Université de Strasbourg

Etude des interactions protéine-protéine et protéine-membrane permettant la canalisation des flux métaboliques dans la voie des phénylpropanoïdes chez *Arabidopsis thaliana*, mettant en jeux des cytochromes P450 des familles CYP73As et CYP98As

Study of protein-protein and protein-membrane interactions leading to the channeling of metabolic fluxes in phenylpropanoid metabolism in *Arabidopsis thaliana*, involving cytochromes P450 from CYP73A and CYP98A families

Analyse fonctionnelle des CYP73As et CYP98As chez *Nicotiana tabacum*

Functional analysis of CYP98As and CYP73As paralogues in *Nicotiana tabacum*

Thèse Présentée à la FACULTE DES SCIENCES DE LA VIE

Pour obtenir le titre de **DOCTEUR DE L'UNIVERSITE de STRASBOURG**

Discipline : **Biologie Cellulaire et Moléculaire**

Présentée et soutenue publiquement le **13 Décembre 2010**

Par **Jean-Etienne BASSARD**

Devant la commision d'Examen composée de :

Pr. Birger Lindberg MØLLER	University of Copenhagen	Rapporteur externe
Pr. Alain PUGIN	Université de Dijon - INRA	Rapporteur externe
Pr. Catherine LAPIERRE	AgroParistech - INRA	Examinateur externe
Pr. Léon OTTEN	Université de Strasbourg - IBMP	Examinateur interne
Dr. Christophe RITZENTHALER	IBMP-CNRS	Examinateur interne
Dr. Danièle WERCK	IBMP-CNRS	Directeur de thèse

Thèse préparée au sein de l'UPR2357, Institut de Biologie Moléculaire des Plantes du CNRS

(IBMP, Strasbourg)

Université de Strasbourg

Etude des interactions protéine-protéine et protéine-membrane permettant la canalisation des flux métaboliques dans la voie des phénylpropanoïdes chez *Arabidopsis thaliana*, mettant en jeux des cytochromes P450 des familles CYP73As et CYP98As

Study of protein-protein and protein-membrane interactions leading to the channeling of metabolic fluxes in phenylpropanoid metabolism in *Arabidopsis thaliana*, involving cytochromes P450 from CYP73A and CYP98A families

Analyse fonctionnelle des CYP73As et CYP98As chez *Nicotiana tabacum*

Functional analysis of CYP98As and CYP73As paralogues in *Nicotiana tabacum*

Thèse Présentée à la FACULTE DES SCIENCES DE LA VIE

Pour obtenir le titre de **DOCTEUR DE L'UNIVERSITE de STRASBOURG**

Discipline : **Biologie Cellulaire et Moléculaire**

Présentée et soutenue publiquement le **13 Décembre 2010**

Par **Jean-Etienne BASSARD**

Devant la commision d'Examen composée de :

Pr. Birger Lindberg MØLLER	University of Copenhagen	Rapporteur externe
Pr. Alain PUGIN	Université de Dijon - INRA	Rapporteur externe
Pr. Catherine LAPIERRE	AgroParistech - INRA	Examinateur externe
Pr. Léon OTTEN	Université de Strasbourg - IBMP	Examinateur interne
Dr. Christophe RITZENTHALER	IBMP-CNRS	Examinateur interne
Dr. Danièle WERCK	IBMP-CNRS	Directeur de thèse

Thèse préparée au sein de l'UPR2357, Institut de Biologie Moléculaire des Plantes du CNRS

(IBMP, Strasbourg)

A toi…

Remerciements

Je tiens à remercier tous ceux qui, de près comme de loin, m'ont aidé à réaliser cette étape, par leur confiance, leur soutien, leur sourire, leur gentillesse…

Je tiens tout d'abord à remercier Danièle Werck-Reichhart, pour m'avoir accepté en stage il y a déjà quelques années… Merci, pour la confiance qu'elle m'a accordée et qu'elle m'a renouvelé de nombreuses fois.

Merci à l'ensemble des membres du jury, Pr Birger Lindberg Møller, Pr Alain Pugin, Pr Catherine Lapierre, Pr Léon Otten, et Dr Christophe Ritzenthaler, pour avoir accepté de juger mon travail et d'avoir pris de leur temps dans cette tâche.

Merci Agnès pour son soutien, son écoute, et la préparation des microsomes de tabac, qui ont été si souvent récalcitrant. Merci à Pascaline pour sa gentillesse, sa disponibilité. Merci à Franck pour les discussions qui sortaient des sujets du labo durant les repas du midi. Merci à tous les membres de l'équipe P450, les anciens membres m'ayant côtoyé comme les membres actuels : J-F, Vincent, Michiyo, Alex, Sébastien, et bien d'autres, (la liste commence à être longue, je ne vous cite pas tous, mais je pense à vous…)

Merci à Dimitri, qui a croulé sous le grand nombre d'analyses que je lui ai demandé… mais qui a toujours été disponible et souriant.

Merci à Jérome Mutterer, sans qui, mon sujet principal de thèse ne se serait pas ce qu'il est, grâce à son idée « simple » (mais judicieuse) de mesurer la re-localisation des protéines solubles sans chercher à modéliser l'ensemble du RE.

Merci Marta pour le très grand nombre de bentha semés et repiqués, ce fut l'usine pendant 2 ans. (Merci d'ailleurs aux benthas, qui ont subi un vrai génocide dans leurs rangs).

Merci à Claire et Rodolphe qui m'ont changé de ma routine quotidienne et donné un petit air de fraicheur durant les quelques mois de stage qu'ils ont passé avec moi. En espérant que j'ai pu vous servir un peu à quelque chose…

Merci à l'ensemble des membres du 5ième et du 4ème de Bota que je n'ai pas cité précédemment. Vous avez toujours été disponible lorsque je suis venu vous voir pour une chose ou un autre ; et je vous en suis reconnaissant.

Merci à l'équipe du Laboratoire d'Innovation Thérapeutique de la Faculté de Pharmacie, pour m'avoir si gentiment accueilli, en particulier Jean-Jacques Bourguignon et Martine Schmitt.

Merci à Ludovic Richert, pour m'avoir permis d'utiliser un appareillage de FLIM clef en main, et merci également pour son écoute et ses réponses à mes questions au sujet du FLIM et d'autres techniques de pointe en microscopie.

Un grand Merci à ma famille, mes ami(e)s (la meute et la triade), pour avoir toujours cru en moi, et m'avoir soutenu toutes ces années…

Merci.

Table of Contents

List of figures

Abbreviations

4CL-1	4-Coumarate: CoenzymeA Ligase isoform 1
ABA	Abscissic acid
ADC	Arginine decarboxylase
ATR1	*Arabidopsis thaliana* Reductase 1
C3'H	Coumarate 3'-Hydroxylase
C4H	Cinnamate 4-Hydroxylase
cDNA	complementary DNA
CGA	Chlorogenic Acid
CO	Carbone Monoxide
CoA	Coenzyme A
CPR	Cytochrome P450 Reductase
CQ	*p*-coumaroyl quinate
CS	*p*-coumaroyl shikimate
CYP	cytochrome P450
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
eGFP / GFP	enhanced Green Fluorescent Protein
EDTA	Ethylene diamine tetraacetic acid
EMS	Ethane Methyl Sulfonate
ER	Endoplasmic Reticulum
EST	Expressed Sequence Tag
F5H	Ferulate 5-Hydroxylase
FLIM	Fluorescence Lifetime Imaging Microscopy
FRAP	Fluorescence Recovery After Photobleaching
FRET	Fluorescence/Föster Resonance Energy Transfer
ER-FWHM	ER- Full Width at Half Maximum
G unit	Guaiacyl unit
H unit	*p*-Hydroxyphenyl unit
HCT	HydroxycinnamoylCoA: shikimate/quinate hydroxycinnamoyl Transferase
HQT	HydroxycinnamoylCoA: quinate hydroxycinnamoyl Transferase
HPLC	High Pressure Liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LC-MS	Liquid Chromatography coupled to Mass Spectrometry
MeJA	Methyl Jasmonate
mRFP / RFP	monomeric Red Fluorescent protein
MS	Mass Spectrometry
MSP	Membrane Scaffold Protein

NADPH	Nicotinamide Adenine Dinucleotide Phosphate
OD	Optical Density
ODC	Ornithine decarboxylase
OMT	*O*-methyl transferase
P450	Cytochrome P450
PAL	L-Phenylalanine Ammoniac Lyase
PCR	Polymerase Chain Reaction
PDE	Proline Dehydrogenase
PTGS	Post-Transcriptional Gene Silencing
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
RT-qPCR	Real Time quantitative PCR
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SIR	Single Ion Recording
SA	Salicylic Acid
SHT	Spermidine Hydroxycinnamoyl Transferase
SRS	Substrate Recognition Site
SPDS	Spermidine synthase
SPR	Surface Plasmon Resonance
sqPCR	semi-quantitative PCR
S unit	Syringyl unit
T-DNA	transfer Deoxyribonucleic Acid
TMV	Tobacco Mosaic Virus
tRNAs	Transfer Ribonucleic Acid
UPLC	Ultra Performance/Pressure Liquid Chromatography
USER	Uracil-Specific Excision Reagent
VIGS	Virus Induced Gene Silencing

Standard amino acid abbreviations

One-letter code Three-letter code, Amino acid			
A Ala, Alanine	**E** Glu, Glutamic acid	**L** Leu, Leucine	**S** Ser, Serine
R Arg, Arginine	**Q** Gln, Glutamine	**K** Lys, Lysine	**T** Thr, Threonine
N Asn, Asparagine	**G** Gly, Glycine	**M** Met, Methionine	**W** Trp, Tryptophan
D Asp, Aspartic acid	**H** His, Histidine	**F** Phe, Phenylalanine	**Y** Tyr, Tyrosine
C Cys, Cysteine	**I** Ile, Isoleucine	**P** Pro, Proline	**V** Val, Valine

Résumé étendu en français

1. Introduction

Les plantes étant des organismes sessiles, pour s'adapter à leur environnement elles ont évolué afin de synthétiser une très grande variété de molécules dites du métabolisme secondaire. Ces molécules leur permettent de lutter contre les stress abiotiques et biotiques, mais aussi de se reproduire, de communiquer entre elles ou avec d'autres organismes.

La diversité en métabolites secondaires produits résulte d'une cascade de réactions conduites par diverses familles d'enzymes telles que des ligases, des réductases, des oxygénases ou des transférases. Par ces processus, les plantes produisent plus de 200 000 métabolites secondaires qui d'après leurs voies de biosynthèse et leurs structures, peuvent être classés en trois principales familles : les terpénoïdes, les phénylpropanoïdes et les alcaloïdes.

La famille de molécules qualifiée de phénylpropanoïdes dérive d'une structure de base en C6-C3 formée à partir de la phénylalanine. Ces molécules forment la famille de molécules la plus abondante et représentent environ 40 % du carbone organique de la biosphère. Les phénylpropanoïdes ont joué un rôle prépondérant dans l'adaptation des plantes au milieu terrestre. Ils ont un large panel de fonctions qui sont essentielles à la structure et au maintien de la plante, mais également, à sa défense, son développement et à sa reproduction. La voie des phénylpropanoïdes canalise jusqu'à 30 % du carbone fixé par les plantes vers la synthèse de la lignine, de la subérine, et d'une grande diversité de pigments, d'arômes, de molécules de défense et d'antioxidants. Les composés ainsi obtenus présentent le plus souvent un fort intérêt économique comme la lignine par exemple.

L'apparition de la lignine a permis à la plante de rigidifier ses tissus, mais a aussi permis la circulation de l'eau des racines jusqu'aux feuilles à la cime des arbres. Ainsi, la lignine est l'un des biopolymères les plus essentiels à l'évolution des plantes au milieu terrestre. Ce biopolymère résistant, est constitué d'un réseau complexe de composés aromatiques, propre aux plantes vasculaires (ptéridophytes, gymnospermes et angiospermes). La lignine est principalement déposée au niveau des parois pecto-cellulosiques des cellules ayant des rôles de soutien mécanique ou de conductance de l'eau.

La lignine a un intérêt économique important pour l'Homme en particulier pour l'industrie papetière, l'industrie forestière, l'industrie des biocarburants et l'agriculture. En effet, la lignine est le composé ayant le plus fort pouvoir calorifique et est donc recherchée dans le bois de chauffage. Ce polymère permet également la portée mécanique du bois. Cependant, la présence de lignine est problématique lors de l'utilisation de la biomasse végétale. La lignine n'est pas digérée par les ruminants : diminuer son taux dans les fourrages permettrait d'accroître l'efficacité énergétique de la digestion. De plus, le papier est fabriqué principalement à partir de fibres cellulosiques, la lignine doit ainsi être éliminée de la matière végétale.

Ainsi une meilleure compréhension de la voie de biosynthèse des phénylpropanoïdes, conduisant entre autres à la formation de la lignine est un préalable pour envisager de réaliser de l'ingénierie métabolique pour la synthèse de molécules d'intérêt. Dans ce sens, depuis 80 ans, des modifications de la voie de biosynthèse conduisant à la formation de la lignine ont été essayées et étudiées.

2. L'étude des interactions protéine-protéine

Le nombre d'enzymes codées par les génomes de plantes (plusieurs centaines) ne permet pas d'expliquer la diversité des molécules qui y sont synthétisées (de l'ordre de plusieurs milliers). L'hypothèse étudiée est que cette diversité résulte de la formation combinatoire de complexes multienzymatiques (métabolons) qui permettent une canalisation des flux métaboliques vers diverses branches du métabolisme et divers compartiments cellulaires.

Le segment en amont de la voie des phénylpropanoïdes chez *Arabidopsis thaliana* est notre modèle au laboratoire pour développer de nouvelles approches permettant l'analyse des interactions protéine-protéine qui gouvernent l'allocation du carbone vers la synthèse des différents produits de cette voie de biosynthèse. Ce segment en amont de la voie des phénylpropanoïdes est essentiel à cette voie et il fait intervenir une séquence d'enzymes qui sont pour la plupart solubles, ainsi que deux monooxygénases qui sont des cytochromes P450 membranaires (CYP73A5 et CYP98A3 chez *Arabidopsis thaliana)*. La mobilité restreinte de ces enzymes membranaires a conduit à l'hypothèse selon laquelle elles serviraient d'ancrage à des partenaires solubles pour la formation de métabolons. Divers indices qui confortent une telle hypothèse ont été obtenus durant ces quarantes dernières années par des expériences d'incorporation de précurseur radiomarqués, de co-immunoprécipitation, co-isolation, colocalisation, par exemple.

Les objectifs du projet étaient de tester le potentiel des nouvelles approches utilisées, de prouver que des interactions protéine-protéine se produisent entre les quatre protéines et plus particulièrement entre 4CL-1 et CYP73A5, et entre HCT et CYP98A3 dans un premier temps. L'objectif final était de prouver l'existence d'un métabolon constitué de ces quatre enzymes précédemment citées.

Cette thèse fait partie d'un projet mis en œuvre avec le soutien de l'organisme « Human Frontier of Science Program », celui-ci est basé sur un partenariat, à la fois multinational et multidisciplinaire.

Les interactions protéine-protéine ont été analysées par des approches complémentaires *in vitro* et *in vivo*.

Tout d'abord des approches *in vitro*. Pour cela, j'ai été chargé de produire et purifier les cytochromes P450 (CYP73A5 et CYP98A3) et leurs partenaires solubles potentiels la 4-coumaroyl CoA ligase isoforme 1 (4CL-1) et l'hydroxycinnamoylCoA : shikimate/quinate hydroxycinnamoyl transférase (HCT). L'expression des cytochromes P450 a été réalisée dans la levure *Saccharomyces cerevisiae*, qui offre un contexte membranaire adapté à l'expression des P450s. Les partenaires solubles, ne nécessitant pas un environnement membranaire particulier, ont été exprimés dans la bactérie *Escherichia coli*. Une fois produites les différentes protéines ont été purifiées par diverses techniques de purification par affinité. Ces outils biologiques ont ensuite été utilisés par nos partenaires.

Nos collaborateurs américains ont incorporé les P450s dans des nanostructures appelées « nanodisc ». Cette technologie originale permet l'ancrage des cytochromes P450 membranaires au sein de nanodisques lipidiques de manière à ce qu'ils restent fonctionnels pour des tests d'interaction. Ensuite, nos partenaires danois ont utilisé ces nanodisques pour piéger les partenaires solubles formant les métabolons et tester les interactions entre P450s et protéines solubles, après greffage des nanodisques sur des puces de Biacore ou par co-immunoprécipitation.

Parallèlement à cette étude *in vitro*, les interactions protéine-protéine sont confirmées dans la plante par expression transitoire des gènes d'intérêt dans des feuilles de *Nicotiana benthamiana* avec des constructions permettant l'expression des protéines en fusion avec des protéines fluorescentes. Les interactions protéine-protéine sont ensuite analysées par microscopie confocale. J'ai tout d'abord validé la localisation subcellulaire des différentes enzymes exprimées individuellement. Puis j'ai entrepris des expériences de colocalisation entre les différents interactants supposés. J'ai également cherché à observer la relocalisation des protéines solubles, en présence des P450s appropriés et sous l'action de différentes combinaisons de substrats de ces enzymes. J'ai ensuite cherché à quantifier ces observations. J'ai enfin entrepris des expériences de FLIM (Fluorescence Lifetime Imaging Microscopy) pour démontrer les interactions protéine-protéine.

Dans un premier temps, pour l'analyse *in vitro* des interactions protéine-protéine, j'ai optimisé l'expression et la purification des différentes enzymes recombinantes afin d'avoir des quantités, qualités et concentrations suffisantes pour satisfaire aux exigences 1) d'insertion efficace des P450s dans les nanodisques, et 2) de détection des interactions par Biacore ou co-immunoprécipitation. J'ai pu produire CYP73A5 et CYP98A3 à des concentrations en P450 de l'ordre de 18 nmole.L^{-1} et de 51 nmole.L^{-1} de culture, respectivement. Après optimisation de la purification, j'ai pu obtenir des cytochromes P450 fonctionnels à des concentrations de l'ordre de 1,5 µmole.mL^{-1} pour CYP98A3 et de 1 µmole.mL^{-1} pour CYP73A5. La quantité de P450s pures et fonctionelles après purification a été doublée après optimisation du protocole de préparation. **La pureté des préparations et les quantités de protéines purifiées étaient suffisantes pour l'incorporation de ces P450s dans les nanostructures.**

En ce qui concerne les partenaires solubles (4CL-1 et HCT), le paramètre qui apparut le plus important et le plus limitant est la souche de bactérie utilisée pour l'expression des protéines. Après optimisation de l'expression, j'ai pu obtenir des concentrations en protéine de l'ordre de 12-13 mg.L^{-1} de culture pour 4CL-1 et de moins de 1 mg.L^{-1} pour HCT. Après purification, j'ai obtenu des concentrations de l'ordre de 5 mg.mL^{-1} pour 4CL-1 et 2 mg.mL^{-1} pour HCT. J'ai amélioré d'un facteur dix la quantité de 4CL-1 produite et j'ai multiplié par deux la concentration en 4CL-1 purifiée. Pour HCT, je n'ai pas pu améliorer la production, mais j'ai amélioré d'un facteur 2 la concentration en protéine obtenue après purification**. Les concentrations, quantités et qualités obtenues avec 4CL-1 ont été suffisantes pour pouvoir continuer le projet avec nos partenaires. En revanche, la concentration obtenue avec HCT ne permet pas la détection d'interaction dans certains cas.**

J'ai également produit différents extraits bruts d'*Arabidopsis thaliana*, pour tenter de lier et d'identifier directement d'éventuelles protéines partenaires susceptibles d'interagir avec les cytochromes P450 (sous forme de nanodisques fixés sur Biacore). **Aucun partenaire n'a pu être identifié.** Seules des protéines majoritaires comme la Rubisco ont été détectées comme interagissant avec les nanostructures.

A l'aide de ces outils biologiques, nos partenaires ont pu démontrer *in vitro* que **les enzymes solubles purifiées (la 4CL-1 et l'HCT) ont la capacité d'interagir avec une bicouche lipidique (sous forme de nanodisques ou de liposomes) en absence de P450**. Ceci conforte ce que j'ai pu observer *in vivo*. De plus, **l'interaction de la 4CL-1 avec les membranes semble dépendante des types de phospholipide constituant cette membrane**. Le mécanisme d'interaction des protéines solubles vis-à-vis des membranes n'a pas été étudié. **L'interaction de 4CL-1 avec les cytochromes P450 (CYP98A3 et CYP73A5) fut démontrée** par co-immunoprécipitation. En revanche, aucune interaction évidente n'a été observée entre HCT et les P450s.

Des approches *in vivo* étaient parallèlement mises en œuvre pour étudier les interactions protéine-protéine au laboratoire. Pour cela, j'ai réalisé une batterie de constructions permettant l'expression dans la plante de protéines fluorescentes en fusion avec nos protéines d'intérêt. L'ensemble de ces constructions a été exprimé de façon transitoire dans des feuilles de *Nicotiana benthamiana* pour une observation au microscope confocal. Ce travail a nécessité la mise au point de nouvelles approches techniques, notamment pour quantifier la relocalisation des protéines solubles.

Un point important pour l'étude des interactions protéine-protéine est la fonctionnalité des protéines utilisées lors de l'étude. Il est admis que des protéines d'intérêt pourraient être déstabilisées par leur fusion avec une protéine fluorescente et par conséquent leurs propriétés d'interaction pourraient être modifiées. Dans un premier temps, la fonctionnalité des protéines constituant la protéine de fusion a été vérifiée. Les constructions préparées ont été démontrées enzymatiquement actives et la protéine fluorescente rapportrice fut également fonctionnelle.

Une méthode a été mise au point pour quantifier la relocalisation des enzymes solubles sous l'influence de la présence de protéines membranaires partenaires. Cette méthode est basée sur la mesure de la répartition des constructions fluorescentes dans un volume tridimensionnel autour des membranes du réticulum endoplasmique. Elle permet de comparer les différentes conditions provoquant une relocalisation, elle n'est pas impactée par les mouvements des composants cellulaires et elle paraît robuste et efficace.

Les résultats obtenus peuvent se résumer de la manière suivante :

- CYP73A5 et CYP98A3 sont manifestement localisés au niveau du Réticulum Endoplasmique (RE).
- Le RE des cellules de plantes est extrêmement mobile en comparaison au RE des cellules animales, ce qui rend l'ensemble du système extrêmement dynamique.
- Les P450s liés aux membranes du RE semblent retenus par des éléments cellulaires non déterminés.
- Les protéines solubles sont localisées dans le cytoplasme, mais sont concentrées localement à la périphérie du RE, même en l'absence d'une co-expression des P450s partenaires. L'HCT est beaucoup plus proche des membranes du RE que la 4CL-1. Le mécanisme conditionnant la localisation des protéines solubles n'a pas été identifié.
- J'ai pu observer une relocalisation partielle au niveau du RE de 4CL-1, mais plus spectaculairement de l'HCT, en présence de CYP73A5 ou de CYP98A3.
- CYP98A3 avait un effet plus marqué que CYP73A5 sur la relocalisation de 4CL-1 et de l'HCT. Ceci laisse supposer un rôle clé de CYP98A3 dans la formation de ce métabolon.
- Une homo-oligomérisation ainsi qu'une hétéro-oligomérisation des deux cytochromes P450 ont été observées.
- L'interaction d'un couple P450-enzyme soluble est renforcée par l'expression de l'autre couple. Par exemple, l'interaction entre HCT et CYP98A3 est observée lors de leur co-expression avec le couple CYP73A5 et 4CL-1.
- L'interaction entre enzymes non consécutives dans la voie de biosynthèse (CYP98A3 avec 4CL-1, par exemple) a été détectée et favorisée lors de la co-expression de l'ensemble des partenaires. Ce qui suggère fortement la formation d'un métabolon.
- Dans les cellules où visuellement la relocalisation des deux enzymes solubles se produit, il est possible de détecter, à la surface du RE, l'interaction des deux enzymes solubles entre elles.
- Finalement, les interactions protéine-protéine n'ont été généralement détectées que lors de la co-expression de l'ensemble des partenaires. Ce qui est sans doute un point critique pour détecter la formation des métabolons, qui sont des complexes transitoires et fragiles.

Tous ces résultats suggèrent la formation d'un complexe supramoléculaire auquel participent ces quatre enzymes. La formation du métabolon est peut-être favorisée par la forte mobilité du RE et des P450s sur celui-ci, ainsi que par la proximité des enzymes solubles avec la membrane du RE. La dimérisation des deux P450s peut offrir une assise plus importante pour le complexe et la réductase indispensable au fonctionnement des P450s peut également avoir un rôle. Ces données mettent en évidence un modèle de métabolons végétaux extrêmement dynamiques, avec des enzymes solubles et membranaires qui sont associées au très versatile réseau endoplasmique.

Les approches expérimentales utilisées lors de ce projet, validées, permettent d'envisager d'appliquer cette procédure à l'étude d'autres voies de biosynthèse. Il serait maintenant intéressant d'étudier la formation de métabolons constitués par les paralogues des quatre enzymes utilisées ici. Une analyse de l'interaction avec d'autres enzymes de la voie métabolique (ou des protéines co-exprimées en diverses conditions) et une analyse plus fine de l'impact des intermédiaires métaboliques peut également être envisagée. Ceci devrait permettre de comprendre le réseau complexe de métabolons se formant dans cette voie de biosynthèse et d'analyser les mécanismes de canalisation du flux métabolique entre les diverses branches de ce métabolisme. La compréhension des mécanismes d'interactions entre protéines et de formation de métabolons est une étape critique pour le succès des approches d'ingénierie métabolique envisagées à l'échelle expérimentale comme à l'échelle industrielle. L'un des enjeux de cette ingénierie métabolique est la production de molécules d'intérêt, comme des molécules thérapeutiques.

3. Caractérisation fonctionnelle des deux clades de chacune des familles CYP73A et CYP98A chez le tabac

Un second projet a été abordé, destiné à élucider les fonctions respectives des différents membres des deux familles de P450s, CYP73As et CYP98As, chez *Nicotiana tabacum*. Chez le tabac, chacune de ces deux familles de P450s est composée de deux clades de paralogues comme chez la plupart des autres espèces végétales (sauf *Arabidopsis*). Ceci semble indiquer que nous sommes devant une plus grande complexité fonctionnelle et métabolique que ce qui a été décrit jusqu'ici. Deux hypothèses peuvent être proposées : soit il y a une différence de fonction catalytique entre les membres des deux clades, soit les différents membres de chacune des deux familles ont une même fonction enzymatique, mais une spécificité d'action résultant de leur expression dans certains organes ou tissus (ou dans certains compartiments subcellulaires).

Afin de déterminer le rôle des sept gènes d'intérêt (4 pour la famille CYP73As et 3 pour CYP98As) différentes approches ont été abordées, parmi celles-ci :

- Les séquences codantes de ces protéines ont été exprimées dans la levure pour déterminer *in vitro* leur activité catalytique avec divers substrats potentiels. Leurs activités catalytiques ont été comparées à celles des enzymes des deux familles chez *A. thaliana*. Une série de substrats appropriés a été synthétisé à cet effet (synthèse chimique avec l'aide du Laboratoire d'Innovation Thérapeutique, UMR 7200, Strasbourg-Illkirch),
- le patron d'expression des gènes a été déterminé par RT-qPCR dans la plante durant le développement, dans les différents organes et en réponse à différents stress biotiques et abiotiques,

- cette expression a été mise en relation avec les profils métaboliques des tissus concernés déterminés par Chromatographie Liquide Ultra Performante couplée à un spectrométre de masse (UPLC-MS/MS),
- des mutants de surexpression ont été générés afin de caractériser l'impact de la surexpression de ces gènes sur le phénotype et le métabolisme de la plante.

3.1. Résultats obtenus au sujet des paralogues de CYP98As

Les résultats obtenus indiquent :

- que deux clades de CYP98As semblent coexister chez certaines plantes comme le tabac par exemple. Les résultats obtenus avec un test d'évolution laissent suggérer qu'une phase d'évolution rapide suivie d'une phase de stabilisation s'est produite dans une branche conduisant aux protéines du clade 2 du tabac.
- que les patrons d'expression des différents gènes chez le tabac sont distincts durant le développement de la fleur et dans les divers organes du tabac.
- que *CYP98A30* du clade 2, qui s'exprime très faiblement dans les tissus sains, est activé durant la défense contre le virus de la mosaïque du tabac. Une induction de son expression fut également observée lors de divers stress.
- que l'expression de *CYP98A30* est la plupart du temps corrélée à celle de *HQT* qui est impliquée dans la biosynthèse de l'acide chlorogénique. L'expression de *CYP98A31* (clade 1) a été le plus souvent corrélée à celle de *HCT* ou de *CYP98A33* (clade 2).
- des activités enzymatiques conformes aux attentes et assez peu différenciées pour les différentes protéines, du moins en ce qui concerne les substrats testés jusqu'ici. Le substrat préférentiel des trois CYP98As de tabac est le *p*-coumaroyl shikimate. De plus, CYP98A31 et CYP98A33 (représentant les deux classes) peuvent également métaboliser des esters et amides très variés de l'acide *p*-coumarique, de la tri-*p*-coumaroyl spermidine au *p*-coumaroyl shikimate. Finalement, les enzymes appartenant au clade 2 semblent avoir la capacité de métaboliser la *p*-coumaroyl octopamine.
- qu'une surexpression dans la plante permet une surexpression des transcrits qui ne se traduit pas par une expression accrue de protéine. Il doit donc exister un mécanisme post-transcriptionnel (non déterminé) qui régule l'expression des enzymes de la voie métabolique.

D'après ces données, CYP98A31 (clade 1) est censé agir dans la voie de biosynthèse de la lignine. CYP98A30 (clade 2), faiblement exprimé dans les tissus sains et induit par certains stress, pourrait être impliqué dans la synthèse de l'acide chlorogénique lors des mécanismes de défense. Finalement, CYP98A33 (clade 2) semble plus versatile.

3.2. Résultats obtenus au sujet des paralogues de CYP73As

Les résultats obtenus indiquent :

- que deux clades de CYP73As semblent coexister chez certaines plantes comme le tabac par exemple.
- qu'une duplication d'un gène ancestral de CYP73 s'est produite avant la séparation des monocotylédones et dicotylédones. Cette duplication a été maintenue chez les Angiospermes avec une divergence significative du clade II.
- que ces divergences correspondent à une ancre N-terminale atypique, une insertion de 3-4 acides aminés en amont de la cystéine servant à la fixation de l'hème, à un motif HXP au niveau du site actif et à une importante modification des acides aminés en surface de la protéine.
- une expression subcellulaire au niveau du RE dans tous les cas, en dépit des prédictions contraires pour les membres du clade II de la famille CYP73A qui présente une séquence d'adressage de type plastidiale.
- une distinction flagrante d'expression entre les deux classes, durant le développement de la fleur et dans les différents organes de la plante. Les gènes des membres du clade II sont exprimés dans les organes reproducteurs (fleurs, fruits). D'autre part, ces gènes ne sont pas induits par les mêmes stress. La plupart du temps, l'expression des CYP73As du clade I était associée à l'HCT et aux CYP98As du clade 1. Ainsi, ces membres semblent liés à la lignification.
- Que les activités enzymatiques sont conformes aux attentes et assez peu différenciées pour les différents membres, du moins en ce qui concerne les substrats testés jusqu'ici. Le substrat préférentiel des trois est le cinnamate. Cette analyse de fonction catalytique révèle cependant une très forte sélectivité de substrat pour les CYP73s qui s'oppose à une très grande plasticité pour les CYP98s.
- qu'une surexpression dans la plante permet une surexpression des transcrits qui ne se traduit pas par une expression accrue de protéine. Ceci pourrait indiquer l'action d'un mécanisme post-transcriptionnel non déterminé, qui régule l'expression de ces enzymes.

D'après ces données, les CYP73As du clade I sont censés agir dans la voie de biosynthèse de la lignine, alors que ceux du clade II pourraient avoir un rôle au niveau de la fleur, son développement, sa protection ou pour la reproduction.

Les différences de fonction des deux clades de CYP73s et CYP98s chez le tabac n'ont donc pas pu être totalement élucidées. En revanche, l'ensemble du travail conduit sur ce second projet est une première étape vers la compréhension du rôle des deux clades de cytochromes P450 de chacune des familles CYP73As et CYP98As chez le tabac, mais également chez d'autres Angiospermes.

Preamble

Plants are sessile organisms, which means they are continuously exposed to varying and often harmful environmental conditions.

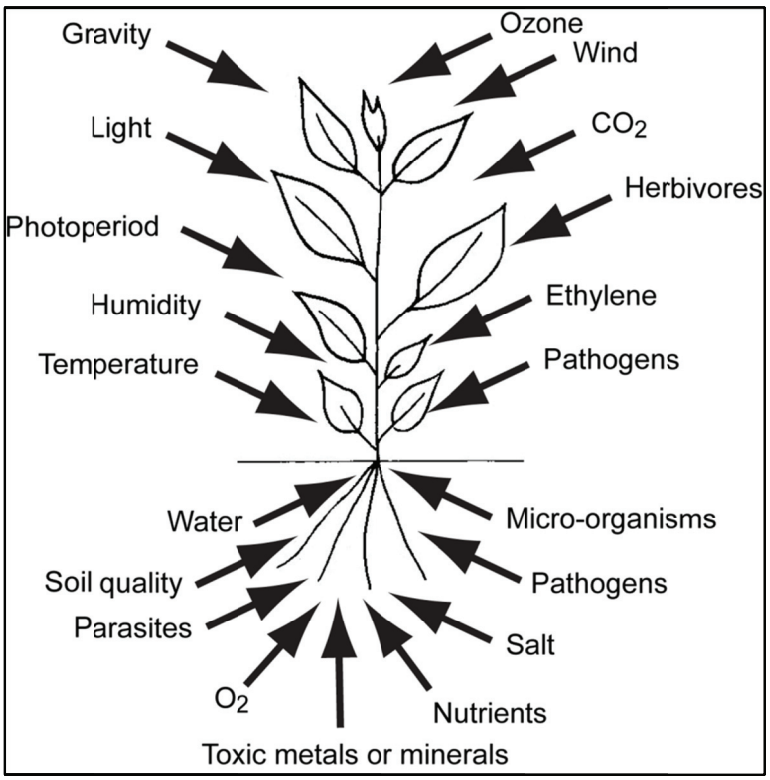


Figure 1. Environmental signals of physical, chemical or biological nature, being able to affect growth and development of plants.

As a result, they must be able to resist many stresses such as climatic variations, wounds, infections by pathogens, aggressions by herbivores, darkness, high light, UV radiation, drought, high salt… (Figure 1). But during evolution, plants were able to develop a whole panel of strategies enabling them to benefit from their environment and to adapt to a great number of situations while living in a sedentary way. One of the most outstanding examples of plant adaptation is the evolution of lignin synthesis allowing water to circulate in a plant and conferring rigidity to the tissues. Lignin was thus essential for evolution of land plants. Another example is found in reproduction: plants were able to cross-pollinate and to spread by attracting pollinating insects and disseminator animals, by developing protective systems of their seeds against cold, desiccation…

Under the effect of selection pressure, plants permanently adapt to their environment. Metabolism and synthesis of specific molecules play an essential role in this adaptation. Some of these molecules are entirely specific to the plant kingdom and essential for the plant life. It is the case for example for molecules acting during photosynthesis. Plants were also able to divert the basal processes, to produce from universal precursors (from primary metabolism) tens of thousands of molecules allowing to react to different situations. This metabolism called "secondary" is of a complexity without equal. It led to the evolution of new enzymes (by duplications/transfers/mutations), giving birth to new molecules with new functions. Some of them were conserved because they conferred adaptative advantages, in a specific environment, over the rest of the population. These beneficial compounds are signals for intra-plant, inter-plants, or plant-micro-organisms or sometimes plant-insects or plant-herbivores communication. These signaling chemicals can be for example: alkaloids, terpenoids, or phenylpropanoids.

The aim of the laboratory is to study a family of enzymes, the cytochromes P450 (P450s), to acquire a better understanding of the cellular and molecular mechanisms which govern the biosynthetic pathways, mainly by biochemical and molecular approaches and by functional genomics. The potential applications of these researches are for example: identification of new chemical mediators of plant development, improvement of the plant defense systems, improvement of the organoleptic properties of the plants or phytoremediation.

The work presented in this report focuses on P450s involved in the phenylpropanoid pathway. This work was supported by the Human Frontier Science Program and by a PhD stipend from the French Ministry of Research.

This thesis manuscript starts with a general introduction on the P450 family of enzymes. Then I will present the phenylpropanoid pathway, the role of some of the resulting metabolites, and the enzymes involved in the upstream section of this pathway. I will then explain the background of my project.

The three next chapters describe the three main lines of my investigations: 1) the protein-protein interactions for the formation of metabolons in the upstream phenolic pathway, 2), the duplications in the CYP98 family in *Nicotiana tabacum* 3) the role of CYP73 duplication also in tocacco. A final section will synthesize the principal results.

GENERAL INTRODUCTION

"…You're waiting for a train, a train that will take you far away. You know where you hope this train will take you, but you can't be sure. …" C. Nolan

1. The cytochromes P450 (P450s)

P450s are membrane hemoproteins. They constitute one of the largest classes of enzymes. P450s are present in all the reigns (prokaryotes, protists, fungi, plants and animals). Nevertheless certain primitive species of bacteria do not contain any P450 enzymes. This dates the appearance of the first cytochromes P450 about 3.5 billion years ago (Nelson *et al.*, 1993). P450s are particularly diversified in plant, where they are involved in many processes: the biosynthesis of secondary metabolites, of hormones, of structural macromolecules constituting the cell wall or the protective surface layers, the detoxification of exogenous molecules or the activation of molecules into toxic compounds.

If all plant species are considered, thousands of P450s provide opportunities for many applications in agronomical and pharmaceutical industries (Morant *et al.*, 2003).

1.1. History of the P450s

The study of cytochromes P450 started in 1955, by detection in mammalian liver of an enzymatic system able to metabolize xenobiotics (Brodie *et al.*, 1955). Then in 1958, Garfinkel and Klingenberg discovered pigments which, associated with a carbon monoxide molecule, absorb the light at 450 nm (Garfinkel, 1958; Klingenberg, 1958). Thereafter, Omura and Sato showed that these pigments were hemoproteins and called them cytochromes P450, P meaning pigment and 450 corresponding to the maximum of absorption of this pigment in nanometer in its reduced and CO-bound form (Omura & Sato, 1964a and b). These discoveries initiated extensive investigations on the properties of the cytochromes P450 and their physiological functions. P450s were then discovered in prokaryotes (bacteria) and many other eukaryotes (fungi, insects, fishes and mammals). Monooxygenase activity of P450 was first described in 1965 (Cooper *et al.*, 1965).

It was in 1969 that the presence of cytochromes P450 was first documented in plants. Murphy and West demonstrated the involvement of P450 in the oxidation of kaurene in *Echinocystis macrocarpa* (Murphy & West, 1969), while Frear and coworkers described the same year P450-dependent demethylation of the phenylurea herbicide monuron in cotton seedlings (Frear *et al.*, 1969). Several other biochemical studies undertaken on microsomal membranes from other plants confirmed the existence of P450s in plants (Markham *et al.*, 1972).

Today, approximately 18.400 P450 sequences are identified, including approximately 4.267 named in plants (Nelson, 2010), with for example 249 genes and 25 pseudogenes coding for P450s in the *Arabidopsis thaliana* genome (http://arabidopsis-P450.biotec.uiuc.edu), 366 genes and 99 pseudogenes in the rice genome, 225 genes in *Selaginella moellendorffii*, 71 genes in *Physcomitrella patens* and 39 genes in *Chlamidomonas* (http://drnelson.uthsc.edu/cytochromeP450.html).

For comparison to the number of P450 genes found in plants, only 57 are found in the human genome and 3 in *S. cerevisiae*.

In plants, an average 1% of annotated genes code for P450s. The functions of less than 30% of them are described in model plants, and these functions are only partially described.

1.2. Characteristics of the cytochromes P450

Because of their high number and diversity, a nomenclature to classify these enzymes soon became necessary. As soon as 1985, a nomenclature based on the identity between peptidic sequences was established (Nebert *et al.*, 1987; Nelson *et al.*, 1993). This classification system, which proved to be effective, evolved to take into account phylogenetic criteria (Nelson, 2004).

P450 classification is mainly based on peptidic sequence similarities. P450s genes are named "CYP" for "Cytochrome P450", followed by a number for the family, of one to two letter(s) indicating the subfamily. Each subfamily can be made up of several genes, each one being then identified by a number. Numbers are assigned by chronological order of annotation or submission (Figure 2). This classification is established and individual genes named by D. Nelson (University of Tennessee, USA): http://drnelson.utmem.edu/CytochromesP450.html. In this system, two enzymes sharing more than 40% of identity usually belong to the same family. Beyond 55% of identity, they belong to the same subfamily. Genes sharing more than 97% sequence identity are considered as allelic variants.

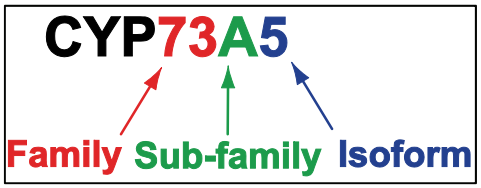


Figure 2. Nomenclature of P450s.

For plant P450s, the families are numbered from CYP71 to CYP99, and then from CYP701 to CYP999 (we are already in the CYP800s). For now, more than 140 families of plant P450s are known. Considering the rapidly increasing number of genes emerging from sequencing, a new degree of organization was recently introduced. The families belonging to a common evolutionary clade are now associated in clans (Nelson, 2004). Each clan is named according to its lowest-numbered family member. This identity rule has some exceptions, especially in plants, where gene duplication and shuffling sometimes makes a straightforward nomenclature difficult. In this case, family assignment is based on other information, in particular phylogeny and gene organization (place and number of introns for the sequenced organisms). Despite these new criteria, in the case of the plants, nomenclature becomes complex and is sometimes reassessed: in very big families some P450s names had to be changed (only three in 2004).

P450s present few common points: they are hemoproteins. Certain residues of their primary protein sequence are well preserved: the Phe-x-x-Gly-x-Arg/His-x-Cys-x-Gly motif on the C-terminal side (helix L) is preserved in most of P450s. It includes the cystein (Cys) which serves as fifth ligand with the iron of heme. The motif Ala/Gly-Gly-x-Asp/Glu-Thr-Thr/Ser is another one. It is located in helix I and corresponds to the site of oxygen binding (Figure 3) at the distal side of the heme (Werck-Reichhart & Feyereisen, 2000; Nelson, 2004). Finally some specific amino-acids are conserved, known as the E-R-R triad consisting of the Glu and Arg of the K-helix consensus (K**E**TL**R**) and the Arg in the "PE**R**F" consensus. The E-R-R triad is generally thought to be involved in locking the heme pockets into position and to assure stabilization of the conserved core structure. Of these conserved domains, only the E-R-R triad and the cysteine in the heme-binding domain are conserved in all plant P450 sequences.

All eukaryotic P450s described so far are anchored in membranes: in animals and plants, they are mainly located onto the smooth endoplasmic reticulum (ER). These cytochromes P450 are named microsomal P450s (Poulos & Johnson, 2005). They are associated with the membranes by a hydrophobic N-terminus sequence of 20 to 30 amino-acids. The globular part of the protein is in the cytosolic side. However some animal or plant P450s were localized in cellular organelles, mitochondria in mammals or insects (Bureik *et al.*, 2002), or chloroplasts in plants (Watson *et al.*, 2001; Siqueira-Junior *et al*., 2008).

1.3. P450 tridimensional structure

The heme core and whole three-dimensional structure of P450s are conserved. However large divergences are found in their primary sequences and some proteins can share less than 15% of identity.

All of them have the same globular three-dimensional structure formed by two domains, one primarily made up of α helixes, the other of β layers (Werck-Reichhart & Feyereisen, 2000). An N-terminal α-helix of about twenty hydrophobic residues anchors P450s in the lipidic bilayer. This helix is followed by a region rich in basic residues (which stabilize interaction with membrane) followed by a proline-rich stretch (PPXP) that favors the correct folding of the globular domain on the external surface of ER by conferring flexibility.

The first studied three-dimensional P450 structures were obtained from prokaryotic cytochromes P450. Crystal resolution provided the structure of P450cam in 1987 and of P450BM3 in 1993 (Ravichandran *et al.*, 1993). Purification and subsequent crystallization were facilitated by solubility of the prokaryotic enzymes. Among most significant advances in structure resolution of membrane-bound enzymes were the crystallographic structure of a cytochrome P450 reductase (CPR) with the first elucidated structure of mammalian P450. This combination of two structures provided better understanding of the interactions between the two partners and of the transfer of electrons from the NADPH for the reaction (Poulos & Johnson 2005; Figure 3).

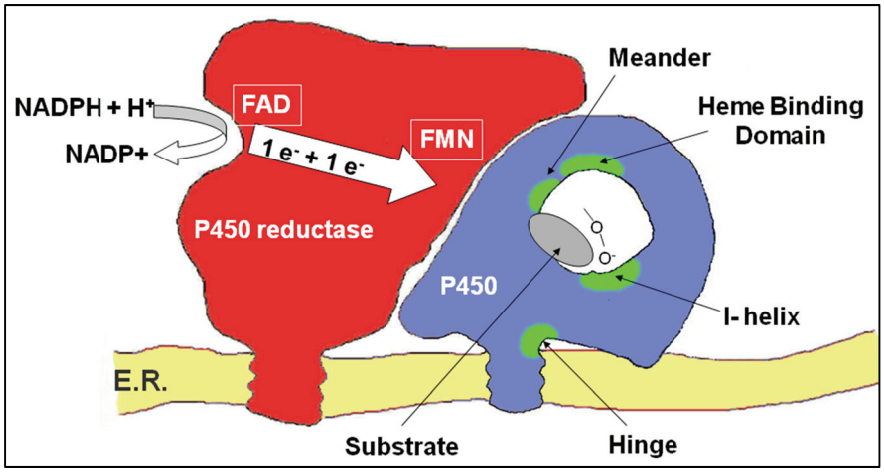


Figure 3. Catalytic functioning of cytochrome P450 (from Werck-Reichart & Feyereisen, 2000). Transfer of two electrons from NADPH is dependent on a P450 reductase essential for the reaction.

1.4. Catalytic function of the P450s

The P450-catalyzed reactions are extremely diverse, but usually based on activation and heterolytic cleavage of molecular oxygen with insertion of one of its atoms into the substrate and reduction of the other to form water (Mansuy, 1998; Werck-Reichhart & Feyereisen, 2000). They are classified as monooxygenases.

Typical reaction: $\mathbf{RH + O_2 + NAD(P)H{+}H^+ \rightarrow ROH + H_2O + NAD(P)^+}$

The reactions catalyzed by P450s correspond to a stereospecific and regiospecific oxygenation of a non activated atom (carbon, sulfur, nitrogen…). In the cellular environment (pH and physiological temperature) these reactions are highly energy consuming. Energy necessary to the reaction is brought by a cofactor: commonly NADPH. This cofactor is used as an electron donor for the reaction. In eukaryotes the transfer of electrons from NADPH is usually dependent on a NADPH:Cytochromes P450 Reductase (CPR) essential for the reaction (Bernhardt, 2006). Just like P450, CPR is anchored in the ER membrane via its N-terminus (Mansuy, 1998; Werck-Reichhart & Feyereisen, 2000). In *A. thaliana,* two CPR were isolated. The expression of *AtATR1* was found to be highest in roots and stems whereas *AtATR2* expression was highest in leaves, stems and flowers (Mizutani & Ohta, 1998).

The heme cofactor of P450s is a protoporphyrin IX, with one of its side-chains, and the heme iron, anchored on the monomeric apoprotein. When the active site of the enzyme is in the resting state, the heme iron is hexacoordinated. Iron is tetracoordinated with the nitrogen atoms of the tetrapyrrole. A fifth coordination forms with the sulfur atom from cysteine residue of a conserved sequence (the heme-binding loop) from the apoprotein. The sixth coordination is with a water molecule present in the substrate pocket, or one with residues of the apoprotein (for example threonine of the helix I; Figure 3).

The substrate comes to position precisely in the middle of the active site, in interaction with amino-acid residues of the protein, and moves one or more water molecules, which causes a dissymmetry around the metal atom. The iron, which had until now a maximum of symmetry compared to the plan of heme, shifts slightly. This mode of binding of the substrate induces a typical type 1 shift of the Soret band of UV-visible absorption from 395 nm to 420 nm. This property is characteristic of P450 enzymes. The redox potential of the Fe^{III}/Fe^{II} couple is then increased (Jefcoate, 1978) by an hundred millivolts, which allows the transfer of one electron and the reduction of iron in Fe^{II+} pentacoordinated. A dioxygen molecule can then come to take the sixth iron coordinance (in the place of the water molecule) and the hydroxylation of the substrate from one of the atoms of dioxygen can be done on a precise position after the use of a second electron. The second oxygen atom is reduced into water molecule. The expulsion of the product leads to the initial form of the active site at the resting state (Werck-Reichhart & Feyereisen, 2000; Figure 4).

A typical complete catalytic cycle requires one molecule of dioxygen and one molecule of NADPH. In some cases, if there is uncoupling between the consumption of NADPH and the formation of the product, abortive cycles can take place leading to the formation of superoxide ion or H_2O_2. Such oxidizing compounds can inactivate the protein and are poisonous for the cell. Uncoupling can occur when the position of the substrate is not optimal in the active site (Perret & Pompom, 1998). A cellular role related to the formation of reactive oxygen species (ROS) catalyzed by P450s was proposed during apoptose (Davydov, 2001).

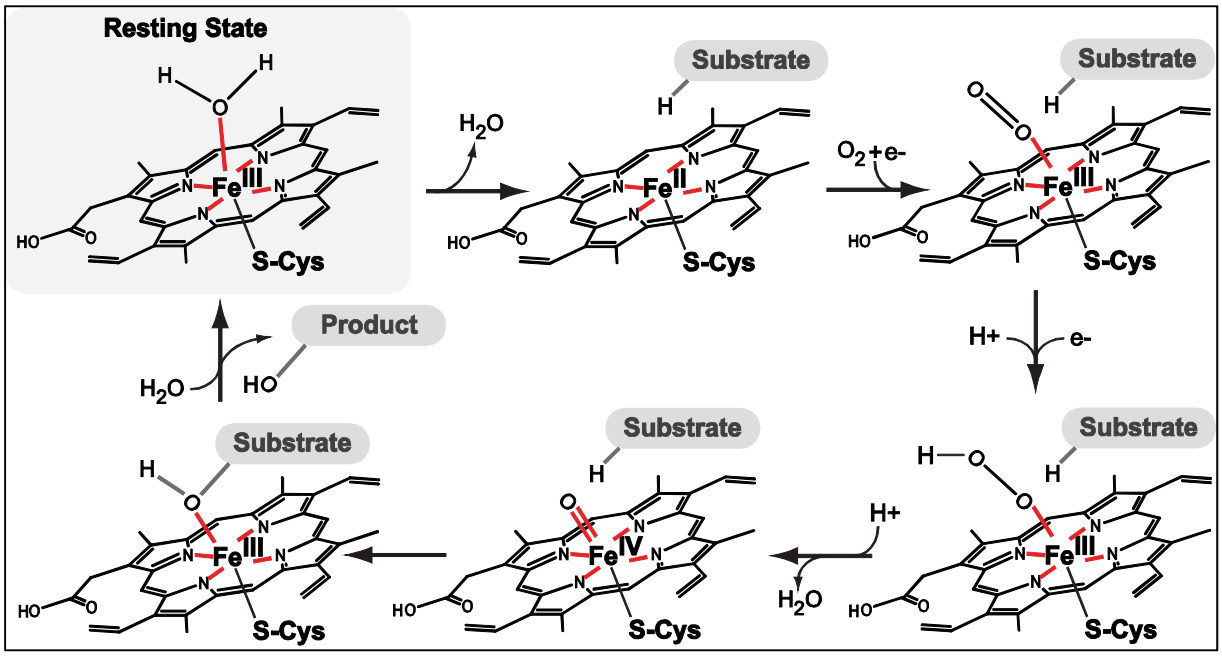


Figure 4. Catalytic cycle of P450 enzymes. The entrance of the substrate in the active site provokes a movment of water molecule(s). After reduction of the iron by one electron, a molecule of dioxygene binds to the iron of the heme. A second electron in addition to an H^+ trigger the hydroxylation of the substrates and the synthesis of one water molecule. The expulsion of the product leads to the resting state of the catalytic cycle.

The most typical reactions catalyzed by P450s are hydroxylations, but P450 activity can also result in a diversity of reactions of epoxydation, demethylation, isomerization, ring extension, ring opening, dehydration, reduction… (Mansuy, 1998; Werck-Reichhart & Feyereisen, 2000; Bernhardt, 2006). More than one hundred different enzymatic reactions described in plants are catalyzed by P450s (Cf. Appendix: List of P450 functions characterized in Arabidopis thaliana p.311).

1.5. Physiological functions of cytochromes P450

The catalytic function of some P450 family is partly preserved, but specificities of substrates also largely diverged, leading to a broad range of functions in different organisms.

Animals P450s were studied more than plants counterparts. They are involved in the biosynthesis or the catabolism of steroids, leukotrienes, prostaglandins, bile acids, vitamin D… They also take part in the detoxification or the activation of exogenous molecules, such as drugs or pesticides.

Plants P450s take part in many reactions, in all the pathways of the secondary metabolism and in few pathways of the primary metabolism, such as biosynthesis of sterols. They can catalyze reactions essential to survival and development, such as the biosynthesis or the catabolism of very diverse metabolites: sterols, hormones and signals (brassinosteroids, auxin, gibberellins, cytokinins, strigolactones, jasmonate…), antioxidants (chlorogenic acid, flavonoids, carotenoids…), structural polymers (lignins, suberins, sporopollenin, cutins), pigments (anthocyanins, carotenoids…), odorant molecules, flavors and defense compounds (cyanogenic glucosides, glucosinolates…). They are also involved in the metabolism of exogenous compounds: detoxication/activation of herbicides, insecticides or pollutants (Schuler, 1996; Werck-Reichhart *et al*., 2002; Schuler & Werck-Reichhart, 2003; Morant *et al.*, 2003; List of P450 functions characterized in Arabidopis thaliana p.311).

In plant various groups of P450s can be distinguished:

- P450s that are essential for the plant survival and that take part in the synthesis of structural polymers, of membrane steroles or of the hormones
- P450s that participate in the metabolisation of xenobiotics
- P450s that are involved in the plant defense reactions or in the communication with other organisms (plants, bacteria, insects...) or with their environment.

I will refer now to some examples in the three next paragraphs.

Some P450s participate in essential processes in plant. For example CYP707 is involved in the catabolism of abscissic acid (terpenoid hormone), itself involved in regulating seed dormancy (Saito *et al*., 2004). They take part in hormonal regulation i.e. in the synthesis of all plant hormones. For example CYP701As catalyze the first and CYP88s the second oxygenation steps in the synthesis of gibberellins, hormones controlling growth and development (Helliwell *et al*., 1998). Other CYPs are involved in the maintenance of the plant. As we will see in the next paragraphs, P450s play a very important role in the biosynthesis of the phenylpropanoids and in particular lignin (with the CYP73s - Cinnamate 4-Hydroxylase, CYP98s - Coumaroyl ester 3'- Hydroxylase and CYP84s - Ferulate 5-Hydroxylase).

P450s play a decisive role in plant defense and communication with its environment. For example, CYP75s take part in the synthesis of the flavonoids as flavonoid 3'-hydroxylase (F3'H) (Brugliera *et al*., 1999) controlling the flower color. This color is important to attract pollinating insects. In plant defense, P450s participate in the biosynthesis of the hydroxylated fatty acids or phenolamides (compounds important in the plant defense against insects and pathogens), in wound healing or abiotic stresses. For example, *A. thaliana* CYP86A2 is required for the biosynthesis of the hydroxylated lipids of cutin that act as defense barrier and/or signaling. Derived signals could repress the expression of bacterial genes (Xiao *et al*., 2004).

Plant P450s can be used to confer a resistance to herbicides in transgenic plants recombinant for one of these genes (Powles & Yu, 2010). For example CYP71A10 from soybean metabolizes four herbicides of the phenylurea family: fluometuron, linuron, diuron, and chlortoluron. Transformation of tobacco lines (*Nicotiana tabacum*) by this gene under the control of a *CaMV-35S* promoter leads to a faster metabolisation of these herbicides and confers a tolerance to linuron more than twelve times higher than that of wild tobacco (Siminszsky *et al*., 1999).

1.6. Economical interests linked to P450s

From a kinetic point of view, P450s are very slow enzymes in comparison to enzymes like alcohol dehydrogenase. Reactions catalyzed by P450s are often difficult to perform by chemical synthesis. They are often limiting, which makes them interesting for metabolic engineering. The set of interesting reactions being catalyzed by P450s and the availability of new genetic engineering techniques allow heterologous expression, improvement and modification of their activity, stability and selectivity. The increasing interest of the industry in life sciences makes them promising candidates for biotechnological applications in the future (Bernhardt, 2006).

An example is the production of transgenic plants such as blue carnations and roses, by over-expression of a CYP75A (Katsumoto *et al.,* 2007; Tanaka *et al.,* 2010). Other examples are the production of taxol precursors (alkaloid used in breast cancer therapy; Ajikumar *et al.*, 2010), or of hydrocortisone (by Sanofi-Aventis) using a designed strain of Baker's yeast expressing a total of four P450s (Szczebara *et al.*, 2003). They can also be used to improve defense against pathogen or insect (Tattersall *et al.*, 2001), or for the production of molecules for perfumery (Morant *et al.,* 2003).

Genes conferring resistance to herbicides could have an economical interest as selectable markers (Werck-Reichhart *et al.,* 2000). The over-expression of CYP76B1 from *Helianthus tuberosus* in *A. thaliana*, under the control of a *CaMV-35S* promoter, confers resistance to the herbicides linuron, isoproturon and chlortoluron, up to twenty times higher than that of wild type plants. This P450 catalyzes the fast oxydative dealkylation of many of the phenylurea herbicides. CYP76B1 expression does not have another phenotypic effect on these transgenic plants (Robineau *et al*., 1998; Didierjean *et al*., 2002).

During the last two decades, numerous experiments were conducted on the over-expression of mammalian (Abhilash *et al.,* 2009) and plants P450s in higher plants such as *Nicotiana tabacum*, *Solanum tuberosum*, *Oryza sativa* or *Arabidopsis thaliana*. The primary objectives of these genetic manipulations were the production of either herbicide resistant plants (e.g. tolerance towards atrazine, simazine) or of plants capable of enhanced metabolization of xenobiotics (pesticides, herbicides, volatile halogenated hydrocarbons, even explosives) for phytoremediation, i.e. removal of toxic compounds from contaminated soil and groundwater.

These various applications make cytochromes P450 an important economic issue.

1.7. Diversity and evolution of P450s

The driving force of P450 diversification in land plants has been closely linked to survival strategy. This diversification confers advantages in the competition with other organisms (Mizutani & Ohta, 2010). The fixing of a new function acquired by an organism can be very fast. For example the degradation of the cyclotrimethylene trinitramine (RDX: explosive) thanks to P450, was acquired in less than 50 years by bacteria since the use of RDX has started (Rylott *et al.,* 2010).

The origin of P450s precedes the appearance of the eukaryotes as well as the accumulation of dioxygen in the atmosphere. The first possible function of P450s could have been a reductase activity evolving to peroxidase activity for adaptation to O_2 appearance (3 to 3.6 billion years ago). When oxygen appeared, a possible critical function for P450s could have been the detoxication of these traces of O_2, by catalyzing a concerted transfer of four electrons to form H_2O (Kahn & Durst, 2000).

A second critical stage in their evolution dates 2.5 billion years when the atmosphere became hyperoxic, switching from reductant to oxidant (Sessions *et al.,* 2009). Oxygen insertion into endogenous or xenobiotic substrates could then be integrated in the catalytic mechanism. However, no function is completely conserved across the phyla, which makes phylogenetic interpretations quite difficult. Only the CYP51 family is well maintained across fungi, animals, plants and *Mycobacterium tuberculosis* (possibly as a result of lateral gene transfer from its host). *CYP51* codes for a sterol 14-α-demethylase. On the other hand, this gene was lost in insects and nematodes which are sterol heterotrophs (Werck-Reichhart & Feyereisen, 2000). This function probably developed with the ancestral prokaryote, more than 2.7 billion years ago (appearance of eukaryotes; Cavalier-Smith, 2009). As sterols and fatty-acids are basic components of the membranes and are tightly related to the cell primary metabolism, the oxygenation of these compounds could be one of the first oxygenase functions of P450s.

To date, P450s were isolated from less than 100 plant species from the more than 300.000 known existing species. For example, very few data are available for the Gymnosperms, where only P450s from conifers were annotated and characterized. Plant P450 families can be categorized into four classes as follows: essential reactions conserved in the plant kingdom, core reactions/pathways conserved in all land plants, essential reactions/pathways that emerged during flowering plant evolution (e.g. plant hormone homeostasis), and specialized reactions/pathways unique to some plant species (Mizutani & Ohta, 2010). Some P450 genes are anterior to the divergence between Gymnosperms and Angiosperms (estimated at 360 million years). Those belong to families *CYP85*, *CYP86*, *CYP73*, *CYP78*, *CYP98* (Nelson, 2006; Figure 5).

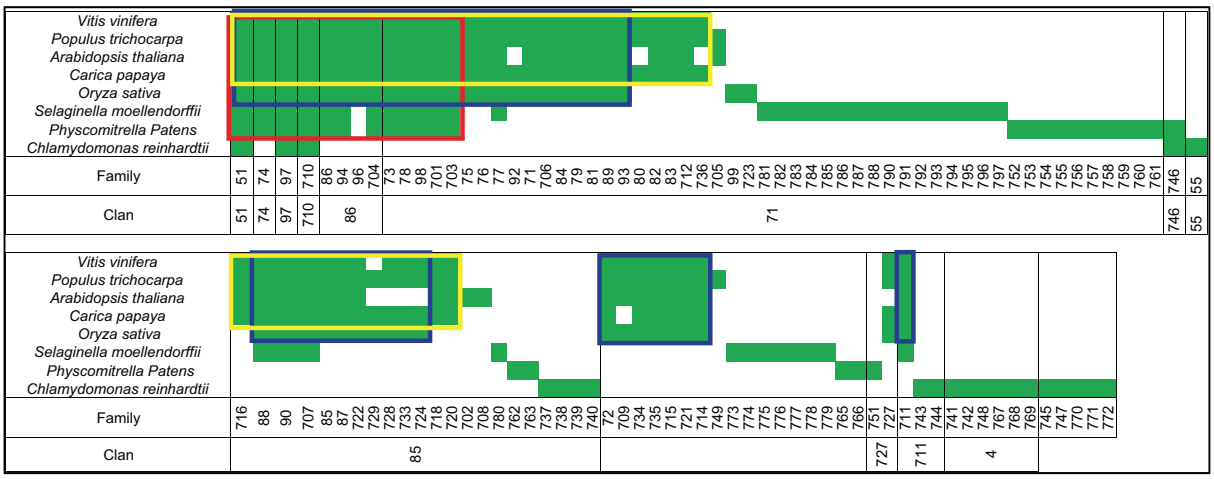


Figure 5. Repartition and evolution of P450 families in sequenced plants representing the evolution.
(Based on D.R. Nelson oral presentation at The 9th International Symposium on Cytochrome P450 Biodiversity and Biotechnology).

1.8. Heterologous expression of plant P450s

The systematic sequencing of the genome of various organisms led to the detection of a large number of genes potentially coding for cytochromes P450. Annotation of the genomic sequences is the first step leading to the biochemical characterization of the enzymes. The next steps are the cloning of the target coding sequences and their expression in the most suitable expression system. Very often, heterologous expression is a limiting step for the biochemical characterization of the enzymes. Indeed, it is very frequent that the system used is not optimal, and in certain cases, it can lead to a low protein expression or to no detectable expression.

1.8.1. Expression of P450s in the yeast system

Bozak *et al.* (1992) were the first to report heterologous expression of a plant P450 in yeast. CYP71A1 from avocado was shown to have a *para*-chloro-*N*-methylaniline demethylase activity but its physiological substrate is still unknown. For heterologous expression the yeast *S. cerevisiae* was transformed by a plasmid containing the *CYP71A1* coding sequence under the control of a galactose-inducible promoter. This system had already been used for the expression of non- plant P450s.

Thereafter, other plant P450s were expressed using this system. Various modifications and optimizations were achieved for obtaining a better production of enzymes.

In 1994, Urban *et al.* described yeast improvement for the expression of CYP73A1 (Cinnamate 4-Hydroxylase from Jerusalem artichoke). For these experiments, various yeast lines were used. These lines differed in the level and type of CPR expressed (endogenous yeast CPR under the control of its native promoter; plant CPR under the control of a galactose-inducible promoter; human CPR under the control of a galactose-inducible promoter).

Finally in 1996, Pompon *et al.* described an optimized system for P450 expression in *S. cerevisiae*. Two new lines were engineered, WAT11 and WAT21 expressing respectively CPRs of *A. thaliana* ATR1 and ATR2 under the control of a galactose inducible promoter (Urban *et al.*, 1997). These yeast lines are still today the most frequently used recombinant organisms for the production of plant P450s.

1.8.2. Main issues during the expression in yeast

The use of yeasts to produce plant proteins is not always an optimal solution. Indeed for plant P450s many attempts of plant P450 heterologous expression proved to be unsuccessful, resulting in no production of the enzyme or production of an inactive enzyme.

These problems can have various causes. For example it was shown that the codon usage bias could have an important impact on the amount of expressed protein. A partial optimization of the coding sequence according to the yeast codon preference can, in some cases, significantly improve the level of expressed protein (Hehn *et al.*, 2002; Kandel *et al.*, 2005).

Another successful strategy of optimization is the exchange of the membrane anchor of the P450 of interest with the anchor of a P450 having a strong level of expression in yeast. Amongst tested P450s, this strategy allowed the expression of CYP73A15 (Cinnamate 4-Hydroxylase of *Phaseolus vulgaris*) or of CYP71AJ1 (a psoralen synthase of *Ammi majus*) (Nedelkina *et al.*, 1999; Larbat *et al.*, 2007). In another case, the native plant membrane anchoring sequence of CYP81B1 was successfully replaced with an amphipathic peptitergent for increased protein production (Cabello-Hurtado *et al.*, 1998).

Lastly, the CPR co-expressed with the P450 can also have importance. For example it was shown that CYP73A1 has a better activity in yeast microsomes when it is co-expressed with the yeast CPR that when it is co-expressed with the human CPR (Urban *et al.*, 1994) while *H. tuberosus* CYP81B1 showed a better functional coupling in the presence of ATR2.

1.8.3. Use of different production systems

In spite of some improvement of protein expression as mentioned previously, some sequences remained recalcitrant to the expression in the yeast.

Other strategies can be considered. Among the systems proposed, expression in *Escherichia coli* can be considered. This system was for example reported for the production and the characterization of CYP71E1 in 1998 (Bak *et al.*, 1998). However, as functional analysis of *E. coli* expressed enzyme requires reconstitution of functional membranes with purified P450 reductase, it appears more appropriate to use an eukaryotic recombinant system with ER membranes.

In this way, an efficient system commonly used in many laboratories is insect cells transfected with baculovirus. This organism provides a different lipidic environment in some cases allowing a better anchoring or stability of the enzyme. Moreover, contrary to yeast, it was shown that the bias in codon usage of the Sf9 cells of *Spodoptera frugiperda* is limited (Landais *et al.*, 2003). Another advantage of this system is that it allows a better flexibility for co-expressions of two proteins or more since it is possible to transfect a same cell culture with several recombinant viruses encoding different proteins. Finally, the dosage of virus used for transfection also makes it possible to control the production of protein (Duan *et al.*, 2004).

In conclusion, cytochromes P450 are membrane-bound monooxygenases involved in a plethora of reactions, mainly due to their high number and high diversity in all reigns. These enzymes catalyze critical steps in all metabolic pathways especially in the highly diverse plant secondary metabolism.

2. Secondary metabolism

By definition, secondary metabolites are molecules that do not participate directly to the normal growth, development or reproduction of the plant. They rather play part in survival against biotic and abiotic stresses, in fecundity and they are species specific. On the contrary, primary metabolites are considered as essential for the plant/cell development and are present in all species. This differentiation between secondary and primary metabolites is however regarded as more and more obsolete and arbitrary (Hartmann, 2007). There is increasing evidence that essential genes of primary metabolism have been recruited for evolution of secondary metabolism by duplication. The duplicated genes have acquired new functions and have been optimized and diversified for their roles in new pathways (Boudet, 2007).

All secondary metabolites arise from universal precursors, present in all plants, by universal mechanisms of the primary metabolism (Figure 6). For example, all plants synthesize phenylpropanes starting from phenylalanine, and all thus necessarily contain all the basic elements involved in this complex process. However, certain plants underwent adaptive divergences; thus, only some of them can accumulate these metabolites in large amounts or deeply modify the structure of these universal elements. Thus all the phenylpropanoid classes are not present in all plant species.

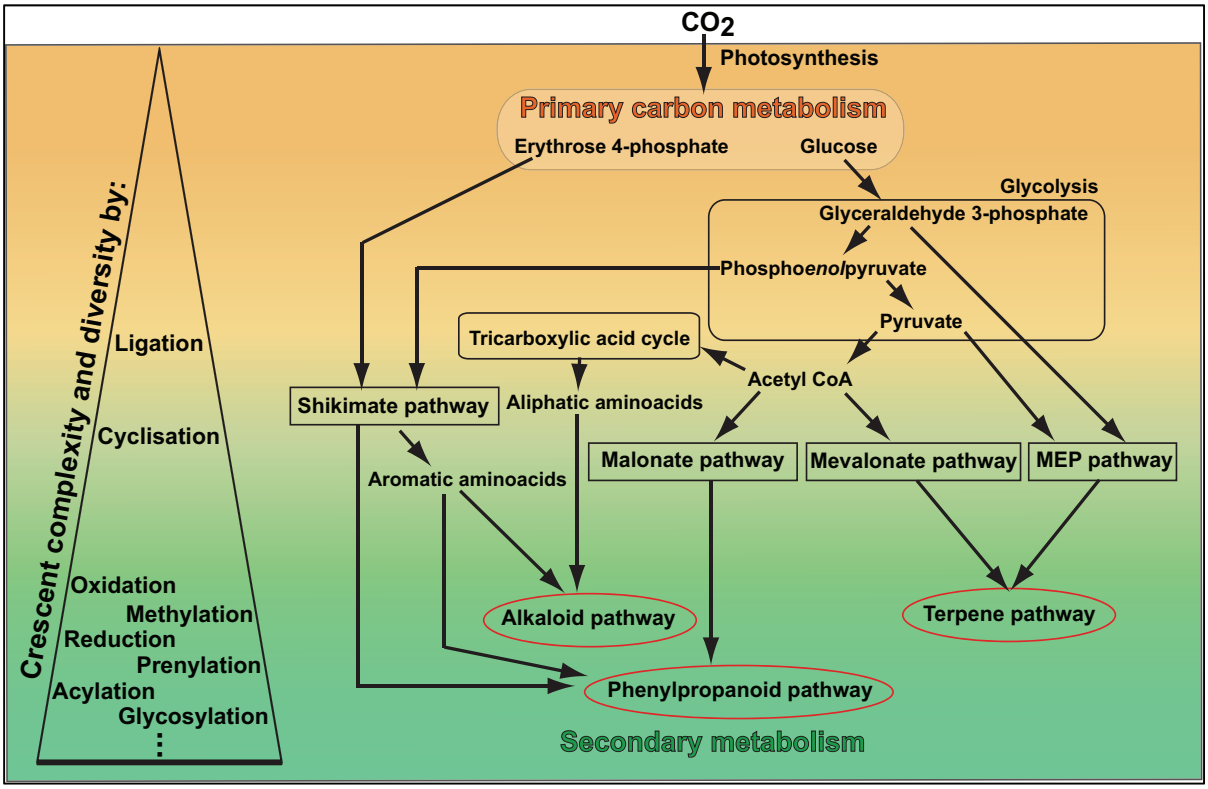


Figure 6. Schematic view of the major secondary metabolites pathways and their interconnections to primary metabolism. Abbreviation: MEP pathway, Methylerythritol Phosphate pathway.

As plants are sessile organisms they have evolved to synthesize a vast number of secondary metabolites to adapt to abiotic and biotic stresses as well as to secure communication among plants. The diversity of metabolites results in several cascades by a combination of ligase, reductases, oxygenases and transferases (Figure 6). Thus, plants produce more than 200.000 secondary metabolites which can be divided into three principal classes according to their biosynthesis pathway and their structure: terpenoid, phenylpropanoid, alkaloid. These various classes have very diverse functions.

To fulfill their function(s) and to avoid toxicity, the accumulation of these secondary metabolites is generally compartmentalized in specific tissues or certain cells types, in which sub-cellular localization is usually tightly controlled. They are often transported from sink cells to other cells, tissues or organs. This transport is specific and strongly controlled for each secondary metabolite and is carried out by membrane transporters, located on plasma membrane or tonoplast (Yazaki, 2005; 2006).

Activation of the secondary metabolite biosynthesis results from an extracellular or intracellular signal perceived by receptors. This leads to the activation or to the *de novo* synthesis of the transcription factors controlling expression of biosynthetic genes involved in the target pathway. Much work is targeted at identification of genes coding for biosynthesis proteins and the transcription factors to increase or decrease the production of the specific secondary metabolites of economical interest (Zhao *et al.*, 2005).

2.1. The terpenoids / isoprenoids ≈25.000 characterized molecules

The name terpenoids was given because the first molecules of this family were isolated from terpentine. The name isoprenoid comes from the fact that their thermal degradation releases isoprene. They form the class of the most diversified metabolites in the plant kingdom. These compounds form by condensation of C5 isoprenoids units and are thus classified according to the number of these units in their skeleton (monoterpenoids: C10; sesquiterpenoids: C15; diterpenoids: C20; triterpenoids: C30). They derive either from the cytosolic mevalonate pathway, or the plastidic MEP (2-C-methyl D-erythritol-4-phosphate) pathway. The terpenoid pathway provides essential compounds for the growth and the survival of the plant: phytohormones such as gibberellins, brassinosteroids, abscissic acid; carotenoids (involved in photosynthesis); ubiquinones (important for cellular respiration); sterols (constituent of the membranes). Moreover terpenoids are involved in resistance to insects, in interactions with other plants, micro-organisms, herbivores or pollinating arthropods. They are also involved in resistance to oxidative stress (Aharoni *et al*., 2005; D'Auria & Gershenzon, 2005). Many molecules with pharmaceutical applications belong to this family, for example taxol (anti-cancer drug).

2.2. **Phenylpropanoids** ≈11.000 characterized molecules.

These compounds represent the most abundant class of secondary metabolites. They have a large panel of functions: they are involved in lignification, pigmentation (anthocyanins), protection against biotic or abiotic stresses, signaling during plant development, fertility (Weisshaar & Jenkins, 1998). For example eugenol, a major compound in the clove essence, is very toxic for the coleopterous insects (Gang, 2005). Most of the aromas (eugenol, vanillin, piperin…), odorous molecules and pigments are phenolic compounds. Many antioxidants as well: sesamin (*Sesamum indicum*), chlorogenic acid, flavonoids (Figure 7)… Some drugs arise from this pathway: podophyllotoxin (*Podophyllum peltatum*: anti-cancer drug), mescaline (*Peyote cactus*: psychoactive compounds)…

	Antioxidants	Sources	Antioxidant activity (mM)
Vitamins	Vitamin C	Fruits and vegetables	1
	Vitamin E	Grains, nuts, oils	2
Flavonoids	Oenin	Black grapes/Redwine	1,8
	Cyanidin	Grapes, raspberries, strawberries	4,4
	Delphinidin	Aubergin skin	4,4
	Quercetin	Onion, apple skin, berries, black grapes, tea, brocoli	4,7
	Kaempferol	Endive, leek, broccoli, grapefruit and tea	1,3
Flavones	Rutin	Onion, apple skin, berries, black grapes, tea, broccoli	2,4
	Luteolin	Lemon, olive, celery, red pepper	2,1
	Chrysin	Fruit skin	1,4
	Apigenin	Celery, parsley	1,5
Hydroxy-cinnamates	Caffeate	White grapes, olive, cabbage, asparagus	1,3
	Chlorogenic acid	Apple, pear, cherry, tomato, peach, aubergine, green tea, robusta coffee	1,3
	Ferulate	Grains, tomato, cabbage, asparagus	1,9
	p-coumarate	White grapes, tomato, cabbage, asparagus	2,2

Figure 7. Relative antioxidant activities of phenylpropanoids compared to vitamins C and E (Based on Rice-Evans *et al.*, 1996).

Antioxidant activity is expressed in concentration of Trolox (Vitamin E derivative) solution having the same antioxidant activity as a solution at 1 mM of the studied solution molecule.

2.3. **Alkaloids** ≈12 000 characterized molecules.

Alkaloids are N-containing heterocyclic compounds, the structure of which varies from very simple to highly complex. The best known examples are nicotine and caffeine. The alkaloids are prevalent in the families of Leguminoseae (lupin) and Solanaceae (tobacco, tomato, potato), but seldom met in the Gymnosperms or Monocotyledons. The alkaloids constitute a significant source of pharmaceutical molecules of interest. For example: quinine which enters in the composition of drugs against paludism, morphine which is the reference analgesic, codeine, another analgesic (both isolated from poppy). The alkaloids are also used to fight uric acid excess (colchicine), as paralyzing/stimulative substance (curare, caffeine), as poisons (strychnine, nicotine), as narcotics (cocaine), as cholinergic (pilocarpine), vasodilator agent (atropine) or anti-cancer drugs (vinblastine, vincristine)…. Moreover, these molecules are important for plant defense against pathogens and herbivores.

Secondary metabolites represent a broad class of natural substances, which, although known as "secondary", are in fact essential for plant adaptation to their natural environment. They are synthesized by the plants to attract pollinators, to attract enemies of herbivores or to defend directly against their environment. Finally, secondary metabolites represent an important source of medicinal products, food additives, and perfumes.

My work focused on phenylpropanoids.

3. The phenylpropanoid pathway

Phenylpropanoid metabolism is specific to the plant kingdom and especially important in land plants (amounts of phenolics in algae and bacteria are minimal). The successful evolutionary plant adaptation to land was achieved largely by a massive formation of phenolic compounds. These molecules represent about 40% of organic carbon in the biosphere (Buchanan *et al.*, 2001).

The shikimate pathway is the entry point to the biosynthesis of phenylpropanoids, all derived from phenylalanine. Phenylpropanoids all share the same carbon skeleton: an aromatic ring and a side chain of three carbons (structure C6-C3) (Figure 8).

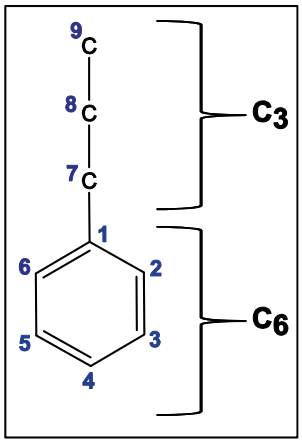


Figure 8. Core structure of phenylpropanoid compounds.

The phenylpropanoid pathway is organized in:

- a core pathway, which begins with phenylalanine to provide coumaroyl CoA via Phenylalanine Ammonia-Lyase (PAL), Cinnamate 4-Hydroxylase (C4H) and 4-Coumarate:CoenzymeA Ligase (4CL).
- several branches leading from 4-coumaroyl CoA to diverse sub-families of molecules, such as flavonoids, stilbenes, coumarins, lignin monomers (Figure 9).

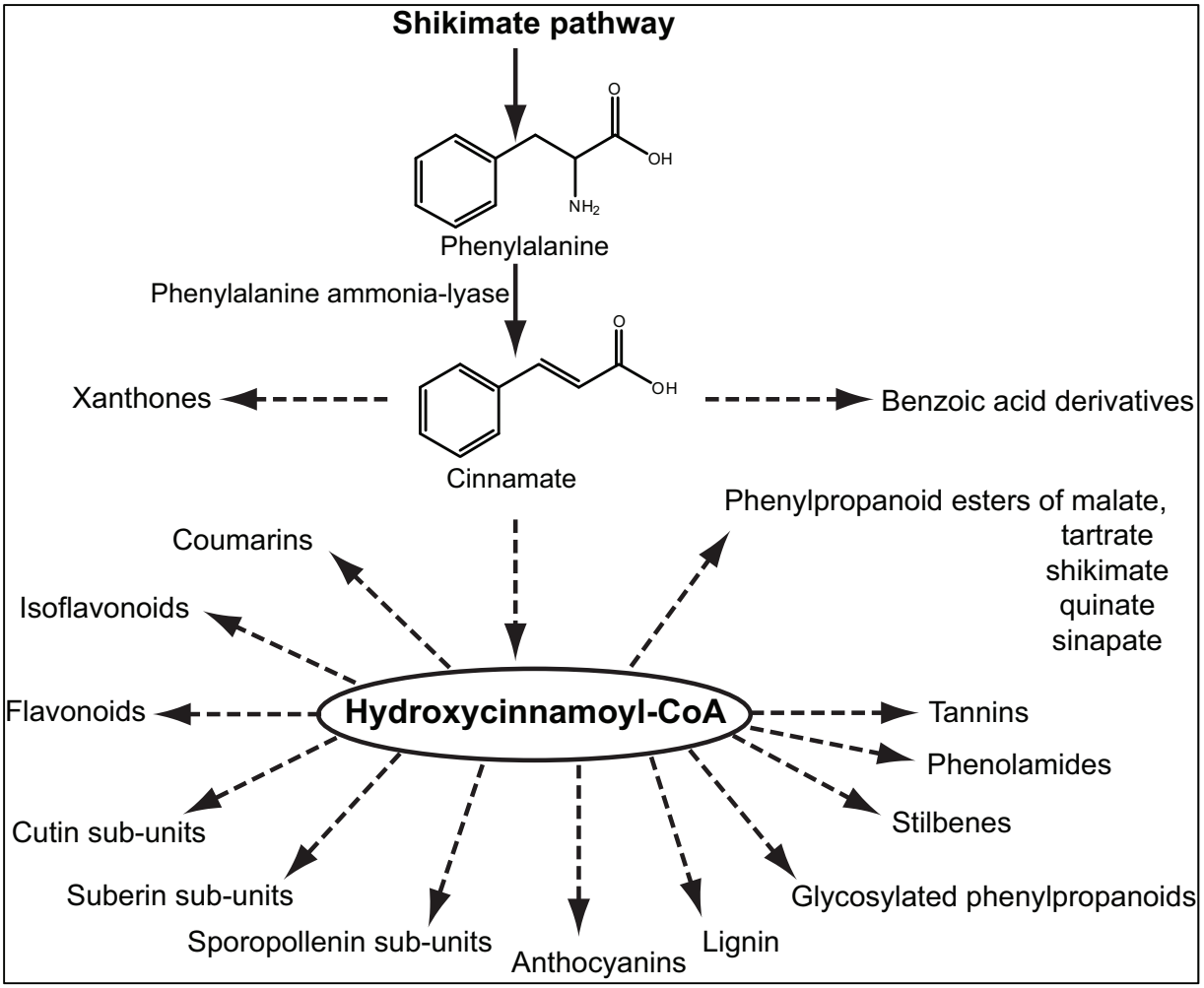


Figure 9. Overview of the phenylpropanoid pathway.

To form the large diversity of phenylpropanoids compounds, the carbon skeleton undergoes, in regio-specific manner, hydroxylation, methylation, glycosylation, acylation, prenylation… by multigenic enzyme families, such as cytochromes P450, *O*-methyltransferases, glycosyltransferases or acyltransferases. Their structural variability reflects the various functions of these compounds.

Phenylpropanoids contribute to all aspects of plant responses to abiotic and biotic stimuli (Figure 10) from signaling of variation of light or mineral environment to mediating pest resistance.

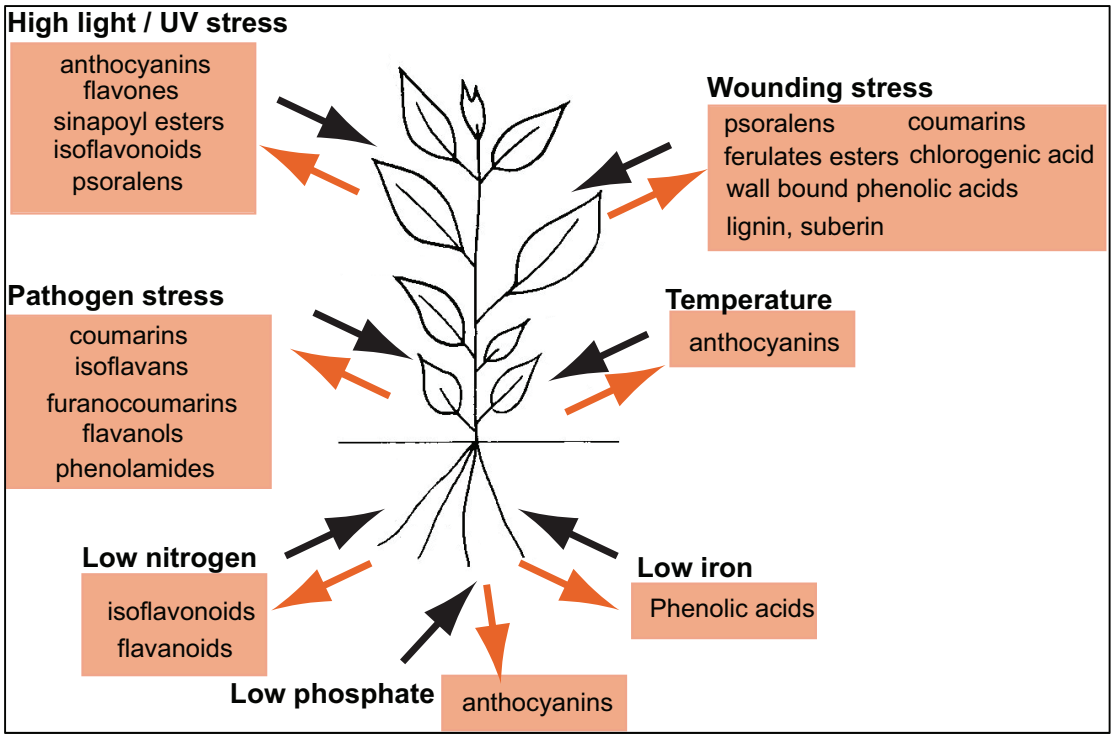


Figure 10. Examples of stress-induced phenylpropanoids (Based on Dixon & Paiva, 1995).

Clearly, phenylpropanoids play an important role in inter and intra-species signaling, as allelochemicals, pigments, flavor and scent, or as signaling molecules in symbiosis. Phenylpropanoids can be classified in different sub-families. A brief description of these sub-families will be given now.

3.1. Hydroxybenzoic derivatives

Derivatives of the benzoic acids (C6-C1), although not having a carbon skeleton in C6-C3, were included in the discussions about the phenylpropanoids, due to their supposed biosynthetic origin via a reduction of the side chain of hydroxycinnamic acids. But, it may not be their only pathway of synthesis. It was recently shown that some *A. thaliana* benzoic acid derivatives are synthesized directly from chorismate by isochorismate synthase (Wildermuth *et al.*, 2001).

Variations in the structure of hydroxybenzoic acids result from hydroxylations and methoxylations of the aromatic cycle (Ribnicky *et al.*, 1998). Four acids seem universal in Angiosperms: *p*-hydroxybenzoic acid, vanillic acid, syringic acid and protocatechuic acids. The first three are components of lignins (Fleuriet & Macheix, 2003; Figures 11 A, B, C).

2-hydroxybenzoic acid or salicylic acid (Figure 11 D), known for long as a key component of plant innate immunity, is a major agent of systemic acquired resistance (SAR) to pathogens and other stresses (Ribnicky *et al.*, 1998). In the fruits, salicylic acid is present in very low amounts in Solanaceae (tomato, eggplant, pepper …), in Cucurbitaceae (melon, cucumber …) and other species (kiwi, grape …).

Benzoic acids are often present as conjugated and complex structures such as hydrolysable tannins or simple molecules combined with sugars or organic acids.

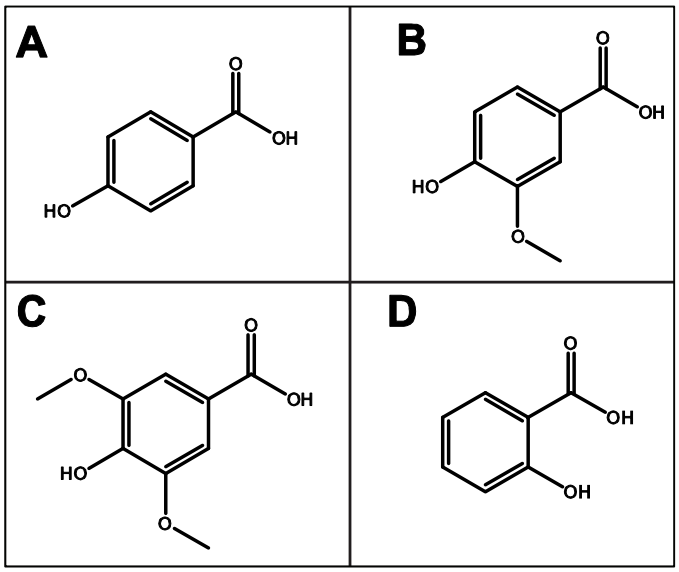


Figure 11. Chemical structures of four hydroxybenzoic acid derivatives.

A. *p*-hydroxybenzoic acid;
B. vanillic acid;
C. syringic acid;
D. salicylic acid.

3.2. Stilbenes, tannins, flavonoids, isoflavonoids and their derivatives

Approximately 9.000 flavonoids were characterized and 300 stilbenoids are known, present in Bryophytes, Pteridophytes, Gymnosperms, and Angiosperms. They all share the same carbon backbone in C15 (C6-C3-C6) resulting from the condensation of *p*-coumaroyl CoA with three molecules of malonyl-CoA. Their initial skeletons are formed either by chalcone synthase for flavonoids and isoflavonoids or by stilbene synthase in the case of stilbenes (Shirley, 1996).

Flavonoids are found in all land plants. Isoflavonoids and stilbenes are present only in specific families, the number of which is limited. For examples isoflavonoids are found mainly in Fabaceae (Leguminoseae) (Dixon *et al.*, 2002). Various functions have been described for these classes of molecules. Flavonoids are involved in UV protection, transport of auxin, seed dormancy, pigmentation, defense against micro-organisms and insects, but also in signaling during reproduction or symbiosis, and most probably in allelopathy (D' Auria & Gershenzon, 2005; Shirley, 1996). For example, medicarpin is a defense compound (*Medicago sativa*; Figure 12 A), pelargonidin (*Pelargonium*) and cyanidin (in red berries) are attractant pigments, kaempferol is a UV-B protectant (*Glycine max*; Figure 12 B), proanthocyanidin an antifeedant (*Sorghum*). Stilbenes, like resveratrol (Figure 12 C), are also phytoalexins of vine and groundnut. During symbiosis with *Rhizobium*, flavonoids are essential in chemotactism.

Stilbenes and isoflavonoids are limited to a few species, but have attracted major interest due to potential health-promoting effects on mammals (Barnes & Prasain, 2005). For example, various health-promoting activities have been reported for proanthocyanidin (non hydrolizable tannins): like epigallocatechin-gallate (EGCG) (Rezai-Zadeh et al., 2005), which seems to slow down the progression of cerebral amyloidosis in Alzheimer disease, or combretastatin which apparently exhibits antineoplastic activity.

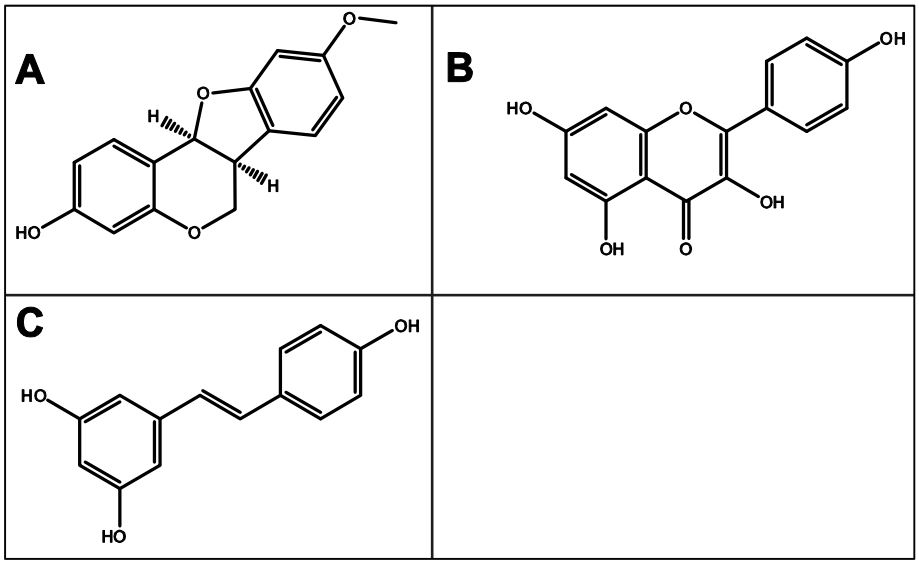


Figure 12. Chemical structures of one flavonoid, one isoflavonoid and one stilbene.
A. medicarpin (isoflavonoid); B. kaempferol (flavonoid); C. resveratrol (stilbene).

3.3. Hydroxycinnamic derivatives

These compounds all derive from cinnamic acid. Four molecules are mainly found: *p*-coumarate, caffeate, ferulate and sinapate (Figures 13 A, B, C, D). For each of these derivatives, the presence of a double bound on the side chain leads to the existence of the 2 isomers *Z* and *E*. However, these compounds are naturally mainly found in *E* configuration.

Hydroxycinnamic acids are present in all land plants, but some are specific to some species or families (Dixon *et al.*, 2002). Chlorogenic acid, for example, only accumulates in some Solanaceae (tobacco), Rosaceae (apple) and Rutaceae (coffee). Hydroxycinnamic acids and their derivatives are rarely found in free form but are esterified in the plant, except upon brutal conditions of extraction, physiological disturbances and contamination by microorganisms.

Because of their abundance and diversity, hydroxycinnamic acid derivatives play very important roles: some are involved in antioxidant reactions, UV-screening, defense against pathogens, plants/micro-organism interactions. Compounds such as acetosyringone (Figure 13 E) induce the virulence genes of *Agrobacterium tumefaciens* (Lee *et al.*, 1995).

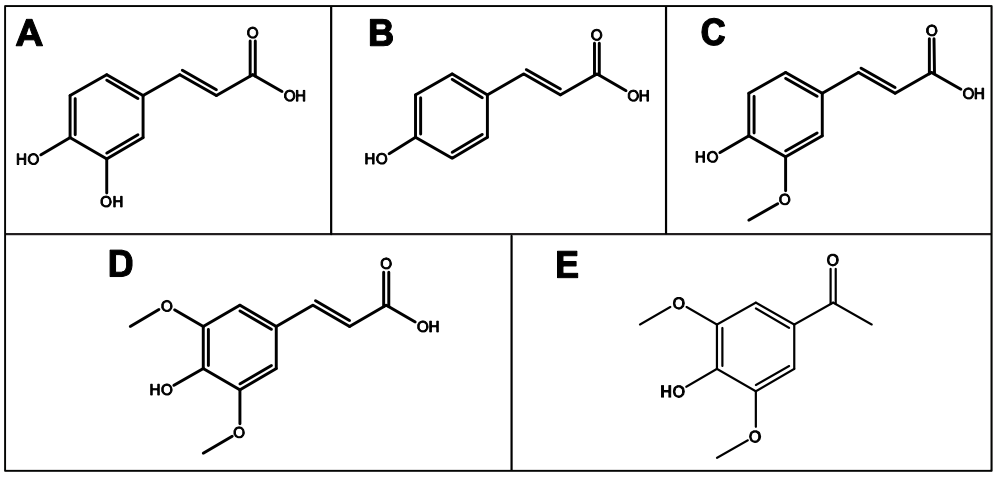


Figure 13. Chemical structures of five hydroxycinnamic acids derivatives.
A. *p*-coumarate; B. caffeate; C. ferulate; D. sinapate; E. acetosyringone.

However, the most important function of these compounds is in cell wall structure: they are building-blocks for lignin synthesis and they influence the physicochemical properties of cell walls (Boerjan *et al.*, 2003).

From the pharmaceutical point of view, it was also shown, *in vitro*, that hydroxycinnamic acids and their conjugates have important antioxidant capacities (Rice-Evans *et al.*, 1996). Indeed, ingestion of these molecules via coffee or fruits, for example, seems to constitute an important way to fight free radicals for health benefit.

3.4. Esters of tartaric, shikimic, quinic, sinapic acids

The esters of tartaric acid were initially detected as a chicoric acid (dicaffeoyl *L*-tartaric acid) in endive (*Cichorium intybus*). In grapes, tartaric acid esters are mainly found as caftaric acid (monocaffeoyl *L*-tartaric acid), coutaric acid (mono-*p*-coumaroyl *L*-tartaric acid) and fertaric acid (monoferuloyl *L*-tartaric) (Herrmann, 1989).

Shikimic esters, 5-*O*-cafeoylshikimic acid for example, are usually not frequent in plants, but found in very large amounts during ripening of dates (*Pheonix dactylifera*), from where comes the name of dactylifric acid.

The most common quinic acid ester is chlorogenic acid (5-*O*-caffeoyl quinic acid; Figure 14 A), in tobacco for example, but found also in very low amount in *Arabidopsis thaliana*. Part of the introduction of chapter II will be devoted to this compound. Two chlorogenic acid isomers exist in plants: the neochlorogenic acid (3-*O*-caffeoyl quinic acid) and cryptochlorogenic acid (4-*O*-caffeoyl quinic acid). Chlorogenic acid is a soluble phenol that plays the role of antioxidant. Other quinic esters are also found in some plants: the isomers of the *p*-coumaroyl quinic acid in apple and of 5-*O*-feruloyl quinate in tomato.

Finally, sinapic acid esters serve as energy-rich metabolites during seed maturation and germination (Mock & Strack, 1992). Sinapoyl malate was described as an UV screening molecule in *A. thaliana* (Landry *et al.*, 1995; Figure 14 B) and is a caracteristic compound from Brassicaceae.

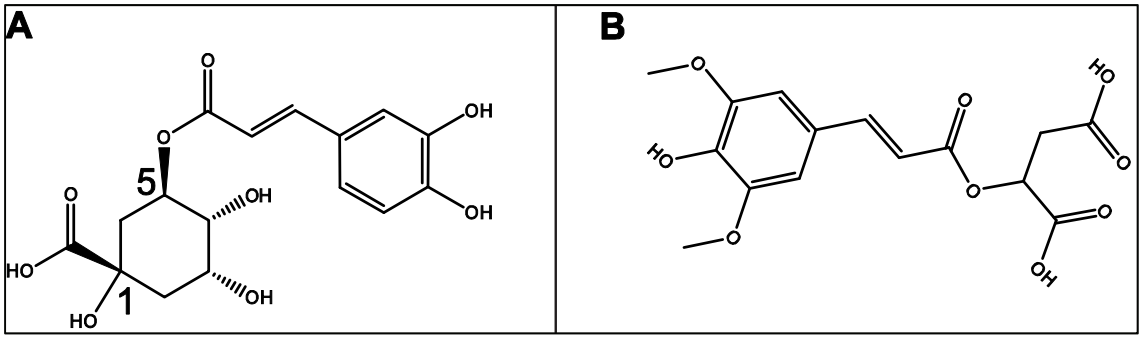


Figure 14. Chemical structures of chlorogenic acid (A) and sinapoyl malate (B).

3.5. Glycosylated derivatives

Large amounts of glycosides were identified in many plants (Mølgaard & Ravn, 1988). They can be classified in two groups.

- Glucose esters (as *p*-coumaroyl glucose; Figure 15 A) are derivatives with an ester bond between the carboxylic function of the phenolic acid and a glucose hydroxyl.
- The glucosides have one of the phenolic hydroxyls of the hydroxycinnamoyl molecule forming an glycosidic bound to glucose (as glucoside of *p*-coumarate; Figure 15 B).

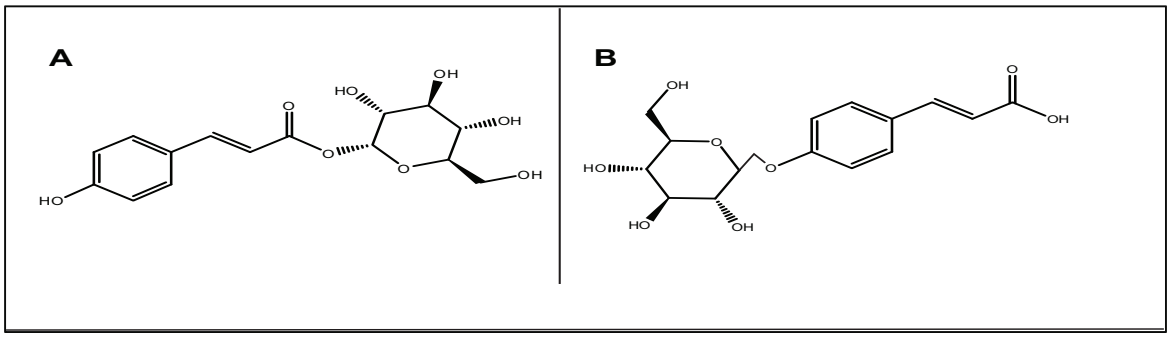


Figure 15. Chemical structures of *p*-coumaroyl glucose (A) and glucoside of *p*-coumarate (B).

Glycosylation of hydroxycinnamic acids to form both glycosides and glucose esters is catalyzed by a group of enzymes named glycosyltransferases (Bowles *et al.*, 2005). An important role of glycosylation was reported in stabilization, modulation of reactivity (by blocking reactive hydroxyl groups), storage or transport of secondary metabolites as phenylpropanoids (Yazaki, 2005; Payyavula *et al.*, 2009). Various volatile phenolics are stabilized by glycoconjugation and released after cell damage by endogenous glycosidases. For example, in tomato fruit, many volatile compounds are bound as glycosides, and represent an aroma reserve that can be liberated when cell compartmentation is destroyed, as happens on consumption of fresh fruits or during late ripening stages. This glycosylation process controls the emission of phenylpropanoids volatiles from tomato fruit (Tikunov *et al.*, 2010). Glycosylation can also influence the organoleptic properties of a compound. In *Citrus*, glycosyl moiety conjugated to naringenin (flavanone) modulate the bitter flavor of the fruit (Frydman *et al.*, 2004)

Phenolic glycosides from strawberry were also reported as antioxidant and antiproliferative of cancer cell (Zhang *et al.*, 2008).

3.6. Coumarins

Approximately a thousand of coumarins are characterized today. Coumarins derive from *O*-hydroxycinnamic acid by *cis-trans* isomerization and lactonization. This reaction was recently shown to be catalyzed by a 2-oxoglutarate-dependent dioxygenase in *A. thaliana* (Kai *et al.*, 2008).

Coumarins are present in almost all the plant taxa, but are found in large amounts in many Monocotyledons, like Poaceae and Orchids, and in some Dicotyledons, such as Rutaceae (lemon) and Apiaceae. In plants, coumarins can occur in seed coat, fruits, flowers, roots, leaves and stems. The highest concentrations are generally found in fruits and flowers. Their roles in plants seem to be mainly defense-related, given their antimicrobial, antifeedant, and germination inhibitor properties.

Coumarins can be classified in three groups: simple coumarins, furanocoumarins (condensation of a furan cycle on a coumarin core) and pyranocoumarins. The furanocoumarins are intensively studied due to their therapeutic interests. Psoralen is used in the healing of skin diseases, such psoriasis and eczema (Figure 16 A). There is also a long list of toxic compounds from this family of compounds. 8-methoxypsoralen present in leaf tissues of *Heracleum mantegazzianum*, causes photophytodermititis and warfarin (a synthetic coumarin) is used as rodenticide, causing internal bleeding in mammals.

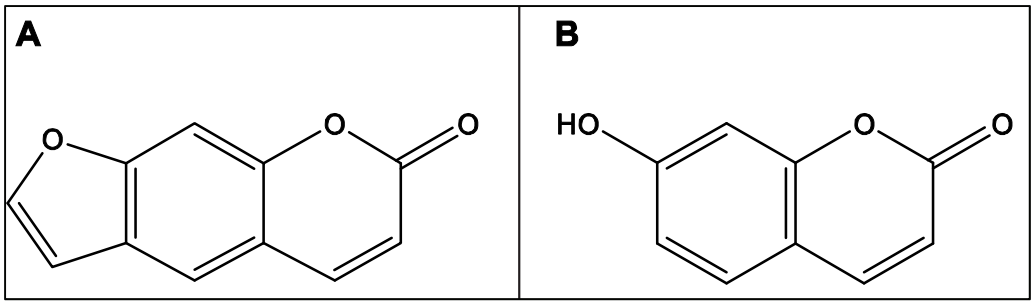


Figure 16. Chemical structures of one furanocoumarin and one coumarin.
A. psoralen; B. umbelliferone.

3.7. Lignins, suberin, cutin and sporopolenin

The four types of biopolymers are very different but nevertheless share complex and highly rigid structures based on two types of covalently bound building blocks, aliphatic (oxygenated fatty acid) and aromatic (phenolic derivatives). For cutin, suberin and sporopollenin, the contribution of the latter is apparently lower than in the case of lignin (Pollard *et al.*, 2008). Only 1% of the cutin constituents are phenolics, and principally ferulates while up to 10% of the suberin is constituted by phenolics, mainly ferulates, but also coumarate, sinapate, caffeate (Pollard *et al.*, 2008).

During normal plant growth, the cell walls of numerous cells undergo modifications to acquire specific functions. For example, the cell walls of xylem vessels are lignified to support the long distance transport of water. In a similar way, the epidermal tissues of underground parts of the plants (roots, tubers, stolons), endoderm and bark cells are suberized to form protection barriers. Suberin plays roles in the waterproofing of the root Casparian strip, in apoplastic water transport and defense against pathogens. From an evolutionary perspective, suberization was of great importance in plant kingdom and may have preceded lignification (Buchanan *et al.*, 2001). Lignification and suberization are specific to some tissues and specific cells (Bernards & Razem, 2001).

Cutin is also a structural biopolymer covering epidermal cells. It is a component of the cuticle, characterisitic of the aerial part of the plants. This biopolymer acts as barrier to control movement of gases, water and solutes and to impair pathogen attacks. The ether-linked phenolics were assumed to rigidify the biopolymer and prevent depolymerization. Linkages have been also proposed between suberin and the cell wall through the polymerized phenolic domain composed of hydroxycinnamic acid (Pollard *et al.*, 2008).

Sporopollenin, found in pollen wall, is essential to prevent pollen desiccation and to provide wall pattern and strength. Sporopollenin is a major constituent of the pollen exine. It contains building blocks similar to those found in cutin and suberin (phenylpropanoids and hydroxylated fatty acids). Compared to cutin and suberin, sporopollenin is more resistant to chemical degradation. Although advances in chemical analysis have enabled partial identification of its constituents, the exact structure of sporopollenin is still not known. The function of phenylpropanoids in sporopollenin, besides the contribution to UV-light protection, may also reside in protection against pathogens (Vogt, 2010). Sporopollenin is the most resistant known biopolymer and is found intact in fossils.

Lignin results essentially from the polymerization of three phenolic units. Formation, components and roles will be detailed in a next paragraph of this introduction.

3.8. Hydroxycinnamic acid amides (Phenolamides)

Phenolamides, at the intersection of phenylpropanoids and polyamines will be treated in the review introducing the chapter II of this thesis report.

Briefly, the existence of phenolamides has been described for several decades and their widespread occurrence, predominantly in flowers and pollen grains, has been attributed to UV-protection, fertility, structure, as well as defense (Walters, 2003; Edreva *et al.*, 2007).

The recent identification of phenylpropanoid polyamine conjugates in different organs of wild-type *Arabidopsis* and insertion lines by LC–MS analysis was accompanied by simultaneous identification of specific genes and enzymes in the phenolamide formation. In the tapetum of *Arabidopsis* flower buds, a new pathway for the biosynthesis of hydroxycinnamic spermidine conjugates was established (Matsuno *et al.*, 2009; Grienenberger *et al.*, 2009).

3.9. Exotic phenylpropanoids

These compounds are extremely variable, with a diversity caused by the evolution of specialized enzymes in some plant species. In these plants, the core phenylpropanoid pathway is not different from the pathway from model plants but with specific branches and enzymes that were recruited for new biosynthetic steps (Vogt, 2010).

Several of these phenylpropanoid-derived metabolites are restricted to few species or families. Those include the curcumins (in *Oryza sativa*, *Curcuma longa*), benzophenones (in *Hypericaceae* and *Maloideae*), biphenyls (in *Maloideae*), or phenylphenalenones. They are considered as phytoalexins. Some are reported as having pharmacological and health-promoting effects, which makes them an interesting target for biosynthetic analyses.

3.10. Economical interests for the study of the phenylpropanoids pathway

Plant phenylpropanoids contribute to the agronomic, industrial and nutritional performances of agricultural and forest resources (Vogt, 2010).

There are many epidemiological evidences that diets rich in fruit and vegetables can reduce the incidence of diseases such as cardiovascular diseases, diabetes, cancer and stroke, and in a more efficient way than expensive dietary supplements (Liu, 2004; Crozier *et al.*, 2009). These protective effects are attributed, in part, to phenylpropanoids. Those are present in high amount in tea, red wine, coffee and chocolate for example, and in lower amounts, in vegetables and other fruits. These protective roles are undoubtedly due to antioxidant activity by scavenging hydroxyl radicals, superoxide radical anions, several organic radicals, peroxyl radicals, peroxinitrites and singlet oxygen. The highest antioxidant activity was observed for caffeic acid (Shahidi & Chandrasekara, 2010).

In plants under stress conditions, phenylpropanoids biosynthesis may represent an alternative way for photochemical energy dissipation that has the benefit to also enhance the antioxidant capacity of the cell (Grace & Logan, 2000).

A complex species-specific bouquet of compounds, including phenylpropanoids, is emitted as floral attractant to pollinators, as defense against pathogens and predators. It can also have allelopathic effects on other plants (Dudareva *et al.*, 2004). Engineering phenolic volatiles has a major impact on flower scent, on the taste of fruits, vegetables and spices.

Finally, the cell wall lignin polymers strengthen the stems of major crops, and lignin composition and content influences tree wood or forage quality. Engineering of lignin quality and content is thus a major challenge for the production of biofuels, for the improvement of paper pulping and of forage quality (Ralph *et al.*, 2004; Vanholme *et al.*, 2008; Simmons *et al.*, 2010).

The phenylpropanoid metabolism leads to the formation of a broad array of compounds with economical interest for plants or Humans. In consequence, engineering this highly diverse pathway offers the opportunity to enhance valuable properties for a large number of applications.

4. The cell wall

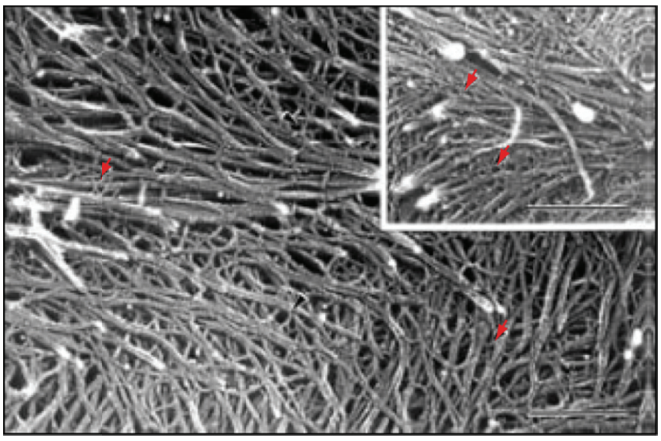


Figure 17. Electron micrograph of the outer cell walls.

Cellulose microfibrils and their cross-links are indicated by red arrowhead. Scale bars, 200 nm (from Fujino *et al.*, 2000).

Plant cells are encased in rigid walls composed of cellulose (Figure 17), hemicellulose, pectin, proteins, and sometimes lignin, which vary in amount depending on cell types, and even in different microdomains of the wall of a given cell. (Zhong & Ye, 2007). A number of factors, including hormones, cytoskeleton, proteins, phosphoinositides, and sugar supply, have been implicated in the regulation of cell wall biosynthesis or deposition (Zhong & Ye, 2007).

