successive aliquots in the same RNeasy spin column. Discard the flow-through after each centrifugation.*

   Optional: If performing optional on-column DNase digestion (see "Eliminating genomic DNA contamination", page 23), follow steps D1–D4 (page 69) after performing this step.

7. **Add 700 µl Buffer RW1 to the RNeasy spin column. Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane. Discard the flow-through.***

   Reuse the collection tube in step 8.

8. **Add 500 µl Buffer RPE to the RNeasy spin column. Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane. Discard the flow-through.**

   Reuse the collection tube in step 9.

   Note: Buffer RPE is supplied as a concentrate. Ensure that ethanol is added to Buffer RPE before use (see "Things to do before starting").

9. **Add 500 µl Buffer RPE to the RNeasy spin column. Close the lid gently, and centrifuge for 2 min at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane.**

   The long centrifugation dries the spin column membrane, ensuring that no ethanol is carried over during RNA elution. Residual ethanol may interfere with downstream reactions.

   Note: After centrifugation, carefully remove the RNeasy spin column from the collection tube so that the column does not contact the flow-through. Otherwise, carryover of ethanol will occur.

10. **Optional: Place the RNeasy spin column in a new 2 ml collection tube (supplied), and discard the old collection tube with the flow-through. Close the lid gently, and centrifuge at full speed for 1 min.**

    Perform this step to eliminate any possible carryover of Buffer RPE, or if residual flow-through remains on the outside of the RNeasy spin column after step 9.

11. **Place the RNeasy spin column in a new 1.5 ml collection tube (supplied). Add 30–50 µl RNase-free water directly to the spin column membrane. Close the lid gently, and centrifuge for 1 min at ≥8000 x *g* (≥10,000 rpm) to elute the RNA.**

12. **If the expected RNA yield is >30 µg, repeat step 11 using another 30–50 µl RNase-free water, or using the eluate from step 11 (if high RNA concentration is required). Reuse the collection tube from step 11.**

    If using the eluate from step 11, the RNA yield will be 15–30% less than that obtained using a second volume of RNase-free water, but the final RNA concentration will be higher.

Annexe 3 : Préparation de cDNA par reverse transcription suivant le protocol « high capacity RNA-to-cDNA master mix » d'Applied Biosystems

1. Place the Master Mix (or No-RT Control) components on ice, then mix and briefly centrifuge them.

2. Calculate the total volume of components needed to prepare the required number of reactions. Use the table below.

Component	Volume per Reaction	Final Concentration
Master Mix (or No-RT Control)	4.0 µL	1X
RNA template	up to 16 µL	1 pg - 1 µg
Nuclease-free H_2O	sufficient to 20 µL	—
Total	20.0 µL	—

IMPORTANT! Include additional reactions in the calculations to provide excess volume for the loss that occurs during reagent transfers.

3. Prepare the reaction mix according to your calculations in step 2.

WARNING CHEMICAL HAZARD. High Capacity RNA-to-cDNA Master Mix may cause eye, skin, and respiratory tract irritation. Read the MSDS, and follow the handling instructions. Wear appropriate eyewear, clothing, and gloves.

IMPORTANT! Prepare the reaction mix on ice.

4. Seal the plate or tubes.

5. Briefly centrifuge the plate or tubes to spin down the contents and to eliminate air bubbles.

6. Place the plate or tubes on ice until you are ready to load the thermal cycler.

Annexe 4 : RT-PCR en une seule étape suivant le protocole « SuperScript™ One-Step RT-PCR with Platinum® *Taq* » d'Invitrogène

1. Program the thermal cycler so that cDNA synthesis is followed immediately with PCR amplification automatically.

Note: The following cycling conditions were established using a DNA Thermal Cycler 9600 or 2400 (Perkin Elmer) and may have to be altered for other thermal cyclers. Efficient cDNA synthesis can be achieved in a 15-30 min incubation at 45-55°C. It is recommended that a 30-min incubation at 50°C be used as a general starting point. The optimal temperature for reverse transcription will depend on primer and target sequences. Remember that cycling conditions may have to be further optimized for different sequences. Annealing and extension steps are separate (three-step cycling).

A: cDNA synthesis and pre-denaturation	B: PCR amplification	C: Final extension-(optional)
Perform 1 cycle of: 45-55°C for 15-30 min 94°C for 2 min	Perform 35-40 cycles of: Denature, 94°C for 15 s Anneal, 55-60°C for 30 s Extend, 68-72°C for 1 min/kb	1 cycle of 72°C for 5-10 min For use in Perkin-Elmer Model 480 cycler, use 30 s denaturation instead of 15 s.

2. Add the following to the microcentrifuge tubes placed on ice. Reaction cocktails can be made when multiple reactions are being assembled.

Components	Volume / 50 µl	Final Concentration
2X Reaction Mix	25 µl	1X
Template RNA	x µl	10 pg – 1 µg
Sense Primer (10 µM)	1 µl	0.2 µM
Anti-sense Primer (10 µM)	1 µl	0.2 µM
RT / Platinum® *Taq* Mix	1 µl	—
Autoclaved distilled water	to 50 µl	—

Note: Absence of genomic DNA in RNA preparations can be verified by omitting the RT / Platinum® *Taq* Mix and substituting 2 units of Platinum® *Taq* DNA polymerase in the reaction.

3. Gently mix and make sure that all the components are at the bottom of the amplification tube. Centrifuge briefly if needed. Depending on the thermal cycler used, overlay with silicone oil, if necessary.

4. Analyze the amplification product.

Annexe 5 : Extraction d'ADN à partir d'agarose suivant le protocle QIAquick Gel Extaraction Kit de Qiagen

1. **Excise the DNA fragment from the agarose gel with a clean, sharp scalpel.**

   Minimize the size of the gel slice by removing extra agarose.

2. **Weigh the gel slice in a colorless tube. Add 3 volumes of Buffer QG to 1 volume of gel (100 mg ~ 100 µl).**

   For example, add 300 µl of Buffer QG to each 100 mg of gel. For >2% agarose gels, add 6 volumes of Buffer QG. The maximum amount of gel slice per QIAquick column is 400 mg; for gel slices >400 mg use more than one QIAquick column.

3. **Incubate at 50°C for 10 min (or until the gel slice has completely dissolved). To help dissolve gel, mix by vortexing the tube every 2–3 min during the incubation.**

   IMPORTANT: Solubilize agarose completely. For >2% gels, increase incubation time.

4. **After the gel slice has dissolved completely, check that the color of the mixture is yellow (similar to Buffer QG without dissolved agarose).**

   If the color of the mixture is orange or violet, add 10 µl of 3 M sodium acetate, pH 5.0, and mix. The color of the mixture will turn to yellow.

   The adsorption of DNA to the QIAquick membrane is efficient only at pH ≤7.5. Buffer QG contains a pH indicator which is yellow at pH ≤7.5 and orange or violet at higher pH, allowing easy determination of the optimal pH for DNA binding.

5. **Add 1 gel volume of isopropanol to the sample and mix.**

   For example, if the agarose gel slice is 100 mg, add 100 µl isopropanol. This step increases the yield of DNA fragments <500 bp and >4 kb. For DNA fragments between 500 bp and 4 kb, addition of isopropanol has no effect on yield. Do not centrifuge the sample at this stage.

6. **Place a QIAquick spin column in a provided 2 ml collection tube.**

7. **To bind DNA, apply the sample to the QIAquick column, and centrifuge for 1 min.**

   The maximum volume of the column reservoir is 800 µl. For sample volumes of more than 800 µl, simply load and spin again.

8. **Discard flow-through and place QIAquick column back in the same collection tube.**

   Collection tubes are re-used to reduce plastic waste.

9. **(Optional): Add 0.5 ml of Buffer QG to QIAquick column and centrifuge for 1 min.**

   This step will remove all traces of agarose. It is only required when the DNA will subsequently be used for direct sequencing, in vitro transcription or microinjection.

10. **To wash, add 0.75 ml of Buffer PE to QIAquick column and centrifuge for 1 min.**

    Note: If the DNA will be used for salt sensitive applications, such as blunt-end ligation and direct sequencing, let the column stand 2–5 min after addition of Buffer PE, before centrifuging.

11. **Discard the flow-through and centrifuge the QIAquick column for an additional 1 min at ≥10,000 x *g* (~13,000 rpm).**

    IMPORTANT: Residual ethanol from Buffer PE will not be completely removed unless the flow-through is discarded before this additional centrifugation.

12. **Place QIAquick column into a clean 1.5 ml microcentrifuge tube.**

13. **To elute DNA, add 50 µl of Buffer EB (10 mM Tris-Cl, pH 8.5) or H_2O to the center of the QIAquick membrane and centrifuge the column for 1 min at maximum speed. Alternatively, for increased DNA concentration, add 30 µl elution buffer to the center of the QIAquick membrane, let the column stand for 1 min, and then centrifuge for 1 min.**

    IMPORTANT: Ensure that the elution buffer is dispensed directly onto the QIAquick membrane for complete elution of bound DNA. The average eluate volume is 48 µl from 50 µl elution buffer volume, and 28 µl from 30 µl.

    Elution efficiency is dependent on pH. The maximum elution efficiency is achieved between pH 7.0 and 8.5. When using water, make sure that the pH value is within this range, and store DNA at –20°C as DNA may degrade in the absence of a buffering agent. The purified DNA can also be eluted in TE (10 mM Tris-Cl, 1 mM EDTA, pH 8.0), but the EDTA may inhibit subsequent enzymatic reactions.

Annexe 6 : Extraction plasmidique suivant le protocol du « kit Genelute™ plamside miniprep » de Sigma-Aldrich

1 Completely resuspend the bacterial pellet with 200 µl of the Resuspension Solution. Vortex or pipette up and down to thoroughly resuspend the cells until homogeneous. Incomplete resuspension will result in poor recovery.

Another rapid way to resuspend the cell pellets is to scrape the bottoms of the microcentrifuge tubes back and forth 5 times across the surface of a polypropylene microcentrifuge tube storage rack with 5 × 16 holes.[3]

2 Lyse the resuspended cells by adding 200 µl of the Lysis Solution. Immediately mix the contents by gentle inversion (6–8 times) until the mixture becomes clear and viscous. **Do not vortex**. Harsh mixing will shear genomic DNA, resulting in chromosomal DNA contamination in the final recovered plasmid DNA. **Do not allow the lysis reaction to exceed 5 minutes**. Prolonged alkaline lysis may permanently denature supercoiled plasmid DNA that may render it unsuitable for most downstream applications.

3 Precipitate the cell debris by adding 350 µl of the Neutralization/Binding Solution. Gently invert the tube 4–6 times. Pellet the cell debris by centrifuging at ≥12,000 × *g* or maximum speed for 10 minutes. Cell debris, proteins, lipids, SDS, and chromosomal DNA should fall out of solution as a cloudy, viscous precipitate. If the supernatant contains a large amount of floating particulates after centrifugation, recentrifuge the supernatant before proceeding to step 6.

4 Insert a GenElute Miniprep Binding Column into a provided microcentrifuge tube, if not already assembled. Add 500 µl of the Column Preparation Solution to each miniprep column and centrifuge at 12,000 × *g* for 30 seconds to 1 minute. Discard the flow-through liquid.

Note: The Column Preparation Solution maximizes binding of DNA to the membrane resulting in more consistent yields.

5 Transfer the cleared lysate from step 3 to the column prepared in step 4 and centrifuge at ≥12,000 × *g* for 30 seconds to 1 minute. Discard the flow-through liquid.

6 Add 500 µl of the Optional Wash Solution to the column. Centrifuge at ≥12,000 × *g* for 30 seconds to 1 minute. Discard the flow-through liquid.

Note: When working with bacterial strains containing the wild-type EndA$^+$ gene, such as HB101, JM101, and the NM and PR series, the Optional Wash step is necessary to avoid nuclease contamination of the final plasmid DNA product.

7 Add 750 µl of the diluted Wash Solution to the column. Centrifuge at ≥12,000 × *g* for 30 seconds to 1 minute. The column wash step removes residual salt and other contaminants introduced during the column load. Discard the flow-through liquid and centrifuge again at maximum speed for 1 to 2 minutes without any additional Wash Solution to remove excess ethanol.

8 Transfer the column to a fresh collection tube. Add 100 µl of Elution Solution or molecular biology reagent water to the column. For DNA sequencing and other enzymatic applications, use water or 5 mM Tris-HCl, pH 8.0, as an eluant. Centrifuge at ≥12,000 × *g* for 1 minute. The DNA is now present in the eluate and is ready for immediate use or storage at –20 °C.

Note: If a more concentrated plasmid DNA preparation is required, the elution volume may be reduced to a minimum of 50 µl. However, this may result in a reduction in the total plasmid DNA yield.

Annexe 7 : Synthèse de la sonde d'hybridation suivant le protocole décrit dans le kit «Amersham Gene Images AlkPhos Direct Labelling and Detection System ».

Protocol	Notes
1. Dilute 20 µl cross-linker solution with 80 µl of the water supplied to give the working concentration.	**1.** The working solution can be kept in a refrigerator at 2–8°C for one week.
2. Dilute DNA (or RNA) to be labelled to a concentration of 10ng/µl using the water supplied.	**2.** The concentration of salt in the sample of nucleic acid should be kept as low as possible and must not exceed 50 mM.
3. Place 10 µl of the diluted DNA sample in a microcentrifuge tube and denature by heating for 5 minutes in a vigorously boiling water bath.	**3.** The labelling reaction can be scaled up by increasing the volume of all the components of the labelling reaction pro rata; DNA, reaction buffer, labelling reagent and cross-linker working solution.
4. Immediately cool the DNA on ice for 5 minutes. Spin briefly in a microcentrifuge to collect the contents at the bottom of the tube.	
5. Add 10 µl of reaction buffer to the cooled DNA. Mix thoroughly but gently.	**5.** The reaction should be kept on ice.
6. Add 2 µl labelling reagent. Mix thoroughly but gently.	
7. Add 10 µl of the cross-linker working solution. Mix thoroughly. Spin briefly in a microcentrifuge to collect the contents at the bottom of the tube.	
8. Incubate the reaction for 30 minutes at 37°C.	**8.** For color detection this temperature can be increased (see protocol on page 22)
9. The probe can be used immediately or kept on ice for up to 2 hours. For long term storage, labelled probes may be stored in 50% (v/v) Glycerol.	**9.** Further treatment of the probe before use is **not** required even after long term storage.

Annexe 8 : Purification des protéines suivant le protocole décrit dans le kit « Ni-NTA Purification System » de Qiagen (Annexe 8).

1 **Equilibrate the Ni-NTA spin column with 600 µl Buffer NPI-10. Centrifuge for 2 min at 890 x *g* (approx. 2900 rpm).**

The spin columns should be centrifuged with an open lid to ensure that the centrifugation step is completed after 2 min.

By adding 10 mM imidazole, the binding of nontagged contaminating proteins is inhibited, leading to greater purity in fewer steps. If the tagged protein does not bind under these conditions the amount of imidazole should be reduced to 1–5 mM.

2 **Load up to 600 µl of the cleared lysate containing the 6xHis-tagged protein onto the pre-equilibrated Ni-NTA spin column. Centrifuge for 5 min at 270 x *g* (approx. 1600 rpm), and collect the flow-through.**

To ensure efficient binding it is important not to exceed 270 x *g* (approx. 1600 rpm) when centrifuging Ni-NTA spin columns. At higher forces, even if the binding kinetics are high, the time the lysate is in contact with the resin is not sufficient for effective binding.

The spin columns can be centrifuged with an open lid to ensure that the centrifugation step is completed after 5 min, but under native conditions, it may be preferable to centrifuge with a closed lid to reduce the flow rate thereby extending binding time.

Please take into account that the time needed for the centrifugation step during protein binding is influenced by the viscosity of the cleared lysate. For very concentrated cell lysates, it may be necessary to extend the centrifugation time to 5 or 10 min at 270 x *g* (approx. 1600 rpm).

Save the flow-through for analysis by SDS-PAGE to check binding efficiency.

3 **Wash the Ni-NTA spin column twice with 600 µl Buffer NPI-20. Centrifuge for 2 min at 890 x *g* (approx. 2900 rpm).**

The number of wash steps required to obtain highly pure protein is determined primarily by the expression level of the 6xHis-tagged protein. When the expression level is high, 2 washes are usually sufficient for removal of contaminants. For very low expression levels or highly concentrated lysates, 3 wash steps may be required to achieve high purity.

Save the flow-through (wash fractions) for analysis by SDS-PAGE to check the stringency of the wash conditions.

4 **Elute the protein twice with 300 µl Buffer NPI-500. Centrifuge for 2 min at 890 x *g* (approx. 2900 rpm), and collect the eluate.**

Most of the 6xHis-tagged protein (>80%) should elute in the first 300 µl eluate. The remainder will elute in the second 300 µl. If higher protein concentrations are desired, do not combine the eluates or, alternatively, elute in 100–200 µl aliquots.

Annexe 9 : Criblage métabolique des différents P450 candidats

Familles	Substrats	CYP71D97*	CYP71AZ1-nt*	CYP76B8	CYP82H1	CYP71D98-nt	CYP71AJ5
Coumarines	Umbelliférone	-	-	-	-	-	-
	Démetylsubérosine	-	-	-	-	-	-
	0-Prenylumbelliférone	-	-	-	-	-	-
Furocoumarines linéaires	Marmésine	-	-	-	-	-	-
	Psoralène	-	-	-	-	-	-
	Xanthoxol	-	-	-	-	-	-
	Bergaptol	-	-	-	-	-	+
	Xanthoxine	-	-	-	-	-	-
	Bergaptène	-	-	-	-	-	+
	Isopimpinélline	-	-	-	-	-	-
Furocoumarines angulaires	Columbianétine	-	-	-	-	-	-
	Angélicine	-	-	-	-	-	-
	Sphondine	-	-	-	-	-	-
	Pimpinélline	-	-	-	-	-	-
Flavanones	Liquiritigénine	-				-	
	Naringénine	-				-	
	Eriodictyole	-				-	
Flavones	Apigénine	-				-	
	Lutéoline	-				-	
	Flavone	-				-	
	5-OH Flavone	+				-	
	3-OH Flavone	-				-	
	6-OH Flavone	-				-	
	7-OH Flavone	-				-	
	7'8-OH Flavone	-				-	

Les différents substrats incubés avec les microsomes fraichement préparé de CYP71D97, CYP71AZ1-nt, CYP76B8, CYP82H1, CYP82H1 et CYP75AJ5. Le symbole * correspond aux microsomes dont l'expression a été observé par un spectre à 450 nm avant de réaliser les tests de métabolisations. Une activité a été détecté avec le 5-OH flavone pour CYP71D97, également avec le bergaptol et le bergaptène pour CYP71AJ5.