The walls function as plant "exoskeleton" and define cell shape, and ultimately organ and whole plant morphology. They not only regulate cell growth and provide structural and mechanical support to plants, but also act as a physical barrier to biotic and abiotic stresses.

Plant cell walls are generally categorized in two types: primary and secondary walls (Buchanan *et al.*, 2001). Primary walls are formed during cytokinesis and further modified during cell expansion and secondary walls are laid down after cessation of cell expansion in some specialized cells. The secondary walls can be modified to answer special needs of certain cell types, such as water conducting tracheary elements, guard cells, and endodermis… These modifications can be for example lignin deposition.

On the basis of the wall characteristics, plant cells are classified in three basic types: parenchyma, collenchyma, and sclerenchyma. Parenchyma and collenchyma cells have only primary walls and sclerenchyma cells contain both primary and secondary walls. Sclerenchyma can be further divided into sclereids, fibers, and tracheary elements.

Plant cell walls are important in human life, providing major dietary fibers and essential raw materials for textile, lumber, paper, and potentially for biofuels.

5. Lignins

Lignins are structural biopolymers, rigid and resistant, formed of a complex of aromatic compounds, and characteristic of the Tracheophytes. These polymers are found mainly in the pecto-cellulosic wall of the cells having mechanical or conducting functions.

After cellulose, lignins are the most abundant known organic carbon source, accounting for as much as 30% of the biosphere (Sommerville *et al.*, 2004).

5.1. Role of lignins

It is in the middle Silurian, approximately 420 million years ago, that the progressive withdrawal of the oceans has supported the adaptation of aquatic and photosynthetic organisms to the terrestrial environment. This adaptation required transformations of the vegetative and reproductive apparatus of the plants mainly to face the absence of water and to grow in the aerial medium, fighting terrestrial attraction. During evolution, this transition, in its structural and physiological complexity, is regarded as being one from the most significant events of the terrestrial plant history.

For several million years, terrestrial plants remained extremely small and of simple shape. Then, the vegetative apparatus gradually evolved into a system of crawling axes and raised shape, having structurally, functionally specialized and well differentiated tissues. The construction of an erected shape of large size required water transport in plant. This requirement led to the formation of specialized tissues, the vascular system, which ensures the transport of water starting from the interface plant-ground towards the aerial parts of the plant. Within the vascular system, two types of conducting tissues allowed irrigation of all the parts of the plant: xylem vessels transporting water and minerals (crude sap) and phloem vessels responsible for the distribution of organic materials resulting from photosynthesis (elaborate sap).

The conductance of the tracheids was considered as being 10^7 fold superior than that of equivalent live cells with their cytoplasmic contents. This physiological property is directly connected to the synthesis and deposit of a complex structural polymer, rigid and hydrophobic: lignins.

The appearance of suberin, and then of lignins was determining for the colonization of land by the plants (Buchanan *et al.*, 2001). Lignin is therefore an essential structural polymer of the secondary walls of specialized plant cells, like tracheids, vessels and sclereids. It ensures critical functions in the plant stiffness and vascular tissue impermeability (Boerjan *et al.*, 2003).

Cell walls impregnated with lignins (known as lignified) are found around all cells of tissues required for rigidification of the plant (collenchymas) or for the transport of water and mineral (xylem). In general, the lignified cells that became waterproof lost their cytoplasm and acquired their function once dead.

Finally, lignin synthesis is also a defense barrier. Lignin deposition and changes in lignin composition contribute to adaptation to various abiotic and biotic (Moura *et al.*, 2010).

5.2. Chemical nature of lignins

Lignins derive from the oxydative polymerization of hydroxycinnamoyl alcohols called monolignols (Boerjan *et al.*, 2003; Lewis & Yamamoto, 1990). The monomeric units of lignin are *p*-Hydroxyphenyl (H), Guaiacyl (G) and Syringyl (S) units, respectively derived from *p*-coumaroyl, feruloyl and synapoyl residues (Douglas, 1996; Figure 18). The monolignols differ by their degree of methoxylation.

The amount and composition of lignins vary according to the taxonomic group, to the plant species, to the analyzed tissues, to the cell types (and even to the part of the cell wall considered), to the developmental stage, to the environmental conditions and the stressors. They also vary with the season and the physical constraints due for example to wind, stem misalignment from its vertical axis, or in response to wounding.

The G unit is predominant in Gymnosperms lignins, whereas subunits G and S are the most abundant in Angiosperms. The H units are preponderant in monocotyledons and are present only in small amounts in dicotyledons.

Three hydroxylation steps are necessary for the formation of monolignols. These reactions are catalyzed by three different P450s: Cinnamate 4-Hydroxylase or C4H (CYP73), Coumaroyl ester 3'-Hydroxylase or C3'H (CYP98) and Ferulate 5-Hydroxylase or F5H (CYP94) (Humphreys & Chapple, 2002).

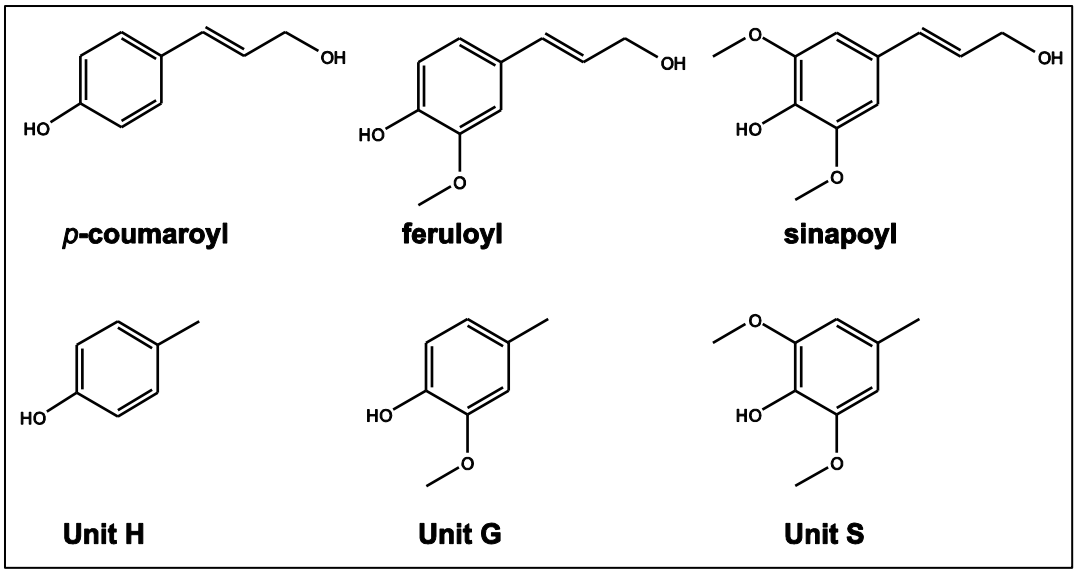


Figure 18. Monolignols subunits and their precursors.
***p*-Hydroxyphenyl (H), Guaiacyl (G) and Syringyl (S) units, respectively derived from *p*-coumaroyl, feruloyl and synapoyl residues**

5.3. The chemical bonds

The synthesis of lignin is carried out by the dehydrogenative polymerization of the three different monolignols (Figure 19). The polymer takes up the shape of a complex three-dimensional network in which the monomeric subunits are connected by ether (C-O-C) and carbon-carbon bonds.

The ratio of S-to-G units in lignin dictates the degree and nature of polymeric cross-linking. Increased G content leads to more highly crosslinked lignin featuring a greater proportion of carbon-carbon bonds, whereas S subunits are typically linked through more labile ether bonds at the 4-hydroxyl position.

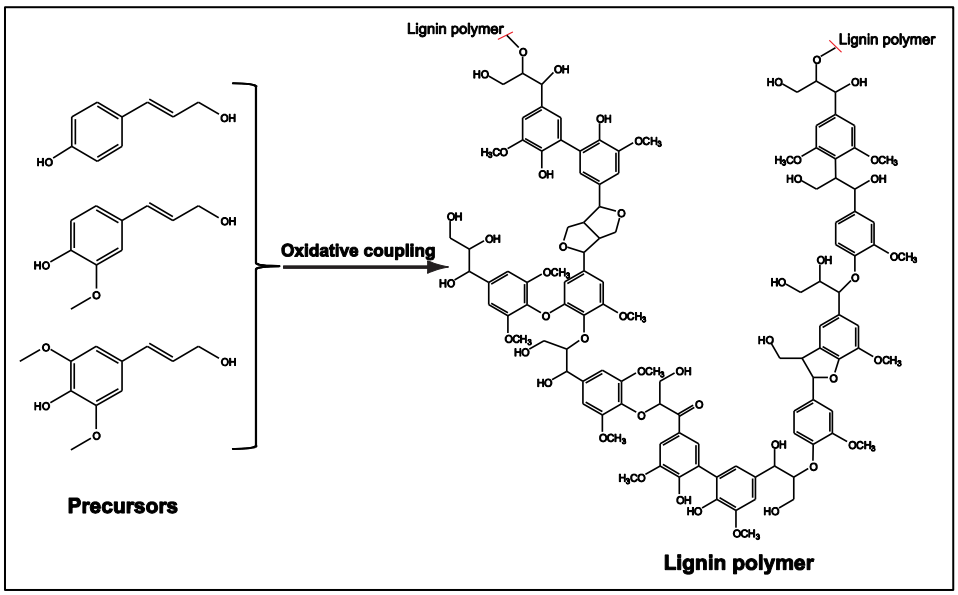


Figure 19. Polymerization of the monolignols precursors to form lignin.

Lignin is constituted of phenolic units but it can also establish bonds with other macromolecules such as polysaccharides, or hydroxycinnamoyl acids via ester or ether bonds (Figures 20 A, B) or be connected to glycoproteins and tannins. Phenolic polymers can be cross-linked to the feruloylated arabinoxylanes for example.

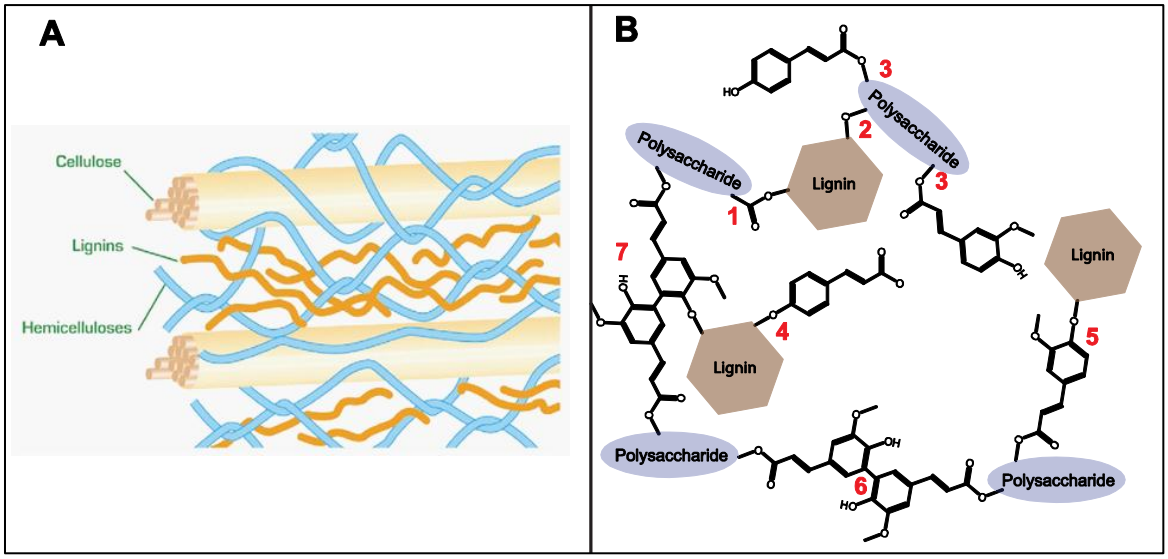


Figure 20. Lignin chemical bonds.

A. Schematic representation of the lignified secondary wall. In addition to cellulose, lignins and hemicelluloses, other cell wall constituents of minor abundance, including proteins and phenolics, are not indicated (From Boudet *et al.*, 2003).

B. Schematic possible bonds. 1- direct ester linkage, 2- direct ether linkage, 3- hydroxycinnamic acid ester, 4- hydroxycinnamic acid ether, 5- ferulic acid bridge, 6- dehydrodiferulic acid diester bridge, 7- dehydrodiferulic acid diester-ether bridge (Based on Buchanan *et al.*, 2001).

Lignin can also react with various exogenous molecules, such as pollutants and pesticides.

5.4. Lignins deposition

The deposit and the modification of the compounds present in the plant cell wall play crucial roles during plant growth and development. Indeed, two types of plant cell walls can be distinguished: the primary wall and the secondary wall that are formed at different stages of cell growth.

The primary walls are settled during cell multiplication. They are necessary for the mechanical stability of the cells and they confer a sufficient extensibility to allow cellular expansion without risk of plasma membrane rupture in turgescent cells. These primary walls are mainly made up of polysaccharides, namely cellulose, hemicelluloses and pectins, interacting to form what is called the extracellular matrix.

Lignin deposition is associated with the formation of the secondary wall, i.e. when the cell reaches its final size (Figure 21). When cell growth is stopped, the mechanical properties of the wall change due to the deposition of a new matrix primarily made up of cellulose and hemicellulose impregnated with lignin.

The secondary wall is thicker, more resistant to mechanical constraints and more hydrophobic due to the enrichment in phenolic compounds (Bolwell, 1988; Iiyama *et al.*, 1994).

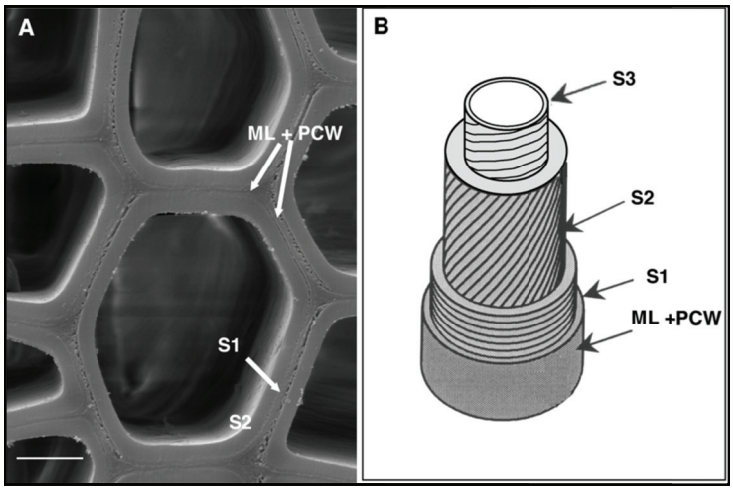


Figure 21. Cell wall deposition and its identified layers.

A. Scanning electron microsocopy.

B. Scheme of the cell wall layers. ML: Middle lamella, PCW: Primary cell wall, S1, S2, S3: layers of the secondary cell wall. Lines on the scheme indicate preferred orientation of the cellulose microfibrils. (Déjardin *et al.*, 2010).

There is no generic lignin but lignins, with diverse compositions and chemical bonds, depending on the plant and due to the heterogeneous deposition. Thus, lignin deposition seems to be a very well ordered process (Déjardin *et al.*, 2010), but still not totally understood.

5.5. Evolution of the lignin pathway

Among species investigated, the complete lignin biosynthesis pathway first appeared in Pteridophytes (Figure 22). The expansion of the gene families mostly occurred after the divergence of Monocots and Dicots.

Emergence of the monolignols biosynthetic pathway occurred by a progressive recruitment of different enzymes from the primary metabolism, followed by their specialization. For example, phenylalanine ammonia-lyase (PAL), the enzyme that converts phenylalanine to *t*-cinnamic acid at the entry point of the phenylpropanoid pathway, is homologous to histidine ammonia-lyase (HAL). This last is involved in histidine degradation in primary metabolism. PAL exhibits structural conservation with HAL, and mirrors HAL in catalytic mechanism (Weng & Chapple, 2010; Xu *et al.*, 2009). Horizontal gene transfers from symbiotic bacteria or fungi also seem to play an important role in the evolution of this pathway.

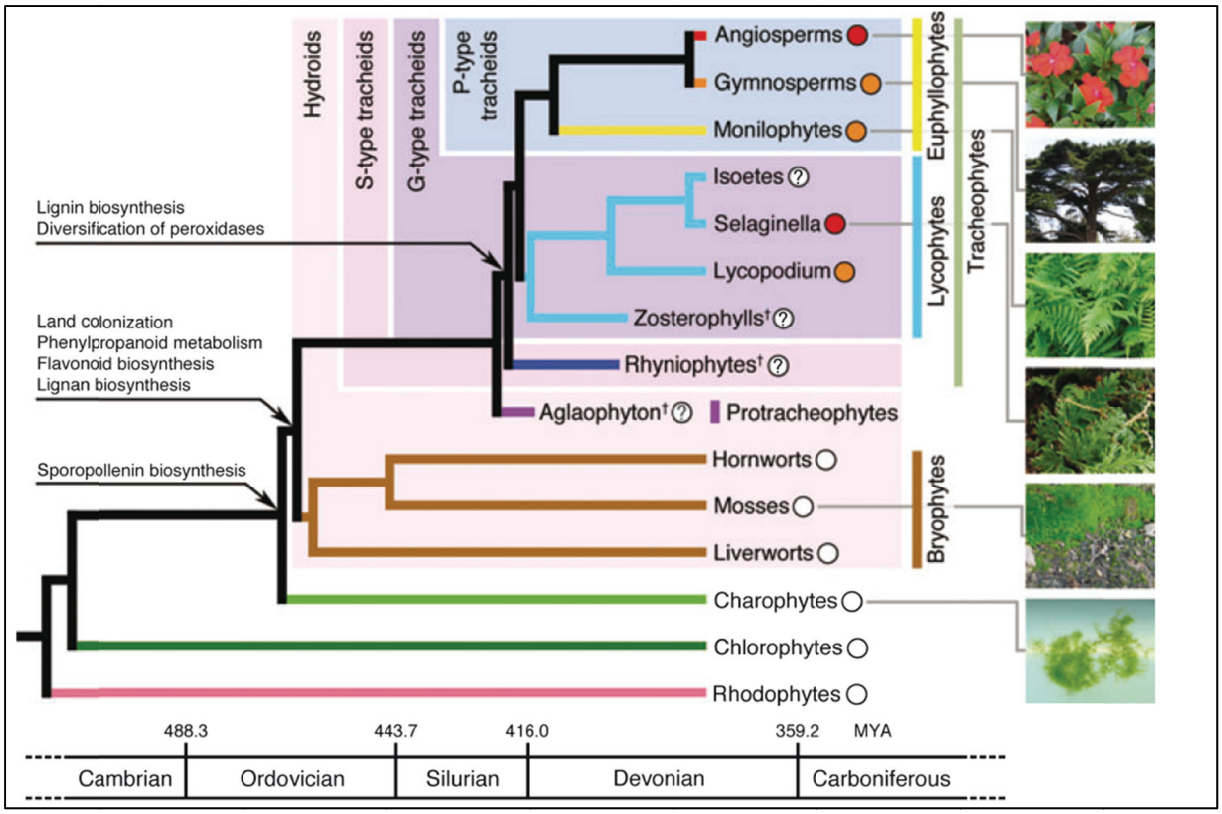


Figure 22. Plant phylogenetic tree marked with the major milestones of evolution of lignin biosynthesis. The distribution of lignin and its monomeric composition across major plant lineages are indicated by a circle at each branch. Open circle, no lignin; orange circle, presence of H and G lignin; red circle, presence of S lignin in addition to H and G lignin; circle with question mark, unknown. †Extinct lineage. (From Weng & Chapple, 2010).

5.6. Economical interests of lignins

Lignin derivatives have a great economical interest, in particular for paper industry, and for the control of the fodder digestibility in agriculture. Lignin is an essential polymer of the wall of all the vascular plants. It is the second most important biopolymer, after cellulose. Thus, it plays a major role in the carbon cycle in the biosphere. The study of this pathway is fundamental.

Of all the components of wood, lignin is that which produces the most energy, so it is strongly required in the firewood. However, the presence of lignin has adverse consequences for the use of biomass. Indeed, lignin is not digested by ruminants. Fodder feeding with reduced lignin improves energetic efficiency and digestion. Moreover, lignin is a problem for the paper manufacturing. Indeed, paper pulp is prepared from cellulose fibers.

Modification of the content and/or structure of lignin would thus have economical consequences. The first observation referring to a natural modification of the biosynthesis of lignin was reported 80 years ago, by Jorgensen. He described, in 1931, the natural mutant *brown midrib* (bm1) in corn affecting the expression of the cinnamyl alcohol dehydrogenase (CAD) (Anterola & Lewis, 2002). In the last fifteen years, many strategies based on genetic manipulation of lignin content and composition have been tested. Many enzymatic steps in the phenylpropanoid metabolism can constitute targets for lignin modification. Results were obtained with PALs, C4H, 4CLs, and C3'H, in the upstream pathway, or, with OMTs and F5H. The enzymes more specifically involved in the biosynthesis of monolignol, namely Cinnamoyl CoA Reductase (CCR), Cinnamyl Alcohol Dehydrogenase (CAD), peroxidases and laccases were also investigated. Perturbation of lignin synthesis was also observed in studies focused on transcription factors, but transcriptional regulation of lignin biosynthesis is still not fully understood (Zhong & Ye, 2009).

Drastic impacts of the modifications in lignin synthesis were reported in several cases. For example:

- In *Arabidopsis* plants silenced for hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT), lignin repression leads to the re-direction of the metabolic flux into flavonoids through chalcone synthase activity (Besseau *et al.*, 2007). Correlation with reduction in plant growth initially attributed to flavonoid accumulation was later shown to depend on inhibition of monolignols biosynthesis (Li *et al.*, 2010).
- Overexpression of F5H, also termed coniferylaldehyde hydroxylase, leads nearly exclusively to change to syringyl lignin in poplar (Stewart *et al.*, 2009),
- Down-regulation of F5H, results in lignin essentially composed of G units in *Brassica napus* seeds (Nair *et al.*, 2000)
- Nitrate supply contributes to the transcriptional regulation of the phenylpropanoid, and nitrate depletion inhibits large sectors of the phenylpropanoid pathway (Fritz *et al.*, 2006)

The effects of knockout mutations in the genes involved in phenylpropanoid metabolism and lignin monomer composition have been analyzed and reviewed by Vanholme *et al.* (2008).

Lignin is thus a complex compound having a distinctive structure for each type of plant but sharing common properties. This biopolymer plays an important role in plant as major structural component. It is used as material for many applications and was thus the target of intensive engineering.

Due to their location upstream in the phenylpropanoid pathway, PALs, C4H, 4CLs, HCT and C3'H are enzymes of interest and will be described in more details below.

6. Enzymes involved in the core phenylpropanoid pathway

To answer cellular requirements, the enzymes of the pathway can be either constitutively, or inducible expressed. Their expression is activated at various stages of plant development, in a tissue- and cell- specific manner (Douglas, 1996). The pathway can also be induced in answer to environmental stresses (UV irradiation, wounding, and pathogen attacks). Several *cis*-elements in the gene promoters were identified (Bell-Lelong *et al.*, 1997; Mizutani *et al.*, 1997; Whitbred & Schuler, 2000). Those are the P, A, L, H, E, G boxes recognized by different transcription factors of the *MYB* or *LIM* families (Zhong & Ye, 2009).

Phenylpropanoids derive from the chloroplastic shikimate pathway that generates the aromatic amino-acids. Thus the branchpoint between primary metabolism and secondary metabolism is here the conversion of phenylalanine to cinnamic acid by PAL.

6.1. Phenylalanine ammonia-lyase (PAL)

PAL is a cytosolic soluble enzyme, but that was also found associated to ER membranes (Achnine *et al.*, 2004).

6.1.1. Reaction catalyzed by PAL

The final products of the shikimate pathway are the three aromatic amino-acids: tyrosine, phenylanine and tryptophan. Phenylalanine and tryptophan are synthesized only by micro-organisms and plants. Tyrosine can be synthesized in the animals also from a phenylalanine precursor (Helmstaedt *et al.*, 2005).

Phenylalanine ammonia-lyase is a key enzyme of the phenylpropanoid pathway: it catalyzes the non-oxydative deamination of phenylalanine to produce *trans*-cinnamic acid (Figure 23). The enzyme is a tetramer and its sub-units are encoded by a multigenic family in most species (Achnine *et al.*, 2004).

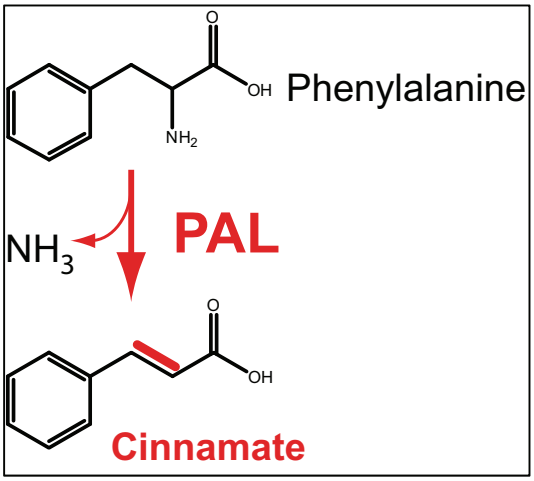


Figure 23. Reaction catalyzed by PAL

PAL genes were cloned from a large number of plant species belonging to Gymnosperms and Angiosperms (Kumar & Ellis, 2001). PAL activity was also characterized in fungi (Nambudiri *et al.*, 1970; Kim *et al.*, 2001). Tyrosine ammonia-lyase (TAL) activity was also reported in Gramineae. In corn, it was shown *in vitro* that PAL had a TAL activity (Rösler *et al.*, 1997; Whetten *et al.*, 1998). This tyrosine deamination reaction would lead directly to *p*-coumaric acid thereby by-passing *t*-cinnamic acid intermediate and C4H. Notably, some species of fungi and bacteria have a tyrosine ammonia lyase (TAL). But it is not clear if PAL and TAL can coexist *in vivo*. From an evolutionary perspective, plant and microbial PALs and TALs are part of a superfamily of enzymes found in plants, fungi, and bacteria, and are derived from a precursor of the widespread histidine ammonia lyase (HAL) family in the histidine degradation pathway (Vogt, 2010).

6.1.2. Role of PAL

PAL plays an important role in the control of the metabolic flux at the entry point of the phenylpropanoid pathway. For example in tobacco, over-expression of PAL causes a clear increase in chlorogenic acid (Howles *et al.*, 1996; Shadle *et al.*, 2003). The enzyme is tightly regulated. In addition to transcriptional regulation, various regulatory mechanisms are involved, including enzyme feedback control and post-translational modifications (Blount *et al.*, 2000; Cheng *et al.*, 2001)

PAL also plays a role in plant defense. It was shown that after infection of tobacco by tobacco mosaic virus (TMV) and in the presence of an inhibitor of the PAL, necrotic lesions were larger than that in non-inhibited tobacco (Shadle *et al.*, 2003). The treatment results in a major reduction in plant resistance.

6.2. Cinnamate 4-Hydroxylase (C4H)

t-cinnamic acid hydroxylase (*t*-C4H) or cinnamate 4-hydroxylase (C4H) was one of the first cytochromes P450 characterized in plants (Nair & Vining, 1965). It is one of the most abundant and constitutively expressed monooxygenases present in all plants and is encoded by *CYP73A* genes.

6.2.1. Discovery of C4H

In 1993, cDNAs coding for the C4Hs from *Vigna radiata* (*CYP73A2*) (Mizutani *et al.*, 1993), *Helianthus tuberosus* (*CYP73A1*) (Teutsch *et al.*, 1993) and *Medicago sativa* (*CYP73A3*) (Fahrendorf & Dixon, 1993) were isolated. Then, starting from these three sequences, other *CYP73As* were identified in many species, for example, *CYP73A5* from *Arabidopsis thaliana* (Urban *et al.*, 1997). To date more than 110 CYP73As are listed in databases. The majority of the sequences coding for known CYP73As are very similar. Plant expression of fusion constructs of C4H with fluorescent proteins indicated localization on endoplasmic reticulum (Ro *et al.*, 2001).

CYP73 cDNAs were also used to determine the enzymatic properties of the recombinant proteins after heterologous expression in yeast. CYP73As are considered as the best characterized plant P450s from an enzymatic point of view. They all catalyze the 4-hydroxylation of cinnamate with a high efficiency (*K*m ranging from 1 to 8.9 µM) (Urban *et al.*, 1994; Pierrel *et al.*, 1994; Gravot *et al.*, 2004). They were the first plant P450s specifically associated to a physiological function.

6.2.2. Reactions catalyzed by the C4H

The enzyme catalyzes the second step of the phenylpropanoid pathway (Figure 24). C4H converts (by a 4-hydroxylation) *t*-cinnamate into *p*-coumarate. But recently, substrate specificity was explored after expression in yeast (Chen *et al.*, 2007) to show that 3-coumaric acid and various *ortho*-substituted cinnamate analogues were also metabolized by the enzyme (with less efficiency). C4H also metabolizes a range of small planar and negatively charged molecules with structural similarity to cinnamic acid, some of them with high efficiency (Pierrel *et al.*, 1994; Schoch *et al.*, 2002; Schalk *et al.*, 1997a; 1998). Some of these molecules were shown to behave as potent and specific inhibitors.

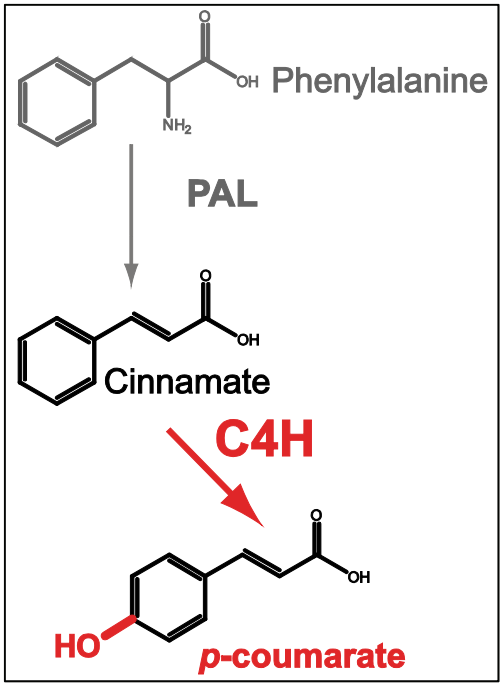


Figure 24. Reaction catalyzed by C4H.

6.2.3. Gene expression and role of the C4H

The *CYP73A5* gene and its promoter have been isolated by two groups (Bell-Lelong *et al.*, 1997; Mizutani *et al.*, 1997). Both groups investigated *CYP73A5* regulation and showed that its expression pattern was essentially the same as that of other genes in the core phenylpropanoid pathway, such as those coding for phenylalanine ammonia-lyase and 4-coumarate CoA ligase. The *CYP73A5* transcripts were detected in all plant tissues, but were highest in lignified tissues, in stems, roots and siliques. Tissue-specific expression, analyzed with *GUS* translational fusions, detected strong expression in all vascular tissues including stems, roots and leaves, in young shoot meristems and in reproductive tissues (Bell-Lelong *et al.*, 1997). Expression of *CYP73A5* also appeared to increase upon light exposure for the production of feruloyl-, sinapoyl- glucose conjugates and sinapoyl malate (Landry *et al.*, 1995; Ruegger *et al.*, 2001; Hemm & Chapple, 2004; Schilmiller *et al.*, 2009).

C4H contributes to plant defense. Indeed, C4H expression is induced by wounding of mesophyll and parenchyma cells from mature leaves (Teutsch *et al.*, 1993), by elicitation (Koopmann *et al.*, 1999; Nedelkina *et al.*, 1999) or by chemical stresses, for example herbicides (Batard *et al.*, 1997).

In 2000, Blount *et al.* showed using sense and antisense constructs that there is a negative feedback loop between C4H and PAL. This feedback modulates the expression and the enzymatic activity of PAL i.e. a reduction in the expression of the C4H in tobacco, by silencing, causes a reduction in PAL expression and activity. This causes a decrease in the concentrations of the products of the pathway, such as chlorogenic acid (Blount *et al.,* 2000), and in a reduction in G and especially S units in lignin (Sewalt *et al.,* 1997).

The *CYP73A5* promoter has been isolated by two groups (Bell-Lelong *et al.*, 1997; Mizutani *et al.*, 1997). Various boxes (P, A, G, H, L…) previously identified in other genes of the phenylpropanoid pathway, were also found in the promoter of *CYP73A5*. This promoter, as well as other promoters of phenylpropanoid genes, is thus expected to be target of the same transcription factors. *CYP73A5* promoter is the principal target of *AtMYB4*, a transcription repressor down-regulated in plants exposed to UV-B (Jin *et al.,* 2000). Regulation of the synthesis of sinapate ester sunscreens may thus occur at the level of *CYP73A5*.

The *CYP73A5* promoter had been repeatedly found to be much more efficient than *CaMV-35S* promoter for driving the expression of genes in the lignin biosynthesis such as coniferaldehyde 5-hydroxylase in different plant species (Meyer *et al.*, 1998; Franke *et al.*, 2000) which identified it as useful tool for lignin biotechnology.

C4H mutants from *Arabidopsis* affected in sinapoyl malate production have reduced epidermal fluorescence phenotype (*ref* phenotype). Due to the essential function of *CYP73* for plant development, mutants for this gene have only been described recently (Schilmiller *et al.,* 2009). As mentioned by these authors, the null-mutants *ref3* display a seedling lethal phenotype with seedlings germinating, but primary leaves failing to expand. A series of allelic EMS mis-sense and leaky mutants were therefore isolated based on reduced epidermal fluorescence (Schilmiller *et al.,* 2009). All these mutations induced single amino acid substitutions in conserved regions and substrate recognition sites (SRS) of the protein, with a resulting unstable enzyme. Severity of the phenotype depends on the allele, ranging from mere loss of apical dominance to severe dwarfism and complete male sterility. This sterility is mainly due to complete pollen abortion, probably due to a deficiency in sporopollenin and pollen coat deposition. At the molecular level, mutants display a strong decrease in lignin content with increased syringyl to guaiacyl ratio. Stronger alleles result in collapsed xylem in tracheary elements. Analysis of soluble metabolites also revealed that mutants accumulate cinnamoyl malate instead of sinapoyl malate.

Such data are in agreement with the upstream position of CYP73 in the phenylpropanoid pathway and with a fundamental role in the regulation and the redirection of the carbon flux from the primary metabolism towards the phenylpropanoids pathway.

6.3. 4-coumarate :CoenzymeA ligase (4CL)

6.3.1. Reaction catalyzed by the 4CL

4-coumarate:coenzyme A ligase is a cytosolic soluble enzyme that catalyzes the ATP-dependent formation of CoA thioesters of hydroxycinnamic acids (Figure 25). This enzyme plays a key function in the phenylpropanoid metabolism: it is located at a metabolic crossroad leading either to lignins or to flavonoids, isoflavonoids and stillbenes... These secondary aromatic metabolites are thus all derived from thioesters of CoA.

4CL was described in many plants and, depending on the plant species, one or more isoforms are found (Xu *et al.*, 2009).

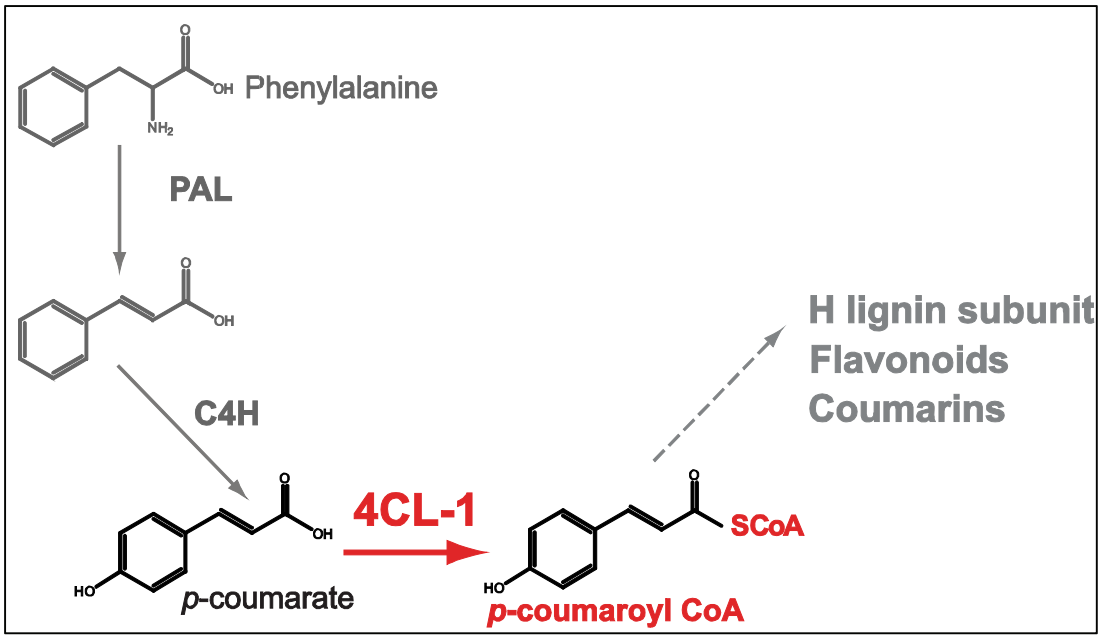


Figure 25. Reaction catalyzed by 4CL: biosynthesis of hydroxycinnamoyl CoA esters.

6.3.2. Role of 4CL

Expression studies of 4CL in various species showed that it is expressed in stems, roots, flowers and, at lower levels, in leaves for the majority of plants. 4CL expression is increased by wounding, elicitation or fungal infection (Soltani *et al.*, 2006).

The existence of several isoforms of 4CL in many plants could suggest a distinct role for each of them. Ehlting *et al.* (1999) studied the expression of three 4CL genes in *Arabidopsis*: *At4CL-1*, *At4CL-2* and *At4CL-3*. They showed that *At4CL-1* and *At4CL-2* have similar expression profiles: induction by fungal infection, wounding and UV light, with *At4CL-1* also strongly expressed in healthy stem and root during the development. The *At4CL-3* gene, on the other hand is induced only by UV light and its expression is more specific to flower and leaves where flavonoids play a major role in UV protection. Their results thus suggested the implication of *At4CL-3* in the redirectioning of *p*-coumarate to flavonoids biosynthesis while the two other genes seemed more specifically involved in lignification. A fourth 4CL from *A. thaliana* was reported. Biochemical characterization of At4CL-4 showed its unusual substrate preference *in vitro* towards sinapate. At4CL-4 was assumed to have its major role elsewhere in phenylpropanoid metabolism (Hamberger & Hahlbrock, 2004). *At4CL-4* expression was shown to be specific to root vasculature during development and to be very rapidly induced by wounding (Soltani *et al.*, 2006)

Lindermayr *et al.* (2002) isolated four cDNAs coding for 4CL from soybean. They showed that, in addition to the strong divergence in their substrate specificities, the proteins of this small family showed distinct expression profiles.

For our experiments we have chosen to analyze more specifically the interactions of 4CL-1 in the lignin metabolon. This isoform was selected since it was shown to have an essential role in the biosynthesis of guaiacyl units (G). Its role was shown by the generation of 4CL-1 antisense lines of *Arabidopsis.* These lines had only 8% of residual activity for 4CL-1. The decrease in activity was correlated with a reduction in G subunit to 7-15% with no change in S subunit, thus leading to a reduction in the G/S ratio (Anterola & Lewis, 2002). Suggesting that, in *Arabidopsis*, the S unit is synthesized by a pathway independent of 4CL-1. This experiment also shows that lignin composition can be modified in a foreseeable manner by exploiting the specific properties of 4CL-1 (Lee *et al.*, 1997).

6.4. Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase and Hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase (HCT and HQT)

HCT and HQT are cytosolic soluble enzymes belonging to the family of BAHD acyltransferases involved in the biosynthesis of various secondary metabolites. The BAHD family was named according to the first letter of the first four characterized members of the family (BEAT, AHCT, HCBT, DAT) (St-Pierre *et al.*, 1998). BEAT (Benzylalcohol *O*-acetyltransferase) from *Clarkia breweri* is responsible for producing the floral volatile benzylacetate (Dudareva *et al.*, 1998). DAT (deacetylvindoline 4-*O*-acetyltransferase) from *Catharanthus roseus* is involved in the synthesis of the alkaloid vindoline (St-Pierre *et al.*, 1998). HCBT (*N*-hydroxycinnamoyl/benzoyltransferase) from *Dianthus caryophyllus* produces a class of phytoalexins known as anthramides (Yang *et al.*, 1997). AHCT (anthocyanin *O*-hydroxycinnamoyltransferase) from *Gentiana triflora* is involved in acylation of anthocyanins (Fujiwara *et al.*, 1997)

BAHD acyltransferases catalyze the transfer of an acyl moiety to a wide range of acceptor molecules and are therefore involved in the biosynthesis of a large array of natural plant compounds such as phenylpropanoids (D'Auria, 2006).

6.4.1. Discovery of HCT and HQT

In 1979, investigation on chlorogenic acid biosynthesis led to the description of HQT in tomato (Rhodes & Wooltorton, 1976). This enzyme was then characterized in *Nicotiana alata* and *Stevia rebaudiana* and potato (Rhodes et al., 1979; Ulbrich & Zenk, 1979). The same year, the experiments of Ulbrich and Zenk on *Stevia rebaudiana* led to the discovery of a second enzyme that they initially named *p*-coumaroyl-CoA shikimate *p*-coumaroyl transferase (Ulbrich & Zenk, 1980), then hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (or CST = HCT).

Only recently, the first HCT was purified from tobacco and the corresponding gene was isolated (Hoffmann *et al.*, 2003). In *Arabidopsis*, a gene (At5g48930) homologous to tobacco HCT was then identified (Hoffmann *et al.*, 2004). This gene codes for an acyltransferase with a catalytic activity similar to that of tobacco HCT. These enzymes accept *p*-coumaroyl-CoA and caffeoyl- CoA as substrates and transfer the acyl group on shikimate or quinate.

It is interesting to note that although the *Arabidopsis* genome contains a gene coding for a protein very similar to HCT, no gene similar to HQT was found. This suggests that only one shikimate/quinate hydroxycinnamoyl transferase is present in *Arabidopsis* (plant not accumulating chlorogenic acid).

Since the discovery of HCT, five other genes of *A. thaliana* have been biochemically identified to encode BAHD acyltransferases. Three of these have been shown to acylate anthocyanin substrates (D'Auria *et al.*, 2007a; Luo *et al.*, 2007), the fourth to catalyze the synthesis of a volatile ester induced in leaves upon wounding (D'Auria *et al.*, 2007b). The fifth, named spermidine hydroxycinnamoyl transferase (SHT) is phylogenetically closely related to HCT. SHT was characterized as an acyltransferase, that uses spermidine as a substrate and various hydroxycinnamoyl-CoA esters as acyl donors (Grienenberger *et al.*, 2009). SHT is described in the review introducing chapter II.

6.4.2. Reactions catalyzed by HCT and HQT

In the plant, HCT and HQT catalyze the conversion of *p*-coumaroyl CoA (donor of acyl group) and shikimic and/or quinic acids (acceptors) into *p*-coumaroyl shikimate and/or *p*-coumaroyl quinate, respectively. They can also catalyze the conversion of caffeoyl CoA into caffeoyl shikimate and caffeoyl quinate (chlorogenic acid) (Figure 26).

Whereas HCT clearly prefers shikimate (Km=750 µM) over quinate (Km=70.10^3 µM) (Hoffmann *et al.*, 2003), HQT has a strong preference for quinate (Km=871 µM) over shikimate (Km=11 820 µM) (Niggeweg *et al.*, 2004). HCT was shown to be the most active with *p*-coumaroyl-CoA and caffeoyl-CoA esters as acyl donors, but it could also metabolize cinnamoyl-CoA (efficiently), feruloyl-CoA (less efficiently), and sinapoyl-CoA, with a very low activity for the last donor (Hoffmann *et al.*, 2003).

Moreover, like most plant acyltransferases, HCT and HQT also catalyze the reverse reaction, i.e. the formation of caffeoyl-CoA (precursor of the G and S lignin subunits) from caffeoyl shikimate.

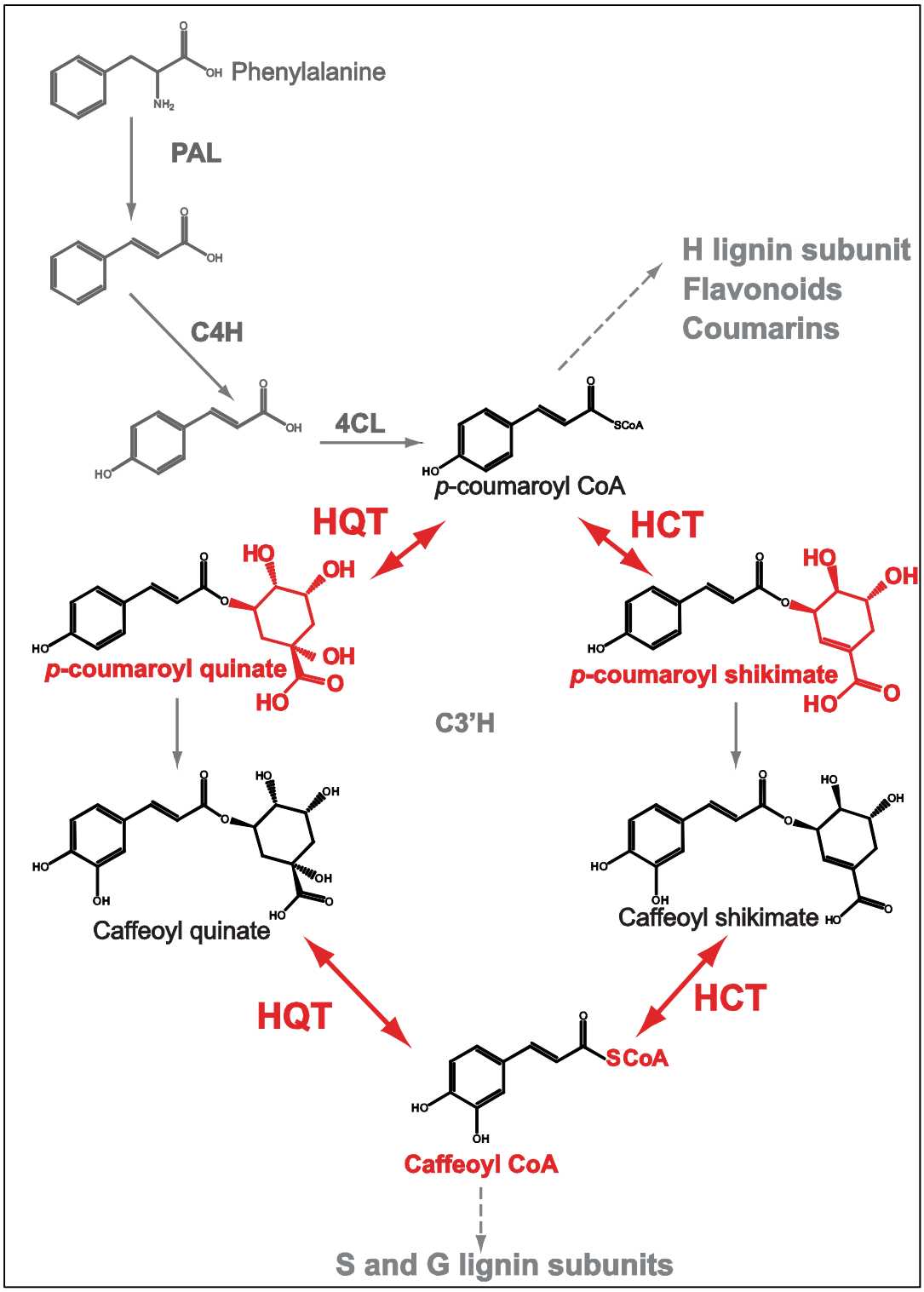


Figure 26. Reactions catalyzed by HCT / HQT.

6.4.3. Study of the expression and the role of HCT and HQT

Study of HCT expression in various tobacco tissues (Hoffmann *et al.*, 2004) showed that highest expression is in the stem internodes. A lower expression was detected in roots and petioles. These experiments showed that HCT is expressed in lignified tissues. In these stems and roots tissues, HCT is more specifically located in phloem cells (external and internal) and at lower levels in the cambial zone.

Silencing of tobacco HCT demonstrated the participation of this enzyme in lignin biosynthesis. Indeed, the silenced plants show deep changes in their development (slower than wild type), in their lignin composition (reduction in HCT accumulation in the xylem parallels a decrease in S unit and an increase of H unit), but also in the wall structure. Cellulase degradation is easier in the silenced plants with reduced lignification. In the silenced plants, enzyme loss in the vascular cells and the impact on lignin structure were clearly correlated (Hoffman *et al.*, 2004).

The suppression of HQT in tobacco or tomato does not have a detectable effect on lignin accumulation. This suggests that it is not redundant with HCT for the biosynthesis of lignin. Silencing and over-expression (stable or transient) of *HQT* in tomato showed that the *in vivo* HQT activity is most likely devoted to the formation of chlorogenic acid (CGA). Indeed, a transient over-expression of *HQT* causes an approximately 60% increase in CGA compared to the wild type. Conversely, stable silencing of *HQT* causes a 98% reduction in the level of CGA and its activity appears to determine the rates of CGA accumulation in the plant (Niggeweg *et al.*, 2004). This could explain why, in *Arabidospsis* that is devoid of HQT, CGA does not accumulate. On the other hand, HQT expression profiles in tomato and tobacco are similar to the distribution profile of CGA described in potato. In this plant, CGA accumulates in large amounts in aerial tissues but is very low in roots (Niggeweg *et al.*, 2004).

Thus, two hydroxycinnamoyltransferases are present in some plants (such as tobacco): HCT, that seems to have a predominant role in lignin biosynthesis, and HQT, that seems to be required mainly for the formation of chlorogenic acid.

6.5. *p*-coumaroyl ester 3'-hydroxylase (C3'H)

C3'H is a cytochrome P450 belonging to the CYP98 family.

6.5.1. Discovery of C3'H and its specificity of substrate

Involvement of a P450 enzyme in the 3'-hydroxylation of *p*-coumaroyl shikimate/quinate esters was first shown in 1985 in parsley and carrot cell cultures (Heller & Kühnl, 1985; Kühnl *et al.*, 1987). However, none of these enzymes had been characterized, until ten years ago.

The high frequency of CYP98As ESTs (Expressed Sequence Tag) found in many plant species, especially in lignifying tissues (roots, inflorescences, siliques), initially drew attention to this gene family. Location and high expression of CYP98A3 in *Arabidopsis* also suggested its involvement in the lignin pathway.

In addition, phylogenetic analysis indicated evolutionary clustering of CYP73 and CYP98 families (Figure 27). So it was proposed that 3'-hydroxylation of *p*-coumaric acid was carried out by CYP98A3 in *A. thaliana*. To test this hypothesis, it was expressed in yeast (Schoch *et al.*, 2001). The recombinant C3'H was shown not to hydroxylate *p*-coumarate in its free acid form as expected, but the esters of shikimate or quinate (Schoch *et al.*, 2001; Franke *et al.*, 2002a; Nair *et al.*, 2002). Indeed, biochemical and kinetic characterization of CYP98As expressed in yeast showed that this enzyme very efficiently converts 5-*O*-(4-coumaroyl)shikimate and 5-*O*-(4-coumaroyl)quinate in their corresponding caffeic esters.

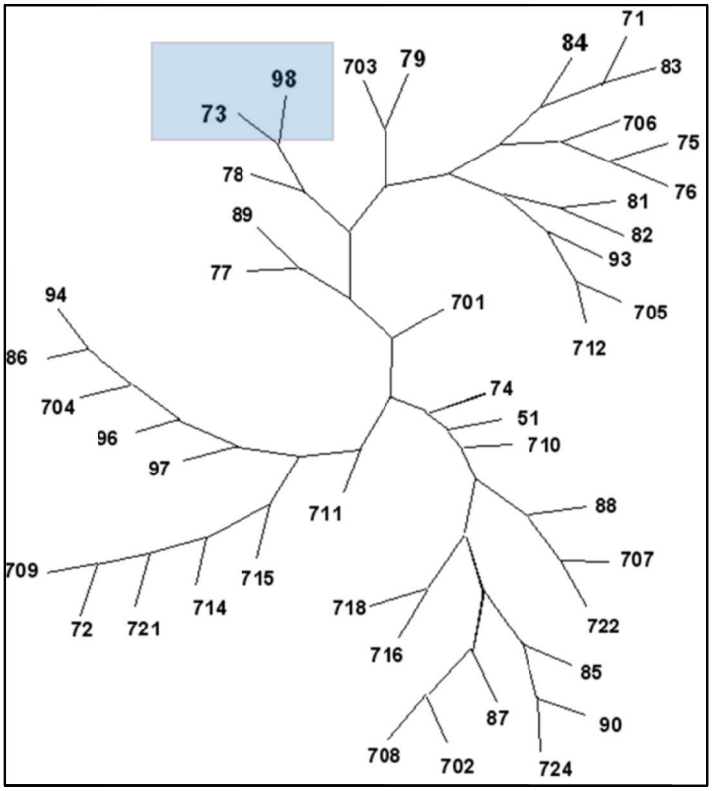


Figure 27. CYP98As in a simplified phylogenetic tree of *A. thaliana* P450s (Schoch *et al.*, 2001)

To date more than thirty-five *CYP98As* from different plants have been reported. Two paralogous CYP98As (CYP98A8 and CYP98A9) from *Arabidopsis thaliana* were recently characterized. The functions of these two CYP98As (Matsuno *et al.*, 2009) will be discussed in the introduction of chapter II.

6.5.2. Reactions catalyzed by C3'H

Recombinant CYP98As did not metabolize free *p*-coumaric acid, nor its CoA or glucose conjugates, but was found to very efficiently hydroxylate the shikimate and quinate esters of *p*-coumaric acid in the 3 position of the phenolic ring (Figure 28). CYP98As was thus capable to catalyze the conversion of *p*-coumaroyl quinate into chlorogenic acid (caffeoyl quinate), and was also able to catalyze the *meta*-hydroxylation step for the formation of lignin monomers. Positional but not *cis* isomers of the substrates were also converted by CYP98As (Schoch *et al.*, 2001). More thorough investigation carried out with recombinant enzymes of wheat or coffee tree, for example, further indicated that the shikimate and quinate esters were not the sole CYP98 substrates. This subject will be discussed in the introduction of chapter II.

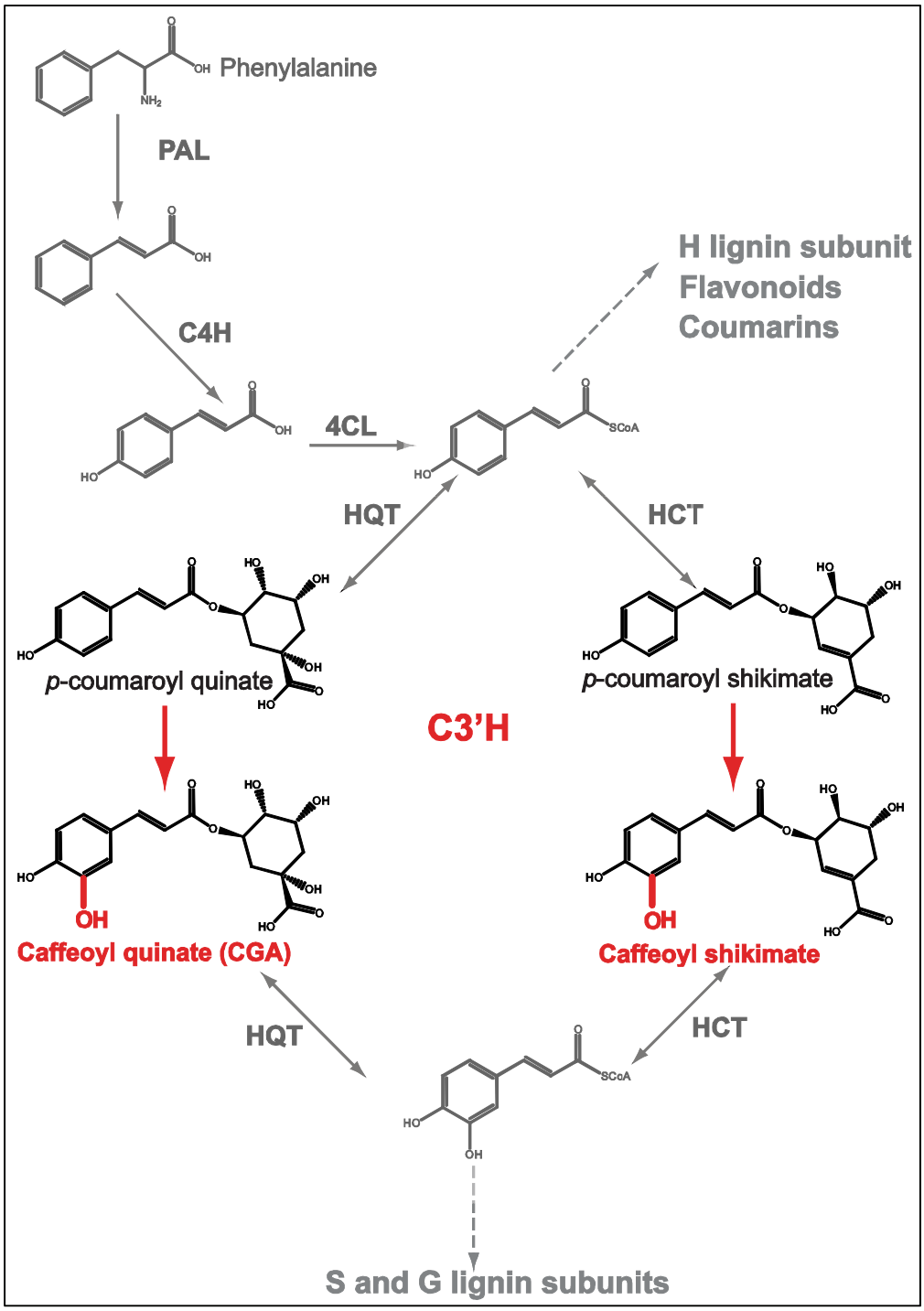


Figure 28. Reactions catalyzed by C3'H in the core phenylpropanoid pathway.

6.5.3. Role of C3'H

CYP98A3 was found to be constitutively expressed in all *Arabidopsis* organs, but at higher levels in inflorescence, stems, roots and siliques (Schoch *et al.*, 2001). Immunolocalization of the protein on tissue prints indicated that its expression was confined to tissues undergoing active lignification in stems and roots. Expression was also induced by wounding (Schoch *et al.*, 2001). These data support a role in the biosynthesis of lignin.

Meanwhile a forward genetics approach led to the *ref8* lines (reduced epidermal fluorescence) that were shown to be mutated in *CYP98A3* (Franke *et al.*, 2002 a and b). Analysis of these mutants confirmed the above result, demonstrating that *ref8* mutants had very reduced lignin content (only 20 to 40% of wild-type), and that this lignin, partially functional, was primarily made up from *p*-coumaroyl alcohol (unit H) that is only found in small amounts in the wild-type plants. This resulted in increased lignin degradability and vascular bundle collapse. The *ref8* mutant shows reduced growth, and is essentially male sterile and prone to fungus attacks, which suggests that the products of the pathway, downstream from *REF8*, are necessary for normal plant development and disease resistance. Furthermore, this mutant was defective in sinapoyl malate accumulation in rosette leaves and was shown to be unable to synthesize caffeoyl residues and downstream compounds by feeding radiolabelled precursors (Franke *et al.*, 2002a). *ref8* mutation was successfully complemented by the wild-type gene, confirming that *CYP98A3* is involved in *meta*-hydroxylation of phenolic precursors.

These data were also confirmed by the isolation of a T-DNA-insertion mutant in Ws background and the creation of co-suppressed lines (Abdulrazzak *et al.*, 2006; Ralph *et al.*, 2006). Severity of the growth inhibition phenotype was found to correlate with the degree of co-suppression, with an extremely severe phenotype in the T-DNA insertion mutant. Decrease in *CYP98A3* gene expression was associated with limp stems and inflorescences, bushy appearance, accumulation of flavonoid glycosides and anthocyanins, male sterility and an ectopic lignification phenotype. While starting to bolt, T-DNA mutants never developed fertile flowers and stems never exceeded 3 cm. The reduced cell growth was associated with an alteration in cell wall polysaccharides, a decrease in crystalline cellulose and profound modification in gene expression. Despite strong reduction in lignin content and accumulation of hydroxycinnamoyl units in the null mutant, significant amounts of G and S residues were still detected, in particular in the roots, and the purple coloration of the cauline leaves was related to the accumulation of substantial amounts of sinapoylated anthocyanins (Abdulrazzak *et al.*, 2006). This last observation led to hypothesize an alternative pathway leading to sinapoyl derivatives by an alternative non-redundant 3'-hydroxylation.

Thus in plants, the phenylpropanoid pathway provides precursors for the biosynthesis of lignin and many soluble secondary metabolites, such as chlorogenic acid. Modified expression of the participating enzymes and a better understanding of this pathway are required to engineer the composition in lignin monomers in plants of agronomic interests, or for efficient production of compounds of industrial interest. Among the candidate enzymes, several P450s have important roles: C4H, C3'H and F5H (F5H enzymatic step is not studied in this manuscript).

7. Issues and objectives of my thesis

7.1. To study protein-protein and protein-membrane interactions leading to the channeling of metabolic fluxes in the upper lignin pathway in *Arabidopsis thaliana*

Plants synthesize a huge variety of compounds that cannot be explained solely by the number of enzymes encoded by their genomes. It is now believed that this diversity at least partially arises from the combination of multi-enzymatic complexes (metabolons) which favor the channeling of metabolic fluxes into specific subpathways (Jørgensen *et al.*, 2005). The phenylpropanoid pathway has to channel up to 30% of the carbon fixed by the plant to the synthesis of lignin, suberin, and to the formation of a variety of pigments, aromas, defense molecules and antioxidants, most of them with economical interests. The upstream segment of this major plant pathway was used as a model to develop new approaches to study the protein-protein interactions in a metabolic network. The upstream phenylpropanoid pathway mainly involves two soluble enzymes and two membrane-bound cytochrome P450 monoxygenases, the latter respectively catalyze the 4- and 3-hydroxylation of the phenolic ring (CYP73A5 and CYP98A3 in *Arabidopsis thaliana*) (Figure 29). P450s are restricted in their mobility, and are, therefore, expected to anchor the soluble partners of the metabolons. Existence of such metabolon in this pathway was already suggested by: precursor feeding experiments, co-precipitation, co-isolation and co-localization, for example.

My thesis took part in a project developed with the support of the Human Frontier of Science Programme, and involving an international (France, Denmark and USA) and multidisciplinary (plant biology, chemistry, nanotechnologies and biophysics) collaboration. The purpose of the project was to demonstrate the formation of metabolons anchored by the two cytochromes P450s: CYP73A5 and CYP98A3, controlling the upstream limiting hydroxylation steps in the lignin biosynthetic pathway. Interactions between proteins of this putative metabolon were studied using several *in vitro* and *in vivo* approaches.

The project took advantage of an original technology allowing anchoring of the membrane cytochrome P450 proteins into lipid nanodiscs that can be used to trap soluble partners forming the metabolon. *In vitro* identification of the partner proteins were carried out via co-immunoprecipitation and surface plasmon resonance (SPR) coupled to mass spectrometry.

In this *in vitro* first part of the project I had to produce pure and functional proteins, to test the interactions. These proteins were CYP73A5, CYP98A3, HCT and 4CL-1 (Figure 29).

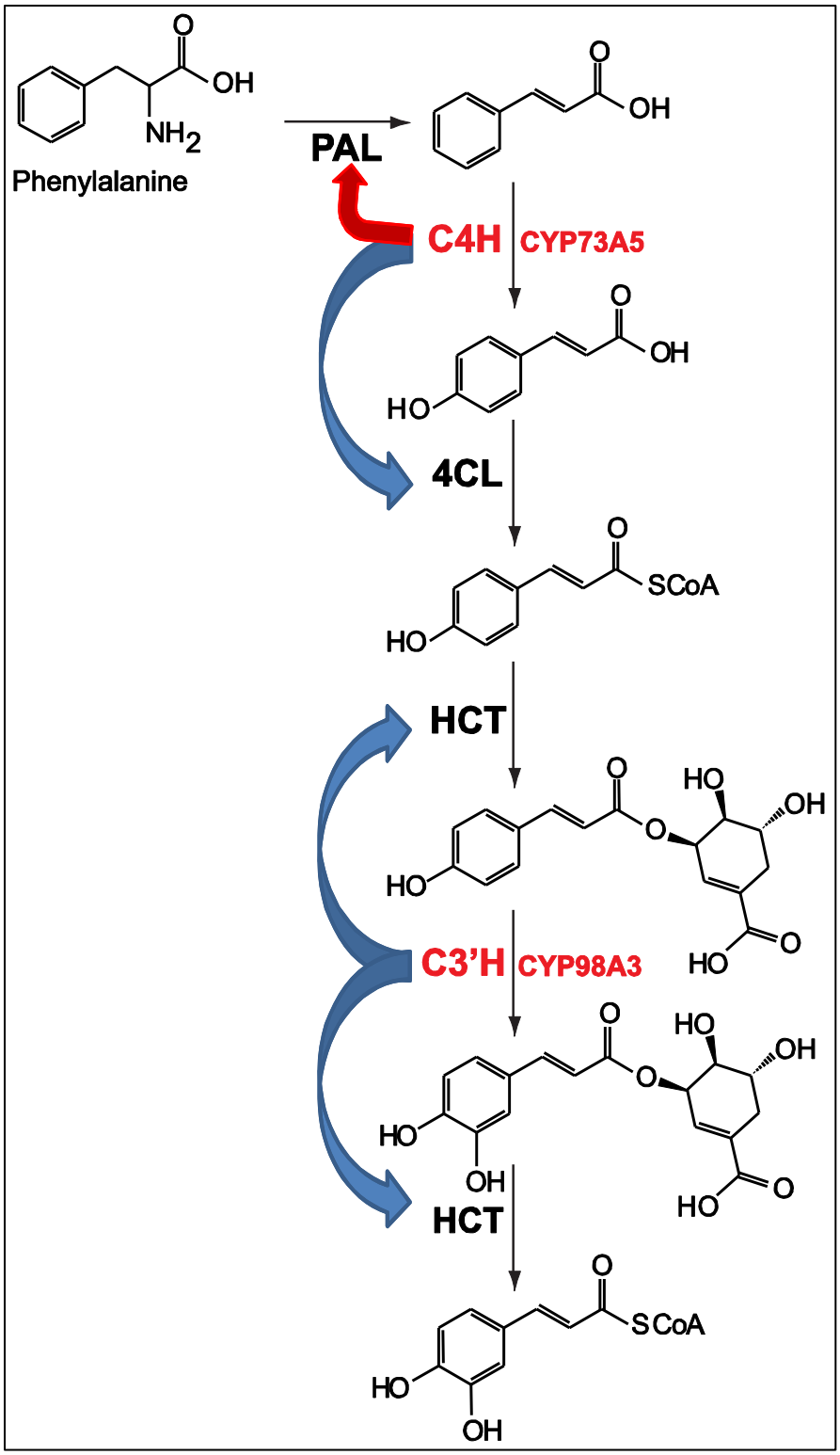


Figure 29. The core phenylpropanoid pathway.

PAL: phenylalanine ammonia-lyase; C4H: cinnamate 4-hydroxylase; 4CL: 4 coumarate CoA Ligase; HCT: hydroxycinnamoyl transferase; C3'H: coumaroyl ester 3'-hydroxylase. Red arrow indicates already proven interaction (Achnine *et al.*, 2004) and blue arrows indicate the supposed interactions studied in this project. P450s are written in red.

In parallel to these *in vitro* approaches, protein interactions were confirmed *in vivo* by transient expression of fluorescent fusion protein constructs in *N. benthamiana* leaves. Protein-protein interactions were then analyzed by enzyme activity assays and confocal microscopy, including co-localization, mobility, re-localization and interaction by Fluorescence Lifetime Imaging Microscopy (FLIM).

In this *in vivo* part of the project I conducted all the experiments.

The aims of this project were to:

- **test the potential of these new approaches,**
- **prove that protein-protein interactions occur between the four proteins and between 4CL-1 and CYP73A5 and between HCT and CYP98A3 in the first place** (Figure 29),
- **prove the existence of a metabolon composed by the four proteins as mentioned above** (Figure 30).

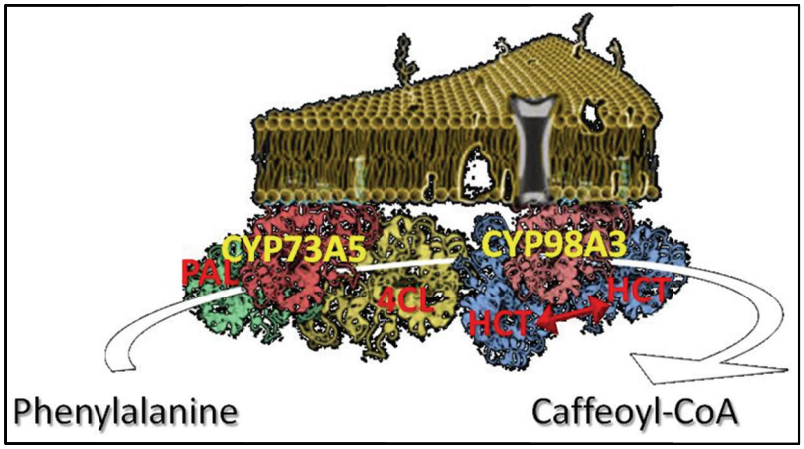


Figure 30. Hypothetical representation of the ER-localized metabolon involved upstream of the phenylpropanoid pathway.

Understanding how such protein interactions take place is an essential prerequisite to control the formation of the metabolons. This will be critical to ensure the success of the metabolic engineering approaches undertaken at experimental and industrial scale in the next decades. One of the metabolic engineering targets is the improvement of the content in molecules of pharmaceutical interest. Indeed, more than 25% of the prescriptions contain at least one active natural component as principal compound (Oksman-Caldentey & Hiltunen, 1996). This is due to the fact that some compounds are very difficult (even impossible) to synthesize chemically and their cost is much less when they can be produced by plants. However, their abundance is often very low in plants, making their extraction complex. Metabolic engineering then makes it possible to increase the yield, and control of metabolons formation is a prerequisite for it.

7.2. Functions of CYP98As paralogues in *Nicotiana tabacum*

HCTs catalyze the formation of *p*-coumaroyl shikimate and *p*-coumaroyl quinate. The preferred substrate of HCT is shikimate. In *A. thaliana* only this hydroxycinnamoyl transferase is present.

C3'H catalyzes the 3'-hydroxylation of HCT products, *p*-coumaroyl shikimate and *p*-coumaroyl quinate, into caffeoyl shikimate and caffeoyl quinate respectively. *A. thaliana* HCT and CYP98A3 show a strong substrate preference and metabolize the shikimate ester more efficiently. In *A. thaliana* only CYP98A3 (the C3'H) hydroxylates this ester.

Tobacco accumulates chlorogenic acid contrary to *A. thaliana*. To this end, it seems to have implemented specific enzymes such as HQT (hydroxycinnamoyl-CoA quinate hydroxycinnamoyltransferase). This transferase is present in tobacco beside HCT and preferentially catalyzes the formation of *p*-coumaroyl quinate. This *p*-coumaroyl quinate is then hydroxylated to caffeoyl quinate (chlorogenic acid) by a C3'H.

The presence in tobacco (as in many plants) of two or more hydroxycinnamoyl transferases, one taking part in the synthesis of the lignin monomers, another involved in the formation of chlorogenic acid, or in the synthesis of other species-specific esters (e.g. rosmarinic acid), and on the other hand the existence of three CYP98As, isolated in tobacco, led us to ask the following questions:

- **Do the different CYP98As have specific functions?**
- **What are the functions of these various isoforms?**
- **Is one of them specialized in the synthesis of the monomers of lignins and another in the formation of chlorogenic acid?**
- **Tobacco being an allotetraploid plant, are there redundancies between these isoforms?**

This part of my work thus consisted in studying the properties of the CYP98 paralogues from tobacco and answering the questions above. To this end, I have isolated one of these genes (others being already available in the laboratory) and tested the *in vitro* activities of the recombinant proteins expressed in yeast. I also attempted to obtain information on their physiological functions *in planta*.

7.3. Functions of CYP73As paralogues in *Nicotiana tabacum*

Four divergent CYP73A can be identified in the *Nicotiana tabacum* genome. CYP73A27 and CYP73A28 were previously isolated from elicitated tobacco plants (Ralston *et al.*, 2001). In addition, two others, CYP73A47 and CYP73A85, were identified in the available EST collections. They were apparently constitutively expressed. This seems to suggest duplication and specialization of the CYP73A family, at least in tobacco.

Gene duplication in the phenolic pathway was proposed to support metabolic channeling, differential cellular and subcellular localization, response specialization (Dixon *et al.*, 2001), or involvement of duplicated genes in parallel pathways using different substrates after subfunctionalization (Matsuno *et al.*, 2009).

CYP73As duplication in tobacco led us to set down hypotheses about the role of maintenance of this duplication:

- **Are the different paralogues differentially expressed and do they have specific tissue or sub-cellular localizations?**
- **Do the CYP73A paralogues have different substrate preferences?**
- **What are their specific functions in the plant? Are they redundant?**

Thus this part of my work consisted in studying CYP73As in *Nicotiana tabacum* and in answering these questions. To this end, I have cloned the different genes. I tested *in vitro* activities of the recombinant paralogues and I tried to gather information that may indicate their physiological functions *in planta*.

Chapter I

Study of protein-protein and protein-membrane interactions leading to the channeling of metabolic fluxes in phenylpropanoid metabolism in *Arabidopsis thaliana*, involving cytochromes P450 from CYP73A and CYP98A families

« En essayant continuellement on finit par réussir. Donc : Plus ça rate, plus on a de chances que ça marche ». Devise Shadocks.

A. Introduction

In spite of protein diversification in plant, their number does not explain the chemical diversity observed in the plant kingdom. Several evidences indicate that many cellular reactions within metabolic pathways are catalyzed not by free-floating soluble enzymes, but via membrane-associated multienzyme complexes. This type of macromolecular organization has important implications for the overall efficiency, specificity, and regulation of metabolic pathways. It was proposed that the metabolic diversity partly results from the combinative association of various enzymes to form these supramolecular complexes called metabolons (Winkel-Shirley, 1999; Jørgensen *et al.*, 2005).

Direct evidence for the existence of metabolons in plant secondary metabolism is difficult to obtain, due to dynamic, transient and low affinity interactions between enzymes. These properties usually prevent the direct isolation of these complexes by purification of plant extracts. Due to all these difficulties, the idea of metabolon formation was mainly based on indirect evidence. Thus sometimes, it led to great scepticism of the scientific community (Ro & Douglas, 2004). By contrast, some authors consider that the evolutionary result of metabolon formation is represented by the fusion of enzymes catalyzing a catalytic sequence in a single polypeptide, like tryptophan synthase or fatty acid synthase (Srere & Ovadi, 1990).

1. From protein-protein interactions to metabolon formation

Study of protein-protein interactions is fundamental to understand how proteins function within the cell. The ultimate goal of protein-protein interactions study is to be able to draw the map of all interactions, in a given proteome, to pass to a higher integration level: the interactome. The characterization of the interactome will be the next big step for the comprehension of cell biochemistry. Within the population of expressed proteins, only a small portion functions alone, the majority is supposed to function in unison with others to orchestrate the myriad of processes controlling functions and structure of the cell.

Interactions between proteins can:

- alter enzyme kinetic properties. This being the result of subtle changes in substrate binding or an allosteric effect.
- authorize substrate channeling by moving substrate between or among sub-units.
- create new site of binding, mainly for small effector molecules.
- change protein specificity toward its substrates by binding with various partners, or even have a new function.
- inactivate or destroy a protein.

A key capacity of all eukaryotic cells is the compartmentalization of biological processes. This compartmentalization allows to optimize concentrations in enzymes and their substrates, in order to increase overall metabolic efficiency. At molecular level, consecutive enzymes organization of a pathway in macromolecular complexes is seen like an important cellular metabolism property.

2. Metabolon and metabolic channeling

Channeling is the process during which, an intermediate produced by a first enzyme is transferred to followers without diffusion in the cytosol (Miles *et al.*, 1999). A substrate metabolic channeling supposes physical organization of successive enzymes. This organization of metabolic pathways at the molecular level is expected to have several advantages (Srere, 1987; Jørgensen *et al.*, 2005; Ralston & Yu, 2006).

For example channeling allows (Srere, 1987):

- insulation of intermediates from other competitive reactions
- effective control of metabolic flux
- securing unstable or poison intermediates by sequestering them and by preventing their diffusion
- to increase regulation signals sensitivity
- to improve catalytic efficiency: by increasing locally substrates concentration, by decreasing intermediate displacement time, by avoiding kinetic constraints of intermediates dilution in the cytosol
- to prevent action of compounds being able to exert an inhibitory effect on enzyme
- to control and coordinate metabolic interconnections carried out by enzymes involved in various pathways, or by common intermediates
- to quickly control metabolic processes by controlling assembly and localization of metabolons
- to control competition between various branches of biosynthesic pathways
- to provide new regulation levels by modulating enzymes associations.

3. Metabolon formation

Metabolons have been defined as supramolecular complexes of sequential metabolic enzymes and cellular structural elements (Srere, 1987). Assembly of a metabolon is expected to be controlled by the expression of the genes encoding its different components, to match developmental or environmental requirements. Fluxes in the branches of a pathway may thus be coordinated by the association or dissociation of different metabolons.

The formation of metabolons implies a transfer of information between proteins, various cellular processes and structures can be involved in metabolon formation. A substrate binding in the cytochrome P450 catalytic site causes a protein conformational change, and allows NADPH-cytochrome P450 reductase to interact with the cytochrome P450 (Paine *et al.*, 2005). A same mechanism can also be involved in metabolon formation. Non-catalytic proteins can also intervene for formation and organization of the metabolon (Winkel, 2004).
Lipid rafts and cytoskeleton could be implied in the formation of metabolons by stabilizing or by controlling soluble protein re-localization. Thus, a multi-proteic complex which intervenes in cell wall reorganization, is bound and controlled by microtubules (Chuong *et al.*, 2004). Moreover, metabolic regulation of the mitochondrial complexes is carried out partly by the cytoskeleton (Appaix *et al.*, 2003). Lipid rafts on biological membranes were proposed to play role in platform constitution for metabolons assembly (Zajchowski *et al.*, 2002). This lipidic organization allows a correct anchoring onto the membranes, and thus, facilitates assembly of components for metabolon formation. In answer to a stress, the lipid rafts could convey metabolons towards a precise cell location, while being guided by the cytoskeleton (Chuong *et al.*, 2004).

Well documented examples of metabolons are found in the primary metabolism (Winkel, 2004). In the latter, protein-protein associations can be very stable. The structure of the *β*-oxydation complex (Ishikawa *et al.*, 2004) has been resolved by X-ray crystallography. Studies carried out on the GAPDH/PRK/CP12 complex of the Benson-Calvin cycle involved in CO_2 assimilation demonstrated that the complex assembly/dissociation and regulation can be dependent on redox status (regulated by thioredoxin and NADPH) and on a small unstructured protein CP12 (Graciet *et al.*, 2004). This work also showed that protein-protein interaction was responsible for kinetic and conformational changes in the different constituents that may remain even after dissociation of the supramolecular structure. A work carried out on pyruvate synthesis via metabolon formation demonstrates the protection of pyruvate from competing other biosynthetic pathways, and the transient metabolon formation occurs in response to respiratory demand (Graham *et al.*, 2007).

4. Metabolons in secondary metabolism

Association of metabolons in the secondary metabolism was proposed to support the formation of the huge diversity of plant natural products by a limited number of enzymes; some of them have broad substrate specificity and combine in a great number of metabolons (Jørgensen *et al.*, 2005). As mentioned above, some constituents of the metabolon are proposed to play an important structural role in the assembly or anchoring of the metabolon. This can be the case for small unstructured proteins, cytoskeletal elements or membrane proteins (Srere, 2000; Graciet *et al.*, 2004). In the case of secondary metabolism, membrane-bound proteins, in particular P450 enzymes, have been proposed to play a critical role in metabolon nucleation (Jørgensen *et al.*, 2005; Ralston & Yu, 2006). Plant cytochromes P450 are usually anchored via their N-terminus on the cytoplasmic surface of the endoplasmic reticulum (in a few cases on the plastidial envelope) with the main protein fold protruding on the surface of the membrane (Bayburt & Sligar, 2002). Mobility of P450s is thus restricted to two dimensions and they can be associated to membrane domains with specific lipid composition. They offer a significant surface for protein-protein association. Evidence for P450 anchoring of plant metabolons has been obtained in flavonoid and cyanogenic glucoside metabolism (Winkel, 2004; Jørgensen *et al.*, 2005; Nielsen *et al*., 2008), but the best documented example is found at the entry point of the phenylpropanoid pathway (Ralston & Yu, 2006).

5. Evidence of channeling and presence of metabolons upstream of the phenylpropanoid pathway

The phenolic metabolism mobilizes a large proportion of the carbon and energy generated by photosynthesis for the synthesis of biopolymers such as lignin and diverse major soluble metabolites such as flavonoids, phenolic esters and amides (Vogt, 2010). Expression of the genes in the phenylpropanoid metabolism is therefore tightly coordinated in each branch pathway in competition. Since 1974, Helen Stafford suggests that the enzymes of this pathway must be organized in one or more enzymatic complexes to facilitate the synthesis of the various compounds (Stafford, 1974). In this system, the metabolic channeling is not only considered like a means of sequestering the unstable or toxic intermediates, but it controls flow between the multiple branches (Winkel-Shirley, 1999; Jorgensen *et al.*, 2005).

The phenylpropanoid pathway branches on phenylalanine that is deaminated into cinnamic acid by phenylalanine ammonia lyases (PAL). Cinnamic acid is then hydroxylated at the 4-position of the phenolic ring by cinnamic acid hydroxylase (C4H), a single copy cytochrome P450 CYP73A5 in *Arabidopsis thaliana*. The last step of the core phenylpropanoid pathway is catalyzed by a 4-coumaric acid CoA ligase (4CL) that generates 4-coumaroyl CoA. This activated CoA ester is then distributed into a number of branch pathways, including flavonoids, stilbenes, phenolic esters and amides.

The first evidence of metabolic channeling in the phenylpropanoid metabolism was obtained for the couple PAL-C4H as early as 1975 when Czichi and Kindl reported partial PAL association with a purified microsomal fraction from potato tuber, together with C4H. Protein-protein association and resulting substrate channeling was further supported by incomplete equilibration of the pool of exogenously supplied cinnamate with the cinnamate formed by phenylalanine deamination: [^{3}H]-radiolabelled phenylalanine was more efficiently converted into *p*-coumaric acid than competing [^{14}C]-cinnamic acid by the purified microsomal fraction. These observations were later confirmed in the case of cucumber cotyledons where coupling appeared increased by white light irradiation (Czichi & Kindl, 1977). Hrazdina and Wagner (1985) brought further support to the supramolecular complex hypothesis by demonstrating that it could be recovered in the excluded fraction by gel filtration. Other enzymes were found associated with these large particles, including 4CL. Further work differentiated the behaviour of PAL1 and PAL2 in tobacco. PAL1, but not PAL2 was found by gel blot analysis to localize to the ER in the wild-type tobacco cells in tobacco cell cultures and in microsomal membranes isolated from cells or stems, exogenous radiolabelled cinnamate did not equilibrate with the pool of cinnamate directly produced from PAL (Rasmussen & Dixon, 1999). Both GFP-tagged PAL isoforms were partially relocalized to the ER upon C4H overexpression (Achnine *et al.*, 2004). Loss of ER-localized fluorescence was observed upon co-expression of PAL1 with PAL2-eGFP, but not of PAL2 with PAL1-eGFP, which confirms greater affinity of PAL1 for ER binding sites. Co-localization with C4H was confirmed by double immunolabelling and FRET, although data were indicative of loose interaction.

4-CL would be expected to be the downstream partner of C4H in the phenylpropanoid metabolon. So far, the only experimental indication for the participation of 4CL in this metabolon is the detection of 4CL activity in the membrane-associated enzyme aggregates isolated by gel filtration reported by Hradzina and Wagner (1985). 4CL is encoded by 4 paralogues in the genome of *A. thaliana* (Hamberger & Hahlbrock, 2004). On the basis of their expression patterns and substrate specificities, 4CL-1 and 4CL-2 have been predicted to be involved in lignin synthesis, while 4CL-3 was proposed to contribute mainly to the synthesis of flavonoids (Ehlting *et al.*, 1999; Raes *et al.*, 2003). Given its strong expression in stem and root and efficient conversion of *p*-coumaric acid into its CoA ester at the expected entry point into the pathway (Ehlting *et al.*, 1999; Costa *et al.*, 2005), 4CL-1 in fact appears as the best candidate to drive the flux of metabolites into the monolignols biosynthesis and to participate in a core lignin metabolon.

HCT involved directly upstream and downstream of the C3'H in the phenylpropanoid pathway is expected to interact with this cytochrome P450 (Hoffmann *et al.*, 2004). Furthermore, these enzymes seem to be co-regulated and be involved in lignin synthesis (Raes *et al.*, 2003).

6. Objectives of the project

With the exception of the polyamine metabolon involved in the biosynthesis of spermidine (Panicot *et al.*, 2002), supramolecular complexes in secondary metabolism usually turned out more elusive and seem to involve less stable interactions than in primary metabolism. This might be related to secondary metabolism plasticity. In addition, as mentioned above, labile membrane-anchored cytochrome P450 enzymes are suspected to serve as nucleation sites for metabolon assembly. Novel and more specific approaches have to be developed to detect transient and more elusive interactions with functional membrane-bound proteins.

We describe here novel approaches to investigate protein-membrane and protein-protein interactions and their application to the phenolic metabolism, with main focus on the interaction between C4H, 4CL-1, HCT and C3'H, to investigate supramolecular protein association in the lignin metabolon.

The nanodisc technology recently developed to maintain membrane-anchored proteins in a stable and functionally active conformation, water soluble and monodisperse form (Borch & Hamann, 2009, Bayburt & Sligar, 2010), was applied for testing the *in vitro* binding of soluble protein to membrane and membrane-bound P450. Protein-protein and protein-membrane interactions were also evaluated using *in vivo* complementary approaches based on confocal microscopy using fluorescent protein fusion constructs.

B. Production of tools suitable for *in vitro* and *in vivo* studies of protein-protein interactions

Prerequisite to *in vitro* tests of protein interactions is the production and generally also the purification of the hypothetical interactants. Thus, an important step was the production of sufficient amounts of pure enough proteins. The protocols already used in the laboratory for the production of these proteins were optimized to this end.

This part of my work had two main objectives:

- The first was to produce the two purified P450s (CYP73A5 and CYP98A3) in order to allow the construction of nanodiscs incorporating them, by the team of Pr. Steven Sligar, University of Illinois, Urbana-Champaign, USA.
- The second was to optimize the production of their soluble partners, 4CL-1 and HCT, and to purify these enzymes in order to be able to analyze their interactions with CYP73A5 and CYP98A3 in partnership with the team of Pr. Peter Roepstorff, Danish Biotechnology Instrument Center, Odense, Denmark.

1. Production of CYP73A5 and CYP98A3

1.1. Expression and purification of P450s (CYP73A5 and CYP98A3)

The expression of P450s was carried out in the yeast *Saccharomyces cerevisiae* which offers a membrane context favorable for the expression of functional P450s and cytochrome P450-reductase.

The expression and purification steps of the two P450s of interest were already implemented in the laboratory before my arrival. Thus I started this project with the "regular" protocols. Expression was done by transformation of a WAT11 yeast strain, expressing one of the cytochromes P450-reductase of *A. thaliana* (ATR1: *Arabidopsis thaliana* Reductase 1). This gene was inserted into the yeast chromosome. A high copy plasmid (pYeDP60) was used to express the studied P450 in yeast. Both P450 and reductase expressions were under the control of a galactose inducible promoter: *GAL10-CYC1*.

P450 purification was carried out starting from microsomal membranes prepared from yeasts over-expressing each of the P450s fused to a 4-Histidine tag at their C-terminal end. The 4-His tag, which is small, was added to the C-terminus for not interfering with membrane insertion and thus correct folding of the P450s (Van Dyke *et al.*, 1992). The microsomal proteins were first solubilized using Emulgen 911, a non-ionic detergent which theoretically does not destabilize the P450 structure.

P450 purification was carried out in two combined steps using two columns connected in series: an anion exchange column and a nickel affinity column. The first one binds and retrieves most of acidic proteins, the second selects among proteins not retained on the first one, those with affinity for the Ni^{2+} ions. Column washes eliminated low Ni^{2+}-affinity proteins, and the proteins of interest were eluted by competition with the histidine in the elution buffer. Samples collected for each purification step were quantified by differential spectrophotometry and analyzed by SDS-PAGE (Figure 31).

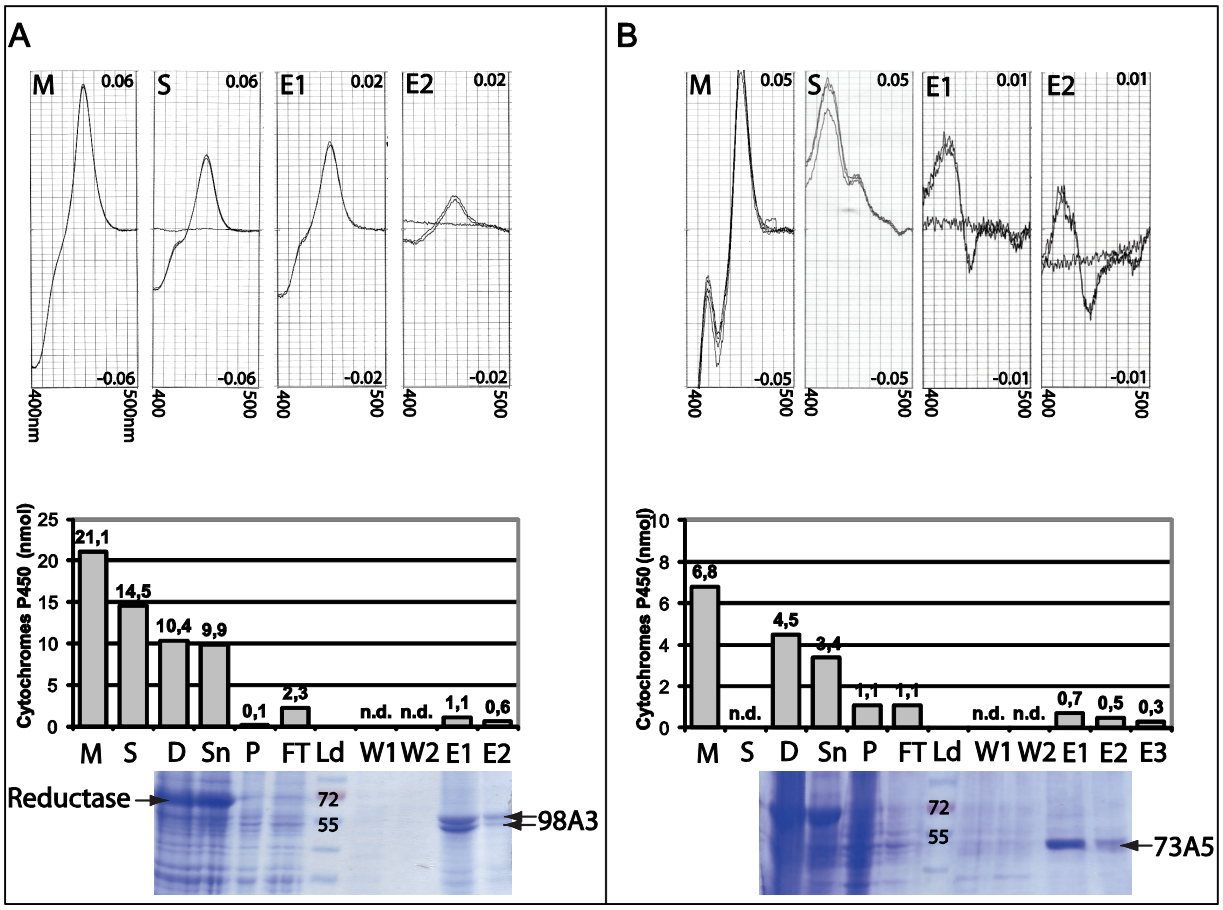


Figure 31. P450 purification follow-up.

A. CYP98A3 B. CYP73A5.

P450s were expressed with the regular protocol in *S. cerevisiae* strain (strain WAT11). P450 content was calculated from (CO-reduced/reduced) difference spectra. Proteins were resolved on 10% polyacrylamide gel and stained with Coomassie Blue. Annotated spots on the SDS-PAGE scan were checked by Mass Spectrometry. Abbrevations: M. Microsomes before purification; S. Solubilized protein ; D. Diluted solubilized protein ; Sn. Supernatant of ultracentrifugation; P. Pellet of ultracentrifugation; FT. Flow Through of the column loading; Ld. Protein Ladder in kDaltons; W1. First Wash; W2. Second Wash; E1. First Elution; E2. Second Elution; E3. Third Elution. N.d. not determined. These data are representative of the different purifications.

The first observation was that expression of CYP73A5 and CYP98A3 was variable between preparations that showed very different P450 content and often the presence of a peak at 420 nm on differential spectra. This usually indicates a mis-folding of the protein or a destabilization of the heme-protein connection. CYP73A5 seemed less stable than CYP98A3 and most often the peak at 420 nm was predominant in membranes and during purification (Figure 31 B). However, P450s in microsomes were functional when tested for enzyme activity (the substrate disappeared) even when only a peak at 420 nm was observed (Figure 32).

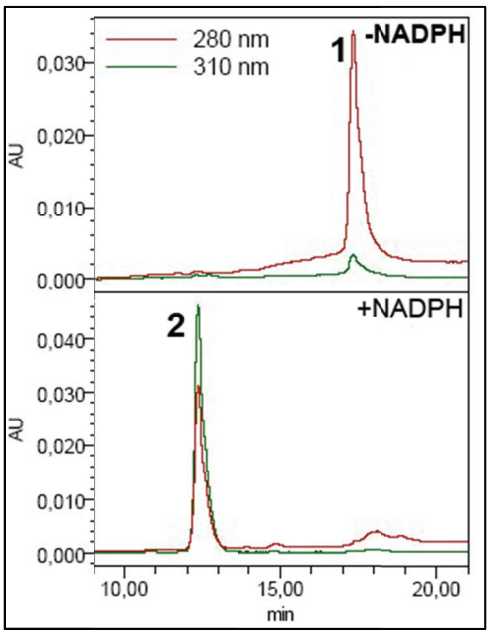


Figure 32. HPLC chromatograms of the CYP73A5 activity test on recombinant yeast microsomes used for enzyme purification.

Negative control: -NADPH.
3.2 nmoles of P450 with only a peak at 420 nm in the CO difference spectrum were incubated for 30 minutes at 27°C with 20 nmoles of cinnamic acid.
Peak number 1: cinnamate (substrate) and peak number 2: *p*-coumarate (product).

According to the SDS-PAGE analyses, the major protein in the solubilized fractions and the supernatant, was not the over-expressed P450: CYP98A3 and CYP73A5 have the respective theoretical molecular masses of 57.9 kDa and 57.7 kDa. The size of the major protein suggests that it could be the reductase, with an expected molecular mass of 76.7 kDa. This major contaminant was eliminated by the anion exchange column.

P450 quantification in the different fractions (from the CO-reduced/reduced difference spectra) showed that small amounts of P450 were lost in the centrifugation pellet after the solubilization step and in the flow-through from the purification columns. This was in agreement with the protein profiles detected on SDS-PAGE (Figure 31). Moreover, a significant P450 loss was observed during solubilization, then during dilution of the solubilized fraction. This seems to indicate partial denaturation of the enzyme, perhaps resulting from heme loss during solubilization.

CYP98A3 was usually found in two forms in purified fraction according to SDS-PAGE analyses. This could result from a partial degradation of the protein, from two different conformations of the same protein, or from a partial post-translational modification in yeast. According to differential spectroscopy monitoring, a significant proportion of the enzyme however remained correctly folded after purification (Figure 31 A). In the eluted fractions, only a very small proportion of enzyme presented an absorption peak at 420 nm. This was considered as important, since exploitation of this purified enzyme for the protein-protein interaction studies, would require P450 in a functional configuration.

According to spectrophotometric P450 quantification, purification of CYP98A3 provided a total amount of approximately 4 nmoles of P450 per Liter of culture. Approximately 2 nmoles of P450 per Liter of culture were obtained after purification of CYP73A5. Yields were around 8% of the total P450 in the starting microsomes.

In conclusion, protein of a sufficient purity for nanodiscs assembly seemed to be obtained. A fast and effective purification of P450s starting from solubilized microsomes was achieved. The main optimizations consisted in producing more P450s as starting materials and in improving the critical steps of the purification that seemed essential for CYP73A5 production.

1.2. Optimizations of P450s productions

Due to the loss of active and correctly folded protein during all purification steps, it was essential to start purification with large amounts of proteins and to limit losses during purification. Three possibilities were considered: improve the expression, increase culture volume, or optimize the purification procedure.

Optimization steps are summarized in the flow-chart (Figure 33). Four different production conditions and eleven conditions of purification were tested. Some buffer additives were tested during solubilization and purification: 10 mM DTT, 1 mM PMSF, 0.012% β-mercapthoethanol. After many trials, no significant improvements were obtained. Moreover, even using the same protocol, both expression and purification were poorly reproducible: protein quantities and concentrations were variable at all stages.

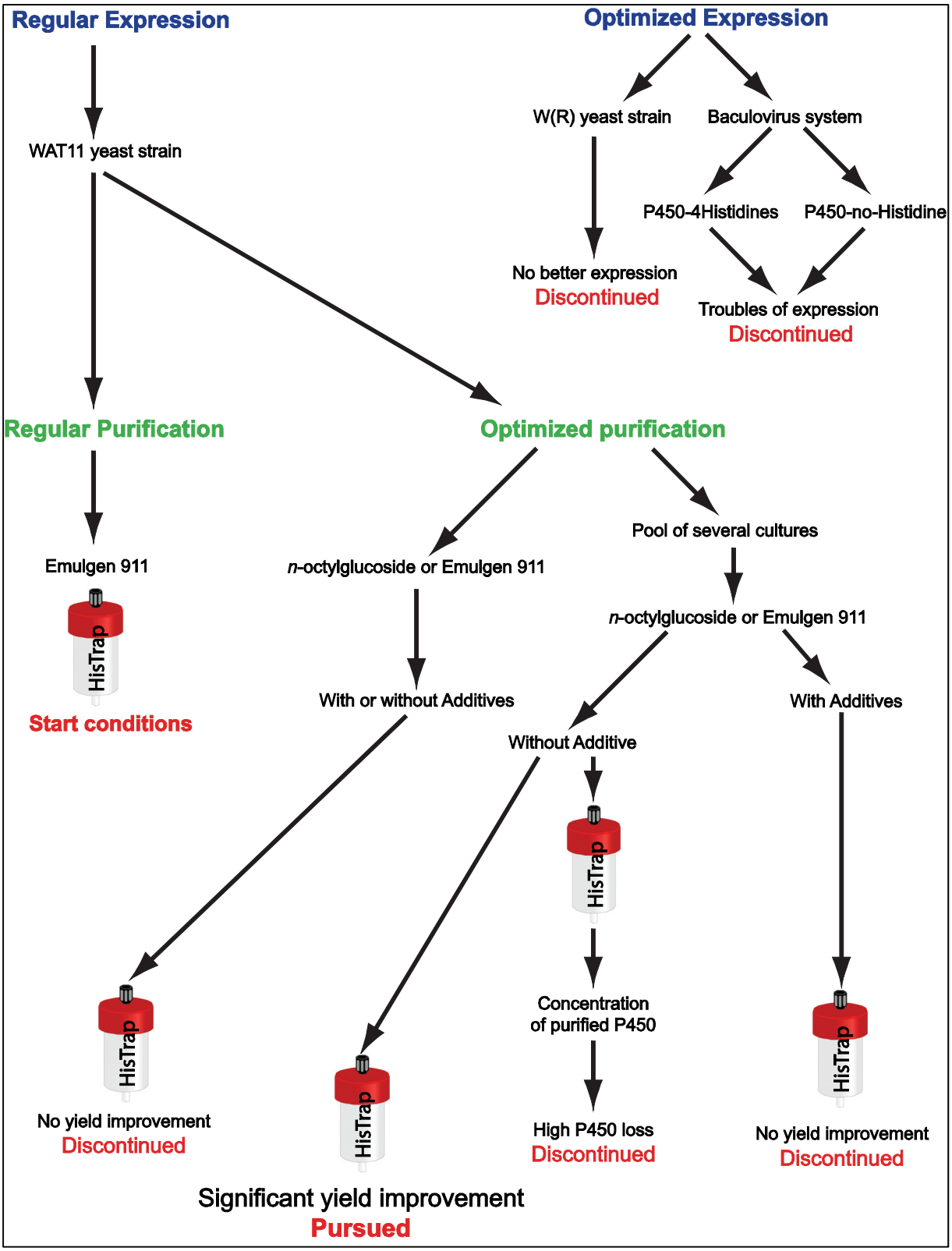


Figure 33. Flow-chart showing the different strategies tested for P450s production. Purifications were performed with HisTrap affinity column.

It was first of all attempted to increase protein expression. Two main approaches were tested:

1. Expression of P450s in another yeast strain. Coupling efficiency and thus P450 protein stability, sometimes varies with the coexpressed P450 reductases (Urban *et al.*, 1997). Thus, instead of using WAT11 strain, P450 expression was tested in the W(R) strain. This strain over-expresses the yeast reductase under the control of the galactose inducible promoter *GAL10-CYC1*. No significant difference in P450 produced was observed between WAT11 and W(R).

2. Expression of P450s was tested in insect cells that were reported to provide a good context for plant P450 expression (Duan *et al.*, 2004). In our hands, this system was complex to handle, in particular due to problems of contamination and quality of the cells, and did not provide large amounts of P450. (Results not shown).

In a second step, several modifications of the purification protocol were tested. Solubilization and dilution of the microsomes were critical: according to P450 quantification, up to 50 % of the P450s was lost at this stage. Tests of solubilization were carried out with another mild non-ionic detergent described as non-denaturing, *n*-octylglucoside. This did not improve the integrity of this unstable fraction of enzyme.

Finally, several batches of culture were pooled to start purification with large amounts of protein and to decrease relative P450 loss at each stage by adsorption on columns, plastic and glassware. Columns seemed never overloaded and an increase of the yields in eluted P450s was consistently observed.

A concentration of CYP98A3 in microsomes of about 51±19 nmole.L^{-1} of culture, and of CYP73A5 about 18±5 nmole.L^{-1} was obtained after optimization. Amount of produced P450 was increased by 1.5 fold. After purification optimization, the best yields in functional P450s were around 7±2 nmole.L^{-1} of culture for CYP98A3 and 5±2 nmole.L^{-1} for CYP73A5. So about 86% of the CYP98A3, and about 72 % of the CYP73A5 were still lost during purification. P450 quantification by differential spectrophotometry and evaluation of total proteins by a Bradford assay on the eluted fractions indicated that functional P450 (characterized by a peak at 450 nm) represented in the best case 4.6 % of the total proteins. Given the SDS-PAGE profiles it thus seemed likely that the major part of purified protein was misfolded or hemeless P450.

Nevertheless purity of the preparations and amount of P450s were finally sufficient for their incorporation into nanostructures. Purified enzyme was sent to our American partners for P450 nanodiscs construction.

1.3. Discussion

Two critical steps appeared in the expression and purification process. **A shift of the temperature for inducing protein expression at 20°C** (instead of 24°C) improved the amout of produced protein. **Yields in pure and functional P450 was doubled** when more proteins was used as starting material for purification.

Nevertheless, solubilization of the microsomes would still deserve optimization. Indeed, about half of the enzyme seems to lose its functional conformation during this step. Use of commercial purified detergents, in particular free of hydroperoxides could lead to significant improvement.

Although the fractions that we provided to the American partners of the project, did not contain a completely pure protein, and most likely in the case of CYP98A3 heterogeneous protein, **the assembly of nanodiscs and the purification of nanodiscs containing P450 was successful**. So, several batches of purified CYP98A3 and CYP73A5 were produced from recombinant yeast. Unfortunately, repeated shipment problems were encountered (slow shipping and thawing, lost parcels). P450 production in the laboratory was thus stopped.

Consequently, insect cell expression constructs and viral particles were transferred to US partners who took care of P450 proteins production, purification and incorporation into Nanodiscs. Once included into nanodiscs, P450s are more stable. They were safely sent to Danish partners for *in vitro* interaction tests.

1.4. Materials and Methods

1.4.1. Biological materials: *Saccharomyces cerevisiae*

The budding yeast *Saccharomyces cerevisiae* is an ascomycete fungus and a widely used system for the over-expression and functional analysis of plant membrane proteins.

1.4.1.1. Different used strains

The three yeast strains used, W(R), WAT11 and WAT21, were built by the group of Pompon, CGM, Gif-sur-Yvette (Pompon *et al.*, 1996) for a higher activity of heterologous expressed P450s. These strains derive from a parental strain W303-1B and are thus isogenic with the exception of the CPR1 (NADPH-cytochrome P450 reductase) locus.

- W (R) strain: (Truan *et al.*, 1993)

This yeast strain (*MAT*a; *ade*2-1; *his*3-11,-15; *leu*2-3,-112; *trp*1-1; can^R; cyr^+) over-expresses the yeast endogenous NADPH-cytochrome P450 reductase. It was constructed by substitution of the natural promoter of the *CPR1* gene by the galactose-inducible and glucose-repressed promoter *pGAL10-CYC1* (Guarente *et al.*, 1982).

- WAT11 and WAT21 strains (Urban *et al.*, 1997):

These strains over-express the *Arabidopsis thaliana* NADPH-cytochrome P450-reductases (ATR1 in WAT11 or ATR2 in WAT21) in the place of the endogenous CPR. *ATR1* and *ATR2* are under the control of the strong promoter *pGAL10-CYC1*.

ATR1 is described as being co-expressed with the proteins from the phenylpropanoid pathway in *A. thaliana* (Mizutani, 1998). Thus this strain was mainly used during this thesis.

1.4.1.2. Yeast culture conditions

Before transformation, yeasts were cultivated in YPGA complete medium. Once transformed, they were grown on selective SGI medium.

SGI medium	Bactocasaminoacid 1 $g.L^{-1}$	Yeast Nitrogen Base 7 $g.L^{-1}$	Glucose 20 $g.L^{-1}$	Tryptophane 20 $mg.L^{-1}$
YPGA medium	Bactopeptone 10 $g.L^{-1}$	Yeast extract 10 $g.L^{-1}$	Glucose 20 $g.L^{-1}$	Adenine 200 $mg.L^{-1}$
Sterilize medium 20 min at 1 bar				

1.4.2. Yeast expression vector: pYeDP60 (Urban *et al.*, 1994)

This replicative vector (9265pb) was used for expression in *Saccharomyces cerevisiae* of P450s under the control of the promoter *pGAL10-CYC1*.

pYeDP60 is a chimeric shuttle vector with yeast and bacterial derived sequences. It comprises the following characteristics:

- Unique origins of replication for each cell type, bacteria and yeast (the yeast origin replication, ori 2µ is an autonomously replicating sequence from native yeast 2-µ plasmid which allows the plasmid to replicate at a medium copy number).

- Different markers for the selection of transformed host cells:
    - The gene *bla* (coding for the *β-lactamase*) confers bacterial resistance to ampicillin.
    - *URA3* and *ADE2* markers allow the complementation of the uracil and adenine auxotrophies of the yeast host strain.
- An expression cassette comprising a strong promoter (glucose-repressed, galactose-inducible *pGAL10-CYC1*), a cloning cassette and the terminator of the yeast phosphoglycerate kinase.

1.4.3. Yeast heterologous expression system

Yeast cells were chemically prepared and transformed by the lithium acetate procedure (Ito *et al.*, 1983) improved and described by Gietz *et al.*, 1992.

1.4.3.1. Competent yeast preparation

A fresh colony of yeast was inoculated into YPGA medium (Cf. § Yeast Culture conditions p.85) and grown overnight at 28°C under agitation (160 rpm). This pre-culture was diluted in 50 mL of the same medium to obtain an OD_{600} of 0.2 and was regrown during 5 hours at 28°C under agitation (160 rpm).

Cells were harvested by centrifugation (5500 *g*, 4°C, 10 min) washed in sterile water (1/10, v/v of culture volume), resuspended in 1mL sterile water, and transferred to 2 mL Eppendorf tubes.

After centrifugation (10 s with a bench centrifuge) cells were washed with 1.5 mL of AcLi/TE solution and concentrated by centrifugation (10 s with a bench centrifuge) in 50 µL of this solution.

AcLi/TE solution	AcLi 100 mM	Tris HCl pH7.5 10 mM	EDTA 1 mM	qsp sterile water
Filtrate the solution with 0.22 µm filter				

1.4.3.2. Transformation of the fresh competent cells

50 µL of competent cells were mixed with 100 µg single strand salmon sperm carrier DNA (denatured during 20 min at 100°C) and with 1-10 µg plasmid of interest. 500 µL of sterile 40% polyethylene glycol (PEG) 4000 in AcLi/TE (40% PEG 4000, 1x TE, 1x AcLi, made fresh from sterile 50% PEG stock, and sterile 10x TE and 10x AcLi) were added before agitation on wheel during 1 hour at 30°C. Then a heat-shock at 42°C during 15 minutes was applied. The yeast cells were then recovered (10 s with a bench centrifuge), washed with 1 mL of sterile water, resuspended in 200 µL of sterile water, and plated onto selective SGI. Plates were left to grow 3-4 days at 28°C.

Solid SGI medium	Bactocasaminoacid 1 $g.L^{-1}$	Yeast Nitrogen Base 7 $g.L^{-1}$	Glucose 20 $g.L^{-1}$	Tryptophane 20 $mg.L^{-1}$	Pastagar 15 $g.L^{-1}$
Sterilize medium 20 min at 1bar before to spread on plates					

1.4.3.3. Yeast microsome preparation (adapted from Pompon *et al.*, 1996)

1.4.3.3.1. Protein expression

One transformed yeast colony was put for growth in 30 ml of SGI medium (Cf. § Yeast Culture conditions p.85) during 24 hours at 28°C under agitation (160 rpm). Then one units of OD_{600} (generally 5 mL of culture) was transferred in 200 mL of YPGE and put under agitation (160 rpm) at 28°C.

After 30 hours of growth, galactose (20 g/L final concentration) was added to induce the protein expression.

After 17 hours of induction at 20°C the cells were harvested by 10 min of centrifugation at 7500 *g* and at 4°C. From this point, all steps were done on ice.

YPGE	Yeast extract 10 $g.L^{-1}$	Bactopeptone 10 $g.L^{-1}$	Glucose 5 $g.L^{-1}$
Sterilize medium 20 min at 1 bar, before addition of ethanol 3 $mL.L^{-1}$ (v/v final)			

1.4.3.3.2. Subcellular fractionation

The pellet were washed with an ice-cold TEK buffer (2 $mL.g^{-1}$ of cells) and centrifuged 10 minutes at 7500 *g*, 4°C. The pellet was resuspended in 1 mL of ice-cold TES buffer and one extra milliliter was used to wash the centrifugation pots. This suspension was transferred in a 50 mL Falcon tube; glass beads (0.45-0.5 mm, Mannheim) were added until their level reaches the surface of the suspension. The cells were mechanically broken by vigorous hand shaking, five times one minute, while letting rest on ice between two agitations. The extract was collected and the beads were washed with ice-cold TES buffer until the supernatant was clear (4 washings of 5 mL of TES). The extract was then centrifuged during 15 minutes at 7500 *g*, 4°C. As soon as the centrifuge stopped, the supernatant was filtered on Miracloth (Calbiochem) then centrifuged during one hour at 100 000 *g*, 4°C. The supernatant was discarded and pellet was washed with 1 mL of ice-cold TEG buffer, then resuspended in 2 mL of ice-cold TEG buffer and finally put in suspension with ice-cold Potter.

The microsomes were sampled and stored at -30°C until use for up to two years (after quantification of P450 by differential spectrophotometry, Cf. § Quantification of P450s by differential spectrophotometry p.90).

TEK buffer	Tris HCl 50 mM pH7.5	EDTA 1 mM	KCl 100 mM	qsp MilliQ water
TES buffer	Tris HCl 50 mM pH7.5	EDTA 1 mM	Sorbitol 100 mM	qsp MilliQ water
TEG buffer	Tris HCl 50 mM pH7.5	EDTA 1 mM	Glycerol 30 % (v/v)	qsp MilliQ water

1.4.4. Heterologous expression system in insect cells

1.4.4.1. Recombinant bacmid generation and production of the viral particles

These bacmids and these viral particles were prepared at the laboratory by V. Sauveplane and F. Duval.

Briefly, after PCR amplification of the P450s and *ATR1* for addition of adequate restriction sites immediately upstream and downstream the sequences of interest, the coding sequences were inserted each one in the pFastBac 1 vector (Invitrogen) according to Invitrogen procedures. Recombinant bacmids were generated using homologous recombination in the *DH10Bac E. coli* strain transformed with the pFastBac 1 vectors following the Invitrogen protocols.

For the viral particles generation, *Sf9* cells were transfected using Lipofectin reagent (Invitrogen) and according to manufacturer instructions. Lipofectin allows the formation of micelle-containing viral DNA. The mix of bacmid and lipofectin was added to Sf9 cells (at 10^6 cells.mL^{-1}) in 5 mL in SF900 SFM medium (Invitrogen) in box of 25 cm^2. Virus particles in the culture media were harvested ten days after transfection and stored at -80°C in presence of 10% Fetal Bovine Serum.

These particles were directly used to infect 200 mL of Sf9 cells in order to produce the recombinant proteins.

1.4.4.2. Insect cells transfection and recombinant protein production

The virus was not titered or purified for the production of the recombinant proteins. Cells were grown in SF900 SFM medium until concentration of 3.10^6 cells.mL^{-1} and they were co-transfected with recombinant virus containing the P450 and the reductase.
Four days after the transfection, insect cells were harvested at 500 *g*, 5 minutes, 4°C.

1.4.4.3. Microsomes preparation

The pellet was washed twice with same volume of the Phosphate Buffer Saline (NaCl: 8 g.L^{-1}; KCl: 0.2 g.L^{-1}; Na_2HPO_4: 1.44 g.L^{-1}; KH_2PO_4: 0.24 g.L^{-1}, pH 7.4) as culture volume (200 mL). Pellet was lastly resuspended in 1/10 (of culture volume) with cold cell lysate buffer (PBS pH 7.4; EDTA 1 mM, DTT 0.1 mM, glycerol 20% (v/v), complete protease inhibitor cocktail EDTA free (Roche)). Cells were then lysed on ice using sonicator (2x 30 s) (sonicator 250TS.20K, Ultrasons Annemasse S.A.) and microsomes were harvested by differential centrifugation. Lysate was centrifuged at 10 000 *g*, 10 minutes, 4°C and the supernatant was then centrifuged at 100 000 *g*, 1 hour, 4°C. Microsomal pellets were resuspended in cold cell lysate buffer (2 mL of buffer per gram of microsomal pellet) with ice-cold Potter.

The microsomes were sampled and stored at -30°C (after quantification of P450 by differential spectrophotometry) until use.

1.4.5. Affinity purification of P450 (adapted from Gabriac *et al.*, 1991)

The purification of cytochrome P450 protein is articulated around two steps, which consist in solubilizing P450s, and purifying them by exploiting their specific properties:
1) their basic pI, 2) the affinity of the 4-histidines tag (added during the construction of the expression vectors) for the divalent cations, such as Ni^{2+}. The proteins can be eluted by competition with histidine (or any other molecules containing a core imidazol).

This purification was made starting from a microsomal preparation from transformed yeast, expressing P450 of interest having a 4-histidine tag at the C-terminus. The tag was placed in the C-terminus because of the presence of the P450 anchor in the N-terminus.

All the buffers used during purification of P450 were cleaned by filter with pores of 0.22 µm, to avoid clogging of the columns.

Each purification step was followed up by differential spectrophotometry, Bradford assay and SDS-PAGE, to check the course and the quality of the purification.

1.4.5.1. Solubilization of microsomes

Generally for each purification about 5 mL of microsomes were solubilized by adding drop by drop 1 mL of Emulgen 911 20% (or *n*-octylglucoside 20%), under gentle agitation over ice. Agitation and addition of detergent were pursued until the solution became translucent. Agitation was maintained for ten extra minutes over ice.

The solution being too viscous, it was diluted three-times by a dilution buffer (Tris 20 mM pH8, 15% glycerol and 0.05% of detergent). An ultracentrifugation during one hour at 100 000 *g*, 4°C allowed to discard not solubilized membranes (in the pellet).

1.4.5.2. Purification of the cytochromes P450 proteins

Before their use, columns were equilibrated by two washes of 50 mL: the first one with MilliQ water, the second one with the equilibration buffer (Tris 20 mM pH8, 15% glycerol, 0.05% detergent).

The supernatant of the ultracentrifugation was loaded over the two purification columns in series. The first column used was an anion exchange column Q Sepharose (HiTrap Q XL 5 ml, GE Healthcare) connected in series with a nickel affinity column (HisTrap HP 1 mL, GE Healthcare). The anion exchange column which retains acid proteins (negatively charged at pH8), was used to avoid the overload of the HisTrap column.

Supernatant was loaded with a flow of 0.18 $ml.min^{-1}$. Once the supernatant was loaded, the dilution buffer was used to push the supernatant until the HisTrap column. The anion exchange column was disconnected from the HisTrap column and washes were performed. The flow was fixed at 1.5 $mL.min^{-1}$. A wash with 50 mL of buffer (Tris 20 mM pH8, 10 mM Histidine, 15% glycerol, 0.05% of detergent) was carried out.

Elution was done at 1.5 $mL.min^{-1}$, generally with 15 mL of elution buffer (Tris 20 mM pH8, 50 mM histidine, 15% glycerol). Eluted fractions as well as fractions for the purification follow-up were stored at -30°C

After their use, columns were regenerated by two different buffers. For the anion exchange column, 50 mL of buffer (Tris 20 mM pH8, NaCl 1M) were used. And for the HisTrap column, 2 washes were done: a wash by 10 mL of MilliQ water followed by a wash with a solution of 50 mM $NiSO_4$.

Finally, columns were stored in 20% ethanol.

1.4.5.3. P450s concentration

Concentrations were carried out with centricon (Amicon Ultra 30K 15 mL). All the buffers used during concentration of P450s were clarified by filter with pores of 0.22 µm.

Before use, centricon was washed and equilibrated by different steps:

- 15 mL of MilliQ water.
- 2x15 mL of Ethanol 20%
- 3x15 mL of MilliQ water
- 2x15 mL of buffer (Tris 20 mM pH8, 50 mM NaCl)

Centrifugations were carried out at 5000 *g*, 4°C, 10-20 minutes.

The purification eluates were loaded progressively, a first centrifugation was carried out. When volume in centricon reach half of the total volume, buffer (Tris 20 mM pH8, 0.5 M NaCl) was added. These steps were repeated until loading of all the purification eluates and concentration was achieved in buffer (Tris 20 mM pH8, 50 mM NaCl).

During the concentration, centrifugations were done at 3000 *g*, 4°C. At the end, glycerol was added to a final concentration of 15%. Concentrated fractions were stored at -30°C.

1.4.6. Quantification of P450s by differential spectrophotometry (Omura & Sato, 1964 a,b)

Before carrying out biochemical tests on the microsomes, and during each purification step, it is necessary to check the presence of P450s and to determine their concentration.

Total cytochrome P450 content was quantified by the dithionite-reduced carbon monoxide differential spectra, modified from Omura & Sato (1964 a,b). This method is based on strong affinity of the reduced iron of P450 heme toward carbon monoxide (CO). The establishment of a coordinance bound between CO and iron induces a displacement of the Soret band of the heme from 390 nm to 450 nm.

For each replicate, 2 mL of a diluted microsomal suspension placed in a glass tube, were reduced in the presence of sodium dithionite ($Na_2S_2O_4$). The content of the tube was then equally divided into two spectrophotometer UV-cuvettes placed in a dual beam UV spectrophotometer (Cary 300Scan, Shimadzu MPS-2000, Safas Monaco) and background corrected. Carbon monoxide was bubbled into the test cuvette, and both cuvettes were then scanned from 500-400 nm. Differential spectrum (type 2 spectrum) presenting a maximum of absorption at 450 nm, in case of presence of well folded P450, was observed. Peak at 420 nm indicates a mis-folding of the protein or a destabilization of the heme-protein connection by binding of the heme to another amino acid residue for example.

The difference in absorbance between 490 and 450 nm permits to calculate the P450 concentration, using an extinction molar coefficient which was determined by Omura and Sato (1964): $\varepsilon_{450\text{-}490\ nm} = 91\ cm^{-1}.mM^{-1}$.

1.4.7. Dosage of total proteins: Bradford assay

The proteins were quantified in microsomes or in samples at each step of the purification process, using the Biorad protein assay kit (Biorad). This kit is based on the Bradford spectrophotometric method (Bradford, 1976). A dilution series of a standard of known protein sample concentration (Bovine Serum Albumin, BSA) was used to draw a standard curve. The measures were carried out in microplates in a final volume of 250 µL, and the absorbance of the standards and unknown samples was determined 5-10 minutes after preparation, with a microplate reader (Synergy HT, Biotek) at 595 nm.

1.4.8. Detection of proteins on polyacrylamide gel

1.4.8.1. Chromatography of proteins on polyacrylamide gel

Discontinuous Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis (discontinuous SDS-PAGE) is a widely used electrophoresis technique in denaturing and reducing conditions (Laemmli, 1970). This technique allows the separation of proteins according to their mass. Indeed, in the presence of a reducing agent (Dithiothreitol: DTT) and of Sodium Dodecyl Sulfate, the proteins are dissociated in their constitutive sub-units (action of the reducer) and charged negatively (by association with the SDS).

To obtain optimal resolution of proteins, polyacrylamide gels in discontinuous system consist of 2 parts:

- A stacking gel cast over the top of the resolving gel (this allows the proteins to be concentrated): 4.5% (v/v)[1] acrylamide / bisacrylamide (37.5 / 1); in buffer 125 mM Tris-HCl pH6.8; 0.1% (w/v)[2] SDS; 0.3% (w/v) ammonium persulfate (APS); 0.2% (v/v) tetramethylethylenediamine (TEMED).

[1] % (v/v) : 1 mL/100 mL

[2] % (w/v) : 1 g/100 mL

- A resolving gel (inferior part): X% acrylamide / bisacrylamide (37.5 /1); in buffer 125 mM Tris-HCl pH8.8; 0.1% (w/v) SDS; 0.1% (v/v) APS; 0.1 (v /v) TEMED, (v/v) glycerol 8%.

The resolving gels were gels with uniform X concentration, according to the mass of the protein of interest. Gels with a low percentage of acrylamide are typically used to resolve large proteins. A gel with 12% of acrylamide permits to resolve the proteins with molecular weights of about 50 kDa.

SDS-PAGE can be carried out after freezing of the samples at -30°C.

Prior to loading samples on the gel, the protein samples were diluted with loading buffer in order to have a mixture with a final composition of Tris-HCl 12 mM pH6.8; DTT 1 mM; SDS 0.4% (w/v); glycerol 10% (v/v); bromophenol blue 0.02% (w/v). Denaturation of proteins was carried out by heating 5 minutes at 100°C, followed by centrifugation at 13 000 *g*, 5 minutes to sediment insoluble debris. Depositions of samples were of 20 µL per well.

Proteins of reference (Fermentas, PageRuler Prestained Protein Ladder) of known molecular masses were also deposited on the gel for the determination of the molecular masses of polypeptides to analyze. The electrophoresis was carried out in the device MiniProtean3 (Biorad), in Laemmli buffer: Tris 25 mM; glycine 192 mM; SDS 0.1% (w/v); pH 8.6, with a voltage of 110 V. Migration of approximately 2h was necessary to obtain a good separation of proteins.

1.4.8.2. Staining of the gel by the Coomassie blue

After migration, the proteins separated on polyacrylamide gel can be fixed and revealed by staining with Coomassie blue. This method of staining has a limit of detection about 50-100 ng of proteins per spot.

After electrophoresis, the gel was washed with MilliQ water, then immersed approximately one hour, under agitation, in a staining solution. The excess dye was eliminated by an overnight wash with the washing solution under constant agitation. The gel was then deposited in a conservation solution containing glycerol. Acetic acid of all these solutions is a staining fixative.

The gel can then be dried between two sheets of cellophane. Drying was carried out by vacuum at 65°C and during approximately 3 hours. The gel stained and fixed was digitized with a Canonscan LIDE35 scanner (Canon) by an image acquisition software provided with the scanner.

Staining Solution	Coomassie Blue R250 0.2% (w/v)	Acetic Acid 10%	Ethanol 50%	qsp MilliQ water
Washing Solution	-	Acetic Acid 10%	Ethanol 30%	qsp MilliQ water
Conservation Solution	-	Acetic Acid 10%	Glycerol 10%	qsp MilliQ water

1.4.9. Test of activity

1.4.9.1. Substrate preparation for CYP98A3 (according to Morant *et al.*, 2007)

The substrate of CYP98A3, *p*-coumaroyl shikimate is not commercial. It was prepared at the laboratory by P. Ullmann.

Briefly, this ester was enzymatically prepared in two steps, using purified recombinant 4-coumaroyl-CoA ligase and hydroxycinnamoyl transferase. The vectors for the recombinant expression in *Escherichia coli* of the 4-coumaroyl-CoA ligase 1 (4-CL1) (Ehlting *et al.*, 1999) and the hydroxycinnamoyl transferase (HCT) (Hoffmann *et al.*, 2003) from *A. thaliana*, were kindly provided by C. Douglas (University of British Columbia, Vancouver, Canada) and M. Legrand (IBMP, Strasbourg), respectively. Each protein was purified according to these authors with some modifications (P. Ullmann, unpublished data). For the synthesis of the CoA ester, the reaction contained in a total volume of 1 ml of potassium phosphate buffer 50 mM pH 7, *p*-coumaric acid 0.4 mM (Sigma), $MgCl_2$ 2.5 mM, DTT 1 mM, ATP 2.5 mM, CoASH 0.4 mM and 0.025 mg of recombinant purified 4-CL1.

The first reaction (*p*-coumaroyl-CoA synthesis) was monitored spectrophotometrically at 333 nm for 5 to 20 minutes at room temperature and in darkness. When the absorbance value indicated a complete substrate conversion into the CoA ester, the second reaction (conversion into *p*-coumaroyl shikimate) was achieved in the same tube.

4 mM shikimic acid and 0.02 mg of recombinant purified HCT were added in a final volume of 1.2 mL. After 1 hour of incubation at 30°C in the dark, the reaction was stopped by addition of 4 M HCl (Total Volume/20). The products were extracted twice with one volume of ethyl acetate, dried under argon flow before resuspension in buffer (potassium phosphate 50 mM pH 7,4) and analysed by High Pressure Layer Chromatography (HPLC).

The final concentration of the *p*-coumaroyl shikimate was determined by spectrophotometry using the extinction coefficients determined by Kühnl *et al*. (1987): *p*-coumaroyl shikimate $\varepsilon_{340\ nm} = 6.2\ mM^{-1}.cm^{-1}$.

1.4.9.2. In vitro activity test

To avoid any risk of adsorption of the substrate on the walls of the tube likely to distort the calculated concentrations, incubations were always carried out in hemolysis tubes with borosilicated glass.

For each reaction a negative control without NADPH was carried out.

Substrate for CYP73A5 is *trans*-cinnamic acid and is commercially available. The substrates used during this thesis were generally in solution in dimethylsulfoxide (DMSO) or methanol.

Each susbtrate (100 µM final) in solution was initially deposited at the bottom of the tube and organic solvent used to solubilized substrate was let to evaporate. The potassium phosphate buffer 50 mM pH 7,4 was added; followed by DTT (20 mM) and finally the microsomes (10 µL). The reaction (total volume of 200 µL) was started by addition of NADPH (0.5 mM final). This mixture was incubated 30 minutes at 28°C, under agitation. Reaction was stopped by addition of HCl 1N (V/10).

1.4.9.3. Liquid/liquid extraction of metabolites

The products were extracted twice with one volume of ethyl acetate, dried under argon flow before resuspension in injection buffer (acetonitrile 10% (v/v), formic acid 0.2% (v/v)) and analyzed on HPLC.

1.4.9.4. High Performance Liquid Chromatography analysis

Metabolites in the injection buffer (injection of 100 µL) were separated by HPLC using a Waters HPLC system coupled to a photodiode array detector.

The analyses were achieved on a C-18 column in reverse phase (LiChrosorb RP-18 Merck, 4x125 mm, 5 µm). Detection was done in UV by a Diode array detector. The mobile phase consisted of a mixture of MilliQ water / acetonitrile and the flow was fixed at 1 $mL.min^{-1}$. The employed mobile phase gradient begins at 10% acetonitrile to finish at 90% acetonitrile supplemented by 0.2% formic acid.

Solvent A : 10%AcN, 0.2% formic acid
Solvent B : 90%AcN, 0.2% formic acid

Gradient

Step	Time (min)	Flow ($mL.min^{-1}$)	%A	%B
1	0,01	1	100	0
2	10	1	100	0
3	30	1	40	60
4	35	1	0	100
5	38	1	0	100
6	42	1	100	0
7	47	1	100	0

Determination of the activities were qualitative and not quantitative, the presence of a major product peak indicating that P450s were active.

2. Production of soluble partners: 4CL-1 and HCT

The expression of the soluble proteins was carried out in *Escherichia coli*. Expression and purification of the two proteins was already implemented in the laboratory before my arrival. I thus started this project using the "regular" protocols. The *4CL-1* gene was available in the pQE30 vector. This vector permits the expression of a His-tagged protein of interest under the control of a promoter inducible with IPTG. This vector was kindly provided by C. Douglas (University of British Columbia, Vancouver). The *HCT* gene was available in pGEX vector, which permits the expression of a GST-tagged protein under the control of a promoter inducible with IPTG. The GST-tag can be cleaved by thrombin due to the presence of a thrombin recognition site between the tag and the protein of interest. This vector was kindly provided by M. Legrand (IBMP).

Accordingly, purification of 4CL-1 was carried out by nickel affinity, due to the affinity of the histidines for divalent cations. HCT, expressed in fusion with a gluthation-*S*-transferase (GST) tag, was purified by exploiting the affinity of this tag for glutathione.

Since the beginning of the project, optimize expression and purification of HCT and 4CL-1 appeared essential. One of the objectives of my Master lab work was to optimize the expression of HCT and 4CL-1 using *E. coli* strain expressing various chaperone proteins. Unfortunately, expressions of both proteins turned out to be worse when co-expressed with any combination of chaperone proteins (including DnaK, DnaJ, GroEL, GroES and GrpE). Over-expression of the chaperone proteins was detected, but led to a decrease in the expression of the protein of interest (Figure 34). Bacteria seemed to preferentially use their machinery to produce chaperone proteins.

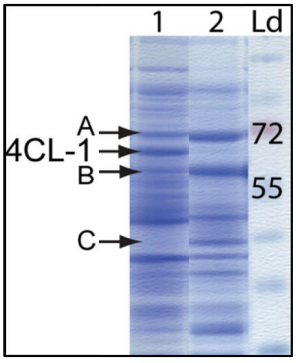


Figure 34. Co-expression of 4CL-1 with a set of chaperone proteins.

1. Expresssion of 4CL-1 without chaperone proteins; 2. Expression of 4CL-1 with chaperone proteins.

Ld. Protein Ladder (in kDaltons); A. DnaK; B. GroEL; C. DnaJ. The 4CL-1 band was checked by Mass Spectrometry. DnaK, GroEL and DnaJ are labeled according to theoretical masses of these proteins.

However, protein availability being essential for the project many efforts were dedicated to the improvement of protein expression and purification. As for P450s it was essential to start purification with large amounts of proteins and to decrease losses during purification steps.

2.1. Optimization of the soluble proteins expressions

In order to increase protein expression, various approaches were tested. They are partly summarized for 4CL-1 and for HCT in Figure 35.

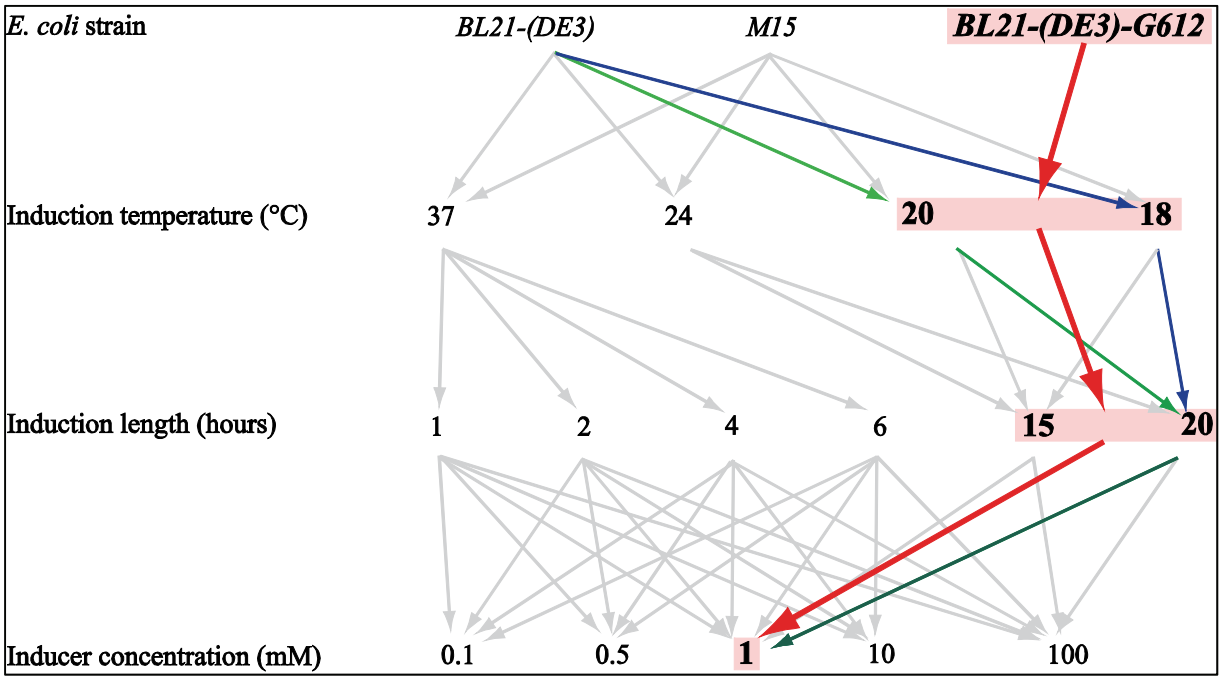


Figure 35. Flow-chart summarizing optimization of 4CL-1 and HCT expression.
In green: conditions used before optimization of 4CL-1 expression. In blue: conditions used before optimization of HCT expression. In grey: all tests leading to no significant improvement in protein expression. In red: tests resulting in significant improvement in protein expression (induction in the range of 18-20°C and 15-20 hours leads to large amounts of produced proteins).

HCT and 4CL-1 were initially expressed in the strain *BL21 (DE3)*. Many adjustments were tested with this strain:

- various kinetics of protein induction from 1 to 20 hours (results not shown).
- induction of protein expression with various concentrations of Isopropyl β-D-1-thiogalactopyranoside (IPTG), from 0.1 to 100 mM (results not shown).
- induction of protein expression at different temperatures 37°C to 18°C. Lower temperatures were expected to slow down protein synthesis, therefore favoring correct folding, and thus limiting formation of inclusion bodies (results not shown).
- induction at different bacterial ODs: from 0.3 to 1 $OD_{600\ nm}$ (data not shown).
- culture at different shaking speeds, from 90 to 170 rpm, before and during the protein expression (data not shown).

Various combinations of these approaches were tested and some of them led to dramatic decrease in expression. But no significant improvement of protein expression was obtained.

Another strain of bacteria was tested with the same approaches, the *M15* strain. It is a modified *BL21(DE3)* strain carrying the *pREP4* plasmid which encodes a repressor of the IPTG-inducible promoter. The later represses the basal expression of the recombinant proteins. No significant improvement was obtained.

Finally, a last strain (BL21-(DE3)-G612) was tested (Figure 35). This strain carries the plasmid *pLysS* and expresses rare prokaryotic tRNAs. *pLysS* is a plasmid that carries the T7 lysozyme gene (lysS). The T7 lysozyme protein has two effects: it degrades the cell wall and it inhibits T7 RNA polymerase. Because T7 lysozyme inhibits T7 RNA polymerase, it helps to silence expression of the *T7* promoter until addition of IPTG. This allows a better control of protein expression. *pLysS* has thus the same role as *pREP4*, consisting to avoid leakage of the promoter in case of expression of a toxic protein. With this strain, **a significant increase in protein production was observed.**

Minor adjustments were also tested with this last strain and sometimes kept, for instance, the cooling of the cultures before the induction. The final optimized protocol is described in Material and Methods.

In all tested conditions, the expression of HCT was always lower, less reproducible, and problematic in comparison to the expression of 4CL-1. Differences in codon usage between *HCT*, *4CL-1* and *E. coli* genome were checked. No striking difference was noticed.

To conclude, the most important parameter to obtain high expression of the two soluble proteins was the *E. coli* strain used for the expression of these proteins. A drastic increase in yield was noticed with the BL21-(DE3)-G612 strain. Other parameters had only minor effects.

2.2. Optimization of protein purification

To monitor purification efficiency, samples were collected at each purification step and analyzed by SDS-PAGE, total protein quantification (Bradford assay), and sometimes, test of activity and Western Blot analysis.

2.2.1. Optimization of 4CL-1 purification

At the start of the project, electrophoretic profiles of affinity purification eluates showed a major protein, devoid of contaminants. The major protein of apparent mass of 69 kDa was 4CL-1 (checked by Mass Spectrometry by J. Borch, Danish Biotechnology Instrument Center, Denmark). The concentration obtained after purification was highly variable from 0.4 to 2 $mg.mL^{-1}$, depending on initial expression of the protein. Large amounts of protein seemed lost at all the purification steps.

Various major and minor optimizations were tested to improve yield and concentration, of purified protein. For instance, bacterial cells were disrupted in liquid nitrogen with a mortar and pestle, then sonication was also tested to improve the protein extraction, but without success. Optimization of the other purification steps is summarized in Figure 36. Different approaches were used: purification with pre-packed columns, purification in batch, or purification with a hybrid system.

The procedure leading to the best protein yield and concentration was finally the hybrid system loaded with bacterial lysate from several 4CL-1 productions (up to 1.5 liters of culture).

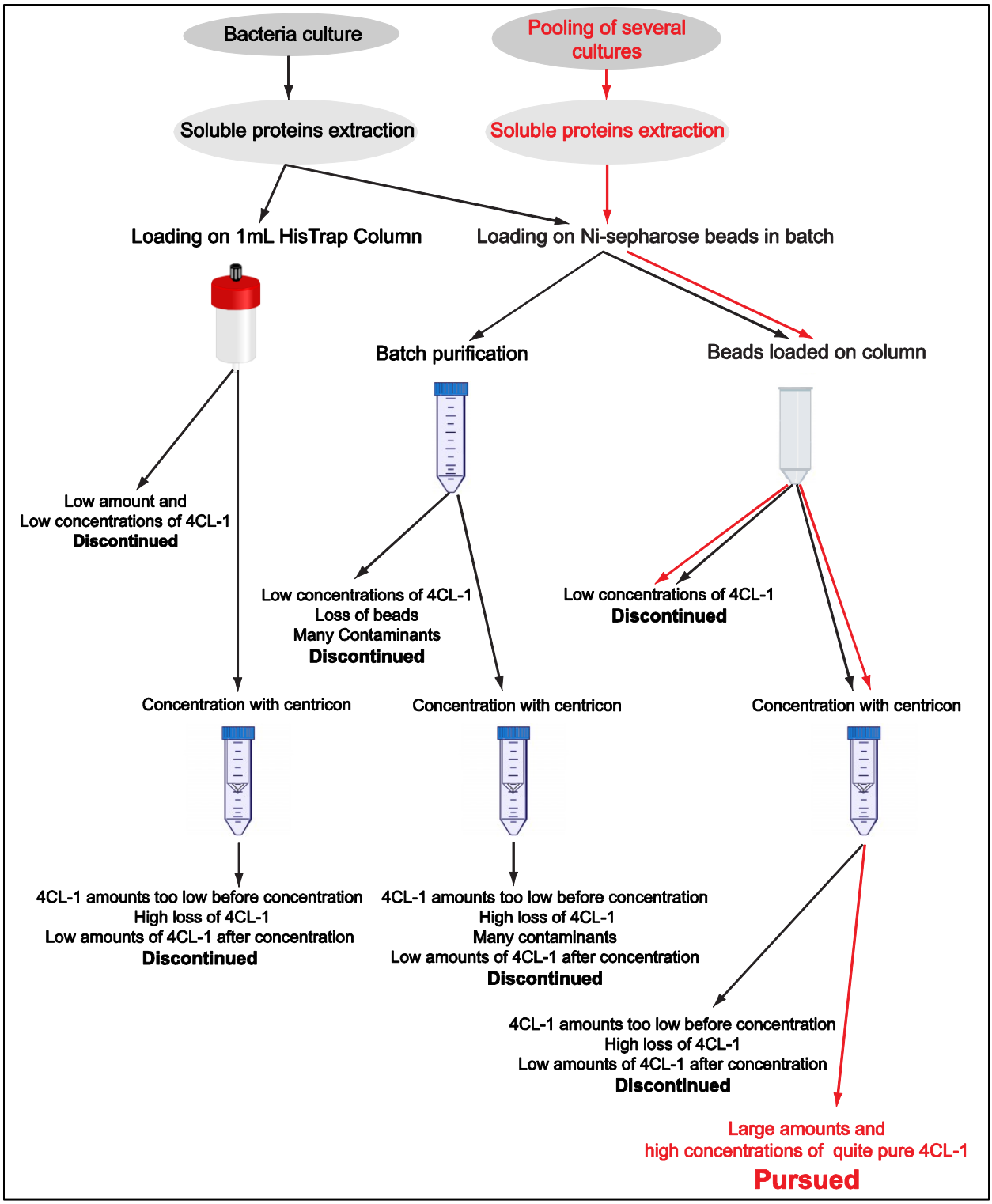
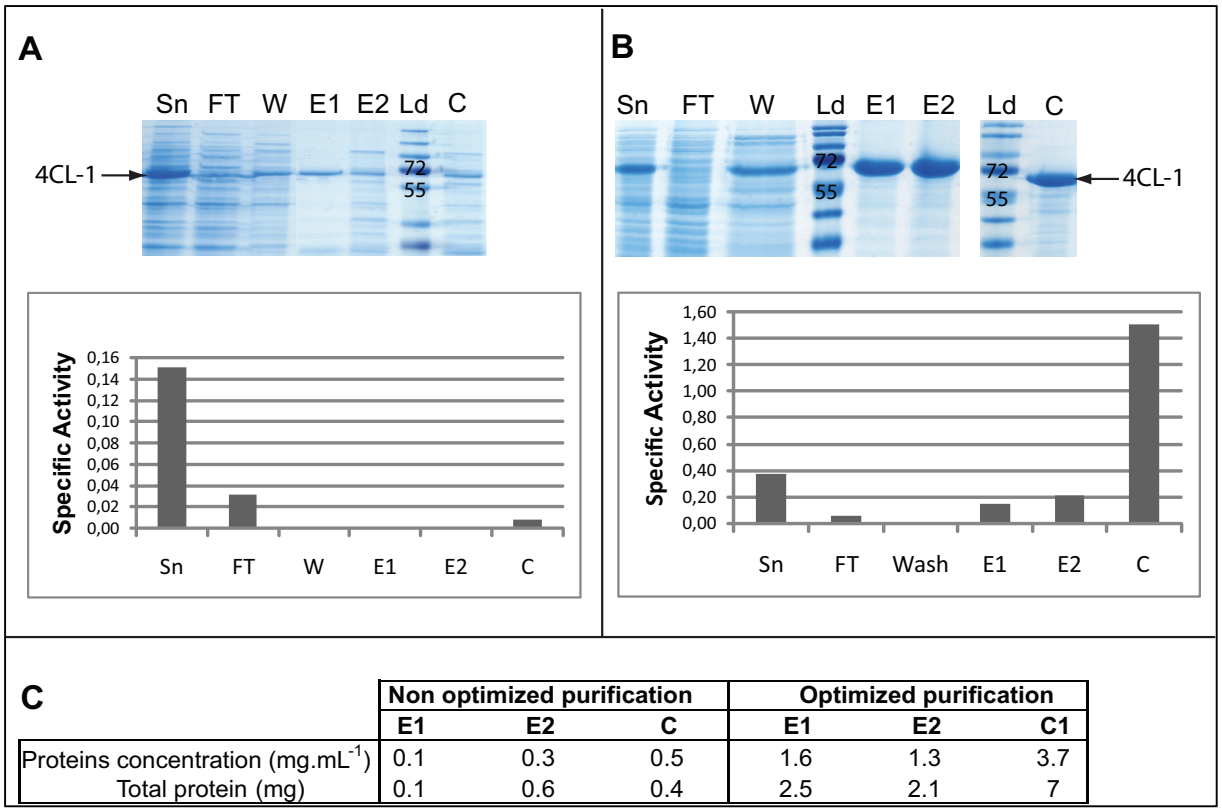


Figure 36. Flow-chart summarizing optimization of 4CL-1 purification.
In red, tests done by increasing the volume of culture used for purification.



C

	Non optimized purification			Optimized purification		
	E1	E2	C	E1	E2	C1
Proteins concentration (mg.mL^{-1})	0.1	0.3	0.5	1.6	1.3	3.7
Total protein (mg)	0.1	0.6	0.4	2.5	2.1	7

Figure 37. Representative 4CL-1 purifications.

A. Non-optimized purification. Production of 4CL-1 using optimized protocol. Lysate from 450 mL of culture was incubated overnight at 4°C with Ni-sepharose beads in batch, and poured onto a column for washing and elution. About 4 µg of total protein were analyzed in each lane by SDS-PAGE.

B. Optimized purification. Production of 4CL-1 using optimized protocol. Lysate from 1500 mL of culture was incubated over night at 4°C with Ni-sepharose beads in batch, and poured onto column. About 5 µg of total protein were analyzed in each lane by SDS-PAGE.

C. Comparison of protocols in A and B.

Gels were stained by Colloidal Blue. Specific Activities were expressed in mmoles of product.milligram^{-1} of protein.minute^{-1}. Abbreviations: Sn: Supernatant, FT: Flow Through, W: Wash, E1 and E2: Eluted fraction, Ld: Protein ladder in kDaltons, C: Concentrated fraction.

Using the same production protocol, with different purifications procedures, the amount of purified proteins can vary dramatically (Figure 37). In Figure 37A, a non-optimized protocol was used to purify 4CL-1. After purification and concentration, SDS-PAGE showed a major band of 4CL-1, but with many contaminants. Active protein was lost during the process and final concentration (after pooling and concentration of elutions) was about 0.5 mg.mL^{-1}. Yield was very low.

After optimization (Figure 37 B) high protein concentration was directly obtained in the elution fractions after affinity purification, e.g. 1.6 mg.mL^{-1} in fraction E1 (Figure 37 C). In one of the concentrated fractions, concentration was up to 3.7 mg.mL^{-1}. According to SDS-PAGE analysis, the major protein was 4CL-1, with very minor contaminants. Loss in active 4CL-1 (Figure 37 B) during the affinity purification was significantly decreased after optimization.

The His-tag used for affinity purification could interfere with the *in vitro* tests of protein-protein interactions. 4CL-1 coding sequence was thus cloned into the pQE30-Xa vector and expressed with the optimized protocol. This vector drives expression of 4CL-1 in fusion to a 6 Histidines tag, cleavable by the Factor Xa protease. Cleavage of the tag was tested directly on column, in batch, or after elution of the 4CL-1. Cleavage was inefficient (Results not shown). ***In vitro* tests of interactions were thus carried out with the His-tagged 4CL-1.**

2.2.2. Optimization of HCT purification

At the start of the project, SDS-PAGE of the GST affinity purified fractions obtained using the lab protocol indicated purification of HCT together with several major contaminants. The protein expected to be HCT at about 50 kDa was first confirmed to be HCT (checked by Mass Spectrometry by J. Borch, Danish Biotechnology Instrument Center, Denmark). Protein concentration obtained after purification was lower than 0.5 $mg.mL^{-1}$, and very low HCT activity was recovered after purification.

HCT purified using a GST-tag showed SDS-PAGE profiles more strongly contaminated than 4CL-1 purified using His-tag affinity. In consequence, the HCT coding sequence was recloned in pQE30-Xa vector in fusion with a His-tag for Ni^{2+} affinity purification. But no condition allowing significant production of protein could be obtained and this approach was discontinued. Optimizations were then focused on purification steps. Bacteria were disrupted in liquid nitrogen with mortar and pestle, or, sonication as for 4CL-1, but no significant difference was observed. Trials and choices for purification are summarized in the Figure 38.

The procedure leading to the best yields and final concentrations of HCT was the system of beads batch loaded onto a column using bacterial lysate from several bacterial cultures (up to 3 Liters of culture).

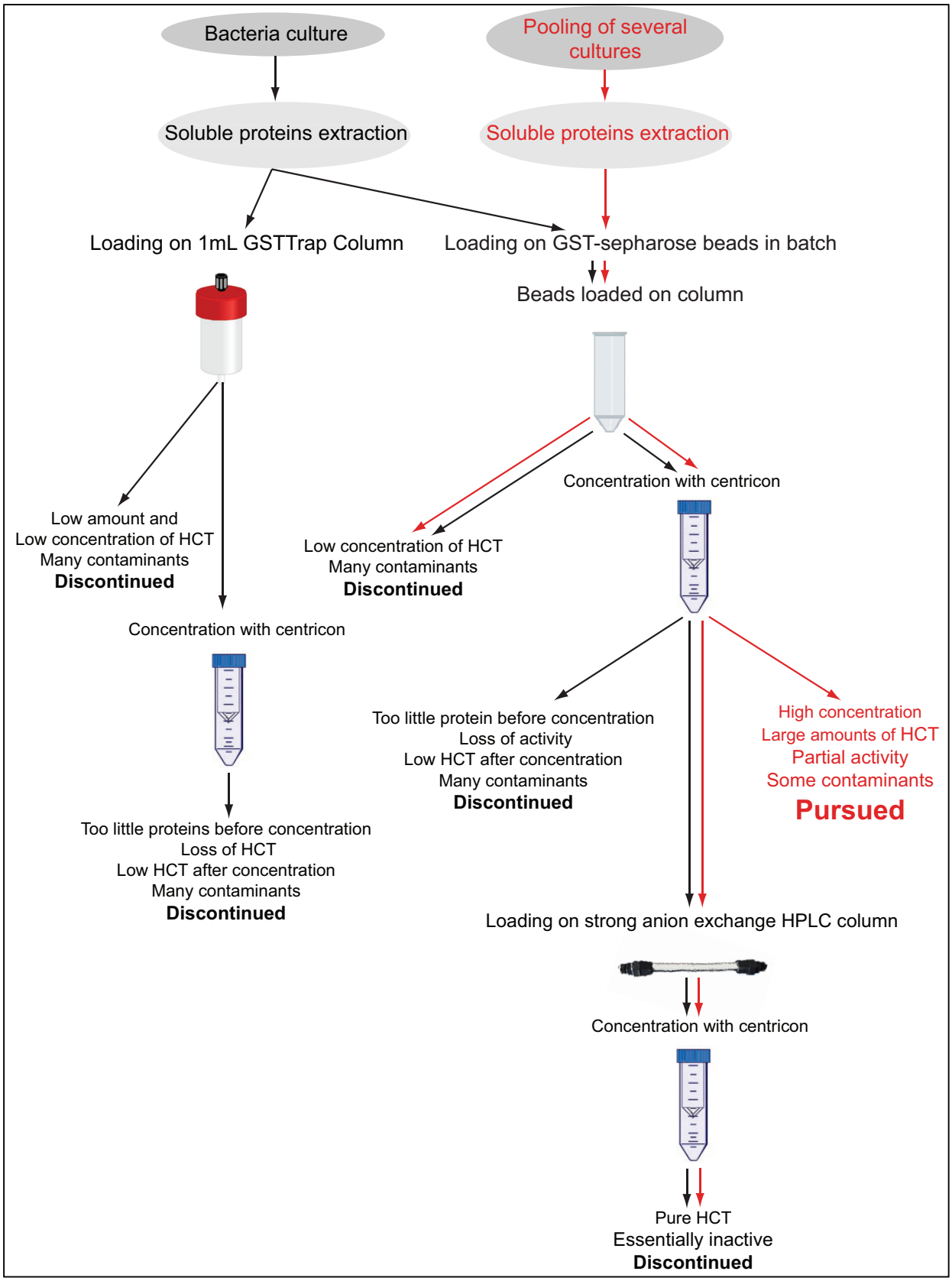
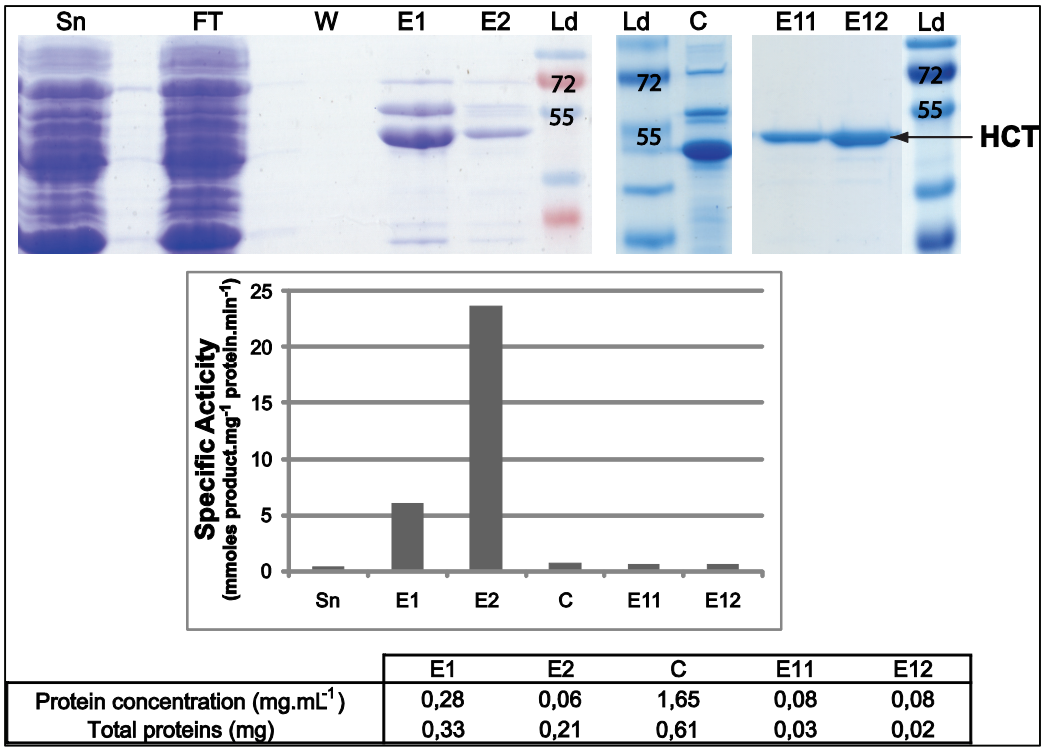


Figure 38. Flow-chart summarizing optimization of HCT purification.
In red, tests done by increasing the volume of culture used for purification.



	E1	E2	C	E11	E12
Protein concentration (mg.mL^{-1})	0,28	0,06	1,65	0,08	0,08
Total proteins (mg)	0,33	0,21	0,61	0,03	0,02

Figure 39. Follow-up of a typical HCT purification.

HCT-expressing bacteria lysate equivalent to 800 mL of culture was incubated overnight at 4°C with GST- sepharose beads in batch and poured onto a column for washing and elution. About 4 µg of total protein were loaded in each lane for SDS-PAGE analysis (except for wash fraction). Gels were stained with Coomassie Blue or Colloidal Blue. Specific activities were expressed in mmoles of product .milligram^{-1}protein.minute^{-1}. Abbreviations: Sn: Supernatant; FT: Flow Through; W: Wash; E1 and E2: Eluted fractions from affinity purification; E11, E12: Eluted fractions from anion exchange purification. Ld: Protein ladder in kDaltons, C: Concentrated fraction.

To improve HCT purity after elution from the affinity column, a second purification step with a strong anion exchanger was attempted (Figure 39). HCT was efficiently purified when two successive purifications were achieved. Unfortunately, enzyme activity was lost during the process. Whereas HCT protein enrichment was noticed in the two eluted fractions from the affinity purification (Figure 39), a loss in activity was observed during concentration and no recovery was detected after anion exchange purification. Proteins and even more activity were lost upon concentration. Total protein recovered after affinity purification was estimated about 842 µg. Only 610 µg were recovered after concentration, and only 85 µg eluted after anion exchange purification. Neither HCT protein, nor HCT activity were detected in the flow-through of the centricon (results not shown). Thus HCT was assumed to be adsorbed onto the centricon membrane. This second purification step was discontinued due to high loss of protein and activity.To pursue the project, *in vitro* interaction tests with active protein seemed essential. Because interactants are checked by Mass Spectrometry, it did not appear essential if potential interactants were not perfectly pure.

HCT expression could still deserve a better optimization to improve its purification output. The whole HCT purification process suffered from its low expression. In addition, this enzyme seemed very unstable and prone to activity loss.

2.3. Discussion-conclusions

The optimization of the expression of the soluble proteins turned out to be more complex than expected. 4CL-1 and HCT appeared poorly expressed in the tested *E. coli* strains except for BL21-(DE3)-G612.

Nevertheless, after optimization, up to **12.6 mg per liter of culture of pure 4CL-1** and **1 mg per liter of culture of quite pure HCT could be obtained**, at **final protein concentrations of 5 mg.ml^{-1} for 4CL-1 and 2 mg.ml^{-1} for HCT.** Those were provided to our partners for *in vitro* interaction tests. Thus, the **protein yield for a purification round was improved by more than ten for 4CL-1 and final protein concentration was increased more than two-fold**. Further concentration is now limited by spontaneous aggregation of the protein upon freezing/defreezing when protein concentrations reach around 5-6 mg.ml^{-1}. **For HCT, protein production was not significantly improved, but the purified protein concentration obtained after purification was increased by two.**

In summary, **the most important parameter for improving production of the pure proteins was the *E. coli* strain used for expression**. HCT expression still requires a further optimization to improve its purification output. The main problem encountered for high HCT expression could be due to the formation of inclusion bodies not correctly solubilized during extraction. This was so far not investigated. In addition, HCT was still highly contaminated after purification. This phenomenon was already described and is typical for proteins purified using the GST/Glutathione affinity system. Bacterial protein contaminants were reported to co-purify, such as chaperone proteins (Rohman & Harrison-Lavoie, 2000). This problem can be solved using another tag system for protein purification. This was tested by expressing HCT with pQE30Xa vector to achieve His-tagged protein purification. But troubles of expression prevented to pursue in this way.

While 4CL-1 concentration, quantity and quality was sufficient to pursue our project, **HCT concentration was possibly to low and limiting to detect interactions**. Considering the time spent for the production optimization of the enzymes, optimization tests were stopped. Main efforts were then focused on protein-protein interactions tests.

2.4. Materials and Methods

Samples of alive materials to analyze, after sampling, were frozen with liquid nitrogen and were immediately stored at -80°C, before grinding with liquid nitrogen and analysis.

2.4.1. Biological materials: *Escherichia coli*

Soluble proteins, which do not require a specific membrane environment, were expressed in *Escherichia coli* bacterial strain.

2.4.1.1. Escherichia coli strains used for the cloning and propagation of the plasmids

+DH5α (Invitrogen)
+TOP10 (Invitrogen)
Blue/White screening can be done with these strains.

2.4.1.2. Escherichia coli strains used for the soluble protein expression (HCT and 4CL-1)

+M15 (Qiagen)
This strain carries *pREP4* plasmid which represses basal expression of recombinant proteins to prevent their toxicity on *E. coli*. A marker of selection (kanamycin) is present in the plasmid.

+BL21 (DE3)

These strains are lysogenic for a λ (DE3) prophage carrying gene coding for the T7 RNA polymerase, under the control of the isopropyl *β*-D-thiogalactoside (IPTG) inducible *LacZ* promoter.

-BL21 (DE3) (Stratagene)
-BL21 (DE3)-G612 (kindly provided by M. Legrand, IBMP, Strasbourg).
Cells carry the plasmid *pLysS* and express rare prokaryotic tRNAs (for codons: AGA, AGG, AUA, CUA, GGA, CCC, CGG). Plasmid *pLysS*, that encodes T7 lysozyme, represses basal basal expression of target genes by a *trans* repression of the T7 RNA polymerase promoter. This strain is resistant to kanamycin.

All the stocks were preserved at -80°C in the culture medium (without antibiotic) containing 20% glycerol.

2.4.1.3. Culture conditions

The bacteria were grown either in LB liquid medium under agitation at 170 rpm, or on a LB agar medium in Petri dishes. Growth of the bacteria for propagation of the plasmid or during transformation was conducted at 37°C. Times of culture depend on the experiment carried out. To the liquid or solid media of culture were added the adequate antibiotics necessary to the selection of the strain and the plasmid used during the transformation.

2.4.2. Bacterial expression vectors

For the three first vectors cited below, expression of proteins of interest is induced by addition of Isopropyl *β*-D-1-thiogalactopyranoside (IPTG). A marker of selection (ampicillin) is present in these three plasmids, and a promoter inducible by IPTG drives the expression of the protein of interest cloned into the multiple cloning site.

+ pGEX-KG (GE Healthcare)

The pGEX-KG vector was used for expression of N-terminally Glutathione-S-Transferase-tagged proteins. This vector encodes a Thrombin protease recognition site on the 3' side of the GST-tag. Thrombin protease treatment results in a recombinant protein without GST-tag at the N-terminus.

+ pQE30 (Qiagen)

The pQE30 vector was used for expression of N-terminally 6xHis-tagged proteins. This vector is designed with the multiple cloning site on the 3' side of the 6xHis-tag coding region.

+ pQE30-Xa (Qiagen)

The pQE30-Xa vector is similar to the pQE30, but also encodes a factor Xa protease recognition site with the 6xHis-tag coding region on the 5' side of the recognition site, and the multiple cloning site on the 3' side. When the gene of interest is directly cloned behind the factor Xa recognition site, Factor Xa Protease treatment results in a recombinant protein without any vector-derived amino acids at the N-terminus. After the digestion, Factor Xa Protease is removed using Xa removal resin (Qiagen).

+ pG-KJE8 (Takara) (11.1kb)

This plasmid was utilized during the improvement expression tests of 4CL-1 and HCT. It can express the chaperone proteins DnaK, DnaJ, GrpE and GroEL, GroES, under the control of the promoters *araB* (induction with the arabinose) and *Pzt1* (induction with tetracycline), respectively. pG-KJE8 derived from the plasmid pG-KJE6 (Nishihara *et al.*, 2000). It carries a resistance gene to chloramphenicol and an origin of bacterial replication (pACYC ori).

2.4.3. Extraction and purification of plasmidic DNA

The bacteria resulting from LB culture were collected by centrifugation and were used to carry out the plasmidic DNA extraction, by using purification kits of the providers Fermentas or Macherey-Nagel. Briefly, these kits are based on the protocol of mini-preparation by alkaline lysis (Birnboim & Doly, 1979) followed by plasmidic DNA purification on affinity columns.

For mini-preparation, bacteria cultures of 5 mL selective medium (LB supplemented with adequate antibiotics) were grown overnight at 37°C and agitation. The bacteria were pelleted during 10 minutes at 5500 *g* before extraction of the DNA. For the elution of plasmidic DNA from the column, a double elution with the elution buffer pre-heated at 72°C was done with a volume of 40 µL, to have the best possible output. The extractions were carried out following the instructions of the providers.

For midi-preparation, colonies were grown in 50 mL LB added with adequate antibiotics at 37°C overnight with agitation. The bacteria were harvested by centrifugation (5500 *g* for 10 minutes) before extraction of the DNA. The extractions were carried out following the instructions of the providers.

The purified plasmidic DNA was quantified by spectrophotometry (Biophotometer, Eppendorf). One unit of absorbance corresponds to a concentration of 50 $\mu g.mL^{-1}$ of double strand DNA. The purified plasmids are then stored at -30°C.

2.4.4. Bacteria transformation

E.coli bacteria were first chemically treated to make them competent. A 42°C heat shock is then required to make competent cells able to take up exogenous DNA.

2.4.4.1. Preparation of chemical (ultra)competent cells (based on Inoue *et al.*, 1990)

Bacteria were grown at 18°C in 250 mL of SOB until obtaining OD_{600} = 0.6. These bacteria were then put on ice during 10 minutes, centrifuged 10 minutes at 2500 *g*. The pellet was resuspended in 80 mL of ice-cold transformation buffer (TB), then left on ice during 10 minutes and centrifuged as previously. This new pellet was resuspended in 20 mL of ice-cold TB added with 1.5 mL of DMSO, left on ice 10 minutes, and then centrifuged as previously. DMSO allows, with the Ca^{2+} ions, to permeabilize the bacterial membrane during the heat-shock of the transformation.

Bacteria solution was divided in 200 µL aliquots, frozen with liquid nitrogen and preserved at -80°C before their use for transformation.

SOB	Bactotryptone 2 $g.L^{-1}$	Yeast extract 0.5$g.L^{-1}$	NaCl 10 mM	$MgSO_4$ 10 mM	$MgCl_2$ 10 mM	KCl 2.5 mM	Adjust pH to 7.5 with KOH	Sterilize at 1 bar, 120°C, 20 min
TB	Piperazine-1,4-bis(2-ethanesulfonic acid) 10 mM	$MnCl_2$ 55 mM	$CaCl_2$ 15 mM	KCl 250 mM	Adjust pH to 6.7 with KOH			

2.4.4.2. Transformation of bacteria

The transformation of the competent bacteria was carried out according to a modification of the protocol of Bergmans (Bergmans *et al.*, 1981). The transformation is done by a heat-shock.

The plasmidic DNA of interest was added (1-10 ng) to 50 µL of ultracompetent bacteria thawed on ice. The mixture was left 10-30 minutes on ice (for the adsorption of the DNA on the surface of the bacteria) before carrying out the heat-shock (1.5 minutes at 42°C). This heat-shock as well as the presence of DMSO and the chemical treatment make the cell membrane porous, allowing DNA to enter in the cell. After 3-4 minutes on ice, 950 µL of LB medium were added; and finally the whole is incubated 1 hour at 37°C under agitation.

This incubation allows the bacteria to restore their metabolism, to initiate their growth, to multiply the plasmid but especially to express antibiotic resistance gene present on the plasmid. The bacteria were centrifuged to discard 800 µL of medium. The remaining 200 µL were used to resuspend the bacteria and then to spread out the bacteria onto solid selective medium (LB supplemented by adequate antibiotics and agar 15 $g.L^{-1}$). The Petri plates were left for growth overnight at 37°C. Under these conditions, only the bacteria having incorporated a recombinant plasmid multiply. The culture plates were then stored at 4°C for less than one month.

2.4.4.3. Screening of the clones

The colonies were subcultured on solid selective LB agar using a sterile tip. After subculture, the tip was rinsed during 3-4 minutes in 20 µL of PCR mix supplemented in Tween-20 0.05% (v/v) and containing the adequate primers (Appendix: List of oligonucleotide primers and PCR Protocols, p. 312-317). The residual bacteria, lysed in the presence of detergent and the heat during the first step of PCR, released their plasmids in the medium which are used as matrix for PCR amplification. The screening of the clones was carried out starting from these amplification products. After PCR, amplicons were resolved by electrophoresis on agarose gel in order to check the size of the amplified fragments, compared to DNAs of known mass (MassRuler DNA ladder - Fermentas).

After identification of the positive clones, liquid cultures were prepared to multiply the plasmids of interest before carrying out an extraction followed by a purification of the plasmidic DNA by mini- or midi-preparation.

2.4.4.3.1. Polymerase Chain Reaction (PCR) on colony

During optimization steps for the production of soluble proteins, PCR amplifications were done on recombinant bacterial colonies, to check the presence of recombinant plasmids of interest. For the screening of the transformed colonies and the checking of the cultures, Taq polymerase (Invitrogen) was used. This Taq polymerase was isolated from the bacterium *Thermus aquaticus*. It has a relatively high error rate ($2,6.10^{-5}$) but is sufficient for screening.

PCR amplifications followed the same general outline: a matrix (cDNA or plasmid) was mixed in the PCR buffer added by 100 µM of each dNTP and 10 pmoles of each selected specific primer. Amplifications proceeded first with one denaturation step at 95°C (during 8 minutes for PCR on colony, 2 minutes for PCR on plasmid), and then followed by amplification cycles.

For each amplification cycle, the protocol was as follows:

- denaturation at 95°C for 30 s
- hybridization of the primers at 55-60°C for 0.5 minute (according to the purpose of the amplification and the optimal temperature of hybridization of the used primers)
- elongation at 72°C for 0.5-2 min (according to the size of the amplicon to obtain).

The reaction is completed by a final elongation at 72°C for 10 min.

Amplifications were done in a thermocycler Master Cycler (Eppendorf). The specific primers used were synthetic oligonucleotides (Sigma). Lists of primers and PCR programs and PCR buffer used are in Appendix: List of oligonucleotide primers and PCR Protocols, p. 315-317.

2.4.4.3.2. Analysis of nucleic acids - Electrophoresis analysis on agarose gel

By this technique DNA fragments were separated according to their size. Loading buffer (6x mass loading dye solution, Fermentas) was added to the PCR mix (1x final concentration), then solution was loaded on agarose gel with 1,5% (w/v) agarose in TAE buffer (40 mM TrisAcetate, 1 mM EDTA), containing ethidium bromide in a final concentration of 0.5 $\mu g.mL^{-1}$. Ethidium bromide is a fluorescent dye, which intercalates between the bases of DNA. The fluorescent yield of the dye/DNA complex is much higher than the unbound dye. DNA is then visualized under ultraviolet rays. The electrophoresis was carried out in TAE buffer under a voltage of 90-110 V. The stained gel was photographed under UV at 254 nm.

The approximate size of the DNA fragments can be estimated by comparison with the migration on the gel of a standard solution containing several DNA fragments of known lengths (MassRuler DNA Ladder, Fermentas).

2.4.5. Production of 4-CL1 (optimized protocol)

Sampling was done at each stage of the production (expression and purification) to follow the good course of the production. Samples were immediately analyzed (for PCR) or frozen and stored at -30°C before analysis (Bradford, SDS-PAGE, Western Blot and test of activity).

2.4.5.1. Heterologous expression of 4-CL1 (optimized protocol based on P. Ullmann, unpublished data)

4CL-1 cloned into pQE30 (gift of C. Douglas, University of British Columbia, Vancouver, Canada) or cloned into pQE30Xa was expressed in BL21 (DE3)-G612 *E. coli* strain. Pre-cultures were inoculated from colonies from a freshly streaked selective plate (no older than 5 days). Pre-cultures and cultures were carried out under agitation at 160 rpm and 125 rpm respectively and contain antibiotics of selection.

The first day, 20 mL of liquid LB supplemented with ampicillin at 100 $mg.L^{-1}$ (for plasmid selection) and kanamycin at 50 $mg.L^{-1}$ (for *E.coli* strain selection) were inoculated with an isolated colony from LB Petri plate. An overnight culture of 16h at 37°C under agitation was carried out. The second day, 6 mL of the pre-culture was sub-cultured in 300 mL of LB selective medium (ampicillin at 50 $mg.L^{-1}$ and kanamycin at 50 $mg.L^{-1}$) in a one liter Erlenmeyer. When culture reached an OD_{600} of 0.4, culture was cooled during 30 minutes in cold room at 6°C. Then the expression of the protein of interest was induced by addition of IPTG 1 mM (final concentration). The culture was incubated at 20°C during 20 hours, and then centrifuged at 4500 *g*, during 10 minutes. Bacterial pellets were frozen with liquid nitrogen and preserved at -80°C until the purification.

2.4.5.2. Purification of 4CL-1

The purification is based on an affinity chromatography for Ni^{2+} of a 6 histidines tag at N-terminal end of the 4CL-1. All the buffers used during 4-CL1 purification were filtered by filters with pores of 0.22 µm, to avoid clogging of the columns.

2.4.5.2.1. Beads preparation

A volume of 2.5 mL of sepharose beads were used for an equivalent of bacterial lysate coming from a culture of 500 mL. Sepharose beads (GE Healthcare) were incubated overnight at 4°C in 15 mL $NiSO_4$ 0.2 M. Then beads were washed by 10 volumes of MilliQ water, 10 volumes of Elution Buffer and 20 volumes of Lysis Buffer. Washes were conducted by centrifugation at 500 *g*, 4°C, during 5 minutes and the supernatants were discarded at each stage.

2.4.5.2.2. Bacteria lysis

The frozen sample of one culture was resuspended in the Lysis Buffer (10 mL of buffer per 100 mL of culture), and the mixture was reduced to the form of a fine powder, under liquid nitrogen with mortar and pestle. Powder was warmed up (gently) under agitation to liquefy the lysate for centrifugation. The lysate was then centrifuged 1h30 at 15000 *g* and 4°C. The supernatant was filtered on Miracloth (200 µm pores) before loading onto the sepharose beads.

2.4.5.2.3. Affinity purification

The filtrate was added to an adequate volume of prepared beads and it was incubated during 20 hours with a wheel (for agitation) at 4°C. The solution was then loaded in a polypropylene column (Polyprep, Biorad). The beads were washed with 50 mL of Washing Buffer. Elution was carried out with 20 mL of Elution Buffer containing 500 mM imidazole. The eluate was concentrated and the buffer changed.

2.4.5.3. Concentration of 4CL-1

Concentrations were carried out immediately after the purification. Concentrations to concentrate the proteins but also to eliminate imidazole were done with Centricon (Amicon ultra 30K 15 mL). Elution fractions of purifications from the equivalent of 1.5 L of initial cultures were used with one centricon. Loading of eluates and six washes (with Conservation Buffer) were done carefully, to prevent local over-concentration which leads to the formation of aggregates. Loading and washes were done by centrifugations at 4000 *g*, 4°C during 1-5 minutes. Proteins were finally concentrated in 4-5 mL and glycerol was added to have a final concentration at 10% (v/v). An aliquot of the concentrate was kept for further analyses, to prevent frost/defrost of the fractions which was used for *in vitro* tests of interaction. The fractions were stored at -30°C.

Lysis Buffer	NaH_2PO_4 50 mM, pH7.4	NaCl 500 mM	Imidazole 5 mM	qsp MilliQ water	1 tablet of Coktail Antiprotease EDTA-free (Roche) for 100 mL buffer
Washing Buffer	NaH_2PO_4 50 mM, pH7.4	NaCl 500 mM	Imidazole 50 mM	qsp MilliQ water	-
Elution Buffer	NaH_2PO_4 50 mM, pH7.4	NaCl 500 mM	Imidazole 500 mM	qsp MilliQ water	-
Conservation Buffer	Potassium phosphate 50 mM pH7.4	-	-	qsp MilliQ water	1 tablet of Coktail Antiprotease EDTA-free (Roche) for 100 mL buffer

2.4.6. Production of HCT (optimized protocol)

Sampling was done at each stage of the production (expression and purification) to follow the quality of these one. Samples were immediately analyzed (for PCR) or frozen and stored at -30°C before analysis (Bradford, SDS-PAGE, Western Blot and test of activity).

2.4.6.1. Heterologous expression of HCT (optimized protocol based on Hoffmann et al., 2003)

HCT cloned into pGEX-KG (kindly provided by M. Legrand, IBMP) was expressed in BL21 (DE3)-G612 *E. coli* strain. Pre-cultures were started from colonies of fresh bacteria transformations (no more five days old). Pre-cultures and cultures were carried out under agitation at 160 rpm and 125 rpm respectively and contained antibiotics of selection.

The first day, 20 mL of liquid LB supplemented with ampicillin at 100 $mg.L^{-1}$ (for plasmid selection) and kanamycin at 50 $mg.L^{-1}$ (for *E. coli* strain selection) were inoculated with an isolated colony from LB Petri plate. An overnight culture of 16h at 37°C under agitation was prepared. The second day, 5 mL of the pre-culture was sub-cultured in 200 mL LB selective medium (ampicillin at 100 $mg.L^{-1}$ and kanamycin at 50 $mg.L^{-1}$) in a one liter Erlenmeyer. When culture reached an OD_{600} of 0.6, culture was cooled during 30 minutes in cold room at 6°C. Then the expression of protein of interest was induced by addition of IPTG 1 mM (final concentration). The culture was incubated at 18°C for 20 hours, and then centrifuged at 4500 *g*, during 10 minutes. Bacterial pellets were frozen with liquid nitrogen and preserved at -80°C until the purification.

2.4.6.2. Affinity purification of HCT

The purification is based on an affinity chromatography for glutathione of a Glutathione *S*-transferase tag at the N-terminal end of HCT. The Glutathione *S*-transferase catalyzes the conjugation of a tripeptide, glutathione (γGlu-Cys-Gly) with an electrophile substrate. GST at the N-terminus of the HCT binds to matrices bound of glutathione on the affinity column and was eluated, in our case, by using thrombin whose site of cleavage is present between the sequences of the GST and the HCT.

All the buffers used during HCT purifications were ultrafiltered (0.22 µm), to avoid clogging of the columns.

2.4.6.2.1. Glutathione-sepharose beads preparation

A volume of 1 mL of Glutathione-Sepharose beads (GE Healthcare) were used for bacterial lysate from a culture volume equivalent to 1 L. Glutathione-Sepharose beads were incubated overnight in 10 mL of PBS at 4°C. Then beads were washed by 10 volumes of MilliQ water, 10 volumes of PBS and 20 volumes of Lysis Buffer. Washes were conducted by centrifugation at 500 *g*, 4°C during 5 minutes. The supernatants were discarded at each stage.

2.4.6.2.2. Bacteria lysis

The pellet of bacteria was washed by 15 mL of PBS and centrifuged 30 minutes at 10 000 *g* and 4°C. Lysis Buffer was added to the bacterial pellet (10 mL of buffer per 300 mL of culture and one tablet of antiprotease cocktail for 100 mL of buffer), to carry out the grinding of the bacteria under liquid nitrogen with mortar and pestle. Powder was warmed up (gently) under agitation to liquefy the lysate for centrifugation. The lysate was then centrifuged 1h30 at 15000 *g* and 4°C. Supernatant was filtered on Miracloth (200 µm pores) before loading and one tablet of EDTA-free antiprotease (Roche) was added for 100 mL of lysate.

2.4.6.2.3. Purification

The filtrate was added to an adequate volume of prepared beads and incubated during 5 hours with a wheel (for agitation) at 4°C. The solution was then loaded in a polypropylene column (Poly-prep, Biorad). The beads were washed with 50 mL of PBS then by 50 mL of phosphate buffer.

Elution was carried out by cutting the GST-tag with thrombin. Two milliliters of phosphate buffer with 80 Units of thrombin were incubated 15 h at 4°C with agitation (on wheel) directly on column. The first eluate (2 mL) was recovered by gravity. To finish the cutting, a new incubation (2 mL phosphate buffer added of 80 U of thrombin) at room temperature during 2 hours (on wheel) was done. The second eluate (2 mL) was recovered by gravity and washed with 10 mL of phosphate buffer to recover completely the HCT.

2.4.6.3. Concentration of HCT

Concentrations were done immediately after the purification. Concentrations were carried out with Centricon (Amicon ultra 30K 15 mL). Eluted fractions of purifications from the equivalent of 3 L of initial cultures were used with one centricon. Loading of eluates and four washes (with Tris Buffer) were done carefully, to prevent local over-concentration which conducts to formation of aggregates. Loading and whashes were done by centrifugations at 4000 *g*, 4°C for 1-8 minutes. Proteins finally were concentrated to 400 µL and glycerol was added to have a final concentration at 10% (v/v). An aliquot of each concentrate was kept for further analysis, to prevent frost/defrost of the fraction which will be used for *in vitro* tests of interaction. The fractions were stored at -30°C.

PBS buffer	NaCl 8 $g.L^{-1}$	KCl 0.2 $g.L^{-1}$	Na_2HPO_4 1.44 $g.L^{-1}$	KH_2PO_4 0.24 $g.L^{-1}$	qsp MilliQ water, pH7,4
Lysis buffer	Triton X100 1% (v/v)	EDTA 1 mM	β-mercaptoethanol 0.1% (v/v)	PBS Buffer	qsp MilliQ water, pH 7,4
Phosphate buffer	77.4 mL Na_2HPO_4 1M	22.6 mL NaH_2PO_4 1M	-	-	qsp MilliQ water, pH 7.4
Tris Buffer	50 mM Tris-HCl pH8	-	-	-	-

2.4.7. Determination of total proteins: Bradford assay

Quantifications of total proteins were done for all the different steps of the production from the lysate to the concentrated fractions. Quantifications were done as for the quantification of total proteins during the P450s production (Cf. § Dosage of total proteins: Bradford assay p.91).

2.4.8. Detection of proteins on polyacrylamide gel

2.4.8.1. Chromatography of proteins on polyacrylamide gel

Polyacrylamide gels were prepared as described in the P450s production paragraph (Cf. § Chromatography of proteins on polyacrylamide gel p.91).

For colloidal blue staining, 5 µg of proteins were deposited on the gel, except for fractions low concentrated. For Western Blot detection of the proteins of interest, 1µg of proteins were deposited on the gel (except for the fractions too low concentrated).

2.4.8.2. Staining of the gel by Colloidal blue (based on Neuhoff *et al.*, 1988)

After migration, the proteins separated on polyacrylamide gel can be fixed and stained with Colloidal Blue G250. This method of staining has a limit of detection about 10-50 ng of proteins by spot.

After electrophoresis, gels were washed with MilliQ water, and then fixed by three Fixing Solution baths of 30 minutes under agitation. The excess of Fixing Solution was eliminated by three successive baths of 20 minutes in 2% (v/v) phosphoric acid. The gels were equilibrated during 30 minutes in the Equilibration Solution and finally 1% (v/v) Staining Solution was added. Stain was carried out during 24-72h.

The gel thus stained and fixed was recorded as digitized image using a Canonscan LIDE35 scanner (Canon) by an image acquisition software provided with the scanner.

Fixing Solution	30% (v/v) Ethanol	2% (v/v) Phosphoric Acid	-	qsp MilliQ water
Equilibration Solution	18% (v/v) Ethanol	2% (v/v) Phosphoric Acid	15% (w/v) Ammonium Sulfate	qsp MilliQ water
Staining Solution	2% (w/v) Brilliant Blue G250	0.2 % (w/v) Sodium Azide	-	qsp MilliQ water

2.4.8.3. Western Blot analysis

The polypeptides separated by electrophoresis on polyacrylamide gel in denaturing conditions were transferred electrically on a membrane of nitrocellulose (Schleicher & Schuell). The primary antibodies raised against the protein of interest were rabbit polyclonal antibodies. The secondary antibodies (monoclonal) from mouse were coupled to peroxidase.

2.4.8.3.1. Transfer of the protein to the nitrocellulose membrane

After a wash of the gel in MilliQ water, the gel was equilibrated during 2x10 minutes in the Transfer Buffer (Laemmli Buffer (Cf. § Chromatography of proteins on polyacrylamide gel p.91) supplemented by methanol 20% (v/v)). During the same time, nitrocellulose membrane and Whatman 3M papers were incubated in the same buffer. The transfer was carried out in a semi-dry system (Biorad) (Towbin *et al.*, 1979). The transfer was done during 25 minutes under a constant voltage of 15 V.

Immediately after transfer, the nitrocellulose membrane was rinsed with MilliQ water, as well as for the source gel. This gel was revealed by colloidal staining to check that polypeptides were efficiently transferred.

2.4.8.3.2. Primary antibody incubation

After rinsing, nitrocellulose membrane were saturated, thus the membrane was incubated one hour at room temperature in the TBST buffer supplemented by powder milk (2% w/v) and Tween 20 (2% v/v). The excess milk and Tween were eliminated by a wash of TBST and a bath of 15 minutes with agitation in TBST buffer.

The proteins of interest fixed on nitrocellulose were detected by rabbit primary antibodies. The membrane was incubated 16 h at 4°C under agitation with the primary antibody diluted at the desired concentration in the TBST buffer supplemented by powder milk 2% (w/v). Then the membrane was rinsed 3x10 minutes at room temperature in the TBST buffer to eliminate excess of primary antibody.

2.4.8.3.3. Secondary antibody incubation

The antibodies recognizing proteins of interest were detected by peroxidase conjugated mouse anti-rabbit secondary antibodies (BIOSYS).

The nitrocellulose membrane was treated during one hour at room temperature with these secondary antibodies (working dilution of 1:10000) in the TBST buffer, then rinsed 2x5 min in the TBST and 2x10min in the TBS buffer.

At this stage, the immobilized complex antigen-primary antibody-secondary antibody was revealed by chemiluminescence in a medium allowing the action of peroxidase. The Kit Immun-Star Chemiluminescent Protein Detection Systems (Biorad) was used. Substrates and membrane were incubated 5 minutes at room temperature before exposure.

The emitted signal was then visualized by exposition of the membrane on an autoradiographic film during 5 to 45 s at room temperature. After development, film was digitized as previously described for the stained gels.

TBS buffer	Tris 10 mM pH7.5	NaCl 9 $g.L^{-1}$	-	qsp MilliQ water
TBST buffer	Tris 10 mM pH7.5	NaCl 9 $g.L^{-1}$	Tween 20 0.05% (v/v)	qsp MilliQ water

2.4.8.3.4. Purification of primary antibodies by acetonic powder

This procedure was applied when antibodies were prepared from, and Western-blots done using purified bacterial proteins. This allows to decrease the aspecific bands due to the elimination of antibodies directed against bacterial proteins other than proteins of interest.

2.4.8.3.4.1. Acetonic powder preparation

A culture of one liter of BL21 (DE3)-G612 strain in LB added by adequate antibiotics was grown during one day at 37°C. Cells were centrifuged 15 min at 4000 *g*, 4°C. The pellet was resuspended in PBS buffer, and then centrifuged 15 min at 4000 *g*, 4°C. This pellet was frozen with liquid nitrogen and bacteria were ground manually in mortar with liquid nitrogen and pestle. The powder was warmed until a filamentous mixture was obtained. Then 50 mL of acetone were added and the grinding was pursued to dissolve the filaments.

A filtration was done to eliminate the solvent (with a vacuum pump); the mixture was ground again with acetone and filtered again. These steps were repeated until the brownish color disappeared.

The powder was recovered and left in an opened tube for evaporation of acetone.

2.4.8.3.4.2. Exhaustion of the antibodies

1.8 mL of PBS was added to 50 mg of acetonic powder in a 2 mL tube. The volume of serum necessary for the Western blot was added to the previous mixture. The whole was incubated on a wheel overnight at room temperature. Then, the tube was centrifuged 5 minutes at 11000 *g* and the supernatant was used directly for the Western blot.

2.4.9. Enzymatic activity assays

2.4.9.1. 4-CL1 activity

The assays were carried out in microplates in a final volume of 250 µL, and the absorbance at 333 nm was measured using a microplate reader (Synergy HT, Biotek).

The assay mixture contained 4 µL of fraction to analyze, DTT 1 mM final, ATP 2.5 mM, $MgCl_2$ 2.5 mM, coumarate 0.2 mM and Potassium Phosphate buffer pH 7.4 50 mM. Reactions were started by addition of Coenzyme A reduced form (CoASH) 0.4 mM. The reactions occurred at 25°C. Kinetics of 3 hours were recorded to measure the parameters of the kinetics. Calculations were done with the product formation curve. The concentration of the product was determined by using the extinction coefficient of *p*-coumaroyl CoA: ε333 nm = 21 $mM^{-1}.cm^{-1}$ (Stökigt & Zenk, 1975)

2.4.9.2. HCT activity

2.4.9.2.1. Substrate preparation

The substrate of HCT, the *p*-coumaroyl CoA was prepared at the laboratory by enzymatic synthesis starting from the *p*-coumaric acid commercially available (Sigma).

p-coumaroyl CoA was prepared by a one-step enzymatic synthesis, using purified recombinant *p* -coumaroyl CoA ligase, as described in the P450s production paragraph (Cf. § Substrate preparation for CYP98A3 p.93).

When the spectrophotometric measure indicated a complete substrate conversion into CoA ester, the reaction was stopped by addition of 4N HCl (V/20) and the products were extracted twice with one volume of ethyl acetate, and then dried under argon flow before resuspension in buffer (Potassium Phosphate 50 mM pH 7,4). The final concentration of the product was determined by spectrophotometry using the extinction coefficient of *p*-coumaroyl CoA: $\varepsilon_{333\ nm} = 21\ mM^{-1}.cm^{-1}$.

2.4.9.2.2. HCT activity assay

100 µL of each enzymatic fraction to analyze were incubated with DTT (1 mM final), shikimate (5 mM final), *p*-coumaroyl CoA freshly prepared (approximately 5 mM final) and Potassium Phosphate buffer pH 7.4 100 mM in a total volume of 1 mL. The reactions were started by addition of the coumaroyl CoA. The reactions occurred at 30°C during 30 minutes. The reactions were stopped by addition of 4N HCl (V/20).

2.4.9.2.3. Liquid/liquid extraction

The products were extracted twice with one volume of ethyl acetate, dried under argon flow before resuspension in 150 µL of injection buffer (Acetonitrile 10% (v/v), formic acid 0.2% (v/v)) and analyzed on HPLC.

2.4.9.2.4. HPLC analysis

Metabolites in injection buffer (injection of 100 µL) were separated by HPLC using a Waters HPLC system coupled to a photodiode array detector.

Chromatographic column used and gradient were described in the P450s production paragraph (Cf. § High Performance Liquid Chromatography analysis p.94)

Determinations of the activities were done by quantification of the synthesized products (values of the corresponding peak area compared to a standard curve).

3. Production of *Arabidopsis thaliana* crude extracts

While validation of partner protein interaction using purified protein was a necessary validation step, the ultimate goal of the project was to trap new interactants of a metabolon on Nanodisc-anchored proteins. To this end, various crude extracts of *Arabidopsis thaliana* were produced.

Extracts were prepared from *Arabidopsis thaliana* ecotypes Col-0 and Ws, stems, roots and flowers. To improve protein quality and concentration in the extracts, different procedures were tested:

- grinding with an ultra-turrax T25 (2x1 min, 6500 rpm) in lysis buffer (4 $mL.g^{-1}$ of tissues),
- grinding in liquid nitrogen with mortar and pestle with or without lysis buffer (1 $mL.\ g^{-1}$ of tissues).

The extracts were more or less concentrated depending on the procedure: protein concentrations ranged from 0.3 to 3.5 $mg.mL^{-1}$. Protein concentration was expected to be critical to mimic *in vivo* conditions since cytosolic protein concentration is extremely high and expected to approach 200 $mg.mL^{-1}$ (Ellis, 2001).

Materials and Methods

1. Plant material: *Arabidopsis thaliana*

Arabidopsis is a model plant because of its small size, its short life cycle, its sequenced genome. In addition, a whole package of tools (mutants, transformation methods…) are available for this plant.

The small dicotyledonous plant *Arabidopsis* belongs to the Brassicaceae family.
Two ecotypes were used during this thesis: Columbia (Col-0) and Wassilevskija (Ws).

The plants in soil were grown in a greenhouse. The temperature of the room was set at 19°C at light with 70% of hygrometry, and at 17°C at night with 85% of hygrometry. The photoperiod was 12 hours with approximately 3000 lux of light during the day. The soil-based compost was treated by an anti-fungus gnat product (Trigard: 0.4 $g.L^{-1}$) before sowing or repicking the plants.

The plants grown in *in vitro* conditions were placed in special culture rooms. The photoperiod was 16 h of light and 8 h of dark, and the temperature was fixed at 23°C. The artificial culture medium was a Murashige & Skoog medium supplemented with 6% (w/v) pastagar.

Arabidopsis thaliana seeds were sterilized before sowing in *in vitro* conditions. They were washed in a sterilization solution (25% (v/v) of sodium hypochlorite, 0.1% (w/v) of sodium dodecyl sulfate in sterile MilliQ water). The sterilization was done during 10 minutes. Then, this solution was discarded and replaced by a solution of 70% ethanol. The seeds were sterilized with this solution during 5 minutes. Finally, ethanol was discarded and the seeds were washed five times in sterile MilliQ water, to remove traces of ethanol, SDS and sodium hypochlorite.

2. Crude extracts preparations

These extractions have been performed with the following Lysis Buffer: EDTA 1 mM, Sucrose 250 mM, Phenylmethanesulfonylfluoride (PMSF) 1 mM, *β*-mercapthoethanol 10 mM, Ascorbate 40 mM in Sodium phosphate buffer 0.1 M pH7,4. This buffer was supplemented by one tablet of antiprotease EDTA-free cocktail (Roche) for 100 mL of buffer.

Roots extracts were prepared from *in vitro* plants, and the other extracts from plants cultivated in pots.

The procedures to prepare crude extract were as follow, *Arabidopsis* tissues were cut, then ground by various procedures previously cited (in the results paragraph). The mixture was centrifuged at 10 000 *g* for 20 minutes at 4°C, and the supernatant was filtered through miracloth (25 µm). The crude extract was then stored at -30°C.

4. Preparation of tools for laser scanning confocal microscopy studies

In vivo detection of protein-protein interactions required the construction of plasmids encoding fusions of candidate enzymes with different fluorescent proteins. Analyses of protein subcellular localization and interactions were carried out using different methods of laser scanning confocal microscopy. *Agrobacterium*-mediated transformation of the constructs into *N. benthamiana* leaves for transient expression was chosen for its speed and versatility.

4.1. Preparation of the constructs

The fluorescent reporter proteins eGFP and mRFP were selected to visualize sub-cellular localization, co-localization and interactions of the different proteins of interest. Thus, various constructs were prepared allowing the expression of fusions with CYP73A5 and CYP98A3, but also to the soluble proteins 4CL-1 and HCT. The P450s are ER membrane proteins, with their N-terminus membrane signal peptide. The latter anchors P450s into the ER membrane. Fluorescent proteins were thus only fused to the C-terminal end of the P450s. The soluble proteins constructs were prepared by fusing fluorescent reporter both at their N- and C-terminal ends.

Different confocal microscopy techniques were used. Protein co-localization was determined with a "regular" Laser Scanning Confocal Microscope, while monitoring interactions required Fluorescence Lifetime Imaging Microscopy (FLIM).

FLIM is based on the Fluorescence/Föster Resonance Energy Transfer (FRET) principle. FRET is a natural process occurring in the antenna complex and photochemical reaction center of the plastidial photosystems (Melkozernov, 2006). Briefly, two compatible fluorescent fluorophores can transfer non-radiative energy in a unidirectional way, from a donor to an acceptor (Figure 40). With proteins as fluorophores, two couples are generally used: eCFP/eYFP or eGFP/mRFP. The last couple permits simpler reconvolution calculations to determine lifetime of the donor fluorescence (the fluorescence decrease is mono-exponential), so this couple was chosen for this project. The transfer occurs when the two fluorophores are in close vicinity, between 1 and 8 nm for the eGFP/mRFP couple. This transfer is also dependent on the three-dimensional organization of the two fluorophores (dipole-dipole functioning). In the case of the eGFP/mRFP couple, the transfer occurs from an excited eGFP (donor) to the mRFP (acceptor) (Figure 40).

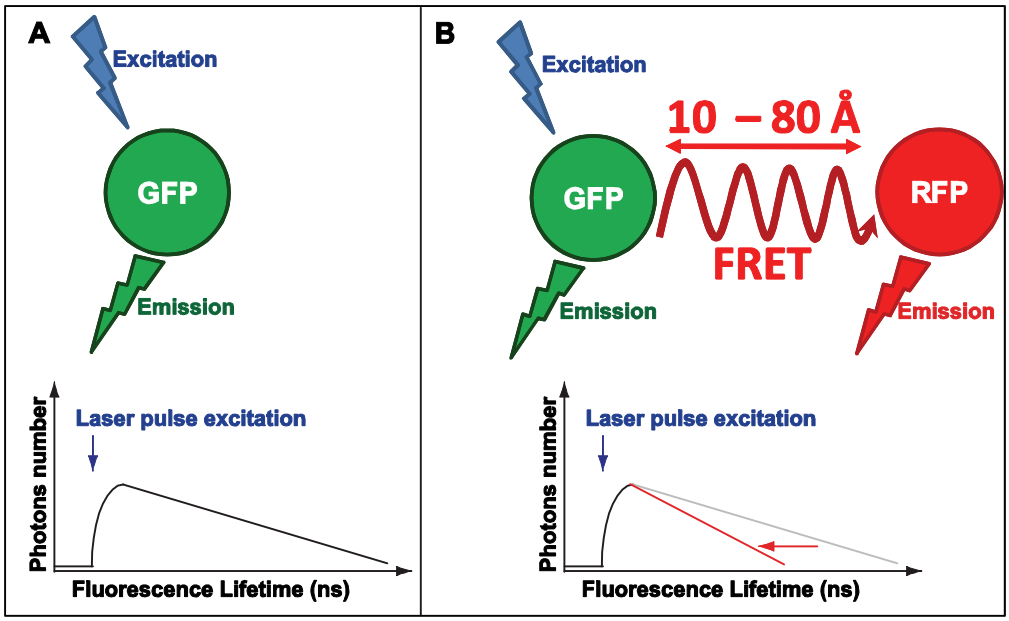


Figure 40. FLIM/FRET principles.
A. eGFP alone or with mRFP located too far away for energy transfer.
B. eGFP close enough to mRFP for energy transfer.
eGFP: energy donor, mRFP: energy acceptor. Abbreviation: ns: nanosecond.

With FLIM, interaction of proteins located in close proximity, so assumed to be interacting, is inferred from the fluorescence lifetime decrease of the donor in presence of the acceptor (Figure 40 B). The nearer is the acceptor, the higher is the transfer of energy, and thus, the higher is the decrease of fluorescence lifetime. But this transfer of energy also depends on the spatial orientation of the donor and acceptor.

For proper folding of the two proteins in the fusion constructs, a linker is usually required that also provides flexibility between the fused proteins. The linker GTGRH was used in this project. The choice of the size and sequence of this linker is essentially arbitrary and empirical. Many studies point to the importance of the linker to ensure good shape and viability of the fusion constructs (Prescott *et al.*, 1999), but no rules could be drawn because the effects largely depend on the studied proteins. Glycine residues have the smallest side chain of all amino acids. Usually this amino acid is repeated in the linker to confer more flexibility to the peptide chain.

All the constructs were prepared by USER™ (Uracil-specific excision reagent) cloning of the full-length sequences according to Geu-Flores *et al.*, (2007) (Flow chart of the USER ™ cloning principle is in appendix: Overview of the USER cloning technique p.318). The constructs were inserted downstream a *CaMV-35S* promoter in the T-DNA of a vector suitable for plant agro-transformation. The list of the constructs is in Figure 41.

Constructs	Vectors
CaMV-35S::CYP73A5::eGFP	pCAMBIA2300U
CaMV-35S::CYP73A5::mRFP	pCAMBIA3300U
CaMV-35S::CYP73A5	pCAMBIA3300U
CaMV-35S::CYP98A3::eGFP	pCAMBIA2300U
CaMV-35S::CYP98A3::mRFP	pCAMBIA3300U
CaMV-35S::CYP98A3	pCAMBIA3300U
CaMV-35S::4CL-1::eGFP	pCAMBIA2300U
CaMV-35S::4CL-1::mRFP	pCAMBIA3300U
CaMV-35S::eGFP::4CL-1	pCAMBIA2300U
CaMV-35S::mRFP::4CL-1	pCAMBIA3300U
CaMV-35S::4CL-1	pCAMBIA2300U
CaMV-35S::HCT::eGFP	pCAMBIA2300U
CaMV-35S::HCT::mRFP	pCAMBIA3300U
CaMV-35S::eGFP::HCT	pCAMBIA2300U
CaMV-35S::mRFP::HCT	pCAMBIA3300U
CaMV-35S::HCT	pCAMBIA2300U
CaMV-35S::eGFP	pCAMBIA2300U
CaMV-35S::mRFP	pCAMBIA3300U

Figure 41. Constructs prepared for sub-cellular localization, re-localization, co-localization and FLIM analysis.

4.2. Fusion proteins are functional

Testing protein-protein interactions is not relevant if fusion to fluorescent protein(s) destabilizes the protein investigated. The first step of this work was therefore to test if the fusion constructs were active when expressed in the plant.

First of all, correct expression of the fused fluorescent proteins was tested, and immediately optimization steps appeared necessary to pursue the project. Images shown below were collected after optimization of the protocol of plant transient transformation (Figure 42).

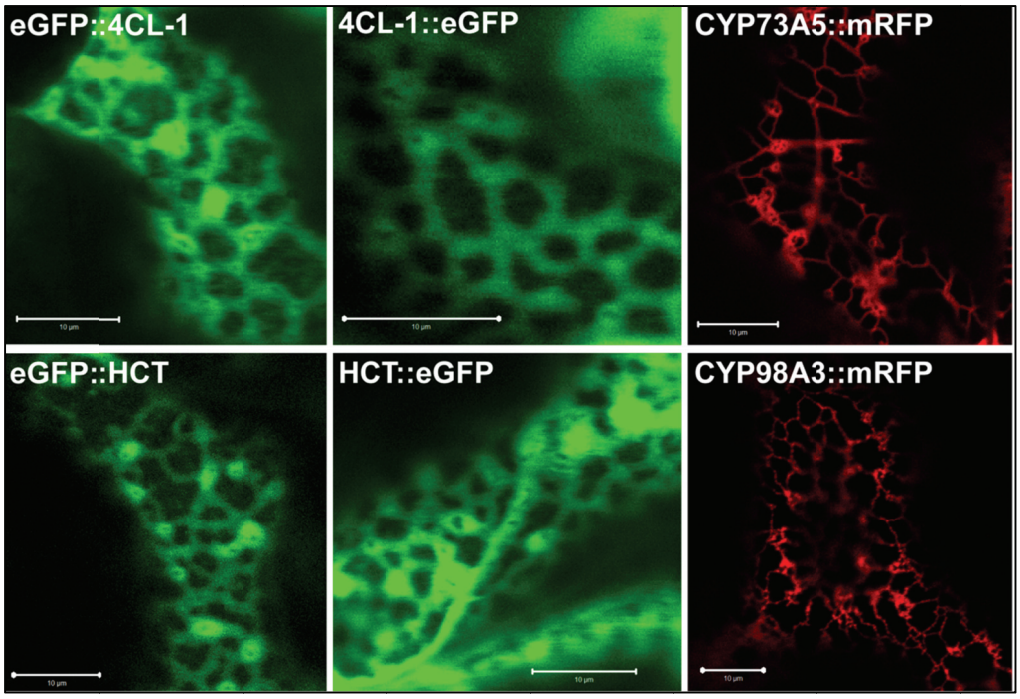


Figure 42. Validation of the expression and fluorescence of the different fluorescent constructs.
Fluorescence detected 5 days after infiltration of *N. benthamiana* leaves with Agrobacteria LBA4404 hypervirulent carrying the vector of interest. Latrunculin B was used (20 µM) to stop movement of the ER. Scale bar: 10 µm.

Expression of fluorescent protein was sufficient and was observed with all the fusion constructs by confocal microscopy (Figure 42). This proves 1) the expression of the fusion proteins in the leaf epidermis, 2) the correct folding of the fluorescent proteins expressed in fusion with the different proteins investigated. It however does not prove correct folding of: CYP73A5, CYP98A3, 4CL-1 and HCT.

To check if the latters were expressed in functional conformation, enzyme activity was measured in the microsomes or soluble fractions isolated from leaves infiltrated with the different constructs and compared to a control expressing only eGFP. Activity of endogenous C4H, C3'H and 4CL, HCT, was detected in eGFP control microsomes and soluble fractions, respectively (controls in Figure 43 and 44). A much higher activity was detected upon expression of fusion proteins, in a ratio around 10 to 25 in all cases, which indicates effective over-expression of functional proteins (Figure 43; Figure 44).

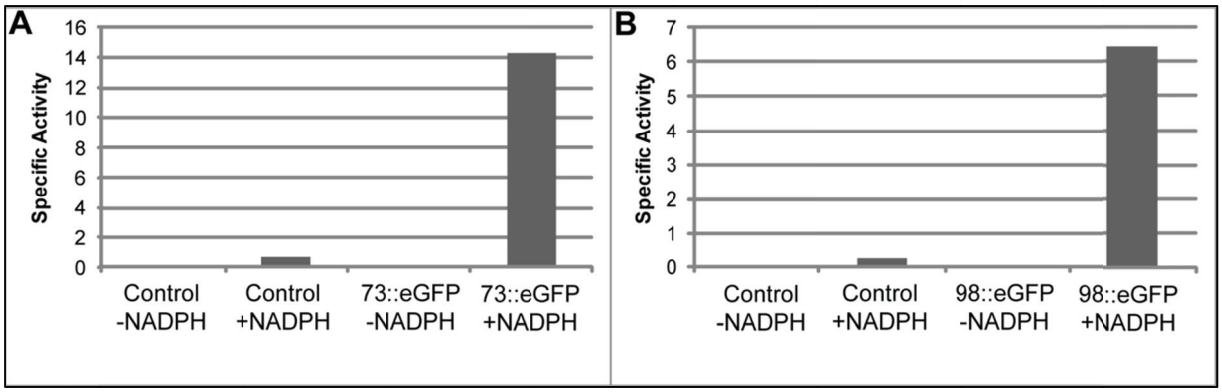


Figure 43. Catalytic activity of the P450s fluorescent protein in microsomal membranes isolated from transfected *N. benthamiana* leaves.

A. Cinnamate 4-hydroxylase activity in leaves transformed with a CYP73A5::eGFP fusion and eGFP control.

B. *p*-coumaroyl shikimate 3'-hydroxylase activity in leaves transformed with a CYP98A3::eGFP fusion and eGFP control.

Controls are *N. benthamiana* plants expressing eGFP. Specific activities are expressed in µmol of product.mg^{-1} protein.min^{-1}.

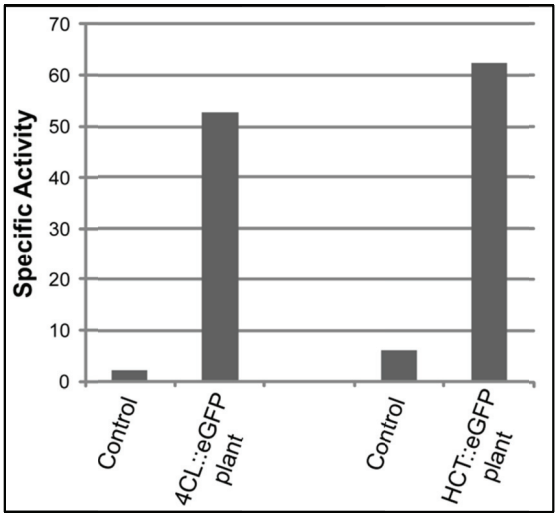


Figure 44. Catalytic activity of the soluble proteins fused to eGFP detected in the crude extract isolated from transfected *N. benthamiana* leaves.

Controls are *N. benthamiana* plants expressing eGFP. Specific activities are expressed in µmol of product.mg^{-1} protein.min^{-1}.

The fusion constructs were thus properly expressed as active enzymes and suitable for *in vivo* interaction assays.

4.3. Optimization of the expression system and of confocal observations

Only scattered cell transformation and low expression rates were observed in our first experiments. Optimization was therefore required. Initially, transfection of epidermal cells was achieved using the *Agrobacterium* LBA4404 strain and *N. tabacum xanthi* for transient expression. With this system, in the best cases, only 50% of the cells were transformed (results not shown) and fluorescence intensity varied markedly between neighboring cells.

Thus, other *Agrobacterium* strains (GV3101 and LBA4404 hypervirulent strains) and expression in *N. benthamiana* epidermal cells were tested. Whereas reproducibility and expression levels were poor with GV3101 and LBA4404 in *N. tabacum* and in *N. benthamiana* (data not shown), expression was higher in terms of cell numbers displaying fluorescence with *Agrobacterium* LBA4404 hypervirulent strain infiltrated into *N. benthamiana*. This system of expression was thus selected for the next steps of the project.

Almost all the epidermal cells expressed the constructs with a strong fluorescence intensity at four days post-infiltration (Figure 45 A). Furthermore, expression and fluorescence were stable until six days post-infiltration. After six days, aggregates of fusion constructs were accumulated all over the cell (Figure 45 B). In consequence, images for co-localization and *in vivo* tests of protein-protein interactions were taken five days after infiltration.

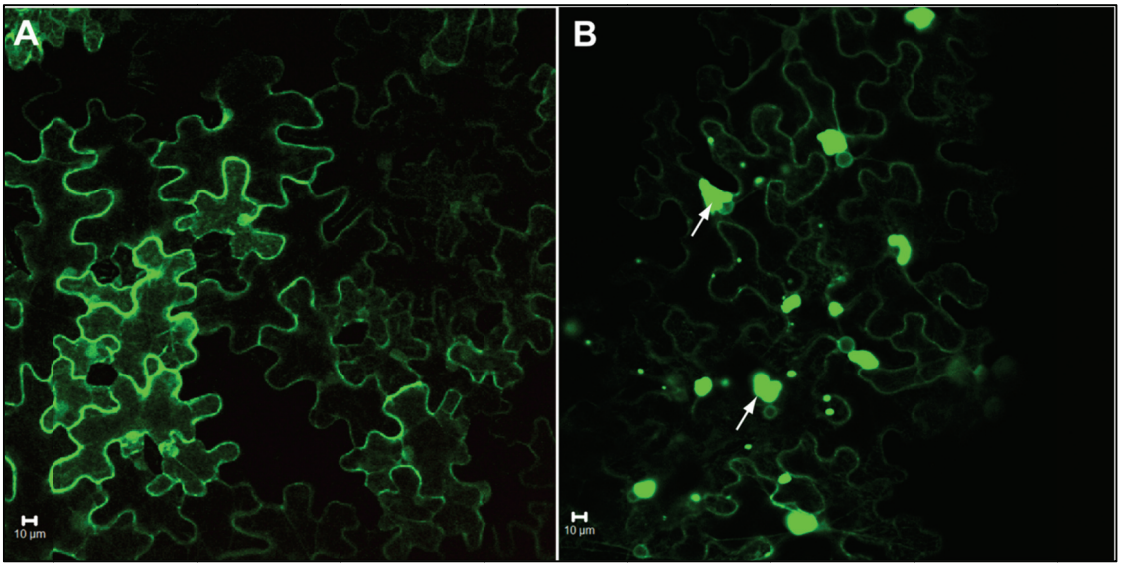


Figure 45. Protein expression in *N. benthamiana* leaf epidermal cells transformed with the *Agrobacterium* LBA4404 hypervirulent strain.
A. Expression of 4CL-1::eGFP. Image collected 4 days post-infiltration.
B. Expression of 73A5::eGFP. Image collected 6 days post- infiltration.
Arrows: aggregates. Images were obtained using a 20x objective. Scale bar: 10 µm.

Another improvement was tested, it consisted in co-expressing the various constructs with the P19 protein. This protein is a virus-encoded suppressor of gene silencing from tomato bushy stunt virus (TBSV). P19 protein prevents the onset of Post-Transcriptional Gene Silencing (PTGS) in the infiltrated tissues and usually allows high level of transient expression (Voinnet *et al.*, 2003). While no obvious improvement was noticed during the first trials, the co-expression of the different constructs was always carried out with the P19 protein.

After optimization, the expression and co-expression of the different constructs were high and highly reproducible.

4.4. Subcellular localization of potential interactants

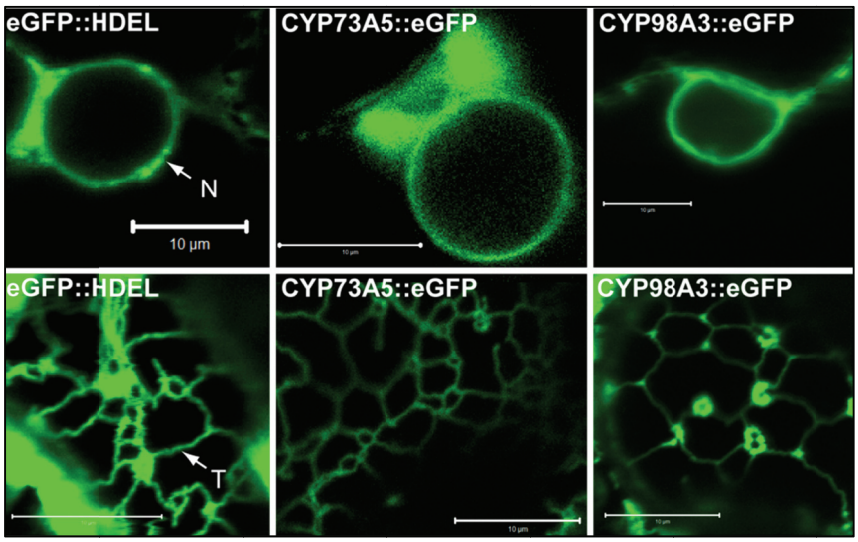


Figure 46. Sub-cellular localization of CYP73A5 and CYP98A3.
Latrunculin B (20 µM) was used for image acquisition.
N: membrane of nucleus. T: ER tubule. Scale bar: 10 µm.

Sub-cellular localization of the different enzymes first had to be confirmed. Images collected with the CYP73A5 and CYP98A3 fusion constructs were compared to a well-characterized ER marker, eGFP with a C-terminal HDEL. The HDEL signal leads to retention of the eGFP protein in the ER. A clear ER localization for both CYP73A5 and CYP98A3 (Figure 46) fused to mRFP or eGFP was observed.

But the first data collections also highlighted the extremely dynamic properties of the plant ER, and the resulting fast movement of both P450s with the ER membranes. Protein movement was a hindrance for further analysis of co-localization and FLIM. The consequences of movement during image acquisition will be explained in the next paragraph (Cf. Confocal microscopy image analysis p.129-136). The movement of the ER tubules is essentially driven by the actin/myosin system (Griffing, 2010). Actin polymerization is involved in tubule growth and movement. It is inhibited by latrunculin B. This molecule is a toxin purified from *Latrunculia magnifica* (Red sea sponge). Latrunculin B inhibits actin polymerization and disrupts microfilament organization, as well as microfilament-mediated processes. As a result, this toxin stabilizes plant ER (Griffing, 2010). Latrunculin B was thus used in all the subsequent experiments.

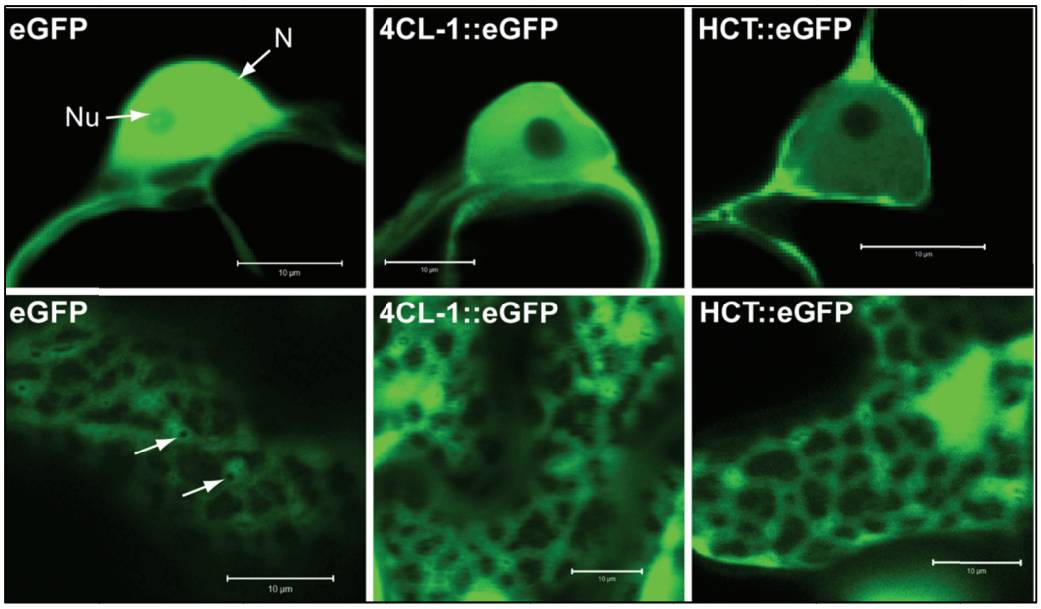


Figure 47. Sub-cellular localization of 4CL-1 and HCT.
Latrunculin B (20 µM) was added before image acquisition. Abbreviations: N. Nucleus. Nu. Nucleolus. Arrows: organelles in negative. Scale bar: 10 µm.

Free eGFP was used as localization control for the soluble proteins. Both HCT and 4CL-1, as free eGFP, had a localization pattern of soluble protein (Figure 47). They were present in the cytoplasm with a reticular pattern when cytoplasm surrounded ER tubules. They were also found in the nucleus. Organelles were visible in negative (arrows in Figure 47), due to cytoplasm encircling the organites. Similar images were obtained with the mRFP constructs and for fluorescent protein fusions at the N-terminal of the soluble proteins.

As expected, CYP73A5 and CYP98A3 were found associated with the ER membranes, while 4CL-1 and HCT were found as cytosolic soluble proteins in *N. benthamiana* epidermal cells. Fusion with fluorescent proteins thus did not interfere with sub-cellular localization of the proteins investigated.

4.5. Confocal microscopy image analysis

Once the sub-cellular localizations confirmed, co-infiltration of the various constructs was undertaken for co-localization analysis. Visualization of proteins residing on intracellular structures by fluorescence microscopy has become a routine approach in cell biology (Bolte & Cordelières, 2006). This technique is used to assess co-localization of proteins of interest with well-characterized markers. When the term co-localization is used in optical microscopy, the objects that are compared are usually pixels or voxels in digital images. A pixel is the smallest element in a two-dimensional image and a voxel is its three-dimensional counterpart. When we say that two molecules are co-localized, we are, in effect, saying "that the voxel illuminated by a fluorophore attached to molecule A is the same voxel that is illuminated by another fluorophore attached to molecule B" (Bolte & Cordelières, 2006). Thus the size of the voxels is important for interpreting the experiments (Bolte & Cordelières, 2006). If the size of the voxel is too large, several elements could be detected as co-localized even if in reality they are not.

In cellular biology, different approaches are used to describe and quantify co-localization. From a rather simplistic qualitative evaluation of overlapping pixels, to a bulk of complex solutions and calculations, most of them are based on global statistic analysis of pixel intensity distributions (Bolte & Cordelières, 2006).

When visualizing co-localization, the most simple method is to present results as an overlay of the different channels, each image being pseudo-coloured using an appropriate colour look-up table. For example, it is commonly accepted that the dual-channel look-up table for green and red give rise to yellow hotspots where the two molecules of interest are present in the same pixels (Bolte & Cordelières, 2006).

This method was used to obtain first images of co-localization with the two couples: 4CL1::eGFP / CYP73A5::mRFP and HCT::eGFP / CYP98A3::mRFP (Figure 48). First glances on the images led to the conclusion that soluble proteins and P450s were not perfectly co-localized, and the GFP signal surrounded the RFP signal along the ER tubules.

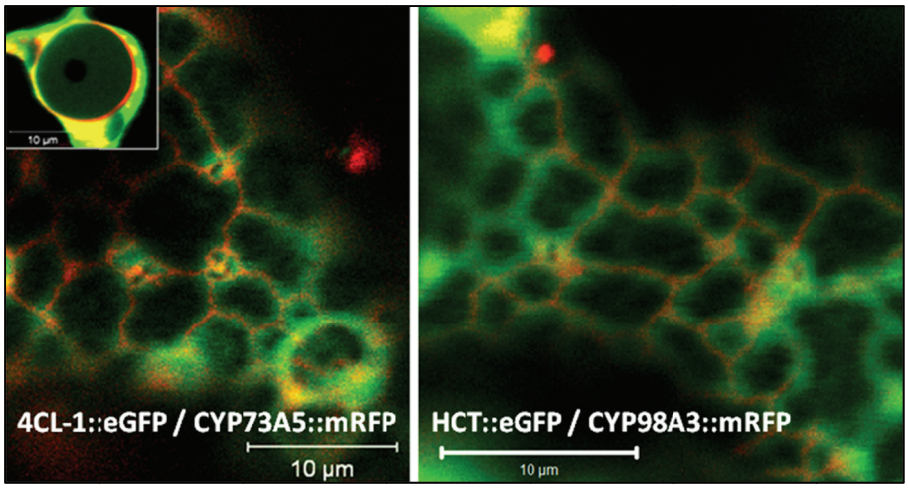


Figure 48. Fluorescence recorded upon co-expression of soluble proteins and partner P450s.
Images of the two channels used for acquisition of eGFP and mRFP fluorescences were merged. Latrunculin B (20 µM) was added before image acquisition. Scale bar: 10µm.

Co-localization analysis was carried out, and, as anticipated from visual inspection, graphical representation of the two signals co-localization was not really conclusive (Figure 49 A). Graphical diagrams obtained from image analyses (Figure 49 A) were very far from the theoretical diagram expected in case of good co-localization (Figure 49 B). Furthemore, comparison between repeated re-localization assays was extremely difficult.

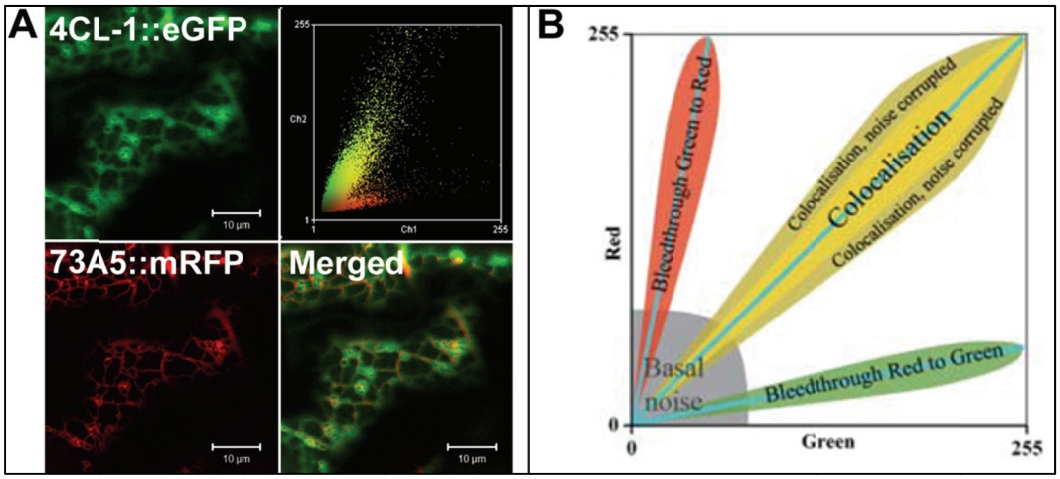


Figure 49. Example of co-localization analysis.
A. Co-localization of 4CL-1::eGFP and CYP73A5::mRFP. Scale bar: 10 µm.
B. Theoretical diagram obtained for co-localization analysis (from Bolte & Cordelières, 2006).

Many others limitations hinder calculation of co-localization and lead to difficulties for evaluation of re-localization. These limitations are discussed below.

The first is the edge effect causing false positive co-localization (Figure 50). In case of a too large optical slice, fluorescent proteins from two parts of a cellular ultrastructure (as an ER tubule) could create artificial co-localization in the projection image of the voxels (Figures 50 B, D). In the best cases, artificial co-localization is only observed on the ultrastructure edges, perpendicular to the optical slice (Figures 50 A, C).