Annexe 10 : Publications scientifiques

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2012) doi: 10.1111/j.1365-313X.2011.04879.x

A 2-oxoglutarate-dependent dioxygenase from *Ruta graveolens* L. exhibits *p*-coumaroyl CoA 2′-hydroxylase activity (C2′H): a missing step in the synthesis of umbelliferone in plants

Guilhem Vialart[1,†], Alain Hehn[1,†,*], Alexandre Olry[1], Kyoko Ito[2], Celia Krieger[1], Romain Larbat[1], Cedric Paris[3], Bun-ichi Shimizu[4], Yukihiro Sugimoto[2], Masaharu Mizutani[1] and Frederic Bourgaud[1]

[1]*UMR 1121 Lorraine University, INRA Agronomie et Environnement Nancy –Colmar, ENSAIA, 2 avenue de la forêt de Haye, 54505 Vandœuvre-lès-Nancy, France,*

[2]*Graduate School of Agricultural Science, Kobe University, Rokkodai, Nada, Kobe 657-8501, Japan,*

[3]*Laboratoire Commun d'Analyse, Lorraine University, ENSAIA, 2 avenue de la forêt de Haye, 54505 Vandœuvre-lès-Nancy, France, and*

[4]*Graduate School of Life Science, Toyo University, Izumino 1-1-1, Itakura, Gunma 374-0193, Japan*

Received 26 July 2011; revised 6 December 2011; accepted 8 December 2011.

*For correspondence (fax +33 3 83 59 57 99; e-mail alain.hehn@ensaia.inpl-nancy.fr).

†These authors contributed equally to this work.

The GenBank accession number for the *RgJF799117* sequence is JF799117.

SUMMARY

Coumarins are important compounds that contribute to the adaptation of plants to biotic or abiotic stresses. Among coumarins, umbelliferone occupies a pivotal position in the plant phenylpropanoid network. Previous studies indicated that umbelliferone is derived from the *ortho*-hydroxylation of *p*-coumaric acid by an unknown biochemical step to yield 2,4-dihydroxycinnamic acid, which then undergoes spontaneous lactonization. Based on a recent report of a gene encoding a 2-oxoglutarate-dependent dioxygenase from *Arabidopsis thaliana* that exhibited feruloyl CoA 6′-hydroxylase activity (Bourgaud *et al.*, 2006), we combined a bioinformatic approach and a cDNA library screen to identify an orthologous ORF (Genbank accession number JF799117) from *Ruta graveolens* L. This ORF shares 59% amino acid identity with feruloyl CoA 6′-hydroxylase, was functionally expressed in *Escherichia coli*, and converted feruloyl CoA into scopoletin and *p*-coumaroyl CoA into umbelliferone with equal activity. Its bi-functionality was further confirmed *in planta*: transient expression of *JF799117* in *Nicotiana benthamiana* yielded plants with leaves containing high levels of umbelliferone and scopoletin when compared to control plants, which contained barely detectable traces of these compounds. The expression of *JF799117* was also tightly correlated to the amount of umbelliferone that was found in UV-elicited *R. graveolens* leaves. Therefore, *JF799117* encodes a *p*-coumaroyl CoA 2′-hydroxylase in *R. graveolens*, which represents a previously uncharacterized step in the synthesis of umbelliferone in plants. Psoralen, which is an important furanocoumarin in *R. graveolens*, was found to be a competitive inhibitor of the enzyme, and it may exert this effect through negative feedback on the enzyme at an upstream position in the pathway.

Keywords: umbelliferone, 2-oxoglutarate-dependent dioxygenase, coumarin, furanocoumarin, *Ruta graveolens*, scopoletin.

INTRODUCTION

Coumarins (1,2-benzopyrones) are ubiquitous in higher plants, where they display important biological and ecological functions. They contribute to the environmental adaptation of plants, and are primarily involved in defense against phytopathogens (Kùc, 1982), limitation of oxidative bursts (Chong *et al.*, 2002) and tolerance to abiotic stresses (Gutierrez *et al.*, 1995). Furanocoumarins (also known as furobenzo-α-pyrones or furocoumarins) constitute a subclass of coumarins that are generally found in four higher plant families: Apiaceae, Rutaceae, Moraceae and Fabaceae.

© 2011 The Authors
The Plant Journal © 2011 Blackwell Publishing Ltd

These compounds demonstrate stronger allelochemical properties than the simple coumarins, and have also been characterized for their contribution to resistance against plant pathogens (Beier and Oertli, 1983), insects (Schuler and Berenbaum, 2003) and environmental stresses (Zobel and Brown, 1993).

Umbelliferone (Figure 1) is a coumarin that is found in all higher plants; it is a pivotal intermediate in the synthesis of more complex derivatives, such as poly-hydroxylated/methoxylated/prenylated coumarins, furanocoumarins and pyranocoumarins (Bourgaud *et al.*, 2006). Early investigations were based on plant feeding and use of radiolabeled precursors; these studies provided data concerning the synthesis of umbelliferone in various model plants. Indeed, the first such experiment, performed on *Hierochloe odorata* and *Melilotus officinalis* (Brown *et al.*, 1960), as well as a complementary study also performed on *M. officinalis* (Ellis and Amrhein, 1971), demonstrated incorporation of cinnamic acid into coumarin and 7-hydroxycoumarins (i.e. umbelliferone derivatives). In addition, other tracer experiments were performed on *Lavandula officinalis*, and clearly showed that *para*-hydroxylation was a prerequisite for *ortho*-hydroxylation (Brown, 1962). This result unequivocally demonstrated that umbelliferone is formed after the *ortho*-hydroxylation of *p*-coumaric that leads to synthesis of 2,4-dihydroxycinnamic acid, which is subsequently transformed into umbelliferone following the spontaneous closure of the lactone ring by a non-enzymatic reaction that occurs under acidic or neutral conditions (Figure 1) (reviewed by Bourgaud *et al.*, 2006). In another set of pioneering experiments, double-labeled (*ortho*-^{3}H, *ring*-1-^{14}C) cinnamic acid was fed to *Melilotus alba* or *Gaultheria*

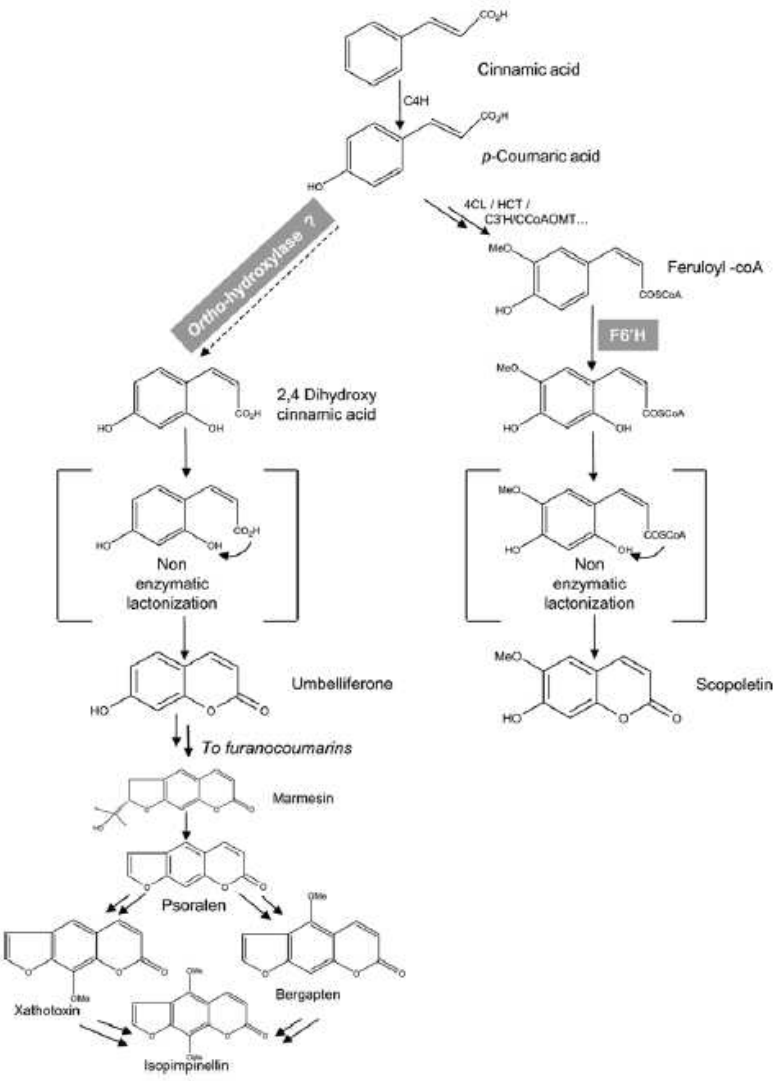


Figure 1. Schematic representation of the umbelliferone/furanocoumarins and scopoletin biosynthetic pathways.
Multiple arrows represent multi-enzymatic steps.
Dashed arrow: *ortho*-hydroxylation step to umbelliferone.

© 2011 The Authors

procumbens, and retention of the label was monitored upon its conversion to *o*-coumaric acid, which is the immediate precursor of coumarin (Ellis and Amrhein, 1971). The significant decrease in the ^{3}H:^{14}C ratio was interpreted as an indication of a cytochrome P450 mono-oxygenase reaction mechanism. Another study of *M. alba* assigned the *ortho*-hydroxylation of cinnamic acid to the chloroplast; again, the results suggested the involvement of a P450-dependent mechanism (Gestetner and Conn, 1974). Unfortunately, these *in vitro* results could not be reproduced, and the enzyme family involved in the *ortho*-hydroxylation of *p*-coumaric acid remained a missing link in the process of phenylpropanoid biosynthesis.

The first evidence demonstrating the pivotal role of a 2-oxoglutarate-dependent dioxygenase (2-OGD) in the *ortho*-hydroxylation of cinnamic acid derivatives was only obtained recently, when the biosynthetic pathway of scopoletin (6-methoxy,7-hydroxy-coumarin) was established in *Arabidopsis thaliana*. The first set of experiments utilized T-DNA insertion mutants within the gene that encodes CYP98A3, which is a P450-dependent enzyme that catalyzes the 3'-hydroxylation of *p*-coumaroyl CoA shikimate or quinate (Kai *et al.*, 2006). Transgenic knockout plants exhibited dramatic decreases in scopoletin, confirming its origin from ferulic acid in *A. thaliana* (Figure 1). The feruloyl CoA 6'-hydroxylase (F6'H) that is responsible for conversion of feruloyl CoA into scopoletin was then isolated and characterized as a 2-OGD enzyme (Kai *et al.*, 2008). Because scopoletin is structurally similar to umbelliferone, it is tempting to assume that the *ortho*-hydroxylation of *p*-coumaric acid that facilitates the formation of umbelliferone may be catalyzed by an enzyme that is similar to F6'H. However, F6'H displayed a narrow specificity; feruloyl CoA was the only substrate that was accepted among cinnamic acid derivatives. This finding left the method of *ortho*-hydroxylation of *p*-coumaric acid unresolved, and the process of synthesis of umbelliferone continued to be debated.

Here we describe the identification of a gene that encodes a 2-OGD in *R. graveolens*, which is a furanocoumarin-producing plant; it catalyzes the conversion of feruloyl CoA into scopoletin and *p*-coumaroyl CoA into 2,4-dihydroxycinnamic acid. This latter activity corresponds to a *p*-coumaroyl CoA 2'-hydroxylase (C2'H), which is the missing step of umbelliferone synthesis in plants. This bi-functional activity was confirmed *in vitro* using a purified enzyme that was heterologously produced in *Escherichia coli* and *in planta* following transient expression of the corresponding gene in *Nicotiana benthamiana*. Further *in vitro* studies revealed that this enzyme is negatively regulated by psoralen, which is a furanocoumarin with high toxicity potential in *R. graveolens*; this result demonstrates possible feedback control by this compound upstream of the pathway.

RESULTS

Involvement of a 2-OGD in the furanocoumarin pathway of *Ruta graveolens* and identification of a candidate gene from this plant

The occurrence of a 2-OGD in a *R. graveolens* furanocoumarin pathway was investigated in plantlets that had been sprayed with prohexadione-calcium, which is an inhibitor of this enzyme family (Puhl *et al.*, 2008). The decrease in furanocoumarin content was statistically significant in the leaves of treated plants when compared to control plants: the levels in the control and treated plants were 22.6 ± 2.6 and 15.1 ± 7.1 mg g^{-1} FW (fresh weight) for bergapten ($P < 0.05$), 1.0 ± 0.2 and 0.6 ± 0.3 mg g^{-1} FW for psoralen ($P < 0.05$) and 6.7 ± 0.5 and 5.9 ± 0.6 mg g^{-1} FW for xanthotoxin ($P < 0.05$), respectively. The umbelliferone content also decreased significantly ($P < 0.05$) between control and treated plants (0.05 ± 0.02 and 0.03 ± 0.01 mg g^{-1} FW, respectively), which are unexpectedly low concentrations.

As the first attempt to identify a putative 2-OGD from *R. graveolens*, the *Arabidopsis thaliana* F6'H nucleotide sequence (*At3g13610*) (Kai *et al.*, 2008) was used as a template to identify *in silico* candidates in the Rutaceae family that shared a high level of homology. The first tBLASTn search performed in the GenBank EST database revealed no *R. graveolens* accessions because only a small number of genes are available for this plant species. However, the same *in silico* analysis identified several candidates belonging to the Rutaceae family, and two gene fragments (GenBank accession numbers CX667249 and CX667579) that corresponded to opposite ends of the same sequence revealed the presence of a putative 2-OGD in a *Citrus paradisi* × *Poncirus trifoliata* hybrid. The ORF was amplified from cDNAs prepared from RNA isolated from *Citrus limetta*, using primers based on the two gene fragments. The resulting sequence (GenBank accession number JN245615) was not completely identical to the gene described in the EST database (93% nucleotide identity and 94% amino acid identity), but it showed 63% nucleotide identity and 61% amino acid identity with *A. thaliana* At3g13610. To determine its activity, the gene was introduced into a bacterial expression system. Although the corresponding protein was expressed efficiently in this heterologous system (data not shown), no reaction products of cinnamic acid derivatives were detected.

Based on its high level of identity with the citrus and *A. thaliana* genes previously described, the *JN245615* sequence was used to screen a bacteriophage library to identify an orthologous 2-OGD-encoding gene. The bacteriophage library contained cDNAs prepared from RNA extracted from *R. graveolens* leaves that had been exposed to UV light. This experiment enabled isolation of a new gene (GenBank accession number JF799117, which is hereafter referred to as *RgJF799117*) (Figure 2).

© 2011 The Authors
The Plant Journal © 2011 Blackwell Publishing Ltd, *The Plant Journal*, (2012), doi: 10.1111/j.1365-313X.2011.04879.x

```
At3g13610    MAPT---LLTTQFSNPAEVTDFVVYKGNGVKGLSETGIKALPEQYIQPLE 47
Ruta 2OGD    MAPTKDSVIHMGAESWDEISEFVTKKGHGVKGLSELGIKTLPKQFHQPLE 50
             ****   ::    ..  *::.**. **:******* ***:**:*: ****

At3g13610    ERLINKFVNETDEAIPVIDMSNPDEDRVAEAVCDAAEKWGFFQVINHGVP 97
Ruta 2OGD    ERFSEKKILER-ASIPLIDMSKWDSPEVVKSICDAAEHWGFFQIVNHGVP 99
             **: :* : *   :**:****: *. .*.:::*****:*****::*****

At3g13610    LEVLDDVKAATHKFFNLPVEEKRKFTKENSLSTTVRFGTSFSPLAEQALE 147
Ruta 2OGD    LETLQRVKEATHRFFALPAEEKNKYSKENSPINNVRFGSSFVPHVEKALE 149
             **.*. ** ***.** **.***.*::****  ..****:** * .*:***

At3g13610    WKDYLSLFFVSEAEAEQFWPDICRNETLEYINKSKKMVRRLLEYLG-KNL 196
Ruta 2OGD    WKDFLSMFYVSEEETNTYWPPICRDEMLEYMRSSEVLIKRLMEVLVVKGL 199
             ***:**:*:*** *:: :** ***.* ***:..*: ::.**:* *  *.*

At3g13610    NVKELDETKESLFMGSIRVNLNYYPICPNPDLTVGVGRHSDVSSLTILLQ 246
Ruta 2OGD    KVKQIDEIREPMLVGSRRINLNYYPKCPNPELTLGVGRHSDISTFTILLQ 249
             :**::** .*.:::** *:****** ****.**:*******:*::*****

At3g13610    DQIGGLHVRSLASG--NWVHVPPVAGSFVINIGDAMQIMSNGLYKSVEHR 294
Ruta 2OGD    DEIGGLHVRKLDDTGNTWVHVTPISGSLIINIGDALQIMSNGRYKSIHHM 299
             *:*******.* .   .****.*::**::******:****** ***:**

At3g13610    VLANGYNRLSVPIFVNPKPESVIGPLPEVIANGEEPIYRDVLYSDYVKY 344
Ruta 2OGD    VVANGTQRLSVPLFVNPKPQAILCPFPEVLANGEKPVYKPVLCSDYSRH 349
             *:*** :.*****:******:::: *:***:****:*:*. ** *** .:

At3g13610    FFRKAHDGKKTVDYAKI- 361
Ruta 2OGD    FYTKPHDGKKTVDFALMN 367
             *: *.********:* :
```

Figure 2. Sequence alignment of the *Ruta graveolens* putative 2-OGD (Rg 2-OGD GenBank accession number *JF799117*) and the *Arabidopsis thaliana* 2-OGD (*At3g13610*). The sequences were aligned using Clustal W. Asterisks indicate identity between the two sequences; colons indicate conservative substitutions; dots indicate semi-conservative substitutions. The dots and triangles above the alignment indicate amino acids (highlighted) that are involved in recognition sites for oxoglutarate and ferrous iron, respectively.

A phylogenic analysis of 2-OGD sequences (Figure 3) revealed that this enzyme family is organized in clusters, and that each cluster is associated with a specific enzymatic function. In this phylogenic tree, the cloned *RgJF799117* sequence belongs to the same cluster as the two *A. thaliana* sequences that were previously described as F6′Hs (Kai *et al.*, 2008). RgJF799117 shares 59% amino acid identity with At3g13610 (Figure 2) and 81% amino acid identity with citrus JN245615.