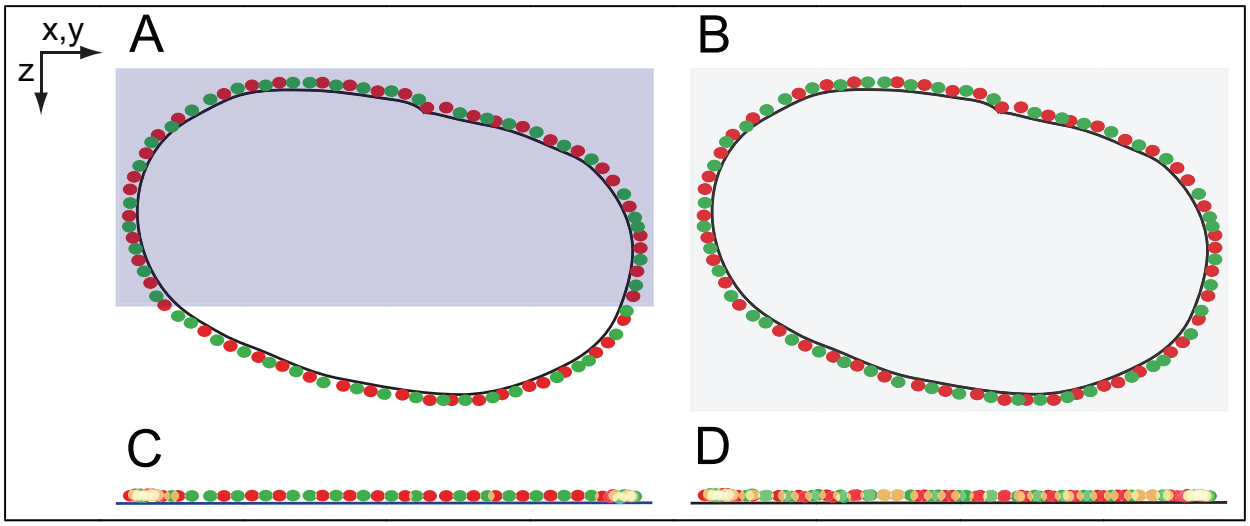


Figure 50. Examples of false positive co-localization.
A. Hypothetical transversal section of the ER tubule with the optical slice shown in blue.
B. Hypothetical transversal section of the ER tubule with the optical slice shown in grey.
C. Projection of fluorescence voxels from image (A).
D. Projection of fluorescence voxels from image (B).
Red spots and green spots represent respectively mRFP and eGFP proteins localized on the ER membrane. Yellow to light yellow areas correspond to co-localization of GFP and RFP.

Artefacts due to edge effects can obviously occur when studying ER tubules. With a regular pinhole used during image acquisition, the best theorical lateral resolution of the confocal microscope is 0.152 µm (resolution = 0.4 x laser wavelength / Numerical Aperture, with Numerical Aperture = 1.32 for the 63x oil immersion objective, and wavelength around 500 nm). The best axial resolution is 0.402 µm (Bolte & Cordelières, 2006). Widths of the ER tubules measured with confocal microscopes are about 0.2 µm (Griffing, 2010), but in fact they seem to be thinner. Tubules width seems to be less than 0.1 µm according to super-resolution STED (stimulated emission depletion) microscopes (Figure 51).

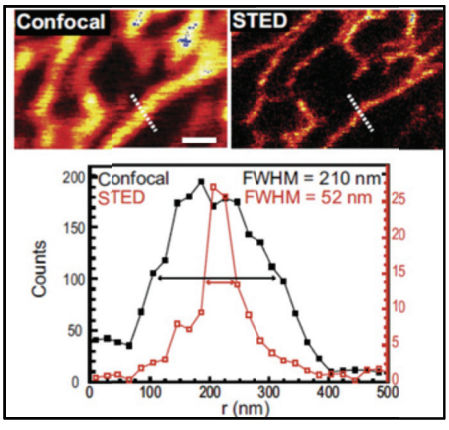


Figure 51. ER tubule width measured by confocal and STED microscopy (Griffing, 2010).

These ultrastructures appear larger than the lateral limit of confocal resolution, but in fact they are thinner. In addition, they are much smaller than the axial resolution limit. In other words, in one image pixel, fluorescence of a voxel about 0.402 µm x 0.152 µm is acquired, and thus several proteins to different locations will be considered as co-localized, even if not.

Another important limitation during co-localization studies is the movement of the cellular components. As mentioned previously, ER moves in the cell, and this movement is driven by the actin-myosin system. The actin-driven movement can be counteracted by latrunculin B, but slight motions are still present. In Figure 52, a global sliding between the two tracks was observed. These motions can be due to vibration of the ER tubules, sliding of the cell, movement driven by myosin… So, motion of ER is not totally annihilated.

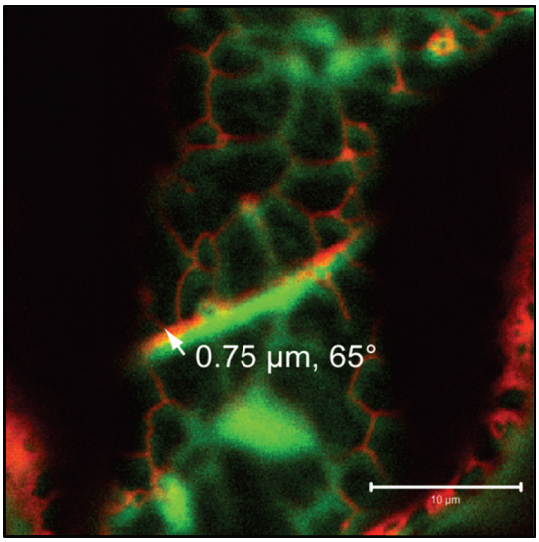


Figure 52. Movement of cellular structures during data acquisition.

Fluorescence emitted by 4CL-1::eGFP and 73A5::mRFP was acquired sequentially by mutli-track frame-by-frame method. One frame is acquired for one fluorophore; the other frame is acquired sequentially. Each channel was acquired during 29s.

Chemical fixation of the tissue (in paraplast for example) would destroy fluorescence and cannot be used. The best manner to overcome movements is to increase scan speed. But by decreasing the pixel scan time, another limitation for co-localization can appear. It is the noise of the images. Noise is another major problem in digital imaging, and depends on the optical system and on the acquisition parameters. Noise can be limited by scanning several times the same voxel and averaging the results (Bolte & Cordelières, 2006). But this increases scan time.

One more important limitation is the saturation of images. Saturated pixels may not be quantified properly because information of the most intense grey level values are lost. It is difficult to judge by eye if an image is saturated, as the human eyes are not sensitive enough. Our eyes can however distinguish between hundreds of colours and, with the image-acquisition software, it is possible to color saturated pixels (with a look-up table palette) and to adjust the dynamics of grey values with the detector (Bolte & Cordelières, 2006). The presence of yellow spots in overlay image is highly dependent on the relative signal intensity collected in both channels; the overlay image will provide a reliable representation of co-localization only in the precise case where both images exhibit similar grey level dynamics. This is rarely the case when imaging two fluorochromes with different signal. Furthermore, these overlay methods are not appropriate for quantification purposes because they may result in misinterpretation of relative proportions of molecules.

Finally, imaging could be impaired by background from either natural fluorescence of the sample or generated when preparing the sample. In most cases, nothing can be done after image acquisition except if an uniform background is observed. Plant cells contain chlorophyll. This molecule is excited with commonly used laser and emited in the red channel. The artifact was limited here by the use of a band pass filter during acquisition of mRFP signal.

To conclude, when visualizing co-localization, a simple method is to present results as an overlay of the different channels. However, limitations are important. To overcome these image analysis problems is crucial. There are two basic ways to evaluate co-localization events, a global statistic approach of intensity correlation coefficient-based analyses and an object-based approach (Bolte & Cordelières, 2006). Many different programs or plugins for ImageJ program were designed to determine co-localization (JACoP, CoLocalizer Pro…). These methods are heavy for studies aiming at re-localization analysis of soluble proteins affected by different treatments.

So, the next paragraphs refer to the development and tests of an easier method which appears in theory less affected by the width and thickness of the optical slice and by the various limitations impacting co-localization analysis.

4.5.1. Theory about re-localization measurements

In case of soluble protein recruitment and re-localization resulting from P450 co-expression, rarefaction of soluble protein in the more distant area surrounding the ER tubules should be detected. As a result, a fluorescence constriction near ER tubules should be observed (Figures 53 A, B).

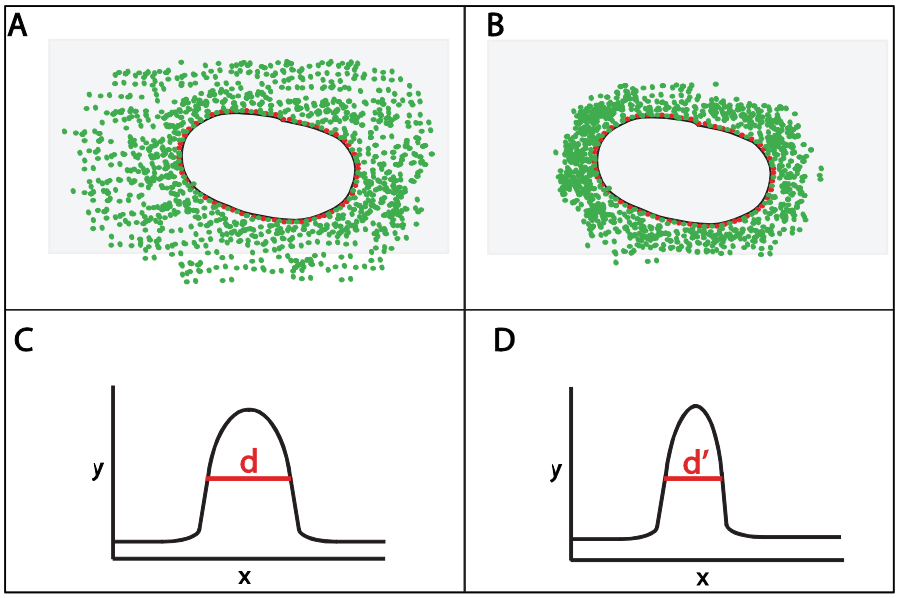


Figure 53. Schematic representation of fluorescence distribution upon protein re-localization near the ER.
A. Schematic representation of a transversal cross-section of an ER tubule and hypothetical distribution of GFP spots of a soluble fusion.
B. Hypothetical redistribution of GFP spots of a soluble fusion if re-localization of this fusion occurs.
C. Hypothetical Gaussian curve corresponding to (A)
D. Hypothetical Gaussian curve corresponding to (B).
The optical slice is represented by the grey shape in A and B. Green spots represent GFP fused to a soluble protein and red spots represent RFP fused to a ER membrane associated protein. d and d' represent Full Width at Half Maximum of the Gaussian curve. X, distance and Y, number of counted pixels. Grey areas in (A) and (B) represent the optical slice thickness.

Distribution of the fluorescent pixels (which represent voxels in the sample) can be modeled by a Gaussian equation (Figures 53 C, D) centered to ER tubule. Once the Gaussian equation is calculated, the Full Width at Half Maximum (FWHM) can be deduced. This value can be compared between the different re-localization trials providing quantification of the fluorescence constriction around ER tubules. The latter reflects re-localization of the soluble proteins near ER.

4.5.2. Method details about re-localization measurements

The samples were prepared and data recorded as follows:

- The pieces of infiltrated *N. benthamiana* leaves were mounted between slide and coverglass in presence of Latrunculin B (20 µM) in water. Other molecules were added when re-localization upon substrate addition was investigated. The sample was then vacuum infiltrated for Latrunculin uptake and to obtain turgescent cells, thus cells in similar conditions, with ER pressed on plasma membrane.
- Then cells were selected randomly for data collection. A fast check indicates if cell and ER are in good shape: clear ER, not too many aggregates of fluorescent proteins in the cell, RE movement almost stopped…
- In case of problem, another cell was randomly taken and checked.

- The pixel size was adjusted on the detector to the minimal justified pixel size, to avoid over or under sampling.
- The time scan was adjusted.
- The pinhole was checked, to have the same parameters between conditions tested.
- The detector gain was fitted (thanks to look-up table panel) to avoid saturation of the image, but also to display the brightest fluorescence and to have similar conditions between constructs and between samples.
- Finally the sample was scanned.
- Files were saved in the .lsm format. This file format can be directly opened by imageJ.

Then, image analyses were carried out with a macro for ImageJ designed by Jérôme Mutterer (Microscopy and Imaging Platform, IBMP, Strasbourg) (Figure 54).

```
macro "fit [F5]" {
roiManager("add");
y=getProfile();
x= newArray(y.length);
getLine(x1, y1, x2, y2, lineWidth);
getPixelSize(unit, pixelWidth, pixelHeight);
d=sqrt(pow(x2-x1,2)+pow(y2-y1,2))*pixelWidth;
for (i=0;i<y.length;i++) x[i]=i*d/y.length;
//Plot.create("Title", "distance "+unit, "intensity", x, y);
Fit.doFit("gaussian",x,y);
Fit.plot();
setResult("a",nResults,Fit.p(0));
setResult("b",nResults-1,Fit.p(1));
setResult("c",nResults-1,Fit.p(2));
setResult("d",nResults-1,Fit.p(3));
setResult("r2",nResults-1,Fit.rSquared);
updateResults;
}
```

Figure 54. Textfile of the Macro for protein re-localization image analysis with ImageJ.

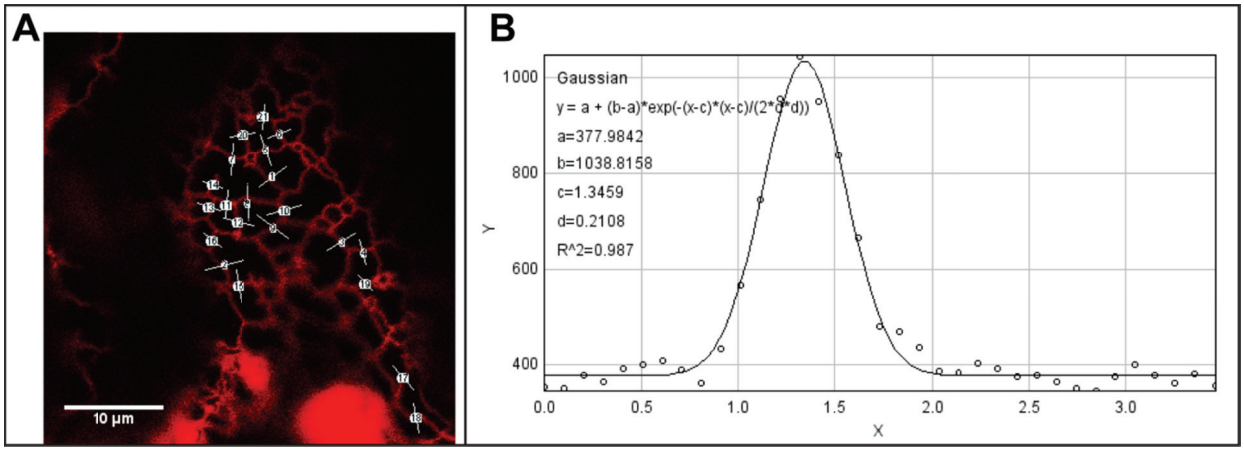


Figure 55. Example of image analysis with CYP73A5::mRFP.
A. Image showing optical slices evaluated for the distribution of fluorescent pixels.
B. Gaussian curve calculated from one measure. X, distance and Y, number of counted pixels.

Fluorescent pixel repartition was determined along lines drawn as perpendicular as possible to the ER tubule. This was a limitation to the method since the angle of this line impacts the pixel repartition determined by the algorithm. At least ten measures were recorded randomly on each image (Figure 55 A). Data from several images from different cells, different plants and different confocal microscopy sessions were pooled for the evaluation of the pixel repartition in one condition.

For each measure, the Gaussian equation representing pixel repartition around the ER tubule was calculated. Only Gaussian curves with an R^2 higher than 0.9 were taken into account, to ensure reliability of the analyses. The "d" value in this Gaussian equation (Figure 55 B) permits calculation of ER-FWHM (ER-Full Width at Half Maximum = $d\ 2\sqrt{(2\ln 2)}$).

The "d" values were collected in an Excel file and analyzed with the plug-in XLtoolbox (http://xltoolbox.sourceforge.net/welcome.html) for statistical analysis. An analysis of variance (ANOVA) with posthoc analysis was carried-out with the Bonferroni-Holm posthoc test to determine consistency of significant difference of the values recorded in different conditions. Assumption for ANOVA test is the equal variance of populations of measures compared. If the variances are different, analysis should be taken with caution. Confidence of the ANOVA results were also determined. A code was used with asterisks from 1 to 4: four asterisks meaning good confidence. During Posthoc tests, a critical value is calculated by the algorithm. If the Probabiliy (P) is over this critical value, the populations of measures are not significantly different. Conversely, if P is under the critical value, the populations of measures are significantly different.

4.5.3. Proof of concept of the method

Reproducibility of the method was tested before bulk analysis of the data. Tests were performed to check sensitivity to image noise and to model the different patterns, i.e. soluble protein pattern (with a flattened Gaussian curve), and ER protein pattern (with sharp Gaussian curve).

4.5.3.1. Tests with P450s, eGFP::HDEL and free eGFP

P450s are ER membrane proteins. The first test conducted to check the validity of the method, was to compare P450s and eGFP::HDEL localization (i.e. distance to ER) and to test validity of this measure by ANOVA.

Proteins	N	Average	SD
Free eGFP	100	1.6	0.5
eGFP::HDEL	100	0.4	0.1
98A3::eGFP	100	0.4	0.1
73A5::eGFP	100	0.4	0.1
Confidence in ANOVA tests	**8,0959E-113**	********	

Posthoc test: Bonferroni-Holm		Critical Value	P	Significant?
Free eGFP	98A3::eGFP	0.01	2.04E-50	**Yes**
Free eGFP	eGFP::HDEL	0.01	1.62E-48	**Yes**
Free eGFP	73A5::eGFP	0.01	7.33E-47	**Yes**
eGFP::HDEL	73A5::eGFP	0.02	0.02	**No**
eGFP::HDEL	98A3::eGFP	0.03	0.04	**No**
98A3::eGFP	73A5::eGFP	0.05	0.86	**No**

Figure 56. Comparison of the localization of ER membrane proteins.
N, Number of measures. SD, standard deviation.

According to the ANOVA test (Figure 56), no significant difference was detected between fluorescence patterns resulting from expression of eGFP::HDEL or from expression of CYP73A5::eGFP and CYP98A3::eGFP. **Thus, as expected, P450s were confirmed as anchored to the ER membranes.**

R^2 values of the Gauss equation calculated for free eGFP were 0.94±0.06 during these preliminary tests. This means that most measures performed with a soluble protein can be used for further analysis. Moreover, **soluble proteins repartition around ER tubules can be modeled by a Gaussian equation.**

4.5.3.2. Localization of CYP98A3 and CYP73A5 under stress conditions

P450 over-expression has been reported to induce cellular stress (Szczesna-Skorupa *et al.*, 2004), in some cases, leading to dramatic Untranslated Protein Response (UPR) resulting in change in genes expression and sometimes cell death. As a result, over-expression of P450s can lead to a dramatic change in ER shape. The project requiring coexpression of several proteins (including two P450s), it was decided to check tubule size stability upon co-expression of several proteins, and addition of cofactors and substrates used in the experiments (cinnamate 5mM, coumarate 5mM, spermidine 5mM, NADPH 5mM, shikimate 5mM) (Figure 57).

Proteins	N	Avg.	SD
73A5::GFP	90	0,4	0,08
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT	90	0.4	0,09
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT + substrates	90	0.4	0.15
73A5::RFP / GFP::4CL1 / 98A3A3 / HCT + substrates	90	0.4	0.13
73A5::RFP / GFP::4CL1 / 98A3A3 / HCT	90	0.4	0.13
Test for equal variance	0.02	equal variance CANNOT be assumed (p <= 0,05)	
Confidence in ANOVA tests	0.04	*	

Posthoc test: Bonferroni-Holm		Critical value	P	Significant?
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT	0.005	0 003	Yes
73A5::RFP / GFP::4CL1 / 98A3A3 / HCT + substrates	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT	0.006	0.021	No
73A5::GFP	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT	0.006	0.024	No
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT + substrates	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT	0.007	0.083	No
73A5::GFP	73A5::RFP / 4CL1::GFP / 98A3A3 / HCT	0.008	0.251	No
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT	73A5::RFP / 4CL1::GFP / 98A3A3 / HCT + substrates	0.010	0.396	No
73A5::GFP	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT + substrates	0.013	0.592	No
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT + substrates	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT + substrates	0.017	0.647	No
73A5::RFP / 4CL1::GFP / 98A3A3 / HCT	73A5::RFP / GFP::4CL1 / 98A3A3 / HCT + substrates	0.025	0.723	No
73A5::GFP	73A5::RFP / 4CL1::GFP / 98A3A3 / HCT + substrates	0.050	0.963	No

Figure 57. Comparison of CYP73A5::mRFP localization under potential stress conditions.
N, Number of measures. Avg., average. SD, standard deviation.

A significant difference in the measures distribution was found by ANOVA between two conditions (CYP73A5::mRFP/4CL1::eGFP/98A3/HCT and CYP73A5::mRFP/eGFP::4CL1/98A3/HCT), but variances were not equal between populations of measures, thus the test must be interpreted with caution. Furthermore confidence in ANOVA analysis was not good, and P value was just under the critical value. For other comparisons, no significant difference was found. Therefore, **ER stability and accuracy of the measure may assume to be correct with these different stress conditions. ANOVA analysis seems to be very sensitive to the variances between populations** and enough measures must be performed for each condition to counteract that.

4.6. Discussion - conclusions

P450s and soluble enzymes were efficiently and reproducibly expressed after optimization. **Intensity of fluorescence was sufficient for further experiments.** Furthermore, **enzymes were functional** as fusion proteins.

As expected, P450s were found associated with the ER membranes. 4CL-1 and HCT behaved as soluble proteins showing however a reticulate pattern.

A novel method to evaluate protein localization and co-localization has been developed. **This method was shown to be robust and effective**. Even if calculated tubule width values were not correct due to limitation of the confocal microscope, **the method permits reliable comparison of fluorescence repartition around the ER tubules. Its most important limitation was the requirement to compare images with same intensity of fluorescence** (without saturation). This can be obtained by adjustment of the detector gain. **This method was not impacted by movement of the cell components** between the acquisition of the two channels (eGFP and mRFP), nor it was limited by the size of the optical slice. **The prerequisite was to use always the same optical slice size between conditions.** A limit of this method, as for other co-localization analyses, was the concentrations of the compared proteins. In this project, soluble proteins were more expressed than membrane proteins. This resulted in an **under-estimation of the re-localization of soluble proteins** since there was not enough P450 to re-localize all the soluble proteins.

Finally, the **ANOVA algorithm seemed to be very sensitive to slight differences between measured populations**. In consequence, this sometimes complicated interpretation of the results in the next paragraphs.

4.7. Materials and Methods

Samples of alive materials (plants) to be analyzed by semi-quantitative PCR or metabolic extraction, were immediately frozen after sampling with liquid nitrogen and stored at -80°C before analyzis.

4.7.1. Biological materials

4.7.1.1. Nicotiana

Nicotiana is a dicotyledonous plant belonging to the family of the Solanaceae.
Two species were used: *Nicotiana benthamiana* for transient expressions and *Nicotiana tabacum* for stable and transient expressions.

For *Nicotiana tabacum*, 2 cultivars were used: Xanthi (line Nc) for the transient expressions, and Xanthi (line SH6) for the stable transformations.

Wild type plants and transformed plants were grown in soil in culture room and subjected to cycles of 16 hours of light (temperature of 24°C) and 8 hours of darkness (temperature of 20°C).

4.7.1.2. Agrobacterium tumefaciens

Agrobacterium tumefaciens is a Gram negative bacterium (family of *Rhizobiaceae*) used to transform plants.

4.7.1.2.1. Used strains

+ LBA4404 strain (ClonTech Laboratories Inc.)
This strain contains a disarmed Ti (Tumor inducing) plasmid, i.e., that it does not have any more its oncogenes. It carries virulence genes (*vir*) necessary for the transfer and the integration of the T-DNA in the cell and the genome of the plant cells. This strain carries in its genome an antibiotic resistance gene (rifampicin).
This strain was used for the transient and stable transformations of tobacco.

+ LBA4404 hypervirulent strain
This strain was kindly provided by Christophe Ritzenthaler (IBMP, Strasbourg).
This strain is a modification of LBA4404 strain, by the insertion of the pMP90 plasmid making it hypervirulent. This strain carries in its genome a resistance gene to rifampicin, and the plasmid of hyper-virulence carries resistance gene to gentamicin.
This strain was used for the transitory transformations of tobacco.

+GV3101 strain
This strain was kindly provided by Christophe Himber (IBMP, Strasbourg).
It also contains a disarmed Ti plasmid. This strain also carries in its genome the resistance gene to rifampicin. This strain was used for the transient transformations of tobacco.

4.7.1.2.2. Culture conditions

The bacteria were cultivated either in LB liquid medium under agitation at 180 rpm or on solid LB medium in Petri plates at 28°C. Adequate antibiotics were added to the media for selection of the strain and of the used plasmid. Bacteria were cultivated at 28°C to have an optimal growth and to remain virulent.

All the stocks of agrobacteria were preserved at -80°C in the culture medium (without antibiotic) containing 30% glycerol.

4.7.1.3. Escherichia coli

Escherichia coli were used for cloning and propagation of the plasmids.

4.7.1.3.1. Used strains

DH5α (Invitrogen) and TOP10 (Invitrogen) strains were used.

4.7.1.3.2. Culture conditions

The bacteria were grown either in LB liquid medium under agitation at 170 rpm, or on LB solid medium in Petri plates. Bacteria cultures were grown at 37°C. To the liquid or solid media of culture the adequate antibiotics were added for the selection of the plasmid of interest.

4.7.2. Plant expression vectors

The two used vectors were suitable for over-expression, sub-cellular localization, co-localization and re-localization studies. Initial pCAMBIA vectors were kindly provided by H. Nour-Eldin (University of Copenhagen, Denmark).

pCAMBIA2300U: This vector has a USER cloning cassette downstream of a *CaMV-35S* promoter, the whole in the T-DNA, which will be integrated in a random way in the genome of the plant cell. In order to allow the selection of the transformed bacteria and transformed plants, this vector carries two genes of resistance. The first, located outside of the T-DNA, allows the selection of the transformed bacteria (*A. tumefaciens*, *E. coli*); the second, located in the T-DNA, allows the selection of the transformed plants. These genes of resistance are both coding for an aminoglycoside phosphotransferase (*NPTII* gene), conferring resistance to kanamycin. The USER cassette was inserted in the pCAMBIA2300 vector by F. Duval from the laboratory.

pCAMBIA3300U: As for the plasmid pCAMBIA2300, this vector has a USER cloning cassette downstream of a *CaMV-35S* promoter, inside a T-DNA. It also carries two markers of seletion for bacteria and plant selection. The first, located outside of the T-DNA, allows the selection of the transformed bacteria (*A. tumefacians*, *E. coli*); the second, located in the T-DNA, allows the selection of the transformed plants. The gene of bacterial resistant is the *nptII* gene. The gene for plant resistance is the *bar* gene, and it confers resistance to phosphinothricin (Basta). The USER cassette was inserted in the pCAMBIA3300 vector by F. Duval from the laboratory.

4.7.3. USER™ Cloning

This technique was used to build the plasmids for plant expression of P450s and soluble proteins with or without fusion of a fluorescent protein.

It is a novel method for direct cloning of PCR fragments in a destination vector. It is a fast and effective cloning, which allows to fuse and to clone simultaneously various products of PCR in various vectors (Geu-Flores *et al.*, 2007). This novel method was developed and marketed in 2003 by New England Biolabs. Then the system was improved by Nour-Eldin *et al.*, (2006) who identified a polymerase with "proof reading" activity compatible with the use of Uracil oligonucleotides during PCR generating the amplicons to clone.

The principal advantage of this technique is that cloning is carried out in only one step.
A flow-chart of the USER cloning is presented in the Appendix: Overview of the USER cloning technique p.318.

4.7.3.1. Preparation of the vectors suitable for USER reaction

Initially, the pCAMBIA2300U or pCAMBIA3300U vectors were linearized. They carry a cloning cassette, which, successively digested by the restriction enzymes *Pac*I and *Nt.Bbv*CI will generate cohesive ends (see Appendix: Overview of the USER cloning technique p.318).

First restriction: 50 Units of PacI enzyme were used to digest 10 µg of plasmid in the NEB4 buffer (1x) and BSA (1x) in a final volume of 200 µL (qsp MiliQ water). The reaction was done at 37°C overnight.

Second restriction: To the previous reaction were added reactants for the second restriction: 50 Units of *Nt.Bbv*CI enzyme, 1 µL of NEB4 buffer and 4 µL of MilliQ water. Reaction was carried out at 37°C during 1h40.

The digestion products were purified on agarose gel by electrophoresis (as mentioned previously, Cf. § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109). The band of interest was cut and extracted from the gel with the kit Nucleospin Extract II (Macherey-Nagel). Procedure was performed according to the manufacturer, except the elution was carried out with 80 µL of MilliQ water warmed at 72°C.
The linearized vectors were stored at -30°C until use.

4.7.3.2. PCR amplification

The various inserts were generated by PCR, before assembling in the linearized vectors. The polymerase: Pfu Turbo Cx (Stratagene) was used for the USER cloning. The buffer used for PCR was the one provided by Stratagene.

The PCR amplification was carried out in a thermocycler (Master Cycler, Eppendorf), and consisted of:

- a stage of denaturation step of 2-5 min at 94°C
- 15 cycles (30 s of denaturation at 94°C; 30 s of hybridization at 55°C (specific temperature of the hybridization zone) and 1min30 of elongation at 72°C)
- 10 similar cycles with a hybridization at 60-64°C (specific temperature of the used primers).
- Amplification ends with an extension step at 72°C during 10 min.

Used primers are listed in Appendix: List of oligonucleotide primers p.315-317.

4.7.3.3. Cloning of the inserts in the destination vectors

The PCR products were resolved on agarose gel by electrophoresis (as mentioned previously, Cf. § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109). The spots of interest were cut and extracted by filtration on Whatman paper by centrifugation at 13 000 rpm for one minute.

Amplicons recovered from the gel as well as the destination vector were incubated in order to "create" the suitable fusion. For example *GFP* amplicons (generated for N-terminal fusion) were incubated with *CYP73A5* amplicons (generated for C-terminal fusion) and linearized pCAMBIA2300U vector to build the construct: *CYP73A5::GFP* in the vector pCAMBIA2300U.

Reaction was carried out in a PCR buffer added with 1 unit of USER enzyme (10 µL final volumes). A thermocycler (MasterCycler, Eppendorf) was used for the two-step reaction: 20 min at 37°C, then 20 min at 25°C.

The primers used during the PCR amplification allow to generate particular amplicons having an uracil base. The USER enzyme, which specifically recognizes this base, removes it and generates amplicons with especially selected cohesive ends for assembling in the linearized vector in a **directional and positional** way. In my case, they were selected so as to insert the fluorescent protein genes at the 3' end of the P450s of interest in order to preserve the N-terminus anchor sequence of P450s. For the soluble partners, fluorescent protein genes were cloned at 5'-end and 3'-end of the gene of interest.

4.7.3.4. Constructs propagation

The constructs generated as described above were used to transform *E. coli* for propagation of the vectors.

Before transformation, constructs were purified by dialysis during 45 min on nitrocellulose membrane with 0.22 µm pores (Millipore). Transformation of DH5α or TOP10 *E. coli* were done by the chemical method as mentioned previously (Cf § Bacteria transformation p.107**). It is important to do not electroporate the newly prepared construct**, which is sensitive to the electric shock, because DNA fragments were not stably bound. The colonies obtained after transformation were checked by PCR on colony as mentioned previously (program in Appendix: PCR Protocols p.312). Transformed bacteria were grown during an overnight liquid culture at 37°C and plasmidic DNA was extracted and purified as mentioned previously (Cf. § Screening of the clones p.108). All constructs were verified by DNA sequencing.

4.7.3.5. Automatic sequencing of DNA

Sequencing is carried out by the sequencing platform of the IBMP, managed by M. ALIOUA. The technique derives from the Sanger sequencing method. It is based on random incorporation of four dideoxyribonucleotides, which causes the stop of the DNA chain elongation. Each dideoxyribonucleotide (ddATP, ddTTP, ddGTP and ddCTP) is coupled with a fluorescent molecule which has a wavelength of excitation which is proper to the dideoxyribonucleotide (Prober *et al.* 1987). The sequencing was done by capillary electrophoresis on a AbiPRISM sequencer (Applied Biosystem).

4.7.4. *Agrobacteria* transformation

Agrobacteria transformation was based on the electroporation transformation method. Firstly, the bacteria were made competent, i.e. able to incorporate exogenous DNA.

4.7.4.1. Preparation of electrocompetent Agrobacteria

This protocol was based on the work of Weigel *et al*. (2006). The strain preserved in glycerol stock at -80°C was grown in LB selective solid medium on a Petri plate, containing adequate selection antibiotics. After 1 to 2 days of culture at 28°C, an isolated colony was used to sow 5 ml of LB liquid medium containing the same antibiotics. This pre-culture was incubated one night at 28°C under agitation, and served to inoculate 500 ml of new LB medium which was also incubated one night at 28°C with agitation. When the OD_{600} reached a value from 0.5 to 0.8, the culture was cooled 1 hour at 4°C then centrifuged 10 min at 4000 *g*, 4°C. The supernatant was discarded and the pellet was resuspended in 5 ml of ice-cold sterile MilliQ water. The volume of the suspension was adjusted to 500 ml with ice-cold sterile MilliQ water. This washing step was repeated three times. After the first repetition, the bacteria were resuspended in 250 ml of ice-cold sterile MilliQ water. After the second repetition, they were resuspended in 50 ml. Lastly, they were resuspended in 5 ml of a 10% (v/v) iced-cold sterile glycerol solution. The preparation was then aliquoted by 200 µL then frozen in liquid nitrogen and preserved at -80°C.

4.7.4.2. Agrobacteria electroporation and recovery

The plasmidic DNA used for transformation was mixed with 50 µL of electrocompetent bacteria thawed on ice. The mixture was transferred in an electroporation cuve and subjected to an electric shock (Capacity 25 µF; 2.5 kV; 400 Ω; Biorad gen Pulser). This electric pulse creates transient pores, allowing incorporation of the plasmidic DNA vector in the bacteria.

The transformed bacteria were diluted in 950 µL LB and incubated 4 hours at 28°C under agitation (170 rpm). This incubation allows the bacteria to reconstitute their wall and to express the resistance gene present on the plasmid.

The bacteria were then spread out over LB solid medium in a Petri plate supplemented with adequate antibiotic(s) for the selection of the transformants. Petri plate was placed 48 h at 28°C. Clones were screened by colony PCR (as mentioned previously § Screening of the clones p.108).

The plate was then preserved at 4°C and used for less than 15 days.

4.7.5. Polymerase Chain Reaction on bacteria

Transformations and cultures of bacteria and agrobacteria were followed by PCR to check the presence of the plasmid of interest in the bacteria. For the screening of the transformed colonies and the checking of the cultures, Taq polymerase (Invitrogen) was used. The protocol previously described was followed (Cf. § Polymerase Chain Reaction (PCR) over colony p.108).

4.7.6. Transitory transformation of *Nicotiana tabacum* or *Nicotiana benthamiana*

This technique allows the *in planta* transient over-expression of a protein of interest. The cDNA coding for this protein is contained in a T-DNA under the control of the *CaMV-35S* promoter in a binary vector. Generally, for the over-expression, to avoid Post Transcriptional Gene Silencing (PTGS) in plant, *Agrobacterium tumefaciens* transformed by a gene of interest were co-infiltrated with another strain of agrobacteria containing a binary vector allowing the expression of a viral protein, P19[1], able to inhibit the phenomena of PTGS.

The strains of agrobacteria were cultivated independently in LB medium containing appropriate antibiotic(s) for 20h at 28°C. Then the bacteria were centrifuged 10 minutes at 4500 *g*, 4°C and resuspended in MilliQ water. They were centrifuged again in the same conditions and resuspended in MilliQ water. This stage was repeated 2 times to wash the bacteria. After the last centrifugation, bacteria were resuspended in MilliQ water in order to obtain an OD_{600} = 0.4. Agrobacteria to co-infiltrate were mix together to equal volume.

[1] P19 protein: is a virus-encoded suppressor of gene silencing from tomato bushy stunt virus (TBSV) (Voinnet *et al.*, 2003).

The solution was infiltrated on *Nicotiana* plants of 4-6 weeks on the lower part of the leaves, using an 1 mL syringe (without needle). Plants were then grown again in the culture room. Generally 4-5 days later the leaves were observed with microscope or collected for analysis.

4.7.7. Microsomes preparation from transformed *Nicotiana benthamiana*

Microsomes were prepared from *N. benthamiana* plants to determine the activity of P450s and soluble enzymes (HCT, 4CL-1) expressed in these plants.

All the processes were done on ice. Fresh tissues were ground in Extraction Buffer: 10 ml of buffer for 1 g of fresh tissue. Grinding was done with an Ultra-turrax T25 by two successive runs of 1 minute at 7000 rpm. The mixture was filtered on miracloth, and then centrifuged 10 minutes at 10000 *g*. The resulting supernatant was filtered on miracloth then ultra-centrifuged 1 hour at 100000 *g* and 4°C. The pellet of microsomes obtained was washed with 3x2 mL of Recovery Buffer, and resuspended carefully in the minimum of Recovery Buffer (to maximize the concentrations of enzymes). The resuspension was done gently in ice-cold Potter. The microsomes and the supernatant were kept at -30°C until analysis.

Extraction Buffer	Sodium Phosphate buffer 0.1 M pH 7.4	10% (v/v) glycerol	Ascorbic acid 40 mM	*β*-mercaptoethanol 15 mM	Phenylmethanesulfonylfluoride (PMSF) 1 mM
Recovery Buffer	Sodium Phosphate buffer 0.1 M pH 7.4	30% (v/v) glycerol	-	-	-
1 tablet of antiprotease cocktail EDTA-free (Roche) was added to the buffers for 100 mL of buffer					

4.7.8. Extration and purification of total RNAs from plants

The plant RNA extraction was carried out with the kit Macherey-Nagel Nucleospin RNA plant, according to the recommendations of the provider.

Sampled tissues for RNA extraction were always directly frozen in liquid nitrogen after their harvesting. The frozen samples were first of all reduced in a fine powder, under liquid nitrogen, to allow their homogenization in the extraction buffer. Grinding was carried out manually in a mortar using a pestle or with a horizontal grinder with metal balls (Mikro-Dismembrator, Sartorius). Powders were dispatched in different 2 mL tubes, one tube per analysis to do. A fraction of the obtained frozen powder (200 mg) was quickly homogenized in the first buffer of the kit. The RNA were finally eluted with 50 µL sterile water (RNAse free) then re-eluted with the previous eluate to obtain the best possible yield.

Total RNAs were quantified by measurement of the absorbance with a spectrophotometer (Biophotometer, Eppendorf). Integrity of RNAs was checked on agarose gel (1.5% (w/v) agarose; TAE buffer 1x; Ethidium bromide 0.5 $\mu g.mL^{-1}$); the gel was then photographed under UVs. The residual presence of DNA in RNA fractions was tested by PCR with primers of the reference gene (program in Appendix: PCR Protocols p.312; primers in Appendix: List of oligonucleotide primers p.315-317).

4.7.9. Reverse-transcriptase Polymerase Chain reaction (RT-PCR)

This technique was used to synthesize the cDNA corresponding to RNAs thanks to a reverse transcriptase (RNA dependent DNA polymerase). The RNAs (2.5 µg for 20 µL reaction) were denatured 7 minutes at 65°C in the presence of 5 pmol of primers (random hexamers for the amplification of all the RNA; Roche) and dNTPs 0.5 mM each (dATP, dGTP, dCTP, dTTP). Then the solution was cooled quickly at 4°C. Polymerase and other reactants were added. The reaction was carried out in the buffer: First Strand Buffer (Invitrogen), in presence of DTT 5 mM, 40 Units of SuperScriptIII Reverse Transcriptase (Invitrogen) and 200 Units of RNaseOUT (Invitrogen). The mixture was preincubated 2 minutes at 4°C, then 10 minutes at 25°C to facilitate the hybridization of the primers, and then incubated 50 minutes at 48°C, the last step being the elongation stage. The enzyme was inactivated by heating at 70°C during 15 minutes. All these temperature steps were done in a thermocycler (Master Cycler, Eppendorf).

The cDNAs thus obtained was preserved at -30°C before further use by qPCR, semi-qPCR or PCR. Before the analysis by semi-qPCR or qPCR, the generated cDNA were diluted or not, according to the mRNA quantity of the gene to analyze.

4.7.10. Semi-quantitative PCR (semi-qPCR)

The cDNA synthesized by RT-PCR were used as matrix for the semi-quantitative PCR. This PCR was done under the following conditions:

- denaturation 2 minutes at 95°C,
- hybridization at the primers specific temperature for 30s,
- elongation 30 s at 72°C.

The hybridization temperature and the number of cycles were adapted for each gene and primers couples. The standard reactions were done with primers of housekeeping genes (i.e. Actin).

The semi-qPCR products were resolved on an agarose gel (3% (w/v) agarose; TAE 1X; Ethidium bromide 0.5 $\mu g.mL^{-1}$). The gel was photographed under UV at 254 nm.
Recipies and list of primers are in Appendix: PCR Protocols p.312; List of oligonucleotide primers p.315-317.

4.7.11. P450s activity

These assays were done as previously described (Cf. § *In vitro* activity test p.93).

Briefly, the substrate for CYP73A5 is *trans*-cinnamic acid and is commercially available, substrates of the CYP98A3 (*p*-coumaroyl shikimate) was enzymaticaly synthetized in the laboratory by P. Ullmann. For each reaction, a negative control without NADPH was carried out. The microsomes (10 µL) were put in the presence of a substrate (100 µM final) and NADPH (0.5 mM final) in potassium phosphate buffer 50 mM pH 7.4 (total volume of reaction 200 µL). This mixture was incubated 30 minutes at 28°C, under agitation. Reaction was stopped by addition of HCl 1N (V/10). A liquid/liquid extraction of the metabolites was done and metabolites were analyzed by HPLC.

Determinations of the activities were based on measure of the peak area of the formed products and the quantities determined by comparison to a reference curve obtained with solutions of known concentration of *p*-coumaric acid and chlorogenic acid, respectively for C4H tests and C3'H tests.

4.7.12. 4-CL1 activity

The assays were done as previously described (Cf. § Enzymatic activity assay p.116).

4.7.13. HCT activity

The HCT activity tests were done as previously described (Cf. § Enzymatic activity assay p.116).

Determinations of the activities were based on the measure of the peak area of the formed products and the quantities of formed products determined by comparison to an reference curve obtained with solutions of known concentration of chlorogenic acid.

4.7.14. Confocal microscopy

Samples of leaves were cut and placed between slide and coverglass. Then a vacuum step was done to have cells in the same shape with ER at the periphery of the cell. In case of the use of molecules (Latrunculin B, cinnamate...) these molecules were applied to the samples before the vacuum step to permit their absorption in the cell. Images acquisition from one preparation were done during a maximum of 1 hour, to prevent dramatic artefacts cause by drying of the tissues.

Cell imaging was performed using a LSM510 confocal laser scanning microscope equipped with an inverted Zeiss axiovert 100 M microscope (Carl Zeiss). For wide-field observations, a 10x Zeiss objective ("Plan-Neofluar") or a 20x Zeiss objective were used. For confocal resolution, images were taken using a 63x, 1.2 numerical aperture water immersion objective ("C-Apochromat").

In order to excite and to specifically observe fluorescent proteins, the microscope has different lasers (argon, laser diode and two Helium-Neon lasers), as well as AOTF filters making it possible to select the various lasers and to regulate their intensity. Dichroic mirrors are used to reflect and separate the beams of excitation and emission, finally a set of filters are used to select the wavelengths of emission and a photomultiplicator detects the signal emitted by the sample.

Images were acquired using LSM510 version 2.8 software; they were processed with the latest version of the Zeiss LSM Image Browser software (Carl Zeiss) and exported as Tiff files. Images were also directly analyzed with ImageJ software version 1.43n3 (Wayne Rasband, NIH, USA; http://rsb.info.nih.gov/ij).

Agrobacteria strains harbouring vector for the expression of eGFP::HDEL or mRFP::HDEL (used as controls) were kindly given by C. Ritzenthaler.

The characteristics of used fluorescent proteins are presented in Figure 58.

Fluorescent Protein	Excitation peak (nm)	Emission peak (nm)	Laser (excitation)	Filter settings
eGFP (enhanced Green Fluorescent Protein)	2 peaks: 395 nm /475 nm	507 nm	Argon laser (488 nm)	"band pass" (505-550nm)
mRFP (monomeric Red Fluorescent Protein)	584 nm	607 nm	Helium-Neon laser (561/584/633 nm)	"band pass" (575-615nm)

Figure 58. Characteristics of used fluorescent proteins.

C. Protein-protein and protein-membrane interactions in the early phenylpropanoid pathway

1. Interactions of 4CL-1 and HCT with membranes

1.1. In vitro interaction of 4CL-1 and HCT with membranes

The first assays to evaluate protein-protein interactions were carried out using nanodiscs-anchored P450s, testing their capacity to bind the purified soluble proteins. These experiments using our purified enzymes were carried out by Jonas Borch (University of Southern Denmark, Odense). The first *in vitro* interactions tests conducted with nanodiscs-anchored P450s led to surprising results (Figure 59).

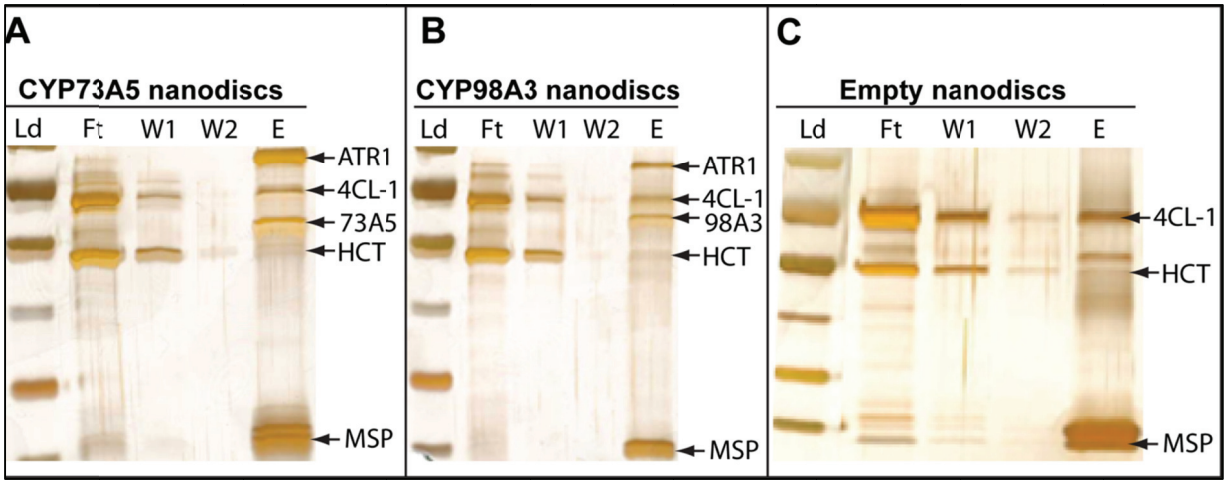


Figure 59. *In vitro* interaction of soluble proteins with nanodiscs and nanodisc-anchored P450s.
A. Interaction of 4CL-1 and HCT with CYP73A5 nanodiscs.
B. Interaction of 4CL-1 and HCT with CYP98A3 nanodiscs.
C. Interaction of 4CL-1 and HCT with empty discs.
Solution at 1.4 mg.mL^{-1} of 4CL-1 and 0.08 mg. mL^{-1} of HCT were used for each test. Five µg of 4CL-1 and one µg of HCT were incubated at 22°C during 2 hours with 2 mM of coumaric acid, 2 mM of Shikimic acid, 0.2 mM of CoA, 4 mM of MgATP and 20 µL of antiFLAG beads. Beads were washed twice 5 minutes with 150 mM NaCl, 50mM Tris-HCl pH 8. Elution was carried out with FLAG peptide at 200 ng.µL^{-1}.
Ld, Protein Mass ladder. In, Input. Ft, Flowthrough. W1 and W2, Washes. E, elution. MSP, Membrane Scaffold Protein. ATR1, *Arabidopsis thaliana* reductase 1 co-incorporated with P450s into nanodiscs. Co-immunoprecipitation experiments were carried out by Jonas Borch (Odense, Danemark).

Nanodiscs (empty or with anchored CYP73A5 or CYP98A3) were tagged with a FLAG tag attached to the membrane scaffold protein (MSP). Anti-FLAG agarose beads were used for co-immunoprecipitation. Purified recombinant 4CL-1 but not HCT, co-precipitated with nanodisc-anchored P450s but also with empty nanodiscs used as negative controls (Figures 59 A, B, C, elution fractions). This raised the following question: is 4CL-1 interacting with phospholipid membranes or with MSP forming the nanodiscs? To answer this question, interactions were tested with liposomes in the abscence of MSP.

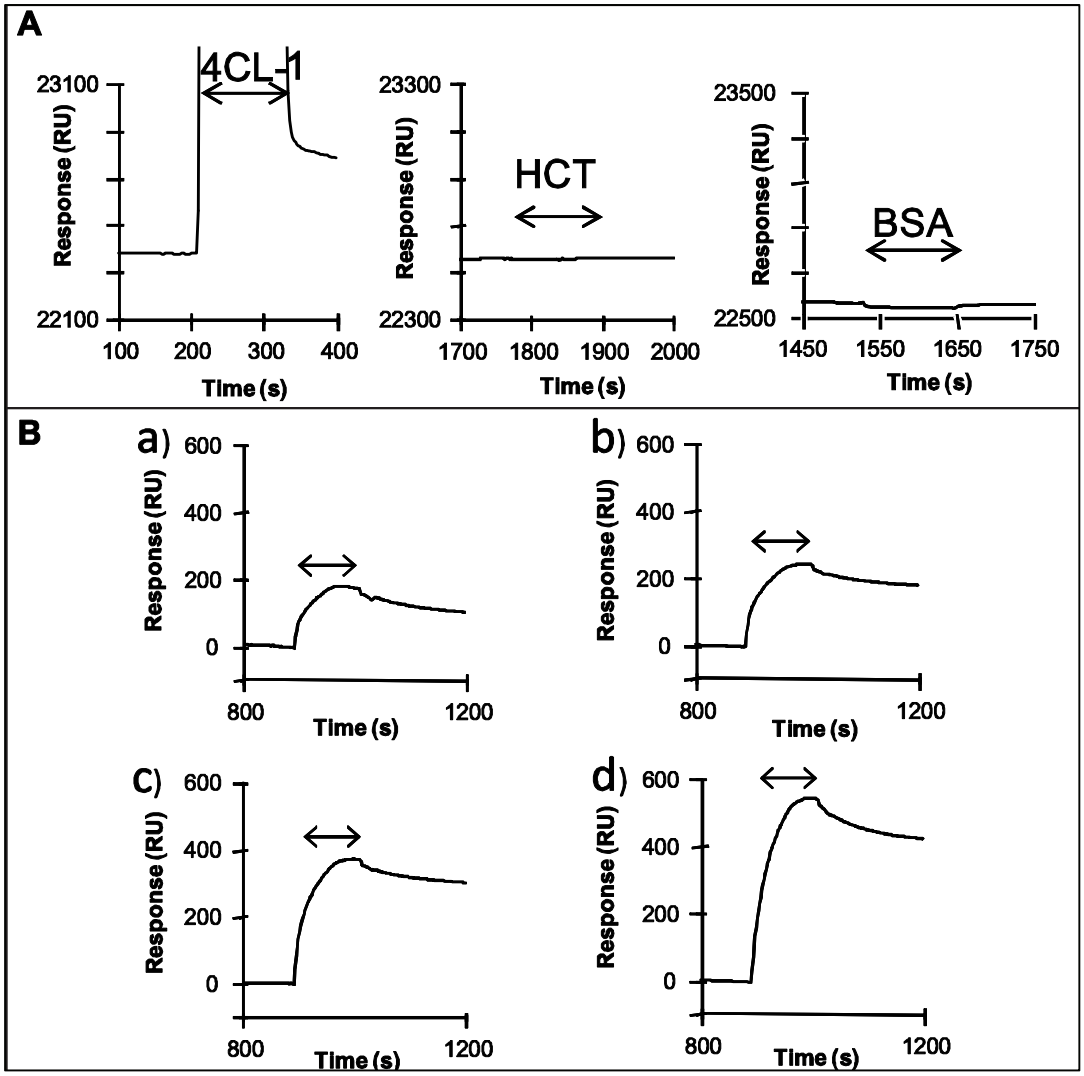


Figure 60. ***In vitro*** **interaction of 4CL-1 with liposomes.**

A. Binding of 4CL-1 and HCT to 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) phospholipid vesicles immobilized on biacore chip. BSA was a negative control.

B. Binding of 4CL-1 to immobilized liposomes of a) phosphatidylcholine; b) phosphatidylcholine + phosphatidylserine; c) phosphatidylcholine + phosphatidylethanolamine, d) asolectin (soybean phospholipid extract from Sigma). RU, resonance units. RU is a measure of refractive index near the surface of the sensorchip. Double arrows indicate time of purified protein injection. The buffer on sensorchip is the same before and after injection of protein, so difference in refractive index after the injection is due to protein binding on the immobilized vesicles. 4CL-1: 50 µg/ml; HCT: 1 µg/ml; BSA: 100 µg/ml. Surface Plasmon Resonance experiment results obtained by Jonas Borch (Odense, Danemark).

To this end, 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) vesicles were captured by hydrophobic anchors in the dextran matrix of a Biacore sensorchip and tested for binding of purified 4CL-1 and HCT (Figure 60 A). Binding of 4CL-1 but not HCT to POPC phospholipids vesicles was detected (Figure 60 A). This result corroborates the nanodiscs experiments and indicates that 4CL-1 is binding phospholipids and not MSP. It must however be kept in mind that the purified protein concentrations were quite different, HCT was fifty fold less concentrated than 4CL-1. So, HCT binding to membrane cannot be completely excluded.

It was then tested if 4CL-1 binding was influenced by phospholipid composition of the liposomes (Figure 60 B). Stronger 4CL-1 interaction was observed with liposomes prepared out of asolectin (complex phospholipid mix extracted from soybean) than with pure phosphatidylcholine, or with a phosphatidylcholine and phosphatidylserine, or a phosphatidylcholine and phosphatidylethanolamine mix. Interestingly, binding curves did not return to zero after injection stop. This indicates that **4CL-1 was strongly bound to phospholipids and adsorbed.**

1.2. In vivo interaction of 4CL-1 and HCT with membranes

4CL-1 and HCT localization was then investigated *in vivo* by confocal image analysis as described in Method details about re-localization measurements (p.134). To this end, 4CL-1 and HCT fused to eGFP were expressed in *N. benthamiana* leaves.

Fluorescent proteins could interfere with membrane binding. GFP fusions at both N- and C-terminus of 4CL-1/HCT were therefore tested.

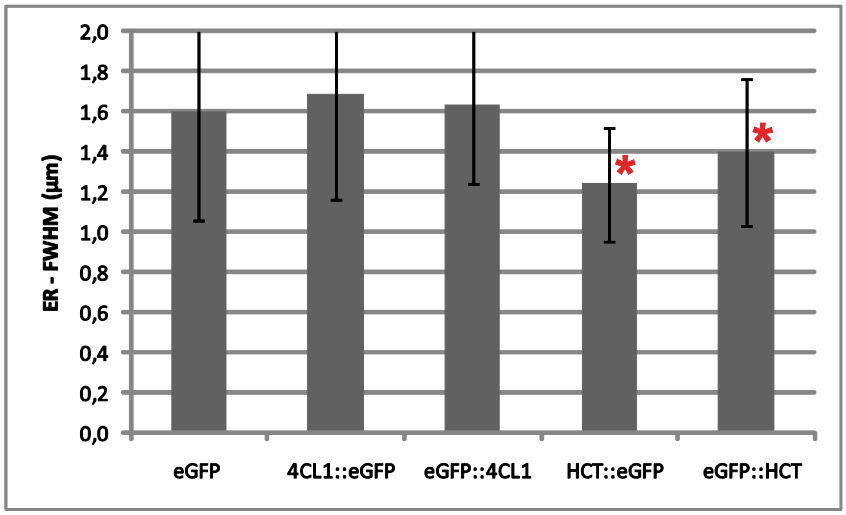


Figure 61. 4CL-1 and HCT *in vivo* localization determined by confocal image analysis.
ER – FWHM: ER-Full Width at Half Maximum. Red asterisks mean significant differences according to ANOVA tests. Details of ANOVA results in appendix p.319.

Image analysis shown in figure 61 indicates that HCT fluorescence (in the absence of co-expressed P450) is detected closer to ER membrane than fluorescence emitted from free eGFP or 4CL-1. A slight difference in the behavior of eGFP::HCT and HCT::eGFP is also detected, but variances in the populations were not equal, and ANOVA test should be interpreted with caution. In these tests, 4CL-1 repartition was similar to that of free eGFP (Figure 61).

These *in vivo* results seem to conflict with the protein behaviors observed *in vitro*. It was therefore decided to measure 4CL and HCT activities associated with isolated microsomal membranes.

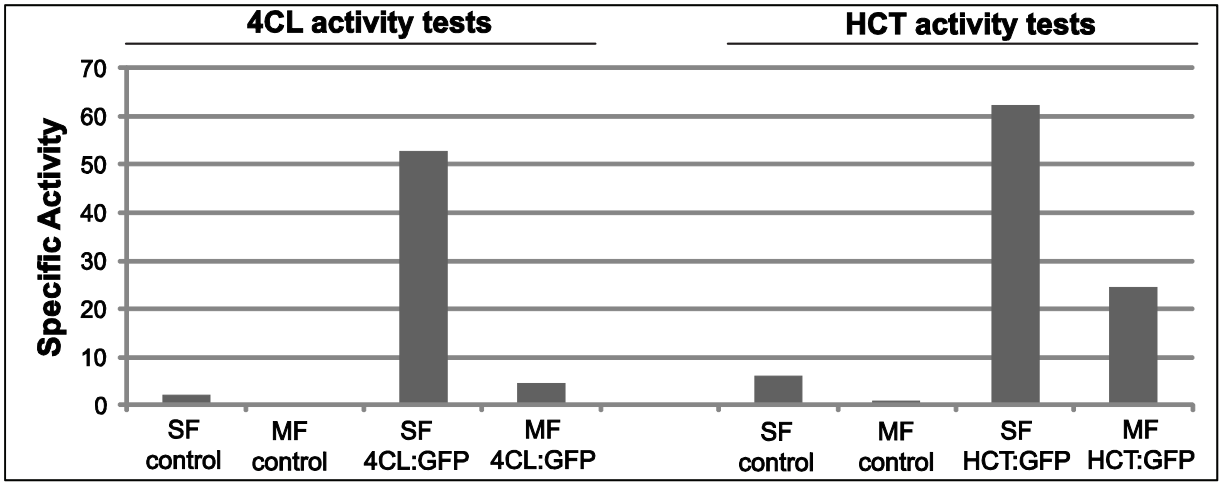


Figure 62. Detection of soluble protein activities associated with isolated microsomal membranes from leaves *of N. benthamiana* transfected with eGFP or with fluorescent fusion constructs.
Activities related to total protein concentrations of each fraction. Specific activity in µmol of product. mg^{-1} proteins.min^{-1}. Abreviations: SF. Soluble fraction; MF. Microsomal fraction.

Activities were measured with microsomes prepared from *N. benthamiana* leaves agro-infiltrated with the different fusion constructs. Microsomes prepared from plants expressing eGFP were used as control. Endogenous 4CL and HCT activities were detected in soluble fractions (Figure 62). Leaf transfection however resulted in a dramatic increase in HCT and 4CL activities. A significant proportion of HCT and 4CL activities was found associated with microsomal fractions (Figure 62). The proportion of membrane-associated activity was much higher for HCT (up to 28% of total activity) than for 4CL (less than 8%). This corroborates confocal microscopy analyses and suggests that **more HCT than 4CL-1 is associated with microsomal membranes *in vivo*.**

In a last attempt to confirm 4CL-1 and HCT association with the ER membranes *in vivo*, their interaction with ER reporter was assessed using FLIM. With FLIM, the FRET is detected by a decrease in fluorescence half-life of a donor (eGFP) in presence of an acceptor (mRFP). The efficiency of the energy transfer expressed in percentage is dependent on the distance between the two fluorescent proteins: the closer the proteins are, the higher is the percentage of FRET. When FRET is detected, proteins of interest are so close that they can be assumed to interact.

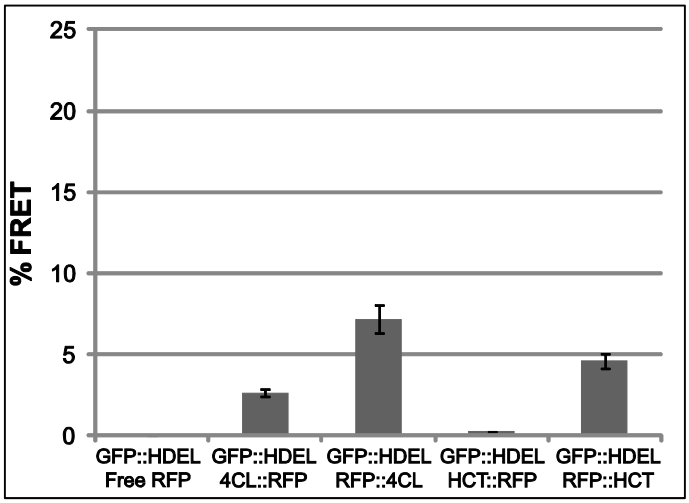


Figure 63. Interaction of soluble proteins and ER protein detected by FLIM.

Details of the FLIM analysis can be found in appendix p.323.

In order to detect 4CL-1 and HCT association with the microsomal membranes, *N. benthamiana* leaves were co-infiltrated with fluorescent mRFP constructs and eGFP::HDEL (inside the ER). Free mRFP did not interact with eGFP::HDEL, as expected for a soluble protein (Figure 63). Conversely, energy transfer was detected between the 4CL-1 or HCT constructs and eGFP::HDEL (Figure 63). **This indicates that a significant proportion of HCT and 4CL-1 is enough close to the ER** for allowing energy transfer from the ER-located eGFP. The difference in FRET observed for N- or C-terminus fusions are expected to reflect 1) the position and the orientation of the mRFP upon interaction of the proteins with the ER-located eGFP and resulting energy transfer, 2) the possible interference of the mRFP protein with 4CL-1 or HCT interaction with the membrane.

1.3. Research of potential post-translational modification sites in the peptide sequence of 4CL-1 and HCT

In addition to integral membrane proteins, a number of proteins bind to the membrane surface. Reversible binding to the membrane can result from electrostatic interaction with phospholipids, as in the case of annexins that bind to phosphatidylserine (Lemmon, 2008). Up to 10% of the proteins bind membranes due to covalent modification with a variety of hydrophobic lipid moieties (Brunsveld *et al.*, 2009), including myristate (C14), palmitate (C16), farnesyl (C15), geranylgeranyl (C20) and glycosylphosphatidylinositol (GPI) (Nadolski & Linder, 2007). Although most lipid modifications are irreversible, protein *S*-palmitoylation, also called *S*-acylation, can reversibly link palmitate (C16) to specific cysteine residues in protein substrates through thioester linkages (Nadolski & Linder, 2007). Palmitoylation enhances the surface hydrophobicity and membrane affinity of proteins, and plays important roles in modulating protein trafficking and protein stability (Linder & Deschenes, 2007). Various programs listed in the website: www.expasy.org, were used to determine if specific post-translational modification sites are predicted on the peptide sequence of 4CL-1 and HCT.

For HCT were predicted: no *O*-glycosylated residues, no GPI-modification site (GPI-SOM; big-PI Predictor programs), no *N*-myristoylation site (NMT – the predictor program), palmitoylation site (CSS-palm program), no farnesylation or geranylgeranylation site (PrePS-prenylation prediction suite).

For 4CL-1 were predicted: no GPI anchor signal (GPI-SOM; big-PI Predictor programs), *N*-myristoylation site (NMT – the predictor program), farnesylation, geranylgeranylation site (PrePS-prenylation prediction suite), but **one possible palmitoylation site was detected** at the surface of the big sub-part (position 381: SGA**C**GTV) with a probability of 0.6 (CSS-palm program). *E. coli* however lacks the enzymes for palmitoylation. The binding of 4CL-1 to membranes detected *in vitro* with recombinant purified enzyme expressed in *E. coli* thus cannot be explained by its palmitoylation.

Taken together all these data suggested that 4CL-1 and HCT are to some extent associated with the ER membrane. Clear evidence was obtained that 4CL-1 can bind phospholipids *in vitro*. HCT binding to the membranes is detected only *in vivo*. This might be due to the low concentrations of purified functional protein that can be obtained *in vitro*. It is also possible that membrane binding of HCT requires the presence of a ligand that is present in the plant membranes. A potential candidate could be endogenous tobacco CYP98As. In the same way 4CL-1 could bind tobacco CYP73As. The role of the endogenous P450s should however be minimal considering the large amounts of soluble proteins expressed under the control of the *CaMV-35S* promoter.

2. Interactions of 4CL1 with CYP73A5 and HCT with CYP98A3

It was next investigated if co-expression of CYP73A5 and CYP98A3 influenced the localization of 4CL-1 and HCT *in vivo* when fluorescent fusion constructs were co-expressed in the leaves of *N. benthamiana* (Figure 64).

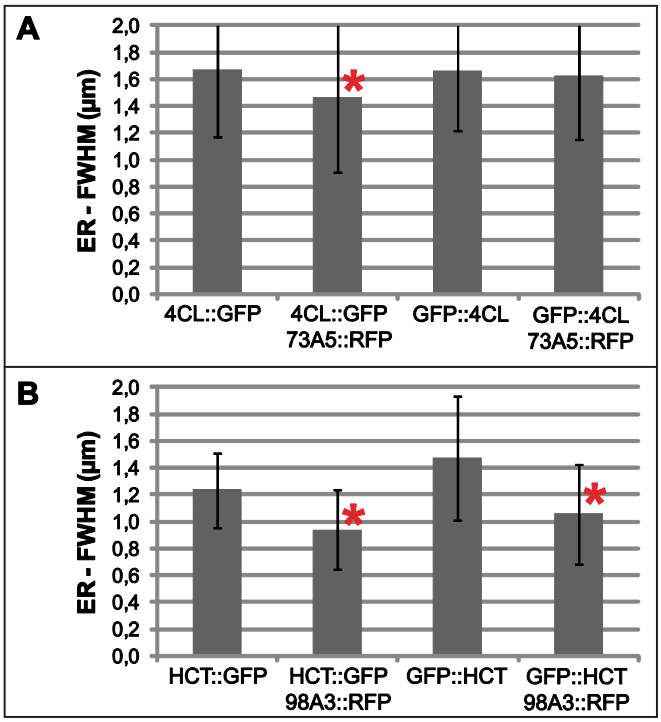


Figure 64. Re-localization of soluble proteins upon co-expression of partner P450s detected by confocal microscopy.
A. 4CL1 re-localization.
B. HCT re-localization.

Red asterisk means significant difference compared to control according to ANOVA tests. Detailed ANOVA results can be found in appendix p. 319-320. Details on the experimental protocol can be found in § 4.5.1 and 4.5.2 (p.133-136).

When 4CL-1 was co-expressed with CYP73A5, a slight re-localization was detected, with the 4CL-1::eGFP C-terminal fusion construct (Figure 64 A). When HCT was co-expressed with CYP98A3, a significant re-localization was detected both with the N- or C-terminal eGFP fusion constructs (Figure 64 B).

Direct interaction of the proteins was then further tested by FLIM.

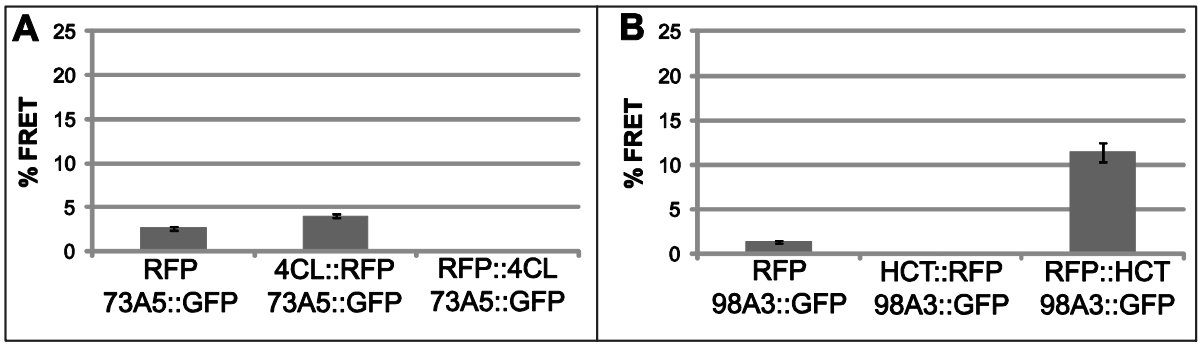


Figure 65. Interaction of 4CL-1 and HCT with their partner P450s detected by FLIM.
A. **4CL-1 interactions with CYP73A5. B. HCT interactions with CYP98A3.**
Details of the FLIM analysis can be found in appendix p.323.

The FRET detected with free RFP and the GFP-fused P450s, is in the range usually considered as non significant (<5%). So, the FRET detected with 4CL-1::mRFP/mRP::4CL-1 and CYP73A5::eGFP couples is also non significant (Figure 65).
In the case of HCT, very significant FRET (11 %) was detected between mRFP::HCT and CYP98A3::eGFP. Probably due to spatial organization, no FRET was however observed with the HCT::mRFP construct.

Thus, taken together, re-localization and FRET data clearly suggest interaction between HCT and CYP98A3. In the case of co-expression of 4CL-1 and CYP73A5 both re-localization and FRET data are too elusive to conclude to protein-protein interaction.

3. Interactions of 4CL-1 with CYP98A3 and HCT with CYP73A5

Formation of a lignin metabolon onto the ER membrane would imply bridging of the two P450-anchored units and their metabolic partners. Some proteins might play a more significant role in the nucleation of this metabolon. This was investigated by testing interactions of the 4CL1 / CYP98A3 and HCT / CYP73A5 couples.

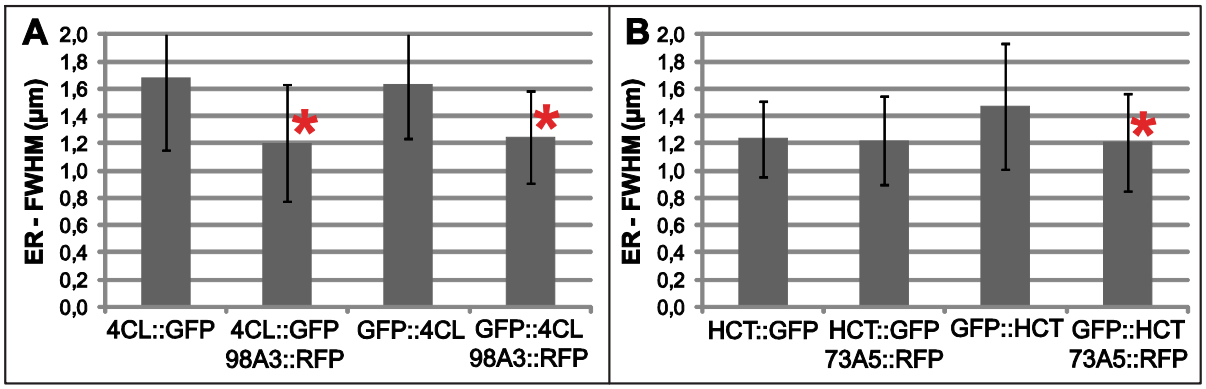


Figure 66. Re-localization of soluble proteins upon co-expression of non-direct partner P450s detected by confocal microscopy.

A. **4CL1 re-localization. B. HCT re-localization.**

Red asterisk means significant difference compared to control according to ANOVA tests. Detailed ANOVA analysis can be found in appendix p.320-321.

Interestingly, 4CL-1 co-expression with CYP98A3 led to a clear re-localization near the membrane whether tested with eGFP at the N- or the C-terminus of 4CL-1 (Figure 66 A). More limited re-localization of the eGFP::HCT fusion upon co-expression with CYP73A5 was also observed (Figure 66 B). Potential cross-interactions were thus tested by FLIM.

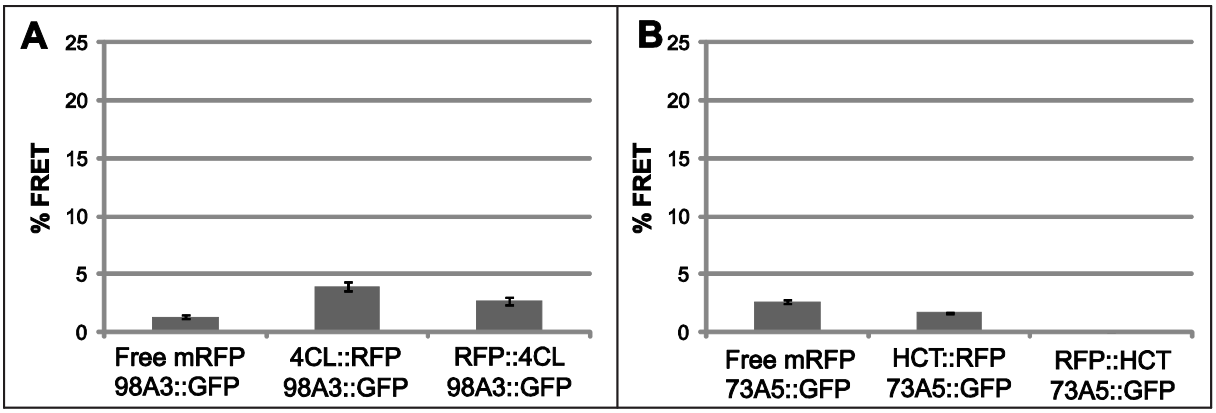


Figure 67. Interaction of 4CL-1 and HCT with their non-partner P450s detected by FLIM.
A. 4CL-1 interactions with CYP98A3. B. HCT interactions with CYP73A5.
Details of the FLIM analysis can be found in appendix p.324.

Only low FRET was detected between CYP73A5 and HCT fusions, that was at the lower limit to be considered as significant (Figure 67). No FRET was detected between CYP98A3 and 4CL-1 fusion proteins. As the absence of FRET may result from mis-orientation and mis-positioning of fluorescent proteins, this does not allow to conclude to the absence of interactions.

Our date thus indicate a redistribution of the 4CL-1 nearer to the membrane upon co-expression of CYP98A3, but do not allow to conclude to a direct interaction.

4. Homo- and hetero- oligomerization of the P450s

Evidence for mammalian P450 homo- and hetero-dimerization and oligomerization has been repeatedly reported (Subramanian *et al.*, 2010). P450 oligomers were thus considered as a potential platform for nucleation of a lignin metabolon.

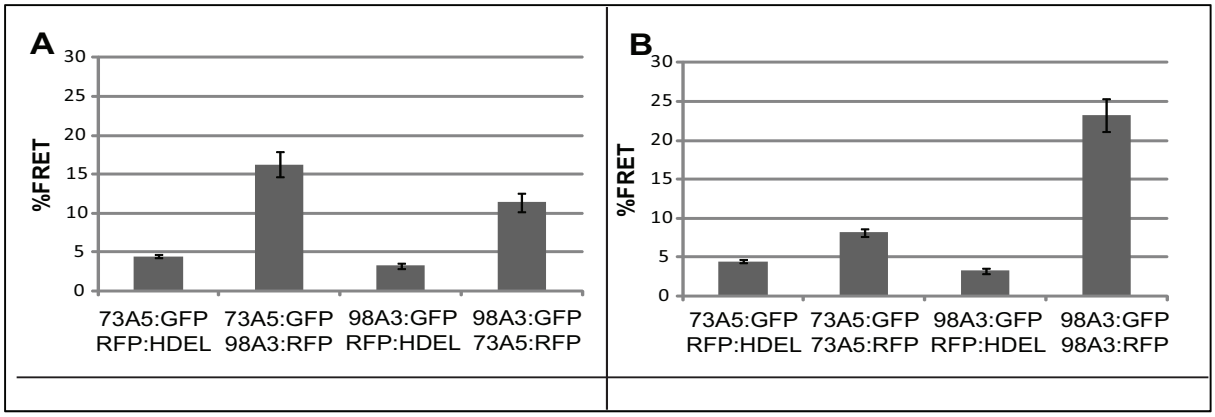


Figure 68. P450 oligomerization detected by FLIM.
A. P450 Hetero-oligomerization.
B. P450 Homo-oligomerization.
Details of the FLIM analysis can be found in appendix p.324.

As shown in Figure 68, **very significant FRET was observed indicating both homo- and hetero-di(oligo)merization of CYP73A5 and CYP98A3**. The largest decrease in fluorescence half-life (corresponding to around 23% FRET) was observed for co-expression of CYP98A3::eGFP and CYP98A3::mRFP, that strongly suggested homo-oligomerization of CYP98A3 in the plant ER membranes. **Interestingly, CYP73A5 appeared proner to form heterodimers with CYP98A3 than homodimers.**

5. Interactions upon co-expression of the four potential members of the metabolon

As suggested by the impact of CYP98A3 on the localization of 4CL-1, the formation of a metabolon is not expected to only involve interactions between sequential enzymes, but implies complex interactions between multiple partner proteins. Some of them might not have any enzyme activity, but just a structuring or regulatory function. Co-expression of multiple partners might thus enhance interaction and help to stabilize the complex. To test this hypothesis, all four candidate partner proteins were transiently co-expressed and their localization and interactions again investigated by confocal microscopy and FLIM.

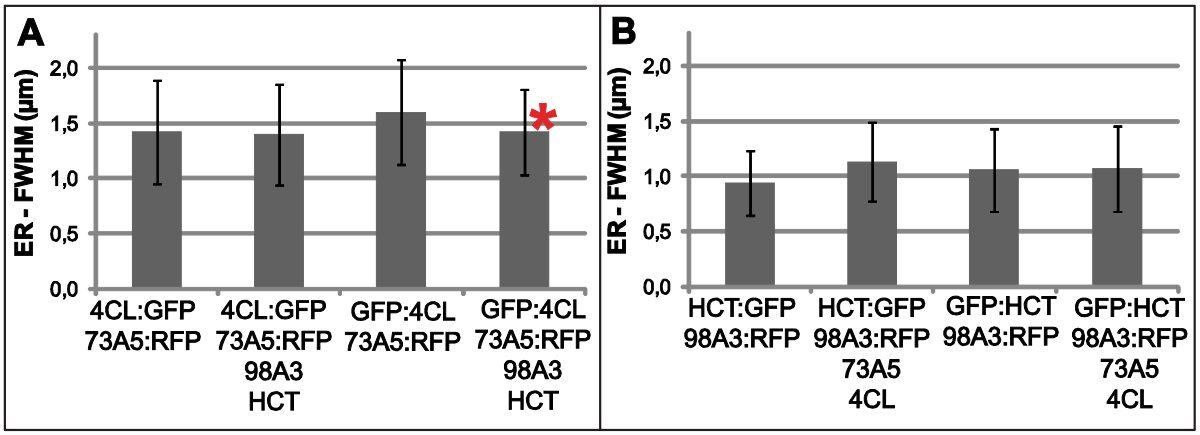


Figure 69. Re-localization of the 4CL-1 and HCT upon co-expression of the four potential partners.
A. Re-localization of 4CL-1. B. Re-localization of HCT.
In each case, one couple was expressed without fluorescent marker for this experiment. Red asterisk means significant differences compared to control according to ANOVA tests. Detailed ANOVA analysis can be found in appendix p. 321-322.

In our previous experiments, an impact of the expression of CYP98A3 on the repartition of HCT and of 4CL-1 has been observed (Figures 64 B and 66 A). Some redistribution of 4CL-1 and HCT upon co-expression of CYP73A5 was also recorded. Co-expression of all four proteins however did not trigger further re-localization of the soluble proteins nearer to the membrane (Figure 69). Change in protein repartition observed with a single partner was nevertheless confirmed.

To sum up, soluble protein localization observed upon co-expression of two partners was not further modified upon co-expression of four enzymes.

Protein interaction upon co-expression of all four enzymes was then investigated by FLIM. In each case one of the couples was expressed without fluorescent marker for this experiment.

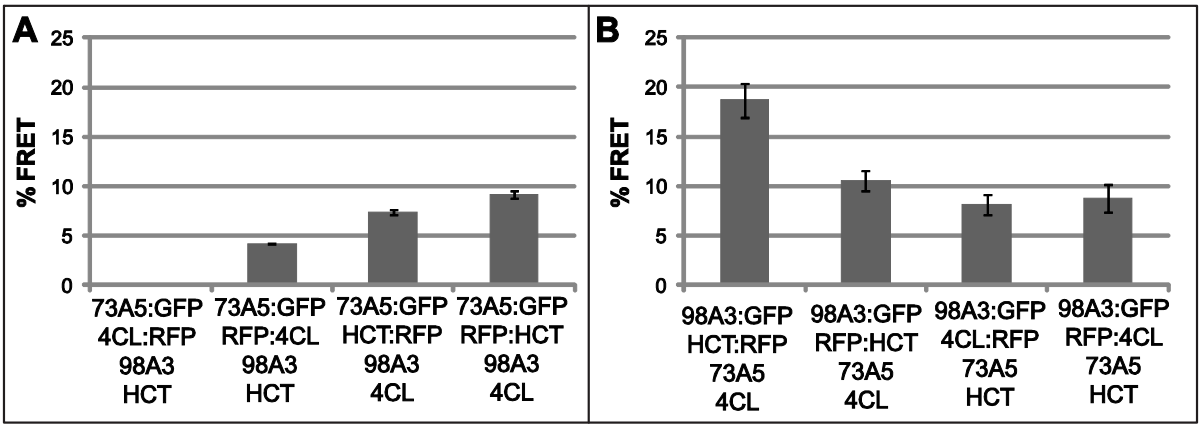


Figure 70. Protein interactions detected upon co-expression of 4CL-1, HCT, CYP73A5 and CYP98A3.

A. Interaction with CYP73A5

B. Interaction with CYP98A3

In each case, one couple was expressed without fluorescent marker for this experiment. Details of the FLIM analysis can be found in appendix p.324. Expression of non GFP-tagged protein was checked by semi-quantitative PCR.

Co-expression of all four proteins led to a dramatic increase in the FRET detected between CYP73A5::eGFP and HCT (compare Figures 70 and 67). No significantly enhanced interaction was detected for CYP73A5::eGFP with 4CL-1, conversely interaction between 4CL-1 and CYP98A3::eGFP were improved. Interaction of CYP98A3::eGFP with the HCT C-terminal mRFP fusion construct was improved, and from undetectable became very significant (compare Figures 70 and 65). FRET up to 19% was observed in the latter case for C-terminal HCT::mRFP fusion construct.

FRET data indicate that cross-interactions are very significantly increased upon co-expression of all four hypothetical enzymes partners of the lignin metabolon. CYP98A3 seems to play an essential role in the nucleation of this metabolon.

6. Interaction between 4CL-1 and HCT

4CL-1 and HCT are consecutive enzymes in the phenylpropanoid pathway, with 4CL-1 producing the CoA-esters used as substrates by HCT. As data described in the last paragraph indicate that both 4CL-1 and HCT seem to interact with CYP98A3 upon co-expression of all four enzymes, their direct interaction could be expected. FLIM experiments were undertaken to test this hypothesis.

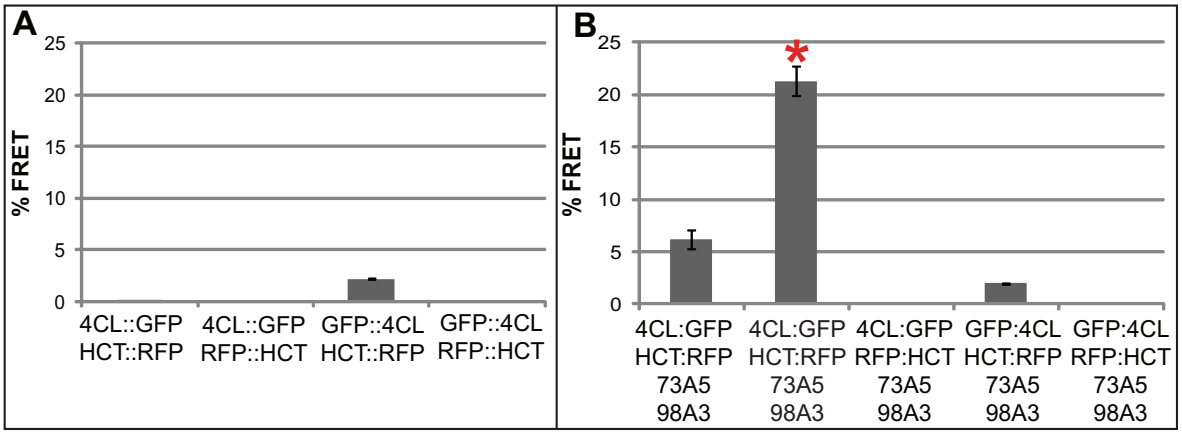


Figure 71. Interaction of 4CL-1 and HCT detected by FLIM.

A. Interaction in absence of co-expressed P450. B. Interaction when CYP73A5 and CYP98A3 were co-expressed. Expression of non GFP-tagged proteins was checked by semi-quantitative PCR. Measures performed when clear re-localization of the soluble protein was observed are marked by a red asterisk. Similar results were obtained (results not shown) by inversion of the fluorescent reporter between soluble proteins i.e. HCT fused to eGFP and 4CL-1 fused to mRFP. Details of the FLIM analysis can be found in appendix p.325.

In the absence of co-expressed P450s (Figure 71 A), no FRET between 4CL-1 and HCT cosntructs was detected even by switching the fluorophore (results not shown). No clear re-localization of these proteins to the periphery of the ER membrane was observed during FLIM acquisition.

When co-expressing the four proteins, very significant FRET was detected only when clear re-localization was observed during acquisition (with the combination: 4CL-1::eGFP / HCT::mRFP / CYP73A5 / CYP98A3). In this case, up-to 21% of FRET was recorded (Figure 71 B). Re-localization of the soluble proteins was optimal in these conditions. This however only occurred in a small number of cells. For the majority of the tested combinations, no clear FRET was detected (Figure 71 B). But, re-localization of the soluble constructs was not observed by eye during FLIM experiments.

The high FRET observed in a limited number of cells between 4CL-1 and HCT upon co-expression with CYP73A5 and CYP98A3 has to be confirmed before being considered as a reliable result.

7. Mobility of the proteins constituting the metabolon

A parameter not taken into account at the start of the project was the extremely dynamic properties of the plant ER, implying fast movement of both P450s with(in) the ER membranes (Figure 72).

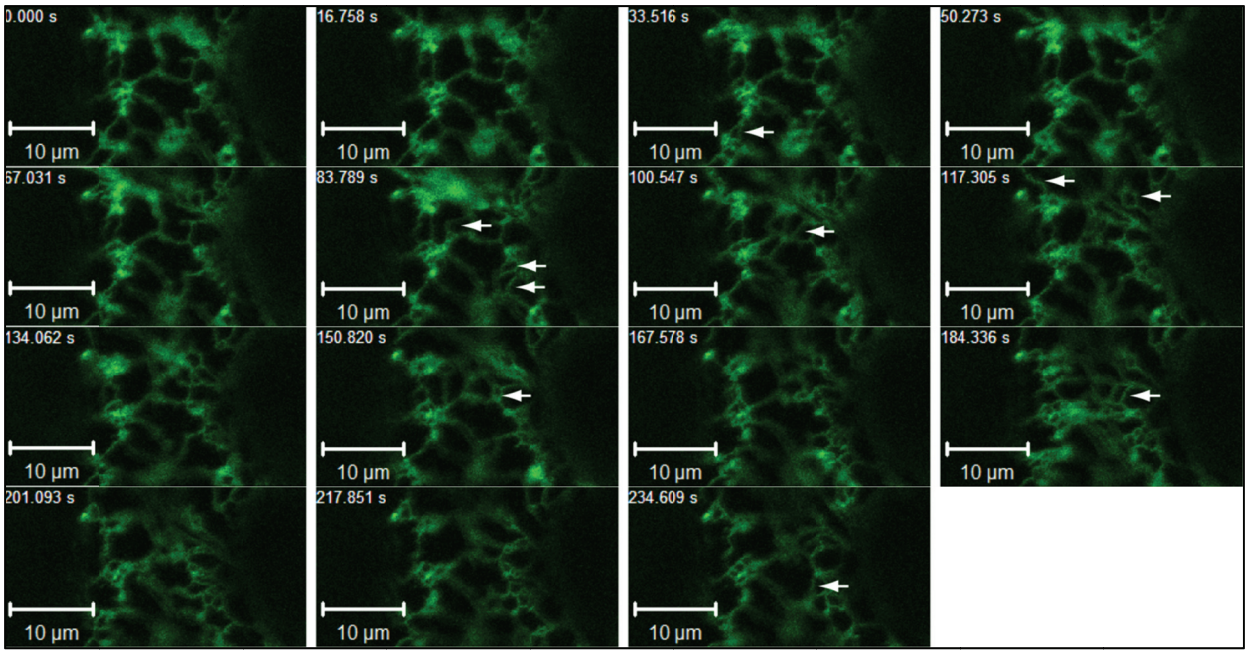


Figure 72. Movement of the CYP73A5::eGFP within the ER of *N. benthamiana* leaf epidermal cells. Latrunculin B was <u>not</u> used during this experiment. Acquisition during 234.609 s every 16,758 s. Important remodeling of the ER structures were highlighted by arrows. Scale bar: 10 µm.

ER was proposed to act as a trafficking network, delivering lipid, protein, calcium and signalling molecules to different regions of the cell (plastids, plasma membrane, nucleus…)(Griffing, 2010). Rates of tubules growth and shrinkage were reported to be of about 1.5-2 $\mu m.s^{-1}$, thus providing a strong stirring of the cytosol (Griffing, 2010).

This property could favor the recruitment of the soluble proteins onto ER membranes. In the same way, fast motion of P450s on the fluid membranes could play a similar role. Conversely, trapping of P450s in membrane lipid rafts or membrane structures could slow down their movement and help to stabilize a complex. To obtain more information on P450 mobility in the ER membranes, Fluorescence Recovery After Photobleaching (FRAP) experiments were carried out.

FRAP is a technique in which a cell region of interest is selectively bleached with a high-intensity laser. The fluorescence recovery occurs when non-bleached fluorescent proteins move into the bleached region. This recovery allows determination of several parameters of protein diffusion, protein transport, mobile fraction… (Pawley, 2006).

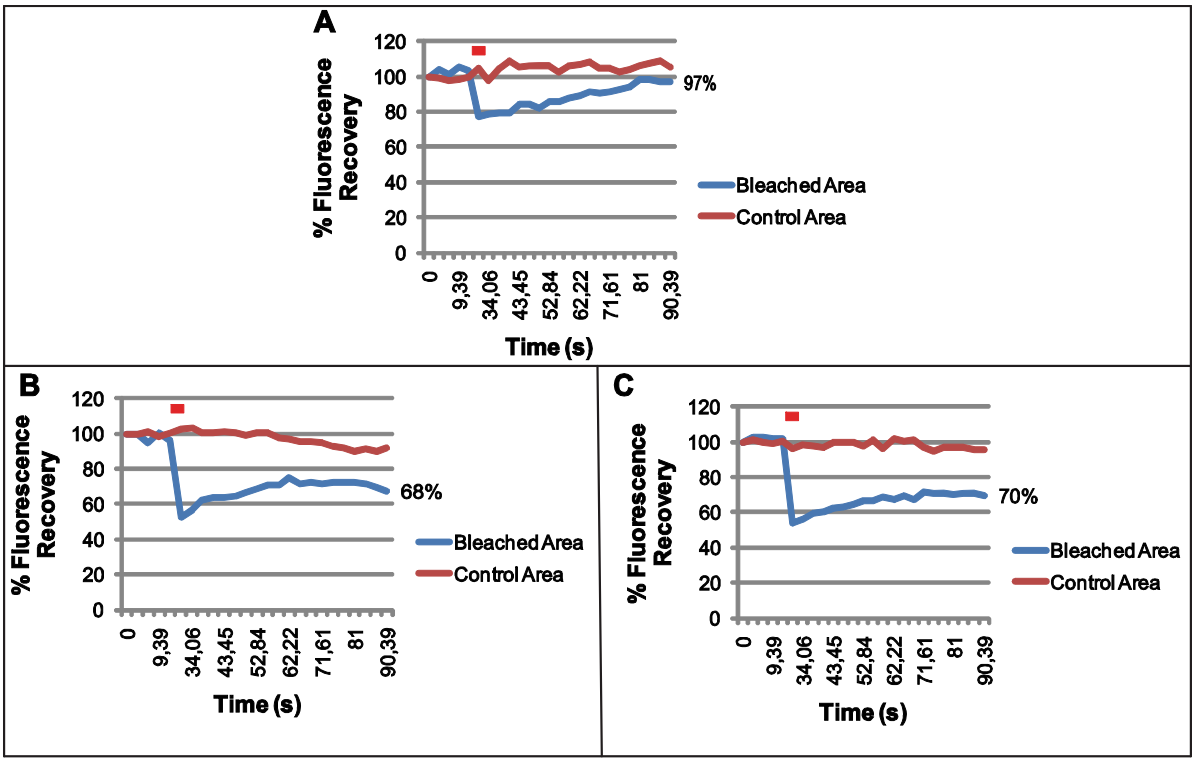


Figure 73. Movement of P450s detected by FRAP experiments.
A. Control eGFP::HDEL. B. CYP73A5::eGFP. C. CYP98A3::eGFP.
Red bar indicates the bleach. Experiment was performed with Latrunculin B (20µM).

The rates of fluorescence recovery vary among cells and ER areas. Nevertheless, a general tendency reflecting ten independent experiments for each combination can be delineated. Compared to the fluorescence recovery observed for ER eGFP::HDEL, recovery observed with CYP73A5::eGFP or CYP98A3::eGFP was not total (Figure 73). A simple equation for determining the mobile fraction (Mf) from FRAP experiments is: Mf = (Ff – Fo) / (Fi – Fo), with Ff: Intensity of fluorescence after total recovery, Fo: fluorescence just after the bleach, and Fi: fluorescence before the bleach. For experiments with eGFP::HDEL, the mobile fraction represented 100% of the eGFP::HDEL. For experiments with P450s, up to 70% of the fusion proteins were immobile. This would suggest that P450s are associated with a membrane sub-structure that limits their mobility.

Metabolons in secondary metabolism are assumed to be transient, dynamic and not stable. Metabolon dynamic was dependent on the dynamic properties of its constituants. Soluble proteins must be re-localized to ER membrane to interact with P450s. Thus the soluble proteins are expected to be very mobile, to move from the cytoplasm to the ER and vice versa.

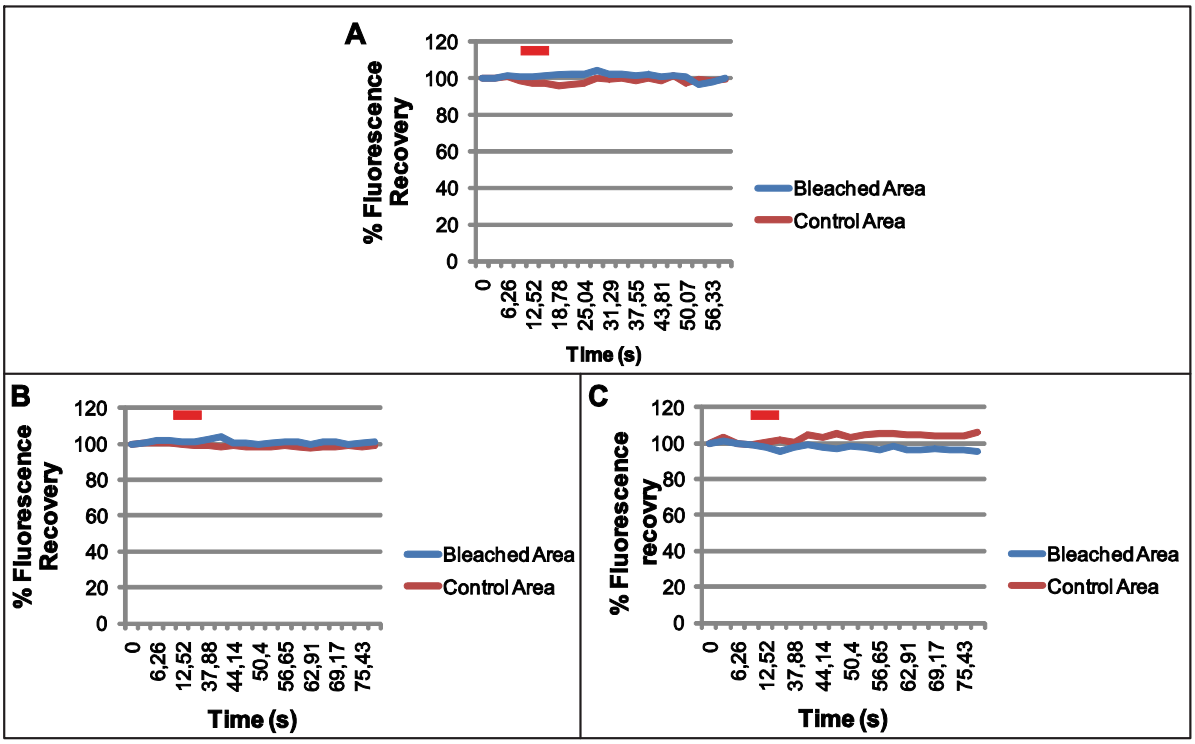


Figure 74. Movement of 4CL-1 and HCT detected by FRAP experiments.
A. **Control eGFP. B. 4CL-1::eGFP. C. HCT::eGFP.**
Red bar indicates the bleach. Experiment was performed with Latrunculin B (20µM).

As expected, movement of the soluble proteins (eGFP, HCT::eGFP and 4CL-1::eGFP), in the cytoplasm at the periphery of the ER membrane, is so fast that no decrease in fluorescence can be detected during the bleach (Figure 74). Even when increasing the laser power to the upper limit, and increasing the bleach duration.

Upon co-expression of the four enzyme partners, soluble enzymes seemed to be re-localized on ER membranes to interact with P450s. Therefore, FRAP experiments were carried out after co-expression of all four enzymes on cells where clear re-localization was detected. In each case one of the couples was expressed without fluorescent marker for this experiment. FRAP was done as much as possible on the ER tubule membranes by bleaching an area closer to the ER tubule.

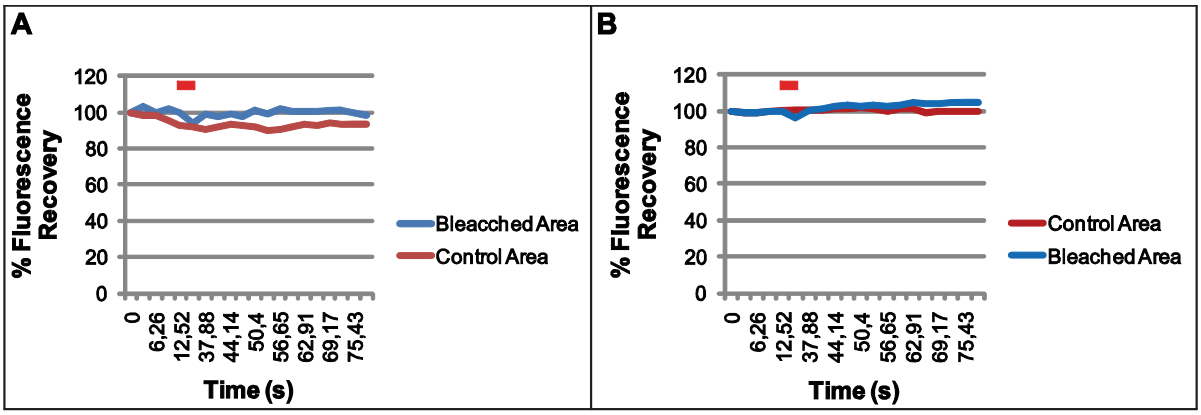


Figure 75. Movement of 4CL-1 and HCT co-expressed with P450s evaluated by FRAP experiments.
A. 4CL-1::eGFP co-expressed with HCT, CYP73A5::mRFP and CYP98A3. B. HCT::eGFP co-expressed with 4CL-1, CYP73A5 and CYP98A3::eGFP
Red bar indicates the bleach. Experiment was performed with Latrunculin B (20µM).

A very slight decrease (less than 10%) in fluorescence was detected during the bleach (Figure 75 A, B). Recovery was immediate at the end of the bleach, even when longer bleach or higher laser intensities were tested. While faint, this decrease in fluorescence, not observed when HCT::eGFP and 4CL-1::eGFP were expressed alone, could be indicative of the involvement of the soluble proteins in a metabolon. Targeting of only ER membrane was difficult, and cytosolic fraction was also recorded during FRAP experiments, thus limiting the detection of the immobilized soluble proteins.

P450 proteins seems partly immobilized on the membrane by association with a substructure (actin, raft, metabolon). Soluble proteins are highly mobile, with slight slow-down when co-expressed with all four partner proteins.

8. Discussion

The aims of this project were to test the potential of novel approaches (Immobilized Nanodiscs onto SPR chip and FLIM experiments), to prove that protein-protein interactions occur between the four proteins (4CL-1, CYP73A5, HCT, CYP98A3) in the core phenylpropanoid pathway, and to prove the existence of a metabolon.

The first striking result that emerged from this work was that **4CL-1 and HCT associate with artificial and ER membranes *in vitro* or *in vivo.*** This is important since binding of the "soluble" component to the membrane limits their mobility to two dimensions and thus increase the probability of protein-protein interaction. In addition to integral membrane proteins, a number of proteins bind to the membrane surface and can associate with it by electrostatic interaction with phospholipids or post-translational modifications. One of the important post-translational modifications is the insertion of hydrophobic lipid moieties. This modification facilitates membrane binding, thereby restricting also the diffusional mobility of the proteins from three to two dimensions, with concomitant significant effects on the interactions with other cellular components (Brunsveld *et al.*, 2009). Numerous proteins could be covalently modified by a variety of lipids and one palmitoylation site was detected at 4CL-1 surface. Palmitoylation is however not likely to be responsible for the phospholipid affinity of 4CL-1 that was demonstrated *in vitro* with the *E. coli* expressed enzyme. *In vivo* interaction of HCT and 4CL-1 transgenes from *A. thaliana* with endogenous P450s from *N. benthamiana* can not be excluded. Nevertheless, implications of endogenous P450s must play a minimal role in comparison to the high quantities of soluble proteins expressed under the control of the *CaMV-35S* promoter of the fusion constructions.

Another interesting result was that **CYP73A5 and CYP98A3 form homo- and hetero-dimers or oligomers.** The formations of both homo- and hetero-meric P450 complexes were already reported for mammalian enzymes (Kelley *et al.*, 2006; Subramanian *et al.*, 2010) and these complexes were shown important for the catalytic activity (Hazai *et al.*, 2005; Praporski *et al.*, 2009). P450 oligomerization may help to scaffold and stabilize the metabolon by formation of a larger nucleation site.

Simultaneous expression of all four proteins increased FRET with direct partners and with other members of the metabolon. This strongly supports the formation of a metabolon. This structure seems to stabilize interactions and decrease the distances between partner proteins. Existence of metabolon in the phenylpropanoid pathway was already suggested by various approaches, including precursor feeding experiments, co-precipitation, co-isolation and co-localization. But no direct evidence for the formation of the metabolon has been obtained, leading sometimes to a great scepticism on the existence of such metabolons. For example, channeling was confirmed by double-labeling with *in vitro* isolated microsomes from tobacco stems or cell suspension cultures (Rasmussen & Dixon, 1999). But, the same double-labeling (using [^{3}H]Phe and [^{14}C]cinnamate) approach failed to demonstrate channeling of endogenously formed *p*-coumarate in yeast (*Saccharomyces cerevisiae*) expressing recombinant poplar PAL and C4H (Ro & Douglas, 2004). These studies however did not examine whether PAL was co-localized with C4H on the yeast endomembrane. This could be important when coexpressing enzymes from a plant that encodes multiple isoforms of each enzyme.

In our experiments, a major problem was that soluble proteins were not re-localized in all cells. But **when re-localization was visible, FRET was clearly detected**. Thus, to detect FRET between soluble protein constructions, these proteins must be re-localized together on their nucleation site. This could depend on the presence of additional partner proteins or on the presence of susbtrates and co-factors. In 2004, a study demonstrated *in planta*, that the sub-cellular localization of two PAL isoforms is different and dependent on the amount of C4H available to organize a complex with PAL1 (Achnine *et al.*, 2004). PAL1 associated with the endoplasmic reticulum was thought responsible for channeling cinnamic acid. In this study, FRET measures to detect direct physical interaction of PAL1 with C4H were not conclusive. This is in agreement with our results. When only two members of the metabolon were co-expressed the FRET were elusive. But they were improved by co-expressing all the four partners. Metabolon formation may also require still unidentified plant-specific factors, such as nucleating proteins, which could explain the difficulties to detect *in vitro* interactions or interactions in yeast, as previously discussed.

Substrates, cofactors, intermediates or stress might contribute to stabilize the formation of the metabolon. We thus investigated the impact of cinnamate, coumarate, NADPH, shikimate, spermidine and wounding on the protein re-localization. Spermidine, not directly involved in this biosynthetic pathway, has already been described as modulator of the electrostatic protein-protein interactions involving CYP11A1 (Berwanger *et al.*, 2010). Eleven combinations of these molecules and stress were tested with all the combinations of co-expressed enzymes (results not shown). No clear improvement of soluble enzyme re-localization was noticed, due to the high variability of the response to these various combinations.

Another way to promote naturally or artificially the formation of the metabolon could be the co-expression of scaffold proteins which are non-enzymatical stabilizing proteins. Indeed, **a significant proportion of the P450s appeared to be anchored to an unknown membrane component**, which could indicate involvement of such proteins in the formation of this metabolon. Recent studies demonstrated the importance of such scaffold proteins. Interaction of CYP2C2 with BAP31 is important for its ER retention and for its expression level (Szczesna-Skorupa & Kemper, 2006). On the other hand, microarray analyses suggested that the transcription pattern of *ERG28* is closely correlates to genes involved in sterol synthesis (Gachotte *et al.*, 2001), and it was finally shown to function as an endoplasmic reticulum transmembrane protein, acting as a scaffold to tether the ergosterol biosynthetic enzymes (Mo & Bard, 2005). More recently, it was demonstrated that synthetic scaffold protein can also improve efficiency of metabolic engineering (Moon *et al.*, 2010).

To identify novel partners of the metabolon, nanodiscs-anchored P450s were used to trap other interactants from *Arabidopsis thaliana* cellular crude extracts by SPR-MS. **Only quantitatively major proteins (such as Rubisco) were found interacting with the nanodiscs** (Results not shown)**.** This experiment might be just impaired by the lack of scaffold proteins inside the nanodiscs. Transcriptomic analysis indicates that the expression of different proteins is closely correlated to the expression of the genes encoding enzymes of the phenylpropanoid pathway. For example, *At1g52760* an uncharacterized coding-protein sequence belonging to the esterase/lipase/thioesterase family according to its sequence, is highly correlated to CYP73A5 and CYP98A3 in CYPEDIA (http://www-ibmp.u-strasbg.fr/~CYPedia/) (r-value of 0.779 for CYP73A5 and 0.575 for CYP98A3). Another difficulty for mimicking the interactions observed *in vivo*, is the protein concentration achieved in the cytoplasm that is extremely high (Ellis, 2001). Small molecules, pH and physicochemical parameters also have to be taken into account. Finally, cytoskeleton could also be important for the recruitment of the soluble enzymes. As latrunculin B was used for all experiments, actin cytoskeleton seemed not to mediate protein recruitement and interactions. Preliminary tests were achieved without Latrunculin B and no improvement in the re-localization was observed (Results not shown). Actin may however contribute to P450 immobilization onto the membrane.

It is still difficult to model protein organization in the metabolon. P450 activity requires participation of a CPR. This last protein is thus expected to contribute to the formation of the metabolon. Human CPR was shown not selective for human P450, and several P450s compete for a single CPR (Watanabe *et al.*, 1993). Ability of the CPRs from different species or even from different kingdoms to functionally complement each other suggests that their P450-interaction domains are well conserved (Jensen & Møller, 2010). Accordingly, P450s could be organized around a CPR, or since we showed that P450s form oligomers, the metabolon could structure like a web on the ER membrane.

However, this structure seems slack and highly mobile. Achnine *et al.* in 2004 already mentioned that the channeling between PAL1 and C4H did not result from a tight physical interaction. The channeling of the phenylpropanoid intermediates might not result from physical contact between the enzymes, but rather from a proximity effect of the active sites in a highly viscous gel-like environnement that maintains low diffusion (the cytoplasm). This proximity effect of the active sites of the different enzymes could also permits a higher flexibility of the system.

Taken together all **our data strongly support the existence of a lignin metabolon**. Indeed, FRET between non-consecutive enzymes in the pathway was detected. They were improved upon co-expression of four consecutive enzymes in the pathway. **CYP98A3 which is the first P450 specific to the lignin branch pathway seems to have an important nucleation role in the formation of this metabolon. Our data also validate the methods that have been proposed and suggest that they could be used to investigate the formation of metabolons in other biosynthesis pathways.**

9. Prospects

New evidence has been obtained for the formation of a metabolon in the phenylpropanoid pathway. Nevertheless, the requirements for the formation of this metabolon are still not understood. Different approaches could be proposed.

A clear limit appeared during the re-localization analyses: the soluble proteins expressed at much higher levels than the P450s. Thus, re-localization was clearly highly under-estimated. Indeed, there was not enough P450 on the ER membranes to re-localize all soluble proteins. A novel strategy could be envisaged using a single transgene to coordinate expression of multiple proteins encoded as a polycistronic messager. This system is based on the activity of the 2A peptide of foot-and-mouth disease virus (FMDV) that acts co-translationally to cleave the polyprotein. So, multiple proteins could be expressed in a ratiomeric manner under the control of the same promoter (El Amrani *et al.*, 2004; Geu-Flores *et al.*, 2009). But the limit of this method is that sometimes the co-translational cleavage is not effective. Preliminary tests were conclusive, and led to unprecedent re-localization of HCT co-expressed with CYP98A3 (results not shown). HCT was still soluble proving that co-translational cleavage was effective.

FRAP experiments should be repeated, to confirm if the movement slow-down of the soluble proteins co-expressed with all the other partners was statistically relevant. This assay carried out with a more favourable protein ratio could be also more informative.

In vitro and *in vivo* interaction and re-localization trials should be tested with candidate scaffold proteins, to determine if such proteins support the formation of a metabolon. In the same way, interactions with the CPR and the impact of the CPR co-expression on metabolon formation have to be performed. In addition, a third P450 (F5H) is involved in the lignin biosynthesis pathway. This is an additional player expected to interact with CYP73A5, CYP98A3 and soluble enzymes of the pathway. We showed that P450s are immobilized on the ER membranes by undetermined structures. It would be interesting to determine if P450s are associated to lipid raft, after such raft purification. Actin cytoskeleton does not seem involved in the recruitment of the soluble enzymes. It might thus rely on the microtubule cytoskeleton. Microtubule inhibitor as aminoprophos could be tried to prevent re-localization of the soluble enzymes. Native gel and Tandem Affinity Purification tag (TAP-TAG) could be tested to identify unknown non-enzymatical partners.

Finally, to understand the structure and formation of the metabolon, determination of the interacting domains of the protein seems a prerequisite. This information could be obtained by chemical cross-linking and surface labelling followed by mass spectrometry. These data could be interpreted in light of resolved structure or homology model and confirmed by site-directed mutagenesis.

A more systematic approach for elucidating protein-protein interactions in the phenylpropanoid pathway can also be considered. This would be to explore the specific roles of the different isoforms of each enzyme (as PALs, 4CLs, CYP98A8/CYP98A9 versus CYP98A3). Such an approach could be very informative since protein-protein interactions in the phenylpropanoid pathway would be likely to support partitioning of the metabolic fluxes between the different branches of the pathway.

10. Materials and Methods

Preparation of the USER constructs for tobacco transient over-expression, agro-infiltration of tobacco leaves, microsomal preparations and enzymatic assays, semi-quantitative-PCR and confocal laser scanning microscopy analyses, and finally production of hypothetical interactants, were done as described in the previous parts of this chapter.

10.1. Fluorescence lifetime imaging microscopy (FLIM)

The tissues were loaded as described above (Cf. § Confocal microscopy p.149). FLIM experiments were performed by Time-Correlated Single-Photon Counting (TCSPC) as described in Brandner *et al.*, 2008.

Measurements were done on a home-built two-photon system based on an Olympus IX70 microscope with an Olympus 60x 1.2NA water immersion objective. Two-photon excitation was provided by a titanium: sapphire laser (Tsunami, Spectra Physics), which was setup to an emission wavelength of 900 nm (for eGFP excitation). Imaging was performed with a laser scanning system using two fast galvo mirrors (Model 6210; Cambridge Technology). Photons were collected using a two-photon short pass filter with a cut-off wavelength of 680 nm (F75–680; AHF), and a band-pass filter 520±17 nm (F37–520; AHF). Fluorescence was analysed by a fiber-coupled avalanche photodiode detector (SPCM-AQR-14-FC; Perkin Elmer), which was connected to a time-correlated single photon-counting module (SPC830; Becker & Hickl), which operates in reversed start–stop mode.

The samples were scanned continuously for about 1-4 minutes to obtain appropriate photon numbers for reliable statistics for the fluorescence decays. Data were analyzed using a commercial software (SPCimage V2.9; Becker & Hickl), which uses an iterative reconvolution method to recover the lifetimes from the fluorescence decays. Measures were done on the whole image, and for each condition, average value of all taken images were used to calculate FRET efficiency.

FRET calculations were done according to the equation: $FRET = 1 - (\tau FRET / \tau ALONE) = (R_0^6 / (R_0^6 + R))$ with R_0 the Förster radius, R the distance between donor and acceptor, τFRET, the lifetime of the donor in the presence of the acceptor, and τALONE the lifetime of the donor in the absence of acceptor. Standard deviation was calculated as follow: $SD_{FRET} = Average_{FRET} \sqrt{((SD_{FRET}/Average_{FRET})^2+(SD_{ALONE}/Average_{ALONE})^2)}$

10.2. Fluorescence Recovery After Photobleaching (FRAP)

The tissues were loaded as described above (Cf. § Confocal microscopy p.149). Images were collected using Zeiss LSM510 laser scanning confocal microscope, with a C-Apo-chromat 63x water objective. FRAP on the ER was performed within regions of interest using a high-intensity bleach mode (with Argon laser at 488 nm).

Followed parameters were used during scanning: unidirectional scans, intensity at 40 mA, line average 2, 3.15 s scan time. Region of interest to bleach measure 90 pixels (x) over 10 pixel (y). Five prebleach images were recorded, and area to bleach was photobleached during 200 scans (300 to 400 during tests with soluble enzymes). After the bleach, a series of postbleaching scans were performed. The postbleach fluorescence was sampled every 3.15 s during approximately 1 min 30 s.

Chapter II

Functional analysis of the *CYP98* family in *Nicotiana tabacum*

« … la diversité doit être considérée comme un don,
puisqu'il faut être divers pour échanger et échanger pour progresser. »
Axel Kahn - Et l'Homme dans tout ça ?

1. Introduction

1.1. The phenylpropanoid pathway: more complicated than expected

The phenylpropanoid pathway was initially described as a linear succession of hydroxylations, methylations and reductions on a simple phenylpropanoid backbone deriving from phenylalanine.

The first unexpected reports that started to modify the linear lignin pathway showed that caffeoyl methylation occurred on the coenzyme A conjugate and not on the free acid (Maury *et al*., 1999; Ye *et al*., 2001). A grid-like pattern then started to emerge. At the same moment, another surprising discovery emerged when the 3-hydroxylation of the phenolic ring was shown to be catalyzed on the shikimate or quinate esters of *p*-coumaric acid by a cytochrome P450 from the *CYP98* family, *CYP98A3* in *Arabidopsis thaliana* (Schoch *et al*., 2001; Franke *et al*., 2002a). This result suggested an essential role for HCT in lignification and this role was soon confirmed (Hoffmann *et al*., 2003; 2004; Niggeweg *et al*., 2004). Altogether, these new data implied a new regulatory key in the pathway, by modulation of the flux towards lignin precursors by the shikimate concentration in the plant cell (Schoch *et al*., 2006), which means dependence on photosynthetic activity.

An additional level of complexity in the phenolic pathway had been suspected for a long time, considering the unexplained data reported by several laboratories attempting to engineer lignin content and composition (Adbulrazzak *et al*., 2006; Besseau *et al*., 2007; Coleman *et al*., 2008). The hypothesis of an alternative *meta*-hydroxylation pathway of phenolic compounds was proposed (Abdulrazzak *et al*., 2006; Schoch *et al*., 2006), regarding changes in soluble phenolics or lignin composition observed in the tissues of plants genetically modified for targeted steps in the pathway, that were not always those expected from the current pathway model. Indeed, a new branch in the pathway recently emerged in *Arabidopsis thaliana* from *in silico* predictions (Ehlting *et al*., 2008). This analysis provided a list of genes likely to contribute to a parallel, and perhaps, redundant branch pathway. The prediction was validated by the converging work of several laboratories (Fellenberg *et al*., 2008; Grienenberger *et al*., 2009; Matsuno *et al*., 2009). This branch pathway is described in the review in Appendix p.331.

1.2. *CYP98As* in plant kingdom

CYP98As are known as hydroxycinnamoyl ester 3'-hydroxylases since the discovery and the identification of *CYP98A3* from *A. thaliana*. Ten years later, about forty sequences of *CYP98As* were annotated from moss to *Arabidopis*. Based on these data, it appears that several *CYP98A* paralogues are often present in a given species. In tobacco, a southern blot analysis by Rachel Million-Rousseau (PhD Thesis, 2006) suggested that only three *CYP98As* were present. Subsequently, three *CYP98As* sequences were detected in EST collections.

All biochemically characterized CYP98As have been shown to metabolize *p*-coumaroyl shikimate (CS) or *p*-coumaroyl quinate (CQ) with a preference for *p*-coumaroyl shikimate, except CYP98A8 and CYP98A9 (the two paralogues of *CYP98A3* in *Arabidopsis*) which have lost these activities.

In classical phylogenetic reconstructions, based on protein sequence comparison, CYP98A8 and CYP98A9 do not belong to the monophyletic clade that includes all confirmed *p*-coumaroyl ester hydroxylases from all Angiosperms, and even CYP98As from Gymnosperm and moss. This led Matsuno *et al.* (2009) to perform a codon substitution analysis, from which they concluded that *CYP98A8* and *CYP98A9* result from the accelerated evolution of an ancestral *CYP98A* duplicated in Brassicaceae or one of their close ancestors. It was shown with recombinant enzymes and *in vivo* that while both CYP98A8 and CYP98A9 share the ability to catalyze the triple *meta*-hydroxylation of N^1,N^5,N^{10}-*tri*(*p*-coumaroyl)spermidine, only CYP98A8 is able to catalyze the second *meta*-hydroxylation of the three phenolic rings to convert N^1,N^5,N^{10}-triferuloylspermidine into N^1,N^5,N^{10}-tri(hydroxyferuloyl) spermidine. The final product of this pathway, N^1,N^5-di(hydroxyferuloyl)-N^{10}-sinapoyl spermidine is a major constituent of the pollen coat and possibly contributes to the pollen wall structure. This alternative phenolic pathway thus offers an opportunity for the formation of syringyl and sinapoyl polyamine-conjugated units.

In addition to CS and CQ, the metabolism of other molecules was tested. As a result, a new level of complexity in phenolic metabolism appeared with the functional description of several CYP98A enzymes isolated from plant species known to accumulate specific phenolic esters, like chlorogenic acid in coffee and rosmarinic acid in *Ocimum basilicum*, *Lithospermum erythrorhizon,* or *Coleus blumei*. Some of these CYP98A enzymes were shown to be able to catalyze 3-hydroxylation of 4-coumaroyl-4’-hydroxyphenyllactate (CYP98A13, Gang *et al*., 2002; CYP98A6, Matsuno *et al*., 2002; CYP98A14, Eberle *et al*., 2009) or *p*-coumaroyl quinate (CYP98A35, Mahesh *et al*., 2007), and in most cases also 3’-hydroxylation of the shikimate ester. The hydroxyphenylactate ester is an intermediate for the rosmarinic acid biosynthesis. This acid presents several biological activities: antioxidant, anti-inflammatory drug, anti-mutagenic, bactericide and antiviral (Petersen & Simmonds, 2003). It is very interesting to note that the 4-coumaroyl-4'-hydroxyphenyllactate is hydroxylated only by CYP98s of plant species accumulating rosmarinic acid.

Chlorogenic acid and rosmarinic acids were considered as final soluble metabolites accumulated in some plant species, rather than possible intermediates in lignin biosynthesis. But 3’-hydroxylation of different esters in the same plant (or in different plants) started to raise the hypothesis of an additional branching in the phenolic grid pathway.

In the same way, extensive screening permitted to isolate eight *CYP98As* from *Triticum aestivum*. It was shown that three of these CYP98A enzymes preferentially metabolize *p*-coumaroyl shikimate. However, CYP98A11 and CYP98A12 can also hydroxylate *p*-coumaroyl tyramine (Morant *et al*., 2007), with kinetic properties suggesting a possible physiological role in the synthesis of feruloyl tyramine, a component of the cell wall and of suberin deposited on the wound areas.

Thus, in addition to the shikimate and quinate esters, other *p*-coumaroyl esters or amides seem to be *meta*-hydroxylated by CYP98As (Figure 76), in particular in plant species accumulating specific compounds such as rosmarinic acid or specific phenolamides derivatives.

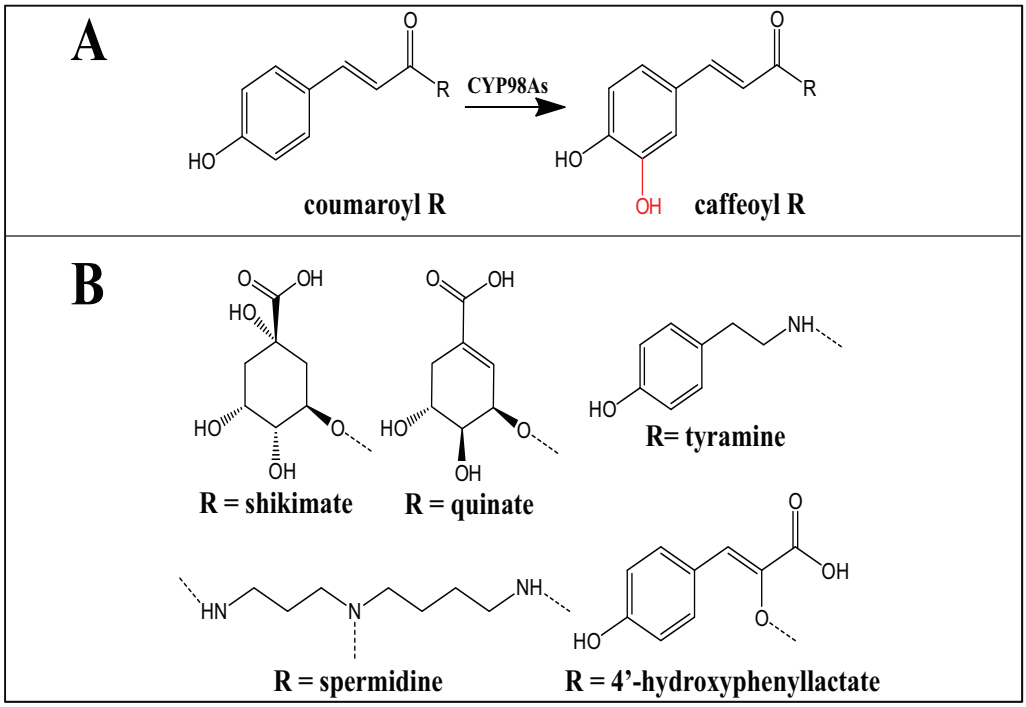


Figure 76. Substrates for the *meta*-hydroxylations by CYP98As.
A. The catalyzed reaction.
B. Ester or amides R side chains of the different substrates.

Consequently, a set of potential substrates were assayed with the three tobacco CYP98As, such as phenolamides, *p*-coumaroyl shikimate and *p*-coumaroyl quinate, precursors of respectively phenolamides, lignin and chlorogenic acid.

1.3. Chlorogenic acid

Chlorogenic acid (*trans*-5-O-caffeoyl quinic acid; CGA) is a major soluble phenolic compound and an efficient antioxidant. It is present in large amounts in some tissues from Solanaceae (tobacco, tomato, potato) and Rosaceae (apple). CGA is however not accumulated in the model plant *Arabidopsis thaliana*.

Despite many studies in different plants, its synthesis and the reasons for its accumulation in some plants are still not well understood.

1.3.1. Discovery of chlorogenic acid biosynthesis pathway

Biosynthesis of chlorogenic acid in potato tuber discs is stimulated by treatment with *L*-phenylalanine and *t*-cinnamic acid. In addition to chlorogenic acid, two other quinic acid esters accumulate in treated tissues: the cinnamoyl quinic acid and *p*-coumaroyl quinic acid (Levy & Zucker, 1960). Levy and Zucker initially thought that these two esters were direct precursors of chlorogenic acid. Moreover, they demonstrated the conversion of these two esters into chlorogenic acid *in vivo* and *in vitro*. But in 1974 it was shown that chlorogenic acid formation depends completely on the presence of ATP and CoA, and that cellular suspensions of *Nicotiana alata* catalyze chlorogenic acid synthesis from caffeoyl CoA and quinic acid (Stockigt & Zenk, 1974).

Rhodes and Wooltorton then studied the enzymology of chlorogenic acid biosynthesis. They purified HQT and suggested that two CGA biosynthetic pathways were possible: one via *p*-coumaroyl CoA, the other via caffeoyl CoA (Rhodes & Wooltorton, 1976). However, higher affinities of *p*-coumarate:CoA ligase for *p*-coumarate (compared to caffeate) and of HQT for *p*-coumaroyl CoA (compared to caffeoyl CoA) suggest that the synthesis occurs via *p*-coumaroyl CoA and *p*-coumaroyl quinate (Figure 77). The same year, a correlation between the activity of phenylalanine ammonia-lyase, C4H and accumulation of CGA in potato tuber was observed (Lamb & Rubery, 1976).

Finally in 1987, it was shown that a cytochrome P450 from carrot catalyses the conversion of *p*-coumaroyl quinate into caffeoyl quinate (Kühnl *et al*., 1987).

1.3.2. Roles of chlorogenic acid

Chlorogenic acid seems to act as an antioxidant in the plants in which it is present (Solanaceae and Rosaceae). In *Coffea canephora* (also known as Robusta coffee), hydroxycinnamoyl quinic acids accumulate in the beans where their level can reach up to 10% of dry weight. Among them, chlorogenic acid represents up to 68% of the hydroxycinnamoyl esters, and it impacts caffeine levels and coffee bitterness.

Inhibition of CGA accumulation in tobacco causes the acceleration of cellular death in mature leaves (typical reaction of oxidative stress) and increases oxidized lipids level (Tamagnone *et al.*, 1998). CGA acts as a potent hydrogen-donating antioxidant (Grace & Logan, 2000), even more effective than ascorbic acid. UV irradiations of cells increase accumulation of CGA (Koch *et al.*, 1990).

In some plants, CGA plays a role in pathogen resistance. Indeed, CGA formation in potato tuber increases clearly after wounding. It was thus proposed to accumulate in the cell wall, forming a physical barrier against disease-causing agents (Matsuda *et al.*, 2003). Furthermore, in tobacco, a reduction in CGA causes an increase in the susceptibility of the plant to pathogens such as *Cercospora nicotianae* (Maher *et al.*, 1994). After infection by the bacterium *Pseudomonas syringae*, tomato plants with high level in CGA show a slower and more limited cellular death than control plants (Niggeweg *et al.*, 2004).

Recent work highlights the interest of CGA antioxidant activity for human health. Epidemiological studies show that a diet rich in fruits and vegetables reduces the incidence of cardiovascular illness and cancers (Bazzano *et al.*, 2002). CGA contributes to this reduction, among fruits and vegetables, some such as potato, tomato or eggplant produce large amounts of CGA. It is also accumulated, at a lower level, in apples, pears, plums. *In vitro*, CGA limits oxidation of low density lipoproteins (LDL), a major factor of atherosclerosis (Laranjinha *et al.*, 1994). Moreover, it is reported to quench reactive species like alkylperoxyl radical or peroxynitrite and to prevent carcinogenesis by reducing their damage on DNA (Grace *et al.*, 1998; Sawa *et al.*, 1999).

CGA thus acts like an important antioxidant in plant and animal. For this reason, strategies are developed for increasing its level in food crops. However, the CGA pathway is not yet understood in the relevant plants.

1.3.3. Recent work on chlorogenic acid

CGA accumulation in cotyledons of *Coffea arabica* was shown to precede the synthesis of phenolic polymers. Therefore it had been suggested that bean lignification could use CGA accumulated during seed development as carbon supply (Aerts & Baumann, 1994).

Today, several biosynthetic pathways are still considered. In some species, such as sweet potato, CGA would be formed via UDP glucose:cinnamate glucosyl transferase (UGCT) and hydroxycinnamoyl D-glucose:quinate hydroxycinnamoyl transferase (HCGQT) (Niggeweg *et al.*, 2004). In other species such as tomato, tobacco and potato, HQT is involved in its formation. According to the data published, HQT can be involved in the different branches of the phenolic pathway (Figure 77):

- For direct production of CGA from caffeoyl-CoA and quinic acid by HQT,
- To synthesize *p*-coumaroyl quinate (from *p*-coumaroyl-CoA and quinic acid), *p*-coumaroyl quinate being then converted into chlorogenic acid by a C3'H.

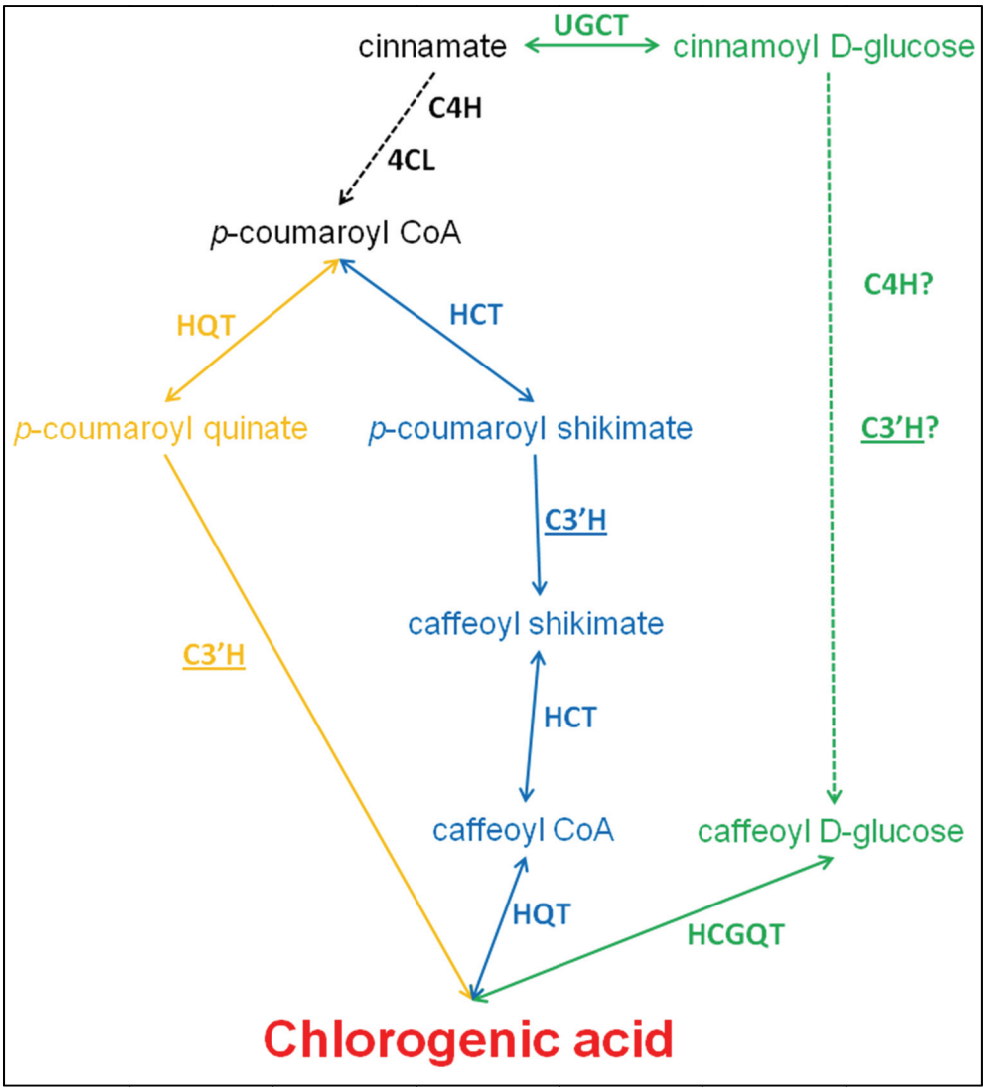


Figure 77. Proposed pathways for the biosynthesis of chlorogenic acid.
UGCT: UDP glucose:cinnamate glucosyl transferase; HCGQT: hydroxycinnamoyl D-glucose:quinate hydroxycinnamoyl transferase; C4H: cinnamate 4-hydroxylase; 4CL: 4-coumaroyl-CoA ligase; HQT: hydroxycinnamoyl-CoA quinate :hydroxycinnamoyl transferase; HCT: hydroxycinnamoyl-CoA shikimate :hydroxycinnamoyl transferase; C3'H: *p*-coumaroyl ester 3'- hydroxylase.

The second possibility is supported by the fact that several CYP98s enzymes from *Arabidopsis*, sweet basil, and wheat, were reported to catalyze hydroxylation of *p*-coumaroyl quinate, although with lower efficiency than hydroxylation of *p*-coumaroyl shikimate (Morant *et al.*, 2007). Furthemore in recent work on *Coffea canephora* (plant accumulating CGA), two *CYP98A* cDNAs were isolated and enzymes characterized. One of them metabolized efficiently *p*-coumaroyl shikimate, but did not metabolize *p*-coumaroyl quinate. The second catalyzed 3'-hydroxylation of *p*-coumaroyl shikimate and *p*-coumaroyl quinate with the same efficiency. It is the only reported CYP98A efficiently converting both substrates. This suggests that *CYP98As* are involved in direct formation of chlorogenic acid (Mahesh *et al.*, 2007). Preliminary work on tobacco by Rachel Million-Rousseau in the laboratory suggested that CYP98A30 and CYP98A31 could be involved in the biosynthesis of chlorogenic acid (PhD manuscript of R. Million-Rousseau, 2006).

In fact, the different pathways might coexist, and the main flux in each tissue may just depend on the local concentrations in quinate, shikimate, *p*-coumaroyl CoA and caffeoyl CoA (Niggeweg *et al.*, 2004) but also in HCT and HQT. In all cases, CGA biosynthesis will require the concerted action of C3'H and HQT.

To conclude, chlorogenic acid is one of the most abundant phenolic compounds found in higher plants. It could serve as carbon storage, phytoanticipin, antioxidant, and it is accumulated at high levels during fruit maturation (for its protection?). Its biosynthetic pathway is still not fully elucidated. Nevertheless, PAL, C4H, 4CL, HQT and a C3'H seem to be required.

1.4. Phenolamides

Phenolamides are considered as a major class of secondary metabolites. They are bioactive compounds sharing properties of both polyamines and phenylpropanoids. Both their phenolic and polyamine moieties can contribute to the biogenesis and consolidation of the cell-wall. Phenolamides accumulation was shown to be correlated with fertility, as well as cell-wall growth and strengthening both in monocots and dicots. Di- and tri-conjugated spermidines are major constituents of the anthers or pollen in all Angiosperms. While the exact role and mode of action of phenolamides is not elucidated, they are considered as growth regulators, translocated or storage forms of phenolics and polyamines that can be remobilized when required. Accumulation of phenolamides was also correlated to stress responses. They were reported to have antiviral and antifungal activities. Interestingly, an alternative phenylpropanoid route based on phenolamides instead of phenolic esters should be influenced by polyamine rather than by shikimate/quinate concentrations. Phenolamides might thus be actors of the cross-talk between phenolic and nitrogen metabolism (Figure 78).
Review in Appendix p.331 for an overview of the phenolamides.

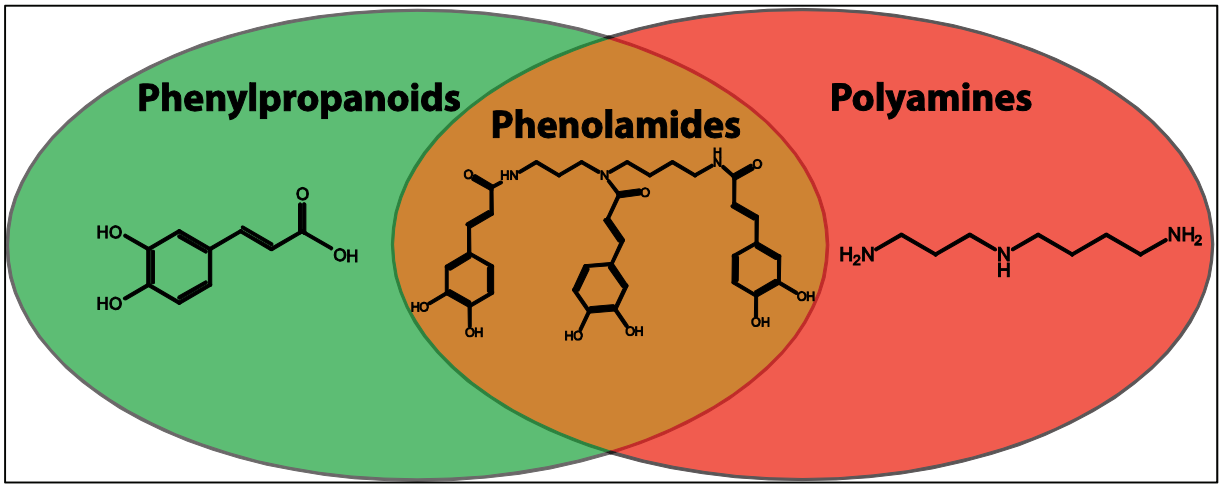


Figure 78. Phenolamides: Bridging polyamines to the phenolic metabolism.

1.5. Questions asked and evolution of the project

During my thesis, the main questions of the project slightly evolved concomitantly with the new insight in phenylpropanoid metabolism and the discovery of the phenolamides branch in the phenylpropanoid pathway.

Initial questions were:

- **What are the preferred substrates of the different CYP98As tobacco isoforms?**
- **What are the functions of these different isoforms?**
- **Which CYP98As are involved in the synthesis of lignin monomers and the formation of chlorogenic acid?**
- **Tobacco being an allotetraploid plant, are there redundancies between these isoforms?**

A new question emerged:

- **Did some of the CYP98As paralogues evolved to preferentially or more efficiently metabolize phenolamides?**

To answer these questions different approaches have been undertaken. The expression patterns of the different *CYP98As* were studied in wild type tobacco and in response to elicitation, wounding, inoculation with Tobacco Mosaic Virus, and several others stresses. Metabolic profiling of wild type tobacco tissues was carried out for correlation with expression patterns of the *CYP98As*. The different *CYP98As* were heterologously expressed in yeast for *in vitro* analysis of substrate specificity. Finally, stable or transient over-expressors were obtained and analysed.

2. Results

2.1. Phylogenetic analysis of the CYP98A family

Since the cloning of *CYP98A3* from *Arabidopsis thaliana* and the characterization of this enzyme as a 4-coumaroyl shikimate/quinate 3'-hydroxylase (C3'H), a large number of sequences coding for CYP98As have been reported. Most of the CYP98A functionally characterized were typical hydroxycinnamoyl ester hydroxylases, but both CYP98A8 and CYP98A9 were shown to have lost this function (Matsuno *et al.,* 2009). In phylogenetic reconstructions based on full length amino acid sequences (Figure 79), CYP98A8 and CYP98A9 form a clade at the base of all CYP98As, when more divergent P450s from the CYP73, CYP75, and CYP83 families were used as an outgroup. They were separated from characterized C3'H by members from conifers and mosses, indicating a large phylogenetic divergence. The CYP98A8/A9 clade appeared to be specific to the Brassicaceae since no orthologs were found in other plants with fully sequenced genomes, and only one orthologue from *Brassica napus* was found. *CYP98A8*/*A9* have diverged from *bona fide* C3'H and have acquired novel Brassicaceae specific functions specialized in the phenolamides pathway.

Another important information is emerging from phylogenetic reconstitutions. While the CYP98As clusters, as expected, could be classified according to the evolutionary history of plant taxa, they are also distributed in two clades within monocots and dicots (Figure 79).

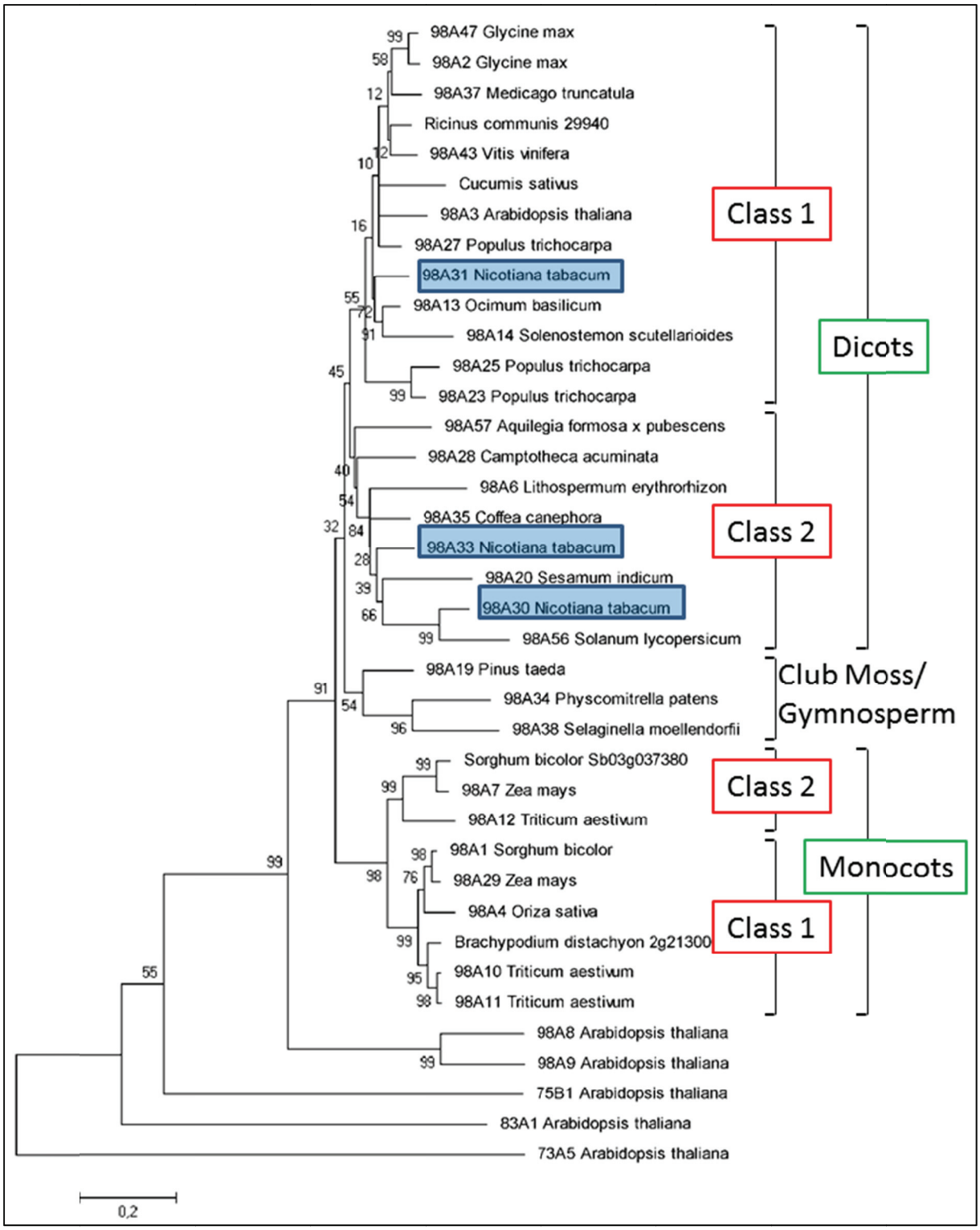


Figure 79. Molecular phylogenetic analysis of the CYP98A family by the Maximum Likelihood method. Sequences from *Nicotiana tabacum* were highlighted in blue.

It thus appears that two independent but early events of *CYP98A* duplication occurred in both monocots and dicots. In each case, duplication occurred and was stabilized in the plant genomes, which very strongly suggests important and independent roles with most likely subfunctionalization of the two classes.

While selected members of each clade have been shown to catalyze dual or specific activities in relation with plant-specific metabolites (Gang *et al.*, 2002; Matsuno *et al.*, 2002; Mahesh *et al.*, 2007; Eberle *et al.*, 2009), the reason for the main gene duplication and role of each clade was so far not investigated. Only one study in monocots compared metabolism of *p*-coumaroyl tyramine and *p*-coumaroyl shikimate/quinate by wheat CYP98s in connection with protein structure and phylogeny (Morant *et al.*, 2007). In conclusion to this report, gene duplication in *CYP98A* family has resulted in a metabolic diversification, but the physiological role of each clade was not investigated.

Interestingly, tobacco sequences are found in the two clades from dicots, with *CYP98A31* belonging to the class 1 and *CYP98A30*, *CYP98A33* belonging to the class 2 (Figure 80).

	Organisms	Class 1	Class 2	Others
Dicotyledons	*Nicotiana tabacum*	1	2	-
	Solanum lycopersicum	1	1	-
	Glycine max	2	-	-
	Medicago truncatula	1	-	-
	Arabidopsis thaliana	1	-	2
	Carica papaya	2	-	-
	Populus trichocarpa	3	-	-
	Vitis vinifera	1	-	-
	Cucumis sativus	1	-	-
	Ricinus communis	1	-	1
Monocotyledons	*Sorghum bicolor*	1	1	-
	Zea mays	1	1	-
	Triticum aestivum	2	1	-
	Brachypodium distachyon	1	-	-
	Oryza sativa	1	-	-

Figure 80. Distribution of CYP98As in different clades in fourteen sequenced genomes plus *Nicotiana tabacum* isolated sequences.

Except if *CYP98As* genes are located in condensed DNA, in the roughly sequenced centromere regions, the distribution of the two *CYP98As* in the plant kingdom indicates that class 2 is not found in all plants, contrary to class 1 (Figures 79 and 80). Consequently the existence of the two classes in the same organism could be either due to specialization or redundancy. Apparently, class 2 is not essential, or redundancy/subfunctionalization is limited to some plant families. For example, until now, two *CYP98As* were isolated in *Coffea canephora*: *CYP98A35* and *CYP98A36*. Both coded proteins metabolize *p*-coumaroyl shikimate, but only CYP98A35 metabolizes *p*-coumaroyl quinate (with unusual efficiency), favoring the production of chlorogenic acid (Mahesh *et al.*, 2007). CYP98A36 belongs to class 1 and CYP98A35 to class 2. It is thus possible that class 2 CYP98As evolved to produce more taxa-specific phenolic derivatives, such as chlorogenic acid or rosmarinic acid.

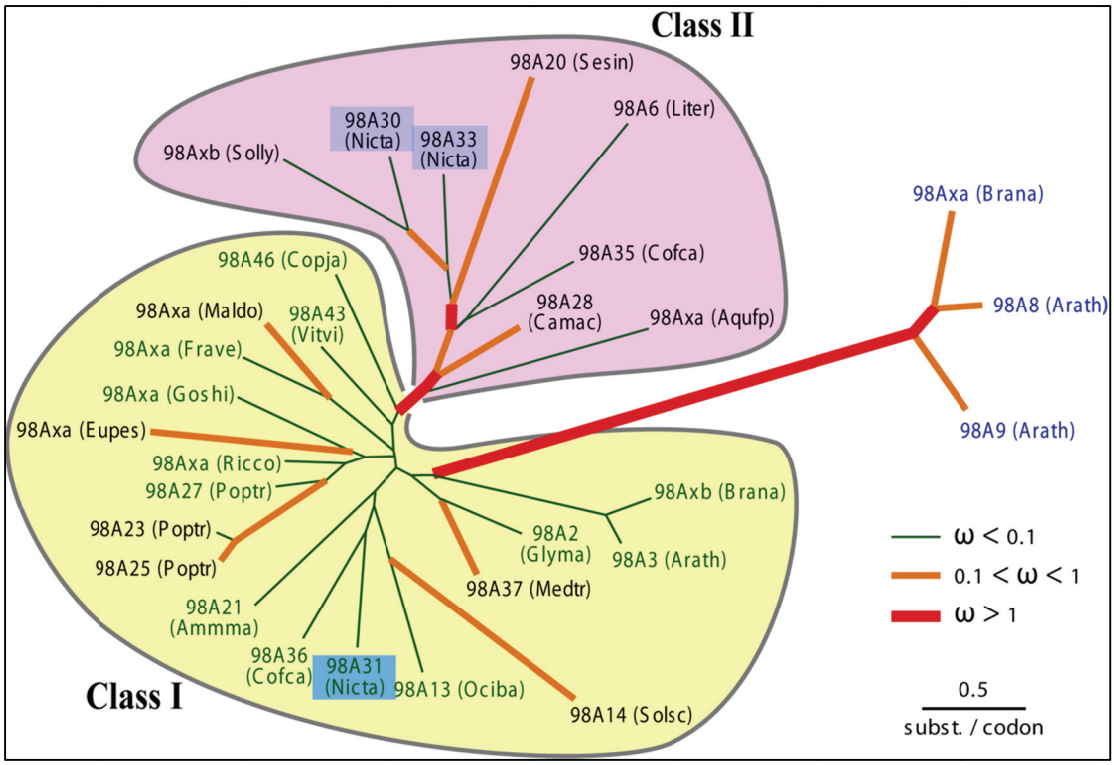


Figure 81. Estimation of synonymous over non-synonymous substitution rates across dicotyledon CYP98As.
A maximum likelihood tree (100 replicates) was performed using the dnaML program from the Phylip package. The branch site model A, implemented in codeml of the paml package was used to estimate the substitution rates per codon. Branch lengths were drawn to scale the substitutions per codon (ω ratios, which are indicative of positive and purifying selection, are greater than one or close to zero, respectively). Species abbreviations: *Ammi majus* (Ammma), *Aquilegia formosa x pubescens* (Aqufp), *Arabidopsis thaliana* (Arath), *Brassica napus* (Brana), *Camptotheca acuminata* (Camac), *Coffea canephora* (Cofca), *Coptis japonica* (Copja), *Euphorbia esula* (Eupes), *Fragaria vesca* (Frave), *Glycine max* (Glyma), *Gossypium hirsutum* (Goshi), *Lithospermum erythrorhizon* (Liter), *Malus domestica* (Maldo), *Medicago truncatula* (Medtr), *Nicotiana tabacum* (Nicta), *Ocimum basilicum* (Ociba), *Populus trichocarpa* (Poptr), *Ricinus communis* (Ricco), *Sesamum indicum* (Sesin), *Solanum lycopersicum* (Solly), *Solenostemon scutellarioides* (Solsc). *Nicotiana tabacum* sequences were highlighted in blue. Analyses were carried out in the laboratory by Jürgen Ehlting.

The potential significance of the gene duplication in tobacco was tested in the laboratory by Jürgen Ehlting, by estimation of synonymous over non-synonymous substitution rates (Figure 81). Results indicated that accelerated evolution (ω>1) and purifying selection (ω≈1) led to the functionalization of the CYP98A8/CYP98A9 clade (Matsuno *et al.*, 2009). Furthermore, two other fast evolution steps followed by purifying selection occurred, leading to the branch represented by tobacco, tomato and sesame class 2 sequences. This could indicate acquisition of new function.

Thus, the existence of two *CYP98As* clades with a branch resulting from fast evolution and purifying selection in Solanaceae, led us to hypothesize that *CYP98A33* and *CYP98A30* in *Nicotiana tabacum* could have acquired specific function, possibly for the synthesis of chlorogenic acid that is accumulated in large amounts in Solanaceae. If so, the tissue-specific expression patterns of these two genes (and of HQT) should be correlated to the pattern of accumulation of CGA.

2.2. Tissue-specific *CYP98As* expression pattern in wild type tobacco

Co-expression analysis is usually a good method to spot genes involved in a same biological process. Here, co-expression analyses were carried-out by quantitative real-time PCR (RT-qPCR) on forty different tissues. These tissues are representative of all organs of the plant (roots, stems, leaves, flowers) and of specific developemental stages of the flowers (Figure 82). Flower developmental stages (stage 1, 3-4, 7, 10 and 12) were selected and named as described by Koltunow *et al.*, 1990. Stage A corresponds to senescing flowers and stage B to maturing one (Figure 82).

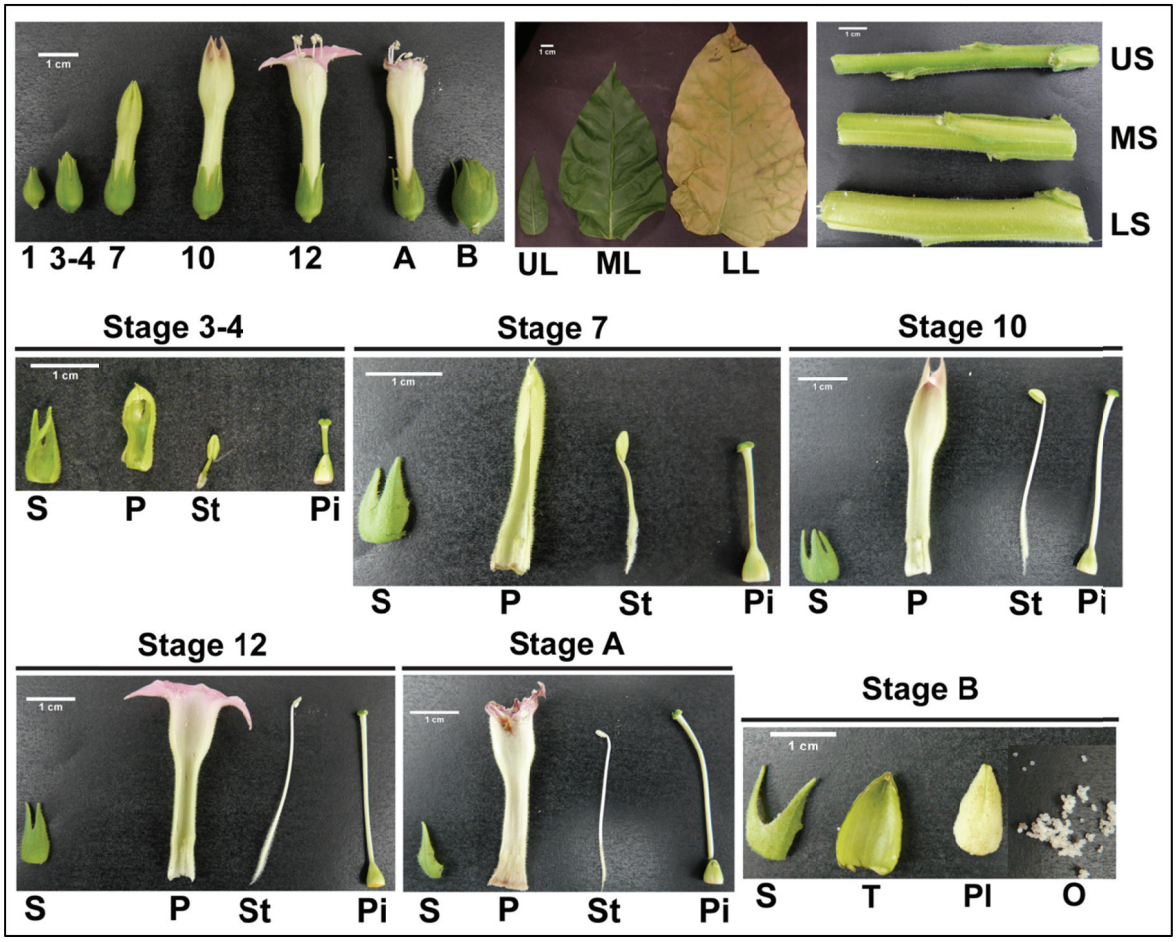


Figure 82. Pictures of analyzed organs.
Abbreviations: 1, stage 1; 3-4: stage 3-4; 7, stage 7; 10, stage 10; 12, stage 12; A, stage A; B, stage B; UL, Upper Leaf; ML, Middle Leaf; LL, Lower Leaf; LS, Lower Stem; MS, Middle Stem; US, Upper Stem; S, sepal; P, Petal; St, Stamen; Pi, Pistil; T, Tegument; Pl, Placenta; O, Ovules. Scale bar: 1 cm.

With the aim to understand the respective functions of CYP98As, their expression were monitored via RT-qPCR in different tissues and compared to the expression of *HCT* and *HQT*, which are more specifically involved in lignin or chlorogenic acid biosynthesis, respectively.

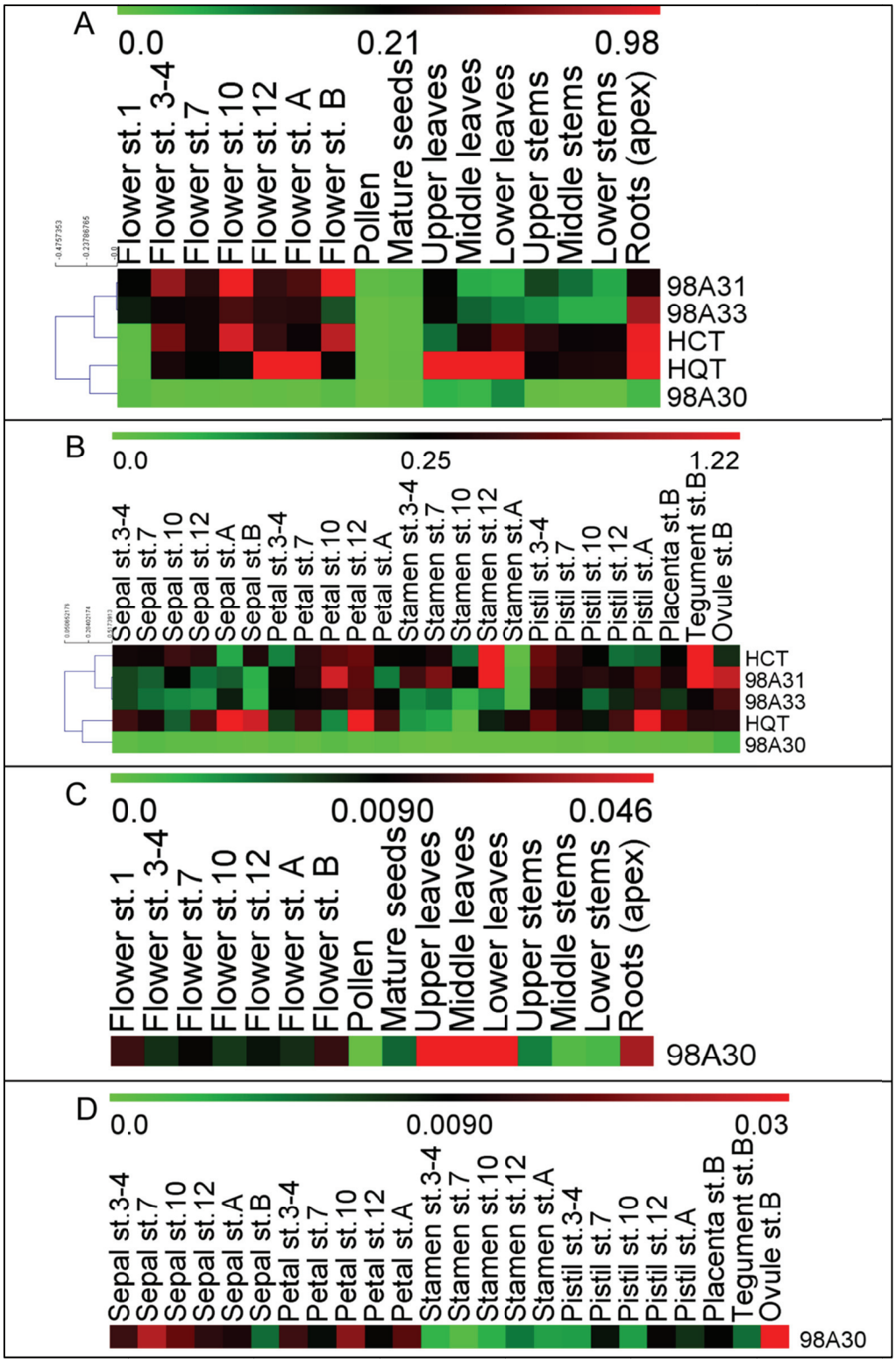


Figure 83. Heatmaps showing *CYP98As* expression in wild type tobacco organs.

A. Organs expression.

B. Flower parts expression.

C. Expression of *CYP98A30* in organs.

D. Expression of *CYP98A30* in flower parts.

RT-qPCR data were analyzed by the δCt method. In C and D, CYP98A30 expression values were extracted from other expression patterns to highlight location of CYP98A30 transcription. Heatmap generation and hierarchical clustering of gene was carried out using MEV software. Experiments were repeated twice, and the results shown here are representative.

δCt analyses were performed and expression values of *CYP98As* were only normalized with values obtained with the internal reference genes. PCR efficiency was 1.92 ± 0.03, so assumed to be similar, allowing direct comparison between genes and organs on the heatmaps.

The first emerging feature was the low expression of *CYP98A30* (class 2) in all organs. The second is the lack of any obvious correlation in the expression of the *CYP98s* with either *HCT* or *HQT* (Figures 83 A, B). However by isolating the *CYP98A30* pattern from others (Figure 83 C, D), *CYP98A30* expression seemed higher in roots, leaves, sepals, petals, and also maturating ovules.

Expression of *CY98A31* (class 1) and *CYP98A33* (class 2) is low in leaves, but higher in roots and in flower organs such as pistils and stamens. *CYP98A31* is remarkably expressed in stamens of stage 12 flowers (mature stamens of the opened flower), in tegument (protecting the maturating ovules) and ovules.

Accordingly, CYP98A31 and CYP98A33 seem to have a specific function in flowers and perhaps different roles due to their differential expression in flower organs. Conversely, CYP98A30 could have a more specific function in roots and leaves, or could be induced in response to certain stimuli.

Hierarchical clustering by Spearman rank correlation analysis, associates *HQT* and *CYP98A30* on one side, and *CYP98A31*, *CYP98A33* and *HCT* on the other side, with a high level of correlation between *CYP98A31* and *CYP98A33*. This suggests that CYP98A30 (class 2) could be involved in CGA biosynthesis. *CYP98A33* (class 2) and *CYP98A31* (class 1) are associated to *HCT* and highly expressed in flower development. This suggests a redundancy or complementary roles of these two CYP98As in the synthesis of lignin, but also in the synthesis of flower specific compounds (lignin being not a major compound in flower tissues).

In other words, the three *CYP98As* of tobacco have distinct expression patterns during flower development and in the different organs of the plants. But the same time, one of the two class 2 *CYP98As* shows a pattern of expression similar to that of the class 1 *CYP98As*. Due to the very low expression level of *CYP98A30* in the wild type plant, it was hypothesized that this P450 could play a role during stress response. To test this hypothesis, different biotic and abiotic stresses were applied to the plants.

2.3. *CYP98As* expression in response to stress

Expression of a large subset of P450s is induced by environmental cues and takes part in specific stresses responses, usually for the synthesis of specific defense compounds. We thus tried to take advantage of differential response to stress to discriminate the function of CYP98As family members.

2.3.1. Biotic stresses

Biotic stress results from bacterial, viral, or fungal pathogens or from damage caused by other living organisms such as herbivores or insects.

Two kinds of stresses were applied to tobacco in this study, in an attempt to elicit different plant responses: elicitation by a fungal elicitor *β*-megaspermin and inoculation of leaves with Tobacco Mosaic Virus (TMV).

The stress responses can be distinguished as early response (immediate or local) and late response (in some cases systemic). Generally, these two types of responses involve different compounds and genes.

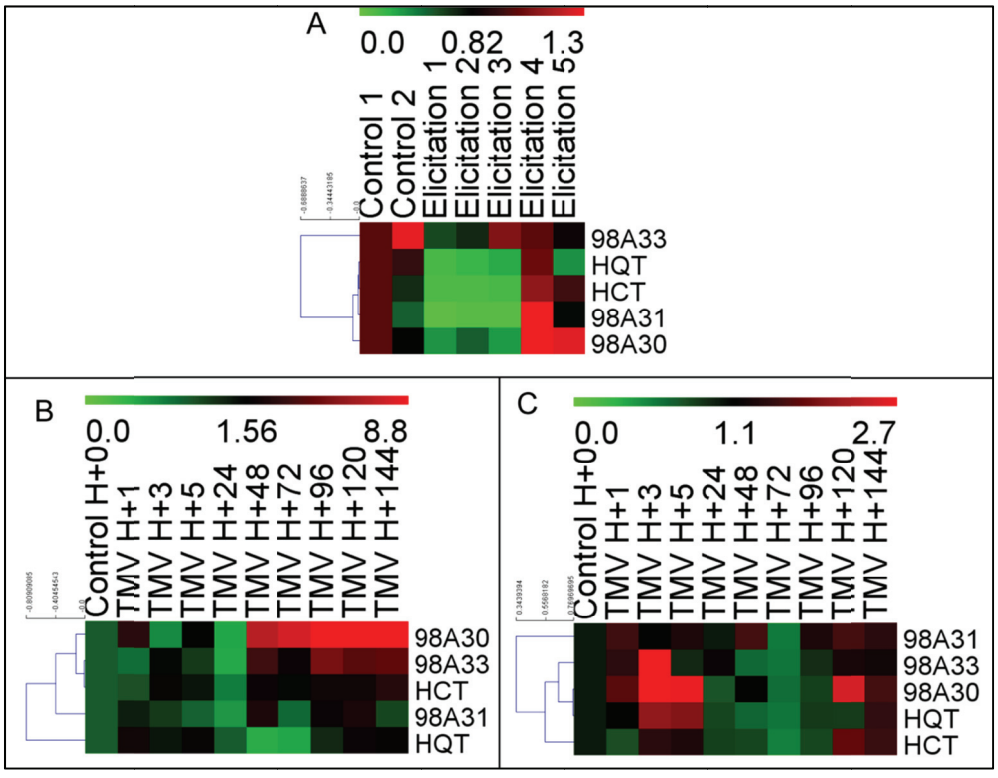


Figure 84. *CYP98As* expression heatmaps of two-month wild type tobacco leaves under biotic stress treatments during 144 hours.

A. Tobacco cell cultures elicited by *β*–megaspermine.

B. Local response of two month-old tobacco plants inoculated by Tobacco Mosaic Virus (TMV).

C. Systemic response of two month-old tobacco plants inoculated with TMV.

RT-qPCR data were analyzed by the δδCt method. VMT experiment was done two times, and the results shown here are representative.

Elicitation with *β*-megaspermin (Figure 84 A) cannot be regarded as significant, since variations between the two controls have the same magnitude as the response triggered by elicitation. **Apparently, *CYP98As* from tobacco were not implied in response to elicitation by *β*-megaspermin.**

Upon TMV inoculation, *CYP98A30* and to some extent also *CYP98A33* expression were induced locally, from 48h after infection with a maximum at 96h, which correlated with moderate response of *HCT* (Figure 84 B). A systemic response was conversely observed very early, with concomitant response of *HQT* with *CYP98A33*, and in particular *CYP98A30* between 3h and 5h post inoculation, that would predict production of CGA (Figure 84 C). A late systemic response at 120h post inoculation involved concerted increase in expression of *HCT* with *CYP98A30*, potentially for cell wall consolidation.

Hierarchical clustering by Spearman rank correlation was carried out, indicating *CYP98A33* expression highly correlated to expression of *HCT* local response to TMV (Figure 84 B). *CYP98A30* expression was clearly associated to *HQT* expression during the systemic response (Figure 84 C).

To conclude, *CYP98A30* that is expressed at very low levels in healthy tissues seems activated upon defense against TMV. It could have a specific role in plant defense.

2.3.2. Environmental and hormonal stresses

Abiotic stress may arise from water excess, drought, low or high temperature, salinity, minerals, but also from low or high light. Most of them cause oxidative stress through formation of reactive oxygen species (ROS) that damage or kill the cell. A typical ROS is H_2O_2, which is involved in signal transduction in abiotic or biotic stress responses.

Among potential mediators of stress response, abscissic acid (ABA), hydrogen peroxide (H_2O_2), methyl jasmonate (MeJA), salicylic acid (SA) and salt (NaCl) were tested in this study. In addition, environnemental cues inducing phenylpropanoids or polyamine biosynthesis, such as wound, drought and cold were also tested. Some of these treatments, such as H_2O_2, wound, SA, MeJA are also relevant to biotic stress.

Interplant variability was observed during preliminary tests, with TMV and with profiling done on wild-type tobacco. This variability interferes with data interpretation, especially when stress only induces slight adjustments of the expression. A test of variability has thus been conducted with four non-stressed plants (Figure 85 A). Variation in gene expression of about two fold was observed in test plants, for all the genes of interest.

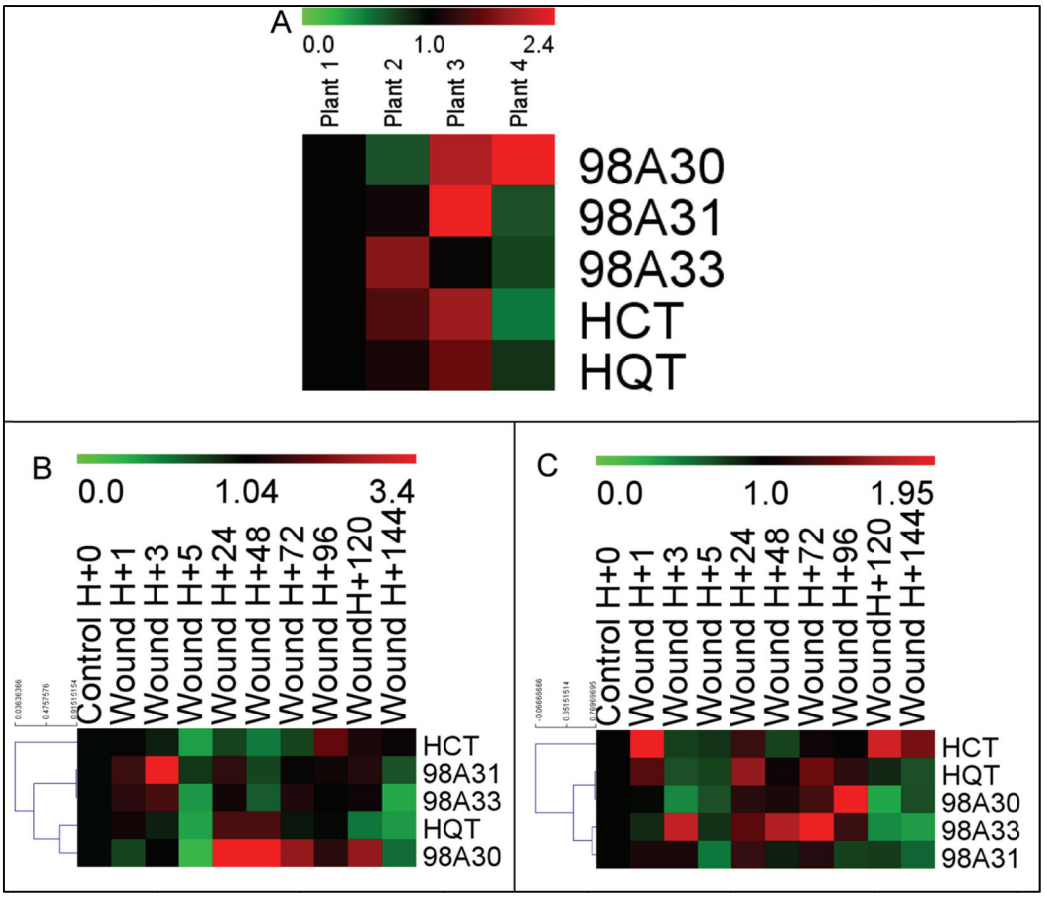


Figure 85. *CYP98As* expression heatmaps of two month-old wild type tobacco leaves.
A. Inter-plants expression variability.
B. Local response to leaf abrasions used for TMV inoculation experiment.
C. Systemic response to leaf abrasions used for TMV inoculation experiment
RT-qPCR data were analyzed by the δδCt method.

During TMV experiments, a wound stress was applied through the TMV inoculation procedure. Leaves were rubbed and thus trichomes were broken. This stress was applied as a control (Figures 85 B, C). A decrease in the expression of all genes of interest was observed five hours after the rubbing (locally and systemically). Locally, *CYP98A31* (class 1) and *CYP98A33* (class 2) appeared slightly induced during early response, and *CYP98A30* (class 2) was clearly induced between 24 and 120 hours after rubbing. In this response, *HQT* and *CYP98A30* expression patterns were clearly associated, as for *CYP98A31* and *CYP98A33*.

The systemic response was not conclusive. Indeed, changes in gene expression did not exceed two fold, which was similar to the inter-plants variability. However, a slight *CYP98A33*, *CYP98A30* and *HQT* response between 24 and 96 hours after rubbing cannot be excluded. Hierarchical clustering once more correlates expressions of *CYP98A30* and *HQT* on one side, and *CYP98A31* and *CYP98A33* on the other side.

A more severe stress was then delivered to the leaves with a needle, so as to generate small holes on the whole leaf surface. Low induction of *HQT* expression was then observed between 5 and 48 hours. Longer-lasting activation of *CYP98A31* and *CYP98A30* was also observed (Figure 86 A). In this experiment, expression of *CYP98A30*, *CYP98A33*, *HQT* and *CYP98A31* appeared correlated.

To further test candidate gene co-expression, their response to a set of other stresses was assessed (Figure 86 B, C). Positive control genes were monitored in order to attest the efficiency of the stresses applied (Figure 86C). Seven control genes were used: gluthation-*S*-transferase class-phi (*GST*), spermidine synthase (*SPDS*), ornithine decarboxylase (*ODC*), arginine decarboxylase isoforms 1 and 2 (*ADC1*, *ADC2*), proline dehydrogenase isoforms 1 and 2 (*PDE1*, *PDE2*). GST is reported to be involved in several biotic and abiotic stress responses, e.g. osmotic and oxidative stress (Dixon *et al.*, 2002). PDE is involved in proline catabolism, thus in osmotic adjustments and redox balance (Szabados & Savouré, 2010). Polyamine biosynthetic genes participate in various defense-related responses: potassium, phosphate, calcium, magnesium, sulfur deficiency, drought, heat stress, chilling, oxidative-related stress… (Alcázar *et al.*, 2010; Takahashi *et al.*, 2010). Apparently all of the treatments were effective: indeed more or less severe expression changes were detected. Notably, high impact of methyl jasmonate (MeJA), salt (NaCl) and cold was observed.

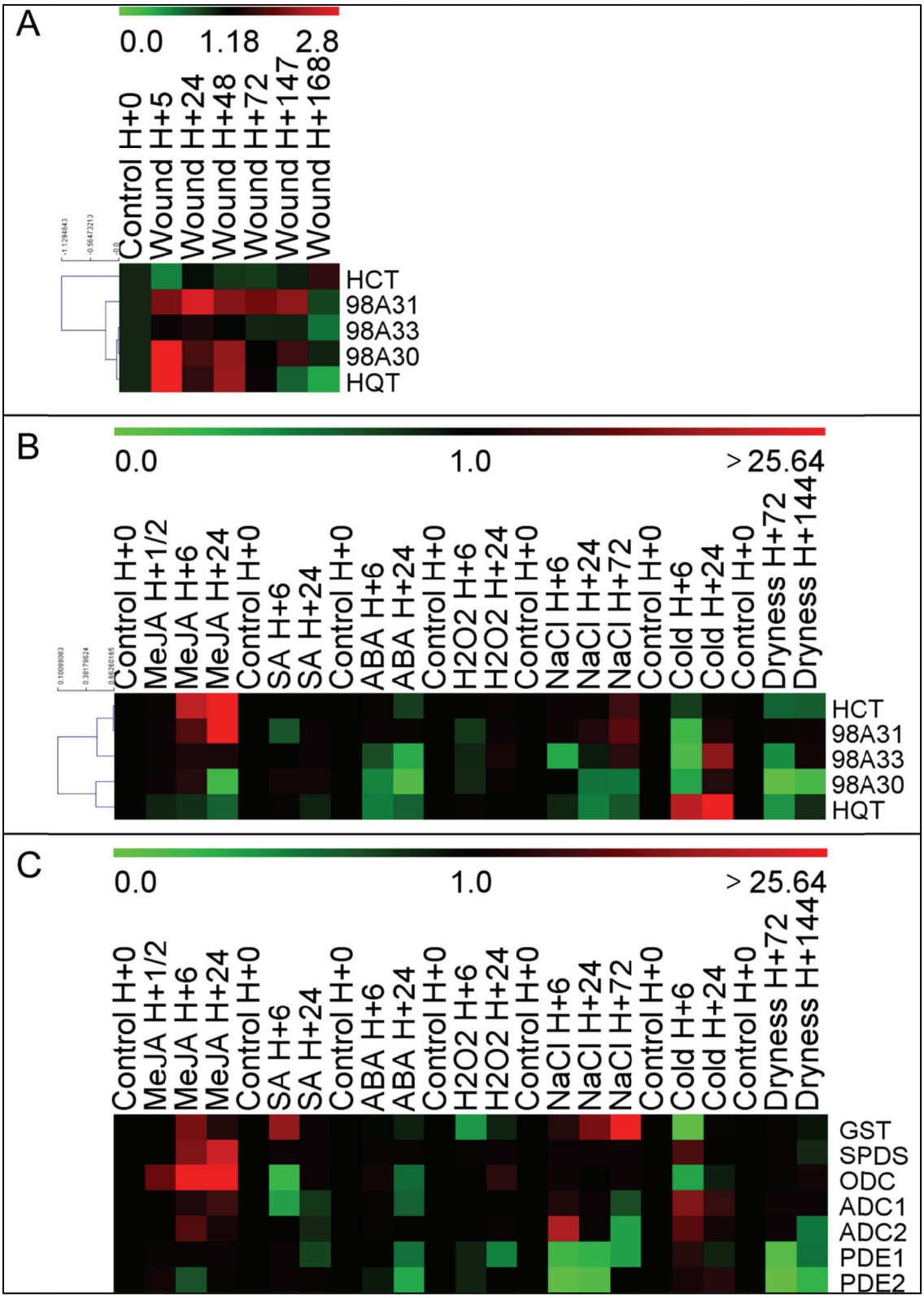


Figure 86. ***CYP98As*** **expression heatmaps of two month-old wild type tobacco leaves treated by various stresses.**

A. Wound stress.

B. Expression of ***CYP98As*** **during various stresses.**

C. Control genes expressions to validate the effects of the stresses.

RT-qPCR data were analyzed by the δδCt method.

GST**, Gluthation-*****S*****-transferase class-phi;** ***SPDS*****, spermidine synthase;** ***ODC*****, ornithine decarboxylase;** ***ADC1*****,** ***ADC2*****, arginine decarboxylase isoforms 1 and 2,** ***PDE1*****,** ***PDE2*****, proline dehydrogenase isoforms 1 and 2.**

As for positive control genes, MeJA, NaCl and cold stress induce the highest response of *CYP98As* and transferases in tobacco. The strongest response (48x) measured in this study was for *CYP98A31* 24 hours after MeJA treatment and for *ODC* (66x), 6 hours after MeJA treatment. *CYP98A30* (class 2) expression was decreased by ABA, NaCl and drought, and increased by SA and MeJA (early response). *CYP98A33* expression (class 2) was activated by MeJA, SA, in addition to a late response to H_2O_2, NaCl and cold. Conversely, its expression decreased in response to ABA. *CYP98A31* expression was induced by MeJA, NaCl, and a late response was detected with ABA, SA and cold. In these experiments, a clear correlation was found for expression of *HCT* and *CYP98A31*. Correlation between *CYP98A30* and *HQT* was lower than in previous experiments.

According to these data, the three members of the *CYP98A* family of tobacco are thus differentially regulated in response to environmental cues.

In conclusion, the expression of *CYP98A30* (class 2), barely detectable in healthy plant tissues, is activated in response to stress. *CYP98A30* expression in several cases correlated with the expression of *HQT*. Thus CYP98A30 is likely to contribute to the biosynthesis of chlorogenic acid. Obviously, its expression pattern differs from those obtained for *CYP98A31* (class 1) (or *CYP98A33*). Expression of *CYP98A31* was most often associated with *HCT* and/or *CYP98A33*. Accordingly, CYP98A31 is expected to function in lignin biosynthesis. Finally, CYP98A33 (class 2) expression seems more versatile and sometimes redundant with the two other enzymes. The *in planta* co-expressions of *CYP98A30* and *HQT*, and of *CYP98A31* and *HCT*, suggest that CYP98A30 could preferentially metabolize *p*-coumaroyl quinate to form chlorogenic acid, while CYP98A31 could generate monolignols from *p*-coumaroyl shikimate.

2.4. Yeast heterologous expression of the three tobacco CYP98s

To test the hypothesis formulated above, *CYP98As* coding sequences were expressed in yeast and their enzymatic activities assayed over a range of potential substrates.

Coding sequences of *CYP98A30* and *CYP98A31* respectively representative of the class 2 and 1, were previously isolated and cloned in the pYeDP60 vector by Rachel Million-Rousseau (PhD Thesis, 2006). The *CYP98A33* sequence was identified in databanks and then isolated by RT-PCR from healthy leaf-extracted mRNA, and finally cloned into the pYeDP60 vector suitable for yeast transformation and inducible co-expression of the P450s with the P450-reductase ATR1 from *A. thaliana* in the WAT11 yeast strain. Microsomal membrane fraction was isolated from galactose-induced yeast cultures. Cytochrome P450 expression was evaluated in microsomes by differential UV-visible spectrophotometry, before *in vitro* enzyme assays.

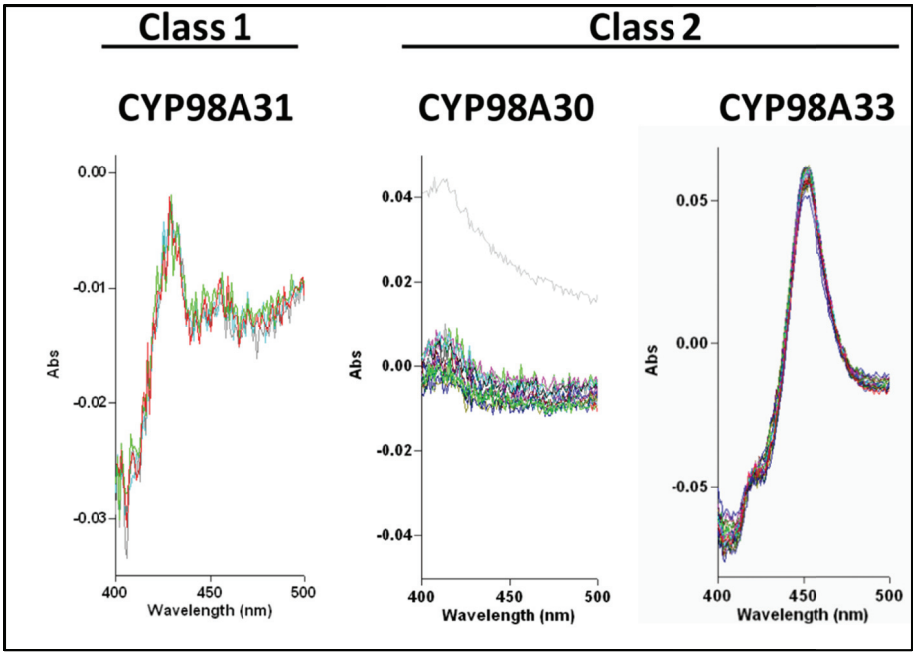


Figure 87. CYP98As differential CO spectra obtained after preparation of microsomes from transformed WAT11 strain and expression induction at 20°C.

Abs, Absorbance.

Fairly good levels of expression of CYP98A33 were obtained in yeast. Its concentration in yeast microsomes was estimated around 4 µM by differential UV-vis spectroscopy (Figure 87). Expression of CYP98A30 and CYP98A31 was however very low. In the best case, a minor peak at 450 nm and a larger at 420 nm were detected for CYP98A31 (Figure 87). Observation of a peak at 420 nm instead of 450 nm indicates troubles in protein folding, or a loosening of the heme bond to the apoprotein, and therefore perhaps a partly inactive protein. The expression of CYP98A30 was not detectable. Low expressions of CYP98A30 and CYP98A31 in yeast were already reported by Rachel Million-Rousseau. She implemented several strategies to optimize their production, but without success. I initiated new trials to improve tobacco CYP98s expression, since reliable comparison in catalytic activities requires active CYPs and measurable P450 concentrations.

Several approaches were tested:

- A decrease in induction temperature from 24°C to 20°C was tested, in order to decrease the expression rate and to improve the folding of the proteins, but without success (results not shown).

- A complete recoding of the cDNA. It has been previously observed in the laboratory, that in some cases a strong accumulation of the transgene messengers after induction by galactose does not result in P450 protein synthesis. The efficiency of message translation depends on the cellular tRNAs concentration, and it is clearly established that the relative concentrations of specific tRNAs vary considerably from one organism to another, and differ between plant and yeast. The cDNA was partially recoded to match yeast codon usage without modifying the peptidic sequence and it significantly improved the P450s expression (Batard *et al.*, 2000). Accordingly, the sequence of *CYP98A30* was completely recoded by full synthesis, so as to match yeast codon preference. Despite codon optimization, no improved expression was observed (results not shown).
- Finally, expression of CYP98A30/31 was tested in the WAT21 yeast strain (expressing ATR2 instead of ATR1). But no improvement in P450 expression was obtained.

To conclude, no improved expression in yeast could be obtained. The enzyme assays were thus carried out with available microsomal preparations that are described below.

Tests were carried out to understand where the expression problem was. Results obtained during these tests showed that (Figure 88):

- Yeast culture grew normally, in similar way for the three P450s. So, CYP98A30 does not seem toxic to cells.
- Plasmid was detected in yeast before preparation of microsomes in similar amounts for each CYP98, which indicates that selection of transformed yeast strain was effective.
- RT-qPCR experiment indicated a higher transcription of *CYP98A30* than *CYP98A31* or *CYP98A33*. In fact, an inverse correlation was found between P450 gene transcription and final protein product.
- Concentrations of microsomal proteins were similar between the different preparations. Final microsomal proteins were higher for CYP98A30 preparation.

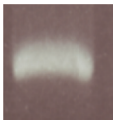
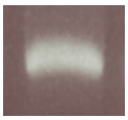
	Class 1	Class 2	
	98A31	98A30	98A33
OD of culture before microsomes preparation	5.8±0.4	5.7±0.3	5.5±0.3
PCR on expression vector (22 cycles)			
RT-qPCR (δδCt method – fold change in expression)	2±0.5	5±0.1	1±0.3
Total proteins in microsomes ($mg.mL^{-1}$)	42±7	42±5	44±5
Approx. total microsomal protein per liter of culture ($mg.L^{-1}$)	230	270	240

Figure 88. Follow up of CYP98As cultures and microsomes productions.
OD measurements and RT-qPCR were preformed on three independent Erlens of cultures per construction. *PDA1* and *ADH6* gene expressions were used as control. Samplings were done before microsomes preparation. Total proteins concentrations were quantified in three microsomes aliquots.

These data indicate a very efficient transcription of *CYP98A30*, even higher than that of other *CYP98s*, which led to production of high levels of microsomal proteins, without detection of 450 nm peak by differential spectrophotometry. It appears that troubles could be either due to an uneffective translation stop, a too fast expression leading to protein mis-folding, or to an intrinsic instability of the protein in yeast (ubiquitination or protease recognition site?). No antibodies were available to detect these P450s, so the hypothesis of a translation stop failure could not be tested.

In animal and yeast, cytochrome P450 overexpression was shown to result in a striking proliferation of the smooth endoplasmic reticulum. The signals mediating this process are not known, but they probably engage signal transduction pathways involved in the unfolded protein response (UPR) or the ER overload response (EOR). Differential effects of P450 expression on apoptosis were observed in human cells. UPR activates pathways that can lead to either cell survival or cell death (Szczesna-Skorupa *et al.*, 2004). Profound gene expression changes and cellular disorders can arise from over-expression of P450s. It is thus possible that CYP98A30 translation was repressed to prevent toxicity.

No improvement of the tobacco CYP98s expression could be obtained in yeast. CYP98A31 expression was very low and CYP98A30 not detectable. Therefore determination of P450 concentrations in yeast microsomes was impossible. As a result, quantitative enzyme assays were impossible for CYP98A30/CYP98A31 and only qualitative tests were carried out.

2.5. CYP98A *in vitro* enzyme assays

A large subset of potential substrates was assayed, including *p*-coumaroyl esters and amides (Mastuno *et al.*, 2009; Morant *et al.*, 2007). Substrates of CYP98As reported so far are molecules with a 4-hydroxylated aromatic cycle. This hydroxyl in position 4 serves to bind and stabilize the molecule in the CYP98A active site.

The following compounds were assayed: *p*-coumarate, *p*-coumaroyl tyramine, *p*-coumaroyl agmatine, di-*p*-coumaroyl putrescine, *p*-coumaroyl putrescine, *p*-coumaroyl shikimate, *p*-coumaroyl quinate, di-*p*-coumaroyl spermidine, tri-*p*-coumaroyl spermidine, di-feruloyl putrescine, diferuloyl spermidine, tri-feruloyl spermidine, di-dihydrocoumaroyl spermidine, tri-dihydrocoumaroyl spermidine and *p*-coumaroyl octopamine.

Since only qualitative tests could be performed, non-limiting incubations using long incubation time (30 minutes) and high NADPH concentration were carried out. An identical volume of each microsomal preparation was used in the assays.

A

	Class 1	Class 2	
	98A31	98A30	98A33
p -coumaroyl shikimate	partial metabolization (1 of 3 isomers)	partial metabolization (1 of 3 isomers)	total metabolization (2 of 3 isomers)
p -coumaroyl quinate	partial metabolization (1 of 2 isomers)	-	total metabolization (1 of 2 isomers)
p -coumaroyl tyramine	low metabolization	-	low metabolization
di-*p* -coumaroyl putrescine	low metabolization	-	low metabolization
di-*p* -coumaroyl spermidine	-	-	low metabolization
tri-*p* -coumaroyl spermidine	low metabolization	-	low metabolization
tri-*p*- dihydrocoumaroyl spermidine	low metabolization	-	-
p -coumaroyl octopamine	low metabolization	low metabolization	partial metabolization

B

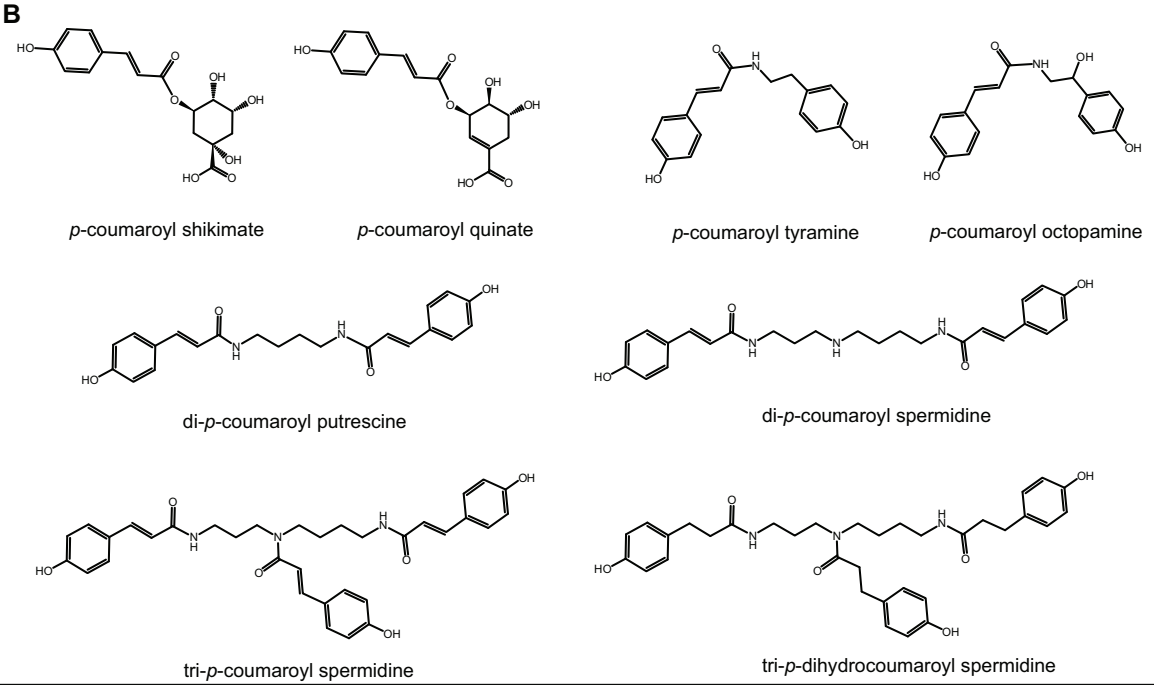


Figure 89. CYP98As *in vitro* metabolization essays.
A. Summarized results. B. Formula of metabolized molecules.

Metabolization was classified in three levels:

- low metabolization, when product was only detected by mass spectrometry single ion monitoring.
- partial metabolization, when product was detectable by mass spectrometry and UV, and a clear decrease in substrates was observable.
- total metabolization, when remaining substrate was not detectable by UV or mass spectrometry.

Our data confirm the low efficiency of tri-*p*-coumaroyl spermidine hydroxylation that was previously reported by Matsuno *et al*. (2009). They further indicate hydroxylation of several other *p*-coumaroyl amide derivatives, never reported before as CYP98s substrates (Figure 89). Those include: *p*-coumaroyl octopamine, di-*p*-coumaroyl putrescine, di-*p*-coumaroyl spermidine and tri-*p*-dihydrocoumaoryl spermidine.

Particularly significant is the conversion of *p*-coumaroyl octopamine. This compound seems to be more efficiently metabolized by the class 2 CYP98As from tobacco than by class 1. Only one hydroxylation product was detected but not further characterized. This hydroxylation could occur either on *p*-coumaroyl or octopamine moieties, which would lead to the formation of caffeoyl octopamine or *p*-coumaroyl noradrenaline, respectively. The feruloyl octopamine derivative of caffeoyl octopamine is ether-linked to the cell-wall of natural and wound potato periderm (Negrel *et al.*, 1996; King & Calhoum, 2005) and seems to constitute major component of the suberin polymer (Graça, 2009). An anti-fungal activity of feruloyl octopamine was reported against *Candida albicans* (*Lee et al.*, 2004) which is consistent with an involvement of this compound in defense.
p-coumaroyl noradrenaline is also involved in defense against pathogen *Pseudomonas syringae* and it was reported to accumulate at high concentrations in tomato (von Roepenack-Lahaye *et al.*, 2003). Interestingly, in phylogenetic reconstructions, CYP98A56 from tomato clusters with CYP98A30 and CYP98A33, at the tip of the branch formed by class 2 CYP98As (Figure 81). This could be indicative of enzyme evolution favouring the synthesis of a taxa-specific defense compound such as *p*-coumaroyl noradrenaline.

CYP98A30 metabolized only two molecules, but preferentially *p*-coumaroyl shikimate, this possibly reflects the preference of this enzyme for *p*-coumaroyl shikimate. Surprisingly, no hydroxylation of *p*-coumaroyl quinate was detected, but a low metabolization of *p*-coumaroyl octopamine was observed.

Besides the more efficient conversion of *p*-coumaroyl octopamine by class 2 enzymes, the only differential feature that is worthy to comment is the apparent better conversion of tri-*p*-dihydrocoumaroyl spermidine by CYP98A31 (class 1) than CYP98A33 (class 2).

Further analyses for characterization of the products are under way.

***In vitro* assays indicate that *p*-coumaroyl shikimate is the preferred substrate of all three tobacco CYP98s. In addition, our data demonstrate that CYP98A31 and CYP98A33 (representative of classes 1 and 2) can also metabolize different di- or tri-substituted putrescine and spermidine conjugates of *p*-coumaric acid, although with very low efficiency. Interestingly, dihydro derivatives which are present at high levels in tobacco are also substrate of CYP98A31 (class 1). Quite surprisingly, no hydroxylation of *p*-coumaroyl quinate by CYP98A30 was detected. This challenges our working hypothesis that proposed CYP98A30 involvement in the synthesis of CGA, although this activity might simply not be detected because of the low expression of the protein.**

Finally, our tests reveal a rather effective hydroxylation of *p*-coumaroyl octopamine by class 2 CYP98 enzymes.

Metabolic profiling was carried out on the different organs of wild-type tobacco with the aim to explain some of these data.

2.6. Correlation between expressions and metabolites patterns

A non-targeted metabolic profiling was carried out by UPLC-MS, on fifteen tissues from wild-type tobacco out of the forty previously analyzed for gene expression. Phenolic conjugates of interest were identified according to their masses and peak areas in the chromatogram were calculated for each compound of interest. Targeted profiling was then carried out focusing on fifteen of the CYP98 substrates and products confirmed by MS/MS analysis: among them, on one hand the substrates di-*p*-coumaroyl putrescine and tri-*p*-coumaroyl spermidine, and on the other hand the products chlorogenic acid (caffeoyl quinic acid) and mono-caffeoyl di-*p*-coumaroyl spermidine.

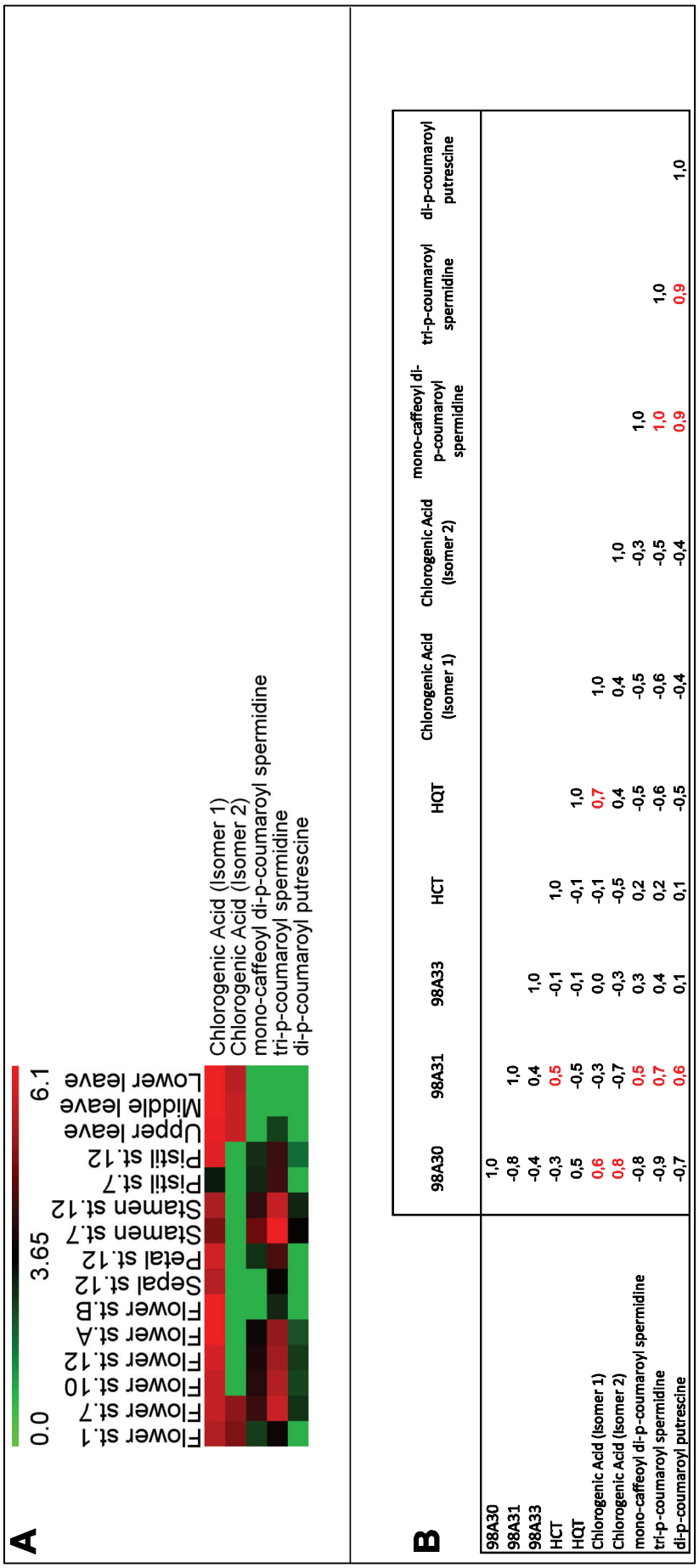
Further analyses are under way to identify CYP98A substrates and products in the non-targeted metabolic profiling.

Figure 90. Transcriptomic and metabolomic correlation tests.

A. Patterns of characterized molecules of interest in organs of wild type tobacco.

B. Spearmann rank correlation matrix of *CYP98As* expression patterns and identified molecules.

Log10 transformation of values collected by mass spectrometry profiling was performed before the heatmap construction and correlation test. Significant correlation coefficients ($p<0.05$) are indicated in red.



B

	98A30	98A31	98A33	HCT	HQT	Chlorogenic Acid (Isomer 1)	Chlorogenic Acid (Isomer 2)	mono-caffeoyl di-p-coumaroyl spermidine	tri-p-coumaroyl spermidine	di-p-coumaroyl putrescine
98A30	1,0									
98A31	-0,8	1,0								
98A33	-0,4	0,4	1,0							
HCT	-0,3	0,5	-0,1	1,0						
HQT	0,5	-0,5	-0,1	-0,1	1,0					
Chlorogenic Acid (Isomer 1)	0,6	-0,3	0,0	-0,1	0,7	1,0				
Chlorogenic Acid (Isomer 2)	0,8	-0,7	-0,3	-0,5	0,4	0,4	1,0			
mono-caffeoyl di-p-coumaroyl spermidine	-0,8	0,5	0,3	0,2	-0,5	-0,5	-0,3	1,0		
tri-p-coumaroyl spermidine	-0,9	0,7	0,4	0,2	-0,6	-0,6	-0,5	1,0	1,0	
di-p-coumaroyl putrescine	-0,7	0,6	0,1	0,1	-0,5	-0,4	-0,4	0,9	0,9	1,0

Two isomers of chlorogenic acid were detected so far, whose characterization by tandem mass spectrometry is planned in a near future (Figure 90 A). One of the two isomers was mainly found in leaves and young buds. The other was evenly distributed, but accumulated in larger amounts in maturating fruit. The targeted phenolamides were mainly present in flower and more precisely in stamen (Figure 90 A). Among these phenolamides, tri-*p*-coumaroyl spermidine was the most abundant.

These five metabolic profiles were tested for match with expression patterns of the three CYP98As and the two hydroxycinnamoyl transferases, by Spearman rank correlation (Figure 90 B). Hypothesizing that metabolites are not translocated in the plants, and that level of transcripts corresponds to the protein expression, *CYP98A30* expression correlates with the accumulation of the two chlorogenic acid isomers, but more specifically with the isomer accumulating only in leaves. Accumulation of the other isomer correlates with the expression of HQT.

Interestingly, tri-*p*-coumaroyl spermidine and its first product of metabolization by CYP98s mono-caffeoyl di-*p*-coumaroyl spermidine also correlate with the expressions of CYP98A31. It is interesting to note that formation of product is related to strong accumulation of substrate, suggesting that this might be a prerequisite for the formation of the product.

In conclusion, even if metabolism of *p*-coumaroyl quinate by CYP98A30 cannot be detected *in vitro*, chlorogenic acid distribution was only correlated with *CYP98*A30 expression. *CYP98A31* expression was clearly correlated with distribution pattern of tri-*p*-coumaroyl/caffeoyl spermidine. No clear correlation was found in the case of *CYP98A33*. To further investigate the respective roles of the different paralogues, over-expressor plants were generated.

2.7. Functional analyses of transient and stable over-expressing plants

Coding sequences of the three *CYP98As* were inserted into pCAMBIA vectors suitable for plant transformation and over-expression under control of the *CaMV-35S* promoter. Transient expressions were carried out in *Nicotiana benthamiana* leaf epidermal cells, and stable transformants were regenerated from transformed *Nicotiana tabacum Xanthi Nc* leaf explants.

The assumption was that an over-expression of each of these P450s could lead to an increase in some downstream metabolites, and perhaps a decrease in the upstream metabolites. Over-expression of each of these CYP98As was checked by semi-qPCR and metabolites were profiled by UPLC-MS/MS. Full scans of metabolites from 100 to 900 Da were carried out in addition to single-ion recording (SIR) of specific molecules of interest, such as chlorogenic acid. Metabolic profiles of stable tobacco transformants were obtained for both leaves and opened flowers.

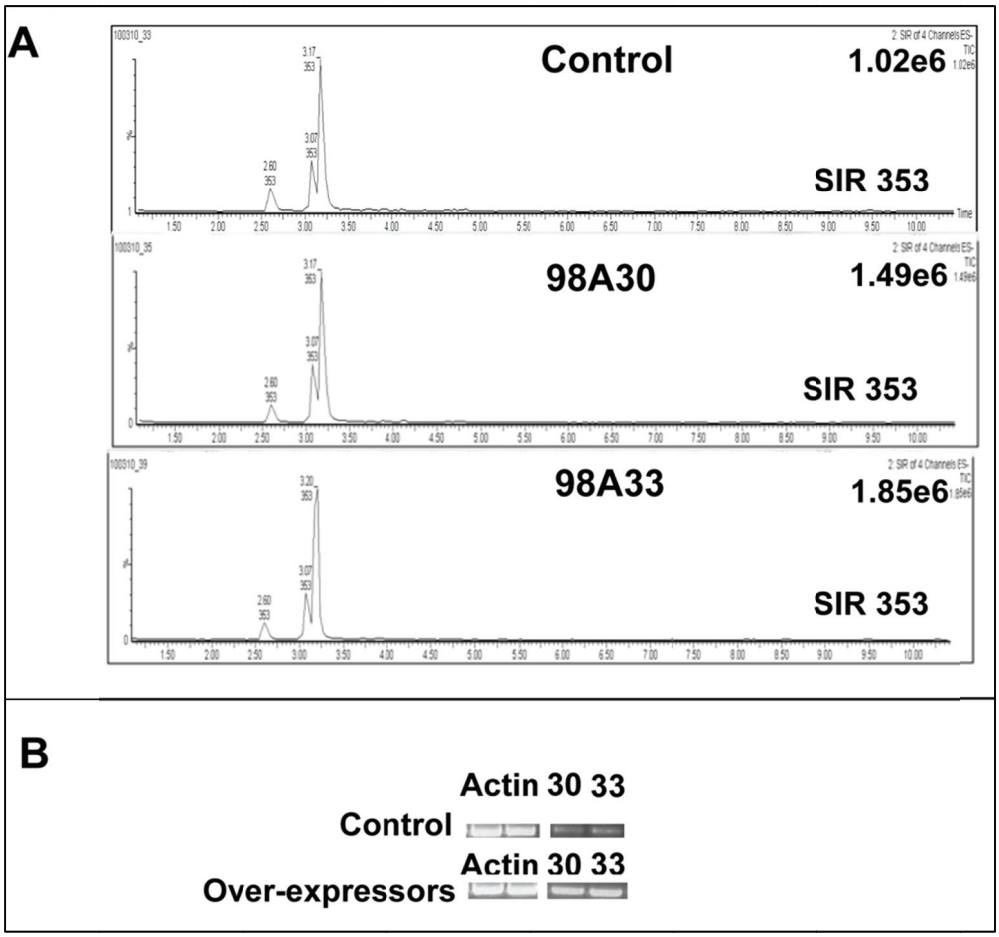


Figure 91. Targeted metabolic study of CYP98As transiently expressed in *Nicotiana benthamiana* leaf epidermal cells.
A. Chromatogram of targeted chlorogenic acid (353M$^-$)
B. sqPCR of control plant expressing eGFP (control) and CYP98A30 (named 30) and CYP98A33 (named 33) transformants. Experiment was repeated twice, and the results shown here are representative.

P450s were clearly stably or transiently over-expressed according to semi-qPCR (Figure 91 B), but no obvious differences were observed in the full scan profiles, neither with the targeted followed metabolites (Figure 91 A, chlorogenic acid targeted). Moreover, no abnormal phenotype was observed in stable over-expressing plants.

This could be due for instance to:

- Tight regulations of the phenylpropanoid pathway preventing toxic accumulation of the target metabolites, or maintaining tight gene co-regulation in the pathway, for example via Post Translational Gene Silencing (PTGS), promoter methylation, extensive degradation of the over-expressed protein, inhibition of protein activity.
- Substrates in limiting amount.
- Storage of target metabolites as glucosides or other conjugates.
- Transport of the metabolites or their incorporation in the cell-wall.
- Unstable and no functional over-expressed protein due to the absence of metabolon partners.

P450 over-expression trouble is not the only possible explanation. Even though infection by disarmed LBA4404 strain was considered mild and did not have any severe effect on the infected leaves and stems, it was still reported to affect phenylpropanoid and flavonoid metabolism (Simoh *et al.*, 2009). The latter are usually induced in the plants in response to biotic and abiotic stresses. Their down-regulation is used by agrobacteria to infect plant. In the case of transient agro-infection it could thus raise problems for the study of phenylpropanoids pathway.

P19 protein was co-expressed with P450s during transient over-expression experiments, which invalidates the PTGS activation hypothesis for this experiment. Accumulation of transcripts also invalidates the hypothesis of promoter methylation. To test the hypothesis of protein degradation by the ubiquitin pathway, an inhibitor of proteasome (MG132) was also tested upon transient expression of P450s, but no improvement was observed (results not shown). Antibodies were not available to control the presence of the protein in the plant. No further analyses were performed for the moment, such as lignin extraction and quantification, glycoside analysis, or addition of substrates during the tests.

To conclude, no over-expression of the tobacco CYP98As seems to occur *in planta*, which could indicate tight translational regulation of the phenylpropanoid pathway, protein degradation or control of enzyme catalytic activities.

3. Discussion

Arabidopsis accumulates sinapoyl derivatives such as sinapoyl malate or sinapoyl choline which have role in defense, adaptation and storage of sinapoyl and choline or malate (Chapple *et al.*, 1992; Landry *et al.*, 1995; Clauss *et al.*, 2008). This plant also evolved a quite divergent pathway for the synthesis of spermidine derivatives that seems to be present only in pollen coat and seeds (Fellenberg *et al.,* 2009; Grienenberger *et al.*, 2009; Luo *et al.,* 2009; Matsuno *et al.,* 2009). But *Arabidopsis* does not accumulate caffeoyl/feruloyl depsides such as chlorogenic acid which are found in many plants. In addition, the class 2 of CYP98 is not represented in *Arabidopsis*, but a recent duplication of the class 1 gene has occurred in Brassicaceae. The accelerated evolution of the resulting *CYP98A8* and *CYP98A9* genes, recently reported by Matsuno *et al*. (2009), together with the loss of a conserved duplication of the CYP73 gene (chapter III of this report), support the hypothesis of a major reorganization of the phenylpropanoid pathway in Brassicaceae. Thus, *Arabidopsis* cannot be regarded as a good model for investigating the phenylpropanoid pathway anymore, now that other plant genomes are available.

Tobacco is representative of the Solanaceae and thus of many plants with economical interest like tomato, potato, eggplant or pepper. Specific phenylpropanoids accumulate in these plants, in particular chlorogenic acid (CGA). In tobacco, three CYP98As and two hydroxycinnamoyl transferases (HCT and HQT) are present. HCT is essential for the synthesis of S and G lignin subunits. HQT was proposed to contribute to the synthesis of CGA in both tobacco and tomato. CGA can result from hydroxylation of *p*-coumaroyl quinate by a C3'H or from caffeoyl CoA transfer to a quinate moiety (Niggeweg *et al.,* 2004). The first route being less energy consuming, it seems to be most favorable for the plant. This implies the involvement of paralogues in the hydroxylation of *p*-coumaroyl quinate and in the hydroxylation of *p*-coumaroyl shikimate (for monolignols synthesis). Here the role of the CYP98A paralogues from tobacco was investigated with focus on compounds that are present in significant amounts in the plant.

The *CYP98As* family is present in all land plants, from moss, fern, Gymnosperms to Angiosperms. CYP98As function as C3'H of the phenolic ring of *p*-coumaric acid esters and involvement in lignin biosynthesis was reported ten years ago (Schoch *et al.*, 2001; Franke *et al.*, 2002 a and b). Then, novel information has emerged suggesting additional complexity in the phenolic metabolism. In the case of CYP98s, **two classes of CYP98As independently evolved in both monocots and dicots (Figure 79)**. CYP98-mediated hydroxylation of a diversity of phenolic conjugates has been recently demonstrated (Mahesh *et al.*, 2007; Morant *et al.*, 2007; Eberle *et al.*, 2009; Fellenberg *et al.,* 2009; Matsuno *et al.*, 2009). These data strongly suggest that one or several alternative, and possibly competing, phenolic pathways may operate in different plants and/or different tissues depending on developmental needs and adaptative responses.

The starting idea of this project was to identify relevant substrates of tobacco CYP98A paralogues to be tested *in vitro* by correlating the expression of the different CYP98As to the metabolic profiles of the major accumulated phenylpropanoids derivatives. But correlations were not obvious. In addition, most of the molecules were not available commercially for reliable identification and/or enzyme assays. Thus, the approach was modified, and *in vitro* metabolism of available analogues of the well documented CYP98 substrates was tested. Molecules previously demonstrated to be metabolized by CYP98s from various plants were used in priority. Additional compounds found in tobacco were synthesized. Even if, in some cases, P450s were poorly expressed in yeast, **all the investigated CYP98A proteins from *N. tabacum* were found to *meta*-hydroxylate, with different efficiencies, the di- and tri-*p*-coumaroyl-conjugated putrescine, spermidine and others aromatic phenolamides**. **Strikingly, our tests revealed a rather effective hydroxylation of *p*-coumaroyl octopamine by class 2 CYP98 enzymes.** Some of these compounds accumulate to high levels in tobacco floral organs, so that, even if phenolamides are not the preferred substrates of tobacco CYP98As when tested *in vitro*, they might be metabolized in particular cellular contexts or on response to adaptative cues.

Expression analyses in this study provided preliminary information on expression patterns of tobacco *CYP98As* from the two different classes. Organ-specific expression and differential stress responses were revealed, indicating that while ***CYP98A30*** **(class 2) could have a role in plant defense, the two other *CYP98As* paralogues have more constitutive functions.** Indeed, **the expression of *CYP98A31* was most often associated with the expression of *HCT* and/or of *CYP98A33*. In consequence, CYP98A31 and CYP98A33 function is expected in lignin biosynthesis.** The expression patterns were **complementary and also sometimes overlapping between the *CYP98As***.

In conclusion, the CYP98s from classe 1 and 2 in tobacco seem to have both redundant and complementary functions, depending on the cell type, developmental stage and environnement. They are capable of hydroxylating a large subset of *p*-coumaroyl-conjugated derivatives.

4. Prospects

New data were obtained about the CYP98A family in *Nicotiana tabacum*. Some of them confirmed previous data obtained by Rachel Million-Rousseau, and others complemented her work. Clearly, *CYP98As* in tobacco present different patterns of expression between organs and response to stress. *CYP98A30* expression barely detectable in plants grown under normal greenhouse conditions was activated in response to tobacco mosaic virus, salicylic acid and methyl jasmonate. Metabolic profiling of the plant tissues and availability of potential susbtrates indicated that phenolamides were most likely susbtrates of CYP98As in the plant. Nevertheless their physiological function(s) are still not fully understood. Different approaches can be proposed and a few ones are considered for a near future.

In situ hybridization experiments are in progress to precise tissue specific expression of the three tobacco CYP98As. Thus, flowers and leaves tissues are already fixed and included in paraplast. More precise expression patterns should bring a new light on the enzyme function(s).

Determination of K_m and V_{max} of CYP98A33 (class 2) with different substrates is possible since this enzyme was correctly expressed in yeast. Due to the bad expression of the other paralogues in yeast, it is still not possible to determine their catalytic parameters. Plant transient expression was thus considered as another possibility to obtain functional enzyme in more suitable cellular context. To this end, the three genes have been transiently expressed in *N. benthamiana* leaf epidermal cells and microsomal membranes were prepared from the leaves. Qualitative enzyme tests will be performed first with CYP98A30. If conclusive, extensive enzyme assays and determination of catalytic parameters will be carried out. *Arabidopsis* P450s expressed in this context were shown to be functional in the project described in the first chapter of this manuscript. These preliminary results seem to be encouraging.

The P450s present in microsomes cannot be properly quantified when no peak at 450 nm is detected by differential spectrophotometry. This precludes comparison of catalytic efficiencies. A solution to this problem can be proposed. It is now possible to estimate relative amounts of a P450 in comparison to a P450 with a known concentration, via label-free peptide mass spectral counting of peptides shared between the two proteins. Furthermore it is also possible to determine accurate concentrations by absolute quantification of the proteins (Langenfeld *et al.*, 2008; Seibert *et al.*, 2009). Therefore, in our case, relative concentration could be obtained for CYP98A30 and CYP98A31 by comparison to CYP98A33. Adjustment of the P450 quantities in the *in vitro* enzyme assays could then be achieved. Consequently, quantitative information obtained with these tests should be more reliable.

Metabolic profiling of wild-type tobacco organs was carried out early in this project, before identification of the most relevant susbtrates in enzyme assays. The different chlorogenic acid isomers were not identified. In addition, octopamine and noradrenaline derivatives were not targeted in the UPLC-MS/MS analysis. New analyses are in progress.

To determine enzyme fonctions in the plant, knock out (KO) or knock down mutants would be required. But, no insertion mutants were available for tobacco. Generation of stable constitutive silenced plants by small hairpin RNAi (shRNAi) were repeatedly attempted in the laboratory with CYP98A30 and CYP98A31, but was unsuccessful. Regeneration of plant from callus (after leaf explants transformation) was not possible. This might be related to the essential role of *CYP98As* in the plant development. A more viable alternative could be inducible or transient silencing. Virus induced gene silencing (VIGS) was considered, but it appeared difficult to handle. Identity of the three CYP98A coding sequences seems too high to independently silence each CYP98A. Shift to another model with open access genome like tomato could be also considered.

Study of protein-protein interaction is an attractive approach in the case of class 1 and 2 CYP98A proteins. Indeed, their divergence could be explained by a preferential coupling with a specific hydroxycinnamoyl transferase, HCT, HQT, or amine transferases. If successful, such an approach could validate the involvement of one paralogue in the synthesis of lignin precursors, of chlorogenic acid or phenolamides.

In conclusion, a better understanding of the functions of *CYP98As* from class 1 and 2 will increase our knowledge of the evolution, complexity and regulation of the phenylpropanoid pathway in plant of high economical interest. It will pave the way for studies of the partition of metabolic fluxes between the different branches of the pathway. This knowledge is essential for the successful of metabolic engineering of the phenylpropanoid pathway in the plants.

5. Materials and methods

Samples of alive materials (plants, yeasts) to analyze by RT-qPCR, PCR, or metabolomic, after taking away, were directly frozen in liquid nitrogen and stored at -80°C before grinding with liquid nitrogen and analysis. If several analyses have to be done for the same sample (mRNA extraction and metabolic extraction), the fine homogeneous powder obtained after crushing was divided into several Eppendorf tubes, one for each experiment to realize.

5.1. Chemicals

Substrates used during this thesis were generally in solution in dimethylsulfoxide (DMSO) or methanol.

p-coumaroyl shikimate and *p*-coumaroyl quinate were prepared by P. Ullmann by 2 enzymatical steps as described previously (Cf. § Substrate preparation for CYP98A3 p.93). *p*-coumarate and caffeate were commercially available from Sigma Aldrich. *p*-coumaroyl putrescine, *p*-coumaroyl tyramine, *p*-coumaroyl agmatine, *p*-coumaroyl octopamine and caffeoyl putrescine were kindly provided by J. Negrel (INRA, Dijon). Tri-*p*-coumaroyl spermidine, di-*p*-coumaroyl spermidine, di-feruloyl putrescine, di-feruloyl spermidine and tri-feruloyl spermidine were chemically synthesized by M. Matsuno (IBMP) and M. Schmitt. Di-dihydrocoumaroyl spermidine and tri-dihydrocoumaroyl spermidine were kindly provided by D. Papaioannou (University of Patras, Greece).

Di-*p*-coumaroyl putrescine was chemically synthesized with gratefull help of M. Schmitt (Faculté de Pharmacie, UDS). The synthesis was done in three steps.

1) Protection of the hydroxyl of the *p*-coumarate by acetylation. To 3 g of *p*-coumarate 135 mg of 4-Dimethylaminopyridine were added and diluted in 10 mL of pyridine. Anhydrous acetic acid was added drop by drop and stirring was continued for one hour (the reaction flask being kept cold by an ice-bath). The aquous phase was acidified to pH2 with HCl 2N. An extraction with 300 mL of Ethyl acetate/Tetrahydrofurane (3:1) was carried out. The organic phase was recovered; the water traces were removed by addition of $NaSO_4$. Then the solution was filtered and evaporated with a rotary evaporator. A whitish powder was obtained; this last was triturated with petroleum ether then filtered.
2) Preparation of chlorure of 4-coumarate. To 2 g of the previous product, 30 mL of dichloromethane were added. Then 2 mL of oxalyte chloride and one drop of dimethylformamide were added drop by drop. Stirring was done during 2 h at -5°C. Then the dichoromethane and the unreacted oxalyte chloride were evaporated with a rotary evaporator. A whitish powder was obtained; this last was triturated with petroleum ether then evaporated.

3) Preparation of di-*p*-coumaroyl putrescine. 3 mmol of putrescine was dissolved in 10 mL of dichloromethane. The solution is cooled at 0°C in an ice box with stirring. Then 7.5 mmol of amine tri-acetate were added. Then 1.5 g of the chloride of 4-coumarate was dissolved in 6 mL of anhydrous dichloromethane. This last solution was added drop by drop, to the first solution. The resulting solution was filtered. Finally, de-acylation of the molecule was done by addition of HCl 3N and acetone, followed by a heating of 10 minutes at 100°C with microwave oven. The solution was cooled; a precipitate appeared and was filtrated on a fritted glass filter, and washed with ether. The powder obtained was characterized as di-coumaroyl putrescine.

During the synthesis, the structures of the product and reaction intermediates were confirmed by ^{1}H- and ^{13}C-NMR analyses and by mass spectrometry.

5.2. Plant materials

Tobacco belongs to the Solanaceae, a large family of dicotyledonous plants. Two cultivars of *Nicotiana tabacum* were used: Xanthi (Nc line) for expressions and metabolic studies and Xanthi (SH6 line) for the preparation of stable transformants.
In addition, *Nicotiana benthamiana* plants were used for transient expression of CYP98As in leaf epidermal cells.

Nicotiana benthamiana and *Nicotiana tabacum* (screened stable transformants or Wild-type plants) were grown on soil in the greenhouse with 16 hours light and 8 hours dark. Day temperatures were approximately 24 ºC, and night temperatures 20 ºC.
Nicotiana tabacum Xanthi SH6 to screen were grown *in vitro* in growth room with 16 hours light and 8 hours dark and temperature of 23.

5.3. Phylogenetic analyses

Phylogenetic analyses were achieved with peptidic sequences. The used sequences were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/) or phytozome (http://www.phytozome.net/). Sequences were treated with the program MEGA5 (Tamura *et al.*, 2007; http://www.megasoftware.net/beta/index.php). Peptidic sequences were aligned with ClustalW algorithm, with the modified standard parameters: PAM Protein Weight Matrix, Multiple alignment gap opening penality set up to 3 and multiple alignment gap extension penality set up to 1.8.

Evolutionary analyses were conducted also in MEGA5. The evolutionary history was inferred by using the Maximum Likelihood method based on the Whelan and Goldman + Freq. model (Whelan & Goldman, 2001). The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) was shown next to the branches. Initial tree for the heuristic search was obtained with Nearest-Neighbor-Interchange method. The final tree was drawn to scale, with branch lengths measured in the number of substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 452 positions in the final dataset.