In vitro functional characterization

The coding region of the candidate gene *RgJF799117* was cloned and fused with a histidine tag in *E. coli*. Protein extracts were purified by affinity chromatography and incubated in the presence of cinnamic acid, caffeic acid, *p*-coumaric acid, ferulic acid, cinnamoyl CoA, caffeoyl CoA, *p*-coumaroyl CoA or feruloyl CoA. HPLC analysis of the incubations revealed that no conversion occurred with either the four acids or the cinnamoyl and caffeoyl esters. However, the incubations with *p*-coumaroyl CoA and feruloyl CoA led to the accumulation of two products, P1 and P2 (Figure 4a,d), that were absent from the control samples. The UV and MS/MS spectra of P1, which formed from *p*-coumaroyl CoA, perfectly matched the spectra from a commercial umbelliferone standard (Figure 4b,c), whereas P2, which formed from feruloyl CoA, was identified as scopoletin (Figure 4e,f).

Kinetic assays revealed a similar affinity and efficiency for *p*-coumaroyl CoA (K_m = 50 ± 10 μM, $k_{cat\ coum}/K_m$ = 0.85 μmol min^{-1}) and feruloyl CoA (K_m = 37 ± 13 μM, $k_{cat\ fer}/K_m$ = 0.75 μmol min^{-1}) (Table 1). To complete the kinetic study, inhibition tests were performed using umbelliferone, psoralen, xanthotoxin, bergapten and isopimpinellin. Only psoralen had a competitive inhibitory effect on the reaction, characterized by an increase in the K_m in a concentration-dependent manner, and no modification of the V_{max} (Table 1). The estimated K_i values were 111 μM for *p*-coumaroyl CoA and 122 μM for feruloyl CoA.

In planta functional characterization

Further functional characterization of *RgJF799117* was achieved by its transient expression in *Nicotiana benthamiana*. The *RgJF799117* ORF was introduced into a pBIN binary plasmid under the control of the CaMV 35S promoter. The recombinant plasmid was then introduced into *Agrobacterium tumefaciens* strain LBA4404, and was co-inoculated into *N. benthamiana* leaves with an *Agrobacterium* strain expressing the P19 silencing suppressor viral protein of tomato bushy stunt virus (Voinnet *et al.*, 2003). Phenolic extracts from 3-day-old infiltrated leaves contained significant

© 2011 The Authors

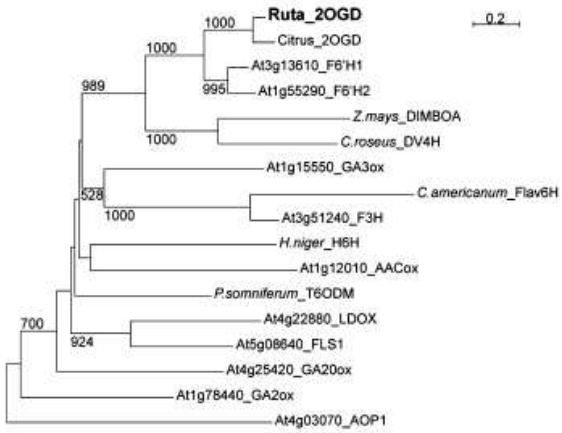


Figure 3. Phylogenetic tree of the 2-OGD family.
The accession numbers in GenBank/EMBL/DDBJ database are given in parentheses: Ruta_2OGD (JF799117); Citrus_2OGD (JN245615); At1g78440_GA2ox (NM_106491); At1g15550_GA3ox (NM_101424); At4g25420_GA20ox (NM_118674); At1g12010_ACCox (NM_101073); At3g13610_F6'H1 (NM_112207); At1g55290_F6'H2 (NM_104404); At3g51240_F3H (NM_114983); At4g22880_LDOX (NM_118417); At5g08640_FLS1 (NM_120951); At4g03070_AOP1 (AF417855); *Zea mays* DIBOA 7-hydroxylase (NM_001111630) (Frey *et al.*, 2003); *Catharanthus roseus* desacetoxyvindoline-4-hydroxylase (AF008597) (Vazquez-Flota *et al.*, 1997); *Chrysosplenium americanum* flavonol 6-hydroxylase (AY605048) (Anzelotti and Ibrahim, 2004); *Papaver somniferum* thebaine 6-*O*-demethylase (GQ500139) (Hagel and Facchini, 2010) and *Hyoscyamus niger* hyoscyamine 6β-hydroxylase (HYSH6H) (Matsuda *et al.*, 1991).

levels of umbelliferone and scopoletin, as revealed by UV and mass spectral analyses; only traces of these products were detected in the control plants (Figure 5). These results from *N. benthamiana* are consistent with the *in vitro* enzymatic investigations, and support the bi-functionality of the enzyme encoded by *RgJF799117* in synthesis of umbelliferone and scopoletin.

Given the potential bi-functionality of *RgJF799117*, the umbelliferone/scopoletin balance was investigated in *R. graveolens* tissues. Scopoletin was not detected by HPLC-DAD in any tissue, but umbelliferone was present in the leaves (0.4 ± 0.1 mg g^{-1} FW) and petioles (0.5 ± 0.1 mg g^{-1} FW), and in trace amounts (<0.05 mg g^{-1} FW) in the stems and roots. Complementary experiments using HPLC-MS on the same plant tissues indicated that scopoletin was present at low concentrations (<10 ng g^{-1} FW) in addition to scopolin and skimmin, which are the glycosylated derivatives of scopoletin and umbelliferone.

Previous experiments indicated that UV light can induce the production of furanocoumarin in *R. graveolens* seedlings (Zobel and Brown, 1993; Zobel *et al.*, 1998). To determine whether there was a correlation between the accumulation of umbelliferone/furanocoumarins and the expression of *RgJF99117*, 3-month-old plantlets were irradiated with UV-B light (312 nm) for 24 h, and compared to non-irradiated control plants. Quantification of the transcripts by real-time PCR revealed a significant induction of *RgJF799117* in the treated leaves and petioles (Figure 6), but no change was detected in the other vegetative tissues (roots and stem). Similarly, the umbelliferone concentration was significantly enhanced ($P < 0.01$) in the UV-treated leaves (1.4 ± 0.1 mg g^{-1} FW) compared to the control plants (0.4 ± 0.1 mg g^{-1} FW). However, the levels were only slightly increased ($0.1 > P > 0.05$) in the UV-treated petioles (0.8 ± 0.2 mg g^{-1} FW) compared to the control plants (0.5 ± 0.1 mg g^{-1} FW), and did not change in the roots and stems (<0.05 mg g^{-1} FW). No significant difference was observed for the levels of furanocoumarins between the UV-treated and control plants. Together, these results show a correlation between the expression of *RgJF799117* and the increase in umbelliferone content in *R. graveolens* leaves and petioles. In a second set of experiments, we analyzed the effect of prohexadion-calcium in UV-treated plants. The results confirmed a decrease in umbelliferone and furanocoumarins for plants treated with both UV and prohexadion-calcium compared to plants with a single UV treatment (umbelliferone, 0.04 ± 0.02 versus 0.08 ± 0.01 mg g^{-1} FW, $P < 0.01$; xanthotoxin, 6.1 ± 0.7 versus 6.8 ± 0.4 mg g^{-1} FW, $P < 0.05$; bergapten, 16.8 ± 5.9 versus 24.4 ± 1.5 mg g^{-1} FW, $P < 0.01$; marmesin, 1.2 ± 0.3 versus 1.8 ± 0.4 mg g^{-1} FW, $P < 0.01$; psoralen, 0.7 ± 0.3 versus 1.1 ± 0.2 mg g^{-1} FW, $P < 0.01$).

DISCUSSION

Umbelliferone is an important coumarin that is ubiquitously produced in plants. Previous studies have demonstrated that umbelliferone is formed following the *ortho*-hydroxylation of *p*-coumaric acid, but molecular evidence regarding its synthesis was unavailable until recently. Elucidation of this *ortho*-hydroxylation step began in 2008 with studies demonstrating the role of a 2-OGD in hydroxylation of feruloyl CoA (F6'H), leading to the formation of scopoletin in *A. thaliana* (Kai *et al.*, 2008). However, it was shown experimentally that *A. thaliana* F6'H is not committed to the formation of umbelliferone (Kai *et al.*, 2008), and the method of biosynthesis of this compound remains unresolved.

Due to obvious structural similarities between scopoletin and umbelliferone, we decided to examine the possible involvement of a 2-OGD in the synthesis of umbelliferone. Investigations were performed in *R. graveolens*, a plant that synthesizes large amounts of furanocoumarins (Milesi *et al.*, 2001), and for which umbelliferone is an important intermediate (Figure 1). The occurrence of a 2-OGD in the furanocoumarin pathway of *R. graveolens* plants was investigated in experiments that utilized prohexadione-calcium, a 2-OGD inhibitor (Puhl *et al.*, 2008), with or without UV treatment. The prohexadione-calcium-treated plants showed a significant decrease in leaf umbelliferone and furanocoumarin

© 2011 The Authors

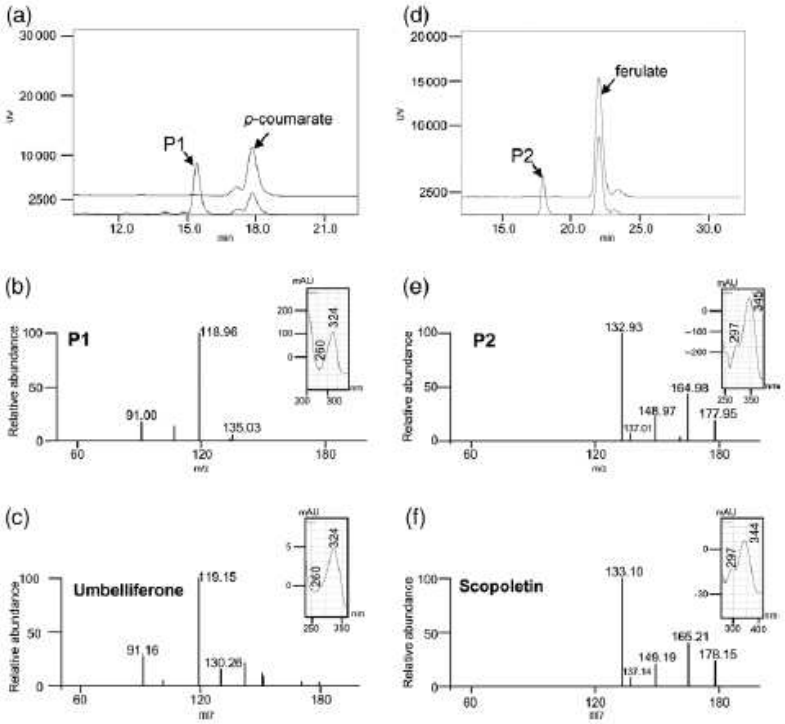


Figure 4. *In vitro* functional characterization of the enzyme encoded by *RgJF799117*.
(a,d) HPLC analysis of *p*-coumaroyl CoA (a) and feruloyl CoA (d) metabolites P1 and P2, respectively.
(b,e) UV and MS spectra of the products P1 (b) and P2 (e).
(c,f) UV and MS spectra analysis of authentic umbelliferone (c) and scopoletin (f) standards.

Table 1 Kinetic constants for Rg 2-OGD heterologously expressed in *Escherichia coli*

	Feruloyl CoA	*p*-coumaroyl CoA
Kinetic constants		
V_{max} (sec^{-1})	0.46 ± 0.05	0.71 ± 0.06
K_m (μM)	37 ± 13	50 ± 10
K_{cat}/K_m (μmol min^{-1})	0.75	0.85
Inhibition with 50 μM psoralen		
$V_{max\ app}$ (sec^{-1})	0.47 ± 0.05	0.69 ± 0.08
$K_{m\ app}$ (μM)	64 ± 14	77 ± 9
Inhibition with 100 μM psoralen		
$V_{max\ app}$ (sec^{-1})	0.45 ± 0.05	0.69 ± 0.08
$K_{m\ app}$ (μM)	73 ± 14	96 ± 9
Inhibition with 200 μM psoralen		
$V_{max\ app}$ (sec^{-1})	0.46 ± 0.06	0.65 ± 0.08
$K_{m\ app}$ (μM)	107 ± 17	108 ± 9

contents compared to untreated plants, supporting the hypothesis that a 2-OGD is involved in this pathway. *In silico* studies and a screen of a *R. graveolens* cDNA library based on the F6′H sequence revealed a putative 2-OGD gene (GenBank accession number JF799117) that was 59% identical to the *A. thaliana* F6′H sequence. Heterologous expression of *RgJF799117* in *E. coli* enabled an *in vitro* screen for enzymatic activity, and clearly showed that the corresponding enzyme accepts two substrates, feruloyl CoA and *p*-coumaroyl CoA, leading to the formation of scopoletin and umbelliferone, respectively, with almost the same efficiency for each substrate (Table 1). The bi-functionality of this enzyme was also tested in *N. benthamiana* transiently expressing the gene under the control of the CaMV 35S promoter. It is well known that tobacco accumulates scopoletin and scopolin, which is the glycoside derivative of scopoletin, but not umbelliferone (Chong *et al.*, 2002; Costet *et al.*, 2002). Treated plants harvested 3 days post-infection had similarly high levels of umbelliferone and scopoletin, and these compounds were both absent in the inoculated control plants. The accumulation of coumarin products observed in *N. benthamiana* clearly indicates that the enzyme encoded by *RgJF799117* is fully functional in plants, and that it can synthesize scopoletin and umbelliferone. The high level of umbelliferone found in tobacco plants expressing *RgJF799117* is particularly interesting with respect to the role of this enzyme in the synthesis of umbelliferone, because this plant does not normally accumulate umbelliferone. In addition, quantification of *RgJF799117* transcripts in *R. graveolens* demonstrated that expression of this gene

© 2011 The Authors

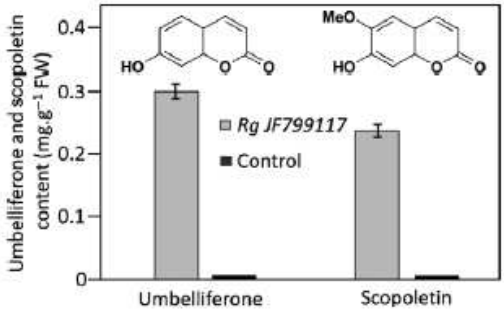


Figure 5. Quantification of umbelliferone and scopoletin in the leaves of *Nicotiana benthamiana* plants transiently expressing *RgJF799117*

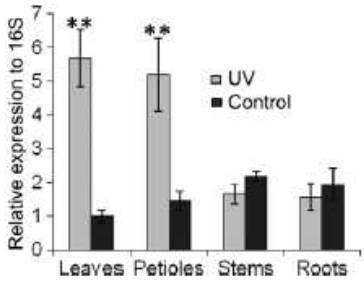


Figure 6. UV-B effects on the *RgJF799117* expression level in *Ruta graveolens* vegetative tissues.
Error bars represent the SE for three replicates. The statistical significance was assessed by Student's *t* test. Asterisks indicate a statistically significant difference ($P < 0.01$) compared to the control.

is positively correlated with the accumulation of umbelliferone in UV-exposed leaves and petioles of this plant. Taken together, the *in planta* results confirm the bi-functional activity of the enzyme. Therefore, *RgJF799117* encodes a 2-OGD that displays both F6'H and C2'H activities in *R. graveolens*. Because the umbelliferone produced by this 2-OGD is derived from *p*-coumaroyl CoA rather than free *p*-coumaric acid, this enzyme probably operates closely with a 4-coumarate CoA-ligase (4CL) that is responsible for synthesis of the *p*-coumaroyl CoA ester. This theory is consistent with other studies in *R. graveolens*, in which two different 4CLs were described (Endler *et al.*, 2008). One of the two *R. graveolens* 4CLs, which is specifically induced by elicitation of the cells, has been suggested to be specifically engaged in coumarin biosynthesis (Endler *et al.*, 2008), and may operate in tandem with the C2'H described above.

Previously, Lukacin *et al.* (2003) also identified a bi-functional 2-OGD that is involved in citrus flavonoid metabolism (Lukacin *et al.*, 2003). Multi-functional enzymes are essential elements that explain the complexity of the metabolic grids found in plants (Schwab, 2003). In addition to F6'H, the presence of C2'H activity in *R. graveolens* allows the generation of umbelliferone and more complex 7-hydroxycoumarin derivatives, such as furanocoumarins, which have important defense functions in plants. These furanocoumarins are absent from *A. thaliana*, which contains only a mono-functional F6'H (Kai *et al.*, 2008). The presence of scopoletin, which is the immediate product of F6'H activity, was investigated, but it was not detected at significant levels in any *R. graveolens* tissues.

The lack of scopoletin in our experiments is puzzling in light of the F6'H activity observed for *RgJF799117* because F6'H was found to be equivalent to C2'H in terms of its activity *in vitro* (heterologous expression in *E. coli*, Table 1) and *in planta* (transient expression in *N. benthamiana*, Figure 5); it would be reasonable to expect co-occurrence of scopoletin and umbelliferone in *R. graveolens*. The formation of associated products downstream of scopoletin may explain the lack of scopoletin itself. However, scopolin (the glycosidic end-product that is synthesized from scopoletin) was also present in only trace amounts in *R. graveolens*, contradicting the above hypothesis. Other phytochemical studies have confirmed the low level of scopoletin in *R. graveolens*: scopoletin has frequently been described to be a minor coumarin in *R. graveolens* (Steck *et al.*, 1971; Paszkiewicz *et al.*, 2008), and a recent report revealed that its concentration is below the HPLC-UV quantification limit (<2.5 $\mu g\ g^{-1}$ FW) (Adam *et al.*, 2009). Although the F6'H activity of the protein that is encoded by *RgJF799117* seems to be fully functional, the reason for the absence of significant levels of scopoletin in *R. graveolens* is currently unresolved, and will require further investigation.

Phylogenetic analyses facilitated the construction of a tree in which 2-OGDs are organized according to their functional diversity (Figure 3). To product of *RgJF99117* shares the highest identity with a peptide sequence isolated from *Citrus limetta* (NCBI accession number JN245615). For an unknown reason, functional characterization of this gene could not be achieved using the *E. coli* heterologous expression system. However, the fact that this citrus gene is the closest neighbor of *RgJF99117* is consistent with the observation that citrus plants synthesize large amounts of furanocoumarins and other umbelliferone derivatives; therefore, it is probably necessary for these plants to possess a C2'H function similar to that in *R. graveolens*. Our sequence corresponding to a C2'H and the two *A. thaliana* F6'H sequences are grouped together and constitute a separate clade.

Kinetic experiments with various coumarins and furanocoumarins revealed that psoralen is a competitive inhibitor of the C2'H/F6'H encoded by *RgJF99117* (Table 1). Furanocoumarins are natural compounds with potent allochemical functions because of their severe cellular toxicity. The severe genotoxicity of linear furanocoumarins (i.e. psoralen derivatives) is a consequence of their ability to intercalate with double-stranded DNA and create covalent

© 2011 The Authors

cross-links between thymine residues (Dall'Acqua *et al.*, 1978). Linear furanocoumarins are therefore potentially toxic to any living organism, including plants that synthesize large quantities of these compounds, such as *R. graveolens*. The K_i that was experimentally determined for psoralen using the *R. graveolens* C2′H/F6′H enzyme was approximately twice the K_m value (Table 1). Although the enzyme affinity is less pronounced for psoralen than for its two substrates (feruloyl and *p*-coumaroyl CoA), psoralen may exert a feedback inhibition on C2′H at higher concentrations, and this may down-regulate the general umbelliferone/furanocoumarin pathway. In *R. graveolens*, this down-regulation pattern is consistent with reports showing constitutive synthesis of furanocoumarins in all the tissues from this plant (Eilert, 1989; Milesi *et al.*, 2001) and the UV elicitation experiments reported here, in which furanocoumarin concentrations in vegetative organs did not significantly increase. In contrast to most plants that accumulate furanocoumarins and exhibit a highly inducible biosynthetic pathway (Hamerski and Matern, 1988) based on differentially expressed genes (Larbat *et al.*, 2007), *R. graveolens* may have developed an alternative regulation system, such as enzyme retro-inhibition.

Finally, two independent studies from the 1960s and 1970s hypothesized that the *ortho*-hydroxylation of *p*-coumaric acid leading to formation of umbelliferone was controlled by a P450-dependent mono-oxygenase. Here, we provide evidence that a 2-OGD is committed to this synthesis in *R. graveolens*, and describe the first *p*-coumaroyl CoA 2′-hydroxylase to have been characterized in plants. In addition to the formation of umbelliferone, this enzyme converts feruloyl CoA into scopoletin (F6′H activity) with equivalent efficiency, and the enzyme is therefore fully bi-functional. The C2′H and F6′H activities are equally repressed by psoralen, which is a competitive inhibitor located downstream of the umbelliferone pathway. Psoralen down-regulates the C2′H function, subsequently reducing the metabolic flux through the umbelliferone/furanocoumarin pathway and preserving the plant tissues from excessive psoralen concentrations. Some higher plants display enzymatic function that may be controlled by a P450 or a 2-OGD enzyme. In Apiaceae (Martens *et al.*, 2001) and *Equisetum arvense* (Bredebach *et al.*, 2011), [2*S*]-naringenin is converted into apigenin by a flavone synthase I that belongs to the oxoglutarate-dependent dioxygenase family. However, in Fabaceae species such as *Glycyrrhiza echinita* (Akashi *et al.*, 1998) and soybean (*Glycine max*) (Fliegmann *et al.*, 2010), this flavone formation is performed by enzymes of the CYP93B P450 family. Further investigations are required to determine whether C2′H is encoded by the 2-oxaloglutarate-dependent dioxygenase superfamily alone, or whether other mono-oxygenase enzymes (such as P450s) are also involved, as previously assumed.

EXPERIMENTAL PROCEDURES

Plant material

Ruta graveolens and *Nicotiana benthamiana* seeds were sown and germinated in a greenhouse on potting soil. The seedlings were grown for 1–2 months before being used in the experiments.

UV-B light elicitation of the plants was achieved as follows: 15 cm tall *R. graveolens* plants were cultured under greenhouse conditions, and were irradiated for 8 h in the dark using a monochromatic lamp (312 nm) that was placed 25 cm from the plants. The plants were rotated every 2 h within the irradiation area to minimize any positional effects prior to RNA extraction.

In vivo inhibition of 2-oxoglutarate-dependent dioxygenases

Prohexadione-calcium (Regalis; BASF, http://www.basf.fr) is a non-specific inhibitor of 2-oxoglutarate-dependent dioxygenases (2-OGD) (Puhl *et al.*, 2008). Prohexadione-calcium was diluted with ultrapure water to a final concentration of 500 ppm, and was administered to the aerial parts of the plant by spraying. Two treatments were administered with a 6 h interval. The control plants were sprayed with ultrapure water. Furanocoumarins were extracted from leaves 4 days post-inoculation, and analyzed by HPLC-DAD and LC-MS. Each treatment was performed on six plants.

Substrate synthesis

Enzymatic synthesis of the CoA substrates cinnamoyl CoA, caffeoyl CoA, feruloyl CoA and *p*-coumaroyl CoA was performed using *A. thaliana* 4-coumarate CoA ligase (4CL) produced in *E. coli* M15 transformed with the pQE30/4CL plasmid (Lee and Douglas, 1996). LB medium (Bacto-Tryptone, 10 g/l; Bacto-yeast extract, 5 g/l; NaCl, 10 g/l) (100 ml) was supplemented with ampicillin (100 mg L^{-1}), and then inoculated at 37°C with transformed bacteria until the OD_{600} reached 0.7. Induction of 4CL synthesis was achieved by adding isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM for 20 h at 30°C. 4CL was purified using the Ni-NTA purification system (Qiagen, http://www.qiagen.com/) according to the manufacturer's instructions.

The CoA esters were synthesized in a final volume of 1 ml at room temperature from the corresponding acids (0.2 mM): ferulic acid, *p*-coumaric acid, caffeic acid and cinnamic acid (Sigma, http://www.sigmaaldrich.com/). The reaction mixture contained 0.2 mM coenzyme A, 2.5 mM ATP, 1 mM DTT, 2.5 mM $MgCl_2$ and purified enzyme (20 mg protein ml^{-1}). The kinetics of the chemical reaction were followed spectrophotometrically at 346 nm for feruloyl CoA and caffeoyl CoA, at 333 nm for *p*-coumaroyl CoA, and at 311 nm for cinnamoyl CoA.

cDNA library screening

Total RNA was extracted from *R. graveolens* leaves that had been exposed to UV light for 24 h using an RNeasy plant mini kit (Qiagen). The total RNA (600 ng) was reverse-transcribed and used to construct a Lambda TriplEx2 cDNA library using a SMART cDNA library construction kit (Clontech, http://www.clontech.com/) according to the manufacturer's instructions.

The protein sequence of *A. thaliana* F6′H identified by Kai *et al.* (2008) was used to perform a tBLASTn search on the Expressed Sequence Tags Databases available in GenBank. This *in silico* screen permitted the reconstruction of a citrus gene that had a high degree of homology with *A. thaliana* F6′H. Primers were designed to amplify this gene from RNA extracted from *Citrus limetta*

© 2011 The Authors

(CitrusFwd, 5′-CCATATGGCACCAACAAATGGCT-3′; CitrusRev, 5′-CGGATCCTCACCTATCTTTGCAAAGCT-3′). Additional *Nde*I and *Bam*HI restriction sites (underlined) were added at the 5′ and 3′ ends to facilitate subsequent cloning into an expression plasmid. Total RNA was extracted from *Citrus limetta* leaves using an RNeasy plant extraction kit (Qiagen) according to the manufacturer's instructions. The gene was amplified using SuperScript™ One-Step RT-PCR with Platinum® *Taq* polymerase (Invitrogen, http://www.invitrogen.com/). The reverse transcription step was performed at 45°C for 30 min. The PCR amplification was performed with an initial step of 94°C for 5 min, five cycles of 94°C for 15 sec, 40°C for 30 sec and 72°C for 2 min, then 25 cycles of 94°C for 15 sec, 52°C for 30 sec and 72°C for 2 min, and a final step at 72°C for 10 min. The resulting amplified product was cloned into the pCR8® plasmid (Invitrogen) and sequenced (accession number JN245615). This gene was then used to synthesize a probe using an AlkPhos Direct™ labeling and detection systems labeling kit (GE Healthcare, http://www.gelifesciences.com). Library screening was performed using the cDNA library from *R. graveolens*. The bacteriophages were inoculated into *E. coli* XL1 Blue, and spread on an LB plate that contained 10 mM $MgSO_4$ to generate 200 000 lysis plaques (150 mm Petri dishes). The bacteriophages were blotted onto Amersham Hybond™ membrane, denatured for 3 min in 0.5 M NaOH and 1.5 M NaCl, neutralized for 3 min in 10 mM Tris/HCl and 1.5 M NaCl, and washed in 2×SSC for 3 min. Finally, the phage DNA was linked to the membrane using a BIO-LINK BLX254 crosslinker (Fisher Scientific - http://www.fr.fishersci.com) at 254 nm and 0.12 J for 30 sec. The probe was hybridized at 52°C overnight, and two washes with 2x SSC for 10 min were performed at 55°C. Detection was achieved by chemiluminescence using CDP Star (Amersham). The plasmids were extracted from lysis plates that were identified as positive as described in the protocol for the SMART cDNA library construction kit using a Cre/Lox excision system in *E. coli* BM25.8. The isolated gene was sequenced (GenBank accession number *JF799117*).

Bacterial heterologous expression of Rg 2-OGD

The coding sequence for the *R. graveolens RgJF799117* gene was amplified from the pTriplEx plasmid using PCR with primers Rg2OGDFor (5′-CGGATCCATGGCACCGACAAAAGA-3′) and Rg2OGDRev (5′-CCTCGAGTTAATTCATCAAAGCGAAGTCGACAGTTTTC-3′). An additional *Eco*RI site at the 5′ end and an additional *Xho*I site at the 3′ end (underlined) were added to facilitate sub-cloning into the bacterial expression vector. PCR amplification was performed with an initial step of 94°C for 5 min, five cycles of 94°C for 15 sec, 40°C for 30 sec and 72°C for 2 min, then 25 cycles of 94°C for 15 sec, 52°C for 30 sec and 72°C for 2 min, and a final step at 72°C for 10 min. The gene was cloned into pET28a plasmid (Clontech) to produce a gene that allowed expression of a fusion protein with a C-terminal histidine tag. *JN245615* was sub-cloned from the pCR8 plasmid and introduced into the pET3a plasmid using the *Nde*I and *Bam*HI restriction sites that had been added to the 5′ and 3′ ends during the RT-PCR step. The recombinant plasmids were finally introduced into the *E. coli* BL21 (DE3) strain.

To produce the recombinant proteins, the bacteria were cultured in 100 ml LB medium supplemented with kanamycin (50 μg ml^{-1}) until the OD_{600} reached 0.7. Protein synthesis was then induced by adding isopropyl-thio-β-D-galactopyranoside to a concentration of 1 mM for 20 h at 26°C. The culture was then centrifuged at 5000 ***g*** for 10 min at 4°C. The cell pellets were resuspended in 5 ml cold PBS, sonicated using a Bandelin Sonopuls HD 2070 ultrasonic homogenizer (Sigma) type MS 73 at a sound intensity of 200 W cm^{-2}, three times for 30 sec each on ice, and centrifuged at 16 100 ***g*** for 10 min at 4°C. The His-tagged RgJF799117 proteins present in the supernatant were purified using an Ni-NTA purification system kit (Invitrogen) according to the manufacturer's instructions. After three column washes, the adsorbed proteins were eluted in 300 μl of a solution containing 500 mM imidazole, 300 mM NaCl and 50 mM NaH_2PO_4 at pH 8. The concentration of the purified proteins was determined by fluorometry using the Invitrogen Qubit system. The purified recombinant proteins were visualized by SDS–PAGE. The proteins were revealed by staining the gel using Coomassie brilliant blue R-250. The proteins were stored at −80°C for further analyses.

Enzymatic assays and HPLC analysis

The enzyme assay reaction contained 1 μg of protein, 0.1 mM Tris/HCl pH 7.2, 0.2 mM $FeSO_4$, 5 mM α-ketoglutarate, 10 mM sodium ascorbate and substrate at concentrations from 2 to 200 μM to determinate the kinetic parameters. For the inhibition assays, three concentrations of psoralen (50, 100 and 200 μM) were added to this enzymatic assay reaction. The kinetic and inhibition assays were performed in triplicate. The reaction was performed for 3 min at 30°C, stopped by addition of 20 μl of 3 mM NaOH, before neutralization with 10 μl glacial acetic acid. This treatment resulted in lysis of the CoA esters and therefore did not permit visualization of these esters; only the corresponding acids could be visualized. The enzymatic mixture was centrifuged for 10 min at 16 000 ***g***, and the supernatant was analyzed by HPLC-DAD on a Cosmosil 5C18-AR-II column (Nacalai Tesque Inc., http://www.nacalai.co.jp) on a linear gradient of 10–70% methanol for 35 min at 1 ml min^{-1}. The UV spectra and retention times of the products formed were compared to commercial standards. The products formed were also analyzed by mass spectrometry. The kinetic parameters were calculated using the SIGMAPLOT software program (Systat Software Inc., http://www.sigmaplot.com/)

Heterologous expression in *N. benthamiana*

The pBIN-GW plasmid was obtained using pBIN-mGFP5ER (Haseloff *et al.*, 1997) in which the *mGFP5ER* gene was replaced by a Multi Cloning Site linker that contained *Bam*HI, *Sma*I, *Eco*RI, *Kpn*I and *Eco*RI sites, producing the plasmid pBIN-Linker. To obtain pBIN-GW, the pBIN-Linker was digested with *Bam*HI, and an RfB cassette was introduced to generate a Gateway destination vector as described by the manufacturer (Invitrogen). The gene encoding Rg 2-OGD was first cloned into the pCR8®/GW/TOPO® TA plasmid (Invitrogen, http://www.invitrogen.com). It was then transferred into the destination plasmid pBIN-GW by LR recombination using Gateway® technology according to the manufacturer's instructions, to produce pBIN-Rg 2-OGD.

The three recombinant plasmids pBIN-Rg 2-OGD, pBIN-GW and pBIN-mGFP5ER (Haseloff *et al.*, 1997) were introduced into *Agrobacterium* strain LBA4404 to produce three *Agrobacterium* strains. In parallel, *Agrobacterium* strain C5851 containing pBIN61-P19 was generated (Voinnet *et al.*, 2003). All of these strains were individually grown in YEB culture medium (5.0 g sucrose. 5.0 g peptone. 5.0 g beef extract. 1.0 g yeast extract. 0.049 g MgSO4 . 7H20. pH to 7.2) for 24 h at 30°C. The bacterial cultures were centrifuged for 5 min at 5000 g, and the pellets were washed three times with sterile distilled water. Finally, the bacterial pellet was resuspended in water to obtain an OD_{600} between 0.3 and 0.4. To avoid gene silencing during the transient expression experiment, each LBA4404 strain (containing the various plasmids) was infiltrated together with the C5851 pBIN61-P19 strain in *N. benthamiana* leaves. Leaves transformed with the three plasmid combinations were harvested 3 days post-inoculation. The leaves of the GFP-transformed plant were used as a positive control to ascertain the agroinfection efficiency, and were observed under UV light. The leaves from

© 2011 The Authors

plants transformed with the empty pBIN-GW plasmid (used as a negative control) or pBIN-Rg 2-OGD were analyzed for their coumarin/furanocoumarin content. All the infiltrations were reproduced in triplicate.

Coumarin extraction and analysis

The leaves, petioles, stems and roots of *R. graveolens* seedlings and the leaves of *N. benthamiana* were harvested, frozen in liquid nitrogen and crushed immediately. Fresh material (200 mg) was collected for each organ. Extraction of the compounds was performed in 1.5 ml of 80% ethanol, followed by an initial grinding step in liquid nitrogen and a subsequent grinding step using a benchtop homogenizer Polytron PT2100 (Kinematica, http://www.kinematica.ch) for 1 min. After centrifugation at 16 000 ***g*** for 10 min, the supernatant was collected and concentrated by overnight ethanol evaporation in an RC10.10 speed vacuum (Jouan, http://www.thermoscientific.com). The pellet was resuspended in 200 µl methanol and analyzed by HPLC-DAD (Shimadzu, http://www.shimadzu.fr) on a Cosmosil 5C18-AR-II column (Nacalai Tesque Inc., http://www.nacalai.co.jp). The mobile phase consisted of a binary solvent system of water (solvent A) and methanol (solvent B), each containing 0.1% v/v formic acid. The gradient (0–5 min, 10% solvent B; 5–45 min, 20–60% solvent B; 45–75 min, 60–75% solvent B; 75–76 min, 75–100% solvent B) was eluted at a flow rate of 0.7 ml min^{-1}. Scopoletin was detected at 346 nm and umbelliferone was detected at 330 nm. To determine statistical significance, Student's *t* test was performed on the data using the MS Office Excel software program (Microsoft, http://www.microsoft.com/).

Real-time PCR

Total RNA was extracted from the leaves, roots, stems and petioles of *R. graveolens*. Extractions were performed on three independent plants. The extraction was performed using the RNeasy plant mini kit (Qiagen) according to the manufacturer's instructions. High-capacity RNA-to-cDNA master mix (Applied Biosystems, http://www.appliedbiosystems.com/) was used for first-strand cDNA synthesis with oligo(dT) as the primer. The resulting cDNAs were diluted 100-fold before use for real-time PCR under the conditions described below. Ribosomal 16S RNA from *R. graveolens* was used as an internal control for data normalization (forward primer, 5′-CATTCGGCCCGTCTTGAA-3′; reverse primer, 5′-CCGTTGACTCGCACACATGT-3′). Real-time PCR was performed using the primer pair qPCRD4Dir (5′-CCACCTCTCAGAGCGACTATCA-3′) and qPCRD4Rev (5′-CCCATGTGAATGACGGAGTCT-3′) designed using the Primer Express Software program (Applied Biosystems, http://www.appliedbiosystems.com). The efficiency of the probes was checked using a standard linear curve method. The relative transcript level of the same gene in the various organs was expressed as the $\log_2$ ratio of the 16S normalized transcript levels in a given organ relative to the mean expression of the gene in organs tested. Student's *t* test was performed on the data using MS Office Excel.