5.4. Cloning of *CYP98A33* in pYeDP60 and pBD132 expression vectors

CYP98A33 was isolated starting from cDNA synthesized from leaves mRNA of *Nicotiana tabacum Xanthi Nc* and cloned in pYeDP60 and in pBD132 vectors suitable for yeast over-expression and plant over-expression, respectively.

5.4.1. pYeDP60 vector

This vector was already described in the chapter one (Cf. § Yeast expression vector p.85).

5.4.2. pBD132 vector

This vector has a cloning cassette downstream of a CaMV-35S promoter, the whole in a T-DNA, which will be integrated in a random way in the genome of the plant cell. In order to allow the selection of the transformed bacteria and transformed plants, this vector carries two genes of resistance. The first, located out of the T-DNA, allows the selection of the transformed bacteria (*A. tumefacians*, *E. coli*); the second, located in the T-DNA, allows the selection of the transformed plants. In both cases, the gene of resistance is *nptII*, coding for an aminoglycoside phosphotransferase conferring resistance to the kanamycin.

5.4.3. PCR for isolation of CYP98A30 cDNA

This PCR was achieved with a high fidelity Taq to avoid mutations, and by primers with 5'-end extensions for addition of restriction sites used during the next sub-cloning.
The used polymerase was the Pfu Turbo Cx (Stratagene) supplemented by Taq polymerase (Invitrogen). Taq polymerase adds T/A extensions to the amplicon, for T/A sub-cloning into TOPO-pCRII vector (or pGEM-T vector). Forward primer was designed with an extension recognized by *Bam*HI restriction enzyme and reverse primer an extension recognized by *Kpn*I. These oligonucleotides were synthesized by Sigma-Aldrich. The obtained amplicons were purified by agarose gel electrophoresis, bands were cut and amplicons were extracted as previously described (Cf. § Extraction and purification of plasmidic DNA p.106). These purified amplicons were then sub-cloned into TOPO-pCRII vectors.

5.4.4. TOPO TA Cloning

This cloning is often used in case of sub-cloning before transfer of the coding sequence of interest in an expression vector. One of the advantages of this vector is that it allows (as the vector pGEM-T) an integration of the PCR product based on the use of the T/A cohesive ends generated during the PCR. Protocol given by the provider was followed for the cloning.

TOPO-pCRII (Invitrogen) (3973pb): This vector allows the fast and direct cloning of PCR products. It is provided on a linearized form with the TOPO TA cloning kit (Invitrogen). A topoisomerase I from *Vaccinia* virus is fixed on vector ends and permit direct ligation of the PCR product which contains cohesive adenine(A)-sorting ends complementary to the thymidine(T)-sorting ends of the vector. The adenine(A)-sorting ends were generated by the Taq DNA polymerase independently of the PCR product by transferase final action of the Taq. This plasmid carries a bacterial replication origin, a selection marker for the screening of the transformants, the *lac* Z' gene and a cloning cassette. The selection marker is the ampicillin resistance gene. The *lac* Z' gene codes for the peptide α of the *β*-galactosidase, wich allow *α*-complementation (selection white/blue). The polylinker cloning site is between opposing promotors SP6 and T7 RNA polymerase promoters.

The PCR products purified on gel were added of 1µL of TOPO-pCRII plasmid and of 1µL of saline solution provided by the manufacturer. Reaction was achieved in 30 minutes at room temperature. Two microliters of the reaction volume was then used to transform competent *E. coli* bacteria for construction propagation. Transformed bacteria were screened by PCR on colony, and sequencing was achieved to verify the cloned sequence.

5.4.5. "Traditional" cloning of CYP98A33 in pYeDP60 and pBD132 vectors

This type of cloning is based on the generation of DNA fragment having cohesive ends generated by restriction enzymes, which allow the integration of the digest fragment in the destination vector having the same cohesive ends.

5.4.5.1. Preparation of the destination vector and of the insert

DNA digestions by a restriction enzyme were carried out according to the indications of the enzyme provider.

To avoid the re-circularization of the vector on itself after digestion and the problems associated to insert orientation, the cloning sites were chosen to achieve directional cloning. In this case two restriction enzymes were used (*Bam*HI and *Kpn*I). Restrictions sites were chosen according to the destination vector, the sequence of interest and to permit double digestion in the same buffer. The optimal combination of restriction enzymes and restriction buffer to use was obtained thanks to the application for double digestion present on Internet sites of the providers: Neb (http://www.neb.com) or Fermentas (www.fermentas.com).

The destination vector and the insert (PCR products or cDNA from other vector) undergo a double digestion by the two suitable restrictions enzymes. A digestion contains 1-2 µg of DNA, restriction buffer 1x and 1 unit of the restriction enzyme in 20 µL; the reaction was incubated one hour at 37°C.
The digested products were then deposited on 1.5% agarose gel (§ Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109) to check the efficient digestion. The fragments of interest were extracted from the gel with the kit Nucleopsin extract II (Macherey-Nagel) and ligation of the insert was done in the destination vector.

5.4.5.2. Ligation

The linearized destination vector and the insert to clone were mixed, in order to obtain an insert / vector ratio equal to 3/1 in a ligation buffer (T4 DNA ligase buffer, Fermentas) to which, one unit of T4 DNA ligase was added (Fermentas).
The reaction was proceeded in Eppendorf tube in a 20 µL final volume at 16°C overnight. Two microliters of this ligation reaction were used to transform bacteria for plasmid propagation. Transformed colony were screened by PCR over colony, as previously described (Cf. § Screening of the clones p.108), extraction and purification of plasmids were done (Cf. § Screening of the clones p.108) and sequencing was achieved to verify the cloned sequence.

5.5. Stable transformation of *Nicotiana tabacum* (adapted from Horsch *et al.*, 1985)

Transformations were done on leaf discs of *in vitro* cultivated tobacco, with the grateful help of M. Schmitz (IBMP).

5.5.1. Preparation of the infection medium

An overnight LB culture (5 mL) of agrobacteria grown with adequate antibiotic(s) was centrifuged for 20 minutes at 5 500 *g* and 4°C. The pellet was washed with a sterile solution of $MgSO_4$ 10 mM, then pelleted again and resuspended in sterile $MgSO_4$ 10 mM in order to have OD_{600}=10. This agrobacteria suspension was then diluted $1/10^e$ with DF medium and used for the infection.

5.5.2. Infection of the leaf discs

From 4-6 weeks-old tobaccos cultivated *in vitro*, leaves explants were cut out and soaked two hours in the infection medium extemporaneously prepared. The leaf discs were then recovered and the excess of agrobacteria was removed on sterile Whatman paper. These explants were then laid out on Petri plates with solid DF medium supplemented by hormones, without bacteriostatic nor selection agents. Plates were sealed with Parafilm and were placed 2-3 days in growth room.

5.5.3. Regeneration and selection of transformants

The explants were transferred on Petri plates with solid DF medium supplemented by hormones, antibiotic and a bacteriostatic agent. The plates were sealed with Parafilm and were put in growth room during 5-6 weeks.

5.5.4. Rooting of transformants and repicking in soil

The seedlings with leaves were transplanted in glass bottles containing the DF medium supplemented by adequate selectable antibiotic, without growth hormones. When the tobacco seedlings developed a sufficient root system (at least one more month), the young plants were transferred to pots and were acclimatized to the greenhouse.

DF medium	Murashige & Skoog powder (M0221 Duchefa) 4.3 $g.L^{-1}$	Vitamins/Glycine stock solution 1% (v/v)	Sucrose 3% (w/v)	Agar 0.8% (w/v)
Vitamins/Glycine stock solution	Glycine 2 $g.L^{-1}$	Nicotinic acid 0.5 $g.L^{-1}$	Thiamine-HCl 0.5 $g.L^{-1}$	Pyridoxine-HCl 0.5 $g.L^{-1}$
DF medium was sterilized without vitamins at 1 bar 20 minutes. Vitamins/Glycine stock solution, hormones stock solution and antibiotics solutions were filter-sterilized and added after autoclaving.				
Supplemented hormones (Naphtalene acetic acid and 6-benzylaminopurine) were added at 0.05 mg. L^{-1} and at 2 mg. L^{-1} respectively. Kanamycine was the antibiotic of selection with the vector used for transformation. Selection was done with 150 $mg.L^{-1}$ of antibiotic. Cefotaxime was used as a bacteriostatic agent at a concentration of 500 $mg.L^{-1}$.				

5.6. Transient expression in *Nicotiana benthamiana* leaf epidermal cells

Agrobacterium LBA4404 hypervirulent strain was grown at 28ºC and shaking of 180 rpm for 24 hours in LB media with adequate antibiotic(s) (kanamycin 50 $mg.L^{-1}$ for vector selection (pBD132) and rifampicin 25 $mg.L^{-1}$ plus gentamycin 50 $mg.L^{-1}$ for *Agrobacterium* selection). Cells were harvested by centrifugation for 10 min at 4500 *g* and 15ºC and then resuspended in MilliQ water. This wash was repeated two times. After the last centrifugation, cells were resuspended in MilliQ water to a final OD_{600} of 0.5. Before infiltration, equal volumes of the *Agrobacterium* strains were mixed with P19 strain carrying a vector for the plant expression of P19. In all experiments, the *Agrobacterium* strain harbouring P19 was added to maximize protein production by suppression of PTGS. Strains mixtures were infiltrated into leaves of 4-6-week-old *N. benthamiana* plants by pressing a 1 mL syringe without metal needle towards the abaxial side of the leaf and slowly injecting the bacterium suspension into the leaf. The plants were grown under greenhouse conditions for five days until sampling and analyses.

When proteasome inhibitor (MG132) was tried, the leaves were cut 6 hours before analysis, and the leaves were immerged during this time in a solution with the inhibitor (100 µM in MES).

5.7. Relative expression quantification

5.7.1. Primers design

Primers were designed by Roche software: LightCycler Probe Design Software version 2. Couples of primers used for RT-qPCR were designed starting from the coding sequences described in the public databanks and according to the following criteria: a size of 20-25 nucleotides, a melting temperature of 60°C and amplicon size of 100-150 nucleotides. These oligonucleotide probes were synthesized by Sigma-Aldrich (Cf. appendix p.315-317 for primers sequences).

In order to obtain an accurate internal normalizer, the expression stabilities of potential reference genes have been compared. Two references were chosen for yeast (*ADH6* and *PDA1*) and for tobacco (*EF2α* and *Act66*).
Probes efficacies were tested by the cDNA dilution method and efficiency of amplification phase. Amplicons were separated on 3% (w/v) agarose gel to check the specificity of the reaction. Probes couples with PCR efficiency near two were used for the experiments.

5.7.2. Complementary deoxyribonucleic acid (cDNA) preparation

For each experiment with plant and yeast materials, sampling was done quickly and taken samples were directly frozen in liquid nitrogen and stored at -80°C.

Extraction and purification of totals RNAs from plants and yeast were done as previously described (Cf. § Extraction and purification of totals RNAs from plants p.147). Briefly, after grinding of the tissues with mortar and pestle under liquid nitrogen or with grinder (Mikro-dismembrator, Sartorius), total RNA were extracted with the Plant total RNA extraction kit (Macherey-Nagel) according to the manufacturer.

Concentrations of total RNA were measured with the spectrophotometer (Biophotometer, Eppendorf) and quality of the RNA were checked by loading RNA on 1.5% (w/v) agarose gel followed by an electrophoresis at 100 V and visualization of band under UV light (as previously described § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109). PCR of 40 cycles on reference genes were done to check the absence of contaminant DNA (PCR program in appendix p. 312).

Finally, cDNAs were synthesized by retro-transcriptase (RT-PCR) as previously described (Cf. § Reverse-transcriptase polymerase chain reaction p.148). The cDNAs thus obtained were preserved at -30°C before future use by RT-qPCR or sqPCR. Prepared cDNAs of one experiment were then tested by PCR in the condition of sqPCR with the probes of the two housekeeping genes, to check the quality of the cDNA.

5.7.3 RT-qPCR

After preparation of cDNAs, the relative expression of genes of interest can be analyzed by quantitative real-time PCR. The principle of these experiments is based on the comparison of expression of genes of interest in various samples. These values of gene expression are standardized using the values obtained with the expression of two constitutively expressed genes, named housekeeping genes. For yeast experiments, the genes: alcohol dehydrogenase 6 (*ADH6*) and pyruvate dehydrogenase E1 α-subunit (*PDA1*) were used as reference, and for tobacco experiments, the two genes: elongation factor subunit 1α (*EF1α*) and actin 66 (*Act66*) were used.

The quantification for each sample is repeated three times. RT-qPCR were done on Biorad (iCycler) or Roche LC-480 machine. Major part of the experiments was done with the Roche machine and plates were prepared with robot Biomek 3000. RT-qPCR were done by the relative quantification method and SYBR Green was used to monitor dsDNA synthesis.

For each quantitative reaction done with the Biorad machine: 5 µL of a five fold diluted cDNA was mixed with 0,25 µM (final) primers, and 1x SYBR Green mix (Eurogentec) in MilliQ water for a final volume of 20 µL.
For each quantitative reaction done with the Roche machine: 1 µL of cDNA was mixed with primers at 0.25 µM (final concentration) and 1x SYBR Green mix (SYBR Green PCR MasterMix Roche) in MilliQ water for a final volume of 10 µL.

SYBR Green mixes were master mixes containing the buffer of reaction, the SYBR Green and the polymerase: Hot Start AmpliTaq Gold. The PCR was initiated by a "hot start", in order to avoid aspecific amplifications. Amplicons were produced with a program of forty cycles of PCR, with two steps: a denaturation at 95°C, then a step at 60°C for hybridization of the primers and polymerization.

Reliability of the reactions was checked by verification of the amplification curve and the melting curve for each reaction. Furthermore, PCR efficiency was estimated for each reaction with LinRegPCR software (**http://LinRegPCR.HFRC.nl**; Ruijter *et al.*, 2009), which uses absolute fluorescence data captured during the exponential phase of amplification to retrieve PCR efficiency, for each reaction. Efficiency was of 1.92±0.03 with all the various used primers and the whole experiments. Efficiency values were taken into account in all subsequent calculations of relative expression rate of target gene transcripts.

Amplicons were analyzed by the comparative threshold cycle (Ct) method (Livak & Schmittgen, 2001; Schmittgen & Livak, 2008), by δδCt or δCt methods.
By δCt method, Ct values of the gene of interest were normalized to Ct values obtained with reference genes. Thus: δCt sample = Ct gene of interest – Ct geometric average of reference genes.
For expression in comparison to a calibrator (stress experiments for example), δδCt method was employed. Equations for the calculations are:

- δδCt = δCt sample - δCt calibrator. The calibrator being the control of the experiment.
- Expression fold change due to treatment: $FC= 2^{-\delta\delta Ct}$. (if last value is below 1, conversion was done: Expression fold change = -1/ FC).

5.7.4 Semi-quantitative PCR (sqPCR) experiments

After preparation of cDNA, the relative expression of genes of interest can be analyzed by sqPCR. Primers designed for RT-qPCR were used in this case for sqPCR.
The principle of these experiments is based on the visual comparison of amplicons separated on 3% (w/v) agarose gel and visualized under UV light. Amplicons were standardized by comparison to a housekeeping gene. In this case, elongation factor subunit 1α (*EF1α*) was used as reference.

For eash reaction done, a "regular" PCR mix was prepared (PCR program in appendix p.312), and one microliter of prepared cDNA was added to the mix. Amplicons were produced with a program of 22-25 cycles of PCR. The number of cycle was determined after condition optimizations; it depends to the level of expression of the studied gene. The PCR program consisted to: a first step at 95°C during 2 minutes to separate the DNA strands, and then followed by the cycles composed by a step at 95°C for 30 s, a step at 60°C during 1 minute (for hybridization), a step at 72°C of 1 minutes (for elongation). After the cycles, an extra-elongation at 72°C during 10 minutes was done.

After the PCR, amplicons were deposited on a 3% (w/v) agarose gel (containing ethidium bromide at 0.5 $\mu g.mL^{-1}$) and separated by an electrophoresis at 100 V. Bands were photographed under UV illumination.

5.8 Stress treatments experiments

Stresses experiments were done on two month-old *Nicotiana tabacum* Xanthi Nc plants grown in growth room.

5.8.1. Tobacco Mosaic Virus (TMV) inoculation experiments (based on Geoffroy *et al.*, 1990)

Two month-old *Nicotiana tabacum* plants having about ten leaves were used for TMV inoculation. The three lower leaves were used for inoculation, and three upper leaves were used to monitor the systemic response (Figure 92).

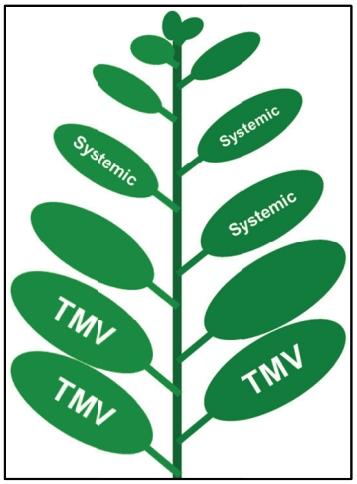


Figure 92. Sampling of TMV inoculated plants.

Inoculations were carried out by rubbing the adaxial part of the leaves with celite (abrasive) moistened by TMV solution (0.4 µg.mL^{-1} of TMV in MilliQ water). Inoculated leaves were washed with MilliQ water and wrung out to remove excess of TMV solution. The inoculated plants were grown in growth room until the sampling. During sampling, three leaves from all the plants were randomly gathered and freezed in liquid nitrogen for each condition (local response and systemic response). Leaves not inoculated by TMV but rubbed, were used as control.

5.8.2 Wound stress experiment

Two month-old *Nicotiana tabacum* plants having about ten leaves were used for wound stress experiment. The three lower leaves were used for wound purpose, and three upper leaves were used to monitor the systemic response (as for TMV experiment Figrure 92). Wound were done by picking the leaves with a little point, to do small holes everywhere on the leaves. The wounded plants were grown in growth room until the sampling. During sampling, three leaves from all the plants were randomly pooled and frozen in liquid nitrogen for each condition (local response and systemic response). Leaves not wounded were used as control.

5.8.3 Abscisic acid (ABA) and salt (NaCl) treatments

Leaves from two month-old *Nicotiana tabacum* plants were detached and immersed during all the kinetic in the solution containing ABA or NaCl. ABA solution at 50 µM in MilliQ water and NaCl solution at 250 mM in MilliQ water were used for the respective treatments. The leaves were let in the greenhouse until the sampling. During sampling, three leaves from all the plants were randomly gathered and frozen in liquid nitrogen. Leaves not treated, immersed in MilliQ water, were used as control.

5.8.4 *Methyl jasmonate (MeJA) treatment* (based on Heitz *et al.*, 1997)

Induction by MeJA was done on two month-old plants placed during the kinetic in sealed Plexiglas boxes of approximately 11 liters, containing MeJA vapors. Cotton wick dipped in 20 µM MeJA solution, before sealing of the box, were deposited near the plants, without direct contact. The elicited plants were grown in greenhouse until the sampling. During sampling, three leaves from all the plants were randomly pooled and frozen in liquid nitrogen. Plants not treated, grown in Plexiglas box during the treatment, were used as control.

5.8.5 *Salicylic acid (SA) and hydrogen peroxide (H_2O_2) treatments* (SA treatment was based on Hong & Hwang, 2009)

For salicylic acid treatment, 5 mM SA in Tween20 0.02% solution was foliar-sprayed onto plants. H_2O_2 treatments were done by foliar-spraying a 3% (v/v) H_2O_2 solution supplemented by Tween20 0.02%. The treated plants were grown in greenhouse until the sampling. During sampling, three leaves from all the plants were randomly pooled and frozen in liquid nitrogen. Plants sprayed with Tween 0.02 % solution, grown in the growth room during the treatment, were used as control.

5.8.6 *Cold treatment*

Two month-old *Nicotiana tabacum* plants were grown during the time of the kinetic in cold room (4°C), thus, in the dark. During sampling, three leaves from all the plants were randomly gathered and frozen in liquid nitrogen. Plants not stressed, grown in dark in the greenhouse, were used as control.

5.8.7 *Drought treatment*

Two month-old *Nicotiana tabacum* plants were dried and not watered during the experiment. They were grown in growth room and samplings were done by randomly taking and gathering three leaves from all the plants. Leaves were frozen in liquid nitrogen. Plants normally watered, grown in the greenhouse during the treatment, were used as control.

5.9 Yeast microsomes preparation

Two *Saccharomyces cerevisiae* strains were used to express CYP98As from tobacco: WAT11 and WAT21 strains (already described in § Different used strains p.85). Culture conditions, chemical yeast preparation and yeast transformation were unchanged in comparison to previous description (Cf. § Biological materials: Saccharomyces cerevisiae p.85).

5.9.1. *Proteins expression*

One transformed yeast colony was put for growth in 30 ml of SGI medium (Cf. § Yeast culture conditions p.85) during 24 hours at 28°C under agitation (160 rpm). Then one unit of OD_{600} was transferred in 200 mL of YPGE and was put under agitation (160 rpm) at 28°C.

After 30 h of growth, sterile galactose (20 $g.L^{-1}$ final concentration) was added to induce protein expression. After 17 h of induction at **20°C** the cells were sedimented by 10 min of centrifugation at 7500 *g* and at 4°C. From this point, all steps were done on ice.

YPGE	Yeast extract 10 $g.L^{-1}$	Bactopeptone 10 $g.L^{-1}$	Glucose 5 $g.L^{-1}$
Sterilize medium at 1 bar, 120°C, 20 min before addition of ethanol 3$mL.L^{-1}$ (v/v final)			

5.9.2 Preparation of the microsomes

Microsomes preparation were done as previously described (Cf. § Yeast microsome preparation p.87).

5.10. *In vitro* metabolization assays

These tests were done as previously described (Cf. § *In vitro* activity test p.93). After stop of the reaction, microsomal proteins were precipitated by addition of acetonitrile 40% (v/v) and centrifugation at 13 000 *g* during 10 minutes 4°C. 200 µL of the supernatant were stored at -30°C until analyses. This incubation medium was analysed by UPLC-MS/MS.

5.11. Methanolic extraction of soluble phenols

The tissues were finely crushed in liquid nitrogen. Crushing was continued in cold 100% methanol (500 µL for 100 mg fresh tissues) then centrifuged 5 minutes at 500 *g* and 4°C. The supernatant was collected and the residual pellet was re-extracted with 70% methanol (200 µL for 100 mg of fresh tissues). The supernatant of the first centrifugation and this second mixture were gathered and agitated vigorously during 2 minutes to support the extraction of phenols. The mixture primarily made up of cell walls was clarified by centrifugation during 15 minutes at 13 000 *g* and 4°C. The supernatant was kept and stored at -30°C until analyzes of the soluble parietal phenols by UPLC-MS/MS.

5.12. Ultra Performance/Pressure Liquid Chromatography coupled with tandem Mass Spectrometry (UPLC-MS/MS) analyses

These analyses were carried out by Dimitri Heintz from the Metabolomic platform of IBMP.

The metabolites, from *in vitro* assays or methanolic extractions, were analyzed by UPLC coupled with a mass spectrometer. An Acquity UPLC system (Waters) coupled to a Quattro Premier XE triple Quadrupole MS system (Waters Micromass) was used. The compounds were resolved on an Acquity UPLC BEH C18 column (hydrophobic stationary phase) at a temperature of 25°C by using a flow at 0.45 $mL.min^{-1}$. Volume injected was of 3 µL. The chromatography was carried out using a gradient of acetonitrile from 5% to 100% supplemented of acetic acid 0,1%. Elution of the compounds was followed in UV using a photodiode array detector and eluates of interest were further analyzed by mass spectrometry.

Molecules were ionized by using a source of electrospray ionization (ESI) in positive and negative modes. Nitrogen was used as the nebulizer gas and for desolvation. For all coumpounds, the cone tension was optimized between 15-25 V. For fragmentation, argon was used as the collision gas. The required masses were analyzed in positive or negative mode, by full scans (search of the masses between 100 and 900 Daltons), SIR (Single Ion Recording: research of precise masses), Daughter Scan (search of the ions fragments of the metabolites of interests) and Parent Scan (search of the parent metabolite of ions fragments).

5.13. Data analyses

Heatmaps were generated with the software MEV release 4.6.1 (Multiple Experiment Viewer) (http://www.tm4.org/mev/; Howe *et al.*, 2010). Spierman's rank correlation algorithm implemented in this software was used to draw gene trees.

The software Statistica release 9.1 was used to generate the Spierman's rank correlation matrix of expression patterns and metabolic patterns.

Chapter III

Functional analysis of the *CYP73* family in *Nicotiana tabacum*

« La réalité est comme un visage qui se reflète sur la lame d'un couteau ; ces caractéristiques dépendent de l'angle sous lequel nous l'envisageons. » Maitre Hsing Yun.

1. Introduction

1.1. *Discovery of* Cinnamate 4-Hydroxylase (C4H) activities and *CYP73As*

The cinnamate 4-hydroxylase enzyme activity was first characterized in the Sixties by Russell & Conn (1967). Its cytochrome P450 nature was demonstrated on *Sorghum bicolor* extracts (Potts *et al.*, 1974). Subsequently, many studies were devoted to the induction of C4H activity in plant tissues with the aim to understand the function, to enhance production of secondary metabolites and to isolate the cDNA encoding the C4H enzyme.

It required almost twenty years and the purification of the C4Hs of *Helianthus tuberosus* (Gabriac *et al.*, 1991) and of *Phaseolus aureus* (Mizutani *et al.*, 1993) until isolation of the first *CYP73A* cDNAs were reported simultaneously from different plant sources (Mizutani *et al.*, 1993; Teutsch *et al.*, 1993; Fahrendorf & Dixon, 1993).

Heterologous expression of these genes made it possible to confirm C4H activity (Urban *et al.*, 1994; Pierrel *et al.*, 1994). *CYP73As* were the first P450 genes associated with a known physiological function to be characterized. After that, many *CYP73A* sequences were isolated or annotated based to their homology with the first isolated sequences. Since these first reports, *CYP73A* sequences have been found in all genomes sequenced from moss to dicots and monocots.

1.2. Extensive studies on CYP73As activity

CYP73s are now the best characterized plant P450s from an enzymological point of view. They were initially shown to be highly specific for conversion of cinnamate into *p*-coumarate when supplied with plant precursors. No activity was found with other phenolics such as 4-coumarate, ferulate, phenylalanine, benzoate, 3- or 4-hydroxybenzoate, phenylacetate, 2- or 3-hydroxy phenylacetate, or salicylate (Pierrel *et al.*, 1994; Overkamp *et al.*, 2000). Furthermore, no metabolization was found on structurally unrelated phenolics natural compounds, which are targets of P450 action, such as, fatty acids and terpenoids (Pierrel *et al.*, 1994).

However, CYP73A1 is capable, although with poor efficiency, of *O*-demethylating natural 7-methoxycoumarin herniarin into umbelliferone, and to metabolize the xenobiotics 7-ethoxycoumarin, *p*-chloro-*N*-methylaniline, 2-naphtoic acid and structurally related compounds, and the herbicide chlorotoluron (Pierrel *et al.*, 1994; Schalk *et al.*, 1997a,b). Thus, CYP73As were shown to metabolize cinnamate analogues, i.e. negatively charged small planar compounds (size of about two adjacent aromatic rings), but usually with a reduced catalytic efficiency, except for 2-naphtoic acid.

A recent extensive study essentially confirmed very old studies carried-out with plant-extracted microsomes and showed that CYP73A5 (from *Arabidopsis thaliana*) hydroxylates cinnamic acid analogues with R1-substitutions of various size and polarity and even some substitutions at R2 position (Chen *et al.*, 2007; Figure 93). According to this study, the catalytic efficiency of C4H towards R1 or R2 substituents decreased when the size of substituents increases.

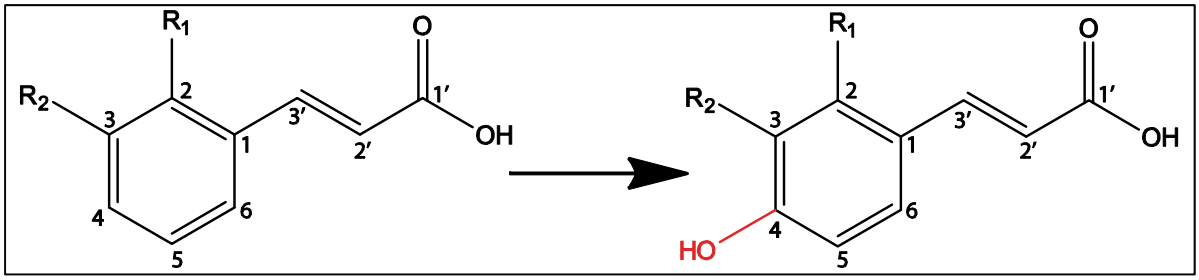


Figure 93. Cinnamate and its analogues hydroxylated by CYP73s.
The 4-hydroxylation position is highlighted in red.

To conclude, all substrates of CYP73A were small planar aromatic molecules. A negative charge (acid) at one extremity is required. Comparisons of the kinetic parameters of CYP73A measured with various substrates showed that cinnamate was the best substrate (Pierrel *et al.*, 1994; Schalk *et al.*, 1997a,b; Chen *et al.*, 2007). This establishes the C4H/CYP73A role in the phenylpropanoid pathway.

1.3. Function of *CYP73As*

Numerous studies have determined C4H expression in different species and C4H expression was detected in most organs analyzed. This expression is predominant in organs with high phenylpropanoid biosynthesis and/or undergoing lignification, such as stems and roots (Mizutani *et al.*, 1997; Urban *et al.*, 1997).

A coordinated expression of *Arabidopsis* C4H with all other genes of the phenylpropanoid pathway is well established (Ehlting *et al.*, 2005). Induction of C4H expression upon stress was also reported in correlation with wounding or with the production of defense compounds (Logemann *et al.*, 1995; Mizutani *et al.*, 1997; Urban *et al.*, 1997; Batard *et al.*, 1997).

To fully understand the role of CYP73As in the plant, mutant analysis was required. Only one copy of the *CYP73* gene is present in the *Arabidopsis* genome (*CYP73A5*; Bell-Lelong *et al.*, 1997). Consequently, a plant homozygous for a null C4H allele was expected to be severely affected. Indeed, analysis of a T-DNA insertion line indicated that the C4H null mutation leads to a seedling lethal phenotype: seedlings germinate, but primary leaves fail to expand (Schilmiller *et al.*, 2009). Only "leaky" C4H alleles were identified in a mutant screen and named *ref3*, for reduced epidermal fluorescence 3.

Analyses of these "leaky" mutants showed reductions in lignin content leading to collapsed vasculature. Furthermore, the *ref3* plants accumulated cinnamoylmalate, which was not found in wild type. Decrease in C4H activity also caused pleiotropic phenotypes, including dwarfism (Schilmiller *et al.*, 2009), and strong *ref3* alleles were affected in pollen development. Sporopollenin, a major structural component of the pollen wall, seemed to be impacted. This compound is a heteropolymer of hydroxycinnamoyl derivatives linked to fatty acids (Guilford *et al.*, 1988; Wehling *et al.*, 1989; Ahlers *et al.*, 1999; Dominguez *et al.*, 1999; Morant *et al.*, 2007). In consequence, C4H activity is necessary for the production of the phenolic moieties of the polymer. C4H defect seems to provoke a deficiency in sporopollenin deposition and leads to male sterility (Schilmiller *et al.*, 2009).

Together, these observations indicate that C4H function is critical for the normal development of *Arabidopsis thaliana*.

In other sequenced plant genomes, however, several *CYP73As* can usually be found. They cluster in two different clades. Thus *Arabidopsis* is not an appropriate model for a representative genetic approach and full-description of *CYP73As* function in crop plants.

1.4. A second class of CYP73As?

Some CYP73s diverge from the other family members by a significant divergence in peptidic sequence (see below). The divergent group of CYP73s was named type/class II C4H in opposition to the first, called type I (or class I) (Nedelkina *et al.*, 1999). This duplication was stabilized in the plant genomes, which strongly suggests functional specialization.

The first class II C4H that has been reported is CYP73A15 from *Phaseolus vulgaris* (Nedelkina *et al.*, 1999). The protein encoded by *CYP73A15* diverges from other CYP73As in its N- and C-terminus and in four internal domains (Nedelkina *et al.*, 1999).
It is characterized as other class II enzymes by:

- A long serine/threonine-rich-N-terminal sequence that suggests plastidial/mitochondrial instead of ER localization.
- A 3-4 amino acid insertion in a region located just upstream of the heme-anchoring cysteine and so called "meander", indicative of a specific property of this class of CYP73s, concerning either its catalytic function, electron donor or other interacting partner protein.
- An isoelectric point lower (8,8) than class I proteins (9,3-10).

In spite of peptidic sequence modification, the class II enzyme has a C4H activity with catalytic constants comparable to class I (Nedelkina *et al.*, 1999).

Antisense and sense expression of cDNA coding for bean CYP73A15 in tobacco, was reported to lead to a slight delayed and reduced production of lignin (Blee *et al.*, 2001). It was then proposed that the class II cinnamate 4-hydroxylase (as class I) might function in lignification in a number of species including French bean and tobacco. Nevertheless, no distinction between the two classes was carried-out and effect on lignin could be attributed to co-suppression of the class I proteins.

Differential expression was also reported for CsC4H1 and CsC4H2 in *C. sinensis* (Betz *et al.*, 2001). CsC4H2 (class I) that was constitutively expressed in orange was suggested to play a role as a 'housekeeping' gene in the phenylpropanoid pathway, whereas CsC4H1 (class II) was expressed only after wounding treatment (Betz *et al.*, 2001). Transcripts for CYP73A15 were shown to be rapidly accumulated after elicitor treatment of suspension-cultured cells of French bean (Nedelkina *et al.*, 1999). A close correlation between the accumulation of PAL and CYP73A15 transcripts was described. CYP73A15 role was reported more to lignification during differentiation than in stress (Nedelkina *et al.*, 1999). Strikingly, accumulation of unspliced transcripts was reported for plant in normal growth conditions. It has never been reported for mRNA coding for class I CYP73As or for other enzymes of the phenylpropanoid pathway. This conditional splicing was hypothesized to be associated with a novel regulatory mechanism, or constituting a stock of available transcripts that can be used for accelerated protein synthesis in response to specific stress (Nedelkina *et al.*, 1999).

Thus, both the evolutionary divergence and the differential expression suggested that different physiological functions were probably acquired in plants which maintained the duplication (Ehlting *et al.*, 2006).

1.5. Questions asked in the CYP73 project

Most of the isolated and well studied CYP73A in plant kingdom belong to the class I. Two class I (*CYP73A47* and *CYP73A8*5) and two class II (*CYP73A27* and *CYP73A28*) CYP73 genes are present in tobacco.

Taking into account previous reports on class II CYP73s, this still raises several questions:

- **Do the CYP73 class I and class II duplications in tobacco result from allotetraploidy?**
- **Are the different paralogues differentially expressed and do they have specific tissue or sub-cellular localizations?**
- **Do the CYP73A paralogues have different substrate preferences?**
- **What are their specific functions in the plant? Are they redundant?**

To answer these questions, I used different approaches. The expression patterns of the different *CYP73As* were studied in wild type tobacco in different tissues during plant development. Their response was also determined to elicitation, wounding inoculation with

TMV, and several abiotic stresses. The different *CYP73As* were heterologously expressed in yeast and *in vitro* enzyme assays were carried out. Homology models were generated in collaboration with Fabrice Klein (IGBMC, Strasbourg) to compare the structural charateristics and emphasize differences between the two classes of *CYP73s*. Finally, stable or transient over-expressors were obtained and analyzed.

2. Results

Four divergent *CYP73A* were identified in *Nicotiana tabacum*. *CYP73A27* and *CYP73A28* were previously isolated from elicitated tobacco plants (Ralston *et al.*, 2001). In addition, two others, *CYP73A47* and *CYP73A85*, were identified in the available EST collections.

2.1. Update phylogenetic analysis of the CYP73 family

Since characterization of the first *CYP73* genes, orthologues were isolated or annotated in numerous species by means of library screening, PCR with degenerate primers, data mining of EST collections or fully sequenced genomes. In total, around 110 *CYP73As* were annotated until now. Less than twenty were identified as class II, and less than fifteen from Gymnosperms, Pteridophytes or moss. The majority of *CYP73A* sequences available to date are highly similar and classified in class I.

In phylogenetic reconstructions based on full-length amino acid sequences the CYP73As form two well-defined clades (Figure 94). The presence of proteins belonging to each class in both mocots and dicots, suggests that the duplication event took place early in plant evolution before monocot/dicot divergence The tree in fact shows branching of class II before emergence of moss and Gymnosperms, but *bona fide* class II sequence is not found in such ancestral land plants. The early branching of class II CYP73s is thus likely to reflect the quite high sequence divergence of class II with class I proteins.

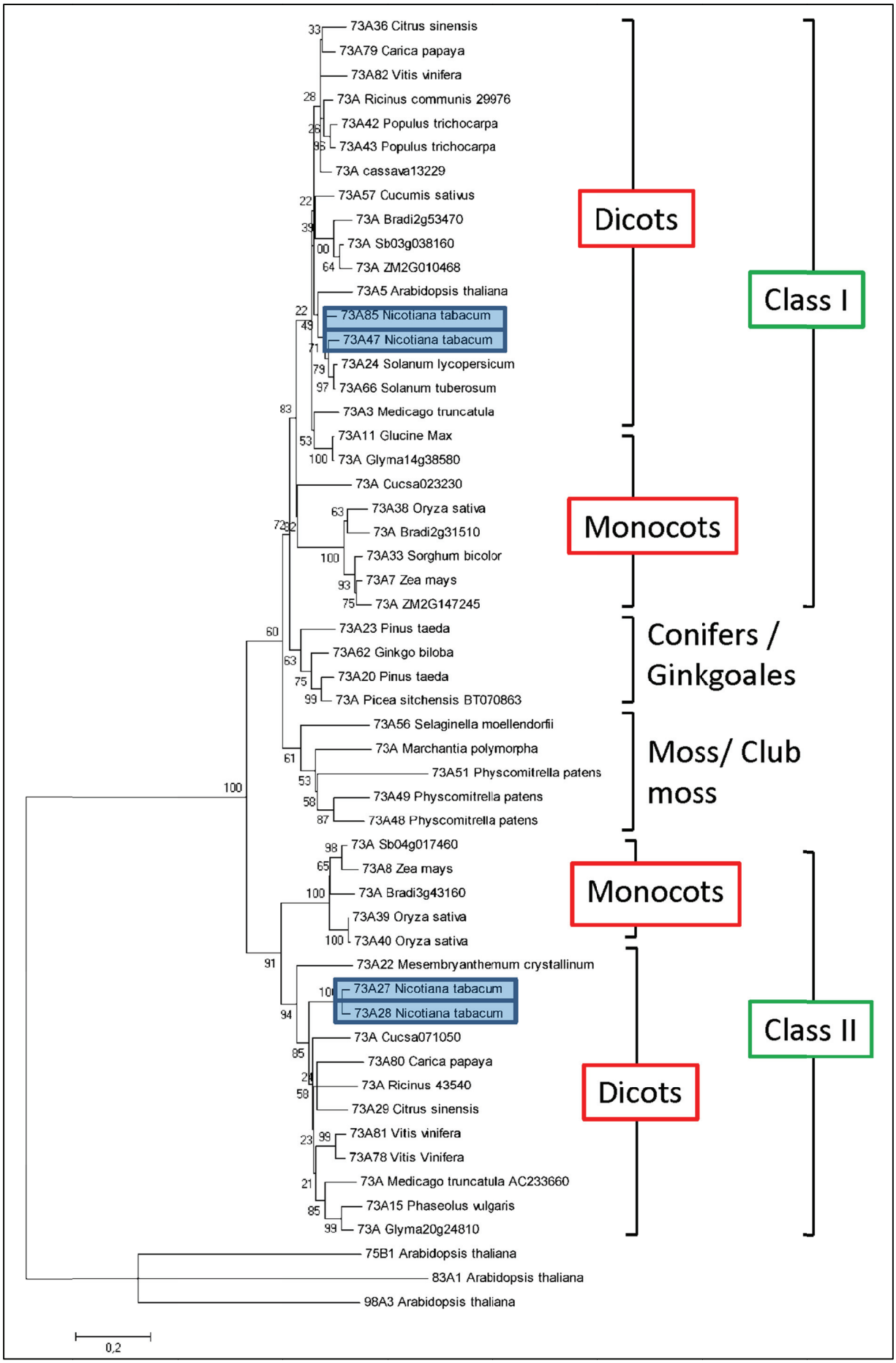

73A36 Citrus sinensis
73A79 Carica papaya
73A82 Vitis vinifera
73A Ricinus communis 29976
73A42 Populus trichocarpa
73A43 Populus trichocarpa
73A cassava13229
73A57 Cucumis sativus
73A Bradi2g53470
73A Sb03g038160
73A ZM2G010468
73A5 Arabidopsis thaliana
73A85 Nicotiana tabacum
73A47 Nicotiana tabacum
73A24 Solanum lycopersicum
73A66 Solanum tuberosum
73A3 Medicago truncatula
73A11 Glucine Max
73A Glyma14g38580
73A Cucsa023230
73A38 Oryza sativa
73A Bradi2g31510
73A33 Sorghum bicolor
73A7 Zea mays
73A ZM2G147245
73A23 Pinus taeda
73A62 Ginkgo biloba
73A20 Pinus taeda
73A Picea sitchensis BT070863
73A56 Selaginella moellendorfii
73A Marchantia polymorpha
73A51 Physcomitrella patens
73A49 Physcomitrella patens
73A48 Physcomitrella patens
73A Sb04g017460
73A8 Zea mays
73A Bradi3g43160
73A39 Oryza sativa
73A40 Oryza sativa
73A22 Mesembryanthemum crystallinum
73A27 Nicotiana tabacum
73A28 Nicotiana tabacum
73A Cucsa071050
73A80 Carica papaya
73A Ricinus 43540
73A29 Citrus sinensis
73A81 Vitis vinifera
73A78 Vitis Vinifera
73A Medicago truncatula AC233660
73A15 Phaseolus vulgaris
73A Glyma20g24810
75B1 Arabidopsis thaliana
83A1 Arabidopsis thaliana
98A3 Arabidopsis thaliana
0,2
Dicots
Monocots
Conifers / Ginkgoales
Moss/ Club moss
Monocots
Dicots
Class I
Class II

Figure 94. Phylogenetic reconstruction of the CYP73A family by the Maximum Likelihood method.
Sequences from *Nicotiana tabacum* were highlighted in blue. Abbreviations: Cucsa. *Cucumis sativus*, Cassava. *Manihot Esculenta*, Bradi. *Brachypodium distachyon*, Sb. *Sorghum bicolor*, ZM. *Zea mays*, Glyma. *Glycine max*.

An increasing number of sequences from class II were annotated in the last two years, as a result of accelerated sequencing of new genomes. Due to the low identity between class I and II (see next paragraphs), it seems that the class II was underestimated until now. It however appears widespread in plant kingdom taking into account the last sequenced genomes (Figure 95).

	Organisms	Class I	Class II
Dicotyledons	*Nicotiana tabacum*	2	2
	Glycine max	3	2
	Medicago truncatula	1	1
	Arabidopsis thaliana	1	-
	Populus tichocarpa	2	-
	Vitis vinifera	1	2
	Cucumis sativus	2	1
	Ricinus communis	1	1
Monocotyledons	*Sorghum bicolor*	1	1
	Zea mays	3	1
	Brachypodium distachyon	2	1
	Oryza sativa	2	1

Figure 95. CYP73As class repartition in eleven sequenced genomes plus *Nicotiana tabacum*.

Several copies of both classes were sometimes identified in the same plant (Figure 95). This could either point to a higher complexity of the phenolic metabolism or reflect recent gene/genome duplication (polyploidization). In counterpart, some species such as Arabidopsis have only one class I gene (Costa *et al.*, 2003). It can be mentioned that all Angiosperms shown to possess a class II enzyme also appear to contain a class I complement, but the reverse is not true. This suggests that while class I would be essential, subfunctionalization most likely occurred in class II.

Classification	P450 Nomenclature	Plant species	Identities in comparison to CYP73A85 (%)	Identities in comparison to CYP73A27 (%)	Classes
Dicots	73A85	*Nicotiana tabacum*	100	58	I
	73A27	*Nicotiana tabacum*	58	100	II
	73A47	*Nicotiana tabacum*	93	58	I
	73A5	*Arabidopsis thaliana*	87	58	I
	73A36	*Citrus sinensis*	84	58	I
	73A42	*Populus trichocarpa*	90	58	I
	73A57	*Cucumis sativus*	89	59	I
	73A79	*Carica papaya*	90	59	I
	73A82	*Vitis vinifera*	86	57	I
	73A11	*Glycine max*	88	59	I
Monocots	73A7	*Zea mays*	77	56	I
	73A38	*Oryza sativa*	75	55	I
	73A33	*Sorghum bicolor*	77	56	I
	unnammed 73A	*Sorghum bicolor*	85	58	I
	unnammed 73A	*Zea mays*	83	57	I
Dicots	73A81	*Vitis vinifera*	58	77	II
	73A80	*Carica papaya*	59	72	II
	73A28	*Nicotiana tabacum*	58	95	II
	73A29	*Citrus sinensis*	58	76	II
	73A15	*Phaseolus vulgaris*	58	72	II
Monocots	73A40	*Oryza sativa*	58	66	II
	73A39	*Oryza sativa*	59	66	II
	73A8	*Zea mays*	58	64	II
	unnammed 73A	*Brachypodium distachyon*	57	64	II
	unnammed 73A	*Sorghum bicolor*	59	66	II
Conifers / Ginkgoales	73A62	*Ginkgo biloba*	79	58	
	73A20	*Pinus taeda*	79	58	
	73A23	*Pinus taeda*	75	58	
	unnammed 73A	*Picea sitchensis*	78	59	
Moss / club moss	73A56	*Selaginella moellendorfii*	69	55	
	73A51	*Physcomitrella patens*	62	49	
	73A48	*Physcomitrella patens*	70	54	
	73A49	*Physcomitrella patens*	71	55	
	unnammed 73A	*Marchantia polymorpha*	70	52	

Figure 96. Amino acid identities of *Nicotiana tabacum* CYP73As and some other known CYP73As. Sequences were sorted by identities to CYP73A27 (C4H class II) and CYP73A85 (C4H class I).

Figure 96 summarizes the protein sequence identities among CYP73 proteins. It emphasizes the divergence between the class I and class II proteins. Only 55-59 % of identity is observed between classes whereas the limit of membership in the same P450 subfamily is at 55 % identity. A second striking feature is that protein identity among class II ranges between 64 and 77 %, while class I proteins are more highly conserved and maintain at least 77 % identity when both monocots and dicots are considered and more that 84% identity within dicotyledonous plants (Figure 96). As expected from the tree structure, CYP73 sequences from conifers and mosses are more similar to class I sequences (72 % identical) than to class II (60 %). Class I can be thus regarded as the ancestral enzyme.

In summary, an early duplication of an ancestral CYP73 gene occurred before the separation of monocots and dicots. This duplication was maintained in Angiosperms with a significant divergence of the clade II proteins. This is likely to have led to a subfunctionalization of the class II enzymes.

2.2. Origin of the class I and II CYP73 duplicates in *N. tabacum*

From the *CYP73As* present in *Nicotiana tabacum,* two belong to class I and two others to class II (Figure 94). *Nicotiana tabacum* is an allotetraploid, hybrid of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. It was thus tempting to hypothesize that the duplication in each class was the result of hybridization. To test this hypothesis, the coding sequences from each ancestor were amplified from leaf genomic DNA, cloned into pBI vector and sequenced.

	CYP73Axx *N. sylvestris*	*CYP73Ayy* *N. sylvestris*	*CYP73Aaa* *N. tomentosiformis*	*CYP73Abb* *N. tomentosiformis*
CYP73A27	0,95	0,18	0,91	0,18
CYP73A28	0,90	0,18	0,94	0,19
CYP73A47	0,19	0,78	0,20	0,78
CYP73A85	0,19	0,91	0,19	0,89

Figure 97. Identity matrix of alignment of partial *CYP73A* sequences from *N. tabacum CYP73As* over *N. tomentosiformis* and *N. sylvestris*.
Assembly of *CYP73As* sequences from *N. tabacum* ancestors was done from several sequenced clones. Alignment is shown in appendix 327.

Figure 97 shows the identity matrix of the *N. sylvestris* and *N. tomentosiformis* coding sequences with the *CYP73s* from *N. tabacum*. These data confirm our hypothesis and clearly show that one of the class II CYP73As (CYP73A28) is derived from *N. tomentosiformis*, while CYP73A85 and CYP73A27 are derived from *N. sylvestris*. Class I CYP73A47 could derive from *N. tomentosiformis* or another ancestor.

This would suggest that *N. tabacum* CYP73 proteins in each class are essentially redundant and could be expected to have the same or very similar enzymatic and physiological functions.

2.3. Structural characteristics of the class II CYP73s

As mentioned in the introduction of this chapter, the class II proteins are characterized by an atypical and long serine/threonine-rich N-terminal sequence that is possibly indicative of a non-ER subcellular localization (Ehlting *et al.*, 2006). It may also be indicative of interaction with different membrane compartiments (rafts), lipids or other partner proteins. In support to the latter hypothesis, recently Subramanian *et al*. (2010) demonstrated interaction of mammalian P450 (CYP2C9 and CYP3A4) via their N-termini. The P450 membrane anchor is only partly responsible for the interaction with the membrane, but can be essential for protein function. N-terminally truncated proteins can be not functionnal. For example, the deletion of the twenty-seven first amino acids of CYP2B4 influences folding of the P450 and electron transfer from the CPR resulting in increased uncoupling (Lehnerer *et al.*, 1998). The role of this major difference in the N-terminal sequence has been further investigated as described in the section: "A second class of CYP73As?" p.232.

Several others sequence differences are also observed between the two classes (small deletions, insertions or mutations). In an attempt to anticipate the impact of these structural differences in the primary sequence of class I and II CYP73 proteins, homology modeling was carried out to map main differences on the protein tridimensional structure. To date, the structures of only few membranar P450s are available. The structure of only one plant P450 (CYP74) was resolved, but it is not representative of CYP71 clan proteins (Chang *et al.*, 2008). Thus, to compare most probable structures of tobacco CYP73As, models were built from tobacco amino acids sequences aligned to the closest non-plant P450s with resolved structure. A similar approach was previously reported for modeling CYP73A5 active site based on crystal structure of a bacterial P450 (CYP102) (Rupasinghe *et al*., 2003).

The work, described here, was carried-out in collaboration with Dr. Fabrice Klein of the IGBMC (Strasbourg, France). The N-terminal helix is hyper variable and does not contribute to globular protein structure. It was thus excluded from homology modeling and removed to generate the template proteins. Peptidic sequences of CYP73A27 and CYP73A85 were blasted and aligned with sequences from protein databank. The sequences having a PPxP segment (assumed to be P450s) were kept. CYP73As sequences were finally aligned on 2F9Q structure of CYP2D6 proved to be the best reference for homology modeling (according to protein sequences alignment). 3D models were generated from the aligned sequences with swiss-model website. Alignments used for generation of the models are in appendix p.326.

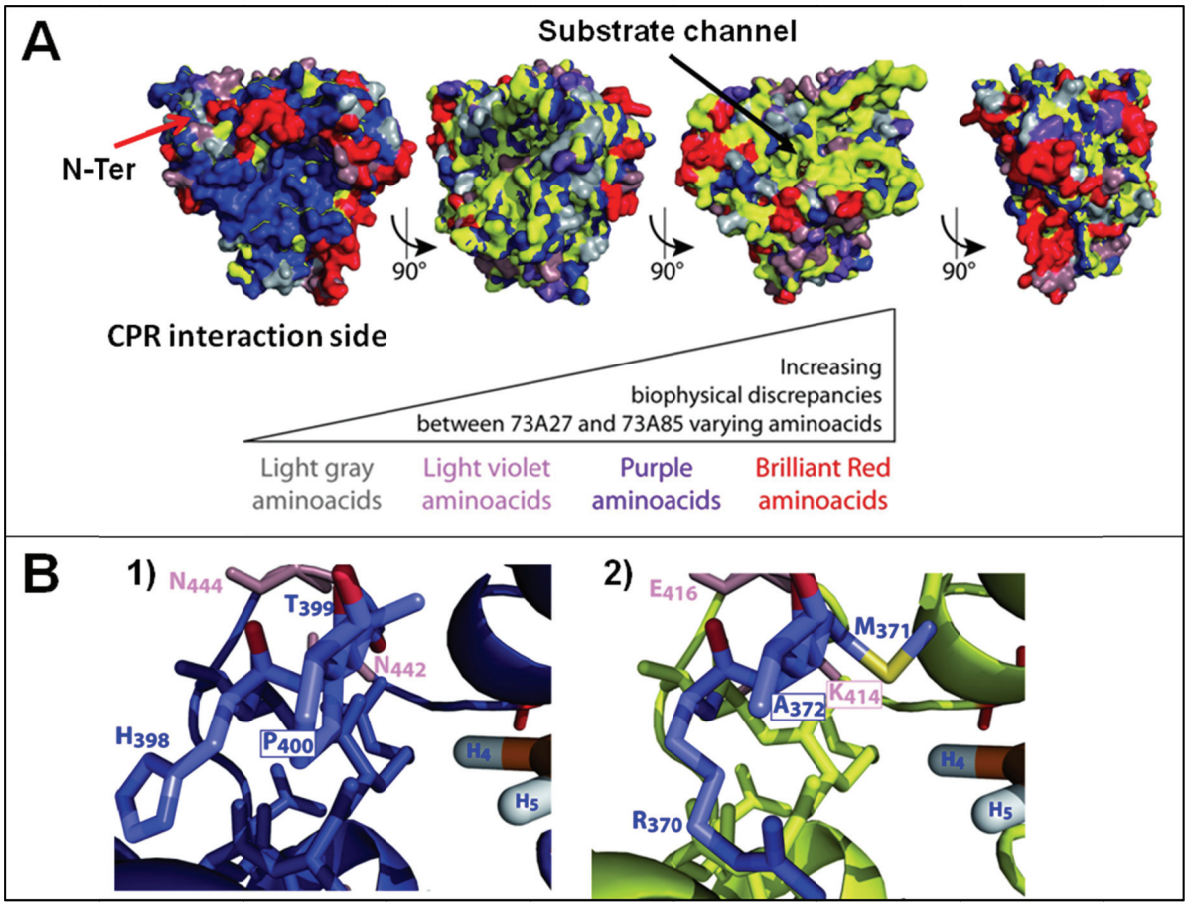


Figure 98. Comparison of CYP73A27 (class II) and CYP73A85 (class I) surface and active site based on homology modeling.

A. Surface variations of superimposed CYP73A27 and CYP73A85 models. Blue and yellow surfaces are conserved aminoacids. Others colors indicate increasing discrepancies between the two models.

B. Details of the substrate pocket. 1) CYP73A27. 2) CYP73A85. The three most variable aminoacids in the pocket - RMA (CYP73A85) and HTP (CYP73A27) - are shown as sticks. The cinnamic acid hydrogens H_4 and H_5 are also shown. Models were generated with swiss-model website from alignement of CYP73As with CYP2D6.

Three striking elements emerge from protein alignments and models.

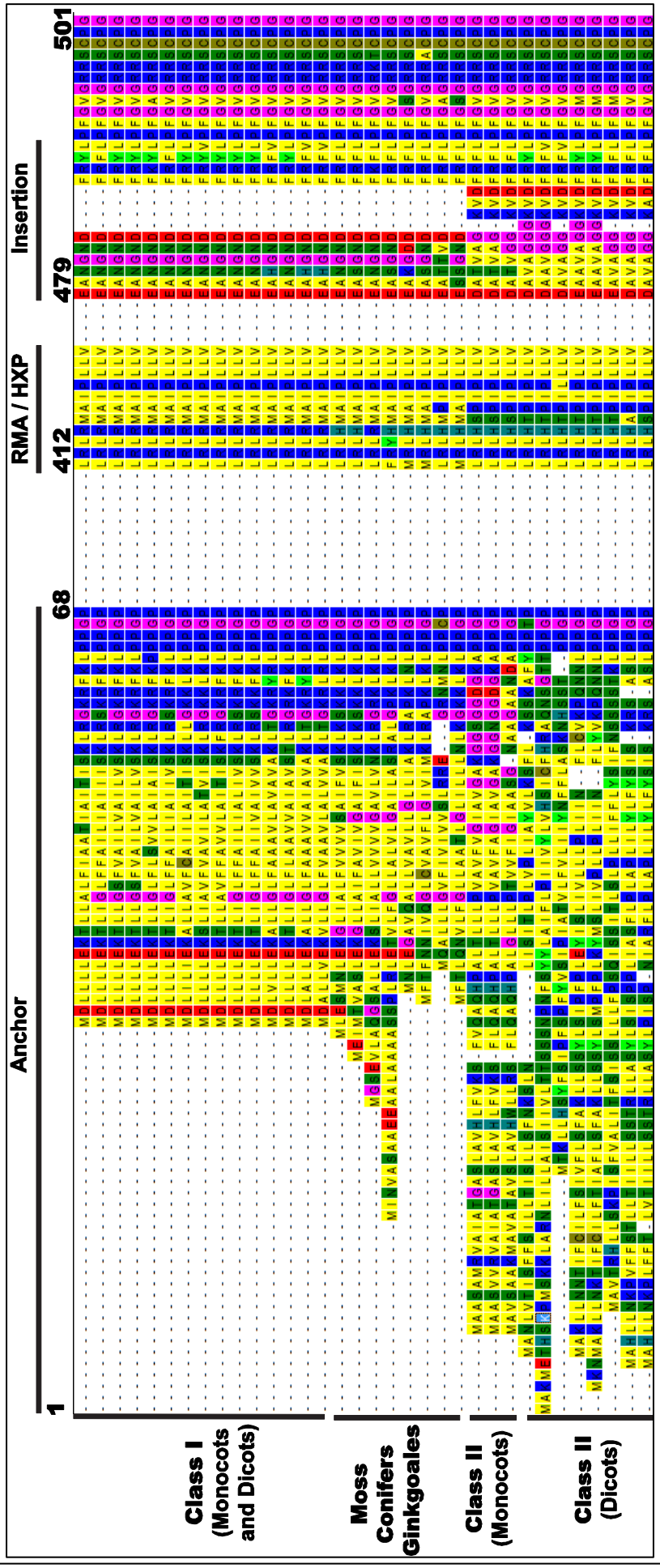
First, the main divergences between class I and II results in **modifications of the protein surface** (Figure 98 A), more precisely on top of the roof of the active site, on the heme proximal surface and close to the substrate access channel. Such surface modifications could indicate interaction with different partner proteins (electron donor, other pathway proteins). Interestingly, cytochrome P450 reductase (CPR) is reported to interact with P450 via the heme proximal surface (Jensen & Møller, 2010). Thus, the divergence observed in this region might reflect interaction with another electron donating protein.

The second striking feature is a **modification of three amino acids: RMA** (371) **in HTP** (400) **located in the heme proximal region of the catalytic site** (Figure 98B). An active site homology model and mutagenesis studies of CYP73A1 have previously demonstrated that asparagine (N302) in helix I (SRS 4) and isoleucine (I371) in SRS5 completely conserved among CYP73s were essential for substrate binding in the active site. The asparagine forms a hydrogen bond with the carboxylate of the substrate and the isoleucine is predicted to form hydrophobic contacts with the aromatic ring of the substrate to position it in close proximity to the heme iron (Schoch *et al.*, 2003). Just next to this isoleucine is present an RMA triplet in CYP73A85. This triplet is replaced by an HTP sequence in *Nicotiana tabacum* class II proteins (Figure 98B). This modification increases slightly the volume of the active site in a region in vicinity to the substrate aromatic ring. In addition, the presence of a proline and of a threonine introduces a new flexibility and a possibility to engage hydrogen bonds that do not exist in type I proteins. This class II HTP(400) motif is reminiscent of the HPPTP sequence present in the SRS5 of the CYP98 proteins. In this motif, T367 in CYP98A3 is described as forming a hydrogen bond with the 4'-OH for stabilizing the phenolic ring in a correct position above heme. It is found in all meta-ring hydroxylases such as flavonoid 3',5'-hydroxylases (CY75; Seitz *et al.*, 2007). The "RMA/HTP" motif could play role in substrate binding and stabilization. Interestingly, the motif RMA is perfectly conserved in class I protein, whereas the motif H-(T/S/A)-P is found in CYP73As class II protein (Figre 99), which suggests functional variations within this class of proteins and between the two classes. Annoted CYP73As from conifers, mosses and ginkgoales possess an intermediate motif: (H/R)-M-(A/P).

A third striking elements is **an insertion of 3-4 amino acids in the meander region upstream the CPG** (Figure 99) and it was expected to have a significant impact on the catalytic site structure. But according to homology model, this insertion form a protuberance outside the P450 and it seems most likely involved in interaction with an electron-donating protein, which could be different to canonical class I protein. Annoted CYP73As from conifers, mosses and ginkgoales do not share the insertion of the class II proteins.

Several others divergences exist between the two classes (small deletions, insertions or mutations) but they are not well conserved inside one class.

Figure 99. Main differences between class I and class II CYP73As sequences.


Anchor
RMA / HXP
Insertion
1
68
412
479
501
Class I
(Monocots
and Dicots)
Moss
Conifers
Ginkgoales
Class II
(Monocots)
Class II
(Dicots)

Significant sequence divergences between class I and II CYP73As can be identified including atypical N-terminal membrane anchor, 3-4 amino acids insertion, and HXP motif in the active site. These results suggest either different sub-cellular localizations or interaction with different partner proteins.

2.4. Subcellular localization of the CYP73 protein

Analysis of a few *Arabidopsis* and other plant P450 sequences, predicts potential signal peptides in the N-terminus that should target them to plastids or mitochondria (Schuler *et al.*, 2006). This prediction had been confirmed in few cases. Of these P450s: CYP74A1 (allene oxide synthase), CYP74B2 (hydroperoxide lyase), CYP97A3 (carotene β-hydroxylase), CYP97C1 (carotene ε-hydroxylase), CYP86B1 (fatty acid hydroxylase) and CYP701A3 (kaurene oxidase) have been experimentally demonstrated to be chloroplast-localized (Froehlich *et al.*, 2001; Helliwell *et al.*, 2001; Watson *et al.*, 2001; Tian *et al.*, 2004; Kim & DellaPenna, 2006; Compagnon *et al.*, 2009). All contain a putative cleavage site for a chloroplast transit sequence, and a substantial numbers of serine and threonine (>14%) in their N-terminus. A different sub-cellular localization could explain the hypothetical different electron donor, previously mentioned.

Different programs were thus tested to predict the sub-cellular localization of the CYP73As from *N. tabacum* (ChloroP: http://www.cbs.dtu.dk/services/ChloroP/; Predotar: http://urgi.versailles.inra.fr/predotar/predotar.html; SignalP: http://www.cbs.dtu.dk/services/SignalP/; BaCelLo: http://gpcr.biocomp.unibo.it/bacello/). In most of the cases, class I CYP73As were predicted localized on the ER. For class II CYP73As, the predictions were less clear-cut and indicated localizations from ER to plastids, mitochondria, plasma membrane and cytoplasm.

GFP fusion constructs were thus generated to determine their sub-cellular localization. Fusion proteins were transiently over-expressed in *N. benthamiana* leaf epidermal cells via agroinfiltration and fluorescent signal was observed by laser scanning confocal microscopy.

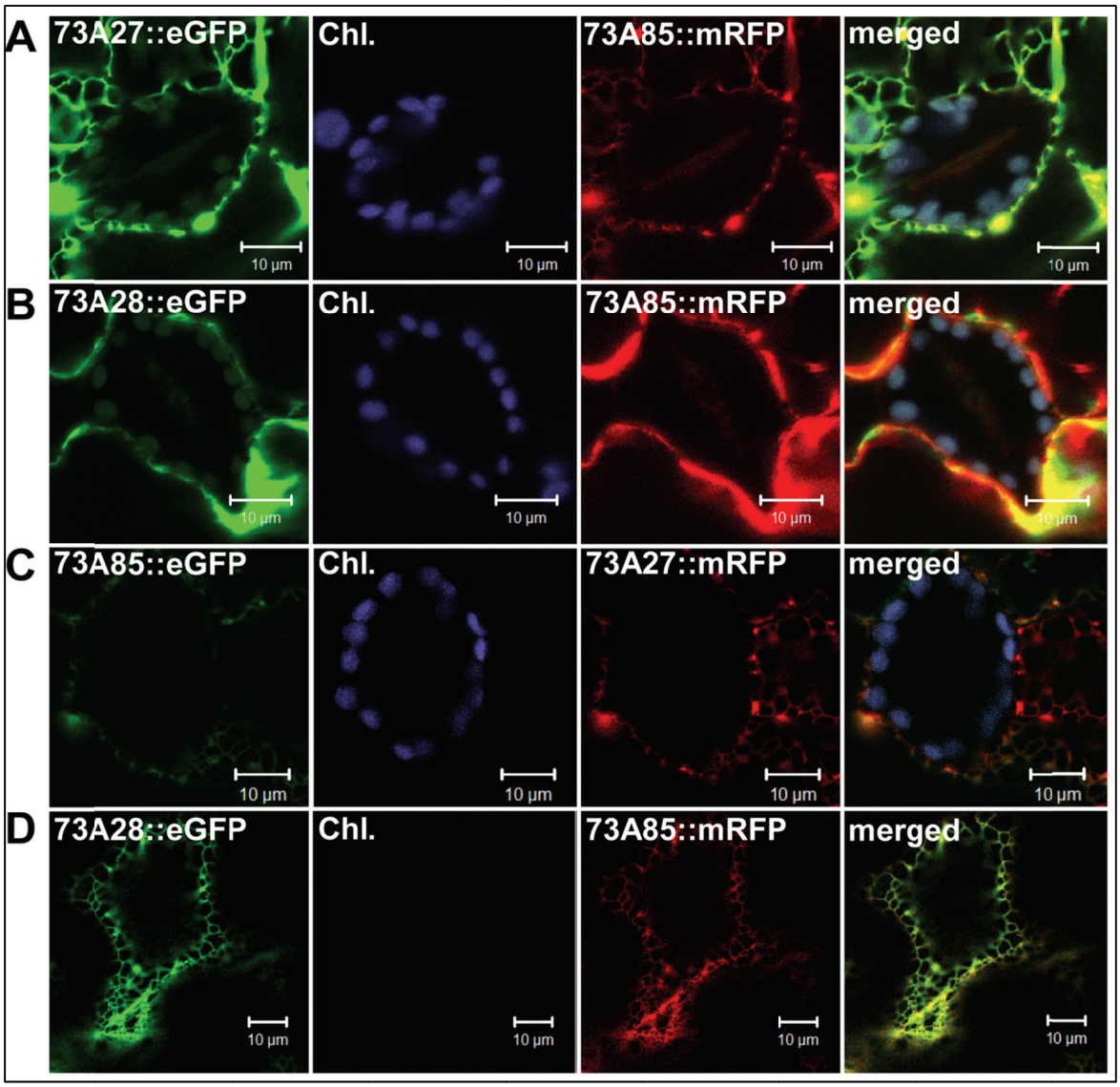


Figure 100. Sub-cellular localization of CYP73A85 (class I) and CYP73A27/28 (class II) in *N. benthamiana* visualized by laser scanning confocal microscopy.

A. Co-expression of CYP73A27::eGFP with CYP73A85::mRFP.

B. and D. Co-expression of CYP73A28::eGFP with CYP73A85::mRFP.

C. Co-expression of CYP73A27::mRFP with CYP73A85::eGFP.

Images were taken 5 days after infiltration. Chl. Chlorophyll was vizualised by a 650 nm long-pass filter. Scale bar: 10 µm.

Most CYP73As available to date are highly similar and belong to class I. These sequences contain a classical N-terminal targeting signal, and it was repeatedly shown that class I C4H were indeed targeted to the ER (Ro *et al.*, 2001; Achnine *et al.*, 2004). In the case of the tobacco CYP73A85, predominant GFP signal was also associated to ER structures (Figure 100 C). A very similar membrane-associated pattern is observed upon expression of CYP73A27 and CYP73A28 class II proteins. Nevertheless, a low signal was detected in chloroplasts of guard cells with both of them when fused with GFP (Figure 100 A, B). This signal was not observed with class I (Figure 100 C). However the signal was not confirmed by inverting the fluorescent proteins, using class II protein-mRFP fusion (Figure 100 C) and could thus be due to artifact caused by pigments in the chloroplasts. To avoid artifact, laser excitation and detector amplification were tightly controlled. The signal intensity was adjusted below the saturation with ER structures as control.

In a previous study (Compagnon *et al.*, 2009) with CYP86B1 known to be targeted to plastids (Watson *et al.*, 2001), fluorescence was only detected in chloroplasts after a prolonged inspection at seven days post-infiltration. Thus, observation was extended overall several days for CYP73A proteins. However, no more signal in plastids was detected until ten days post-infiltration. After ten days of P450 expression, leaves and cells were senescent. Fluorescent fusion proteins were progressively aggregated all over the cell, before disappearance of the fluorescence (Results not shown).

Based on the predominant GFP signal in the reticulate membrane structures, localization of all the *N. tabacum* CYP73As can be considered as mainly ER. Transfer of a minor proportion of the class II protein into plastids however cannot be excluded. Tissue-specific variations in such transfer are also possible. Compartmentation in different ER subdomains or different protein-protein interactions thus have to be considered for the class II CYP73As.

2.5. CYP73As expression patterns in wild type tobacco

Different expression patterns would be expected for *CYP73s* with different functions. Quantitative real-time PCR was thus carried-out on the forty different tissues previously mentioned (Figure 82). These tissues are representative of all organs of the plants (roots, stems, leaves and flowers) and of specific flowers developmental stages.

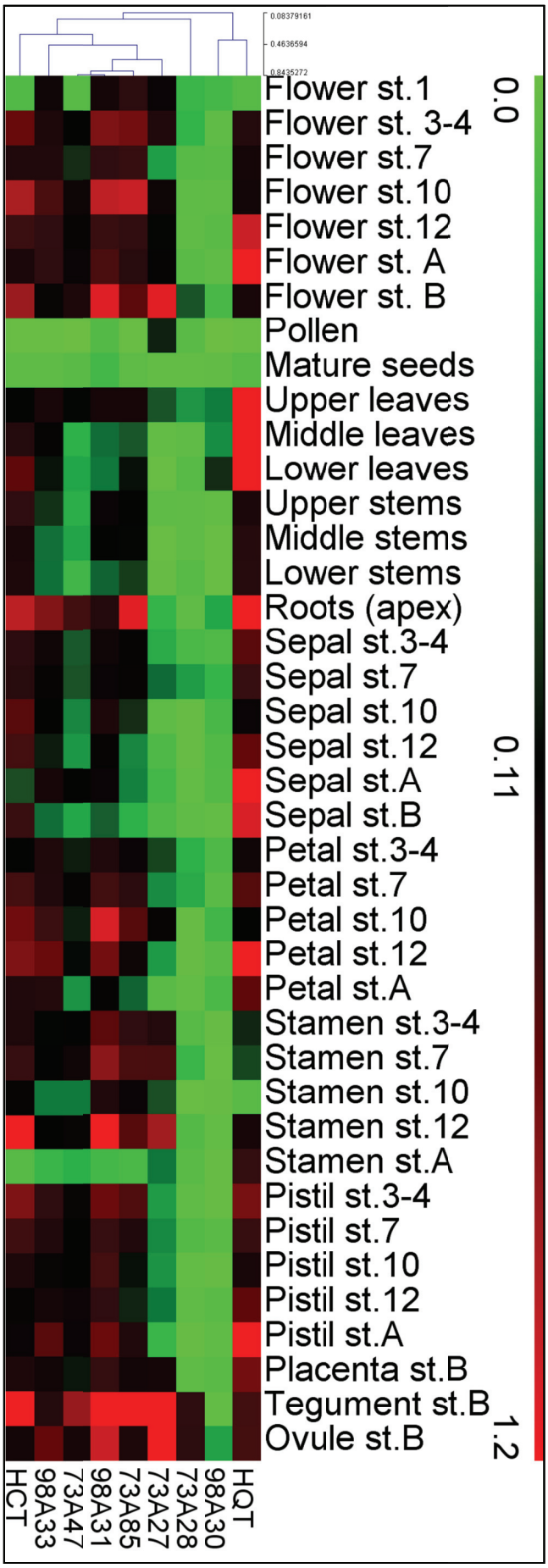


Figure 101. ***CYP73As*** **transcription heatmaps of wild type tobacco organs.**

RT-qPCR data were analyzed by the δCt method. Heatmap generation and gene hierarchical clustering was done with MEV software. Experiment was repeated twice and the results shown here are representative. Flower stages are as shown in chapter II.

In order to identify co-expressed genes and potential partner proteins, their expressions were compared to the expression of *CYP98As*, *HCT* and *HQT* genes involved in lignin or chlorogenic acid biosynthesis.

CYP73A47 and *CYP73A85* (class I) appear constitutively transcribed in lignifying tissues, roots, upper leaves, stems, but also in flowers and fruits, including sepals, petals, stamens, pistils, tegument and ovules (Figure 101). They show very similar expression patterns, with *CYP73A85* stronger expressed than *CYP73A47*. *CYP73A47* transcription is highly correlated to that of *CYP98A31* and *CYP73A85*, which could be likely to reflect an involvement of class I enzymes in lignin biosynthesis. This would be in accordance with typical expression of *CYP73A5* in *A. thaliana* in organs with high phenylpropanoid biosynthesis and undergoing lignification (Mizutani *et al.*, 1997; Urban *et al.*, 1997). A coordinate expression of *CYP73A5* with all other genes of the phenylpropanoid pathway was also clear during development of inflorescence stems (Ehlting *et al.*, 2005).

Expression of tobacco class II *CYP73As* is more restricted (Figure 101). Low *CYP73A28* expression is detected in upper leaves. It is more significant in fruits, more precisely in the tegument and ovule of the fruit. *CYP73A27* is actively transcribed in flowers and fruits, in particular in stamens, placenta, tegument and ovules. Thus, class II *CYP73As* are not expressed in lignifying tissues but in reproductive organs which accumulate specific phenylpropanoids compounds.

Class I and II tobacco *CYP73As* show clearly distinct transcription patterns upon plant and flower development. While expression of class I genes seems related to lignification, class II seems to have specific roles in flower and fruit development.

2.6. *CYP73s* expression in response to stress

Most of the *CYP73As* genes have been isolated from tissues undergoing stress and high C4H expression was reported in correlation with the production of defense compounds (Logemann *et al.*, 1995; Mizutani *et al.*, 1997; Urban *et al.*, 1997; Batard *et al.*, 1997). To further discriminate *CYP73As* functions, their stress response was also compared to that of *CYP98As*, *HCT* and *HQT*.

2.6.1. Effect of biotic stress on CYP73As expressions

TMV was inoculated in young wild type tobacco plant leaves.

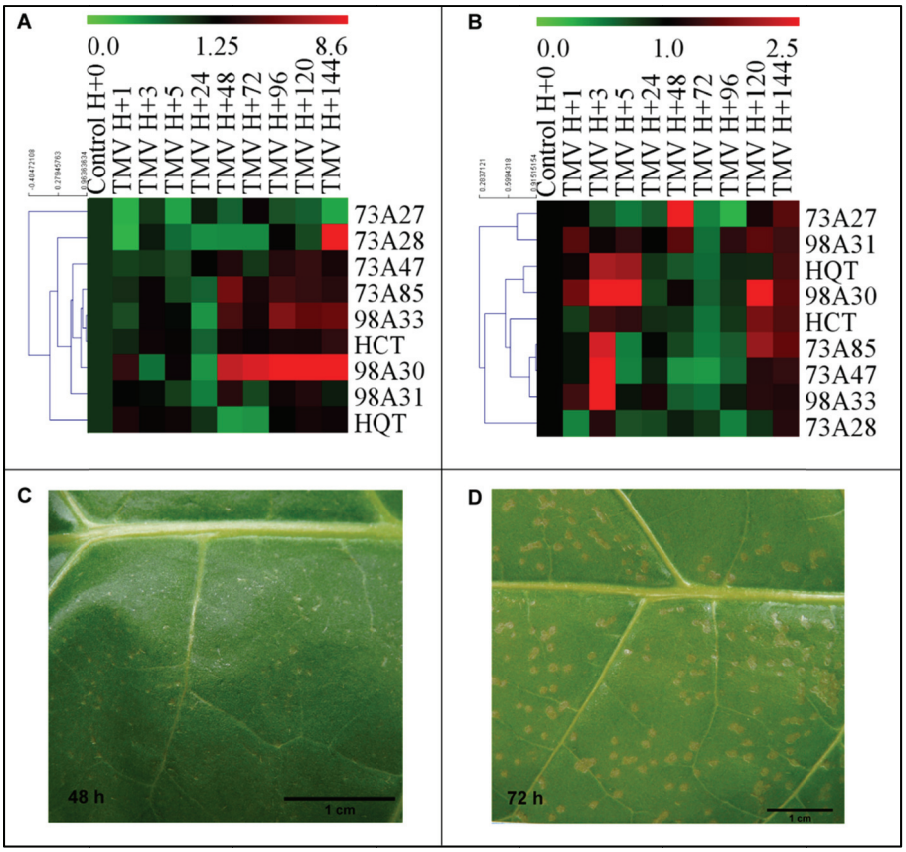


Figure 102. *CYP73As* transcription heatmaps of two month-old wild type tobacco leaves under biotic stress treatments.

A. Local response of two month-old tobacco plants inoculated by Tobacco Mosaic Virus (TMV).

B. Systemic response of two month-old tobacco plants inoculated with TMV.

C. Image of leaf surface 48 hours after infection.

D. Image of leaf surface 72 hours after infection.

Scale bar: 1 cm. RT-qPCR data were analyzed by the δδCt method. VMT experiment was repeated twice, and the results shown here are representative.

Locally, expression of class I *CYP73As* was induced in the late response to TMV together with *HCT* and *CYP98As* (Figure 102 A). Interestingly, this late response corresponds to the appearance of the viral lesions on the leaf surface (Figure 102 C, D). Conversely, *CYP73As* from class II were not activated, and even down regulated during the first hour after the inoculation. Upon such local response *CYP73A85* (class I) transcription was highly correlated to the transcription of *HCT* and *CYP98A33*, which could indicate involvement in lignin biosynthesis in the necrotic lesions.

In the systemic response to TMV infection, due to variation of expression in the range of the inter-plant variability (about 2 for each gene), it was difficult to conclude to systemic activation of the candidate genes (Figure 102 B).

To conclude, class I *CYP73As* seem to contribute to defense against TMV and their expression is associated with *HCT* transcription. No significant activation of class II *CYP73As* was observed.

2.6.2. *Effects of environmental and hormonal stresses*

During VMT experiments, a wound stress was applied as control of the VMT inoculation (Figure 103 A, B). From the local response analysis, change in transcription was detected during the early response for *CYP73A47* (class 1) and *CYP98A31*. No other significant response could be outlined. Nevertheless, hierarchical clustering of the gene transcription patterns correlates expression of the two class II *CYP73As* on one side and *CYP73A85* with *CYP98A31* and *CYP98A33* on the other side.

In the same way, systemic response study was not conclusive since expression fold changes were very low.

An additional wound stress experiment was carried out. A more severe stress was then delivered to the leaves with a needle, so as to generate small holes on the whole leaf surface. Nevertheless, only low variation in transcription level was detected that did not differenciate gene expression (Figre 103 C).

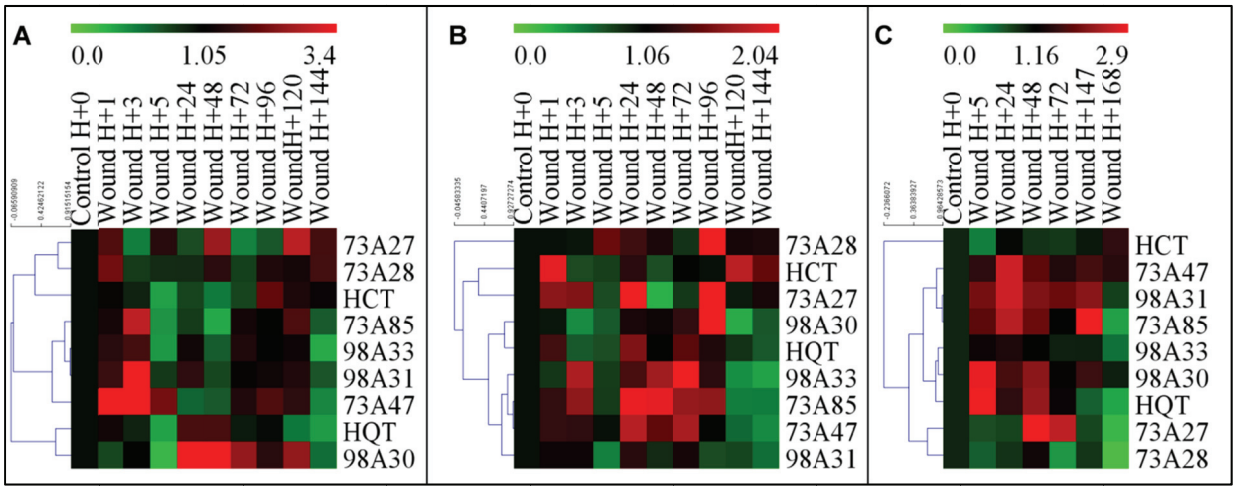


Figure 103. *CYP73As* transcription heatmaps of wounded two month-old wild type tobacco leaves.
A. Local response to leave abrasions used for TMV inoculation.
B. Systemic response to leave abrasions used for TMV inoculation.
C. Wound stress generated by picking with needle the whole surface of the leaf.
RT-qPCR data were analyzed by the δδCt method.

Finally, in an attempt to delinate more conclusive correlations several other stresses were tested with only two or three kinetic points.

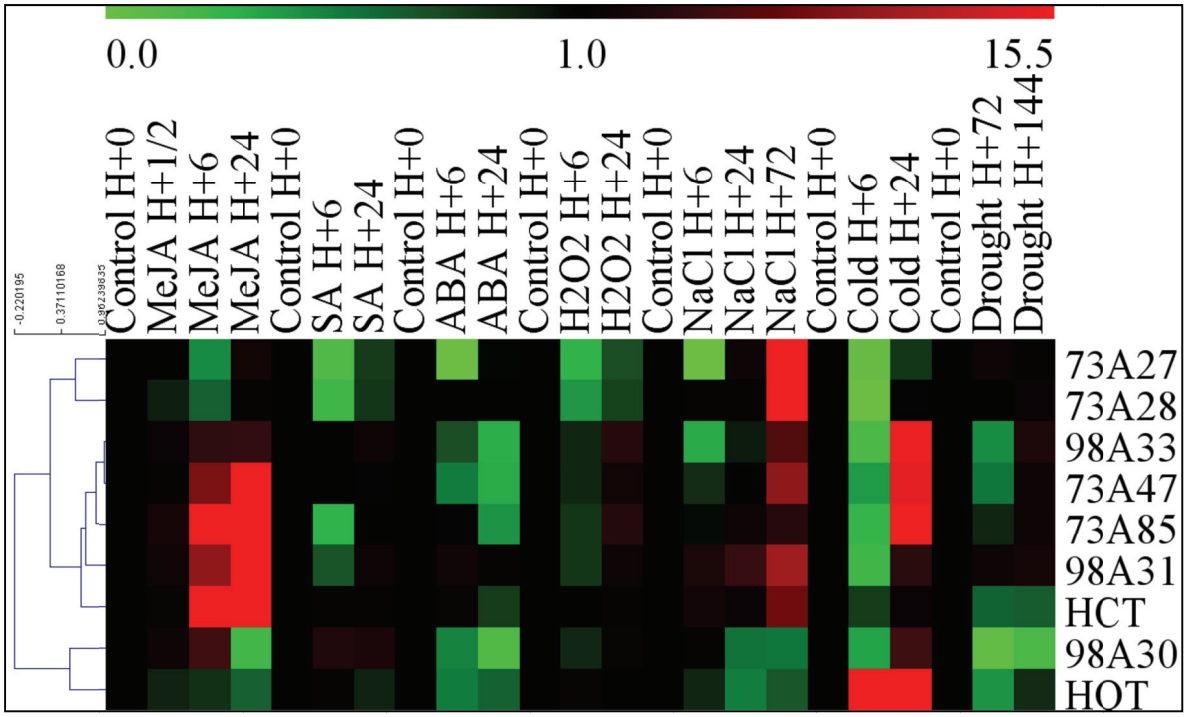


Figure 104. *CYP73As* transcription heatmaps of two month-old wild type tobacco leaves submitted to various stresses. RT-qPCR data were analyzed by the δδCt method.

Interestingly, a very strong and selective response of class II *CYP73* genes was observed upon late salt treatment. The highest activation values (86 fold) were obtained with *CYP73A27* after 72 hours NaCl (Figure 104). Conversely, early repression was observed upon salicylic acid, methyl Jasmonate, hydrogen peroxide and cold treatments for both genes and upon early ABA and NaCl response in the case of *CYP73A27*. Early response to cold stress was accompanied by down-regulation of all *CYP98As* and *CYP73As*.

A very different activation pattern was observed for class I genes with strong and selective activation by MeJA (6-24h) and cold (24h) treatments. Interestingly, MeJA, activation occurred with concomitant expression of *HCT* and *CYP98A31*, while cold response involved *HQT* and *CYP98A33*. *CYP73As* from the class I were also a little induced after 24 hours of hydroperoxyde and drought stresses.

In conclusion, the transcription patterns of the two classes of CYP73As are very clearly different. Class II can be clearly associated with salt stress and flower and fruit development, while expression of class I genes seems related to lignification.

2.7. Optimized heterologous expression

As structure comparison of the two classes of CYP73As indicates difference in the heme proximal region of the active site, comparison of their substrate preference was undertaken after heterologous expression in yeast.

Thus, three *CYP73As* coding sequences (one class I and two class II) were cloned into the pYeDP60 vector, suitable for yeast transformation and inducible expression of P450s in this organism. *CYP73A27* sequence was isolated from flower cDNA. *CYP73A28* and *CYP73A85* cDNA were kindly provided by J. Chapell (University of Kentucky, USA) and P. Bolwell (University of London, UK), respectively. Microsomes were prepared from transformed WAT11 yeast after induction of P450 and P450 reductase proteins expression by galactose.

The first attempts to express tobacco CYP73As in yeast showed that CYP73A85 was well expressed and a P450 concentration of 1 µM in microsomal preparation was calculated from differential CO bound-reduced versus reduced absorption spectrum (Figure 105). For CYP73A27, a peak at 450 nm was detected but also a peak at 420 nm indicating the presence of a proportion of inactive protein. Enzyme concentration in recombinant microsomes evaluated from the peak at 450 nm was 1µM. Expression of CYP73A28 turned out to be more challenging and only a peak at 420 nm in differential CO spectra was observed suggesting the expression of inactive enzyme.

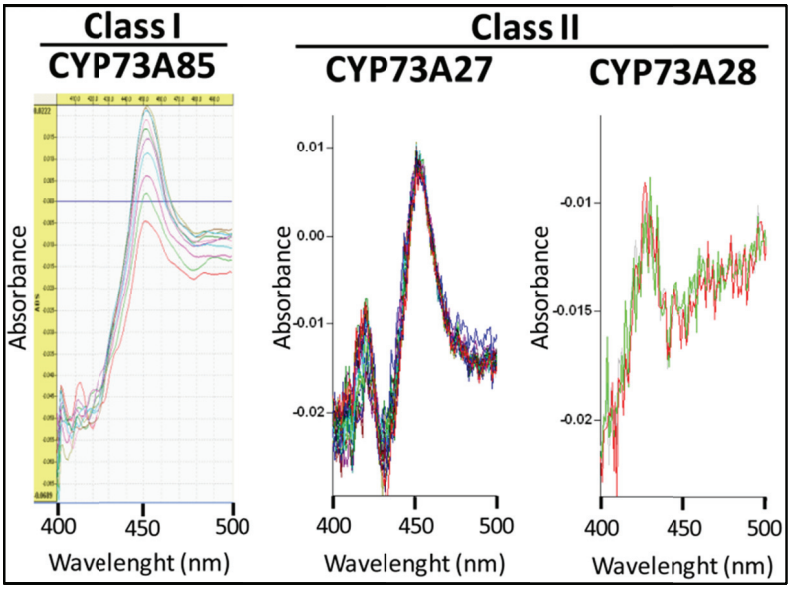


Figure 105. CYP73As differential CO spectra obtained after preparation of microsomes from transformed WAT11 strain and expression induction at 20°C.

In order to improve protein expression, a decrease in the induction temperature from 24°C to 20°C was attempted without success (results not shown).

Previously, exchange of the atypical N-terminal sequence of CYP73A15 improved expression of this P450 and resulted in a small increase in its K_m for cinnamate (Nedelkina *et al.*, l999). We therefore exchanged the sequences encoding signal peptides of the tobacco class II CYP73As for those encoding the N-terminal sequences of CYP73A85 (class I). But no significant improvement was noticed according to differential CO spectra (Results not shown).

Finally, expression of class II enzymes in the WAT21 yeast strain (expressing the cytochrome P450 reductase that also has a long S/T-rich N-terminal signal peptide) was tested. Again, no improvement in protein expression was obtained (results not shown).

To conclude, no improved expression of CYP73A28 could be obtained. So, tests were carried out to understand where the problem was.

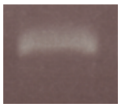
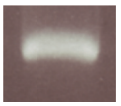
	Class I	**Class II**	
	73A85	**73A27**	**73A28**
OD culture before preparation of microsomes	4.7±0.3	4.6±0.2	3.8±0.2
PCR on expression vector (22 cycles)			
RT-qPCR (δδCt)	1±0.5	14±1.2	25±2.3
Total proteins in microsomes ($mg.mL^{-1}$)	27±3	23±1	19±3
Total microsomal proteins per liter of culture ($mg.L^{-1}$)	140	140	67

Figure 106. Follow-up of cultures and microsome production from recombinant yeast harbouring CYP73As.
Tests were carried-out by induction of P450 and CPR expression with 20 g.L-1 of galactose. OD measurement as RT-qPCR were performed on three independent cultures per construction. Cultures were sampled before preparation of microsomes. Total protein concentration in microsomes is average from three aliquots.

Results obtained during these tests showed that (Figure 106):

- Growth of the yeast cultures expressing CYP73A28 was impacted.
- Plasmid was detected in yeast before preparation of microsomes, thus selection of the transformants was effective.
- RT-qPCR experiment indicates a higher transcription of *CYP73A28* than of *CYP73A85* and *CYP73A27*.
- Concentrations in microsomal protein were similar in preparations from the different recombinant yeasts and a little more microsomes were produced with CYP73A85.
- Less microsomes were produced by the CYP73A28 expression culture obviously due to lower growth of the yeast. This can be possibly related to the high level of the *CYP73A28* transcripts.

Therefore, transcription of *CYP73A28* was higher than transcription of the other P450s, but not followed by a correct translation into protein. This could be due to either an improper translation stop, either to a too fast/too high expression leading to protein mis-folding, or to toxicity of CYP73A28 with resulting problems for the yeast growth. It is however striking that this case is very similar to that of CYP98 proteins mentioned in previous chapter: in all cases high transcript accumulation resulted in a failure in correct protein production. Consequently, a reduction in CYP73A28 expression rate was attempted by decreasing concentration of the expression inducer galactose.

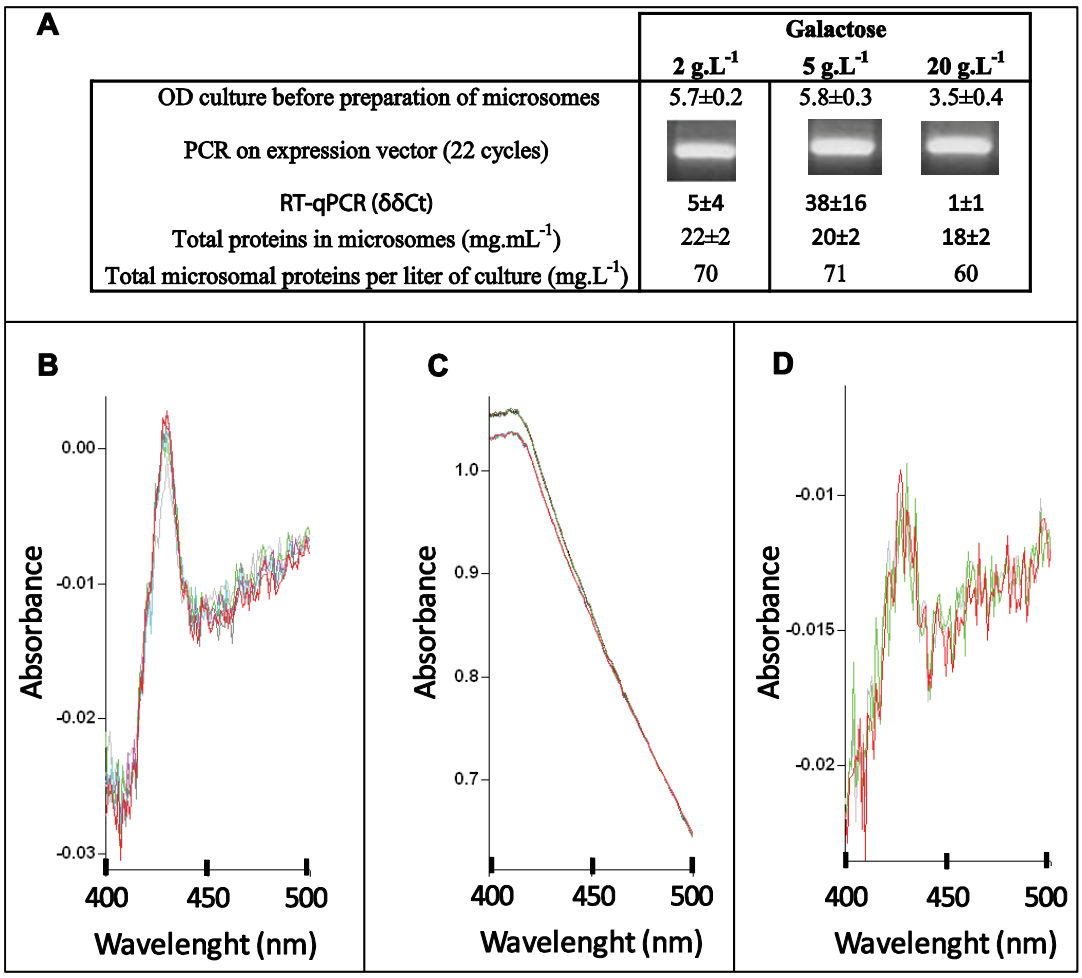


Figure 107. Optimization of CYP73A28 expression by decreasing concentrations in inductor (galactose).

A. Summary of the CYP73A28 induction tests.

B. CO differential spectrum obtained by induction of expression with 2 g.L^{-1} of galactose.

C. CO differential spectrum obtained by induction of expression with 5 g.L^{-1} of galactose.

D. CO differential spectrum obtained by induction of expression with 20 g.L^{-1} of galactose.

OD and RT-qPCR measurements were performed on three independent erlens of culture per construction. Cultures were sampled before microsomes preparation. Total protein concentration in microsomes was determined in three aliquots of microsomes.

Results obtained during these tests showed that (Figure 107 A):

- Cultures growth was restored by decreasing the concentration of inductor.
- Unexpectedly, RT-qPCR experiment indicated a higher transcription of *CYP73A28* whereas the inductor concentration was lower. The highest transcription rate was observed with the intermediate concentration 5 $g.L^{-1}$ of galactose.
- A slight increase in total microsomal proteins concentration was measured by decreasing the inductor concentration.
- More microsomes were obtained per liter of culture due to a better growth of the yeast.

Unfortunately, no significant improvement in the production of functional protein was obtained, still indicating problems of translation or stability of the protein (Figure 107 B, C, D). Again, for the highest transcription rates, differential CO spectrum was worst with even no peak at 420 nm. This seems to indicate trouble in protein translation or folding/stability.

No improvement in protein expression could be achieved. Nevertheless, *in vitro* assays of catalytic activity were carried out with all CYP73A microsomal preparation.

2.8. In vitro determination of catalytic activity

A previous study carried out with CYP73A1, a representative class I enzyme, demonstrated critical need of a carboxylic acid moiety for successful binding of the substrate in the catalytic site. A small size and planar structure of the substrates were also essential (Schalk *et al.*, 1997). However, class II enzymes have some important changes in their active site such as the HTP motif instead of RMA, which suggests potential shift in catalytic function. We considered in particular the possibility to metabolize 4-hydroxylated compounds thanks to a T-anchoring near the heme as mentioned above and the conversion of esters and of phenolamides that recently appeared as relevant intermediates in the phenolic metabolism. Accordingly, our tests included compounds with diverse groups in position 2, 3 and 4 of the phenolic ring, phenolamides, and various conjugates of the carboxylic acid.

Twenty-three different susbtrates were tried: cinnamate, cinnamoyl CoA, cinnamoyl shikimate, cinnamoyl tyramine, di-cinnamoyl putrescine, di-cinnamoyl spermidine, tri-cinnamoyl spermidine, *p*-coumarate, *o*-coumarate, *m*-coumarate, *p*-coumaroyl tyramine, *p*-coumaroyl agmatine, *p*-coumaroyl putrescine, tri-*p*-coumaroyl spermidine, cinnamoyl glucose, acetyl salicylic acid, salicylic acid, benzoic acid, phenylalanine, ethyl cinnamate, phenylpyruvate, dihydrocinnamate.

Qualitative tests were carried out as with CYP98As in the previous chapter. Non-limiting incubations were done, with a long incubation time of thirty minutes, and with an excess of NADPH. The same volume of microsomes was used in all assays.

Substrate conversion was classified in three levels: low, partial, total, as for the previous chapter.

A

	Class I	Class II	
	73A85	73A27	73A28
cinnamate	total	total	-
phenylpyruvate	low	low	-
dihydrocinnamate	partial	partial	-
ortho -coumarate	low	low	-
meta -coumarate	total	low	-

B

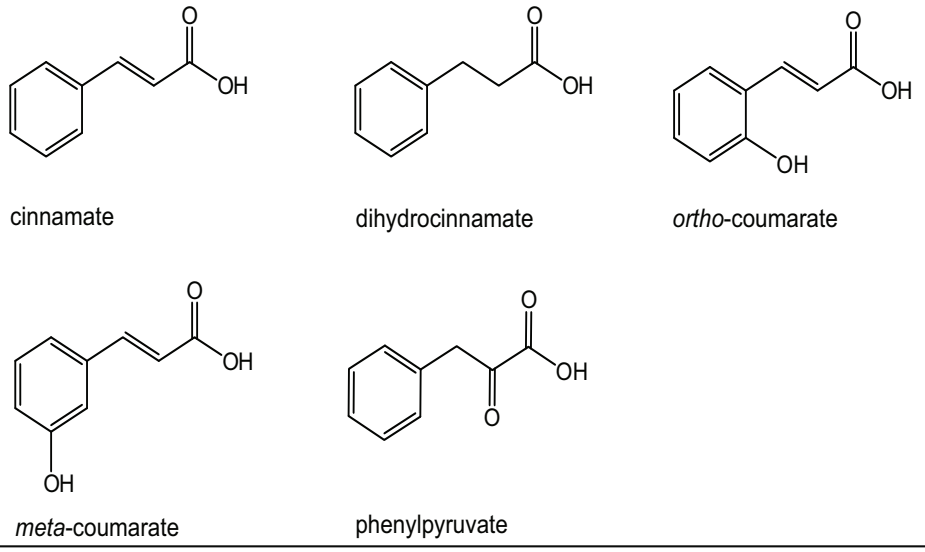


Figure 108. Substrates converted by tobacco CYP73As *in vitro*.
A. Substrate conversion.
B. Structure of the metabolized molecules.

No compounds (not even cinnamate) were metabolized by microsomes of yeast transformed with *CYP73A28* thus confirming that no functional protein was obtained. CYP73A85 (class I) and CYP73A27 (class II) metabolized the same five molecules with similar efficiencies (Figure 108) except for *meta*-coumarate wich appeared more actively converted into caffeate by CYP73A85.

No clear difference in substrate preference was observed between class I and class II CYP73As. No trace of enzyme activity was detected with CYP73A28. This raises the question of possible function loss after alloploidization and evolution into a pseudogene.

2.9. CYP73As over-expression on plants

Coding sequences of the three *CYP73As* were cloned downstream to a *CaMV-35S* promoter in a vector suitable for plant agrotransfection. CYP73As were transiently expressed in *Nicotiana benthamiana* leaf epidermal cells, and stably expressed in *Nicotiana tabacum Xanthi Nc* or *Arabidopsis thaliana* ecotype Columbia-0.

Our assumption was that over-expression of different P450s could lead to increase specific metabolites and perhaps decrease upstream precursors. P450 expression in the transformants was checked by semi-qPCR or RT-qPCR and metabolic profiling by UPLC-MS was carried out. Profiling of stable tobacco transformants was performed on young leaves and opened flowers: full scans of metabolites from 100 to 900 Da and single-ion recording (SIR) of specific molecules of interest such as cinnamate and *p*-coumarate. No metabolic profiling was done for *A. thaliana* transformants.

No significant difference was observed in full scan or targeted analyses between the different transformants (stable or transient transformants) and with the control. Moreover, no phenotypic modification was observed (Results not shown).

Stable plant transformation was performed using the fluorescent protein fusion constructs previously used for sub-cellular localizations. As shown above these constructs led to a strong protein expression in *N. benthamiana*. It was thus possible to check for the expression of the candidate proteins in the stable transformants. Free eGFP and *4CL1::eGFP* constructs were used as controls. Different organs were examined under fluorescence macroscope: roots, flowers, leaves, stems of *A. thaliana* overexpressors.

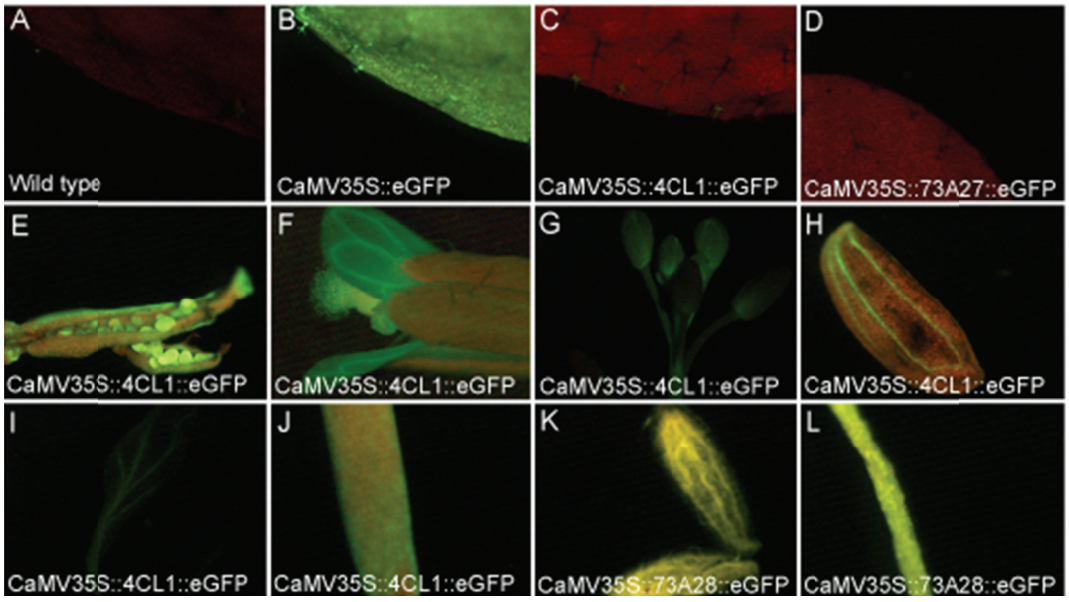


Figure 109. Stable *Arabidopsis thaliana* transformants examined by fluorescence macroscopy. A-D Leaf surface. E. Silique with seeds. F. Flower. G. Inflorescence. H. Sepal. I. Petal. J. Stem. K. Sepal. L. Root.
Observations with long-pass filter to detect GFP and chlorophyll signals. Same fluorescence patterns were observed with transformed tobacco plants.

In our observation conditions, wild type plants exhibited fluorescence in anther, pollen and trichome. Fluorescence was detected in all tissues in *eGFP* control plants. For the *CaMV-35S::4CL-1::eGFP* transformants, fluorescence was observed in all tissues, apart from leaves, but the signal was stronger in vessels and not uniform (Figure 109 C, E-J).

However, in plants transformed with the *CYP73A85::eGFP* and *CYP73A27::eGFP* constructs, no fluorescence was detected in most tissues. When fluorescence was detected, it was not found in all plant organs and observations were not reproducible between different lines. For *CYP73A28::eGFP* transformants, fluorescence was repeatedly observed but only observed in sepals and roots (Figure 109 K,L).

Similar expression patterns were noticed with *N. tabacum* stable over-expressors (Results not shown).

To conclude, over-expression of tobacco CYP73As did not cause any change in phenotype or metabolite content while transgenes were actively transcribed. In most of the cases, the CYP73 transcripts were not translated into protein. This could indicate a tight control of key steps of the phenylpropanoid pathway at the protein translation level.

3. Discussion

Most of the CYP73As characterized so far belong to the class I. But an increasing number of class II *CYP73A* genes were recently annoted, and it is now well established that two classes coexist in most plant genomes. Two class II (*CYP73A27* and *CYP73A28*) in addition to two class I (*CYP73A47* and *CYP73A85*) are present in the genome of *N. tabacum*. Thus tobacco is an interesting model to study the reasons for this gene duplication although polyploidization complicates the task.

Phylogenetic reconstruction shows that the divergence between class I and II results from an **early duplication in the CYP73A family** occurred before the divergence leading to Monocots and Dicots. Class I is present in all plant genomes but class II was lost in *Arabidopsis* and poplar. This would suggest redundancy and that class II is dispensable. However, the overall sequence identity between proteins of class I and II is quite low and indicates subfunctionalization. It is in addition striking to observe a **high conservation among class I proteins while class II is more variable**. This variability is spread all over the sequence but **some features are strongly conserved in class II proteins, including a long and S/T-rich N-ter signal peptide, a 3-4 amino acid insertion upstream of the meander region, an HXP motif in SRS5** which are universal in class II. Such conserved divergences, their properties and location in the protein structure, strongly suggest modifications in catalytic properties, but also different subcellular or membrane subdomain localization and interaction with different partner proteins.

Our subcellular localization analysis *in vivo* shows that in leaf epidermis **all the *N. tabacum* CYP73As are predominantly found anchored onto ER membranes**. It is however not excluded that this localization might vary in other plant tissues. In this line, a plastidial localization has been reported for C4H in grape berries using an immunogold labeling approach (Chen *et al.*, 2006). Due to the present limitations of confocal microscopy and intrinsic fluorescence of some plant organs such as anthers, this was however not possible to experimentally assess. Determination of the protein association to specific ER substructures is not a simple task and was not undertaken. Another interesting hypothesis for the atypical anchor of the class II CYP73s could be interaction of these P450s with other partner proteins.

Protein modeling, on the other hand, indicated that significant and conserved sequence divergences are located at the P450 surface and in the active site (SRS5). **Divergences on the P450 surface** can be considered as indicative of interaction with different partner proteins. This could include different PALs and 4-CLs, but also different redox partner(s). A quite striking similitude between the signal peptide of some P450 reductases such as *Arabidopsis* ATR2 and that from class II CYP73s suggests that interaction between them has to be considered. Modifications of amino acids in the active site, on the other hand, are supportive of a catalytic subfunctionalization or neofunctionalization. Nevertheless, until now, **we were not able to identify strickingly new or differentially metabolized substrates for class II CYP73As.** Only one significant difference was observed concerning the more active metabolism of *m*-coumarate by class I enzymes, although this has to be further validated by catalytic parameters determination. Efficient metabolism of cinnamate by class II CYP73A was unexpected, but already observed for the class II CYP73A15 from French bean (Nedelkina *et al.*, 1999). Hypothesis of neofunctionalization thus seems to be discarded while subfunctionalization still has to be further investigated using an enlarged set of potential substrates.

A clear difference between tobacco class I and II CYP73As genes transcription patterns upon both plant development and stress response was observed. **Class I genes are expressed in most plant organs and responsive to several stress conditions** such as TMV infection, salt, hydrogen peroxide, and mainly MeJA and cold treatments. **Class I expression is most often associated with a correlated upregulation of CYP98A31, HCT or CYP98A33**, although this is not a rule, and cold induces a correlated response of CYP73A47/85, CYP98A33 and HQT. Class II genes were so far most often associated with stress response such as fungal *Colletotrichum lindemuthianum* elicitor (Nedelkina *et al.*, 1999) and wound (Betz *et al.*, 2001). Expression patterns of tobacco **class II CYP73 genes however associates them quite clearly with flower and fruit development and quite specifically to salt stress response.** The class II genes in fact appear quite strongly down-regulated upon most other early (6h) stress responses.

No relevant information concerning enzyme subfunctionalization could be obtained from ectopic gene over-expression experiments. This could be due to the absence of protein expression in the transformants. Our work indeed revealed that while CYP73 genes were very actively transcribed in both tobacco and *Arabidopsis* stable transformants, protein was produced only in a few cell clusters in specific organ such roots and sepals. This expression pattern did not seem to correlate with any developmental or visible stress situation. **This could indicate tight regulation of the phenylpropanoid pathway at the level of protein translation**.

Other inconsistencies upon CYP73 genes over-expression have already been reported. For instance, changes in phenolic and flavonoid contents were observed, but were not predictable and did not correlated with transcript level when C4H over-expression was tested in tomato (Millar *et al.*, 2007). In tobacco, over-expression of C4H in leaves did not consistently result in increased accumulation of chlorogenic acid (Blount *et al.*, 2000), even if chlorogenic acid accumulation appeared to be a usual consequence of the phenylpropanoid metabolism engineering in Solanaceae (Shadle *et al.*, 2003). In the same way, over-expression of alfalfa CYP73 in tobacco was reported to result in a two-fold increase in activity, but in no effect on lignin composition. Meanwhile, its down-regulation (more than two-fold lower activity) resulted in reduced total lignin content (Sewalt *et al.*, 1997). Similarly, in grape, phenylpropanoid enzyme amounts were not correlated with the changes in enzymes activity (Chen *et al.*, 2006). As a whole, all the results obtained in this study demonstrate a highly plasticity and complex regulation of the phenolic pathway, at the transcription, the translation and the post-translational levels.

While no clear catalytic subfunctionalization could be demonstrated in this study, our data demonstrate that:

- *meta*-coumarate is a good substrate of CYP73A85 class I enzyme.
- two new physiologically relevant substrates of the CYP73s (phenylpyruvate and dihydrocinnamate) were identified. In one hand, huge number and amount of dihydro-hydroxycinnamate derivatives were found by metabolic profiling on tobacco organs. These derivatives were shown as important as protectant in potatoes and even humans (by *in vitro* approaches), but without any clearly defined function (Parr *et al.*, 2005; Hwang *et al.*, 2009).
  In the other hand, recently Tzin *et al.*, (2009) hypothezied that a cytochrome P450 could be involved in the synthesis of 4-hydroxyphenylpyruvate via 4- hydroxylation of phenylpyruvate. 4-hydroxyphenylpyruvate directly leads to homogentisate, precursor of γ-tocopherol and γ-tocotrienol.

Finally, CYP73As seem to provide a model of P450 subfamily where high structural divergences can maintain highly conserved function.

4. Prospects

Different approaches can be proposed to further clarify the function of class II CYP73s. Some of them are already considered for a near future.

In situ hybridization experiments are in progress to define more precisely the tissular localization of the *CYP73* transcripts. Thus, flowers and leaves tissues were already fixed and included in paraplast. Probes are also ready to use. It will be interesting to check if expression of one or all of the CYP73As can be found in the tapetum (nutritive tissue of the pollen), in the tegument and in the ovule. It will be also interesting to visualize the differential expression of paralogues in a same tissue to gain more information on their specific functions.

Sub-cellular localization of class II CYP73As is still not clear. New strategies could be used for further investigations, such as *in vitro* import with purified chloroplasts, chloroplast proteomic analysis, use of VIGS constructs to silence phytoene desaturase to bleach chlorophyll background of the chloroplast, or immunogold sub-cellular localization. This latter strategy seems to be a promising technique, not limited by the chlorophyll background, the contamination of RE in purified chloroplast or the restriction of the use of an *in vitro* method. But it would require to raise specific antibodies of each class of CYP73 which are not available so far.

To better evaluate possible enzyme subfunctionalization, we have to determine the catalytic parameters of the CYP73A85 (class I) and CYP73A27 (class II) which were correctly expressed in yeast. This is scheduled but requires good preparations of microsomes which is for the moment difficult to obtain. Plant transient expression seems to be more adapted than yeast expression with recalcitrant P450s. Membrane context and plant cellular machinery could support synthesis of high level of functional enzymes. In consequence, to overcome expression problems in yeast, the different enzymes have been expressed in *N. benthamiana* leaf epidermal cells for preparation of microsomal membranes and activity determination. In addition to the tests with known substrates, a library of fifty-four new aromatic molecules will be tested to determine catalytic activities of the two classes of enzymes. These new compounds have diverse groups on each positions of the phenolic ring, on the carbon chain, or different substituents of the carboxylic group. Of particular interest are the ring and side-chain substituents that are potentially interacting with SRS5.

To further determine paralogue functions, knock down mutants would be ideally suited. No insertion mutants are available for tobacco. Virus induced gene silencing (VIGS) was thus considered. But the coding sequences of CYP73As in each class appeared too similar to select sequences likely to specifically silence each group. In addition, complementation of the *ref3 A. thaliana* mutant (*CYP73A5*) by expressing class II *CYP73As* under the control of the *CYP73A5* promoter might be informative for demonstrating redundancy.

The lack of knock-out mutant collection for tobacco, and tobacco genome not publicly available was a limitation for our work. Shift to a model with sequenced genome and mutant access should help to solve the enigma of the class II gene function. This is the aim of a new project in the laboratory.

In conclusion, understanding the function of class II CYP73As will increase our knowledge on regulation, branching and evolution of the phenylpropanoid pathway, and pave the way for studies of metabolic fluxes partition in this pathway. Such knowledge will be essential for effective engineering of the phenypropanoid metabolism.

5. Materials and Methods

Samples of alive materials (plant, yeast) to analyze by RT-qPCR, PCR, or metabolomic, after taking away, were directly frozen in liquid nitrogen and stored at -80°C before grinding with liquid nitrogen and analysis. If several analyses have to be done for the same sample (mRNA extraction and metabolic extraction for example), the fine homogeneous powder obtained after crushing was divided into several Eppendorf tubes, one for each experiment to realize above.

5.1. Chemicals

Substrates used during this thesis were generally in solution in DMSO or methanol.

Cinnamate, *o*-coumarate, *m*-coumarate, *p*-coumarate, dihydrocinnamate, phenylpyruvate, salicylic acid, acetyl salicylic acid, benzoic acid, phenylalanine and ethyl cinnamate were commercially available from Sigma-Aldrich. Cinnamoyl shikimate and cinnamoyl glucose were kindly enzymatically synthesized by P. Ullmann (IBMP). Cinnamoyl CoA was kindly chemically prepared by P. Goeffroy (IBMP). Tri-cinnamoyl spermidine and tri-*p*-coumaroyl spermidine were chemically synthesized by M. Matsuno (IBMP) with the help of M. Schmitt. Cinnamoyl tyramine, *p*-coumaroyl tyramine, *p*-coumaroyl agmatine and *p*-coumaroyl putrescine were kindly provided by J. Negrel (INRA Dijon).

Di-cinnamoyl putrescine and N^1,N^{10}-di-cinnamoyl spermidine were prepared with the grateful help of M. Schmitt (Faculté de Pharmacie, UDS).

Di-cinnamoyl putrescine was synthesized in two steps:

1) Preparation of the chloride of cinnamate: 3g of cinnamate were dissolved in 40 mL of thionyle chloride ($SOCl_2$) with stirring at 110°C. After cooling of the solution, evaporation with a rotary evaporator (Rotavap) was carried out. The solution was washed with di-chloromethane (to remove HCl traces) and solvent removed from the sample by evaporation with Rotavap.
2) Preparation of the di-cinnamoyl putrescine: 3 nmole of putrescine were dissolved in 10 mL de di-chloromethane. After cooling the mixture in an ice bath with stirring, 7.5 nmole of amine tri-acetate were added. Then 1.1g cinnamate chloride was dissolved in 6 mL anhydrous di-chloromethane. This last solution was added drop by drop to the first one. The mixture was then filtered through a frit glass filter, and the obtained powder was characterized as di-cinnamoyl putrescine.

N^1,N^{10}-di-cinnamoyl spermidine was synthesized in six steps:

1) Preparation of cinnamate chloride: Done as described in the previous paragraph.

2) Protection of the N^1 and N^{10} of the spermidine: N-carbethoxy-phthalimide was dissolved in 50 mL of di-chloromethane and spermidine was dissolved in 25 mL of dichloromethane. The N-carbethoxy-phthalimide solution was added dropwise to the spermidine solution at 25°C. Dichloromethane was evaporated by Rotavap. The obtained mixture was washed with ethyl acetate and water. The solution was decanted, and the organic phase was recovered and dried with Na_2SO_4. Residual solvents were evaporated with Rotavap and the mixture was triturated with di-ethyl ether, and then filtered through a frit glass filter. N^1,N^{10}-Di-phthalimide spermidine was obtained.
3) Protection of the N^5 of the N^1,N^{10}-di-phtalimide spermidine: The last product was dissolved in 12 mL of dichloromethane. Tert-butoxycarbonyl (Boc) and 4-dimethylaminopyridine (DMAP) were dissolved 12 mL of dichloromethane. This last solution was added drop by drop to the first one. The solution was stirred overnight at room temperature. A yellowish solution was obtained, then whashed with water and with HCl 0.5 M (to remove DMAP). The solution was decanted and the organic phase recovered and dried with Na_2SO_4. Solution was evaporated with Rotavap. The obtained mixture was washed with ether then evaporated again with Rotavap. The intermediates: N^1,N^{10}-di-phtalimide N^5-Boc-spermidine was obtained.
4) Deprotection of the N^1 and N^{10} of the Boc-spermidine: 1.5 mmol of the previous intermediate was stirred with 20 mL of MeOH plus 6 mmol of hydrazine (N_2H_4) by heating at 95°C. The mix was cooled, and a precipitate appeared. This last one was filtered; the filtrate was then evaporated by Rotavap. The obtained intermediate was : N^5-Boc-spermidine.
5) Coupling of the last product with cinnamate chloride: N^5-Boc-spermidine was dissolved in 10 mL of dichloromethane. The solution was cooled at 0°C with agitation in ice, and then 240 µL amine tri-acetate were added. 432 mg of cinnamate chloride was dissolved in 6 mL of anhydrous dichloromethane. This last solution was added drop by drop to the first solution. Dichloromethane was evaporated with Rotavap. The mixture was resuspended in ethyl acetate. The solution was washed with sodium bicarbonate ($NaHCO_3$) to remove unreacted cinnamate chloride. The organic phase was recovered by decantation and washed with water and then HCl 0,5 M (to remove amine intermediates). The organic phase was recovered by decantation and dried with sodium sulphate (Na_2SO_4). The solution was filtered and evaporated.

   The product was purified by flash chromatography over silica. For that, the orangey oil was resuspended in ethyl acetate, silica was added, and the mixture was evaporated and then deposited over the column for flash chromatography. A gradient of heptane 100% to acetone 100% was carried out, elution of the product was followed by UV detector. The fractions containing the product of interest were pooled together and evaporated with Rotavap. The obtained intermediate was: N^1,N^{10}-di-cinnamoyl N^5-Boc-spermidine.

6) Deprotection of the N^1,N^{10}-di-cinnamoyl N^5-boc-spermidine: The last product was dissolved in 5 mL of dichloromethane, and 5 mL of tri-fluoroacetic acid was gently added with agitation. After 2 hours of stirring, dichloromethane and tri-fluoroacetic acid were evaporated with Rotavap. Ethyl ether was added and a new evaporation with Rotavap was done. This step was repeated twice to purify the product. The precipitate was dried and the final product was identified as N^1,N^{10}-di-cinnamoyl spermidine.

During the synthesis, the structures of the products and reaction intermediates were confirmed by ^{1}H- and ^{13}C-NMR analyses and by mass spectrometry.

5.2. Plant materials

Two cultivars of *Nicotiana tabacum* were used: Xanthi (Nc line) for expressions and metabolic studies and Xanthi (SH6 line) for the preparation of stable transformants.
In addition, *Nicotiana benthamiana* plants were used for transient expression of CYP73As in leaf epidermal cells.
Nicotiana sylvestris and *Nicotiana tomentosiformis* the ancestors of *Nicotiana tabacum* were grown to clone and identify their *CYP73As*.

Arabidopsis thaliana plants were also stably transformed for over-expression of CYP73As.

Culture conditions were already described (for *Arabidopsis*: § Plant material: Arabidopsis thaliana p.119 and, for tobacco § Nicotiana p.141).

5.3. Phylogenetic analyses

Phylogenetic analyses were achieved with peptidic sequences. The used sequences were retrieved from NCBI or phytozome. Sequences were treated with the program MEGA5 as previously described (Cf. § Phylogenetic analyses p.214).

Briefly, peptidic sequences were aligned with ClustalW algorithm implemented in MEGA5 and evolutionary analyses were conducted. The evolutionary history was inferred by using the Maximum Likelihood method based on the Whelan and Goldman + Freq. model. The bootstrap consensus tree was taken to represent the evolutionary history of the taxa analyzed. The final tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 54 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 456 positions in the final dataset.

5.4. CYP73A27 and CYP73A85 homology model generation

Homology models generation were carried out in collaboration with Dr. Fabrice Klein of the IGBMC (Strasbourg, France).

Peptidic sequences of CYP73A27 and CYP73A85 were blasted and aligned with Pipealign (http://bips.u-strasbg.fr/PipeAlign/) with sequences from protein databank. The sequences having a PPxP segment (assumed to be P450s) were kept. Twenty four sequences were found well aligned with the targets. CYP73As sequences were finally aligned on 2F9Q structure of CYP2D6 proved to be the best reference for homology modeling among the previous twenty-four sequences (according to protein sequences alignment).

3D models were generated with swiss-model (http://swissmodel.expasy.org/) from the aligned sequences. 3D models were calculated using ProModII and energy minimized using the Gromos force field. The quality of the resulting model was assessed using the ANOLEA mean force potential (See the swiss-model website for further information).

5.5. Cloning of *CYP73As* from *Nicotiana tabacum* ancestors

Genomic DNA was extracted from leaves with the kit Nucleospin Plant genomic extraction (Macherey-Nagel) according to the recommendations of the provider. Amplification PCR of the ancestors *CYP73As* (from *N. tomentosiformis* and *N. sylvestris*) were done with oligonucleotide primers designed from *N. tabacum CYP73As* sequences. Amplicons were purified on agarose gel, DNA was extracted (as previously mentioned § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109) and cloned into pBI vector. Finally, sequencing was achieved with oligonucleotide primers specific to the vector (SP6 and T7 primers).

5.6. Cloning of *CYP73A28, CYP73A27* and *CYP73A85* in pYeDP60 and pCAMBIA expression vectors

CYP73A27 was isolated starting from cDNA synthesized from flowers mRNA of *Nicotiana tabacum Xanthi Nc* and cloned in pYeDP60 and pCAMBIA vectors suitable for yeast over-expression and plant over-expression, respectively.

CYP73A28 and *CYP73A85* cDNA were kindly provided by J. Chapell (University of Kentucky, USA) and P. Bolwell (University of London , UK), respectively, and were cloned in pYeDP60 and pCAMBIA vectors.

5.6.1. pYeDP60 vector

This vector was already described in the chapter one (Cf. § Yeast expression vector p.85).

5.6.2. pCAMBIA2300U and pCAMBIA3300U vectors

These vectors were already described in the chapter one (Cf. § Plant expression vectors p.142).

5.6.3. PCR for isolation of CYP73A28 cDNA

This PCR was achieved with a high fidelity Taq to avoid mutations, and by primers with 5'-end extensions for the addition of restriction sites used during the next sub-cloning (PCR program in appendix p.312). The used polymerase was the Pfu Turbo Cx (Stratagene) supplemented by Taq polymerase (Invitrogen). Primers were used to add extension to the amplicons for next sub-cloning. Forward primer was designed with an extension recognized by *Bam*HI restriction enzyme and reverse primer an extension recognized by *Sac*I. These oligonucleotides were synthesized by Sigma-Aldrich. The obtained amplicons were purified by agarose gel electrophoresis, bands were cut and amplicons were extracted. These purified amplicons were then sub-cloned into TOPO-pCRII vectors.

5.6.4. TOPO TA Cloning

This cloning method and the vector TOPO-pCRII were already described in the chapter II (Cf. § TOPO TA Cloning p.216).

The purified PCR products were put in the presence of 1µL of TOPO- pCRII plasmid and of 1µL of saline solution. Reaction was achieved during 30 minutes at room temperature. Two microliters of the reaction volume were then used to transform competent *E. coli* bacteria for construction propagation. Transformed bacteria were screened by PCR colony, and sequencing was achieved to verify the cloned sequence.

5.6.5. "Traditional" cloning of CYP73A27, CYP73A28, CYP73A85 in pYeDP60

5.6.5.1. Preparation of the destination vector and of the insert

DNA digestions by a restriction enzyme were carried out according to the indications of the enzyme provider.

To avoid the re-circularization of the vector on itself after digestion and the problems associated to insert orientation, the cloning sites *Bam*HI and *Sac*I (added during the PCR), were chosen. These enzymes were previously checked to be compatible with the sequences of interest and to the destination vector.

The destination vector and the insert undergo a double digestion by the two suitable restriction enzymes. The digestion contains 1-2 µg of DNA, restriction buffer 1x and 1 unit of each restriction enzyme in 20 µL; the reaction was incubated one hour at 37°C.
The digested products were then deposited on 1.5 % agarose gel (Cf. § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109) to check the efficient digestion. The fragments of interest were extracted from the gel and ligation of the insert was done in the destination vector.

5.6.5.2. Ligation

This reaction as well as bacterial transformation, colony screening, extraction and purification of plasmids and sequencing were done as previously described (Cf. § Ligation p.217).

5.6.6. *USER™ cloning of CYP73A27, CYP73A28 and CYP73A85 in pCAMBIA vectors*

This technique was used to build the plasmids used for P450s over-expression in plants. P450 sequences were cloned in fusion with a fluorescent protein coding sequence (eGFP and mRFP) used as a reporter protein.

5.6.6.1. Preparation of the vectors

Initially, the pCAMBIA2300U or pCAMBIA3300U vectors were linearized (as previously described § Preparation of the vectors suitable for USER reaction p.143) and stored at -30°C until use.

5.6.6.2. Amplification PCR of the P450s and fluorescent proteins sequences

The various inserts were generated by PCR. The polymerase: Pfu Turbo Cx (Stratagene) was used for these PCR. The buffer used for PCR was the one provided by Stratagene.

The amplification PCR was carried out as previously explained (Cf. Amplification PCR p.144) and used primers are listed in appendix p.315-317.

5.6.6.3. Cloning of generated inserts into vectors

The PCR products were separated on agarose gel by electorphoresis (as mentioned previously, Cf. § Analysis of nucleic acids - Electrophoresis analysis on agarose gel p.109). The spots of interest were cut and extracted from the gel.

Amplicons recovered from the gel as well as the destination vectors were incubated with USER enzyme, in order to "create" the suitable fusion. For example *GFP* amplicons (generated for N-terminal fusion) were incubated with *CYP73A85* amplicons (generated for C-terminal fusion) and linearized pCAMBIA vector to build the construct: CYP73A85::GFP. Reaction was carried out as previously described (Cf. § Cloning of the inserts in the destination vectors p.144).

The constructs obtained were used to transform *E. coli* for propagation of the vectors (Cf. §. Bacteria transformation p.108). The obtained colonies were checked by PCR as mentioned previously (Cf. § Polymerase Chain Reaction over colony p.108). Transformed bacteria were grown during an overnight liquid culture at 37°C and plasmidic DNA was extracted and purified as mentioned previously (Cf. § Extraction and purification of plasmidic DNA p.106). For validation of the construct, sequencing was carried out.

5.7. Stable transformation of *Arabidopsis thaliana* (Clough & Bent, 1998)

5.7.1. Transformed Agrobacteria preparation

Transformations of Agrobacteria were done as previously mentioned (Cf. § A*grobacteria* transformation p.145). GV3101 strain was used to stably transform *A. thaliana* plant.

An overnight pre-culture of one agrobacteria colony was grown in 5 mL LB medium supplemented with adequate antibiotic(s) at 28°C with shaking at 180 rpm. This pre-culture was used to sow 200 mL of LB medium supplemented in antibiotic(s). This culture was incubated about 20 hours under shaking at 180 rpm and 28°C. Agrobacteria were centrifuged for 20 minutes at 5500 *g* and 4°C before be resuspended in a 5 % sucrose solution in a way to have a final OD_{600} equal to 1.

5.7.2. Arabidopsis thaliana transformation

Before transformation Silwet L-77 (OSI technologies) was added to the transformation solution in order to have 0.025% (v/v) final of Silwet L-77. Inflorescences were cleared of open flowers and siliques. Then, plants were dipped in the agrobacteria solution during 20 s with a gentle shaking. Plants were a little wring out and were placed in plastic box of about 11 Liters (to keep a high hygrometry which permits to increase the lifetime of the agrobacteria, favouring the transformation mediated by these bacteria). The boxes with the plants were put in dark during 24 hours after the transformation. After 3-4 days, the plastic boxes were gradually opened. The transformation step by dipping was repeated one week after the previous one to increase the rate of transformation.

Finally, when the siliques were well developed, the watering was stopped to favour the maturation of the seeds. And these last were harvested when the plants were well dry.

5.7.3. Selection of Arabidopsis thaliana transformants

Selections of transformed plants were done *in vitro* on plates.
Before sterilization, seeds were put at 4°C during at least 72 hours for stratification.

5.7.3.1. Seeds sterilization

Arabidopsis thaliana seeds were washed in a sterilization solution (25% (v/v) of sodium hypochlorite, 0.1% (w/v) of sodium dodecyl sulfate in sterile MilliQ water). The sterilization was done during 10 minutes. Then, this solution was discarded and replaced by a solution of 70% ethanol. The seeds were sterilized with this last solution during 5 minutes. Finally, the ethanol was through away and the seeds were washed five times in sterile MilliQ water, to remove trace of ethanol, SDS and sodium hypochlorite.
During this process, all the floating seeds were considered dead and were discarded during the remove of the sterilization or washing solutions.

After these treatments, the seeds were sterile and ready to sow in *in vitro* conditions on solid artificial culture medium.

5.7.3.2. *In vitro* culture medium

The artificial culture medium was a Murashige & Skoog medium supplemented with 6% (w/v) Pastagar. Adequate antibiotic(s) and herbicid(s) used for selection were also added to this medium. A bacteriostatic agent, cefotaxim (100 $mg.L^{-1}$) was used to kill the remaining agrobacteria used for the plant transformation.

Once the *in vitro* cultivated plants had enough grown, they were transplanted in soil and let to grow in culture room as previously explained.

5.8. Stable transformation of *Nicotiana tabacum* (adapted from Horsch *et al.*, 1985)

Transformations were done as explained in the chapter II (Cf. § Stable transformation of *Nicotiana tabacum* p.217). Leaf discs of *in vitro* cultivated tobacco were transformed by *Agrobacterium* harbouring the plasmid of interest and plants were regenerated from the transformed explants.

5.9. Transient expression in leaf epidermal cells of *Nicotiana benthamiana*

Transient expressions were done as described in the chapter II (Cf. § Transient expression in *Nicotiana benthamiana* leaf epidermal cells p.218).

When proteasome inhibitor (MG132) was tried, the leaves were cut 6 hours before analysis, and the leaves were immerged during this time in a solution with the inhibitor (100 µM in MES).

5.10. Relative expression quantification

5.10.1. Primers design

Primers were designed by Roche software: LightCycler Probe Design Software version 2. Couples of primers used for RT-qPCR were designed and tested as described in the chapter II (Cf. § Primers design p.219).

5.10.2. Complementary deoxyribonucleic acid (cDNA) preparation

For each experiment with plant and yeast materials, sampling was done very fast and taken samples were directly frozen in liquid nitrogen and store at -80°C.

After cryogenic grinding of the tissues using a mortar and pestle or with the Sartorius (Mikrodismenbrator) grinder, total RNA were extracted with the Plant total RNA extraction kit (Macherey-Nagel) according to the manufacturer.
Concentrations of total RNA were checked and quality of the RNA were verified as previously explained (Cf. § Complementary deoxyribonucleic acid (cDNA) preparation p.219).

cDNA were synthesized by retro-transcriptase (RT-PCR) as previously described (Cf. § Complementary deoxyribonucleic acid (cDNA) preparation p.219). The cDNA thus obtained were preserved at -30°C before future use by RT-qPCR. Prepared cDNA of one experiment were tested by PCR in the condition of sqPCR with the probes of the two housekeeping genes to check the quality of the cDNA.

5.10.3. RT-qPCR

After preparation of cDNA, the relative expressions of gene of interests were analyzed by quantitative real-time PCR. Reference genes, conditions of RT-qPCR, and calculations were done as depicted in the chapter II (Cf. § RT-qPCR p.220).

Reliability of the reactions was checked by verification of the amplification curve and the melting curve for each reaction. PCR efficiency was estimated for each reaction with LinRegPCR software. Efficiency was of 1.92±0.03 with all the various used primers and the whole experiments. Efficiency values were taken into account in all the calculations of relative expression rate.

Amplicons were analyzed by the comparative threshold cycle (Ct) method, by δδCt or δCt methods.

5.10.4. Semi-quantitative PCR (sqPCR) experiments

After preparation of cDNA, the relative expressions of genes of interest were sometimes analyzed by sqPCR. Primers designed for RT-qPCR were used in this case.

The principle of these experiments is based on the visual comparison of amplicons separated on agarose gel and visualized under UV light. Amplicons were standardized by comparison to a housekeeping gene. In this case, elongation factor subunit 1α (*EF1α*) was used as reference.

Semi-quantitative polymerase reactions were done as previously explained in the chapter II (Cf. § Semi-quantitative PCR (sqPCR) experiments p.221).

5.11. Fluorescent fusion constructions observations

5.11.1. Macroscopy observation

Whole plants were scrutinized with a Nikon binocular macroscope coupled with a UV lamp and filters suitable for GFP visualization. A CCD camera was used for image acquisitions.

5.11.2. Laser scanning confocal microscopy

Samples of leaves were cut and placed between slide and coverglass. Then a vacuum infiltration was done to have cells in the same shape and for infiltration of specific molecules used during the tests (Latrunculin B, ATP and GTP).

Cell imaging was performed using a LSM510 confocal laser scanning microscope equipped with an inverted Zeiss axiovert 100 M microscope (Carl Zeiss). For confocal resolution, images were taken using a 63x, 1.2 numerical aperture water immersion objective ("C-Apochromat"). Images were acquired using LSM510 version 2.8 software; they were processed with the latest version of the Zeiss LSM Image Browser software (Carl Zeiss) and exported as Tiff files.

5.12. Stress experiments

Stress experiments were done on two month-old *Nicotiana tabacum* Xanthi Nc plants grown in culture room.

All the treatments were achieved as previously described in the chapter II (Cf. § Stress treatments experiments p.221).

5.13. Yest microsomes preparation

Two *Saccharomyces cerevisiae* strains were used to express CYP73As from tobacco: WAT11 and WAT21 strains (already described in § Different used strains p.85). Culture conditions, yeast expression vector, chemical yeast preparation, yeast transformation, protein expression and microsomes preparation were unchanged in comparison to previous description in the chapter I (Cf. § Yeast microsome preparation p.87).

5.14. *In vitro* metabolization assays

These tests were carried out as described in the chapter I (Cf. § In vitro activity test p.93). After stop of the reaction, microsomal proteins were precipitated by addition of acetonitrile 40% (v/v) and centrifugation at 13 000 *g* during 10 minutes 4°C. 200 µL of the supernatant were stored at -30°C until analyses.This incubation medium was analysed by UPLC-MS/MS.

5.15. Methanolic extraction of soluble phenols

Methanolic extractions were carried out as described in the chapter II (Cf. Methanolic extraction of soluble phenols p.224).

5.16. Ultra Performance/Pressure Liquid Chromatography coupled with tandem Mass Spectrometry (UPLC-MS/MS) analyses

These analyses were carried out as previously described in the chapter II (Cf. § Ultra Performance/Pressure Liquid Chromatography coupled with tandem Mass Spectrometry analyses p.224) by Dimitri Heintz from the Metabolomic platform of IBMP.

5.17. Data analyses

Heatmaps were generated with the software MEV release 4.6.1. Spierman's rank correlation algorithm implemented in this software was used to draw gene trees.

General conclusion

« Savoir que l'on sait ce que l'on sait et que l'on ne sait pas ce que l'on ne sait pas, voilà le vrai savoir. » Confucius.

The molecules resulting from the secondary metabolism represent evolutionary adaptations of the plants to their environment, but they are also an extraordinary source of molecules of interest for Humans. The number of molecules which remain to be discovered is still huge, and understanding the mechanisms involved in their synthesis is primordial. The development of new approaches as metabolomics, proteomics, transcriptomics and deep sequencing is now boosting pathways discovery. Biotechnologies will enable production of these molecules, either in microorganisms or in plants.

This thesis was based on the study of enzymes involved in the upstream part of the phenylpropanoid pathway and leading to the most abundant phenolic derivatives such as monolignols, or phenolics esters and amides.

Results described in the first chapter, investigating formation of a lignin metabolon, pointed to unexpected features of the target enzymes, such as membrane binding of two soluble enzymes (4CL-1 and HCT) independently from the presence of their membrane partners. *In vivo* membrane relocalization of soluble enzymes in the presence of their partner P450 proteins and direct protein-protein interactions between the partners CYP73A5, CYP98A3 and 4CL-1 or HCT were also suggested. These interactions were enhanced when all the proteins were co-expressed in the same cells. This work also revealed P450-P450 homo- and hetero-oligomerization and suggests that CYP98A3 plays an essential role in the nucleation of the lignin metabolon. Taken together our data provide quite strong support to the formation of a lignin metabolon that is favored upon co-expression of multiple enzymes of the pathway.

Chapter two and three were devoted to the functional characterization of the paralogues of the CYP98 and CYP73 families in *Nicotiana tabacum*. A clear function assignment of each paralogue could not be achieved. Nevertheless, different expression patterns in healthy or stressed plants were observed, suggesting different or sometimes redundant functions in the plants. Notably, CYP73As from the clade II were mainly expressed in the flowers and fruits, and CYP98A30 (clade 2) was expressed only upon stress responses. The other members of two families were more constitutively expressed. Interestingly, CYP73As showed a very restricted substrate specificity, while CYP98As were capable to metabolize very diverse molecules from *p*-coumaroyl esters to a range of *p*-coumaroyl-phenolamides. The latters result from the activity of different hydroxycinnamoyl transferases linking hydroxycinnamoyl residues to diverse amine backbones (putrescine, spermidine, tyramine…) to form conjugates that were regarded as final products. But recently, several reports (Morant *et al.*, 2007; Matsuno *et al.*, 2009; Fellenberg *et al.*, 2009) indicated that the phenolamide phenolic ring can be hydroxylated by specific CYP98A enzymes. We showed that this capability is apparently shared by all CYP98As.

This thesis points to our still poor understanding and complexity of the phenylpropanoid pathway with an increasing number of potential ramifications and redundancies. Our data together with previous reports demonstrate its high plasticity and complex regulation, at transcriptional, translational and post-translational levels. While formation of extremely dynamic metabolons is most likely an important factor that controls fluxes in the pathway, some important components of such metabolons most likely remain to be identified.

Pour conclure, l'étude de protéines aussi versatiles que les cytochromes P450s peut se révéler tout aussi passionnante que frustrante. Ces travaux nous ont en effet permis d'aborder plusieurs facettes de plusieurs sujets complexes, en mettant en œuvre des techniques variées de biochimie, de biologie moléculaire, de métabolomique et d'imagerie. Nous avons ainsi pu démontrer dans le premier projet l'existence d'interactions protéine-membrane et protéine-protéine pour la formation de métabolon dans la voie des phénylpropanoïdes. Les résultats des deux autres projets posent les premières pierres pour des recherches ultérieures sur la signification de l'existence de deux clades au sein des familles CYP98A et CYP73A, dans des familles végétales autres que les Brassicacées. Ils démontrent également (s'il y en avait encore besoin) que cette voie de biosynthèse est beaucoup plus complexe que ce qui était pensé il y a encore peu d'années.

En souhaitant encore longue vie à ces thématiques....

Merci,

JEB

References

Abdulrazzak N., Pollet B., Ehlting J., Larsen K., Asnaghi C., Ronseau S., Proux C., Erhardt M., Seltzer V., Renou J.P., Ullmann P., Pauly M., Lapierre C., Werck-Reichhart D. (2006) A coumaroyl-ester-3-hydroxylase insertion mutant reveals the existence of nonredundant meta-hydroxylation pathways and essential roles for phenolic precursors in cell expansion and plant growth. Plant Physiol. 140(1): 30-48.

Abhilash P.C., Jamil S., Singh N. (2009) Transgenic plants for enhanced biodegradation and phytoremediation of organic xenobiotics. Biotechnol Adv. 27(4): 474-488.

Achnine L., Blancaflor E.B., Rasmussen S., Dixon, R.A. (2004) Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. Plant Cell. 16: 3098-3109.

Aerts R.J., Baumann T.W. (1994) Distribution and utilization of chlorogenic acid in *Coffea* seedlings. J Exp Bot. 45(4): 497-503.

Aharoni A., Jongsma M.A., Bouwmeester H.J. (2005) Volatile science? Metabolic engineering of terpenoids in plants. Trends Plant Sci. 10(12): 594-602.

Ahlers F., Thom I., Lambert J., Kuckuk R., Wiermann R. (1999) ^{1}H NMR analysis of sporopollenin from *Typha angustifolia*. Phytochemistry 50(6): 1095-1098.

Ajikumar P.K., Xiao W.H., Tyo K.E., Wang Y., Simeon F., Leonard E., Mucha O., Phon T.H., Pfeifer B., Stephanopoulos G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. Science. 330(6000): 70-74.

Alcázar R., Altabella T., Marco F., Bortolotti C., Reymond M., Koncz C., Carrasco P., Tiburcio A.F. (2010) Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. Planta. 231(6): 1237-1249.

Anterola A.M., Lewis N.G. (2002) Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry. 61(3): 221-294.

Appaix F., Kuznetsov A.V., Usson Y., Kay L., Andrienko T., Olivares J., Kaambre T., Sikk P., Margreiter R., Saks V. (2003) Possible role of cytoskeleton in intracellular arrangement and regulation of mitochondria. Exp Physiol. 88(1): 175-190.

Bak S., Kahn R.A., Nielsen H.L., Møller B.L., Halkier B.A. (1998) Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor (L.) Moench* by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. Plant Mol Biol. 36(3): 393-405.

Barnes S., Prasain J. (2005) Current progress in the use of traditional medicines and nutraceuticals. Curr Opin Plant Biol. 8: 324-328.

Batard Y., Hehn A., Nedelkina S., Schalk M., Pallett K., Schaller H., Werck-Reichhart D. (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. Arch Biochem Biophys. 379(1): 161-169.

Batard Y., Schalk M., Pierrel M.A., Zimmerlin A., Durst F., Werck-Reichhart D. (1997) Regulation of the Cinnamate 4-Hydroxylase (CYP73A1) in Jerusalem Artichoke Tubers in Response to Wounding and Chemical Treatments. Plant Physiol. 113(3): 951-959.

Bayburt T.H., Sligar S.G. (2002) Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. Proc Natl Acad Sci U S A. 99(10): 6725-6730.

Bayburt T.H., Sligar S.G. (2010) Membrane protein assembly into Nanodiscs. FEBS Lett. 584(9): 1721-1727.

Bazzano L.A., He J., Ogden L.G., Loria C.M., Vupputuri S., Myers L., Whelton .PK. (2002) Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. Am J Clin Nutr. 76(1): 93-99.

Bell-Lelong D.A., Cusumano J.C., Meyer K., Chapple C. (1997) Cinnamate-4-hydroxylase expression in *Arabidopsis*. Regulation in response to development and the environment. Plant Physiol. 113(3): 729-738.

Bergmans H.E., van Die I.M., Hoekstra W.P. (1981) Transformation in *Escherichia coli*: stages in the process. J Bacteriol. 146(2): 564-570.

Bernards M.A., Razem F.A. (2001) The poly(phenolic) domain of potato suberin: a non-lignin cell wall bio-polymer. Phytochemistry. 57(7): 1115-1122.

Bernhardt R. (2006) Cytochromes P450 as versatile biocatalysts. J Biotechnol. 124(1): 128-145.

Berwanger A., Eyrisch S., Schuster I., Helms V., Bernhardt R. (2010) Polyamines: naturally occurring small molecule modulators of electrostatic protein-protein interactions. J Inorg Biochem. 104(2): 118-125.

Besseau S., Hoffmann L., Geoffroy P., Lapierre C., Pollet B., Legrand M. (2007) Flavonoid accumulation in *arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. Plant Cell 19: 148-162.

Betz C., McCollum T.G., Mayer R.T. (2001) Differential expression of two cinnamate 4-hydroxylase genes in 'Valencia' orange (*Citrus sinensis* Osbeck). Plant Mol Biol. 46(6): 741-748.

Birnboim H.C., Doly J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7(6): 1513-1523.

Blee K., Choi J.W., O'Connell A.P., Jupe S.C., Schuch W., Lewis N.G., Bolwell G.P. (2001) Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco. Phytochemistry. 57(7): 1159-1166.

Blount J.W., Korth K.L., Masoud S.A., Rasmussen S., Lamb C., Dixon R.A. (2000) Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway. Plant Physiol. 122(1): 107-116.

Boerjan W, Ralph J, Baucher M. (2003) Lignin biosynthesis. Annu Rev Plant Biol. 54: 519-546.

Bolte S., Cordelières F.P. (2006) A guided tour into subcellular colocalization analysis in light microscopy. Journal of Microscopy, Vol. 224, Pt 3, pp. 213-232S.

Bolwell G. P. (1988) Synthesis of cell wall components: aspects of control. Phytochemistry 27: 1235-1253.

Borch J., Hamann T. (2009) The nanodisc: a novel tool for membrane protein studies. Biol Chem. 390: 805-814.

Boudet A.-M. (2007) Evolution and current status of research in phenolic compounds. Phytochemistry. 68(22-24): 2722–2735

Boudet A.M., Kajita S., Grima-Pettenati1 J., Goffner D. (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. Trends Plant Sci. 8(12): 576-581

Bowles D., Isayenkova J., Lim E.K., Poppenberger B. (2005) Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol. 8(3): 254-263.

Bozak K.R., O'keefe D.P., Christoffersen R.E. (1992) Expression of a Ripening-Related Avocado (*Persea americana*) Cytochrome P450 in Yeast. Plant Physiol. 100(4): 1976-1981.

Bradford M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 72: 248-254.

Brandner K., Sambade A., Boutant E., Didier P., Mély Y., Ritzenthaler C., Manfred Heinlein M. (2008) Tobacco Mosaic Virus Movement Protein Interacts with Green Fluorescent Protein-Tagged Microtubule End-Binding Protein 1. Plant Physiol. 147: 611-623.

Brodie B.B., Axelrod J., Cooper J.R., Gaudette L., La Du B.N., Mitoma C., Udenfriend S. (1955) Detoxication of Drugs and Other Foreign Compounds by Liver Microsomes. Science. 121: 603-604.

Brugliera F., Barri-Rewell G., Holton T.A., Mason J.G. (1999) Isolation and characterization of a flavonoid 3'-hydroxylase cDNA clone corresponding to the Ht1 locus of Petunia hybrida. Plant J. 19(4): 441-451.

Brunsveld L., Waldmann H., Huster D. (2009) Membrane binding of lipidated Ras peptides and proteins--the structural point of view. Biochim Biophys Acta. 1788(1): 273-288.

Buchanan B.B., Gruissem W., Jones R.L. (2001) Biochemistry & molecular biology of plants. Edited by Buchanan, B. B., Gruissem, W., Jones, R. L. American Society of Plant Physiologists, Rockville, Maryland US.

Bureik M., Schiffler B., Hiraoka Y., Vogel F., Bernhardt R. (2002) Functional expression of human mitochondrial CYP11B2 in fission yeast and identification of a new internal electron transfer protein, etp1. Biochemistry. 41(7): 2311-2321.

Cabello-Hurtado F., Batard Y., Salaün J.P., Durst F., Pinot F., Werck-Reichhart D. (1998) Cloning, expression in yeast, and functional characterization of CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation of fatty acids. J Biol Chem. 273(13): 7260-7267.

Cavalier-Smith T. (2009) Predation and eukaryote cell origins: a coevolutionary perspective. Int J Biochem Cell Biol. 41(2): 307-722.

Chang Z., Li L., Pan Z., Wang X. (2008) Crystallization and preliminary X-ray analysis of allene oxide synthase, cytochrome P450 CYP74A2, from Parthenium argentatum. Acta Crystallogr Sect F Struct Biol Cryst Commun. 64(Pt 7): 668-670.

Chapple C.C., Vogt T., Ellis B.E., Somerville C.R. (1992) An Arabidopsis mutant defective in the general phenylpropanoid pathway. Plant Cell. 4(11): 1413-24.

Chen H., Jiang H., Morgan J.A. (2007) Non-natural cinnamic acid derivatives as substrates of cinnamate 4-hydroxylase. Phytochemistry. 68(3): 306-311.

Chen J.-C., Wen P.-F., Kong W.-F., Pan Q.-H., Wan S.-B., Huang W.-D. (2006) Changes and subcellular localizations of the enzymes involved in phenylpropanoid metabolism during grape berry development. J. Plant Physiol. 163(2): 115-127.

Cheng S.H., Sheen J., Gerrish C., Bolwell G.P. (2001) Molecular identification of phenylalanine ammonia-lyase as a substrate of a specific constitutively active Arabidopsis CDPK expressed in maize protoplasts. Febs Lett. 503: 185-188.

Chuong S.D., Good A.G., Taylor G.J., Freeman M.C., Moorhead G.B., Muench D.G. (2004) Large-scale identification of tubulin-binding proteins provides insight on subcellular trafficking, metabolic channeling, and signaling in plant cells. Mol Cell Proteomics. 3(10): 970-983.

Clauss K., Baumert A., Nimtz M., Milkowski C., Strack D. (2008) Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae. Plant J. 53(5): 802-813.

Clough S.J., Bent A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 16(6): 735-743.

Coleman H.D., Park J.Y., Nair R., Chapple C., Mansfield S.D. (2008) RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. Proc Natl Acad Sci U S A. 105(11): 4501-4506.

Compagnon V., Diehl P., Benveniste I., Meyer D., Schaller H., Schreiber L., Franke R., Pinot F. (2009) CYP86B1 is required for very long chain omega-hydroxyacid and alpha, omega -dicarboxylic acid synthesis in root and seed suberin polyester. Plant Physiol. 150(4): 1831-1843.

Cooper D.Y., Levin S., Narasimhulu S., Rosenthal O. (1965) Photochimical action spectrum of the oxidase of mixed function oxidase systems. Science. 147: 400-402.

Costa M.A., Bedgar D.L., Moinuddin S.G., Kim K.W., Cardenas C.L., Cochrane F.C., Shockey J.M., Helms G.L., Amakura Y., Takahashi H., Milhollan J.K., Davin L.B., Browse J., Lewis N.G. (2005) Characterization in vitro and in vivo of the putative multigene 4-coumarate:CoA ligase network in *Arabidopsis*: syringyl lignin and sinapate/sinapyl alcohol derivative formation. Phytochemistry. 66: 2072-2091.

Costa M.A., Collins R.E., Anterola A.M., Cochrane F.C., Davin L.B., Lewis N.G. (2003) An in silico assessment of gene function and organization of the phenylpropanoid pathway metabolic networks in *Arabidopsis thaliana* and limitations thereof. Phytochemistry. 64(6): 1097-1112.

Crozier A., Jaganath I.B., Clifford M.N. (2009) Dietary phenolics: chemistry, bioavailability and effects on health. Nat Prod Rep. 26(8): 965-1096.

Czichi U., Kindl H. (1975). Formation of *p*-coumaric acid and *o*-coumaric acid from L-phenylalanine by microsomal membranefractions from potato: Evidence of membrane-bound enzyme complexes. Planta. 125: 115-125.

Czichi U., Kindl H. (1977). Phenylalanine ammonia-lyase and cinnamic acid hydroxylase as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution. Planta. 134: 133-143.

D'Auria, J.C. (2006) Acyltransferases in plants: a good time to be BAHD. Curr Opin Plant Biol. 9(3): 331-340.

D'Auria J.C., Gershenzon J. (2005) The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. Curr Opin Plant Biol. 8(3): 308-316.

D'Auria J.C., Reichelt M., Luck K., Svatos A., Gershenzon J. (2007a) Identification and characterization of the BAHD acyltransferase malonyl CoA: anthocyanidin 5-O-glucoside-6"-O-malonyltransferase (At5MAT) in *Arabidopsis thaliana*. FEBS Lett. 581(5): 872-878.

D'Auria J.C., Pichersky E., Schaub A., Hansel A., Gershenzon J. (2007b) Characterization of a BAHD acyltransferase responsible for producing the green leaf volatile (Z)-3-hexen-1-yl acetate in *Arabidopsis thaliana*. Plant J. 49(2): 194-207.

Davydov D.R. (2001) Microsomal monooxygenase in apoptosis: another target for cytochrome c signaling? Trends Biochem Sci. 26(3): 155-160.

Déjardin A., Laurans F., Arnaud D., Breton C., Pilate G., Leplé J.-C. (2010) Wood formation in Angiosperms. C. R. Biologies 333: 325-334

Didierjean L., Gondet L., Perkins R., Lau S.M., Schaller H., O'Keefe D.P., Werck-Reichhart D. (2002) Engineering herbicide metabolism in tobacco and *Arabidopsis* with CYP76B1, a cytochrome P450 enzyme from Jerusalem artichoke. Plant Physiol. 130(1): 179-189.

Dixon D.P., Lapthorn A., Edwards R. (2002) Plant glutathione transferases. Genome Biol. 3(3): REVIEWS3004.

Dixon R.A., Achnine L., Kota P., Liu C.J., Reddy M.S., Wang L. (2002) The phenylpropanoid pathway and plant defence-a genomics perspective. Mol Plant Pathol. 3(5): 371-390.

Dixon R.A., Chen F., Guo D., Parvathi K. (2001). The biosynthesis of monolignols: a « metabolic grid », or independent pathways to guaiacyl and synringyl units ? Phytochem. 57: 1069-1084.

Dixon R.A., Paiva N.L. (1995) Stress-Induced Phenylpropanoid Metabolism. Plant Cell. 7: 1085-1097.

Domínguez E., Mercado J.A., Quesada M.A., Heredia A. (1999) Pollen sporopollenin: degradation and structural elucidation. Sex Plant Reprod. 12: 171-178.

Douglas C.J. (1996) Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. Trends Plant Sci. 1: 171-178.

Duan H., Civjan N.R., Sligar S.G., Schuler MA. (2004) Co-incorporation of heterologously expressed *Arabidopsis* cytochrome P450 and P450 reductase into soluble nanoscale lipid bilayers. Arch Biochem Biophys. 424(2): 141-153.

Dudareva N., D'Auria J.C., Nam K.H., Raguso R.A., Pichersky E. (1998) Acetyl-CoA:benzylalcohol acetyltransferase — an enzyme involved in floral scent production in *Clarkia breweri*. Plant J. 14(3): 297-304.

Dudareva N., Pichersky E., Gershenzon J. (2004) Biochemistry of plant volatiles. Plant Physiol. 135: 1893-1902.

Eberle D., Ullmann P., Werck-Reichhart D., Petersen M. (2009) cDNA cloning and functional characterisation of CYP98A14 and NADPH:cytochrome P450 reductase from *Coleus blumei* involved in rosmarinic acid biosynthesis. Plant Mol Biol. 69: 239-253.

Edreva A.M., Velikova V.B., Tsonev T.D. (2007) Phenylamides in plants. Russian J. Plant Physiol. 54: 287-301.

Ehlting J., Büttner D., Wang Q., Douglas C. J., Somssich I. E., Kombrink E. (1999) Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. Plant J. 19: 9-20.

Ehlting J., Hamberger B., Rachel Million-Rousseau R., Werck-Reichhart D. (2006) Cytochromes P450 in phenolic metabolism Phytochem Rev. 5: 239-270.

Ehlting J., Mattheus N., Aeschliman D.S., Li E., Hamberger B., Cullis I.F., Zhuang J., Kaneda M., Mansfield S.D., Samuels L., Ritland K., Ellis B.E., Bohlmann J., Douglas C.J. (2005) Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. Plant J. 42(5): 618-640.

Ehlting J., Sauveplane V., Olry A., Ginglinger J.F., Provart N.J., Werck-Reichhart D. (2008) An extensive (co-)expression analysis tool for the cytochrome P450 superfamily in *Arabidopsis thaliana.* BMC Plant Biol. 8:47.

El Amrani A., Barakate A., Askari B.M., Li X., Roberts A.G., Ryan M.D., Halpin C. (2004) Coordinate Expression and Independent Subcellular Targeting of Multiple Proteins from a Single Transgene. Plant Physiol. 135: 16-24.

Ellis R.J. (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci. 26(10): 597-604.

Fahrendorf T., 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. Arch Biochem Biophys. 305(2): 509-515.

Fellenberg C., Böttcher C., Vogt T. (2009) Phenylpropanoid polyamine conjugate biosynthesis in Arabidopsis thaliana flower buds. Phytochemistry. 70(11-12): 1392-400.

Fellenberg C., Milkowski C., Hause B., Lange P.R., Böttcher C., Schmidt J., Vogt T. (2008) Tapetum-specific location of a cation-dependent *O*-methyltransferase in *Arabidopsis thaliana*. Plant J. 56(1): 132-145.

Fleuriet A., Macheix J.-J. (2003) Phenolic acids in fruits and vegetables. Flavonoids in health and disease. Second edition. Rice-Evans C. A., Packer L. Marcel Dekker. Los Angeles, USA.

Franke R., McMichael C.M., Meyer K., Shirley A.M., Cusumano J.C., Chapple C. (2000) Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. Plant J. 22(3): 223-234.

Franke R., Humphreys J.M., Hemm M.R., Denault J.W., Ruegger M.O., Cusumano J.C., Chapple C. (2002a) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J. 30(1): 33-45.

Franke R., Hemm M.R., Denault J.W., Ruegger M.O., Humphreys J.M., Chapple C. (2002b) Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis. Plant J. 30(1): 47-59.

Frear D.S., Swanson H.R., Tanaka F.S. (1969) ***N***-demethylation of substituted 3-(phenyl)-1-methylureas: Isolation and characterization of a microsomal mixed function oxidase from cotton. Phytochemistry. 8: 2157-2169.

Fritz C., Palacios-Rojas N., Feil R., Stitt M. (2006) Regulation of secondary metabolism by the carbon nitrogen status in tobacco: nitrate inhibits large sectors of phenylpropanoid metabolism. Plant J. 46: 533-548

Froehlich J.E., Itoh A., Howe G.A. (2001) Tomato allene oxide synthase and fatty acid hydroperoxide lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes of chloroplast envelope. Plant Physiol. 125(1): 306-317.

Frydman A., Weisshaus O., Bar-Peled M., Huhman D.V., Sumner L.W., Marin F.R., Lewinsohn E., Fluhr R., Gressel J., Eyal Y. (2004) Citrus fruit bitter flavors: isolation and functional characterization of the gene Cm1,2RhaT encoding a 1,2 rhamnosyltransferase, a key enzyme in the biosynthesis of the bitter flavonoids of citrus. Plant J. 40(1): 88-100.

Fujino T., Sone Y., Mitsuishi Y., Itoh T. (2000) Characterization of cross-links between cellulose microfibrils, and their occurrence during elongation growth in pea epicotyl. Plant Cell Physiol. 41(4): 4864-94.

Fujiwara H., Tanaka Y., Fukui Y., Nakao M., Ashikari T., Kusumi T. (1997) Anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. Purification, characterization and its role in anthocyanin biosynthesis. Eur J Biochem. 249(1): 45-51.

Gabriac B., Werck-Reichhart D., Teutsch H., Durst F. (1991) Purification and immunocharacterization of a plant cytochrome P450: the cinnamic acid 4-hydroxylase. Arch Biochem Biophys. 288(1): 302-309.

Gachotte D., Eckstein J., Barbuch R., Hughes T., Roberts C., Bard M. (2001) A novel gene conserved from yeast to humans is involved in sterol biosynthesis. J Lipid Res. 42(1): 150-154.

Gang D.R. (2005) Evolution of flavors and scents. Annu Rev Plant Biol. 56: 301-325.

Gang D.R., Beuerle T., Ullmann P., Werck-Reichhart D., Pichersky E. (2002) Differential production of *meta* hydroxylated phenylpropanoids in sweet basil peltate glandular trichomes and leaves is controlled by the activities of specific acyltransferases and hydroxylases. Plant Physiol. 130(3): 1536-1544.

Garfinkel D. (1958) Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. Arch Biochem Biophys.77(2): 493-509.

Geu-Flores F., Nour-Eldin H.H., Nielsen M.T., Halkier B.A. (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 35(7): e55.

Geu-Flores F., Olsen C.E., Halkier B.A. (2009) Towards engineering glucosinolates into non-cruciferous plants. Planta. 229(2): 261-270.

Gietz R.D., St. Jean A., Woods R.A., Schiestl H. (1992) Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20(6): 1425

Geoffroy P., Legrand M., Fritig B. (1990) Isolation and characterization of a proteinaceous inhibitor of microbial proteinases induced during the hypersensitive reaction of tobacco to tobacco mosaic virus. Mol Plant Microbe Interact. 3(5): 327-333.

Graça J. (2009) Hydroxycinnamates in suberin formation, Phytochem. Rev. 9: 85-91.

Grace S.C., Logan B.A. (2000) Energy dissipation and radical scavenging by the plant phenylpropanoid pathway. Phil Trans R Soc Lond B. (355): 1499-1510.

Grace S.C., Salgo M.G., Pryor W.A. (1998) Scavenging of peroxynitrite by a phenolic/peroxidase system prevents oxidative damage to DNA. FEBS Lett. 426(1): 24-28.

Graciet E., Lebreton S., Gontero B. (2004) Emergence of new regulatory mechanisms in the Benson-Calvin pathway via protein-protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex. J Exp Bot. 55(400): 1245-1254.

Graham J.W., Williams T.C., Morgan M., Fernie A.R., Ratcliffe R.G., Sweetlove L.J. (2007) Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. Plant Cell. 19(11): 3723-3738.

Gravot A., Larbat R., Hehn A., Lièvre K., Gontier E., Goergen J.L., 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. Arch Biochem Biophys. 422(1): 71-80.

Grienenberger E., Besseau S., Geoffroy P., Debayle D., Heintz D., Lapierre C., Pollet B., Heitz T., Legrand M. (2009) A BAHD acyltransferase is expressed in the tapetum of *Arabidopsis* anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. Plant J. 58: 246-259.

Griffing L.R (2010) Networking in the endoplasmic reticulum. Biochem. Soc. Trans. 38: 747-753

Guarente L., Yocum R.R., Gifford P. (1982) A GAL10-CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site. Proc Natl Acad Sci U S A. 79(23): 7410-7414.

Guilford W.J., Schneider D.M., Labovitz J., Opella S.J.. (1988) High resolution solid state C NMR spectroscopy of sporopollenins from different plant taxa. Plant Physiol. 86(1): 134-136.

Hamberger B., Hahlbrock K. (2004) The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. Proc Natl Acad Sci U S A. 101: 2209-2214.

Hartmann T. (2007) From waste products to ecochemicals: 50 years research on plant secondary metabolism. Phytochemistry. 68: 2831–2846.

Hazai E., Bikádi Z., Simonyi M., Kupfer D. (2005) Association of cytochrome P450 enzymes is a determining factor in their catalytic activity. J Comput Aided Mol Des. 19(4): 271-285.

Hehn A., Morant M., Werck-Reichhart D. (2002) Partial recoding of P450 and P450 reductase cDNAs for improved expression in yeast and plants. Methods Enzymol. 357: 343-351.

Heitz T., Bergey D.R., Ryan C.A. (1997) A gene encoding a chloroplast-targeted lipoxygenase in tomato leaves is transiently induced by wounding, systemin, and methyl jasmonate. Plant Physiol. 114(3): 1085-1093.

Heller W., Kühnl T. (1985) Elicitor induction of a microsomal 5-O-(4-coumaroyl)shikimate 3'-hydroxylase in parsley cell suspension cultures. Arch Biochem Biophys. 241(2): 453-460.

Helliwell C.A., Sheldon C.C., Olive M.R., Walker A.R., Zeevaart J.A.D., Peacock W.J., Dennis E.S. (1998) Cloning of the *Arabidopsisent*-kaurene oxidase gene *GA3*. Proc Natl Acad Sci U S A. 95(15): 9019-9024

Helliwell C.A., Sullivan J.A., Mould R.M., Gray J.C., Peacock W.J., Dennis E.S. (2001) A plastid envelope location of *Arabidopsis* ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. Plant J. 28: 201-208.

Helmstaedt K., Strittmatter A., Lipscomb W.N., Braus G.H. (2005) Evolution of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase-encoding genes in the yeast *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 102(28): 9784-9789.

Hemm M.R., Rider S.D., Ogas J., Murry D.J., Chapple C. (2004) Light induces phenylpropanoid metabolism in *Arabidopsis* roots. Plant J. 38(5): 765-778.

Herrmann K. (1989) Occurrence and content of hydroxycinnamic and hydroxybenzoic acid compounds in foods. Crit Rev Food Sci Nutr. 28(4): 315-347.

Hoffmann L., Besseau S., Geoffroy P., Ritzenthaler C., Meyer D., Lapierre C., Pollet B., Legrand M. (2004) Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. Plant Cell. 16(6): 1446-1465.

Hoffmann L., Maury S., Martz F., Geoffroy P., Legrand M. (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. J Biol Chem. 278(1): 95-103.

Hong JK, Hwang BK. (2009) The promoter of the pepper pathogen-induced membrane protein gene CaPIMP1 mediates environmental stress responses in plants. Planta. 229(2): 249-259.

Horsch R.B., Fry J.E., Hoffman N.L., Eichholtz D., Rogers S.G. Fraley R.T. (1985) A simple and general method for transferring genes into plants. Science. 227: 1229-1260.

Howe E., Holton K., Nair S., Schlauch D., Sinha R., Quackenbush J. (2010) MeV: MultiExperiment Viewer. In Biomedical Informatics for Cancer Research. 2010: 267-277.

Howles P.A., Sewalt V., Paiva N.L., Elkind Y., Bate N.J, Lamb C., Dixon R.A. (1996) Overexpression of L-Phenylalanine Ammonia-Lyase in Transgenic Tobacco Plants Reveals Control Points for Flux into Phenylpropanoid Biosynthesis. Plant Physiol. 112(4): 1617-1624.

Hrazdina G., Wagner G.J. (1985) Metabolic as enzyme complexes: Evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. Arch Biochem Biophys. 237(1): 88-100.

Humphreys J.M., Chapple C. (2002) Rewriting the lignin roadmap. Curr Opin Plant Biol. 5(3): 224-229.

Hwang Y.P., Yun H.J., Chun H.K., Chung Y.C., Kim H.K., Jeong M.H., Yoon T.R., Jeong H.G. (2009) Protective mechanisms of 3-caffeoyl, 4-dihydrocaffeoyl quinic acid from Salicornia herbacea against tert-butyl hydroperoxide-induced oxidative damage. Chem Biol Interact. 181(3): 366-376.

Iiyama K., Lam T. B.-T., Stone B. A. (1994) Covalent cross-links in the cell wall. Plant Physiol. 104: 315-320.

Inoue H., Nojima H., Okayama H. (1990) High efficiency transformation of Escherichia coli with plasmids. Gene. 96(1): 23-28.

Ishikawa M., Tsuchiya D., Oyama T., Tsunaka Y., Morikawa K. (2004) Structural basis for channelling mechanism of a fatty acid *beta*-oxidation multienzyme complex. EMBO J. 23(14): 2745-2754.

Ito H., Fukufa Y., Murata K., Kimura A. (1983) J Bacteriol. 153: 163-168.

Jefcoate C.R. (1978) Measurement of substrate and inhibitor binding to microsomal cytochrome P-450 by optical-difference spectroscopy. Methods Enzymol. 52: 258-79.

Jensen K., Møller B.L. (2010) Plant NADPH-cytochrome P450 oxidoreductases Phytochemistry. 71: 132-141.

Jin H., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B., Martin C. (2004) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J. 19(22): 6150-6161.

Jørgensen K., Rasmussen A.V., Morant M., Nielsen A.H., Bjarnholt N., Zagrobelny M., Bak S., Møller B.L. (2005) Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Curr. Opin. Plant Biol. 8: 280-291.

Kahn R.A., Durst F. (2000) Chapter Six: Function and evolution of plant cytochrome P450. Recent Advances in Phytochemistry. 34: 151-189

Kai K., Mizutani M., Kawamura N., Yamamoto R., Tamai M., Yamaguchi H., Sakata K., Shimizu B. (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.

Kandel S., Morant M., Benveniste I., Blée E., Werck-Reichhart D., Pinot F. (2005) Cloning, functional expression, and characterization of CYP709C1, the first sub-terminal hydroxylase of long chain fatty acid in plants. Induction by chemicals and methyl jasmonate. J Biol Chem. 280(43): 35881-35889.

Katsumoto Y., Fukuchi-Mizutani M., Fukui Y., Brugliera F., Holton T.A., Karan M, Nakamura N., Yonekura-Sakakibara K., Togami J., Pigeaire A., Tao G-Q., Nehra N.S., Lu C.-Y., Dyson B.K., Tsuda S., Ashikari T., Kusumi T., Mason J.G.,Tanaka Y. (2007) Engineering of the Rose Flavonoid Biosynthetic Pathway Successfully Generated Blue-Hued Flowers Accumulating Delphinidin. Plant Cell Physiol. 48(11): 1589-1600.

Kelley R.W., Cheng D., Backes W.L. (2006) Heteromeric complex formation between CYP2E1 and CYP1A2: evidence for the involvement of electrostatic interactions. Biochemistry. 45(51): 15807-15816.

Kim J., DellaPenna D. (2006) Defining the primary route for lutein synthesis in plants: the roleof *Arabidopsis* carotenoid beta-ring hydroxylase CYP97A3. Proc. Natl. Acad. Sci. USA
103: 3474-3479.

Kim S.H., Kronstad J.W., Ellis B.E. (2001) Induction of phenylalanine ammonia-lyase activity by tryptophan in *Ustilago maydis*. Phytochemistry. 58(6): 849-857.

King R.R., Calhoun L. (2005) Characterization of cross-linked hydroxycinnamic acid amides isolated from potato common scab lesions, Phytochemistry. 66(20): 2468-2473.

Klingenberg M. (1958). Pigments of rat liver microsomes. Arch. Biochem. Biophys. 75: 376-386.

Koch U., Kühnl T., Conradt W., Wellmann E. (1990) Differential effects of light and fungal elicitor on chlorogenic acid and caffeoylshikimic acid metabolism. Plant Sci. 70(2): 167-174.

Koltunow A.M., Truettner J., Cox K.H., Wallroth M., Goldberg R.B. (1990) Different Temporal and Spatial Gene Expression Patterns Occur during Anther Development. Plant Cell. 2(12): 1201-1224.

Koopmann E., Logemann E., Hahlbrock K. (1999) Regulation and functional expression of cinnamate 4-hydroxylase from parsley. Plant Physiol. 119(1): 49-56.

Kühnl T., Koch U., Heller W., Wellmann E. (1987) Chlorogenic acid biosynthesis: characterization of a light-induced microsomal 5-O-(4-coumaroyl)-D-quinate/shikimate 3'-hydroxylase from carrot (*Daucus carota L.*) cell suspension cultures. Arch Biochem Biophys. 258(1): 226-232.

Kumar A., Ellis B.E. (2001) The phenylalanine ammonia-lyase gene family in raspberry. Structure, expression, and evolution. Plant Physiol. 127(1): 230-239.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227(5259): 680-685.

Lamb C., Rubery P.H. (1976) Photocontrol of chlorogenic acid biosynthesis in potato discs. Phytochemistry. 15: 665-668.

Landais I., Ogliastro M., Mita K., Nohata J., López-Ferber M., Duonor-Cérutti M., Shimada T., Fournier P., Devauchelle G. (2003) Annotation pattern of ESTs from *Spodoptera frugiperda* Sf9 cells and analysis of the ribosomal protein genes reveal insect-specific features and unexpectedly low codon usage bias. Bioinformatics. 19(18): 2343-2350.

Landry L.G., Chapple C.C.S., Last R.L. (1995) *Arabidopsis* Mutants Lacking Phenolic Sunscreens Exhibit Enhanced Ultraviolet-B Injury and Oxidative Damage Plant Physiol. 109 (4): 1159-1166.

Langenfeld E, Meyer HE, Marcus K. (2008) Quantitative analysis of highly homologous proteins: the challenge of assaying the "CYP-ome" by mass spectrometry. Anal Bioanal Chem. 392(6): 1123-1134.

Laranjinha J.A., Almeida L.M., Madeira V.M. (1994) Reactivity of dietary phenolic acids with peroxyl radicals: antioxidant activity upon low density lipoprotein peroxidation. Biochem Pharmacol. 48(3): 487-494.

Larbat R., Kellner S., Specker S., Hehn A., Gontier E., Hans J., Bourgaud F., 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 D.G., Park Y., Kim M.R., Jung H.J., Seu Y.B., Hahm K.S., Woo E.R. (2004) Anti-fungal effects of phenolic amides isolated from the root bark of *Lycium chinense*. Biotechnol Lett. 26(14): 1125-1130.

Lee D, Meyer K, Chapple C, Douglas CJ. (1997) Antisense suppression of 4-coumarate:coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. Plant Cell. 9(11): 1985-1998.

Lee V. W., Jin S., Sim W. S., Nester E. W. (1995) Genetic evidence for direct sensing of phenolic compounds by the VirA protein of *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. USA 92: 12245-12249.

Lehnerer M., Schulze J., Pernecky S.J., Lewis D.F., Eulitz M., Hlavica P. (1998) Influence of mutation of the amino-terminal signal anchor sequence of cytochrome P450 2B4 on the enzyme structure and electron transfer processes. J Biochem.124(2): 396-403.

Lemmon M.A. (2008) Membrane recognition by phospholipid-binding domains. Nat Rev Mol Cell Biol. 9(2): 99-111.

Levy C.C., Zucker M. (1960) Cinnamyl and p-coumaryl esters as intermediates in the biosynthesis of chlorogenic acid. J Biol Chem. 235: 2418-2425.

Lewis N.G., Yamamoto E. (1990) Lignin: occurrence, biogenesis and biodegradation. Annu Rev Plant Physiol Plant Mol Biol. 41: 455-496.

Li X., Bonawitz N.D., Weng J.K., Chapple C. (2010) The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. Plant Cell. 22(5): 1620-1632.

Linder M.E., Deschenes R.J. (2007) Palmitoylation: policing protein stability and traffic. Nat Rev Mol Cell Biol. 8(1): 74-84.

Lindermayr C., Mollers B., Fliegmann J., Uhlmann A., Lottspeich F., Meimberg H., Ebel J. (2002) Divergent members of a soybean (*Glycine max L.*) 4-coumarate:coenzyme A ligase gene family. Eur J Biochem 269: 1304-1315.

Liu R.H. (2004) Potential synergy of phytochemicals in cancer prevention: mechanism of action. The J Nutr. 134: 3479S-3485S.

Livak K.J., Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25(4): 402-408.

Logemann E., Parniske M., Hahlbrock K. (1995) Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proc Natl Acad Sci U S A. 92(13): 5905-5909.

Luo J., Fuell C., Parr A., Hill L., Bailey P., Elliott K., Fairhurst S.A., Martin C., Michael A.J. (2009) A novel polyamine acyltransferase responsible for the accumulation of spermidine conjugates in *Arabidopsis* seed. Plant Cell. 21(1): 318-333.

Luo J., Nishiyama Y., Fuell C., Taguchi G., Elliott K., Hill L., Tanaka Y., Kitayama M., Yamazaki M., Bailey P., Parr A., Michael A.J., Saito K., Martin C. (2007) Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. Plant J.50(4): 678-695.

Maher E.A., Bate N.J., Ni W., Elkind Y., Dixon R.A., Lamb C.J. (1994) Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. Proc Natl Acad Sci U S A. 91(16): 7802-7806.

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., Campa C. (2007) Functional characterization of two p-coumaroyl ester 3'-hydroxylase genes from coffee tree: evidence of a candidate for chlorogenic acid biosynthesis. Plant Mol Biol. 64(1-2): 145-159.

Mansuy D. (1998) The great diversity of reactions catalyzed by cytochromes P450. Comp Biochem Physiol Part C Pharmacol Toxicol Endocrinol. 121(1-3): 5-14.

Markham A., Hartman G.C., Parke D.V. (1972) Spectral evidence for the presence of cytochrome P-450 in microsomal fractions obtained from some higher plants. Biochem J. 130(2): 90P.

Markvart T. (2000) Light harvesting for quantum solar energy conversion. Prog quantum electron. 24:107-186.

Matsuda F., Morino K., Miyashita M., Miyagawa H. (2003) Metabolic flux analysis of the phenylpropanoid pathway in wound-healing potato tuber tissue using stable isotope-labeled tracer and LC-MS spectroscopy. Plant Cell Physiol. 44(5): 510-517.

Matsuno M., Compagnon V., Schoch G., Schmitt M., Debayle D., Bassard J.-E., Pollet B., Hehn A., Heintz D., Ullmann P., Lapierre C., Bernier F., Ehlting J., Werck-Reichhart D., (2009) Retroposition and positive selection led to evolution of a novel phenolic pathway for pollen development. Science. 225: 1688-1692.

Matsuno M., Nagatsu A., Ogihara Y., Ellis B.E., Mizukami H. (2002) CYP98A6 from Lithospermum erythrorhizon encodes 4-coumaroyl-4'-hydroxyphenyllactic acid 3-hydroxylase involved in rosmarinic acid biosynthesis. FEBS Lett. 514(2-3): 219-224.

Maury S., Geoffroy P., Legrand M. (1999) Tobacco O-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. Plant Physiol. 121(1): 215-224.

Melkozernov A.N., Barber J., Blankenship R.E. (2006) Light harvesting in photosystem I supercomplexes. Biochemistry. 45(2): 331-345.

Meyer K., Shirley A.M., Cusumano J.C., Bell-Lelong D.A., Chapple C. (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci U S A. 95(12): 6619-6623.

Miles E.W., Rhee S., Davies D.R. (1999) The molecular basis of substrate channeling. J Biol Chem. 274(18): 12193-12196.

Millar D.J., Long M., Donovan G., Fraser P.D., Boudet A.M., Danoun S., Bramley P.M., Bolwell G.P. (2007) Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids. Phytochemistry. 68(11): 1497-509.

Mizutani M., Ohta D. (1998) Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Gene structure, heterologous expression in insect cells, and differential regulation. Plant Physiol. 116(1): 357-367.

Mizutani M., Ohta D. (2010) Diversification of P450 genes during land plant evolution. Annu Rev Plant Biol. 61: 291-315.

Mizutani M., Ohta D., Sato R. (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta. Plant Physiol. 113(3): 755-763.

Mizutani M., Ward E., DiMaio J., Ohta D., Ryals J., Sato R.(1993) Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. Biochem Biophys Res Commun. 190(3): 875-880.
Mo C., Bard M. (2005) Erg28p is a key protein in the yeast sterol biosynthetic enzyme complex. J Lipid Res. 46(9): 1991-1998.

Mock H.P., Strack D. (1992) Energetics of uridine 5-diphophoglucose: hydroxycinnamic acid glucosyltransferase reaction. Phytochemistry. 32: 575-579.

Mølgaard P., Ravn H.(1988) Evolutionary aspects of caffeoyl ester distribution In Dicotyledons . Phytochemistry. 27(8): 2411-2421

Moon T.S., Dueber J.E., Shiue E., Prather K.L. (2010) Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. Metab Eng. 12(3): 298-305.

Morant M., Bak S., Møller B.L., Werck-Reichhart D. (2003) Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr. Opin. Biotechnol. 14: 151-162.

Morant M., Schoch G.A., Ullmann P., Ertunc T., Little D., Olsen E., Petersen M., Negrel J., Werck-Roichhart D. (2007). Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. Plant Mol. Biol. 63: 1-19.

Moura J.C., Bonine C.A., de Oliveira Fernandes Viana J., Dornelas M.C., Mazzafera P. (2010) Abiotic and biotic stresses and changes in the lignin content and composition in plants. J Integr Plant Biol. 52(4): 360-376.

Murphy PJ, West CA. (1969) The role of mixed function oxidases in kaurene metabolism in *Echinocystis macrocarpa* Greene endosperm. Arch Biochem Biophys. 133(2): 395-407.

Nair R.B., Joy R.W. 4th, Kurylo E., Shi X., Schnaider J., Datla R.S., Keller W.A., Selvaraj G. (2000) Identification of a CYP84 family of cytochrome P450-dependent mono-oxygenase genes in *Brassica napus* and perturbation of their expression for engineering sinapine reduction in the seeds. Plant Physiol. 123(4): 1623-1634.

Nair P.M., Vining L.C. (1965) Cinnamic acid hydroxylase in spinach. Phytochemistry. 4: 161-168.

Nair R.B., Xia Q., Kartha C.J., Kurylo E., Hirji R.N., Datla R., Selvaraj G. (2002) *Arabidopsis* CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. Plant Physiol. 130(1): 210-220.

Nadolski M.J. & Linder M.E. (2007) Protein lipidation. FEBS J. 274(20): 5202-5210.

Nambudiri A. M., Rao P. V., Bhat, J. V. (1970) Metabolism of aromatic compounds by an *Alternaria* species. Phytochemistry 9: 687-693.

Nebert D.W., Adesnik M., Coon M.J., Estabrook R.W., Gonzalez F.J., Guengerich F.P., Gunsalus I.C., Johnson E.F., Kemper B., Levin W., Phillips R.Y., Sato R., Waterman M.R. (1987) The P450 gene superfamily: recommended nomenclature. DNA. 6(1): 1-11.

Nedelkina S., Jupe S.C., Blee K.A., Schalk M., Werck-Reichhart D., Bolwell G.P. (1999) Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: engineering expression in yeast. Plant Mol Biol. 39(6): 1079-1090.

Negrel J., Pollet B., Lapierre C. (1996) Ether-linked ferulic acid amides in natural and wound periderm of potato tuber, Phytochemistry. 43(6): 1195-1199.

Nelson D.R. (2004) Cytochrome P450 Nomenclature. Methods in Molecular Biology, vol. 320: Cytochrome P450 Protocols: Second Edition 1-10.

Nelson D.R. (2006) Cytochrome P450 nomenclature 2004. Methods Mol Biol. 320:1-10.

Nelson D.R. (2010) Progress in tracing the evolutionary paths of cytochrome P450. Biochim Biophys Acta. 2010 (in press)

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., Nebert W. (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol. 12(1): 1-51.

Neuhoff V., Arold N., Taube D., Ehrhardt W. (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis. 9(6): 255-262.

Nielsen K.A., Tattersall D.B., Jones P.R., Møller B.L. (2008) Metabolon formation in dhurrin biosynthesis. Phytochemistry. 69(1): 88-98.

Niggeweg R., Michael A.J., Martin C. (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. Nat Biotechnol. 22(6):746-754.

Nishihara K, Kanemori M, Yanagi H, Yura T. (2000) Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. Appl Environ Microbiol. 66(3): 884-889.

Nour-Eldin H.H., Hansen B.G., Nørholm M.H., Jensen J.K., Halkier B.A. (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. 34(18): e122.

Oksman-Caldentey K.-M., Hiltunen R. (1996) Transgenic crops for improved pharmaceutical products. Field Crops Research 45: 57-69.

Omura T., and Sato R. (1964a). The Carbon Monoxide-binding Pigment of Liver. Microsomes. I. Evidence for its hemoprotein nature. J. Biol. Chem. 239, 2370-2378.

Omura T., and Sato R. (1964b). The Carbon Monoxide-binding Pigment of Liver. II. Solubilization, purification, and properties. J. Biol. Chem. 239, 2379-2385.

Overkamp S., Hein F., Barz W. (2000) Cloning and characterization of eight cytochrome P450 cDNAs from chickpea (*Cicer arietinum L.*) cell suspension cultures. Plant Sci. 155(1): 101-108.

Pain M.J.I., Scrutton N .S., Munro A.W.,Gutierrez A., Roberts G.CK., Wolf C.R. (2005) Electron transfer partners of cytochromes P450. Cytochrome P450: Structure, Mechanism and Biochemistry, 3rd edition. Edited by Ortiz de Montellano. Kluwer Academic/Plenum Publishers: 115-148.

Panicot M., Minguet E.G., Ferrando A., Alcázar R., Blázquez M.A., Carbonell J., Altabella T., Koncz C., Tiburcio A.F. (2002) A polyamine metabolon involving aminopropyl transferase complexes in *Arabidopsis*. Plant Cell 14: 2539-2551.

Parr A.J., Mellon F.A., Colquhoun I.J., Davies H.V. (2005) Dihydrocaffeoyl polyamines (kukoamine and allies) in potato (Solanum tuberosum) tubers detected during metabolite profilingJ Agric Food Chem. 53(13): 5461-5466.

Pawley J. B. (2006) Handbook of biological confocal microscopy. Third edition. Springer science, New York, USA.

Payyavula R.S., Babst B.A., Nelsen M.P., Harding S.A., Tsai C.J. (2009) Glycosylation-mediated phenylpropanoid partitioning in *Populus tremuloides* cell cultures. BMC Plant Biol. 9:151.

Perret A., Pompon D. (1998) Electron shuttle between membrane-bound cytochrome P450 3A4 and b5 rules uncoupling mechanisms. Biochemistry. 37(33): 11412-11424.

Petersen M., Simmonds M.S. (2003) Rosmarinic acid. Phytochemistry. 62(2): 121-125.

Pierrel M.A., Batard Y., Kazmaier M., Mignotte-Vieux C., Durst F., Werck-Reichhart D. (1994) Catalytic properties of the plant cytochrome P450 CYP73 expressed in yeast. Substrate specificity of a cinnamate hydroxylase. Eur J Biochem. 224(3): 835-844.

Pollard M., Beisson F., Li Y., Ohlrogge J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Pl. Sci. 13: 236-246.

Pompon D., Louerat B., Bronine A., Urban P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol. 272: 51-64.

Potts J.R., Weklych R., Conn E.E., Rowell J. (1974) The 4-hydroxylation of cinnamic acid by sorghum microsomes and the requirement for cytochrome P-450. J Biol Chem. 249(16): 5019-5026.

Poulos T.L., Johnson E.F. (2005) Structures of Cytrochrome P450 Enzymes. Cytochrome P450. Structure, Mechanism, and Biochemistry. Third edition. P.R.O.d. Montellano, Springer US.

Powles S.B., Yu Q. (2010) Evolution in action: plants resistant to herbicides. Annu Rev Plant Biol. 61: 317-347.

Praporski S., Ng S.M., Nguyen A.D., Corbin C.J., Mechler A., Zheng J., Conley A.J., Martin L.L. (2009) Organization of cytochrome P450 enzymes involved in sex steroid synthesis: PROTEIN-PROTEIN INTERACTIONS IN LIPID MEMBRANES. J Biol Chem. 284(48): 33224-33232.

Prescott M., Nowakowski S., Nagley P., Devenish R.J. (1999) The Length of Polypeptide Linker Affects the Stability of Green Fluorescent Protein Fusion Proteins. Anal. Biochemistry 273: 305-307

Prober J.M., Trainor G.L., Dam R.J., Hobbs F.W., Robertson C.W., Zagursky R.J., Cocuzza A.J., Jensen M.A., Baumeister K. (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. Science. 238(4825): 336-341.

Raes J., Rohde A., Christensen J.H., Van de Peer Y., Boerjan W. (2003) Genome-wide characterization of the lignification toolbox in *Arabidopsis*. Plant Physiol. 133(3): 1051-1071.

Ralph J., Akiyama T., Kim H., Lu F., Schatz P.F., Marita J.M., Ralph S.A., Reddy M.S., Chen F., Dixon R.A. (2006) Effects of coumarate 3-hydroxylase down-regulation on lignin structure. J Biol Chem. 281(13): 8843-8853.

Ralph J., Guillaumie S., Grabber J.H., Lapierre C., Barrière Y. (2004) Genetic and molecular basis of grass cell-wall biosynthesis and degradability. III. Towards a forage grass ideotype. C R Biol. 327(5): 467-79.

Ralston L., Kwon S.T., Schoenbeck M., Ralston J., Schenk D.J., Coates R.M., Chappell J. (2001) Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1,3-dihydroxylase from tobacco (*Nicotiana tabacum*). Arch Biochem Biophys. 393(2): 222-235.

Ralston L., Yu O. (2006) Metabolons involving plant cytochrome P450s. Phytochem Rev. 5: 459-472

Rasmussen S., Dixon R.A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. Plant Cell 11: 1537-1551.

Ravichandran K.G., Boddupalli S.S., Hasermann C.A., Peterson J.A., Deisenhofer J. (1993) Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. Science. 261(5122): 731-736.

Rezai-Zadeh K., Shytle D., Sun N., Mori T., Hou H., Jeanniton D., Ehrhart J., Townsend K., Zeng J., Morgan D., Hardy J., Town T., Tan J. (2005) Green tea epigallocatechin-3-gallate (EGCG) modulates amyloid precursor protein cleavage and reduces cerebral amyloids in Alzheimer transgenic mice. J Neurosci. 25: 8807-8814.

Rhodes M.J.C. & Wooltorton L.S.C. (1976) The enzymatic conversion of hydroxycinnamic acids to *p*-coumaryliquinic and chlorogenic acids in tomato fruits. Phytochemistry. 15: 947-951.

Rhodes M.J.C., Wooltorton L.S.C., Lourenço E.J. (1979) Purification and properties of hydroxycinnamoyl CoA quinate hydroxycinnamoyl transferase from potatoes. Phytochemistry. 18: 1125-1129.

Ribnicky D.M., Shulaev V.V, Raskin I.I. (1998) Intermediates of salicylic acid biosynthesis in tobacco. Plant Physiol. 118(2): 565-572.

Rice-Evans CA, Miller NJ, Paganga G. (1996) Structure-antioxidant activity relationships of flavonoids and phenolic acids. Free Radic Biol Med. 20(7): 933-956.

Ro D.K., Douglas C.J. (2004) Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (Saccharomyces cerevisiae): implications for control of metabolic flux into the phenylpropanoid pathway. J Biol Chem. 279(4): 2600-2607.

Ro D.K., Mah N., Ellis B.E., Douglas C.J. (2001) Functional characterization and subcellular localization of poplar (*Populus trichocarpa* x *Populus deltoides*) cinnamate 4-hydroxylase. Plant Physiol. 126(1): 317-329.

Robineau T., Batard Y., Nedelkina S., Cabello-Hurtado F., LeRet M., Sorokine O., Didierjean L., Werck-Reichhart D. (1998) The chemically inducible plant cytochrome P450 CYP76B1 actively metabolizes phenylureas and other xenobiotics. Plant Physiol. 118(3): 1049-1056.

Rohman M., Harrison-Lavoie K.J. (2000) Separation of copurifying GroEL from glutathione-S-transferase fusion proteins. Protein Expr Purif. 20(1): 45-47.

Rösler J., Krekel F., Amrhein N., Schmid J. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol. 11: 175-179.

Ruegger M., Chapple C. (2001) Mutations that reduce sinapoylmalate accumulation in Arabidopsis thaliana define loci with diverse roles in phenylpropanoid metabolism. Genetics. 159(4): 1741-1749.

Ruijter J.M., Ramakers C., Hoogaars W.M., Karlen Y., Bakker O., van den Hoff M.J., Moorman A.F. (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. 37(6): e45.

Rupasinghe S., Baudry J., Schuler M.A. (2003) Common active site architecture and binding strategy of four phenylpropanoid P450s from *Arabidopsis thaliana* as revealed by molecular modeling. Protein Eng. 16(10): 721-731.

Russell D.W., Conn E.E. (1967). The cinnamic acid 4-hydroxylase of pea seedlings. Arch Biochem Biophys. 122(1): 256-258.

Rylott E.L., Jackson R.G., Sabbadin F., Seth-Smith H.M., Edwards J., Chong C.S., Strand S.E., Grogan G., Bruce N.C. (2010) The explosive-degrading cytochrome P450 XplA: Biochemistry, structural features and prospects for bioremediation. Biochim Biophys Acta. 2010 (in press).

Saito S., Hirai N., Matsumoto C., Ohigashi H., Ohta D., Sakata K., Mizutani M. (2004) Arabidopsis CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. Plant Physiol. 134(4): 1439-1449.

Sawa T., Nakao M., Akaike T., Ono K., Maeda H. (1999) Alkylperoxyl radical-scavenging activity of various flavonoids and other phenolic compounds: implications for the anti-tumor-promoter effect of vegetables. J Agric Food Chem. 47(2): 397-402.

Sawa T., Nakao M., Akaike T., Ono K., Maeda H. (1999) Alkylperoxyl radical-scavenging activity of various flavonoids and other phenolic compounds: implications for the anti-tumor-promoter effect of vegetables. J Agric Food Chem. 47(2): 397-402.

Schalk, M., Batard, Y., Seyer, A., Nedelkina, S., Durst, F., Werck- Reichhart, D., (1997a) Design of fluorescent substrates and potent inhibitors of CYP73As, P450s that catalyze 4-hydroxylation of cinnamic acid in higher plants. Biochemistry 36: 15253-15261.

Schalk M., Cabello-Hurtado F., Pierrel M.A., Atanossova R., Saindrenan P., 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 Physiol. 118(1): 209-218.

Schalk, M., Pierrel, M.A., Zimmerlin, A., Batard, Y., Durst, F., Werck- Reichhart, D., (1997b) Xenobiotics: substrates and inhibitors of the plant cytochrome P450. Environ. Sci. Pollut. Res. 4: 229-234.

Schilmiller A.L., Stout J., Weng J.K., Humphreys J., Ruegger M.O., Chapple C. (2009) Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. Plant J. 60(5): 771-782.

Schmittgen T.D., Livak K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 3(6): 1101-1108.

Schoch G.A., Attias R., Le Ret M., Werck-Reichhart D. (2003) Key substrate recognition residues in the active site of a plant cytochrome P450, CYP73A1. Homology guided site-directed mutagenesis. Eur J Biochem. 270(18): 3684-3695.

Schoch G., Goepfert S., Morant M., Hehn A., Meyer D., Ullmann P., Werck-Reichhart D. (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem. 276(39): 36566-36574.

Schoch G.A., Morant M., Abdulrazzak N., Asnaghi C., Goepfert S., Petersen M., Ullmann P., Werck-Reichhart D. (2006) The *meta*-hydroxylation step in the phenylpropanoid pathway:
a new level of complexity in the pathway and its regulation. Environ Chem Lett. 4:127-136.

Schoch G.A., Nikov G.N., Alworth W.L., Werck-Reichhart D. (2002) Chemical inactivation of the cinnamate 4-hydroxylase allows for the accumulation of salicylic acid in elicited cells. Plant Physiol. 130(2): 1022-1031.

Schuler M.A. (1996) The role of cytochrome P450 monooxygenases in plant-insect interactions. Plant Physiol. 112(4): 1411-1419.

Schuler M.A., Duan H., Bilgin M., Ali S. (2006) Arabidopsis P450s through the looking glass: a window on plant biochemistry. Phytochem Rev. 5: 205-237.

Schuler M.A. & Werck-Reichhart D. (2003) Functional Genomics of P450s. Annu. Rev. Seedling Biol. 54: 629-667

Scriven D.R.L., Lynch R.M., Moore E.D.W. (2008) Image acquisition for colocalization using optical microscopy. Am J Physiol Cell Physiol. 294: C1119-C1122.

Seibert C., Davidson B.R., Fuller B.J., Patterson L.H., Griffiths W.J., Wang Y. (2009) Multiple-approaches to the identification and quantification of cytochromes P450 in human liver tissue by mass spectrometry. J Proteome Res. 8(4): 1672-1681.

Seitz C., Ameres S., Forkmann G. (2007) Identification of the molecular basis for the functional difference between flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase. FEBS Lett. 581(18): 3429-3434.

Sessions A.L., Doughty D.M., Welander P.V., Summons R.E., Newman D.K. (2009) The continuing puzzle of the great oxidation event. Curr Biol. 19(14): R567-574.

Sewalt V., Ni W., Blount J.W., Jung H.G., Masoud S.A., Howles P.A., Lamb C., Dixon R.A. (1997) Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of *L*-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase. Plant Physiol. 115(1): 41-50.

Shadle G.L., Wesley S.V., Korth K.L., Chen F., Lamb C., Dixon R.A. (2003) Phenylpropanoid compounds and disease resistance in transgenic tobacco with altered expression of *L*-phenylalanine ammonia-lyase. Phytochemistry. 64(1): 153-161.

Shahidi F., Chandrasekara A. (2010) Hydroxycinnamates and their *in vitro* and *in vivo* antioxidant activities. Phytochem Rev. 9: 147-170.

Shirley B.W. (1996) Flavonoid biosynthesis: 'new' functions for an 'old' pathway. Trends Plant Sci. 1: 377-382.

Siminszky B., Corbin F.T., Ward E.R., Fleischmann T.J., Dewey R.E. (1999) Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast and tobacco enhances the metabolism of phenylurea herbicides. Proc Natl Acad Sci U S A. 96(4):1750-1755.

Simmons B.A., Loqué D., Ralph J. (2010) Advances in modifying lignin for enhanced biofuel production. Curr Opin Plant Biol. 13(3): 313-320.

Simoh S., Quintana N., Kim H.K., Choi Y.H., Verpoorte R. (2009) Metabolic changes in *Agrobacterium tumefaciens*-infected *Brassica rapa*. J Plant Physiol. 166(10): 1005-1014.

Siqueira-Júnior C.L., Jardim B.C., Urményi T.P., Vicente A.C., Hansen E., Otsuki K., da Cunha M., Madureira H.C., de Carvalho D.R., Jacinto T. (2008) Wound response in passion fruit (*Passiflora f. edulis flavicarpa*) plants: gene characterization of a novel chloroplast-targeted allene oxide synthase up-regulated by mechanical injury and methyl jasmonate. Plant Cell Rep. 27(2): 387-397.

Soltani B.M., Ehlting J., Hamberger B., Douglas C.J. (2006) Multiple cis-regulatory elements regulate distinct and complex patterns of developmental and wound-induced expression of *Arabidopsis thaliana* 4CL gene family members. Planta. 224(5):1226-1238.

Somerville C., Bauer S., Brininstool G., Facet M., Hamann T., Milne J., Osborne E., Paredez A., Persson S., Raab T., Vorwerk S., Youngs H. (2004) Toward a systems approach to understanding plant cell walls. Science 306: 2206-2211.

Srere P.A. (1987) Complexes of sequential metabolic enzymes. Annu. Rev. Biochem. 56: 89-124.

Srere P.A. (2000) Macromolecular interactions: tracing the roots. Trends Biochem Sci. 25(3): 150-153.

Srere P.A., Ovadi J. (1990) Enzyme-enzyme interactions and their metabolic role. FEBS Lett. 268(2): 360-364.

Stafford H.A. (1974). Possible multienzyme complexes regulating the formation of C6-C3 phenolic compounds and lignins in higher plants. Rec. Adv. Phytochem. 8: 53-79.

Stewart J.J., Akiyama T., Chapple C.C.S., Ralph J., Mansfield S.D. (2009) The effects on lignin structure on overexpression of ferulate 5-hydroxylase in hybrid poplar. Plant Physiol. 150: 621-635.

St-Pierre B., Laflamme P., Alarco A.M., De Luca V. (1998) The terminal *O*-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. Plant J. 14: 703-713.

Stöckigt J., Zenk M.H. (1974) Enzymatic synthesis of chlorogenic acid from caffeoyl coenzyme A and quinic acid. FEBS Lett. 42(2): 131-134.

Subramanian M., Tam H., Zheng H., Tracy T.S. (2010) CYP2C9-CYP3A4 protein-protein interactions: role of the hydrophobic N terminus. Drug Metab Dispos. 38(6): 1003-1009.

Szabados L., Savouré A. (2010) Proline: a multifunctional amino acid. Trends Plant Sci. 15(2): 89-97.

Szczebara F.M., Chandelier C., Villeret C., Masurel A., Bourot S., Duport C., Blanchard S., Groisillier A., Testet E., Costaglioli P., Cauet G., Degryse E., Balbuena D., Winter J., Achstetter T., Spagnoli R., Pompon D., Dumas B. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. Nat Biotechnol. 21(2): 143-149.

Szczesna-Skorupa E., Chen C.D., Liu H., Kemper B. (2004) Gene expression changes associated with the endoplasmic reticulum stress response induced by microsomal cytochrome P450 overproduction. J Biol Chem. 279(14): 13953-13961.

Szczesna-Skorupa E, Kemper B. (2006) BAP31 is involved in the retention of cytochrome P450 2C2 in the endoplasmic reticulum. J Biol Chem. 281(7): 4142-4148.

Takahashi T., Kakehi J. (2010) Polyamines: ubiquitous polycations with unique roles in growth and stress responses. Ann Bot. 105(1): 1-6.

Tamagnone L., Merida A., Stacey N., Plaskitt K., Parr A., Chang C.F., Lynn D., Dow J.M., Roberts K., Martin C. (1998) Inhibition of phenolic acid metabolism results in precocious cell death and altered cell morphology in leaves of transgenic tobacco plant. Plant Cell. 10(11): 1801-1816.

Tamura K., Dudley J., Nei M., Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24(8): 1596-1599.

Tanaka Y., Brugliera F., Kalc G., Senior M., Dyson B., Nakamura N., Katsumoto Y., Chandler S. (2010) Flower color modification by engineering of the flavonoid biosynthetic pathway: practical perspectives. Biosci Biotechnol Biochem. 74(9): 1760-1769.

Tattersall D.B., Bak S., Jones P.R., Olsen C.E., Nielsen J.K., Hansen M.L., Høj P.B., Møller B.L. (2001) Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science. 293(5536): 1826-1828.

Teutsch H.G., Hasenfratz M.P., Lesot A., Stoltz C., Garnier J.M., Jeltsch J.M., Durst F., Werck-Reichhart D. (1993) Isolation and sequence of a cDNA encoding the *Jerusalem artichoke* cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. Proc Natl Acad Sci U S A. 90(9): 4102-4106.

Tian L., Musetti V., Kim J., Magallanes-Lundback M., DellaPenna D. (2004) The *Arabidopsis LUT1* locus encodes a member of the cytochrome P450 family that is required for carotenoid ε-ring hydroxylation activity. Proc. Natl. Acad. Sci. USA 101: 402-407.

Tikunov Y.M., de Vos R.C., González Paramás A.M., Hall R.D., Bovy A.G.. (2010) A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. Plant Physiol. 152(1): 55-70.

Truan G., Cullin C., Reisdorf P., Urban P., Pompon D. (1993) Enhanced in vivo monooxygenase activities of mammalian P450s in engineered yeast cells producing high levels of NADPH-P450 reductase and human cytochrome b5. Gene. 125(1): 49-55.

Tzin V., Malitsky S., Aharoni A., Galili G. (2009) Expression of a bacterial bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in Arabidopsis. Plant J. 60(1): 156-167.

Ulbrich B. & Zenk M.H. (1979) Partial purification and properties of hydroxycinnamoyl-CoA:quinate hydroxycinnamoyl transferase from higher plants. Phytochemistry. 18: 929-933.

Ulbrich B. & Zenk M.H. (1980) Partial purification and properties of hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase from higher plants. Phytochemistry. 19: 1625-1629.

Urban P., Mignotte C., Kazmaier M., Delorme F., Pompon D. (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J Biol Chem. 272(31): 19176-19186.

Urban P., Werck-Reichhart D., Teutsch H.G., Durst F., Regnier S., Kazmaier M., 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. Eur J Biochem. 222(3): 843-850.

Van Dyke M.W., Sirito M., Sawadogo M. (1992) Single-step purification of bacterially expressed polypeptides containing an oligo-histidine domain. Gene 111: 99-104.

Vanholme R., Morreel K., Ralph J., Boerjan W. (2008) Lignin engineering. Curr Opin Plant Biol. 11: 278-285.

Vogt T. (2010) Phenylpropanoid biosynthesis. Mol Plant. 3(1): 2-20.

Voinnet O., Rivas S., Mestre P., 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): 949-56.

Von Roepenack-Lahaye E., Newman M.A., Schornack S., Hammond-Kosack K.E., Lahaye T., Jones J.D., Daniels M.J., Dow J.M. (2003) *p*-Coumaroylnoradrenaline, a novel plant metabolite implicated in tomato defense against pathogens. J Biol Chem. 278(44): 43373-43383.

Walters D.R. (2003) Polyamines and plant disease. Phytochemistry. 64(1): 97-107.

Watanabe J., Asaka Y., Fujimoto S., Kanamura S. (1993) Densities of NADPH-ferrihemoprotein reductase and cytochrome P-450 molecules in the endoplasmic reticulum membrane of rat hepatocytes. J Histochem Cytochem. 41(1): 43-49.

Watson C.J.W., Froehlich J.E., Josefsson C.A., Chapple C., Durst F., Benveniste I, Coolbaugh RC. (2001) Localization of CYP86B1 in the outer envelope of chloroplasts. Plant Cell Physiol. 42(8): 873-878.

Wehling K., Niester C., Boon J.J., Willemse M.T.M., Wiermann, R. (1989) *p*-Coumaric acid – a monomer in the sporopollenin skeleton. Planta. 179(4): 376-380.

Weigel D., Glazebrook J. (2006) Transformation of *Agrobacterium* Using Electroporation. Cold Spring Harbor Protocols.

Weisshaar B., Jenkins G.I. (1998) Phenylpropanoid biosynthesis and its regulation. Curr Opin Plant Biol. 1(3):251-257.

Weng J.K., Chapple C. (2010) The origin and evolution of lignin biosynthesis. New Phytol. 187(2): 273-285.

Werck-Reichhart D., Bak S., Paquette S. (2002) Cytochromes P450. The Arabidopsis Book ed. American Society of Plant Biologists, 1-29.

Werck-Reichhart D., Feyereisen R. (2000) Cytochrome P450: a success story. Genome Biol. 1: 1-9.

Werck-Reichhart D., Hehn A., Didierjean L. (2000) Cytochromes P450 for engineering herbicide tolerance. Trends Plant Sci. 5(3): 116-123.

Whelan S, Goldman N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 18(5): 691-699.

Whetten R. W., MacKay J. J., Sederoff R. R. (1998) Recent advances in understanding lignin biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 585-609.

Whitbred J.M., Schuler M.A. (2000) Molecular characterization of CYP73A9 and CYP82A1 P450 genes involved in plant defense in pea. Plant Physiol. 124(1): 47-58.

Wildermuth M.C., Dewdney J., Wu G., Ausubel F.M. (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature. 414(6863): 562-565.

Winkel-Shirley B. (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Physiol. Plant. 107: 142-149.

Winkel B.S. (2004) Metabolic channeling in plants. Annu Rev Plant Biol. 55: 85-107.

Xiao F., Goodwin S.M., Xiao Y., Sun Z., Baker D., Tang X., Jenks M.A., Zhou J.M. (2004) Arabidopsis CYP86A2 represses *Pseudomonas syringae* type III genes and is required for cuticle development. EMBO J. 23(14): 2903-2913.

Xu Z., Zhang D., Hu J., Zhou X., Ye X., Reichel K.L., Stewart N.R., Syrenne R.D., Yang X., Gao1 P., Shi1W., Doeppke C., Sykes R.W., Burris J.N., Bozell J.J., Cheng Z.-M., Hayes D.G., Labbe N., Davis M., Stewart Jr C.N., Joshua S Yuan J.S. (2009) Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. BMC Bioinformatics. 10(Suppl 11): S3.

Yang Q., Reinhard K., Schiltz E., Matern U. (1997) Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:anthranilate N-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, *Dianthus caryophyllus L*. Plant Mol Biol. 35(6): 777-789.

Yazaki K. (2005) Transporters of secondary metabolites. Curr Opin Plant Biol. 8(3): 301-307.

Yazaki K. (2006) ABC transporters involved in the transport of plant secondary metabolites. FEBS Lett. 580(4): 1183-1191.

Ye Z.-H., Zhong R., Morrison III W.H., Himmelsbach D.S. (2001) Caffeoyl coenzyme A O-methyltransferase and lignin biosynthesis. Phytochemistry. 57: 1177-1185.

Zajchowski L.D., Robbins S.M. (2002) Lipid rafts and little caves. Compartmentalized signalling in membrane microdomains. Eur J Biochem. 269(3): 737-752.

Zhang Y., Seeram N.P., Lee R., Feng L., Heber D. (2008) Isolation and identification of strawberry phenolics with antioxidant and human cancer cell antiproliferative properties. J Agric Food Chem. 56(3): 670-675.

Zhao J., Davis L.C., Verpoorte R. (2005) Elicitor signal transduction leading to production of plant secondary metabolites. Biotechnol Adv. 23(4): 283-333.

Zhong R., Ye Z.H. (2007) Regulation of cell wall biosynthesis Curr Opin in Plant Biol. 10: 564-572.

Zhong R., Ye Z.H. (2009) Transcriptional regulation of lignin biosynthesis. Plant Signal Behav. 4(11): 1028-1034.

Appendix

List of P450 functions characterized in *Arabidopis thaliana*.

P450	activity	pathway
51G1	obtusifoliol 14α-demethylase	sterols/steroids
51G2	obtusifoliol 14α-demethylase	sterols/steroids
71A13	conversion of indole-3-acetaldoxime in indole-3-acetonitrile	camalexin
71B15	conversion of s-dihydrocamalexic acid to camalexin	camalexin
72C1	exact substrate not identified	degradation of brassinosteroids
73A5	cinnamic acid 4-hydroxylase (t-CAH)	phenylpropanoid
74A1	allene oxide synthase (AOS)	oxylipin
74B2	hydroperoxide lyase (HPL)	oxylipin
75B1	3'-hydroxylase for narigenin, dihydrokaempferol (F3'H)	phenylpropanoid
76C1	geraniol 10-hydroxylase	isoprenoid
77A4	epoxidation of free fatty acids	fatty acids
79A2	conversion of phenylalanine to oxime	benzylglucosinolate
79B2	conversion of tryptophan, tryptophan analogs to oxime	indole glucosinolate
79B3	conversion of tryptophan to oxime	indole glucosinolate
79F1	mono to hexahomomethionine in synthesis of short and long chain aliphatic glucosinolates	aliphatic glucosinolate
79F2	long chain penta and hexahomomethionine in synthesis of long chain aliphatic glucosinolates	aliphatic glucosinolate
81F2	indol-3-ylmethylglucosinolate 4-hydroxylase	indole glucosinolate
83A1	oxidation of methionine-derived oximes;oxidation of p-hydroxyphenyl-acetaldoxime, indole-3-acetaldoxime	aliphatic glucosinolate
83B1	oxidation of indole-3-acetaldoxime	indole glucosinolate
84A1	5-hydroxylase for coniferaldehyde, coniferyl alcohol and ferulic acid (F5H)	phenylpropanoid
85A1	C6-oxidase for 6-deoxycastasterone, other steroids	brassinolide
85A2	C6-oxidase for 6-deoxycastasterone, other steroids;conversion of castasterone to brassinolide	brassinolide
86A1	ω-hydroxylase for satur. and unsatur. C12 to C18 fatty acids	fatty acids
86A2	ω-hydroxylase for satur. and unsat. C12 to C18 fatty acids	fatty acids
86A4	ω-hydroxylase for satur. and unsat. C12 to C18 fatty acids	fatty acids
86A7	ω-hydroxylase for lauric acid	fatty acids
86A8	ω-hydroxylase for satur. and unsatur. C12 to C18 fatty acids	fatty acids
86B1	fatty acid ω-hydroxylase	fatty acids
86C3	fatty acid ω-hydroxylase	fatty acids
88A3	multifunctional ent-kaurenoic acid oxidase	gibberellin
88A4	multifunctional ent-kaurenoic acid oxidase	gibberellin
90A1	23α-hydroxylase for 6-oxo-cathasterone	brassinolide
90B1	22α-hydroxylase for 6-oxo-campestanol, campesterol and cholesterol	brassinolide
90C1	C-23 hydroxylation ; conversion of typhasterol to castasterone	brassinolide
90D1	C-23 hydroxylation	brassinolide
94C1	fatty acid ω-hydroxylase, C12 to C16 fatty acids	fatty acids
97A3	ε-ring hydroxylase on carotenes	carotenoid
97B3	β-ring hydroxylase?	carotenoid
97C1	β-ring hydroxylase on carotenes	carotenoid
98A3	3'-hydroxylase for p-coumaryl shikimic/quinic acids (C3'H)	phenylpropanoid
98A8	meta-hydroxylation of N1,N5,N10-tri(p-coumaroyl)spermidine; m-hydroxylation of N1,N5,N10-triferuloylspermidine	phenolamide
98A9	meta-hydroxylation of N1,N5,N10-tri(p-coumaroyl)spermidine	phenolamide
701A3	multifunctional ent-kaurene oxidase	gibberellin
703A2	in-chain monohydroxylation of saturated mid-chain fatty acids	fatty acids
704B1	ω-hydroxylase for satur. and unsatur. C16 and C18 fatty acids	fatty acids
705A5	thalian-diol desaturase	oxigenated triterpens
707A1	8'-hydroxylase for ABA	degradation of abscisic acid
707A2	8'-hydroxylase for ABA	degradation of abscisic acid
707A3	8'-hydroxylase for ABA	degradation of abscisic acid
707A4	8'-hydroxylase for ABA	degradation of abscisic acid
708A2	thalianol hydroxylase	oxigenated triterpens
710A1	C-22 desaturase for β-sitosterol	sterols
710A2	C-22 desaturase on 24-epi-campesterol and β-sitosterol	sterols
724B2	22α-hydroxylase for 6-oxo-campestanol, campesterol and cholesterol	sterols
734A1	26-hydroxylase for brassinolide and castasterone	degradation of brassinolides
735A1	trans-hydroxylase for isopentenyladenine, tri/di/monophosphates	cytokinins
735A2	trans-hydroxylase for isopentenyladenine, tri/di/monophosphates	cytokinins

PCR Protocols

semi-qPCR

Reaction mix	Final concentration
MgCl2	1,5 mM
Taq Buffer	1X
dNTP	250 µM each
Taq polymerase	1U
cDNA matrix	2 µL
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
X cycles	95°C -2 min
	95°C - 30s
	Hydridization temperature - 1 min
	72°C - 1 min
	72°C - 10 min

Colony PCR

Reaction mix	Final concentration
MgCl2	1,5 mM
Tween 20	0,05% (v/v)
Taq Buffer	1X
dNTP	250 µM each
Taq polymerase	1U
Colony	1 colony
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
30 cycles	95°C -8 min
	95°C - 30s
	Hydridization temperature - 30 s
	72°C - Y min
	72°C - 10 min

USER cloning PCR

Reaction mix	Final concentration
Pfu Turbo Cx Buffer	1X
dNTP	250 µM each
Pfu Turbo Cx Buffer	1U
cDNA	n µL
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
15 cycles	95°C - min 95°C - 30s Hydridization sequence temperature - 30 s 72°C - Y min
10 cycles	95°C - 30s Hydridization primer temperature - 30 s 72°C - Y min
	72°C - 10 min

« Regular » Cloning PCR

Reaction mix	Final concentration
MgCl2	1,5 mM
Taq Buffer	1X
dNTP	250 µM each
High Fidelity Polymerase plus Taq polymerase (if TOPO cloning)	1U + 0,1U
cDNA	n µL
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
25 cycles	95°C -2 min 95°C - 30s Hydridization temperature - 1 min 72°C - Y min
	72°C - 10 min

total RNA checking by PCR

Reaction mix	Final concentration
MgCl2	1,5 mM
Taq Buffer	1X
dNTP	250 µM each
Taq polymerase	1U
total RNA	1 µL
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
40 cycles	95°C -2 min
	95°C - 30s
	60°C - 30 s
	72°C - 0,5 min
	72°C - 10 min

cDNA checking by PCR

Reaction mix	Final concentration
MgCl2	1,5 mM
Taq Buffer	1X
dNTP	250 µM each
Taq polymerase	1U
cDNA	1 µL
Forward Primer	0,5 µM
Reverse Primer	0,5 µM
qsp 20µL MilliQ water	

PCR program	
40 cycles	95°C -2 min
	95°C - 30s
	60°C - 1 min
	72°C - 0,5 min
	72°C - 10 min

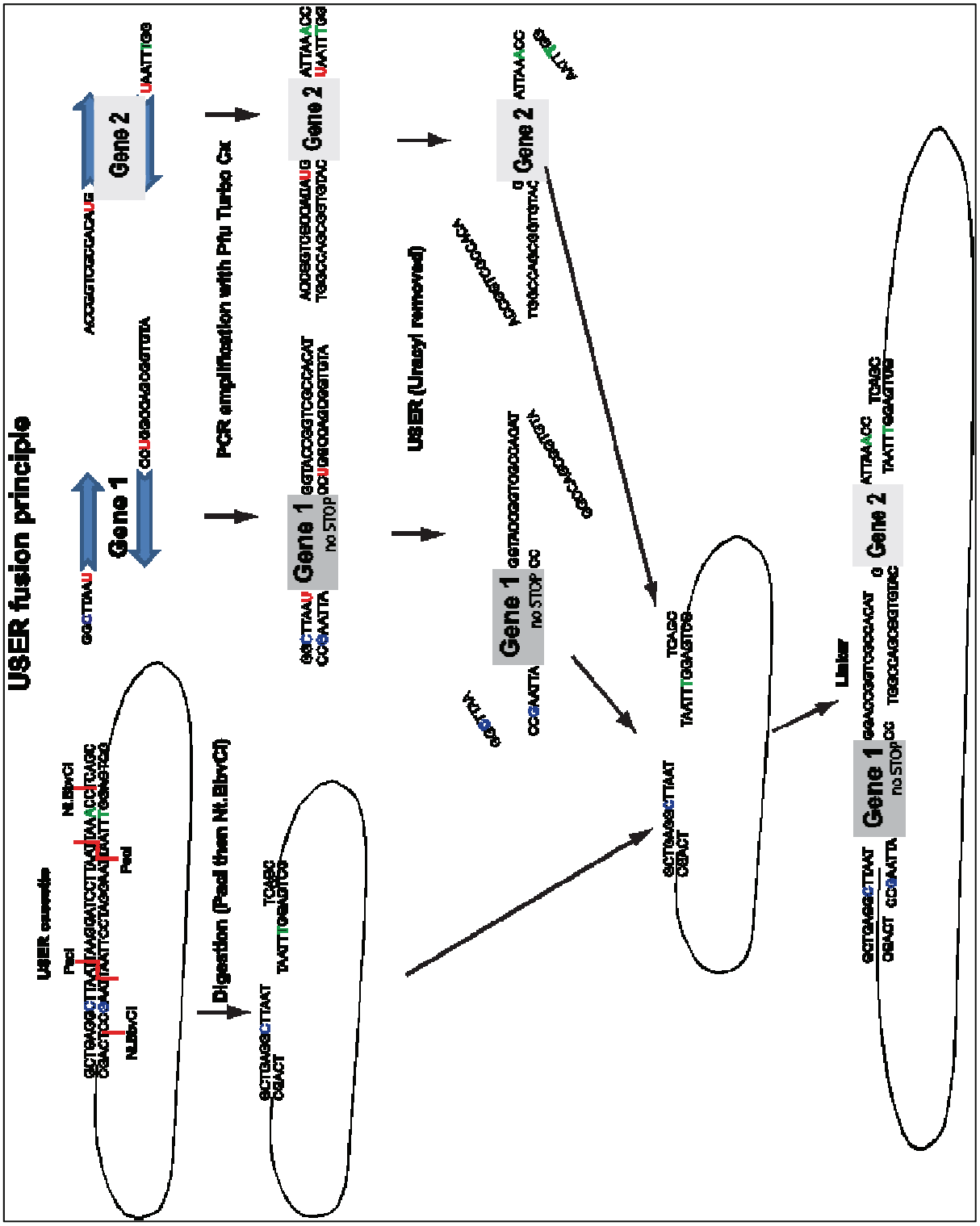
List of oligonucleotide primers

	Target	Sequence (5' - 3')	observation
Oligonucleotidic primers for USER cloning	*AtHCT*	ggcttaaUatgaaaattaacatcagagattcc	forward primer
	AtHCT	ggttaaUtcatatctcaaacaaaaacttctcaaac	reverse primer (STOP)
	AtHCT	accggtcgccacaUgaaaattaacatcag	forward primer with linker (for N-ter USER fusion)
	AtHCT	atgtggcgaccggUacctatctcaaacaaaaac	reverse primer with linker and no STOP (for C-ter USER fusion)
	At4CL-1	atgtggcgaccggUacccaatccatttgc	reverse primer with linker and no STOP (for C-ter USER fusion)
	At4CL-1	ggttaaUtcacaatccatttgctagttttgccc	reverse primer (STOP)
	At4CL-1	ggcttaaUatggcgccacaagaacaagc	forward primer
	At4CL-1	accggtcgccacaUggcgccacaagaacaagcag	forward primer with linker (for N-ter USER fusion)
	At98A3	ggttaaUttacatatcgtaaggcacgcgtttc	reverse primer (STOP)
	At98A3	ggcttaaUatgtcgtcgtggtttctaatagcg	forward primer
	At98A3	atgtggcgaccggUacccatatcgtaaggc	reverse primer with linker and no STOP (for C-ter USER fusion)
	At73A5	ggttaaUttaacagttccttggtttcataacgattatgg	reverse primer (STOP)
	At73A5	ggcttaaUatggacctcctcttgctggag	forward primer
	At73A5	atgtggcgaccggUaccacagttccttgg	reverse primer with linker and no STOP (for C-ter USER fusion)
	Nt73A27	ggcttaaUatgaaaaacatggccaaacttctc	forward primer
	Nt73A27	atgtggcgaccggUacctgcagcaataggcttgaagac	reverse primer with linker and no STOP (for C-ter USER fusion)
	Nt73A28	ggcttaaUatggccaaacttctcaacaacacc	forward primer
	Nt73A28	atgtggcgaccggUacctgcagcaataggcttgaagacaac	reverse primer with linker and no STOP (for C-ter USER fusion)
	Nt73A85	ggcttaaUatggatcttctccttctagagaagacc	forward primer
	Nt73A85	atgtggcgaccggUaccaaaagatcttggtttcatcac	reverse primer with linker and no STOP (for C-ter USER fusion)
	mRFP	atgtggcgaccggUaccggcgccggtggagtg	reverse primer with linker and no STOP (for C-ter USER fusion)
	mRFP	accggtcgccacaUgcctcctccgaggacg	forward primer with linker (for N-ter USER fusion)
	mRFP	ggcttaaUatggcctcctccgaggacg	forward primer
	mRFP	ggttaaUttaggcgccggtggagtgg	reverse primer (STOP)
	eGFP	atgtggcgaccggUacccttgtacagctcgtccatgcc	reverse primer with linker and no STOP (for C-ter USER fusion)
	eGFP	accggtcgccacaUggtgagcaagggcgagg	forward primer with linker (for N-ter USER fusion)
	eGFP	ggcttaaUatggtgagcaagggcgagg	forward primer
	eGFP	ggttaaUttacttgtacagctcgtccatgc	reverse primer (STOP)

	Target	Sequence (5' - 3')	observation
Oligonucleotidic primers for regular cloning	*Nt73A28*	atggccaaacttctcaacaacac	forward primer for cloning of *N. tabacum* parent sequences
	Nt73A28	**tta**tgcagcaataggcttgaagac	reverse primer for cloning of *N. tabacum* parent sequences
	Nt73A85	atggatcttctccttctagagaagacc	forward primer for cloning of *N. tabacum* parent sequences
	Nt73A85	**tta**aaaagatcttggtttcatcacaatgg	reverse primer for cloning of *N. tabacum* parent sequences
	Nt73A47	atggatcttctcttactagagaagacc	forward primer for cloning of *N. tabacum* parent sequences
	Nt73A47	**tca**gaaagaccttggtttcaacac	reverse primer for cloning of *N. tabacum* parent sequences
	At4CL-1	aggcctatggcgccacaagaacaag	forward primer for cloning in pQE30-Xa (StuI)
	At4CL-1	ggtacc**tca**caatccatttgctagttttgc	reverse primer for cloning in pQE30-Xa (BamHI)
	Nt98A33	ggatccatggctctatctttcatat	forward primer for cloning in pYeDP60 and pBD517 (BamHI)
	Nt98A33	ggtacc**tta**catatccactggcac	reverse primer for cloning in pYeDP60 and pBD517 (KpnI)
	Nt73A28	ggatccatggccaaacttctc	forward primer for cloning in pYeDP60 (BamHI)
	Nt73A28	atgagctc**tta**tgcagcaataggcttg	reverse primer for cloning in pYeDP60 (SacI)
	Nt73A85	ttggatccatggatcttctccttctag	forward primer for cloning in pYeDP60 (BamHI)
	Nt73A85	atgagctc**tta**aaaagatcttggtttc	reverse primer for cloning in pYeDP60 (SacI)
Oligonucleotidic primers for RT-qPCR	*NtEF2α (AJ299248)*	tgctggtacacaagctcatcaa	forward primer
	NtEF2α (AJ299248)	agtcactgcctgcttcaaacc	reverse primer
	NtActin66 (U60491)	attccggcgacggtgtctca	forward primer
	NtActin66 (U60491)	acttccggacatctgaacctctctga	reverse primer
	NtHQT	cttctaaagtccaagtccaaaca	forward primer
	NtHQT	gcaacgccaaatatggg	reverse primer
	NtHCT	ttttgccccgacgttagaac	forward primer
	NtHCT	acgattgaattccttgtgagtaatca	reverse primer
	Nt73A27	ccaaacttctcaacaagaccatctt	forward primer
	Nt73A27	aggcatagataggtaggaggacagtaa	reverse primer
	Nt73A28	cgattgcagctgatgctatataaca	forward primer
	Nt73A28	aaggatcatcttgggactcaaact	reverse primer
	Nt73A47	cccggctcattggaagaaa	forward primer
	Nt73A47	ccattggcctcaacargcr	reverse primer
	Nt73A85	ccctcttttagtcccacacatga	forward primer
	Nt73A85	tctccgcgggaatgtca	reverse primer
	Nt98A30	ccgactggtcccaaatatatgg	forward primer
	Nt98A30	tgcattattcacaaccacattcag	reverse primer
	Nt98A31	tcgctggatgttccctattga	forward primer
	Nt98A31	tagctcgagtgagacggtctcta	reverse primer
	Nt98A33	cgactatcggttattgcccttt	forward primer
	Nt98A33	gtgaccaagttgatagcaagttgtg	reverse primer
	ScPDA1	cattcaaacgccaaccatca	forward primer
	ScPDA1	ggtgggagtgcgaagaaca	reverse primer
	ScADH6	tcaagctagggcccaagtca	forward primer
	ScADH6	catgaaaagacttgagcacctacac	reverse primer

	Target	Sequence (5' - 3')	observation
Oligonucleotidic primers for sqPCR	*At73A5*	gggagaaatcaacgaggaca	forward primer
	At73A5	cacgtgcgattcttcttcaa	reverse primer
	At4CL-1	gatgccgttaatgccaagtt	forward primer
	At4CL-1	ggaacttcaccagctgcttc	reverse primer
	AtHCT	tgctggtgttctcttcgttg	forward primer
	AtHCT	gctgacagtgttcccatcct	reverse primer
	At98A3	ccgatcgtcggtaacctcta	forward primer
	At98A3	aacgcaaccgctcctaagta	reverse primer
	NbActin	tacaacgagcttcgtgttgc	forward primer
	NbActin	gaatctctcagctccgatgg	reverse primer
	eGFP	gcatcgacttcaaggagg	forward primer
	eGFP	atatagacgttgtggctgt	reverse primer

Underlined**:** Restriction site.
In orange: Start methionine
In **bold**: Stop codon
In green and blue: single base differences between the generated overhangs, which are responsible for the directional USER cloning



Overview of the USER cloning technique

Extension of user forward primer for Gene 1 : 5’ – GGCTTAAU; Extension of user reverse primer for Gene 1: 5’-ATGTGGCGACCGGUCC; Extension of user forward primer for Gene 2: 5’-ACCGGTCGCCACAUG; Extension of user reverse primer for Gene 2: 5’-GGTTTAAU; Nt.BbvCI and PacI recognition sites are marked in red; green and blue marks are the single base differences between the generated overhangs, which are responsible for the directional cloning. Linker : GGT ACC GGT CGC CAC

G T G R H

Supplemental data of confocal analyses experiments

Proteins	N	Average	SD
Free eGFP	90	1.6	0.5
4CL1::eGFP	90	1.7	0.5
eGFP::4CL1	90	1.6	0.4
HCT::eGFP	90	1.2	0.3
eGFP::HCT	90	1.4	0.4

Confidence in ANOVA tests	**5,423E-13**	****		
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
GFP::4CL1	HCT::eGFP	0.01	9.4E-13	**Yes**
4CL1::eGFP	HCT::eGFP	0.01	1.7E-11	**Yes**
Free eGFP	HCT::eGFP	0.01	8.4E-08	**Yes**
4CL1::eGFP	eGFP::HCT	0.01	2.7E-05	**Yes**
eGFP::4CL1	eGFP::HCT	0.01	6.5E-05	**Yes**
HCT::eGFP	eGFP::HCT	0.01	0.001	**Yes**
Free eGFP	eGFP::HCT	0.01	0.005	**Yes**
Free eGFP	4CL1::eGFP	0.02	0.237	No
4CL1::eGFP	eGFP::4CL1	0.03	0.394	No
Free eGFP	eGFP::4CL1	0.05	0.619	No

4CL-1 and HCT in vivo localization determined by confocal image analysis (Data of Figure 61). N. Number of analyzed images; SD. Standard deviation.