MS analysis

The HPLC-MS system comprised a binary solvent delivery pump and a linear ion trap mass spectrometer (LTQ-MS, Thermo Scientific, http://www.thermoscientific.com). The LTQ linear ion trap was equipped with an atmospheric pressure ionization interface and operated in ESI positive ion mode. The data were processed using the XCALIBUR software program (version 2.1, http://www.thermoscientific.com). For the mass spectrometer, the spray voltage was 5 kV, and the temperature of the heated capillary was set at 300°C. The flow rates for sheath gas, auxiliary gas and sweep gas were set to 40, 10 and 10 arbitrary units min^{-1}, respectively. The capillary voltage was 36 V, the split lens was −44 V, and the front lens was −3.25 V. All of the parameters were optimized by infusing a standard solution of isopimpinellin (0.1 g l^{-1}) in the mobile phase [water + 0.1% acetic acid/methanol + 0.1% acetic acid (ratio 90/10)] at a flow rate of 5 ml min^{-1}. The coumarins and furanocoumarins of interest were monitored by specific MS/MS scans in addition to a full scan (50–2000 *m/z*): MS^2(163) for umbelliferone and MS^2(193) for scopoletin.

Phylogenetic analysis

Multiple sequence alignment of the amino acid sequences of the 2-OGD family members was performed at the Bioinformatics Center, Institute for Chemical Research, Kyoto University, Japan (http://www.genome.jp/) using the ClustalW program (http://www.genome.jp/tools/clustalw/). *A. thaliana* AOP1 was included in the alignment as an outgroup. Based on the alignment, a molecular phylogenetic tree was constructed using the neighbor-joining method. The statistical significance of the neighbor-joining tree topology was evaluated using bootstrap analysis with 1000 iterative tree constructions. The tree was visualized using the TREE VIEW software program (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

ACKNOWLEDGEMENTS

The experiments performed in Japan by G.V. were funded by INRA/JSPS 2009–2011. The authors are grateful to Professor Jim Haselhoff (Cambridge, UK) who provided the pBIN-mGFP5ER vector, Professor David Baulcombe (Cambridge, UK) who provided pBIN61-P19, and to Professor Carl Douglas (Vancouver, Canada) who provided the pQE30/4CL plasmid. The authors would like to acknowledge the technical help of L. Calvier, A. Courteille, N. Delacroix, L. Eliasson, A. Kutrowska, E. Martins and M. Lourdin for their assistance in the Furoc2008 project (ENSAIA).

REFERENCES

Adam, M., Dobias, P., Eisner, A. and Ventura, K. (2009) Extraction of antioxidants from plants using ultrasonic methods and their antioxidant capacity. *J. Sep. Sci.* **32**(2), 288–294.

Akashi, T., Aoki, T. and Ayabe, S.I. (1998) Identification of a cytochrome P450 cDNA encoding (2*S*)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata* L.; Fabaceae) which represents licodione synthase and flavone synthase II. *FEBS Lett.* **431**(2), 287–290.

Anzelotti, D. and Ibrahim, R.K. (2004) Molecular characterization and functional expression of flavonol 6-hydroxylase. *BMC Plant Biol.* **4**, 20.

Beier, R.C. and Oertli, E.H. (1983) Psoralen and other linear furocoumarins as phytoalexins in celery. *Phytochemistry*, **22**(11), 2595–2597.

Bourgaud, F., Hehn, A., Larbat, R., Doerper, S., Gontier, E., Kellner, S. and Matern, U. (2006) Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochem. Rev.* **5**(2–3), 293–308.

Bredebach, M., Matern, U. and Martens, S. (2011) Three 2-oxoglutarate-dependent dioxygenase activities of *Equisetum arvense* L. forming flavone and flavonol from (2*S*)-naringenin. *Phytochemistry*, **72**(7), 557–563.

Brown, S.A. (1962) Biosynthesis of coumarin and hemiarin in lavender. *Science*, **137**, 977–978.

Brown, S.A., Towers, G.H. and Wright, D. (1960) Biosynthesis of the coumarins. Tracer studies on coumarin formation in *Hierochloe odorata* and *Melilotus officinalis*. *Can. J. Biochem. Physiol.* **38**, 143–156.

Chong, J., Baltz, R., Schmitt, C., Beffa, R., Fritig, B. and Saindrenan, P. (2002) Downregulation of a pathogen-responsive tobacco UDP-Glc: phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. *Plant Cell*, **14**(5), 1093–1107.

Costet, L., Fritig, B. and Kauffmann, S. (2002) Scopoletin expression in elicitor-treated and tobacco mosaic virus-infected tobacco plants. *Physiol. Plant.* **115**(2), 228–235.

© 2011 The Authors

Dall'Acqua, F., Vedaldi, D. and Recher, M. (1978) The photoreaction between furocoumarins and various DNA with different base compositions. *Photochem. Photobiol.* **27**, 33–36.

Eilert, U. (1989) Elicitor induction of secondary metabolism in dedifferentiated *in vitro* system of *Ruta graveolens*. In *Primary and Secondary Metabolism of Plant Cell Cultures* (Kurz, W.G.W., ed). Berlin: Springer, pp. 219–228.

Ellis, B.E. and Amrhein, N. (1971) The 'NIH-shift' during aromatic *ortho*-hydroxylation in higher plants. *Phytochemistry*, **10**, 3069–3072.

Endler, A., Martens, S., Wellmann, F. and Matern, U. (2008) Unusually divergent 4-coumarate: CoA-ligases from *Ruta graveolens* L. *Plant Mol. Biol.* **67**(4), 335–346.

Fliegmann, J., Furtwangler, K., Malterer, G., Cantarello, C., Schuler, G., Ebel, J. and Mithofer, A. (2010) Flavone synthase II (CYP93B16) from soybean (*Glycine max* L.). *Phytochemistry*, **71**(5–6), 508–514.

Frey, M., Huber, K., Park, W.J., Sicker, D., Lindberg, P., Meeley, R.B., Simmons, C.R., Yalpani, N. and Gierl, A. (2003) A 2-oxoglutarate-dependent dioxygenase is integrated in DIMBOA-biosynthesis. *Phytochemistry*, **62**(3), 371–376.

Gestetner, B. and Conn, E.E. (1974) 2-hydroxylation of *trans*-cinnamic acid by chloroplasts from *Melilotus alba*. *Arch. Biochem. Biophys.* **163**, 617–624.

Gutierrez, M.-C., Parry, A., Tena, M., Jorrin, J. and Edwards, R. (1995) Abiotic elicitation of coumarin phytoalexins in sunflower. *Phytochemistry*, **38**(5), 1185–1191.

Hagel, J.M. and Facchini, P.J. (2010) Dioxygenases catalyse the *O*-demethylation steps of morphine biosynthesis in opium poppy. *Nat. Chem. Biol.* **6**(4), 273–275.

Hamerski, D. and Matern, U. (1988) Elicitor-induced biosynthesis of psoralens in *Ammi majus* L. suspension cultures. Microsomal conversion of demethylsuberosin into (+)marmesin and psoralen. *Eur. J. Biochem.* **171**(1–2), 369–375.

Haseloff, J., Siemering, K.R., Prasher, D.C. and Hodge, S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. *Proc. Natl Acad. Sci. USA*, **94**(6), 2122–2127.

Kai, K., Shimizu, B.-I., Mizutani, M., Watanabe, K. and Sakata, K. (2006) Accumulation of coumarins in *Arabidopsis thaliana*. *Phytochemistry*, **67**(4), 379–386.

Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K. and Shimizu, B.I. (2008) Scopoletin is biosynthesized via *ortho*-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *Plant J.* **55**(6), 989–999.

Kùc, J. (1982) Phytoalexins from the Solanaceae. In *Phytoalexins* (Bailey, J.A. and Mansfield, J.W., eds). Glasgow: Blackie & Sons, pp. 81–105.

Larbat, R., Kellner, S., Specker, S., Hehn, A., Gontier, E., Hans, J., Bourgaud, F. and Matern, U. (2007) Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis. *J. Biol. Chem.* **282**(1), 542–554.

Lee and Douglas (1996) Two divergent members of a tobacco 4-coumarate: Coenzyme A ligase (4CL) gene family. *Plant Physiol.* **112**, 193–205.

Lukacin, R., Wellmanna, F., Britsch, L., Martens, S. and Matern, U. (2003) Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochemistry*, **62**(3), 287–292.

Martens, S., Forkmann, G., Matern, U. and Lukacin, R. (2001) Cloning of parsley flavone synthase I. *Phytochemistry*, **58**(1), 43–46.

Matsuda, J., Okabe, S., Hashimoto, T. and Yamada, Y. (1991) Molecular cloning of hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of *Hyoscyamus niger*. *J. Biol. Chem.* **266**(15), 9460–9464.

Milesi, S., Massot, B., Gontier, E., Bourgaud, F. and Guckert, A. (2001) *Ruta graveolens* L.: a promising species for the production of furocoumarins. *Plant Sci.* **161**, 189–199.

Paszkiewicz, M., Orlita, A., Dziabas, A., Golebiowski, M., Lojkowska, E., Szafranek, J., Malinski, E. and Stepnowski, P. (2008) Simplex optimized LC analysis of plant coumarins and furanocoumarins. *Chromatographia*, **67** (7–8), 653–657.

Puhl, I., Stadler, F. and Treutter, D. (2008) Alterations of flavonoid biosynthesis in young grapevine (*Vitis vinifera* L.) leaves, flowers, and berries induced by the dioxygenase inhibitor prohexadione-Ca. *J. Agric. Food. Chem.* **56**(7), 2498–2504.

Schuler, M.A. and Berenbaum, M.R. (2003) Diversification of furanocoumarin-metabolizing cytochrome P450 monooxygenases in two papilionids: specificity and substrate encounter rate. *Proc. Natl Acad. Sci. USA*, **100**(Suppl. 2), 14593–14598.

Schwab, W. (2003) Metabolome diversity: too few genes, too many metabolites? *Phytochemistry*, **62**(6), 837–849.

Steck, W., Bailey, B.K., Shyluk, J.P. and Gamborg, O.L. (1971) Coumarins and alkaloids from cell cultures of *Ruta graveolens*. *Phytochemistry*, **10**(1), 191–194.

Vazquez-Flota, F., De Carolis, E., Alarco, A.M. and De Luca, V. (1997) Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus* (L.) G. Don. *Plant Mol. Biol.* **34**(6), 935–948.

Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**(5), 949–956.

Zobel, A.M. and Brown, S.A. (1993) Influence of low-intensity ultraviolet radiation on extrusion of furanocoumarins to the leaf surface. *J. Chem. Ecol.* **19**(5), 939–952.

Zobel, A.M., Zalewski, A., Renke, K.W. and Growniak, K. (1998) Influence of UV-radiation on xanthotoxin and bergapten contents in *Ruta graveolens* L. and *Angelica arachangelica* L. seedlings. *Allelopathy J.* **5**(2), 183–189.

© 2011 The Authors

Publication soumise à BMC Plant Biology

C3'H (CYP98A22) from *Ruta graveolens:* a new tool for enhancing the furanocoumarin concentration in *Ruta graveolens.*

Fazeelat Karamat*, Alexandre Olry *, Sébastien Doerper, Guilhem Vialart, Frédéric Bourgaud, Alain Hehn[&]

UMR 1121 Université de Lorraine (INPL) - INRA Agronomie et Environnement Nancy-Colmar, ENSAIA 2 avenue de la forêt de Haye 54505 Vandœuvre-lès-Nancy, France.

* Equal co-authors.

[&] Corresponding author:

Alain Hehn
UMR1121 Université de Lorraine - INRA
Agronomie et Environnement Nancy-Colmar
ENSAIA, 2 Avenue Forêt de Haye - BP 172 - 54505 Vandœuvre-lès-Nancy
Phone: 0 (33) 3-83-59-58-63
Fax: 0 (33) 3-83-59-57-99
E-mail: alain.hehn@univ-lorraine.fr

Abstract

Background:

Furanocoumarins are molecules with proven therapeutic properties and are produced in only a small number of medicinal plant species such as *Ruta graveolens*. *In vivo*, these molecules play a protective role against phytophageous insect attack. Furanocoumarins are members of the phenylpropanoids family, and their biosynthetic pathway is initiated from *p*-coumaroyle coA. The enzymes belonging to the CYP98A cytochrome P450 family have been widely described as being aromatic meta-hydroxylases of various substrates, such as *p*-coumaroyle ester derivatives. Thus, in furanocoumarin-producing plants, these enzymes catalyze the step directly downstream of the junction with the furanocoumarin biosynthetic pathway and might indirectly impact their synthesis.

Results: In this work, we describe the cloning and functional characterization of the first CYP98A gene isolated from *R. graveolens*. Using *Nicotiana benthamiana* as a heterologous expression system, we have demonstrated that this enzyme adds a 3-OH to *p*-coumaroyle ester derivatives but is more efficient to convert *p*-coumaroyle quinate into chlorogenic acid than to metabolize *p*-coumaroyle shikimate. Plants exposed to UV-B stress showed an enhanced expression level of the corresponding gene. The *Ruta graveolens CYP98A22* open reading frame and the orthologous *Arabidopsis thaliana* CYP98A3 open reading frame were overexpressed in stable transgenic Ruta plants. Both plant series were analyzed for their production of furanocoumarins. A detailed analysis indicates that the metabolic profiles of Ruta plants overexpressing CYP98A22 are different from the plants overexpressing CYP98A3.

Conclusion: The overexpression of CYP98A22 impacts the production of furanocoumarins in *Ruta graveolens* and is therefore a good tool to generate plants that are more resistant to phytophageous insects, which are natural bioreactors for the production of these therapeutic molecules.

Introduction

The adaptation of plants to their environment and their survival under stressed conditions, e.g., pathogenic attacks, requires secondary metabolites, such as polyphenols. These molecules are widely distributed in the plant kingdom with more than 8,000 phenolic structures currently known, ranging from simple molecules, such as phenolic acids, to highly polymerized substances, such as tannins (Pan et al., 2010). The furanocoumarins constitute one of these classes of polyphenols. Despite their importance in plant life, their biosynthesis remains relatively poorly documented at the molecular level. These molecules exist in only 4 plant families: Rutaceae, Apiaceae, Fabaceae and the Moraceae. They play diverse functions in plant adaptation to the environment as phytoalexins in defense systems (Beier and Oertli, 1983) or in plant-insect interactions (Schuler and Berenbaum, 2003). These toxic molecules also display remarkable physical properties. They can interfere in enzymatic reactions through the inhibition of P450 enzymatic activities (Koenigs and Trager, 1998; Gravot et al., 2004). They also interact with nucleic acids through the photocycloaddition of pyrimidic bases (Dall'Acqua et al., 1978a). These physicochemical properties make furanocoumarins attractive candidates for therapeutic use. For example, furanocoumarin derivatives have been used for decades as treatments for skin diseases, such as psoriasis and vitiligo (Adişen et al., 2008). In addition, there are other applications for furanocoumarins in various therapeutic fields, such as the symptomatic treatment of multiple sclerosis (Pereira et al., 2007), photochemotherapy of T cell lymphoma (Plumas et al., 2003), or chemotherapy of multidrug-resistant tumors (Kawase et al., 2005). Thus, it would be beneficial to increase the production of furanocoumarins in plants to match pharmaceutical demand. Therefore, it is essential to understand the biosynthetic pathway of furanocoumarins to determine how the production of these molecules can be enhanced.

Furanocoumarin-producing plants are not model plants for the scientific community. Therefore, little is known about their genomes and the genes that encode the enzymes involved in their biosynthetic pathways. Only four genes have been described so far. Two P450s, psoralen synthase and angelicin synthase, have been described and are specifically involved in the synthesis of these molecules. These synthases catalyze the transformation of marmesin and colombianetin in psoralen and angelicin, respectively (Larbat et al., 2007; Larbat et al., 2009). Another study describes the identification and the characterization of a *o*-methyl transferase in *Ammi majus* that catalyzes the transformation of bergaptol into bergapten (Hehmann et al., 2004). Finally, an Fe^{2+}/α-cetoglutarate-dependent dioxygenase was recently identified in *Ruta graveolens*, which is able to metabolize *p*-coumaroyle coA leading to the synthesis of umbelliferone (Vialart et al., 2012).

An enzyme belonging to the cytochrome P450 enzyme family catalyzes an important step in the phenylpropanoid pathway. CYP98A (C3'H) catalyzes the *meta*-hydroxylation step of *p*-coumarate derivatives in the phenylpropanoid pathway (Figure 1). This *meta*-hydroxylation is not catalyzed on free *p*-coumaric acid, but on esterified *p*-coumate together with shikimic, quinic, tyramine or phenyllactic acids (Schoch et al., 2001a). In most of the studies, these enzymes showed preferential affinity for the 5-*O*-shikimate and 5-*O*-D-quinate esters of trans-*p*-coumaric acid (Schoch et al., 2006; Morant et al., 2007; Sullivan and Zarnowski, 2010), however 4-coumaroyle-3',4'-dihydroxyphenyllactate and *p*-coumaroyle tyramine are also precursor molecules that form rosmarinic

acid and caffeoyltyramine, respectively (Murray et al., 1989; Morant et al., 2007; Matsuno et al., 2009). Recently, another study demonstrated a complex process of evolution of the CYP98A family to acquire new functions implicated in the mechanism of pollen development in plants (Matsuno et al., 2009). Other studies have described roles for these enzymes in the plant stress response to various molecules, such as salicylic acid derivatives and isonicotinic acids (Basson and Dubery, 2007), and its expression level is highly increased after UV-C treatment (Moglia et al., 2009). Therefore, the CYP98A subfamily plays a major role in the phenylpropanoid pathway, and their action is not limited to the direct synthesis of caffeoyl ester derivatives but has a larger impact on plant metabolism.

In the present work, we describe the identification and *in vitro*/*in vivo* enzymatic characterization of a new gene encoding CYP98A22, a *p*-coumaroyle ester 3'-hydroxylase from *Ruta graveolens,* which constitutes the first CYP98 characterized from a furanocoumarin-producing plant. The biochemical characterization was assayed in 3 different systems: i) heterologous expression in yeast using the galactose-inducible strain pYeDP60/WAT11, ii) heterologous transient expression in the leaves of *N. benthamiana* together with the TBSV P19-silencing suppressor, and iii) stable expression in *R. graveolens* plants. Our results clearly show that CYP98A22 efficiently hydroxylates *p*-coumaroyle quinate rather than *p*-coumaroyle shikimate and, therefore seems to be preferentially involved in chlorogenic acid metabolism. Leaf exposure to UV-B light and further analyses of the plant metabolism and the expression level of *CYP98A22* revealed an increased furanocoumarin content and *CYP98A22* mRNA accumulation. Finally, the analyses of the phenylpropanoid extracted from transgenic *R. graveolens* overexpressing *CYP98A22* clearly showed that there is an increase in the production of furanocoumarins. The overexpression of CYP98A22 from *R. graveolens* and its ortholog CYP98A3 from *A. thaliana* led to different results in terms of metabolic profiles in stable transgenic *R. graveolens*. Therefore, the work described here demonstrates that CYP98A22 can be used as a tool to modulate the expression of furanocoumarins in *R. graveolens.*

Results

Identification of a 98A orthologous gene in Ruta graveolens

To identify genes present in furanocoumarin-producing plants, we used a PCR-based approach and the CODEHOP strategy described by Morant et al. (Morant et al., 2002). First, we focused on the identification of genes belonging to the CYP98A subfamily. To achieve this, we performed an alignment of 9 sequences of CYP98A available in databases using the CLUSTAL W freeware. This sequence comparison allowed us to identify two peptidic consensus domains (EWAMAEL and PFGAGRR) and define degenerated primers. The PCR reactions were performed on genomic DNA extracted from young *R. graveolens* seedlings. A DNA fragment of 389 nucleotides corresponding to the internal sequence of a gene encoding a cytochrome P450 was amplified and subsequently cloned. A Genbank (NCBI) homology search showed 89% identity at the amino acid level with a C3'H isolated from *Ocinum basilicum* (AAL99200.1). Cloning of the 3’ and 5’ end of the *R. graveolens* gene was achieved through PCR conducted on a SMART cDNA library using the RNA extracted from the leaves of young *R. graveolens* seedlings (Vialart et al., 2012). To perform the PCR, we used specific primers designed against the isolated sequence and specific primers (CDSIII/3’PCR at the 3’ end and 5LD

PCR at the 5' end) directed against the pTriplex plasmid for which the cDNA library had been constructed. The resulting sequence (GenBank JF799117) was 1527 bp long and encoded for a 508 amino acid protein, which displayed 81% identity with the Arabidopsis CYP98A3. A phylogenetic tree was constructed using 25 different CYP98A isoforms, which showed that CYP98A22 was clustered with CYP98A3 (Figure 1)

In vitro biochemical characterization of CYP98A22

To characterize the activity of CYP98A22, the open reading frame was cloned into the pYeDP60 plasmid and expressed in the *Saccharomyces cerevisae* WAT11 strain as described by Larbat and collaborators (Larbat et al., 2007). The P450 functional expression was determined at 450 nm using a CO differential spectrum. Unfortunately, for the yeast expressing CYP98A22, no peak could be detected with microsomes prepared from cells grown in classical or modified culture conditions (increase of the induction length from 12 to 24 h and a decrease in the induction temperature from 28 to 18℃). Despite undetectable levels of expression , we performed incubations using the conventional substrates previously described for the enzymes of this P450 family. We were unable to detect any significant activity using HPLC measurements with *p*-coumaroyle quinate, *p*-coumaroyle shikimate, *p*-coumaroyle tyramine and *p*-tricoumaroyle spermidine as substrates. Microsomes prepared from CYP98A3 expression yeast were perfectly functional and consistent with the results described elsewhere, which showed that the preparation of microsomes was efficient.

The absence of CYP98A22 activity in the yeast expression system allowed us to express the protein in an *N. benthamiana* plant expression system using the modified protocol of Voinnet *et al.* (Voinnet et al., 2003). In an initial attempt to monitor the expression of the corresponding proteins, the YFP protein was fused to the C terminus of the gene encoding CYP98A22. The recombinant coding sequence was cloned into the binary plasmid pBIN-GW (Vialart et al., 2012) under the control of the constitutive CaMV 35S promoter before being introduced into a LBA4404 *Agrobacterium tumefasciens* strain (Vialart et al., 2012). Young leaves of *N. benthamiana* were infiltrated with *A. tumefaciens* LBA4404 that was transformed with either pBIN-CYP98A22-YFP or the empty pBIN plasmid. The coinfiltration was performed with Agrobacterium transformed with pBIN61-P19. Four days postinoculation, the fluorescence signal for YFP was monitored in leaf discs using confocal microscopy. A strong fluorescence signal was observed at the endoplasmic reticulum level in the case of the coexpression of P19 and CYP98A22-YFP (Figure 3B and 3D), whereas no signal could be detected with the empty plasmid. These experiments demonstrated that *N. benthamiana* was an effective system for the efficient expression of CYP98A22.

To assess the function of CYP98A22, a second set of experiments was conducted using a recombinant pBIN plasmid containing *CYP98A22* as a single sequence (without the presence of YFP). The agrobacteria transformed with this plasmid were coinfiltrated into *N. benthamiana* leaves along with recombinant *A. tumefaciens* containing pBIN61-P19. In parallel, as a control, the inoculation of separate plants was performed with *A. tumefaciens* transformed with pBIN-CYP98A3. Four days

postinoculation, microsomes were prepared from the infiltrated *N. benthamiana* leaves as described in the materials and methods section. These microsomes were tested for enzymatic activity *in vitro* with saturated concentrations of various potential substrates, i.e., *p*-coumaroyle shikimate, *p*-coumaroyle quinate, *p-tri*coumaroyle spermidine or *p*-coumaroyle tyramine. For both proteins (CYP98A22 and CYP98A3), no metabolism could be detected in the presence of *p-tri*coumaroyle spermidine or *p*-coumaroyle tyramine but *p*-coumaroyle shikimate and *p*-coumaroyle quinate, were transformed in two products that shared similar retention time and UV spectrum than caffeoyl shikimate and caffeoyl quinate standards, respectively. The identity of the caffeoyl products was confirmed using LC/MS analyses (data not shown). The quantification of the caffeoyl derivatives for samples involving CYP98A3 showed nearly the same efficiency of metabolism of caffeoyl shikimate and caffeoyl quinate, which is consistent with the previous results of the biochemical characterization represented by Schoch *et al*. (Schoch et al., 2001a) in a yeast expression system. In contrast to CYP98A3, CYP98A22 seems to be more specific to *p*-coumaroyle quinate. There was a ten-fold increase in the production of caffeoyl quinate as compared with caffeoyl shikimate in the incubation mixture (data not shown).

In planta biochemical characterization of CYP98A22

N. benthamiana contains a pool of *p*-coumaroyle shikimate and quinate. Thus, we assumed that the *in vivo* metabolism of endogenous CYP98A22 potential substrates could be examined following transient expression in epidermal cells using LC/MS approaches. This type of a approach has 2 main advantages: the accuracy of the *in vitro* results can be demonstrated and the risk of the loss of enzymatic activity during the protein purification process is reduced. To this end, we infiltrated *N. benthamiana* leaves with recombinant *A. tumefaciens* containing pBIN:CYP98A22, pBIN:CYP98A3 or empty pBIN. All infiltrations were analyzed along with Agrobacterium containing pBIN61-P19. The LC/MS analyses of the extracts prepared four days postinfiltration showed an increase of chlorogenic acid and caffeoyl shikimate derivatives when any CYP98A was transiently overexpressed. This increase was statistically similar in the presence of CYP98A3 (Table 1). For the plants expressing CYP98A22, 50% more chlorogenic acid was detected than caffeoyl shikimate (Table 1). These results are consistent with those obtained for the *in vitro* characterization in the plant microsomes and shows that CYP98A22 is more specific for *p*-coumaroyle quinate (the precursor of chlorogenic acid).

Tissue-specific expression pattern of CYP98A22 in R. graveolens.

Total RNA was extracted from various organs of 3-month-old *R. graveolens plants* (leaves, seeds, petals, pistils, roots, and stems) and was used to establish the relative expression pattern of *CYP98A22* using a real-time PCR approach. Two sets of primers were selected for the quantification of *CYP98A22* mRNA, and the two probes gave similar results. The results showed a broad expression pattern of the mRNA (Figure 4). Indeed, mRNA was detected in all the tissues tested and was not restricted to lignified tissues. The results showed that the expression level is highest in the roots and petals and is also significantly high in the petioles and pistil. For the stems, leaves and seeds, the expression was much lower.

UV-B treatment of R. graveolens

As described by different authors (Zobel and Brown, 1993; Vialart et al., 2012), the biosynthesis of furanocoumarins in *R. graveolens* plants can be enhanced through exposure to UV-B radiation. After a 24-hour exposure to UV-B light (312 nm), the leaves, stems, petioles and roots of *Ruta graveolens* plants were collected and both total RNA and furanocoumarins were extracted. As a control, the same extractions were performed on non-UV-treated plants. The LC-MS quantification was conducted for bergapten, xanthotoxin, psoralen, marmesin and umbelliferone in the treated and nontreated plants. The results showed a 100-fold increase of umbelliferone content in the UV-treated plants as described in Vialart *et al.* (Vialart et al., 2012). Concerning other molecules, no statistically reliable difference could be demonstrated. Real-time PCR experiments were performed on the same set of material as published in Vialart et al. The expression of *CYP98A22* was significantly stimulated in the UV-treated leaves (3-fold increase), whereas no statistically significant change was observed in the petioles, stems or roots (Figure 4).

Over-expression of CYP98A22 in R. graveolens

To identify the function of CYP98A22 *in vivo* in a furanocoumarin-producing plant, transgenic *R. graveolens* were generated. The same recombinant *A. tumefaciens* strains were used for stable transformation (containing either pBIN-CYP98A22 or CYP98A3) and transient expression in *N. benthamiana*. In both cases, the gene was placed under the control of the CaMV constitutive 35S promoter. The transformation protocol was performed as described by Lièvre *et al.* (Lièvre et al., 2005). Eight 35S:CYP98A22 and nine 35S:CYP98A3 plants were generated and characterized at the molecular level using PCR directed against the *35S-CYP98A22* (the endogenous *CYP98A22* gene was not amplified) and the *CYP98A3* sequences. For each plant, the samples were divided into two halves. The first half was used to perform an RNA extraction in parallel and to confirm the expression level of the transgenes using real-time PCR (data not shown). The second half was used to extract phenolic molecules, which were further analyzed using HPLC. These analyses were compared with a set of extractions from wild-type plants. The production of the three major furanocoumarins (psoralen, bergapten and xanthotoxin) present in *R. graveolens* was analyzed. To determine the accuracy of the results, each analysis was performed twice for the CYP98A22 plants (two different extractions delayed by two months were performed for each plant). The results showed a statistically relevant increase in the furanocoumarin content for all the transformed plants (increase of 2.9 and 3.2 times the furanocoumarin content for the 35S:CYP98A22 plants for the first and second attempts, respectively, and 3.2 times for the 35S:CYP98A3 plants) (Table 2). To determine the global impact on the phenylpropanoid pathway, the production of rutin (a flavonoid), umbelliferone and scopoletin (coumarins) was also monitored for each transgenic and wild-type plant. Whereas no impact was observed for the production of scopoletin, an important decrease of umbelliferone (3.2 times) and rutin (2.3 times) could be observed for the 35S:CYP98A22 plants (Table 2). Surprisingly, a different production pattern for the phenylpropanoids was observed for the 35S:CYP98A3 plants. Indeed, a 3.6-fold increase of the umbelliferone content was observed in these plants, which is in contrast to the

expression observed in the 35S:CYP98A22 plants. Moreover, a 2-fold increase in the concentration of scopoletin in the 35S:CYP98A3 plants was observed, but the concentration of this molecule was not affected in 35S:CYP98A22 plants (Table 2). This result likely reflects different functions for both genes in the transgenic plants.

Discussion

The function of the P450 CYP98A gene subfamily has been extensively discussed in literature for over 10 years. More than 50 different genes have been identified so far and are referenced in the international P450 database (http://drnelson.uthsc.edu/cytochromep450.html). Various functions have been attributed to these proteins. Their activity is classically described as a C3'H catalyzing hydroxylation in the 3' position of *p*-coumaroyle shikimate and *p-coumaroyle* quinate (Schoch et al., 2001a; Mahesh et al., 2007; Moglia et al., 2009). However, other substrates have also been described, such as 4-*p*-coumaroyle-3'4'-dihydroxyphenyllactate (Murray et al., 1989; Kriechbaumer et al., 2008), *p*-coumaroyle tyramine (Morant et al., 2007) or spermidine esters (Matsuno et al., 2009). Among the enzymes that metabolize *p*-coumaroyle shikimate, some were demonstrated to hydroxylate *p*-coumaroyle quinate at a lower rate and efficiency (Schoch et al., 2001a), while others were simply unable to metabolize both substrates (Mahesh et al., 2007). In some cases, these enzymes have been described as having a larger effect on metabolism. For instance, Schoch and collaborators (Schoch et al., 2006) indicated that the presence of C3'H is necessary for the biosynthesis of many divergent compounds such as, lignins, UV-absorbing pigments, antioxidants, flavors, fragrances and coumarins. Together, all these data show that this enzyme plays a central (Figure 1) and pivotal role in the phenylpropanoid pathway and that it has an indirect impact on the synthesis of several molecules. In this work, we hypothesized that an orthologous gene isolated from a furanocoumarin-producing plant might have impact the synthesis of these molecules.

Little is known concerning the genome of *Ruta graveolens*, a furanocoumarin-producing plant. Only a few accessions are available in the nucleic acid or protein databases, and no information concerning a putative gene encoding a *CYP98A* orthologous gene was found. However, each P450 subfamily shares several conserved consensus domains. These domains have made it possible to use sequence-based PCR approaches for cloning and identifying plant cytochrome P450 genes (Mizutani et al., 1998). We used a PCR approach with degenerated primers (Morant et al., 2002) to isolate a gene from the *R. graveolens* genome that potentially displays C3'H activity. The translational product of the resulting open reading frame (ORF), *CYP98A22* (Genbank JF799117), was compared with 25 peptidic sequences available in the databases. The results showed that the protein was clustered with enzymes described as classical C3'H, among which the first C3'H CYP98A3 was functionally described. This first set of data seemed to indicate that this enzyme might be a conventional C3'H that potentially metabolized *p*-coumaroyle quinate and *p*-coumaroyle shikimate.

The biochemical characterization of plant cytochrome P450s using conventional approaches is difficult due to the low abundance and instability of these membrane-bound proteins. Thus, several different

heterologous systems for the expression of the P450s, including yeast, *E. coli* and baculovirus, have been described in literature. To perform the enzymatic characterization of CYP98A22, we selected the yeast expression system described by Pompon and collaborators (Pompon et al., 1996b). This system is widely used for the heterologous expression and characterization of enzymatic activities of proteins belonging to the P450 family. This system is efficient for the production of certain enzymes (Urban et al., 1994) but was limited for the production of other plant enzymes. Various improvements have been proposed in literature, including changes in the media and culture conditions (Jiang and Morgan, 2004), the modification or exchange of nucleotide sequences (partial or complete recoding of genes) (Batard et al., 2000; Morant et al., 2002) or the replacement of the membrane anchor (Batard et al., 2000) (Larbat et al., 2009). A carbon monoxide differential spectrum is generally performed to assess the expression level of P450 in such a system. To achieve the expression of CYP98A22, several of these strategies were employed. Although no peak at 450 nm could be observed with a CO spectrum for any strategy tested, we used the prepared microsomes to conduct metabolism tests using the classical substrates (*p*-coumaroyle quinate, *p*-coumaroyle shikimate) and other substrates, such as *p*-coumaroyle tyramine or *p*-tricoumaroyle spermidine. No transformation of any substrate could be detected, whereas the metabolism of *p*-coumaroyle quinate and shikimate was observed in the presence of CYP98A3 yeast microsomes.

To determine the activity of CYP98A22, we used *Nicotiana benthamiana* as an alternative heterologous expression system. The transient transformation of the plant petals with cytochrome P450s using particle bombardment as a technique for transferring a T-DNA has been described in the literature (Su and Hsu, 2003; Su and Hsu, 2010). The use of this system has also been reported for the functional characterization of a P450 belonging to the CYP71A subfamily involved in the biosynthetic pathway of alkaloid phytoalexins (Nafisi et al., 2007).

To test the efficiency of this plant heterologous expression system to allow the expression of CYP98A22, we first constructed a fusion protein comprising CYP98A22 and yellow fluorescent protein (YFP). This recombinant ORF was placed under the control of the CaMV 35S promoter and agroinfiltrated in the epidermal cells of *N. benthamiana*. Because no fluorescence could be observed, a second set of infiltrations was performed in the presence of a plasmid allowing the expression of the Tomato Bushy Stunt Virus silencing suppressor protein P19 as described by Voinnet *et al.* (Voinnet et al., 2003). In this case, the analyses of the infiltrated leaves revealed a strong fluorescence signal localized in the endoplasmic reticulum. This result is consistent with the work of Ro and collaborators who demonstrated that cinnamate 4-hydroxylase (C4H), a cytochrome P450 enzyme of the phenylpropanoid pathway, is bound in the ER membrane (Ro et al., 2001). We demonstrated that this plant system is efficient for the functional expression of CYP98A22 in the presence of the gene silencing suppressor protein.

To realize the functional characterization of the enzyme, a new set of plasmids were constructed containing the ORF of *CYP98A22* and *CYP98A3* under the control of the 35S promoter in the pBIN vector prior to agroinfiltration to *N. benthamiana* in the presence of TBSV P19. As a first attempt, we prepared microsomes from the leaves at 4 days postinfiltration, and the enzymatic tests showed that

both *p*-coumaroyle quinate and *p*-coumaroyle shikimate were hydroxylated to their caffeic counterparts by CYP98A22 and CYP98A3. However, CYP98A22 showed a much higher efficiency with *p*-coumaroyle quinate as a substrate than with *p*-coumaroyle shikimate, whereas CYP98A3 seems to preferentially metabolize *p*-coumaroyle shikimate. This first experiment revealed that the *N. benthamiana* plant system is a more efficient tool for the expression and *in vitro* characterization of CYP98A22 than the yeast system. The use of this plant heterologous system for the expression of CYP98A22 also provides a rapid method to test the *in vivo* activity of the concerned protein. In addition to a pool of *p*-coumaroyle esters of quinate and shikimate, the dominant phenylpropanoids of tobacco are chlorogenic acid and rutin (Wang et al., 2007). The presence of these molecules in plants provides an interesting tool to determine the activity of the overexpressed enzymes and to highlight the natural function of the enzyme. The analysis of the metabolic profile and the quantification of chlorogenic and caffeoyl shikimic acid derivatives were performed on leaves infiltrated with recombinant agrobacteria containing pBIN-CYP98A22 and on noninfiltrated leaves. The results showed that a statistically significant increase of the chlorogenic acid content could be observed.

To complete this study and fully characterize the function of this enzyme we performed quantitative RT-PCR to examine the tissue-specific expression pattern. Indeed, as discussed above, most of the members of the CYP98A P450 subfamily catalyze efficient 3'-hydroxylation of *p*-coumaroyle shikimate and much slower synthesis of caffeoyl quinate (Schoch et al., 2001a; Nair et al., 2002). Because the hydroxylation of *p*-coumaroyle-shikimate is a key step in the formation of the monolignol lignin precursor, this observed activity, along with the relatively high level of expression of most of the described C3'H in stem and vascular bundles, indicates that these genes play an important role in lignification (Schoch et al., 2001a; Nair et al., 2002; Sullivan and Zarnowski, 2010). The results we obtained in our investigation concerning the expression pattern of *CYP98A22* were somewhat different. Unlike other C3'H, *CYP98A22* displays a broad expression pattern. Strong expression was observed in the petals and roots, whereas a relatively weak expression was observed in the petioles, stems and pistils.

Together, these results show that even if CYP98A22 is phylogenetically related to CYP98A3, it might play a more complex role *in planta.* Indeed, caffeoyl shikimate plays an important role in lignification, while the caffeoyl quinate derivatives, such as chlorogenic acid, are described as growth regulators, disease resistance factors, antioxidants and compounds affecting the organoleptic quality of fruits (Ogunkanmi A.B. et al., 1973; Wang et al., 2007; Leiss et al., 2009). The ability of CYP98A22 to be more dedicated to the synthesis of chlorogenic acid could be evidence of its involvement in the stress response and, indirectly, in the synthesis of furanocoumarins, which may be involved in the response to insect attacks.

To explore this hypothesis, we attempted to study a relationship between the synthesis of furanocoumarins and the expression of *CYP98A22* in *R. graveolens*. Plantlets were exposed to UV-B light for 24 hours as described by Vialart (Vialart et al., 2012). Even if the analysis of the furanocoumarin content showed no increase in comparison with the nontreated plants, the UV-B elicitation stimulated the synthesis of umbelliferone, which is a precursor of this pathway. Moreover,

these analyses demonstrated that the expression level of *CYP98A22* gene was strongly increased (3-fold) in UV-treated leaves as compared with nontreated leaves. Recently, the elicitation experiments were utilized in *Cynara cardunculus* to induce the expression of *CYP98A49* using UV-C light (Moglia et al., 2009). Although this does not provide irrefutable evidence, these elements are consistent with the hypothesis that CYP98A22 is involved in responses against several stresses in *R. graveolens*. These preliminary results must now be explored in depth.

As mentioned previously, furanocoumarins are molecules of pharmaceutical interest for humans. *In planta*, these molecules have been described in plant defense against phytophageous insect attacks. The selection of plants synthesizing higher amounts of these molecules would be useful to facilitate the production of these molecules. We therefore explored the possibility of generating transgenic plants using the overexpression of enzymes involved in the stress response as a strategy. According to the results described above, CYP98A22 is a good candidate for this purpose. We compared two plant lines: transgenic plants overexpressing CYP98A22, which is the endogenous C3'H, or CYP98A3, which is the *A. thaliana* C3'H. In both cases, the genes were placed under the control of the CaMV 35S promoter ensuring the constitutive expression of the genes. The analyses indicate that in both cases, there is a significant increase of the production of furanocoumarins. However, the metabolic profiles also highlight a clear decrease of the amount of umbelliferone (a furanocoumarin precursor) and rutin (a flavonoid) in the case of CYP98A22, whereas the overproduction of CYP98A3 appears to have a broader effect on the secondary metabolism of the transgenic plants (increases of the concentrations of scopoletin, umbelliferone and rutin). Therefore, even if the enhancement of the furanocoumarin concentration is nearly the same for both P450s, CYP98A22 seems to be a better candidate for engineering the production of furanocoumarin in *R. graveolens*.

In conclusion, the results described in this work showed that members of the CYP98A family are involved in the phenylpropanoid biosynthesis pathway and indirectly impact the metabolism of *R. graveolens*. These properties can be used in a metabolic engineering strategy to enhance the production of molecules of therapeutic interest. The study of metabolic networks opens a wide field of investigation for the production of plants of interest for pharmaceutical applications.

Methods

cDNA library

The total RNA was extracted from the *Ruta graveolens* plantlets using the Plant Rneasy kit (Qiagen). The cDNA library was prepared as described by Vialart et al. (Vialart et al., 2012).

Amplification and cloning of CYP98A from R. graveolens

Design of CODEHOP primers to amplify CYP98A

Clustal W was used to perform a peptidic alignment between CYP98A1 (AF029856), CYP98A2 (AF022458), CYP98A3 (AC002409), CYP98A6 (AB017418), CYP98A8 (AC011765), CYP98A9 (AC011765), CYP98A13 (AAL99200), CYP98A19 (AY064170), and CYP98A20 (AY065995). The

specific primers were directed against the EWAMAEL CYP98A sequence and the large PFGAGRR P450 consensus domain to generate 98-11DIR: 5'-GAGTGGGCTATGGCTGARHTNRT-3' and 98-1R: 5'-CCTCCTGCCNGCNCCRAANGG-3'. The 5' ends of the primers were not degenerate, and the codon usage was based on *Citrus sinensis,* which also belongs to the *Rutaceae* family. The 3' end was completely degenerate. The PCR conditions were set as described in a previous study (Morant et al., 2002), and the starting hybridization temperature was 70℃ and decreased by 1℃ at each step.

5' and 3' end amplification

The CDS III and SMART IV primers (Smart cDNA Library Construction Kit, Clontech) designed against λTriplEx were used in association with primers designed to the internal sequence of the partial *CYP98A22* gene amplified using the CODEHOP approach. *CYP98A22 was* amplified using PCR (98FLREV: 5'-GGGGTACCTTACAAATCAGCAGCAACACGTTT-3', 98FLDIR: 5'-GGGGTACCATGGGTCTCCCACTCATCCC-3'), and a *Kpn*I restriction site was inserted at the 5' and 3' ends. Dr. Danièle Werck Reichhart (IBMP Strasbourg – France) provided the pYeDP60-CYP98A3 plasmid, and the gene encoding CYP98A3 was amplified using the primers 98A3DIRBam: 5'-GGGATCCATGTCGTGGTTTCTAATAGCG-3' and 98A3REVEco: 5'-GGAATTCTTACATATCGTAAGGCACGCG. The PCR was conducted using the following conditions: 94℃ for 5 min, 30 cycles of 94℃ for 15 sec, 50℃ for 15 sec, and 72℃ for 90 sec followed by a final extension step at 72℃ for 10 min. The PCR fragment s were introduced into the pCR8-GW-TOPO-TA vector according to the manufacturer's recommendations.

Construction of pYeDP60 recombinant plasmids

The recombinant pCR8 plasmids containing the *CYP98A22* sequence were inserted into the pYeDP60 yeast expression plasmid ((Pompon et al., 1996b)). The empty pYeDP60 and pCR8 plasmids containing *CYP98A22* were digested with *Kpn* I and were purified following agarose gel electrophoresis using the MinElute Gel Extraction Kit (Qiagen). The *CYP98A22* inserts and the pYeDP60 vector were mixed and ligated using the T4 DNA ligase from Invitrogen. The resulting clones obtained in *E. coli* cells were examined using colony PCR to determine the insertion and orientation of the coding sequence. Dr. Werck-Reichhart (IBMP – Strasbourg – France) kindly provided the pYeDP60 plasmid, which included the *CYP98A3* open reading frame.

Construction of the YFP fusion proteins and cloning of the CYP98A22 coding sequences in plant vectors

The design of the *CYP98A22* C-terminal fusion in frame with a fluorescent reporter was a two-step procedure. The first set of the PCR reactions was designed to yield *CYP98A22* with a 18-nucleotide 3' end overlapping the first 18 nucleotides of *YFP* (forward primer 98FLDIR, reverse primer 98A22YFPREV: 5'-CCTTGCTCACCATGTGGCGACCGGTACCCCCCAAATCAGCAGCAAC-3') and the *YFP* sequence with an additional sequence at the 5' end corresponding to the last 18 nucleotides from *CYP98A22* (forward primer 98A22YFPDIR: 5'-GTTGCTGCTGATTTGGGGGGTACCGGTCGCCACATGGTGAGCAAGG-3' and reverse primer

YFPREV: 5'-TTACTTGTACAGCTCGTCCATGCC-3'). In a second set of PCR reactions, the products from the previous PCR were mixed, and the amplification was performed using the CYP98A22 direct primer and the YFP reverse primer to generate the chimeric *CYP98A22-YFP* fusion gene. The recombinant sequence was inserted into the pCR8 vector (Invitrogen) according to the manufacturer's recommendations.

Subcloning in the pBIN-GW plasmid

The genes inserted in the pCR8 plasmid were transferred to the pBIN-GW (Vialart et al., 2012) plasmid using LR recombination as recommended by Invitrogen. The pBIN recombinant plasmids were introduced into the LBA4404 *A. tumefaciens* strain ((Trieu et al., 2000)), and the resulting clones were used for the transient expression of *N. benthamiana* leaves or the stable transformation of *R. graveolens.*

Yeast expression

The *Saccharomyces cerevisiae* WAT11 strain were transformed as described by Pompon (Pompon et al., 1996b) using the different pYeDP60 recombinant plasmids. The expression of the corresponding proteins and preparation of microsomes were performed as described by Larbat et al. (Larbat et al., 2007).

Transient expression in N. benthamiana.

For *N. benthamiana* transient protein expression, 4- to 6-week-old plants were transiently transfected with *A. tumefaciens* containing pBIN 35S:98A3, pBIN 35S:98A22, pBIN 35S:98A22-YFP and pBIN61-P19 in accordance with a previously published protocol (Tremblay et al.). A volume of 5 mL of overnight *A. tumefaciens* culture was pelleted, washed three times with water and resuspended in water. The leaves were coinfiltrated with Agrobacterium (at a final OD_{600} of 0.2) containing one of the pBIN-98 constructs and *Agrobacterium* (at a final OD_{600} of 0.4) containing pBIN61-P19. The infiltration was performed in the lower epidermis using a 1-ml syringe (without a needle) and gentle pressure.

Fluorescence analysis in infiltrated N. benthamiana leaves.

Fluorescence (in the form of a fusion protein) was observed 96 hours postinoculation. For the observation of the fluorescent proteins using confocal microscopy, the leaf disks were excised, mounted between slides and coverslips, and vacuum infiltrated using water. The leaf discs were observed using a laser scanning confocal microscope (Fluoview, FV10i, Olympus). The excitation and emission wavelengths were 488 nm and 505 to 545 nm for YFP, respectively. The images presented are single focal sections. The image processing was performed using ImageJ (National Institutes of Health, http://rsb.info.nih.gov/ij/) and Photoshop 6.0 (final image assembly; Adobe Systems, San Jose, CA).

Plant microsomes preparation and enzymatic test activity

Approximately 4 g of fresh leaves were harvested at 96 hours postinoculation, frozen in liquid nitrogen and ground in a mortar. The cell debris was resuspended in 0.1 M of KPi buffer (pH 7.0) containing a cocktail of protease inhibitors (Roche) and centrifuged for 30 min at 10,000 g. The supernatant was filtered through miracloth and centrifuged for 1 h at 100,000 g using a SORVALL WX80 ultracentrifuge. The resulting supernatant was carefully removed, and the pellet was resuspended in 200 µl of 0.1 M KPi buffer (pH 7.0). All the microsomal preparations were prepared at 4℃.

The enzymatic syntheses of *p*-coumaroyle quinate and *p*-coumaroyle shikimate were performed as described by Morant *et al.* (Morant et al., 2007). The other *p*-coumaroyle esters, *p*-coumaroyle tyramine and tri-p-coumaroyle spermidine, were kindly provided by Dr. Werck-Reichhart (IBMP - Strasbourg – France). The enzymatic assays were conducted as described by Larbat and collaborators (Larbat et al., 2007). Microsomes overexpressing CYP98A22 were incubated with 0.2 mM NADPH in 20 mM sodium phosphate buffer (pH 7.0) containing 100 µM *p*-coumaroyle shikimate or *p*-coumaroyle quinate and *p*-coumaroyle tyramine or tricoumaroylespermidine. The reaction was performed for 30 min at 28℃ and 0.1 M HCl were add ed to terminate the reaction. The reaction mix was centrifuged for 30 min at 10,000 g, and 50 µl of the supernatant was analyzed using reverse phase HPLC (LiChrospher 100 RP-18 column). The HPLC program was set as follows: buffer A (H_2O, 0.1% formic acid) and buffer B (100% MeOH, 0.1% formic acid); 5 min with 10% isocratic in buffer B followed by a linear gradient from 10 to 80% buffer B for 15 min and final a washing step. The absorbances of the hydroxylated products were measured at 330 nm with a diode array detector. The identity of the metabolized products caffeoyl shikimate and quinate were analyzed using an electrospray mass spectrometer in the negative ion mode as described by Matsuno et al. (Matsuno et al., 2009).

Chlorogenic acids and caffeoyl shikimic derivatives extracted from N. benthamiana and quantification

Four days postinfiltration, 6 independently infiltrated leaves were collected and crushed separately in liquid nitrogen using a mortar. A total of 0.1 g of each preparation was collected, and the quinic derivative molecules were extracted in 1 ml of 80% MeOH, vortexed vigorously for 1 min and centrifuged at 13,000 g for 20 min. The resulting supernatant was removed and evaporated to dryness. The solid extract was then resuspended in 100% MeOH prior to HPLC analysis. The chlorogenic acids were followed and identified as described by Hoffmann et al. (Hoffmann et al., 2004). The quantification of the esters was performed in comparison with taxifolin, an external standard, which was added at the beginning of the extraction procedure.

LC/MS analyses of the caffeoyl derivatives

The HPLC-MS system consisted of a binary solvent delivery pump and a linear ion trap mass spectrometer (LTQ-MS, Thermo Scientific, San Jose, CA, USA). The LTQ was equipped with an atmospheric pressure ionization interface operating in the ESI negative ion mode. The data were processed using Xcalibur software (version 2.1). The operational parameters of the mass spectrometer are given below. The spray voltage was 4.5 kV and the temperature of the heated capillary was set at 300℃. The flow rates of sheat h, auxiliary and sweep gases were set (in arbitrary

units min-1) to 40, 10, and 10, respectively. The capillary voltage was set at -36 V, the split lens was set at 44 V and the front lens was set at 3.25 V. All the parameters were optimized by infusing a standard solution of isopimpinelline (0.1 g L^{-1}) in mobile phase [water + acetic acid 0.1% / methanol + acetic acid 0.1% (90/10)] at a flow rate of 5 µl min^{-1}. The caffeoylshikimic acid isomers were monitored through MS^2(335)-specific and full (100-400 m/z) scans.

Plant Material and UV-B treatments

The *Ruta graveolens* seeds were germinated on soil. The experimental setup under glasshouse conditions was similar to that described by Vialart and collaborators (Vialart et al., 2012).

Real-Time PCR

To examine tissue-specific gene expression, the total mRNA was extracted from 100 mg of fresh tissues from the leaves, roots, stems, petioles, petals, seeds, and flowers of *Ruta graveolens* plants using the RNeasy plant-mini-kit (Qiagen). A high-capacity RNA-to-cDNA master mix (Applied Biosystems) was used for the first-strand cDNA synthesis. A total of 1 µg of RNA from each organ was reverse-transcribed using oligo$(dT)_{17}$ as a primer. The obtained cDNA samples were diluted 1:100 before use. Real-time PCR was performed using the following procedure. Specific 16S ribosomal RNA primers from *Ruta graveolens* were used as an internal control for data normalization (Rg16SDIR: 5'-CATTCGGCCCGTCTTGAA-3' forward and Rg16SRev: 5'-CCGTTGACTCGCACACATGT-3'). Two pairs of primers were designed for the quantification of *CYP98A22*. The following sequences were used for the first set of primers: 98RTPCRDIR1: 5'-CACGGAGTTGGCGAAGGA-3', 98RTPCRREV1 5'-GGTGCCTGTCAGCCAATTG-3'. The following sequences were used for the second set of primers: 98RTPCRDIR2: 5'-ACAGCAGAGTGGGCAATGG-3', 98RTPCRREV2 5'-CCTGTGCTTTGTGTTGCACTCTA-3'. The three sets of probes were designed using the Primer Express software version 3.0 (Applied Biosystems), and their efficiency was determined using the standard linear curve method. The relative transcript level of the same gene in various organs was expressed as the log2 ratio of the 16S normalized transcript levels in a given organ in comparison with the average expression of the gene in all the organs tested.

Plant stable transformation

The transformation of *Ruta graveolens* was performed as described previously (Lièvre et al., 2005). The regenerated transgenic plants were characterized at the molecular level using a PCR approach. The genomic DNA was extracted using the Plant DNeasy Kit (Qiagen). The PCR amplification was performed using specific primers. The forward primer was designed against the 35S promoter. This primer was the same for all the tested plants (35SDIR100: 5'-CACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAA-3'). To determine the presence of transgenic *CYP98A22* and *CYP98A3*, we used the same reverse primers that were used to perform the cloning steps described above.

Phylogenetic analysis

25 sequences of previously described CYP98A gene were aligned using the Clustal X software. An additional sequence of a divergent CYP was added to the alignment as an outgroup sequence. A molecular phylogenetic tree was generated from this alignment using the neighbor-joining method. The tree was visualized using the TreeView software.

Acknowledgement

The author would like to acknowledge Dr. Daniele Werck Reichhart (IBMP – Strasbourg – France) for providing *p*-coumaroyle tyramine and *p*-tricoumaroyle spermidine and Cedric Paris (LCA- ENSAIA – Nacy – France) for mass spectrometer analyses)

Author contributions

KF performed the yeast and plant transient expression and the enzymatic characterization. OA conducted the metabolomic and transcriptomic analysis and the YFP fusion construction. DS constructed the transgenic *Ruta graveolens* and conducted the corresponding analysis. VG treated the *Ruta graveolens* with UV and performed the corresponding analysis. HA conducted the plylogenetic analysis and the isolation and cloning of the gene. HA and BF conceived the study, participated in its design, and coordinated and finalized the written manuscript. All authors read and approved the final manuscript.

Legends

Table 1: Varying concentrations of chlorogenic acid and caffeoyl shikimate derivatives in infiltrated leaves transiently overexpressing CYP98A3 and CYP98A22. The results are presented as a ratio between the extraction of leaves expressing the P450 and the extraction of leaves transformed with an empty plasmid. *: Statistically tendency, **: Statistically relevant, ***: Statistically highly relevant.

Table 2: Metabolic modifications in *35S:CYP98A3* and *35S:CYP98A22* transgenic *Ruta graveolens* (x: increase factor, /: decrease factor)

Figure 1: The simplified phenylpropanoid biosynthesis pathway (PAL: phenyl ammonia lyase, C4H: cinnamate 4-hydroxylase, 4CL: 4-coumaroyle CoA ligase, HCT: hydroxycinnamoyle ester transferase, C3'H: cinnamoyle ester 3'-hydroxylase). Black dashed arrows: multistep reaction. Grey dashed arrows: putative indirect interaction

Figure 2: Phylogenic analysis of 25 CYP98A sequences. CYP98A1 (AF029856); CYP98A2 (AF022458); CYP98A3 (AC002409); CYP98A6 (AB017418); CYP98A7 (EU967060); CYP98A8 (AC011765); CYP98A9 (AC011765); CYP98A10 (AJ439883); CYP98A11 (AJ583531); CYP98A12 (AJ439885); CYP98A13 (AAL99200); CYP98A14 (AJ427452); CYP98A16 (AC093612); CYP98A19 (AY064170); CYP98A20 (AY065995); CYP98A21 (AY532371.2); CYP98A22 (JF799117); CYP98A29 (EU953568); CYP98A34 (BJ977736.1); CYP98A35 (DQ269126); CYP98A36 (DQ269127); CYP98A37 (DQ335791); CYP98A39 (AJ585988); CYP98A40 (AJ585989); CYP98A46 (AB374410); CYP98A49 (FJ225121); CYP73A1 (Z17369). The alignment was conducted using the Clustal X software. The phylogenetic tree was generated using the neighbor-joining method. The tree was visualized using the TreeView software.

Figure 3: Confocal microscopy images showing subcellular CYP98A22-YFP localization in *Nicotiana benthamiana* leaves coinfiltrated with *Agrobacterium tumefaciens* containing pBIN61-P19 and A and C) the empty pBIN-GW plasmid or B and D) pBIN-CYP98A22-YFP. The images were recorded at 3 days postinfiltration.

Figure 4: Expression level of *CYP98A22* in *Ruta graveolens*. Vertical bar chart showing the results of real-time RT-PCR experiments A) in different organs of the plants B) with or without UV-B stress conditions. I indicates UV-B induced. NI indicates UV-B noninduced. (The data $2^{-\Delta\Delta Ct}$ were deduced using the Ct method; the error bars represent the standard errors of 3 replicates).

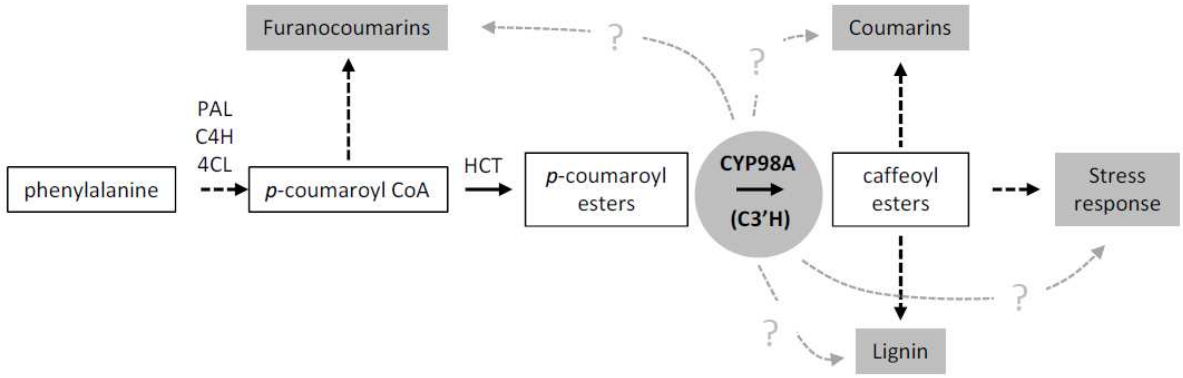


Figure 1

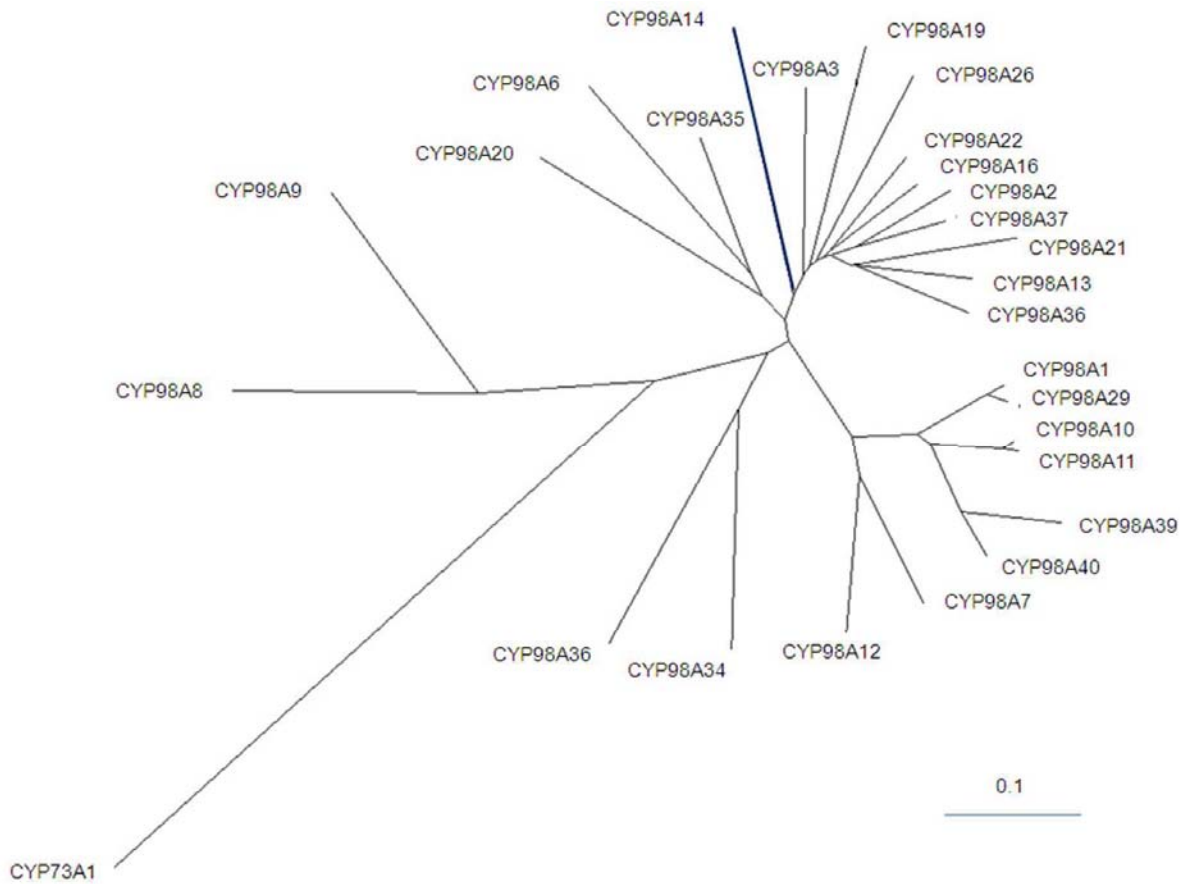


Figure 2

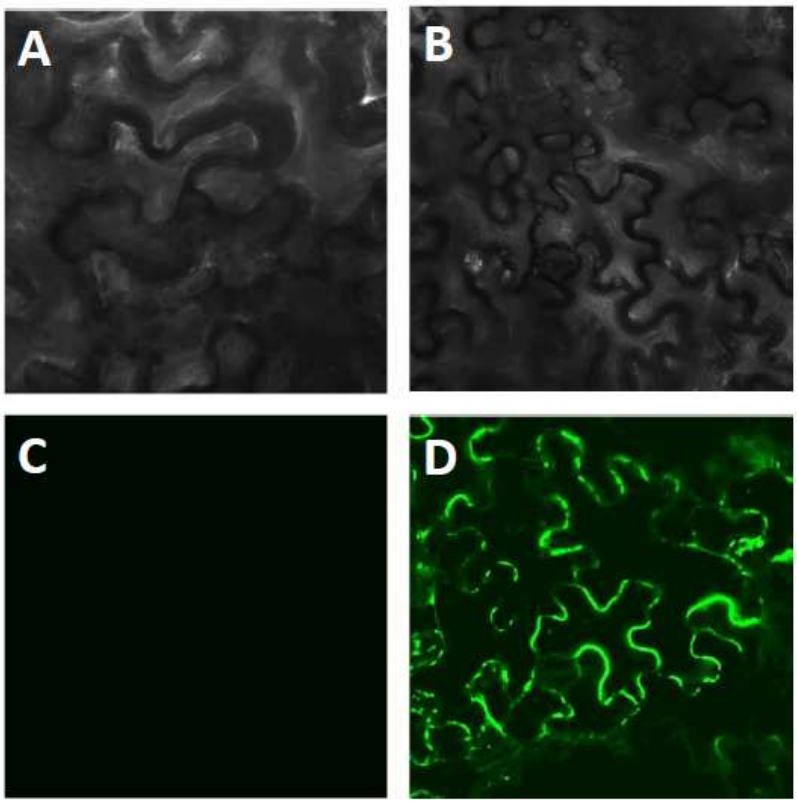


Figure 3

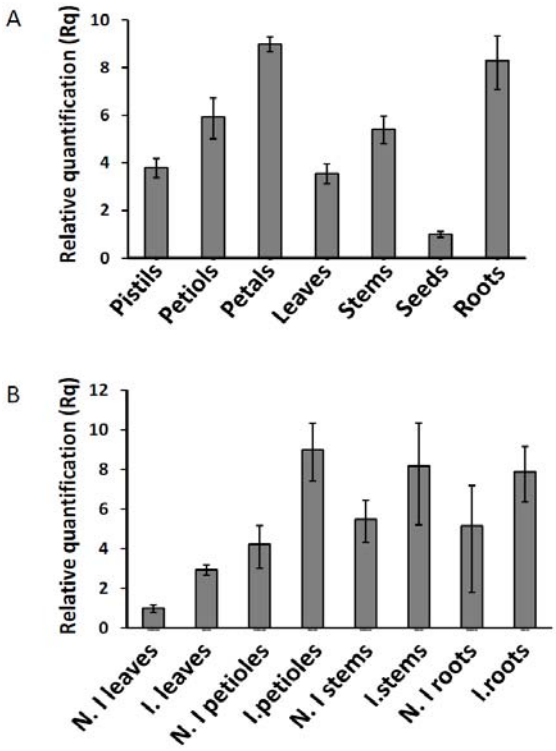


Figure 4

	Chlorogenic acid derivatives		*Caffeoyl shikimate derivatives*	
CYP98A3/void plasmid	**x1.2**	*****	**x1.2**	******
CYP98A22/void plasmid	**x1.5**	*******	**x1**	**Not significant**

Tableau 1

	35S::CYP98A3	*35::CYP98A22*
Furanocoumarins (Psoralen + Xanthotoxin + Bergapten)	**x 3.2**	**x 2.9 – 3.2**
Scopoletin	**x 2.0**	**x 1.0**
Umbelifferone	**x 3.6**	**/ 3.2**
Rutine	**x 1.0**	**/ 2.3**

Tableau 2

Listes de publications

Articles dans des revues internationales à comité de lecture

Vialart G., Hehn A., Olry A., Ito K., Krieger C., Larbat R., Paris C., Shimizu B., Sugimoto Y., Mizutani M and Bourgaud F. (2012). *A 2-oxoglutarate-dependent dioxygenase from Ruta graveolens L. exhibits p-coumaroyle CoA 2'-hydroxylase activity (C2'H): a missing step in the synthesis of umbelliferone in plants*. "The plant Journal" Online depuis le 16 janvier 2012

Karamat F., Olry A., Doerper S., Vialart G., Bourgaud F., Hehn A. (2012). *C3'H (CYP98A22) from Ruta graveolens: a new tool for enhancing the furanocoumarin concentration in Ruta graveolens*. (Soumis dans BMC Plant Biology)

Communications internationales

Présentation orale

Vialart G., Hehn A., Olry A., Ito K., Paris C., Mizutani M. and Bourgaud F (2011*). A 2-oxoglutarate-dependent dioxygenase involved in the biosynthesis of umbelliferone, an intermediate of the furanocoumarin pathway.* Présentation effectuée au "17th International Conference on Cytochrome P450 Biochemistry, Biophysics and Structure" du 26 au 30 Juin 2011, à Manchester, UK.

Par voie d'affiche

Vialart G., Hehn A., Olry A., Ito K., Paris C., Mizutani M. and Bourgaud F (2011). *A 2-oxoglutarate-dependent dioxygenase involved in the biosynthesis of umbelliferone, an intermediate of the furanocoumarin pathway.* "17th International Conference on Cytochrome P450 Biochemistry, Biophysics and Structure" du 26 au 30 Juin 2011, à Manchester, UK

Olry, A., Karamat F., Doerper S., Vialart G., Hehn A., Bourgaud F. *CYP98A22, a cinnamoyle 3'-hydroxylase (C3'H) as a tool for engineering the furanocoumarin pathway in Ruta graveolens.* "17th International Conference on Cytochrome P450 Biochemistry, Biophysics and Structure" du 26 au 30 Juin 2011, à Manchester, UK

Communications nationales

Par voie d'affiche

Vialart G., Hehn A., Mizutani M., Bourgaud F. (2010). *A la recherche de gènes de dioxygénases impliqués dans la biosynthèse de furocoumarines chez Ruta graveolens.* Ecole doctorale RP2E, jeudi 28 janvier 2010, Nancy.

Identification et caractérisation fonctionnelle de gènes impliqués dans la voie de biosynthèse des furocoumarines chez les végétaux supérieurs

Les furocoumarines sont des métabolites secondaires qui dérivent de la voie de biosynthèse des phénylpropanoïdes. Ces phytoalexines interviennent notamment dans les mécanismes de défense des plantes tels que la résistance aux bioagresseurs. Le déterminisme moléculaire de cette voie de biosynthèse est encore mal connu mais il a néanmoins été démontré depuis les années 1960 que les enzymes catalysant les différentes étapes de la voie la synthèse des furocoumarines appartenaient à différentes familles. Les travaux présentés dans ce document se sont focalisés sur deux familles d'enzymes : les dioxygenases oxoglutarate dépendantes et les cytochromes P450s.
La première étape de la voie de biosynthèse des furocoumarines consiste en une *ortho*-hydroxylation du *p*-coumarate qui mène à la formation de l'umbelliférone. Basé sur des travaux récents sur *Arabidopsis*, six gènes codants pour des dioxygénases ont été isolés chez *Ruta graveolens*, *Citrus limetta* et *Pastinaca sativa.* Les protéines correspondantes présentent plus de 58% d'identité avec la Féruloyle 6' Hydroxylase (F6'H) d'*A. thaliana*. La caractérisation fonctionnelle de ces enzymes a été réalisée dans un système d'expression hétérologue procaryote. Sur les 6 enzymes, trois n'ont pu être exprimée efficacement, et deux présentent une activité F6'H similaire à celle décrite pour *A. thaliana*. La dernière enzyme dispose de caractéristiques nouvelles non décrites à ce jour. Elle est en mesure de réaliser l'hydroxylation du féruloyle coA et du *p*-coumaroyle coA. Ces études *in vitro* ont été complétées par une exploration des fonctions de la protéine dans la plante. Une analyse fine du profil d'expression du gène a permis de mettre en évidence une expression qui est en corrélation avec le niveau de production d'umbelliférone. La fonction de la protéine a également été validée par une analyse des produits formés dans des feuilles de *Nicotiana benthamiana* transformées transitoirement.
Les cytochromes P450 catalysent 60% des réactions de la voie de biosynthèse de furocoumarines. Un travail de criblage fonctionnel de cytochromes P450 identifiés au préalable chez *Ammi majus* et *Thapsia garganica* a été entrepris. Les analyses de bio-informatiques et les modifications apportées au niveau du mode opératoire pour l'expression dans la levure ont permis d'émettre des hypothèses concernant le rôle de certains P450 candidats. Ces travaux exploratoires et préliminaires font supposer de nouvelles conjectures relatives à cette voie de biosynthèse.

Mots clés : Dioxygénase Fe/α cétoglutarate-dépendante, cytochrome P450, *p*-coumaroyle CoA, umbelliférone, furocoumarines, expression transitoire, *Ruta graveolens.*

Identification and functional characterization of genes involved in furocoumarines biosynthesis by higher plants.

Furanocoumarins are secondary metabolites deriving from the phenylpropanoid biosynthetic pathway. These phytoalexins are especially involved in plant defense mechanisms against insects or phytopatogenous fungi and bacteria. The molecular control of this biosynthetic pathway is still poorly understood even though it has been demonstrated since the 1960s that enzymes catalyzing the different steps are belonging to different enzymatic families. The work presented here is focused on two enzyme families: oxoglutarate dependent dioxygenases and cytochrome P450s.
The first step in the furanocoumarins biosynthetic pathway is the *ortho*-hydroxylation of *p*-coumarate, which leads to the formation of umbelliferone. Based on a recent work done on *Arabidopsis*, we isolated six genes encoding dioxygenases from *Ruta graveolens*, *Citrus limetta* and *Pastinaca sativa*. The corresponding proteins share more than 58% identity with the *A. thaliana* féruloyle 6 'Hydroxylase (F6'H). The functional characterization of these enzymes was performed in a prokaryotic heterologous expression system. Of the six enzymes, three could not be functionnaly expressed and two exhibited a similar F6'H activity as described for *A. thaliana*. The last enzyme has new properties not described to date. It is able to achieve both hydroxylation of féruloyle CoA and *p*-coumaroyle CoA. These *in vitro* studies were completed by a functional exploration of the protein *in planta*. A detailed analysis of the gene expression pattern highlighted a link with the level of umbelliferone synthesis. The function of the protein was also confirmed by an analysis of the products formed in transiently transformed *Nicotiana benthamiana* leaves.
Cytochrome P450s catalyze 60% of the reactions of the furanocoumarin biosynthetic pathway. Therefore, a functional screening of cytochrome P450 previously identified in *Ammi majus* and *Thapsia garganica* was undertaken. The bioinformatic analyses and the changes undertaken in the procedure for expression in yeast allowed drawing hypotheses on the function of some of these P450 candidates. These exploratory and preliminary experiments allowed suggesting new hypotheses about the biosynthetic pathway.

Key words : Fe/α-oxoglutarate-dependent dioxygenase, cytochrome P450, p-coumaroyle CoA, umbelliferone, furanocoumarins, transient expression, *Ruta graveolens.*