Name	N	Average	SD
4CL1::eGFP	108	1.7	0.5
4CL1::eGFP 73A5::mRFP	108	1.5	0.6
eGFP::4CL1	108	1.7	0.4
eGFP::4CL1 73A5::mRFP	108	1.6	0.5

Confidence in ANOVA tests	**0,008**	**		
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
4CL1::eGFP	4CL1::eGFP 73A5::mRFP	0.01	0.01	**Yes**
4CL1::eGFP 73A5::mRFP	eGFP::4CL1 73A5::mRFP	0.01	0.02	No
eGFP::4CL1	eGFP::4CL1 73A5::mRFP	0.03	0.54	No
4CL1::eGFP	eGFP::4CL1	0.05	0.97	No

Re-localization of soluble proteins upon co-expression of partner P450 detected by confocal microscopy, 4CL1 re-localization (Data of Figure 64A). N. Number of analyzed images; SD. Standard deviation.

Name	N	Average	SD
HCT::eGFP	99	1.2	0.3
HCT::eGFP 93A3::mRFP	99	0.9	0.3
eGFP::HCT	99	1.5	0.5
eGFP::HCT 98A3::mRFP	99	1.1	0.4

Confidence in ANOVA tests	**2,20876E-23**	********		
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
HCT::eGFP	HCT::eGFP 93A3::mRFP	0.01	6.4E-12	**Yes**
eGFP::HCT	eGFP::HCT 98A3::mRFP	0.01	4.2E-11	**Yes**
HCT::eGFP	eGFP::HCT	0.02	1.9E-05	**Yes**
HCT::eGFP 93A3::mRFP	eGFP::HCT 98A3::mRFP	0.05	0.01	**Yes**

Re-localization of soluble proteins upon co-expression of partner P450 detected by confocal microscopy, HCT re-localization (Data of Figure 64B). N. Number of analyzed images; SD. Standard deviation.

Name	N	Average	SD
4CL1::eGFP	87	1.7	0.5
4CL1::eGFP 98A3::mRFP	87	1.2	0.4
eGFP::4CL1	87	1.6	0.4
eGFP::4CL1 98A3::mRFP	87	1.2	0.3

Confidence in ANOVA tests	**2,51216E-17**	********		
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
eGFP::4CL1	eGFP::4CL1 98A3::mRFP	0.01	9,9E-11	**Yes**
4CL1::eGFP	4CL1::eGFP 98A3::mRFP	0.01	4,3E-10	**Yes**
4CL1::eGFP	eGFP::4CL1	0.03	0,4	No
4CL1::eGFP 98A3::mRFP	eGFP::4CL1 98A3::mRFP	0.05	0,4	No

Re-localization of soluble proteins upon co-expression of non direct partner P450 detected by confocal microscopy, 4CL1 re-localization (Data of Figure 66A). N. Number of analyzed images; SD. Standard deviation.

Name	N	Average	SD
HCT::eGFP	99	1.2	0.3
HCT::eGFP 73A5::mRFP	99	1.2	0.3
eGFP::HCT	99	1.5	0.5
eGFP::HCT 73A5::mRFP	99	1.2	0.4

Confidence in ANOVA tests		**1,22922E-07**	****	
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
eGFP::HCT	eGFP::HCT 73A5::mRFP	0.01	9.8E-06	**Yes**
HCT::eGFP	eGFP::HCT	0.01	1.9E-05	**Yes**
HCT::eGFP	HCT::eGFP 73A5::mRFP	0.03	0.7	No
HCT::eGFP 73A5::mRFP	eGFP::HCT 73A5::mRFP	0.05	0.8	No

Re-localization of soluble proteins upon co-expression of non direct partner P450 detected by confocal microscopy, HCT re-localization (Data of Figure 66B). N. Number of analyzed images; SD. Standard deviation.

Proteins	N	Average	SD
4CL1::GFP	90	1.7	0.5
4CL1::GFP 73A5::RFP	90	1.4	0.5
4CL1::GFP 73A5::RFP 98A3 HCT	90	1.4	0.5
GFP::4CL1	90	1.6	0.4
GFP::4CL1 73A5::RFP	90	1.6	0.5
GFP::4CL1 73A5::RFP 98A3 HCT	90	1.4	0.4

Confidence in ANOVA tests		**1,38324E-06**	****	
Posthoc test: Bonferroni-Holm		**Critical value**	**P**	**Significant?**
4CL1::GFP	4CL1::GFP 73A5::RFP 98A3 HCT	0.003	8.0E-05	**Yes**
GFP::4CL1	GFP::4CL1 73A5::RFP 98A3 HCT	0.004	0.00	**Yes**
4CL1::GFP	4CL1::GFP 73A5::RFP	0.005	0.00	**Yes**
GFP::4CL1 73A5::RFP	GFP::4CL1 73A5::RFP 98A3 HCT	0.006	0.01	**Yes**
4CL1::GFP 73A5::RFP	GFP::4CL1 73A5::RFP	0.007	0.01	No
4CL1::GFP	GFP::4CL1	0.010	0.39	No
GFP::4CL1	GFP::4CL1 73A5::RFP	0.013	0.64	No
4CL1::GFP 73A5::RFP	4CL1::GFP 73A5::RFP 98A3 HCT	0.017	0.67	No
4CL1::GFP 73A5::RFP 98A3 HCT	GFP::4CL1 73A5::RFP 98A3 HCT	0.025	0.71	No

Re-localization of the 4CL-1 upon co-expression of the four potential partners. (Data of Figure 69A). N. Number of analyzed images; SD. Standard deviation.

Name	N	Average	SD
HCT::GFP	99	1.2	0.3
HCT::GFP 93A3::RFP	99	0.9	0.3
HCT::GFP 98A3::RFP 73A5 4CL1	99	1.1	0.4
GFP::HCT	99	1.5	0.5
GFP::HCT 98A3::RFP	99	1.1	0.4
GFP::HCT 98A3::RFP 73A5 4CL1	99	1.1	0.4

Confidence in ANOVA tests	1,56907E-23	****

Posthoc test: Bonferroni-Holm		Critical value	P	Significant?
HCT::GFP	HCT::GFP 93A3::RFP	0.004	6.4E-12	**Yes**
GFP::HCT	GFP::HCT 98A3::RFP	0.004	4.2E-11	**Yes**
GFP::HCT	GFP::HCT 98A3::RFP 73A5 4CL1	0.004	2.9E-10	**Yes**
HCT::GFP	GFP::HCT	0.005	1.9E-05	**Yes**
HCT::GFP 93A3::RFP	HCT::GFP 98A3::RFP 73A5 4CL1	0.006	6.4E-05	**Yes**
HCT::GFP	HCT::GFP 98A3::RFP 73A5 4CL1	0.013	0.02	No
HCT::GFP 98A3::RFP 73A5 4CL1	GFP::HCT 98A3::RFP 73A5 4CL1	0.025	0.27	No
GFP::HCT 98A3::RFP	GFP::HCT 98A3::RFP 73A5 4CL1	0.050	0.81	No

Re-localization of HCT upon co-expression of the four potential partners (Data of Figure 69B). N. Number of analyzed images; SD. Standard deviation.

Supplemental data of FLIM experiments

Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
GFP::HDEL	2850	128	60	-	-
GFP::HDEL / Free RFP	2875	44	15	0	0.0
GFP::HDEL / 4CL1::RFP	2775	212	17	3	0.2
GFP::HDEL / RFP::4CL1	2646	305	17	7	0.9
GFP::HDEL / HCT::RFP	2843	216	17	0	0.0
GFP::HDEL / RFP::HCT	2719	222	28	5	0.4

Data of Figure 63: Interaction of soluble proteins and ER protein detected by FLIM. N. Number of analyzed images; SD. Standard deviation.

Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
73A5::GFP	2022	91	81	-	-
73A5::GFP / Free RFP	1969	81	31	3	0.2
73A5::GFP / 4CL1:RFP	1940	79	41	4	0.2
73A5::GFP / RFP::4CL1	2100	55	17	0	0.0
98A3::GFP	2145	153	75	-	-
98A3::GFP / Free RFP	2117	163	35	1	0,1
98A3::GFP / HCT:RFP	2152	120	25	0	0.0
98A3::GFP / RFP::HCT	1900	118	26	11	1.1

Data of Figure 65: Interaction of 4CL-1 and HCT with their partner P450s detected by FLIM. N. Number of analyzed images; SD. Standard deviation.

Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
73A5::GFP	2022	91	81	-	-
73A5::GFP / Free RFP	1969	81	31	3	0.2
73A5::GFP / HCT:RFP	1988	32	17	2	0.1
73A5::GFP / RFP::HCT	2028	71	17	0	0.0
98A3::GFP	2145	153	75	-	-
98A3::GFP / Free RFP	2117	163	35	1	0.1
98A3::GFP / 4CL1:RFP	2061	134	28	4	0.4
98A3::GFP / RFP::4CL1	2088	214	24	3	0.3

Data of Figure 67: Interaction of 4CL-1 and HCT with their non-partner P450s detected by FLIM. N. Number of analyzed images; SD. Standard deviation.

Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
73A5::GFP	2022	91	81	-	-
73A5::GFP / HDEL::RFP	1932	63	32	4	0.2
73A5::GFP / 98A3:RFP	1694	151	26	16	1.6
73A5::GFP / 73A5:RFP	1857	67	26	8	0.5
98A3::GFP	2145	153	75	-	-
98A3::GFP / HDEL::RFP	2076	178	42	3	0.4
98A3::GFP / 73A5:RFP	1901	137	30	11	1.2
98A3::GFP / 98A3:RFP	1647	93	20	23	2.1

Data of Figure 68: P450 oligomerization detected by FLIM. N. Number of analyzed images; SD. Standard deviation.

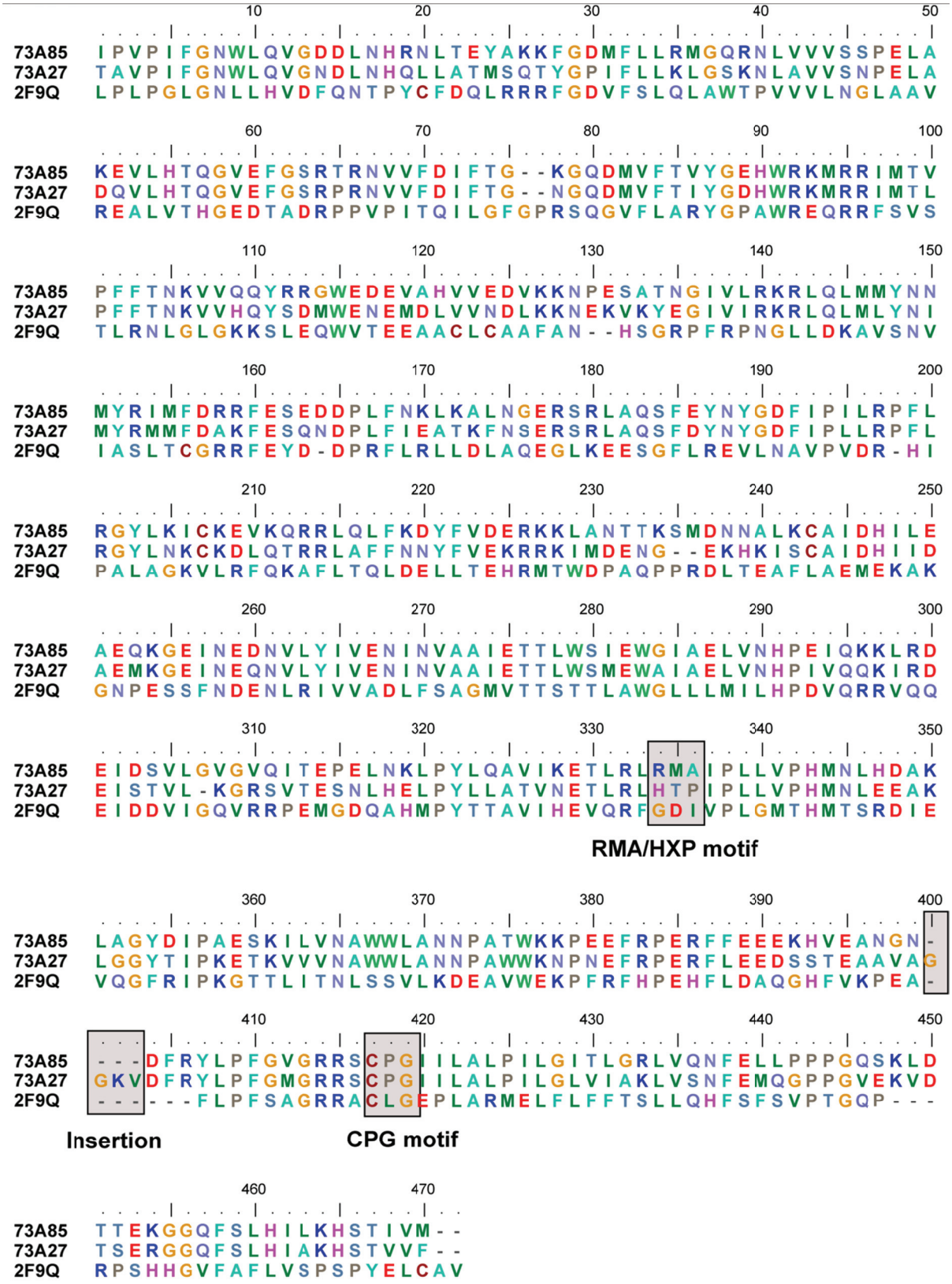
Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
73A5::GFP	2022	91	81	-	-
73A5::GFP / 4CL1:RFP / 98A3 / HCT	2020	49	16	0	0.0
73A5::GFP / RFP::4CL1 / 98A3 / HCT	1937	76	21	4	0.3
73A5::GFP / HCT:RFP / 98A3 / 4CL1	1872	59	17	7	0.4
73A5::GFP / RFP::HCT / 98A3 / 4CL1	1837	74	17	9	0.6
98A3::GFP	2145	153	75	-	-
98A3::GFP / HCT:RFP / 73A5 / 4CL1	1744	106	10	19	1.7
98A3::GFP / RFP::HCT / 73A5 / 4CL1	1919	135	17	11	1.1
98A3::GFP / 4CL1:RFP / 73A5 / HCT	1970	205	30	8	1.0
98A3::GFP / RFP::4CL1 / 73A5 / HCT	1956	268	28	9	1.4

Data of Figure 70: Protein interactions detected upon co-expression of 4CL-1, HCT, CYP73A5 and CYP98A3. N. Number of analyzed images; SD. Standard deviation

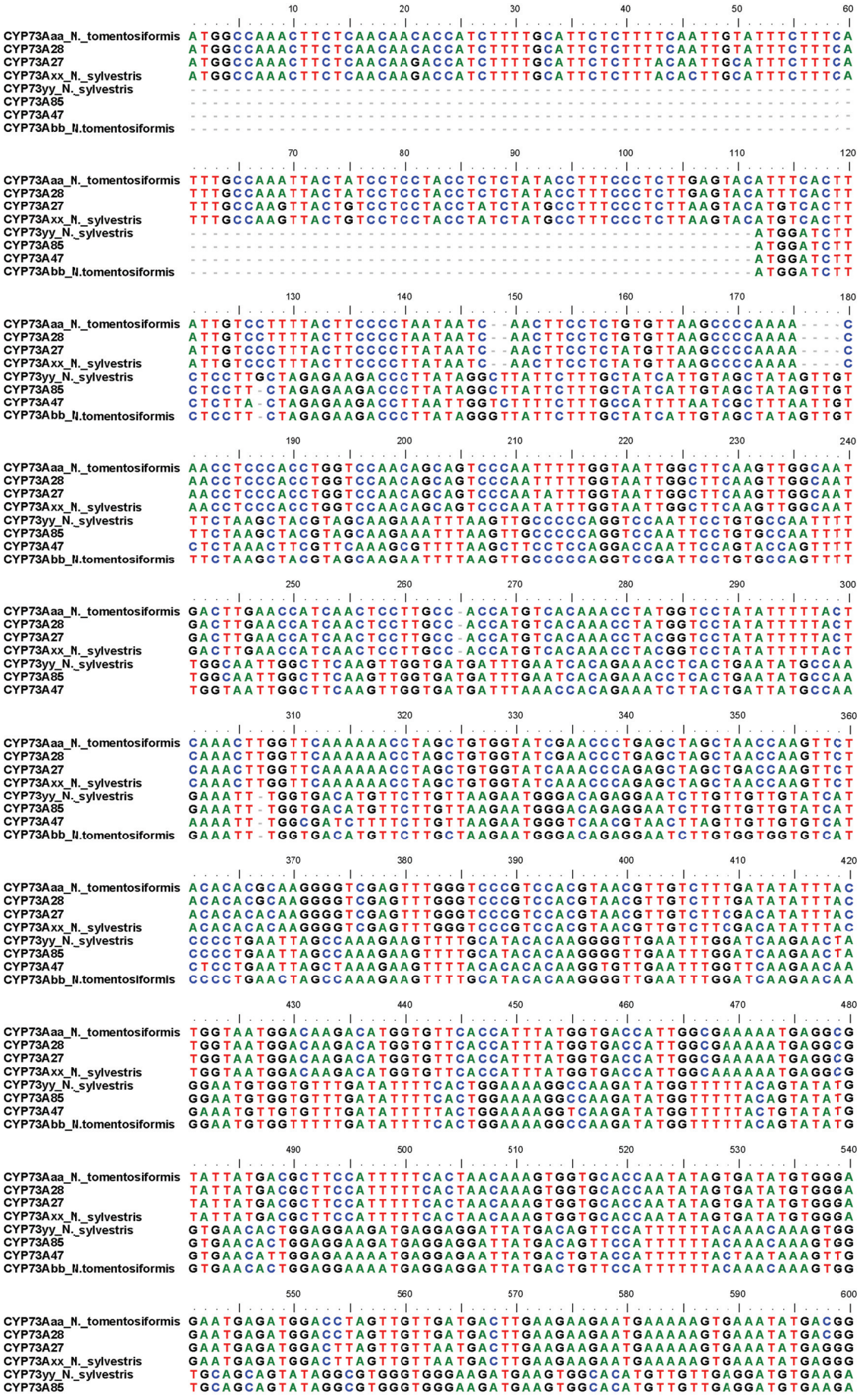
Proteins	Fluorescence half-life (ps)	SD	N	% FRET	SD FRET
4CL1::GFP	2571	164	17	-	-
4CL1::GFP / HCT::RFP	2570	58	17	0	0.0
4CL1::GFP / RFP::HCT	2592	75	17	0	0.0
GFP::4CL1	2643	33	17	-	-
GFP::4CL1 / HCT::RFP	2585	23	19	2	0.0
GFP::4CL1 / RFP::HCT	2651	25	18	0	0.0
4CL1::GFP	2571	164	17	-	-
4CL1::GFP / HCT::RFP / 73A5 / 98A3	2412	301	25	6	0.9
* 4CL1::GFP / HCT::RFP / 73A5 / 98A3	2024	28	9	21	1.4
4CL1::GFP / RFP::HCT / 73A5 / 98A3	2601	33	17	0	0.0
GFP::4CL1	2643	33	17	-	-
GFP::4CL1 / HCT::RFP / 73A5 / 98A3	2590	45	17	2	0.0
GFP::4CL1 / RFP::HCT / 73A5 / 98A3	2642	21	9	0	0.0

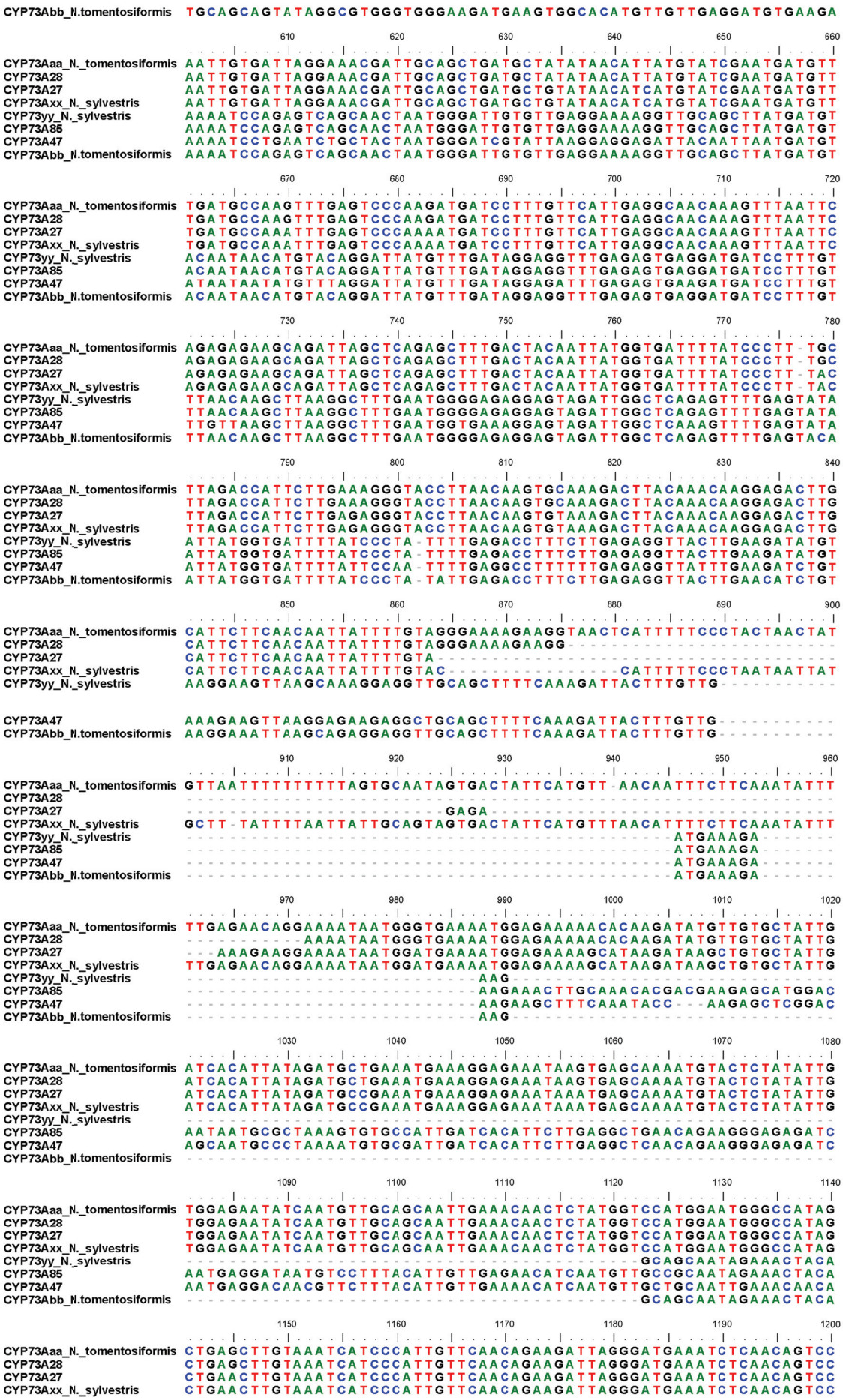
Data of Figure 71: Interaction of 4CL-1 and HCT detected by FLIM. Measurements performed, when clear re-localization of the soluble protein was observed are marked by an asterisk. N. Number of analyzed images; SD. Standard deviation

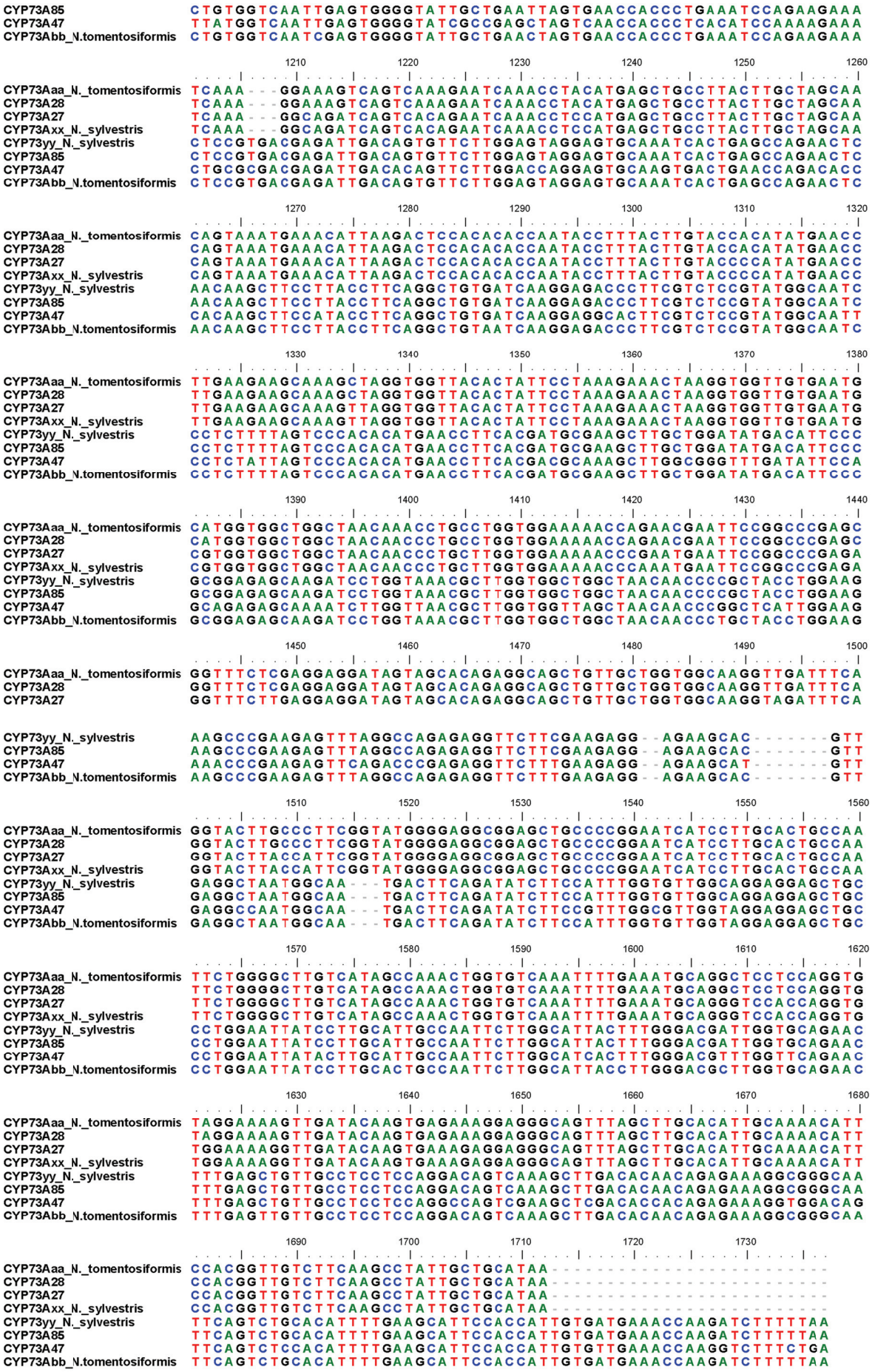
Alignment used for homology model generation



Alignment of *N. tabacum* parent sequences


CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis


CYP73Abb_N.tomentosiformis TGCAGCAGTATAGGCGTGGGTGGGAAGATGAAGTGGCACATGTTGTTGAGGATGTGAAGA
610 620 630 640 650 660
CYP73Aaa_N._tomentosiformis AATTGTGATTAGGAAACGATTGCAGCTGATGCTATATAACATTATGTATCGAATGATGTT
CYP73A28 AATTGTGATTAGGAAACGATTGCAGCTGATGCTATATAACATTATGTATCGAATGATGTT
CYP73A27 AATTGTGATTAGGAAACGATTGCAGCTGATGCTGTATAACATCATGTATCGAATGATGTT
CYP73Axx_N._sylvestris AATTGTGATTAGGAAACGATTGCAGCTGATGCTGTATAACATCATGTATCGAATGATGTT
CYP73yy_N._sylvestris AAAATCCAGAGTCAGCAACTAATGGGATTGTGTTGAGGAAAAGGTTGCAGCTTATGATGT
CYP73A85 AAAATCCAGAGTCAGCAACTAATGGGATTGTGTTGAGGAAAAGGTTGCAGCTTATGATGT
CYP73A47 AAAATCCTGAATCTGCTACTAATGGGATCGTATTAAGGAGGAGATTACAATTAATGATGT
CYP73Abb_N.tomentosiformis AAAATCCAGAGTCAGCAACTAATGGGATTGTGTTGAGGAAAAGGTTGCAGCTTATGATGT
670 680 690 700 710 720
CYP73Aaa_N._tomentosiformis TGATGCCAAGTTTGAGTCCCAAGATGATCCTTTGTTCATTGAGGCAACAAAGTTTAATTC
CYP73A28 TGATGCCAAGTTTGAGTCCCAAGATGATCCTTTGTTCATTGAGGCAACAAAGTTTAATTC
CYP73A27 TGATGCCAAATTTGAGTCCCAAAATGATCCTTTGTTCATTGAGGCAACAAAGTTTAATTC
CYP73Axx_N._sylvestris TGATGCCAAATTTGAGTCCCAAAATGATCCTTTGTTCATTGAGGCAACAAAGTTTAATTC
CYP73yy_N._sylvestris ACAATAACATGTACAGGATTATGTTTGATAGGAGGTTTGAGAGTGAGGATGATCCTTTGT
CYP73A85 ACAATAACATGTACAGGATTATGTTTGATAGGAGGTTTGAGAGTGAGGATGATCCTTTGT
CYP73A47 ATAATAATATGTTTAGGATTATGTTTGATAGGAGATTTGAGAGTGAAGATGATCCTTTGT
CYP73Abb_N.tomentosiformis ACAATAACATGTACAGGATTATGTTTGATAGGAGGTTTGAGAGTGAGGATGATCCTTTGT
730 740 750 760 770 780
CYP73Aaa_N._tomentosiformis AGAGAGAAGCAGATTAGCTCAGAGCTTTGACTACAATTATGGTGATTTTATCCCTT-TGC
CYP73A28 AGAGAGAAGCAGATTAGCTCAGAGCTTTGACTACAATTATGGTGATTTTATCCCTT-TGC
CYP73A27 AGAGAGAAGCAGATTAGCTCAGAGCTTTGACTACAATTATGGTGATTTTATCCCTT-TAC
CYP73Axx_N._sylvestris AGAGAGAAGCAGATTAGCTCAGAGCTTTGACTACAATTATGGTGATTTTATCCCTT-TAC
CYP73yy_N._sylvestris TTAACAAGCTTAAGGCTTTGAATGGGGAGAGGAGTAGATTGGCTCAGAGTTTTGAGTATA
CYP73A85 TTAACAAGCTTAAGGCTTTGAATGGGGAGAGGAGTAGATTGGCTCAGAGTTTTGAGTATA
CYP73A47 TTGTTAAGCTTAAGGCTTTGAATGGTGAAAGGAGTAGATTGGCTCAAAGTTTTGAGTATA
CYP73Abb_N.tomentosiformis TTAACAAGCTTAAGGCTTTGAATGGGGAGAGGAGTAGATTGGCTCAGAGTTTTGAGTACA
790 800 810 820 830 840
CYP73Aaa_N._tomentosiformis TTAGACCATTCTTGAAAGGGTACCTTAACAAGTGCAAAGACTTACAAACAAGGAGACTTG
CYP73A28 TTAGACCATTCTTGAAAGGGTACCTTAACAAGTGCAAAGACTTACAAACAAGGAGACTTG
CYP73A27 TTAGACCATTCTTGAGAGGGTACCTTAACAAGTGTAAAGACTTACAAACAAGGAGACTTG
CYP73Axx_N._sylvestris TTAGACCATTCTTGAGAGGGTACCTTAACAAGTGTAAAGACTTACAAACAAGGAGACTTG
CYP73yy_N._sylvestris ATTATGGTGATTTTATCCCTA-TTTTGAGACCTTTCTTGAGAGGTTACTTGAAGATATGT
CYP73A85 ATTATGGTGATTTTATCCCTA-TTTTGAGACCTTTCTTGAGAGGTTACTTGAAGATATGT
CYP73A47 ATTATGGTGATTTTATCCAA-TTTTGAGGCCTTTTTTGAGAGGTTATTTGAAGATCTGT
CYP73Abb_N.tomentosiformis ATTATGGTGATTTTATCCCTA-TATTGAGACCTTTCTTGAGAGGTTACTTGAACATCTGT
850 860 870 880 890 900
CYP73Aaa_N._tomentosiformis CATTCTTCAACAATTATTTTGTAGGGAAAAGAAGGTAACTCATTTTTCCCTACTAACTAT
CYP73A28 CATTCTTCAACAATTATTTTGTAGGGAAAAGAAGG
CYP73A27 CATTCTTCAACAATTATTTTGTA
CYP73Axx_N._sylvestris CATTCTTCAACAATTATTTTGTAC-----------CATTTTTCCCTAATAATTAT
CYP73yy_N._sylvestris AAGGAAGTTAAGCAAAGGAGGTTGCAGCTTTTCAAAGATTACTTTGTTG
CYP73A47 AAAGAAGTTAAGGAGAAGAGGCTGCAGCTTTTCAAAGATTACTTTGTTG
CYP73Abb_N.tomentosiformis AAGGAAATTAAGCAGAGGAGGTTGCAGCTTTTCAAAGATTACTTTGTTG
910 920 930 940 950 960
CYP73Aaa_N._tomentosiformis GTTAATTTTTTTTTTAGTGCAATAGTGACTATTCATGTT-AACAATTTCTTCAAATATTT
CYP73A28
CYP73A27 GAGA
CYP73Axx_N._sylvestris GCTT-TATTTTAATTATTGCAGTAGTGACTATTCATGTTTAACATTTTCTTCAAATATTT
CYP73yy_N._sylvestris ATGAAAGA
CYP73A85 ATGAAAGA
CYP73A47 ATGAAAGA
CYP73Abb_N.tomentosiformis ATGAAAGA
970 980 990 1000 1010 1020
CYP73Aaa_N._tomentosiformis TTGAGAACAGGAAAATAATGGGTGAAAATGGAGAAAAACACAAGATATGTTGTGCTATTG
CYP73A28 AAAATAATGGGTGAAAATGGAGAAAAACACAAGATATGTTGTGCTATTG
CYP73A27 AAGAAGGAAAATAATGGATGAAAATGGAGAAAAGCATAAGATAAGCTGTGCTATTG
CYP73Axx_N._sylvestris TTGAGAACAGGAAAATAATGGATGAAAATGGAGAAAAGCATAAGATAAGCTGTGCTATTG
CYP73yy_N._sylvestris AAG
CYP73A85 AAGAAACTTGCAAACACGACGAAGAGCATGGAC
CYP73A47 AAGAAGCTTTCAAATACC---AAGAGCTCGGAC
CYP73Abb_N.tomentosiformis AAG
1030 1040 1050 1060 1070 1080
CYP73Aaa_N._tomentosiformis ATCACATTATAGATGCTGAAATGAAAGGAGAAATAAGTGAGCAAAATGTACTCTATATTG
CYP73A28 ATCACATTATAGATGCTGAAATGAAAGGAGAAATAAGTGAGCAAAATGTACTCTATATTG
CYP73A27 ATCACATTATAGATGCCGAAATGAAAGGAGAAATAAATGAGCAAAATGTACTCTATATTG
CYP73Axx_N._sylvestris ATCACATTATAGATGCCGAAATGAAAGGAGAAATAAATGAGCAAAATGTACTCTATATTG
CYP73yy_N._sylvestris
CYP73A85 AATAATGCGCTAAAGTGTGCCATTGATCACATTCTTGAGGCTGAACAGAAGGGAGAGATC
CYP73A47 AGCAATGCCCTAAAATGTGCGATTGATCACATTCTTGAGGCTCAACAGAAGGGAGAGATC
CYP73Abb_N.tomentosiformis
1090 1100 1110 1120 1130 1140
CYP73Aaa_N._tomentosiformis TGGAGAATATCAATGTTGCAGCAATTGAAACAACTCTATGGTCCATGGAATGGGCCATAG
CYP73A28 TGGAGAATATCAATGTTGCAGCAATTGAAACAACTCTATGGTCCATGGAATGGGCCATAG
CYP73A27 TGGAGAATATCAATGTTGCAGCAATTGAAACAACTCTATGGTCCATGGAATGGGCCATAG
CYP73Axx_N._sylvestris TGGAGAATATCAATGTTGCAGCAATTGAAACAACTCTATGGTCCATGGAATGGGCCATAG
CYP73yy_N._sylvestris GCAGCAATAGAAACTACA
CYP73A85 AATGAGGATAATGTCCTTTACATTGTTGAGAACATCAATGTTGCCGCAATAGAAACTACA
CYP73A47 AATGAGGACAACGTTCTTTACATTGTTGAAAACATCAATGTTGCTGCAATTGAAACAACA
CYP73Abb_N.tomentosiformis GCAGCAATAGAAACTACA
1150 1160 1170 1180 1190 1200
CYP73Aaa_N._tomentosiformis CTGAGCTTGTAAATCATCCCATTGTTCAACAGAAGATTAGGGATGAAATCTCAACAGTCC
CYP73A28 CTGAGCTTGTAAATCATCCCATTGTTCAACAGAAGATTAGGGATGAAATCTCAACAGTCC
CYP73A27 CTGAACTTGTAAATCATCCCATTGTTCAACAGAAGATTAGGGATGAAATCTCAACAGTCC
CYP73Axx_N._sylvestris CTGAACTTGTAAATCATCCCATTGTTCAACAGAAGATTAGGGATGAAATCTCAACAGTCC


CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1210 1220 1230 1240 1250 1260
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1270 1280 1290 1300 1310 1320
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1330 1340 1350 1360 1370 1380
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1390 1400 1410 1420 1430 1440
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1450 1460 1470 1480 1490 1500
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1510 1520 1530 1540 1550 1560
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1570 1580 1590 1600 1610 1620
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1630 1640 1650 1660 1670 1680
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis
1690 1700 1710 1720 1730
CYP73Aaa_N._tomentosiformis
CYP73A28
CYP73A27
CYP73Axx_N._sylvestris
CYP73yy_N._sylvestris
CYP73A85
CYP73A47
CYP73Abb_N.tomentosiformis

Review "Phenolamides: Bridging polyamines to the phenolic metabolism"

Phytochemistry 71 (2010) 1808–1824

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Molecules of Interest

Phenolamides: Bridging polyamines to the phenolic metabolism

Jean-Etienne Bassard, Pascaline Ullmann, François Bernier, Danièle Werck-Reichhart *

Institute of Plant Molecular Biology, CNRS UPR 2357, University of Strasbourg, 28 rue Goethe, 67000 Strasbourg, France

ARTICLE INFO

Article history:
Received 26 July 2010
Accepted 3 August 2010
Available online 26 August 2010

Keywords:
Phenolamides
Polyamines
Phenylpropanoid metabolism
Defence
Bioactivity
Flower development
Adaptation

ABSTRACT

Phenolamides constitute a diverse and quantitatively major group of secondary metabolites resulting from the conjugation of a phenolic moiety with polyamines or with deaminated aromatic aminoacids. This review summarizes their bioactivities and their reported roles in plant development, adaptation and defence compared to those of their polyamine precursors. The most conclusive recent developments point to their contribution to cell-wall reinforcement and to direct toxicity for predators and pathogens, either as built-in or inducible defence. Phenolamides were often considered as accumulated end-chain products. Recent data bring a light on their biosynthesis and suggests their possible contribution in the branching of the phenylpropanoid metabolism.

1. Introduction

Phenolamides are frequently referred to as hydroxycinnamic acid amides (HCAA) or phenylamides. They have been reported throughout the plant kingdom, usually as main phenolic constituents of reproductive organs and seeds. They are regarded either as products of polyamine catabolism or as polyamines or phenolics storage forms. However, they seem to have specific functions in plant development and defence, as metabolic intermediates and final products. This review sums-up HCAAs occurence, biosynthesis and potential functions in plants in connection with polyamines and phenolic metabolism.

2. Polyamines

The name "polyamines" refers to aliphatic organic compounds with more than one amino group. Putrescine, spermidine and spermine are the most widespread in all living organisms especially in actively proliferating tissues. They are also the most common in plants, while cadaverine was also reported in legumes. Recent data suggest that the spermine isomer thermospermine might be also widespread and was present before spermine in aerial plants (Kakehi et al., 2008; Minguet et al., 2008). Norspermidine, norspermine and homospermine on the other hand were described as taxonomic markers of Bryophytes, Pteridophytes, Gymnosperms and Fungi (Hamana and Matsuzaki, 1985). The positive charge of polyamines at physiological pH confers them the property to bind negatively charged macromolecules or to modulate the activity of some ion channels. In plants, polyamines are found not only in the cytoplasm but also in vacuoles, plastids and mitochondria (Kumar et al., 1997). Several recent reviews provide a good overview on the current knowledge on polyamine biosynthesis, catabolism and bioactivity, including their roles in plant development and adaptative responses. Those will thus be just summarized briefly.

2.1. Biosynthesis and catabolism

Intracellular concentrations of polyamines are quite high and range from several hundreds of micromolar to a few millimolar. Due to their important biological functions, polyamine concentrations are very tightly controlled. While biosynthesis, catabolism, conjugation and transport contribute to polyamines homeostasis, catabolism also contributes to their bioactivity.

Polyamine biosynthesis involves similar pathways in bacteria, animals and plants (Kusano et al., 2008). Two alternative pathways starting from L-arginine have been confirmed in plants (Fig. 1). The ornithine decarboxylase pathway is favoured in meristematic and dividing cells, while the arginine decarboxylase pathway predominates in mature tissues and in response to environmental stress (Flores, 1991). Only one of these pathways was proposed to be operating in *Arabidopsis thaliana* since no ornithine decarboxylase has been predicted from its annotated genome (Hanfrey et al., 2001). However, Tassoni and coworkers (2003) reported ornithine decarboxylase activity associated with the plastid membranes in *Arabidopsis* leaves. In legumes, cadaverine is derived from lysine

* Corresponding author.
E-mail address: daniele.werck@ibmp-ulp.u-strasbg.fr (D. Werck-Reichhart).

0031-9422/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2010.08.003

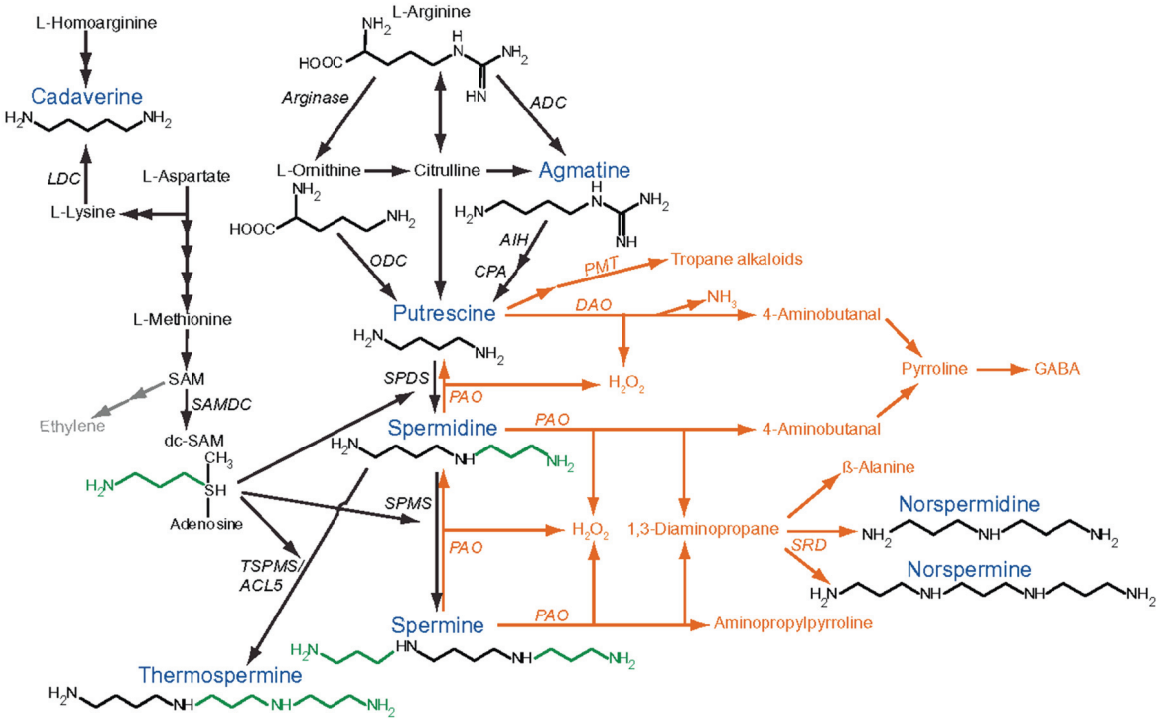


Fig. 1. Polyamine metabolism. Schematic biosynthetic pathways for common polyamines and related metabolites are indicated by black lines and catabolic processes in red. Common polyamines are in blue and enzymes in italics. Abbreviations: ACL5, ACAULIS5; ADC, arginine decarboxylase; AIH, agmatine iminohydrolase; CPA, *N*-carbamoylputrescine amidohydrolase; DAO, diamine oxidase; dc-SAM, decarboxylated *S*-adenosylmethionine; GABA, γ-aminobutyric acid; LDC, lysine decarboxylase; ODC, ornithine decarboxylase; PAO, polyamine oxidase; PMT, putrescine *N*-methyltransferase; SAM, *S*-adenosylmethionine; SAMDC, *S*-adenosylmethionine decarboxylase; SPDS, spermidine synthase; SPMS, spermine synthase; SRD, Schiff-base reductase/decarboxylase; TSPMS, thermospermine synthase.

via a lysine decarboxylase (EC 4.1.1.18). Most of the genes in the polyamine biosynthetic pathway are duplicated in plants. In *Arabidopsis*, two spermidine synthases and a spermine synthase associate to form a metabolon (Panicot et al., 2002). *ACL5*, one of the predicted *A. thaliana* spermine synthase genes was recently shown to encode a protein with thermospermine synthase activity (Knott et al., 2007). This suggested the presence of this spermidine isomer in plants and may explain the characteristic phenotypes in stem elongation and vascular development observed upon *ACL5* defect (Hanzawa et al., 1997; Clay and Nelson, 2005; Kakehi et al., 2008). Interestingly, a duplication of the genes encoding spermidine/spermine synthases has led to the evolution of putrescine *N*-methyltransferases, catalyzing the first step to the secondary metabolites nicotine and tropane alkaloids in Solanales (Minguet et al., 2008).

S-adenosylmethionine carboxylase (SAMDC) seems a critical point of regulation of polyamine homeostasis in all organisms (see e.g. Kumar et al., 1996; Martin-Tanguy, 1997) and might be responsible for antagonism between synthesis of higher polyamines and ethylene. Polyamine-controlled upstream ORF-mediated translational regulation of SAMDC was reported in mammals and plants. In *Arabidopsis*, two uORFs contribute to SAMDC regulation and translation repression at reduced levels of polyamines (Hanfrey et al., 2005).

Plant polyamine catabolism was recently reviewed by Moschou et al. (2008a) and is also summarized in Fig. 1. The main polyamine catabolic pathway, via both diamine and polyamine oxydases, generates H_2O_2. Catabolism of putrescine via diamine oxydase (DAO) also generates γ-aminobutyric acid, often considered as a mediator of plant defence. Plant polyamine oxydases (PAOs) catalyze the retroconversion from spermine and spermidine to spermidine and putrescine, respectively. They do not require acetylation of their substrates as shown for animal enzymes. Several DAO and PAO genes are found in plant genomes. Some of these enzymes were shown to be targeted to peroxisomes (Moschou et al., 2008a; Kamada-Nobusada et al., 2008) or the vacuole (Cervelli et al., 2004). This raises the question of their specific roles in the plant.

Polyamine homeostasis further relies on their conversion into secondary metabolites such as nicotine or tropane alkaloids in Solanaceae (Kusano et al., 2008) or in their conjugation with hydroxycinnamic acids (see below) or with proteins, hemicellulose or lignin (Creuss et al., 1991). Cross-linking mediated by transglutaminases might play a significant role in polyamine bioactivity for flower development and compatibility in reproduction (Serafini-Fracassini and Del Duca, 2008). Extracellular transglutaminases are expected to support organization of the cell-wall and pollen tube growth. For cytosolic enzymes, tubulin and actin have been identified as substrates (Del Duca et al., 1997), while chloroplastic forms would protect thylakoid proteins and Rubisco.

2.2. Role in plant growth and development

Polyamines, in particular spermidine, are essential for plant viability. Double insertional mutants of both arginine decarboxylase genes, both spermidine synthase genes, or of *SAMDC1* and *SAMDC4* *S*-adenosylmethionine carboxylases are embryo lethal in *A. thaliana* (Urano et al., 2005; Imai et al., 2004a; Ge et al., 2006). Mutants defective in spermine synthesis however grow normally (Imai

et al., 2004b). Spermidine is a precursor of deoxyhypusine, required for post-translational modification of the eukaryotic translational initiation factor eIF5A (Park, 2006) which may provide an explanation for the absolute putrescine/spermidine requirement for plant embryogenesis. In addition, single gene mutants in polyamine biosynthesis usually show strong phenotypes suggesting cross-talk with phytohormone control of plant development (Kumar et al., 1997; Kusano et al., 2007). One obvious reason for such a cross-talk is the share of the precursor *S*-adenosylmethionine for both ethylene and spermidine/spermine biosynthesis.

In bacteria and animals, polyamines are described as increasing cell growth and being essential for organ functionality (Kusano et al., 2008). This is however less clearly established in plants, except in the case of thermospermine. *A. thaliana* thermospermine synthase mutant *acl5* shows a very severe dwarf phenotype and xylem proliferation (Hanzawa et al., 2000) that is rescued by exogenous application of thermospermine but not spermine (Kakehi et al., 2008). *ACL5* is preferentially expressed in maturing xylem and is under negative feedback regulation by thermospermine (Clay and Nelson, 2005; Muñiz et al., 2008). Analysis of *acl5* suppressor mutants indicates that thermospermine controls stem elongation via the bHPLP-type transcription factor *SAC51* and one of its upstream open reading frames located in 5′ leader sequence (Takahashi and Kakehi, 2009). Plants seem to have acquired thermospermine biosynthesis ability at an early stage of evolution by horizontal gene transfer from prokaryotes (Minguet et al., 2008).

In addition to signalling, polyamines have other direct and indirect effects on plant development via mechanisms that also impact plant adaptation and defence. Those include electrostatic binding to macromolecules such as DNA, RNA and proteins, developmentally controlled cross-linking mediated by transglutaminases as mentioned above (Serafini-Fracassini and Del Duca, 2008), and interaction with ion channels and receptors, resulting in regulation of Ca^{++}, Na^{+} and K^{+} homeostasis (Kusano et al., 2008). The action of polyamines on cation channels depends on their net positive charge with spermine > spermidine >> putrescine. Production of H_2O_2 upon polyamine catabolism is another level of regulation of plant growth and development. Polyamine oxidases were shown to be developmentally regulated and associated to cell-wall strengthening, lignification and programmed cell death (Moschou et al., 2008a). In plants, senescence of different organs can be delayed by polyamines. Their role in developmental cell death is best documented in the case of *Nicotiana tabacum* flower corolla (Della Mea et al., 2007a,b).

2.3. Polyamines and abiotic stress

An extensive literature depicts the correlation between polyamines levels and physiological perturbations, as well as the protective effects of polyamines observed in response to environmental stress (Bouchereau et al., 1999; Kakkar and Sawhney, 2002; Urano et al., 2003; Alcázar et al., 2006a; Groppa and Benavides, 2008), including heavy metal stress (Groppa et al., 2001, 2008).

In grapevine, cultivars accumulating large amounts of free polyamines, exhibit a higher tolerance to osmotic stress than other cultivars (Paschalidis et al., 2009). In agreement with this observation, exogenous application of polyamines is reported to protect against abiotic stress (Chattopadhayay et al., 2002). Studies using loss-of-function mutants or transgenic plants overexpressing the genes for polyamine biosynthetic enzymes also support a role of polyamines in stress resistance. For instance, *Arabidopsis* or potato lines overexpressing the SPDS gene are tolerant to multiple environmental stresses (Kasukabe et al., 2004, 2006). Overexpression of the arginine decarboxylase (ADC2) gene in *Arabidopsis* results in increased putrescine level and drought tolerance (Alcázar et al., 2010). On the other hand, *ADC2* loss-of-function plants are more sensitive to salt stress (Urano et al., 2004). The *Arabidopsis acl5/spms* mutant is unable to produce spermine and is hypersensitive to salt and drought stresses (Yamaguchi et al., 2007). The symptoms are reversed by exogenous spermine. Remarkably, this mutant exhibits symptoms of Ca^{2+} deficiency (Yamaguchi et al., 2007), which points to an involvement of polyamines in Ca^{2+} regulation (Kusano et al., 2008).

The mode of action of polyamines during abiotic stress is still not well understood. It seems pleiotropic and depends on the stress and the plant species. A common feature of all types of abiotic stresses (salt, osmotic, drought, UV, heavy metals...) is oxidative stress promoted by the formation of reactive oxygen species (ROS). Polyamines induce antioxidative enzymes, increase amount of carotenoids and limit lipid peroxidation (Verma and Mishra, 2005). They are reported to enhance heavy metal tolerance by protecting gluthatione reductase and superoxide dismutase (Groppa et al., 2001). Spermine and spermidine also prevent leakage of amino acids or electrolytes (Chattopadhayay et al., 2002). According to Groppa and Benavides (2008) polyamines stabilize macromolecules, proteins and membranes. They were shown to modulate electrostatic protein–protein interactions (Berwanger et al., 2009), to affect DNA–protein interactions, translocation of protein kinases and gene expression by selective inhibition of cytosine dependent DNA methylases (Kuznetsov et al., 2006). Polyamines also indirectly inhibit plasma membrane and vacuolar H^{+}-ATPase antiporters and a long-term polyamine decrease contributes to maintain cation–anion equilibrium in the cytoplasm (Janicka-Russak et al., 2010). Putrescine is an efficient stimulator of ATP synthesis and causes depolarization of membranes. Conversely spermidine and spermine are described as uncouplers and prevent thylakoid membrane energization and reactive oxygen formation (Ioannidis and Kotzabasis, 2007). Polyamines have also been proposed to play a role in photosynthesis since they are capable of reversing stress-induced damages in photosynthetic apparatus (Sfakianaki et al., 2006). Polyamine conjugation by transglutaminases, especially to Rubisco seems to have an important role in protecting this protein from protease action, thus preserving its photosynthetic efficiency (Serafini-Fracassini et al., 1995).

Abiotic and biotic stresses seem to induce the export of spermidine/spermine into the apoplast for PAO/DAO-mediated catabolism resulting in H_2O_2 production. Accumulation of H_2O_2 results either in the tolerance response or plant cell death (PCD), depending on the levels of intracellular polyamines (Moschou et al., 2008b). When polyamine anabolism predominates, catabolism PCD fails to occur. Complex mechanisms for polyamine signalling and the subsequent responses to generate ROS upon abiotic stress have been described (Moschou et al., 2008a; Toumi et al., 2010). They involve synergetic or antagonist roles of hormones depending on the type of stress and on the responses. For example, putrescine seems to modulate ABA biosynthesis (Cuevas et al., 2008). Reciprocally, ABA modulates polyamine metabolism at transcriptional and metabolite levels (Alcázar et al., 2006b). ABA contributes to the conversion of the bound or conjugated forms to the free soluble form of polyamine (Ben Hassine et al., 2009). Increased putrescine content seems to repress GA biosynthesis (Alcázar et al., 2005). The calcium sensor calcineurin B-like 3 (CBL3) that mediates calcium signalling is described as another modulator of polyamine biosynthesis (Oh et al., 2008), while induction of NO production by spermine and spermidine has been reported in *Arabidopsis* (Tun et al., 2006). Under anoxic conditions, NO can react with polyamines to produce NONOates. Spermine NONOate has been favoured as a chemical NO donor (Yamasaki and Cohen, 2006). Several enzymes involved in polyamine biosynthesis are inhibited by NO-mediated *S*-nitrosylation (Wang et al., 2006). Putrescine,

spermidine and spermine exert different roles during abiotic stress response. The ratio of (spermidine + spermine)/putrescine seems to be fundamental for plant tolerance and survival (Groppa and Benavides, 2008) and modifications of this ratio seems a factor that controls plant response to different environmental cues.

2.4. Polyamines and disease resistance

Since the early 80s, a plethora of experiments have demonstrated implication of polyamines in plant disease resistance (Walters, 2003a,b). However, the molecular mechanisms involved are still unclear. The main problem for outlining clear rules is that total polyamine concentrations and ratios between individual polyamines markedly vary with plant species, with plant tissues and also with pathosystems investigated (Marina et al., 2008).

However, some aspects of this mode of action were recently elucidated (Fig. 2). A rapid biosynthesis and accumulation of polyamines was observed following fungal and viral infection or elicitor treatment. Spermidine and/or spermine levels were higher in the apoplast (Yamakawa et al., 1998; Marini et al., 2001; Cowley and Walters, 2002; Yoda et al., 2003, 2006; Marina et al., 2008; Moschou et al., 2008b). High apoplastic spermidine/spermine concentrations were shown to modulate three signals. One of them is Ca^{2+} influx. Spermine controlled Ca^{2+}/K^{+} channels and calcium-channel blockers attenuated this signal (Kusano et al., 2007). Another signal involved the action of spermidine and spermine as inhibitors of pectin dimerization. In the cell-wall, spermidine and spermine form complexes with pectins that prevent Ca^{2+}-induced pectin oligomerization and thus inhibits their biological activity as endogenous elicitors (Messiaen and Van Cutsem, 1999). Conversely, putrescine stabilizes pectic fragments. A third signal results from the production of H_2O_2 from spermine/spermidine oxidation by polyamine oxidases (Cona et al., 2006). This is a common defensive strategy in host- and non-host-hypersensitive response (Walters, 2003a,b; Mitsuya et al., 2009; Yoda et al., 2003, 2009). The use of DAO/PAO inhibitors before inoculation impaired host defence reactions (Mitsuya et al., 2009). H_2O_2 produced by PAO seems to contribute to the second phase of the oxidative burst (Angelini et al., 2008), and seems decisive for the fate of the cell. Polyamine catabolism-dependent H_2O_2 would trigger the hypersensitive response (HR) through a so-called "spermine-signalling pathway" (Kusano et al., 2007). The latter was suggested to involve mitochondrial dysfunctions, photorespiration and ATP consumption, activation of caspases, activation of protein kinases and of MAPK cascade, ER stress and the Unfold Protein Response, increased expression of HR marker genes and zinc finger genes, with as final result defence responses or HR-like cell death (Takahashi et al., 2003, 2004; Uehara et al., 2005; Kusano et al., 2008; Mitsuya et al., 2009). Many other defence mechanisms triggered by abiotic stress are also activated. For example, increased expression of PAO induces peroxidase, superoxide dismutase, catalase and activities of enzymes, contributing to redox homeostasis and limiting oxidative damage. Furthermore H_2O_2 released by PAO is required for the peroxidase-mediated deposition of lignin and suberin polyphenolic domain in wounded tissues (Angelini et al., 2008).

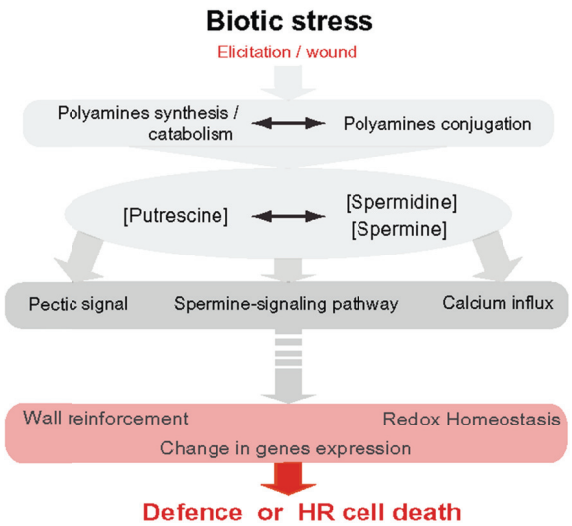


Fig. 2. Polyamine-mediated response in plant defence against biotic stress.

Polyamines thus have a dual function in defence response. A protective effect results from oxygen radical scavenging, increased expression of antioxidant enzymes, and generation of H_2O_2 as a signal. Toxic pro-oxidative effects occur when unrestrained increase in their content leads to intensive degradation, while H_2O_2 levels beyond a given threshold lead to PCD (Kuznetsov et al., 2006). Putrescine can prevent PCD induced by PAO-generated H_2O_2 (Yoda et al., 2009). The ratio of (spermidine + spermine)/putrescine and the ratio between polyamines, pectins and Ca^{2+} also seem to be important. A tight control of the homeostasis of putrescine and spermidine/spermine appears to predetermine cell fate between defence and PCD (Kusano et al., 2008).

3. Phenolamides (also termed phenylamides or hydroxycinnamic acids amides)

A large proportion of the polyamines found in plants are mono-, di- or tri-substitued with phenolic acids such as coumaric, caffeic and ferulic acids. These phenolic acids can also be conjugated with arylmonoamines like tyramine, tryptamine, octopamine or anthranilate. Phenolamides form a large class of plant secondary metabolites that are abundant in plants. A few examples are depicted in Fig. 3. Depending on the presence of a residual free amino group, the resulting conjugate can be basic or neutral which conditions its physicochemical properties (Edreva et al., 2007). Conjugation of polyamines with phenolics significantly reduces their polarity and hydrophilicity. This may favour their translocation, stability and compartmentation. While conjugation can be a mean to regulate the pools of both parent compounds and to store phenolics and bioactive polyamines, conjugates were often regarded as final and accumulated products. Turnover and translocation of conjugates were however quite early described (Martin-Tanguy, 1985, 1997; Havelange et al., 1996) and, more recently, interconversion between free and conjugated precursors has found some support (Luo et al., 2009). In addition, new data suggest that they might also be metabolic intermediates contributing to the complexity of the phenolic metabolism and possibly to its cross-talk with nitrogen metabolism (Morant et al., 2007; Matsuno et al., 2009).

Abundance, diversity and distribution of phenolamides are well documented (Martin-Tanguy et al., 1978; Martin-Tanguy, 1985; Bienz et al., 2005; Rogoza et al., 2005). However, surprisingly little is known about their biological functions. Potential roles of phenolamides in plant development and defence have been summarized in two reviews (Facchini et al., 2002; Edreva et al., 2007).

3.1. Phenolamides associated with plant growth and floral initiation

High concentrations of phenolamides are most often associated with organ growth, in particular floral induction and development (Martin-Tanguy, 1985; Kakkar and Rai, 1993).

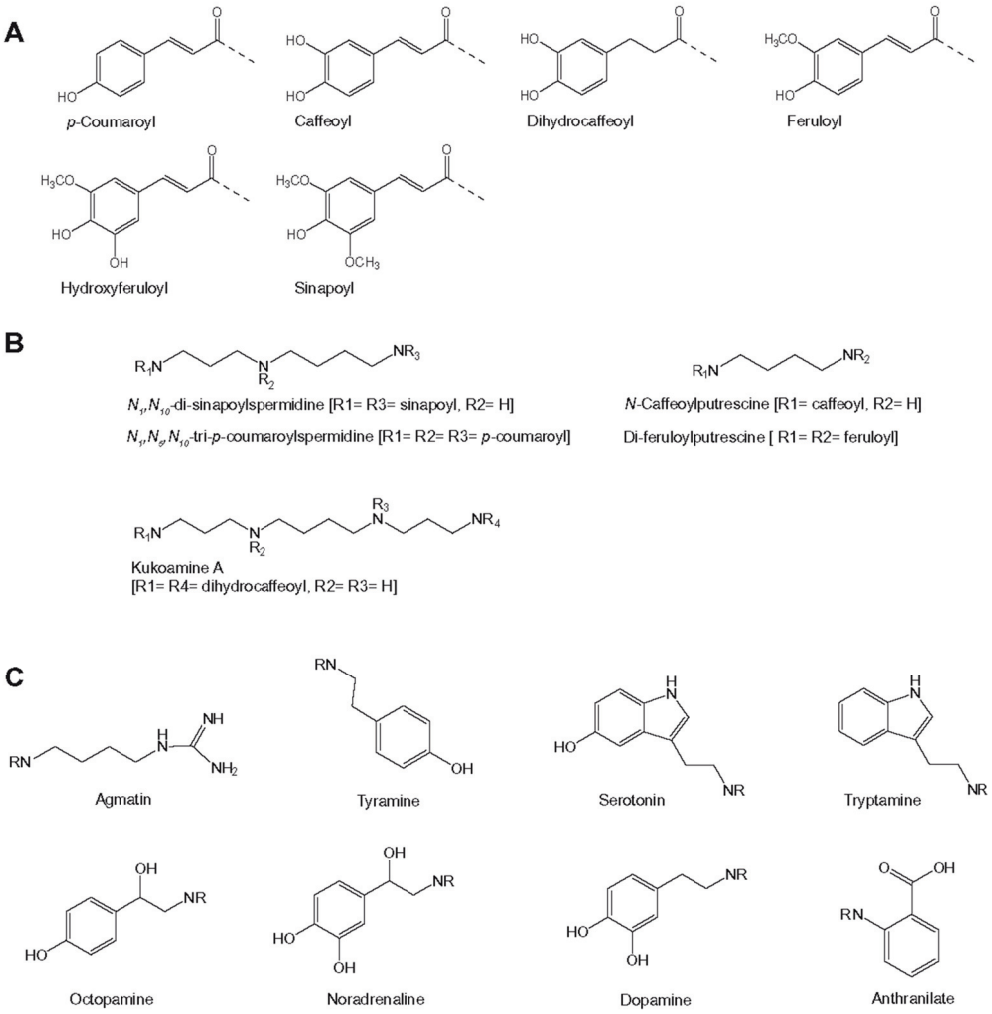


Fig. 3. Chemical structures of common phenolamides and their parent hydroxycinnamic acids. (A) Common acyl substituents found in phenolamides. (B) Examples of phenolamides derived from polyamines. (C) Most common arylamine moieties found in phenolamides.

Polyamides were proposed to be a necessary component of the mobile signal to flower development (Tarenghi and Martin-Tanguy, 1995; Havelange et al., 1996). Even if they are not considered as the florigen itself, their accumulation clearly correlates with flower initiation and development. Flowering plants usually contain little or no phenolamides in leaves. During floral transition however, phenolamides first accumulate in the upper leaves and apices (Martin-Tanguy, 1985, 1997). Subsequent accumulation in floral organs correlates with disappearance from the leaves. In *Sinapis alba*, free and conjugated putrescine were found in phloem leaf exudates after floral induction (Havelange et al., 1996). Their production and flowering response were inhibited with difluoromethylornithine, (an inhibitor of ornitine decarboxylase) and this inhibition could be reversed by application of putrescine to the roots. A role of phenolamides in floral development was also supported by their absence in mutants that do not flower (Martin-Tanguy, 1985) or by altered floral morphologies upon accumulation of spermidine conjugates (Malmberg and McIndoo, 1983).

Different classes of phenolamides are clearly associated with different floral organs (Martin-Tanguy, 1985; Aribaud and Martin-Tanguy, 1994a; Tarenghi and Martin-Tanguy, 1995). In tobacco, basic water soluble caffeoylputrescine and caffeoylspermidine accumulated in female organs are considered as markers of female fertility, while neutral di-*p*-coumaroylputrescines, di-*p*-coumaroylspermidines and *p*-coumaroyltyramine found in the anthers are considered as markers of male fertility (Cabanne et al., 1981). Diferuloyl conjugates are transiently detected after fertilization. In a similar way, high levels of neutral phenolamides were found in male flowers from Araceae while basic phenolamides were detected in female flowers (Ponchet et al., 1980). Sterile flowers of the same plants were devoid of both types of compounds. Male sterility in maize was also linked to the absence of phenolamides in the anthers (Martin-Tanguy et al., 1982). Various studies have been carried out in different plant species to further correlate male sterility with polyamine biosynthesis and accumulation. In most cases, male sterility could be associated with reduced content in polyamines and conjugates, in particular

insoluble conjugates (Aribaud and Martin-Tanguy, 1994b; Li and Li, 1997; Tian et al., 1998; Guo et al., 2003). Higher levels of polyamines were however reported in the male sterile stamenless-2 mutant of tomato and related to abnormal stamen development (Rastogi and Sawhney, 1990).

If accumulation of specific phenolamides seems clearly correlated with specific stages of reproductive development, the accumulated molecules usually differ among plant species and a causal role could only be assessed indirectly. One approach used inhibitors of polyamine biosynthesis, such as DFMO (α-DL-difluoromethylornithine), a specific inhibitor of ornithine decarboxylase, which results in delayed flowering, leaf wrinkling and branching and reduced height (Burtin et al., 1991; Havelange et al., 1996). These effects could be reversed by putrescine feeding. Inhibitors of spermidine biosynthesis did not delay flowering but interfered with anther development and caused flower abortion, lack of pollen, heterostyly, replacement of anthers by petals, and infertility (Burtin et al., 1991). These defects were reversed by supplying spermidine. Both free polyamines and their conjugates were reduced after inhibitor treatments and restored upon putrescine feeding.

Alterations of plant development similar to those observed by treatment with DFMO were obtained by tobacco transformation with the root-inducing (Ri) TL-DNA from *Agrobacterium rhizogenes* and with 35S-*rolA* or 35S-*rolC* constructs (Martin-Tanguy et al., 1996; Martin-Tanguy, 1997). Genes carried by the Ri TL-DNA have specific effects on plant signalling and development. Ri TL-DNA transformation resulted in wrinkled leaves, shortened internodes, increased branching, flower retardation and strongly reduced fertility, all correlated with a decrease in polyamine conjugates. This effect was reversed by putrescine and tyramine feeding that also restored accumulation of polyamine conjugates in the Ri LT-DNA plants. In plants expressing the 35S-*rolA* construct, with abnormal flower and stamen development (anthers contained little or no viable pollen), application of free putrescine and tyramine was not sufficient to restore fertility. However, grafting on a wild-type rootstock that was induced to flower restored flower development, but flowers aborted without additional supply of putrescine and tyramine (Martin-Tanguy et al., 1996). In 35S-*rolC* plants, male sterility seemed more specifically correlated with an inability to form feruloyl amine derivatives due to a lack of putrescine:feruloyl-CoA and tyramine:feruloyl-CoA transferase activities (Martin-Tanguy, 1997). Accordingly, free amines failed to reverse the sterility phenotype.

3.2. Major pollen constituents

From the plethora of analytical data collected over several decades, the presence of neutral phenolamides in the male gametophyte appears as a clear and constant feature. In particular, di- and tri-substituted hydroxycinnamoyl conjugates emerge as the major metabolites detected in the anthers and more specifically in pollen grains. Putrescine and/or spermidine conjugates have been detected in the anthers or pollen of all Angiosperms, including Dicots such Rosideae (Strack et al., 1990; Tarenghi and Martin-Tanguy, 1995), Brassicaceae (Havelange et al., 1996), Solanaceae (Leubner-Metzger and Amrhein, 1993; Kang and Back, 2006), Asteraceae (Aribaud and Martin-Tanguy, 1994b; Werner et al., 1995), Betulaceae, Fagaceae and Juglandaceae (Meurer et al., 1986, 1988; Bokern et al., 1995; Meurer-Grimes, 1995), Acanthaceae, (Werner et al., 1995), and of Monocots like maize (Martin-Tanguy et al., 1982), Liliidae (Youhnovski et al., 1998, 2001) or Araceae (Ponchet et al., 1980). High polyamine conjugate contents were also reported in Gymnosperm sexual buds (Daoudi and Bonnet-Masimbert, 1998; Fraga et al., 2004). Interestingly, while phenolamides were present in all taxa, they differed in the chain length of the polyamine (putrescine or most often spermidine), the degree and pattern of substitution of the polyamine chain (di- or tri-substituted) and the degree of hydroxylation of the phenolic rings (from *p*-coumaroyl to sinapoyl). Phenolamides thus seem essential for pollen development, viability or germination. The reason for their structural variability among taxa is however not understood.

A. thaliana mutants affected in the biosynthesis (detailed below) of the hydrocinnamoylspermidines recently provided tools to investigate their role in pollen development (Fellenberg et al., 2008, 2009; Grienenberger et al., 2009; Matsuno et al., 2009). Promoter-GUS fusions constructs for several genes in the biosynthetic pathway revealed a very high expression in the tapetal cells. N^1,N^5-di(hydroxyferuloyl)-N^{10}-sinapoyl spermidine represented a major constituent in the pollen coat, as shown by autofluorescence loss in mutant and easy methanol wash out (Grienenberger et al., 2009). In addition, total depletion led to occasional pollen grain distortions, which might be indicative of participation to the structure of the pollen wall (Grienenberger et al., 2009). Although more frequent silique abortion and reduced seed set was reported for some mutants (Fellenberg et al., 2008; Matsuno et al., 2009), no strong impact on pollen viability, germination and fertility was reported.

3.3. Other plant tissues and developmental stages

Basic and especially neutral polyamines such as di-feruloylputrescine, di-feruloylspermidine, feruloyl tyramine have been detected in large amounts upon seed development and maturation in Monocots like maize and rice (Martin-Tanguy, 1985; Bonneau et al., 1994a). In maize, quite large fluctuations in specific phenolamide compounds were observed in the embryo and endosperm at different developmental stages (Martin-Tanguy, 1985). Highest phenolamide concentrations correlated with maximal rates of DNA, RNA and protein synthesis. Fertility was reported to be correlated with the level of phenolamides in developing seed. Microanalysis and fluorescence localized the main sites of accumulation of basic phenolamides in the embryo and of the most abundant neutral di-*p*-coumaroylputrescine and di-feruloylputrescine in the pericarp and/or aleurone cells (Sen et al., 1994). In maize, fairly large amounts of phenolamine conjugates were recovered in corn bran or corn fibers extracts (Moreau et al., 2001). More recently, di-sinapoyl spermidine and its glucose conjugate were found as major phenolic derivatives in the seeds of *A. thaliana* (Luo et al., 2009), their synthesis relying on a sinapoyl-CoA-specific acyl transferase (SDT).

Both in rice and *Arabidopsis*, a sharp decline in phenolamides was observed upon seed germination, with a concomitant increase in polyamines in the case of rice (Bonneau et al., 1994a; Luo et al., 2009). Interestingly, the *Arabidopsis* SDT transferase was found to be able to catalyze the reverse reaction, regenerating spermidine and sinapoyl CoA, and was expressed at high level during early stages of seed germination (Luo et al., 2009). It can thus be proposed that phenolamides constitute a polyamine and hydroxycinnamoyl storage unit that can be remobilized upon seed germination. In agreement with this hypothesis, high amine conjugate content in rice seeds was positively correlated with seed viability (Bonneau et al., 1994a). As recently pointed by Luo et al. (2009), accumulation of phenolamides in tissues that express high levels of reversible SDT requires compartmentation of the products. Such compartmentation can be supported by glycosylation of the spermidine conjugates, as observed in *Arabidopsis* seeds. Alternatively, product accumulation may result from transport into sink (e.g. inflorescence) tissues.

In tobacco roots, abundant free and wall-bound feruloyltyramine were reported (Hagel and Facchini, 2005). Conjugated amines were also detected in rice seedling roots (Bonneau et al.,

1994b), and various hydroxycinnamoyl derivatives of tyramine or spermidine in the roots of tropical plants (Lee et al., 2004; Zamble et al., 2006). Interestingly, recent work of Luo et al. (2009) indicates that an *Arabidopsis* polyamine acyltransferase specific for *p*-coumaroyl-CoA and spermidine is specifically expressed in root tips while phenolamides could not be detected in the roots. They may thus be further converted in such plant tissues.

Finally, accumulation of basic phenolamides in the stolons before potato tuber initiation suggested a role in tuberization (Paynot et al., 1983). This was later questioned by work of Leubner-Metzger and Amrhein (1992, 1993) who found no correlation between *in vitro* tuberization and accumulation of phenolamides and distribution of phenolamides in different species of Solanaceae and their tuber development.

3.4. Phenolamides and cell-wall cross-linking

It progressively appeared that cell-wall cross-linking has essential roles in plant development and defence. Those include control cell of elongation and plant growth, nucleation of lignin, cell-wall stiffening and thickening upon ageing, response to wounding and pathogen attack, and biodegradability by microbial and endogenous enzymes (Passardi et al., 2004; Buanafina, 2009). Cross-linking largely results from the formation of diferuloyl bridges between polysaccharides (via arabinoxylans) and lignin (Buanafina et al., 2009). However, polyamines and conjugates, feruloyltyramine in particular, seem to constitute other important bridging agents, especially in grasses and upon wounding or pathogen challenge. Large amounts of polyamines were found strongly bound to cellular structures, especially in roots (Vallée et al., 1983; Hagel and Facchini, 2005). Feruloyltyramine and feruloyloctopamine are ether-linked to the cell-wall of natural and wound potato periderm (Negrel et al., 1996).

In tobacco thin layer cultures, polyamine synthesis inhibitors induced various developmental changes in the cell-wall and middle lamella as well as a loss of cell adhesion (Berta et al., 1997). Those were reversed when the culture was supplemented with polyamines. This suggested a role of polyamines in cell-wall cross-linking but did not imply involvement of phenolamides. Wound-healing and infection are usually associated with deposition of suberin in cell-walls. Feruloylamides such as feruloyltyramine and feruloyloctopamines seem to constitute major components of the suberin polymer (Borg-Olivier and Monties, 1993; Graça, 2009). In addition, phenolamides have been suggested as the preferential substrates for some amine oxidases and peroxidases, respectively involved in H_2O_2 generation in the apoplast and in H_2O_2-dependent polymerization in the cell-wall (Aribaud and Martin-Tanguy, 1994b; Bernards et al., 1999).

Clarke (1982) first described rapid accumulation of highly fluorescent hydroxycinnamoyl tyramine and octopamine conjugates in methanol-soluble granules and their subsequent binding to the cell-wall following exposure of potato tubers to *Phytophthora infestans*. Irreversible incorporation of radiolabelled hydroxycinnamoyltyramine into cell-wall residue of TMV-infected *Nicotiana tabaccum* was then reported by Negrel and Jeandet (1987). Feeding *Nicotiana* cell cultures with radiolabelled tyramine also demonstrated incorporation into polymeric material at higher rate after elicitation with chitosan (Villegas et al., 1990). A similar response was reported for fungal infection of potato cell cultures and leaves that triggered incorporation *p*-coumaroyltyramine and feruloyltyramine in the cell-wall (Keller et al., 1996; Schmidt et al., 1998).

In the case of *Botrytis*-infected onion epidermal cells, several autofluorescent hydroxycinnamoyltyramine derivatives were found accumulated as granular deposits near the penetration site with associated polarization of the actin microfilaments. The major autofluorescent compounds (feruloyltyramine and its 3′-methoxy derivative) were subsequently bound via ether links onto the cell-wall (McLusky et al., 1999). More recently, extraction of suberized potato scab (*Streptomyces*-induced) lesions confirmed the presence of large amounts of feruloyltyramines and feruloyloctopamines, and also revealed the presence of minor components identified as cross-linked dimers of feryloyltyramines and feruloyloctopamines (King and Calhoun, 2005). Elevated expression and activity of the enzymes involved in feruloyltyramine synthesis provided a track to investigate the significance and impact of its insertion in the cell-walls. Engineered constitutive expression of tyrosine decarboxylase and of hydroxycinnamoyl-CoA:tyramine hydroxycinnamoyltransferase were thus shown to lead to an increased incorporation of hydroxycinamoyltyramines in the cell-wall at the wound sites, and to reduced digestibility (Facchini et al., 1999; Hagel and Facchini, 2005; Guillet and De Luca, 2005).

3.5. Bioactive compounds

Phenolamides were often described as bioactive compounds with antiviral, antibacterial antifungal, insecticidal, deterrent or therapeutic activities. For example, *N*-feruloyltyramine was isolated as the most active garlic component, suppressing P-selectin expression (Park, 2009). It is thus expected to play a major role in garlic positive effect on cardiovascular system by the inhibition of platelet activation. *N-trans*-feruloyltyramine was also described as antitumoral (Park and Schoene, 2002), antimycobacterial (Mata et al., 2004), as melanogenesis inhibitor in mouse melanoma cell (Efdi et al., 2007) and inhibitor of cyclooxygenase (Park, 2007). This led to attempts to engineer its production in *Escherichia coli* (Kang et al., 2009) and rice (Park et al., 2009a).

Kukoamines, initially isolated from medicinal plants such as *Lycium chinense*, have attracted attention for their hypotensive effects and anti-trypanosomal activity (Funayama et al., 1980, 1995; Ponasik et al., 1995). Interestingly, kukoamine A (N^1,N^{12}-bis(dihydrocaffeoyl)spermine) and related dihydrocaffeoylated spermidines and spermines have recently also been identified in potato tuber and other Solanaceae such as tomato fruit or *Nicotiana sylvestris* leaves (Parr et al., 2005). Other phenolic amides isolated from the root bark of *L. chinense*, dihydro-*N*-caffeoyltyramine, *trans-N*-caffeoyloctopamine, *cis*- and *trans-N*-caffeoyltyramine displayed antifungal activity against *Candida albicans* (Lee et al., 2004) or anti-inflammatory properties via suppression of cycloxygenase expression (Han et al., 2010). Similarly, the tri-substituted spermidines representing the major *Quercus alba* pollen constituents were shown to decrease mycelial growth of *Pyronophora avenae* and to reduce powdery mildew infection of barley seedlings (Walters et al., 2001).

In other cases antioxydant and radical scavenging activities were reported (e.g. Bors et al., 1989; Calvin et al., 1998; Son and Lewis, 2002; Han et al., 2002; Zamble et al., 2006). Potentiated radical scavenging and quenching of singlet oxygen by phenolamides relative to parent compounds has been evidenced (Bors et al., 1989; Velikova et al., 2007). All (*E*)-spermidine conjugates such as those present in the pollen coat undergo very easy photoisomerization of the phenolic acid side chains at 365 nm. Their strong absorbance in the 270–330 nm range was thus postulated to have a protective function for germinal cells (Hu et al., 1998; Bienz et al., 2005).

3.5.1. For plant defence against microorganisms

While a defensive role of phenolamides against plant pathogens is well documented, in many cases their defensive function is only deduced from a correlation between metabolite accumulation and stronger resistance to the pathogen. For example, jasmonic acid was shown to promote a strong increase in the local and systemic concentrations of phenolic putrescine and spermidine conjugates

in barley leaves. This increase in phenolamide concentrations correlated with a reduction in powdery mildew infection (Walters et al., 2002). More documented is the impact of feruloylamines accumulated in tobacco upon hypersensitive response to TMV (Martin-Tanguy, 1985). Accumulation of these compounds correlated with the hypersensitive reaction. Their application on leaf discs caused a significant reduction of the number of virus-induced lesions (Martin-Tanguy et al., 1976). In addition, TMV inoculation was shown to promote incorporation of radiolabelled tyramine and feruloyltyramine into the acid-insoluble fraction of the cell-wall (Negrel and Jeandet, 1987). Interestingly, it has been proposed that the rare presence of viruses in flowers and seeds is related to their high contents in phenolamides (Martin-Tanguy, 1985; Edreva et al., 2007). Antimicrobial role of feruloyltyramine is often attributed to cell-wall strengthening and inhibition of pathogen penetration. As mentioned above, this is supported by the accumulation of fluorescent methanol-soluble granules observed in various plant-pathogen interactions near the penetration site and prior to papilla formation (Clarke, 1982; McLusky et al., 1999). Newman et al. (2001) also observed *in vitro* inhibition of bacterial growth by coumaroyltyramine and feruloyltyramine that appeared to determine incompatible interaction of the pepper plant with *Xanthomonas campestris* while antifungal activity was reported for the feruloyltyramine isolated from *Allium* roots (Fattorusso et al., 1999).

Putrescine, spermidine or tyramine derivatives appear ubiquitous in higher plants. A diversity of more specific phenolic amides accumulated upon pathogen attack were shown to behave as phytoanticipins or phytoalexins. The most extensively described are probably avenanthramides produced in oat leaves upon crown rust infection (Mayama et al., 1981, 1982; Miyagawa et al., 1995). These *N*-hydroxycinnamoyl anthranilate derivatives form a family of compounds also found in large amounts in seed groats and hulls (Collins, 1989) which concentrations appear tightly correlated with genetic resistance to crown rust (Wise et al., 2008). They are incorporated in the cell-walls of elicited oat leaf segments floated on stable isotope precursors (Okazaki et al., 2004).

In barley, *p*-coumaroyl-hydroxyagmatin accumulates in response to fungal attack by *Erysiphe graminis*. It exhibits antifungal activity both *in vivo* and *in vitro* and seems involved in cross-linking process of papillae (von Röpenack et al., 1998). Oxydative dimerization of *p*-coumaroylagmatin and feruloylagmatin leads to preformed and pathogen-inducible antifungal metabolites called hordatins that exhibit spore germination inhibition activity (Stoessl and Unwin, 1970). Kristensen et al. (2004) extensively discussed their significance and role in plant defence that may also include a contribution to polymerization products in papillae and cell-walls and to bacterial encapsulation. Unexpectedly and although detected in low amounts, *p*-coumaroylagmatine was recently described as a metabolic marker of rosette leaves in *A. thaliana* (Matsuda et al., 2009).

In the case of onion, feruloyl-3′-methoxytyramine was the major product accumulated in response to *Botrytis alii* (McLusky et al., 1999). A direct antifungal activity could not be demonstrated for this compound, but local accumulation as autofluorescent paramural granules at the penetration site was interpreted as participation to the local peroxidative cross-linking of the cell-wall. Other 7′- and 5′,7′-hydroxylated forms of *p*-coumaryoltyramine, *p*-coumaroyloctopamine and *p*-coumaroylnoradrenaline are the major metabolites accumulated by tomato carrying specific pathogen resistance genes (von Roepenack-Lahaye et al., 2003). In this case, *p*-coumaroylnoradrenaline but not *p*-coumaroyloctopamine was shown to have antibacterial activity against *Pseudomonas syringae*. Other examples of the diversification of defence compounds include clovamide (caffeoylDOPA) and the associated tyrosine phenolamide in red clover (Tebayashi et al., 2000) or *N*-*p*-coumaroylserotonin and *N*-feruloylserotonin accumulated in bamboo upon infection with *Phyllostachys bambusoides*, the causative agent of witch broom disease (Tanaka et al., 2003). *N*-*p*-coumaroylserotonin antifungal activity was confirmed.

Transient phenolamide accumulation was also reported upon beneficial symbiotic interactions with arbuscular mycorrhizal fungi in onion and barley and interpreted as the initiation of a defence response (Grandmaison et al., 1993; Peipp et al., 1997). Reduction of hyphal branching and of growth of mycorrhizal fungi has been reported by Grandmaison et al. (1993). In pine seedling, ectomycorrhizal fungus inoculation rather resulted in an early phenolamide accumulation in the needles (Niemi et al., 2006).

3.5.2. For defence against insects

Basic plant phenolamides are very closely related to the polyamine conjugates found in the venoms of predaceous spiders and wasps. It was thus proposed that they might be involved in protection against arthropods or used as natural insecticides. Investigation with synthetic basic phenolamides indicated no antifeedant nor toxic activity toward a variety of lepidopteran larvae using semi-synthetic diets laced with N^1- and N^8-coumaroylspermidine. *In vitro* inhibition of glutamatergic crustacean and mammalian synaptic receptors by such spermidine and spermine derivatives was however observed (Klose et al., 2002; Fixon-Owoo et al., 2003). Oxidative decomposition of the phenolamides in the diet was proposed to explain these contradictory results.

Well supported insect deterrence activity was however recently reported for caffeoylputrescine. Tebayashi et al. (2007) demonstrated that ovipositional deterrence of the leafminer *Liriomyza trifolii* acquired by sweet pepper tissues upon ageing or jasmonic acid treatment was related to their ability to accumulate caffeoylputrescine. Cotyledon treatment with synthetic *p*-coumaroylputrescine also decreased oviposition. These results were recently comforted by the work of Kaur et al. (2010) who demonstrated control of caffeoylputrescine and dicaffeoylspermidine synthesis by the jasmonic acid activated *Na*MYB8 transcription factor in *Nicotiana attenuata*. *Na*MYB8-silenced plant lacked these compounds and allowed better performance of both specialist (*Manduca sexta*) and generalist (*Spodoptera littoralis*) caterpillars than wild-type plants. Consistently, exogenous application of synthetic caffeoylputrescine at physiological doses impaired growth of *M. sexta* caterpillars.

3.5.3. For adaptation to abiotic stress

The role of phenolamides in abiotic stress is difficult to dissociate from that of its phenolic and polyamine constituents (Bouchereau et al., 1999; Groppa and Benavides, 2008). Their more specific functions seem to principally rely on their antioxydant and radical scavenging properties. Those have been extensively discussed by Edreva et al. (2007). Phenolamides that are good substrates for peroxidases may in addition support elimination of H_2O_2 and, as mentioned above, cell-wall strenghtening in the apoplast. Conjugation and turnover may in addition impact polyamine cross-talk with ethylene.

The best documented example is the accumulation of conjugated putrescine observed in O_3-sensitive and O_3-tolerant tobacco lines exposed to ozone (Bors et al., 1989; Langebartels et al., 1991). Leaf injury of the O_3-sensitive line by ozone treatments was prevented to a large extent by root application of polyamines. The titers of soluble free and conjugated polyamines were concomitantly increased and the amounts of polyamines associated with cell-wall or membrane pellet fractions were elevated four to six times above control. Reactivity assays of polyamines and conjugates towards hydroxyl, *tert*-butoxyl, sulphite radicals and superoxide anions indicated high rate constants for putrescine conjugates only (Bors et al., 1989). Accordingly, exposition to ozone doses that did not cause any visible injury resulted in rapid conjugated putrescine accumulation in the O_3-tolerant tobacco lines but only a slow

one in the O_3-sensitive plants. Monocaffeoyl-putrescine accumulated in the apoplastic fluid of the tolerant plants, in agreement with a potential role as an extracellular oxyradical scavenger (Langebartels et al., 1991). Either ethylene or polyamines were found to be early induced, in agreement with their biosynthetic antagonistic routes and shared precursor. Slow phenolamide accumulation correlated with the development of necrotic lesions similar to the hypersensitive response to TMV infection.

Accumulation of phenolic conjugates was also recorded upon a diversity of abiotic stresses, such as K, Ca, Mg and P deficiencies (Delétang, 1974), sulphur starvation (Klapheck, 1983), water excess (Edreva et al., 2007), heat shock (Edreva et al., 1998). Conversely, salt-stress was reported to reduce formation of polyamine conjugates in the roots (Shevyakova et al., 2006). Involvement of phenolamides in plant adaptation to UV-B irradiation was recently suggested by the existence of common regulatory elements with herbivore elicitation (Kaur et al., 2010). This prediction has yet to be confirmed. Interestingly, a spinach-specific phenolamide (*N*-feruloyl-(3-hydroxy-4-methoxyphenyl)ethylamine) accumulates in response to sublethal doses of diphenylether herbicides that promote tetrapyrrole accumulation and oxidative stress (Suzuki et al., 1981). Higher doses also result in necrotic lesions.

3.6. Biosynthesis of phenolamides

Phenolamides are essentially derived from aminoacids. Their biosynthesis branches, on one side, on the core phenylpropanoid pathway at the level of hydroxycinnamoyl-CoA esters (Vogt, 2010). The hydroxycinnamoyl moiety thus derives from deaminated phenylalanine. The amine moiety results either from the ubiquitous polyamine pathway for the most widespread putrescine, spermidine and spermine conjugates found in all seed plants, either from just decarboxylated amino acids (and occasionally from an aminoacid, e.g. from tyrosine) (Fig. 4). The most common conjugate derived from a decarboxylated amino acid is feruloyltyramine that has been found in a broad range of species (Smith, 1977). Accordingly, the tyrosine decarboxylase gene characterized initially in tobacco (Farmer et al., 1999), potato (Schmidt et al., 1999) and opium poppy (Facchini et al., 1999) was since described in a variety of plants such as tomato (von Roepenack-Lahaye et al., 2003), *Camptotheca acuminata* (Lopez-Meyse and Nessler, 1997), *Ophiorrhiza pumila* (Yamazaki et al., 2003), rice (Kang et al., 2007) and *A. thaliana* (Lehmann and Pollmann, 2009). In some plants, more specific agmatine, tryptamine or phenylethylamine conjugates result from arginine, tryptophan or phenylalanine decarboxylation. The evolution, biochemistry and regulation of aminoacids decarboxylases were reviewed by Facchini et al. (2000). Decarboxylation products can be hydroxylated to form, for example, octopamine, dopamine, serotonin and noradrenaline that can also be conjugated to form phenolamide products (Matsuda et al., 2005). The 5-hydroxylation of tryptamine for the formation of serotonin can be catalyzed by a soluble tetrahydrobiopterin-dependent enzyme (Kang et al., 2007) or by a cytochrome P450 CYP71P1 (Fujiwara et al., 2010) in rice, expression of the latter being correlated with defence against rice blast. In pepper, two tryptophan decarboxylases have been described. One of them is only expressed upon fungal attack or ethylene for the synthesis of the hydroxycinnamoylserotonin phytoalexins (Park et al., 2009b). Anthranilate, another common amine moiety, derives from the aromatic amino acid precursor chorismate via an elicitor-inducible anthranilate synthase AS subunit. This enzyme is insensitive to tryptophan feedback and constitutes a paralogue of the enzyme involved in tryptophan biosynthesis (Bohlmann et al., 1996).

The conjugates can be further decorated via species-specific hydroxylation, methylation, cyclisation or coupling reactions. The

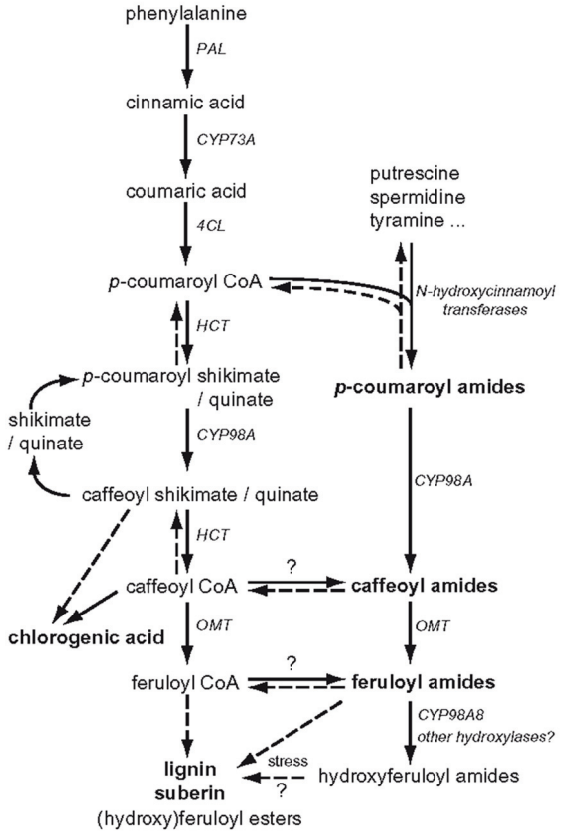


Fig. 4. Connection of phenolamides to the phenolic metabolism. PAL, phenylalanine ammonia lyase; 4CL, 4-coumaroyl-CoA ligase; HCT, hydroxycinnamoyl:shikimate/quinate hydroxycinnamoyl transferase; OMT, *O*-methyl transferase.

latter modifications are poorly described so far, except for the intramolecular phenol-oxidative coupling of N^1,N^{10}-bis(*p*-coumaroyl)spermidine forming the alkaloid lunarine in the seeds of *Lunaria annua*, and the final oxidative coupling of (*S*)-dihydroverbacine leading to the synthesis of macrocyclic spermine alkaloids aphelandrine and orantine from *Aphelandra* sp. (Nezbedova et al., 2001). In both cases, the reaction was proposed to be catalyzed by a cytochrome P450 enzyme, based on its regio/stereoselectivity, on the microsomal localization of the enzyme activity and NADPH-dependence. In the case of alephandrine formation, this assumption was further supported by the O_2-dependence and CO-inhibition of the reaction.

3.6.1. A panel of N-hydrocinnamoyl transferases

Coupling of the hydroxycinnamoyl and amine moieties is a critical step and can be considered as the real entry point to the phenolamide branch pathways. It is catalyzed by a diversity of soluble specific hydroxycinnamoyl transferases, several of them belonging to the superfamily of BAHD acyltransferases (D'Auria, 2006). In this family, five clades have been defined. *O*-hydroxycinnamoyl transferases producing phenolic esters belong to clade V (D'Auria, 2006; Petersen et al., 2009). *N*-hydroxycinnamoyl transferases using aliphatic amines as acyl acceptors (*N*-aliphatic AHCAT) belong to this same group, with the exception of barley agmatine:hydroxycinnamoyl transferase (ACT) found in clade IV (Burhenne et al., 2003) and *Arabidopsis* ACT surprisingly located in clade 1 (Muroi et al., 2009). A number of *N*-aliphatic AHCAT have

been purified from different plants and some related genes cloned and functionally characterized. Their properties are summarized in Table 1.

AHCAT are soluble enzymes and their specificity for the acyl acceptor and acyl donor varies, depending of the plant source. Di- and tri-acylated amines are commonly detected in plants. *Arabidopsis* spermidine transferases SCT (spermidine:dicoumaroyl transferase) and SDT (spermidine: di-sinapoyl transferase) can form di-acylated conjugates by sequential transfer of two hydroxycinnamoyl-CoA to spermidine (Luo et al., 2009), whereas the corresponding tobacco SHT transfers a single acyl group, the product formed *in vitro* identified as N^1-feruloylspermidine (Negrel et al., 1991). After incubation of *Arabidopsis* SHT (spermidine:hydroxycinnamoyl transferase) with spermidine and feruloyl-CoA, triferuloylspermidine was the major reaction product (Grienenberger et al., 2009). *N*-acylating enzymes using aliphatic tetra-amines (agmatine, spermine) or aromatic diamines (serotonin) have not yet been shown to transfer more than a single acyl group. Another group of *N*-hydroxycinnamoyl transferases uses aromatic amines as acyl acceptors. HCBT and HHT (Table 1) acylate anthranilate and 5-hydroxyanthranilate, respectively, for the biosynthesis of specific phytoalexins in carnation and *Arabidopsis*. Both belong to clade V of the BAHD family (Muroi et al., 2009).

All other enzymes described so far catalyze the synthesis of amides from tyramine. Tyramine:*N*-hydroxycinnamoyl transferases (THTs) were extensively studied in Solanaceae. THT enzyme and gene were first identified in tobacco (Negrel and Martin, 1984; Negrel and Javelle, 1997; Farmer et al., 1999). In addition to Solanaceae, THTs were biochemically and/or genetically characterized from Papaveraceae, Piperaceae and Gramineae (Table 1). Plant THTs are not members of the BAHD family. Analysis of their amino acid sequences (Farmer et al., 1999; Kang et al., 2006; Schmidt et al., 1999) revealed substantial homology with mammalian spermidine/spermine acetyltransferases (SSAT) and more generally GCN5-related *N*-acetyltransferases (GNAT). Eukaryotic SSAT (Lu et al., 1996; Pegg, 2008) and several microbial antibiotic *N*-acetyltransferases contain a highly conserved domain (RGFGIGS motif), where three amino acids residues (Arg, Gly, Gly) have been shown to be essential for activity and are conserved in plant THTs. Like human SSAT, plant THT enzymes also seem to be active as dimers (Lu et al., 1996; Schmidt et al., 1999). Analyses of purified proteins by SDS/PAGE and HPLC suggest that tobacco and opium poppy THT are homodimers with subunits of 24–28 kDa each (Farmer et al., 1999; Yu and Facchini, 1999).

Substrate specificities of THT enzymes have been investigated in detail. Their best acyl acceptor was in most cases tyramine. All share a marked preference for feruloyl-CoA compared to other hydroxycinnamoyl-CoAs. CASHT (serotonin:*N*-hydroxycinnamoyl transferase) and THT proteins isolated from pepper have clearly distinct substrate specificities. Their catalytic efficiencies were measured on purified, native or recombinant enzymes (Burhenne et al., 2003; Kang et al., 2006). Recombinant THT does not accept serotonin as substrate but efficiently catalyzes the synthesis of feruloyltyramine and *p*-coumaroyltyramine. When feruloyl-CoA is the acyl donor, CASHT has a sixfold lower catalytic efficiency with tyramine than with serotonin (Jang et al., 2004). A functional analysis using recombinant chimeric CASHT and THT proteins was performed to determine the specificity of the amine-binding domain (Kang et al., 2006) showing that tyrosine 149 is a critical amino acid residue controlling amine substrate specificity. Kinetic studies and product inhibition patterns on THT indicate that the mechanism of catalyzis is ordered (or iso-ordered) bi–bi, with hydroxycinnamoyl-CoA being the first substrate to bind the transferase (Hohlfeld et al., 1995; Negrel and Javelle, 1997).

The apparent K_m values reported for purified or recombinant *N*-hydroxycinnamoyl transferases usually vary between 1 and 10 μM for the best acyl donor (CoA ester) and from 22 to 76 μM for the preferred amine substrate (Table 1). This compares to values ranging from 50 to 600 μM for the acyl donors and 0.75–70 mM for quinate or shikimate in the case of the *O*-hydroxycinnamoyl transferases (Hoffmann et al., 2003; Niggeweg et al., 2004) that are assumed to drive the main flow of precursors to the cell-wall formation. Since phenolamides accumulate often at high concentrations in the plant tissues, an important question is "can they be remobilized to restore the amine and phenolic precursors?" The reversibility of the reaction was seldom evaluated. When feruloyl or *p*-coumaroyltyramine were incubated with potato THT and free coenzyme A, feruloyl or *p*-coumaroyl-CoA and free tyramine were formed (Hohlfeld et al., 1995). The high value calculated for the equilibrium constant $K = 1.3 \times 10^4$ and the relatively high negative value for $\Delta G'^{\circ}_{eq}$ (-23.5 kJ mol^{-1}) are in favour of a reaction mainly in the direction of the acylated products. The major spermidine conjugate found in *Arabidopsis* seeds is a 4′-*O*-glycosyl-di(sinapoyl)-spermidine (Luo et al., 2009). *In vitro*, di-sinapoyl spermidine was predominantly formed when recombinant SDT was incubated ($k_{cat} \approx 5\ s^{-1}$) with spermidine ($K_{mapp} = 37$ μM) and sinapoyl-CoA (8 μM), but SDT also catalyzes di-sinapoyl spermidine hydrolysis in the presence of CoA with a $k_{cat} \approx 38\ s^{-1}$ ($K_{mapp} = 35$ μM for di-sinapoyl spermidine and 84 μM for CoA). During germination, the decreased level of spermidine conjugates correlates well with the higher SDT activity. The available data however were not sufficient to determine flux direction *in planta*, which may depend on metabolite concentrations and physiological context.

Expression of *N*-hydroxycinnamoyl transferases has been associated with different stages of the plant development (Table 1) and is most often activated by viral (Negrel and Martin, 1984), fungal, bacterial infections and chemical elicitation (Hohlfeld et al., 1995; Ishihara et al., 1998; Von Roepenack-Lahaye et al., 2003; Yang et al., 2004; Muroi et al., 2009), or by wounding (Ishihara et al., 2000) and abiotic stress such as UV treatment (Back et al., 2001).

3.6.2. A versatile P450 family for 3′-hydroxylation of phenolic conjugates

For a long time it was taken for granted that hydroxycinnamoyl transferases were assembling *p*-coumaroyl-, caffeoyl-, feruloyl- or sinapoyl- and amine-building blocks to form terminal conjugates. Recent data rather suggest that phenolamides could also be metabolic intermediates contributing to form multiple branches on the phenylpropanoid pathway (Fig. 4). This hypothesis was initially supported by flux analyses in potato tuber indicating that both phenolic esters and amides were turned over, and that turnover of amides of tyramine and octopamine was selectively and strongly increased upon elicitor treatment (Matsuda et al., 2005). These studies also suggested that feruloyltyramine and feruloyloctopamine could result from direct hydroxylation and methylation of *p*-coumaroyltyramine and *p*-coumaroyloctopamine. This assumption has recently found some support with the demonstration that cytochrome P450 enzymes belonging to the CYP98 family can catalyze the *meta*-hydroxylation of phenolamides. It was first shown that coumaroyltyramine was taken as a substrate for *meta*-hydroxylation by CYP98 enzymes from different plants including *Arabidopsis* CYP98A3, sweet basil CYP98A13 and wheat CYP98A11 and CYP98A12 (Morant et al., 2007). The reaction catalytic parameters of wheat CYP98A11 and CYP98A12 indicated that gene duplication might have led to the evolution of more competent isoforms for the hydroxylation of phenolamides.

A striking example of such specialization has recently emerged from a reverse genetics approach of CYP98A8 and CYP98A9 function in *A. thaliana* (Matsuno et al., 2009; Fellenberg et al., 2009). Both of them were shown to result from a retroposition and accel-

Table 1
Characteristics of purified or recombinant hydroxycinnamoyl-CoA:amine *N*-(hydroxycinnamoyl) transferases.

Enzyme	Plant source	Size of monomer/ highest or tissue specific expression	Major acyl acceptors (in decreasing order of catalytic efficiency)	Major acyl donors (in decreasing order of catalytic efficiency)	NCBI Genbank Nucleotide Accession No.	References
ACT (EC 2.3.1.64)	*Hordeum vulgare*	Monomer 48 kDa	Aliphatic amines	Absolute specificity for agmatine (K_m: 5–8 μM)	- Cinnamoyl-CoA - Coumaroyl-CoA($K_m \approx 2$ μM) - Feruloyl-CoA - Caffeoyl-CoA	AY228552 AY234333	Bird and Smith (1981), Bird and Smith (1983) and Burhenne et al., 2003
ACT	*Arabidopsis thaliana*			Agmatine	- *p*-Coumaroyl-CoA - Feruloyl-CoA	BT011800(At5g61160)	Muroi et al. (2009)
SHT	*Nicotiana tabacum*			High specificity for spermidine	- *p*-Coumaroyl-CoA - Feruloyl-caffeoyl CoA		Negrel et al. (1991)
SpmHT/SHT	*Aphelandra tetragona*			Spermidine spermine	- *p*-Coumaroyl-CoA - Feruloyl-CoA		Hedberg et al. (1996)
SCT	*Arabidopsis thaliana*	54 kDa		Absolute specificity for spermidine ($K_m \approx 52$ μM)	- Absolute specificity for *p*-coumaroyl-CoA ($K_m \approx 10$ μM)	NM 128072 (At2g25150)	Luo et al. (2009)
SDT		53 kDa siliques, emerging seedlings		High specificity for spermidine ($K_m \approx 37$ μM)	- High specificity for sinapoyl-CoA ($K_m \approx 8$ μM)	NM 127915 (At2g23510)	
SHT		51 kDa tapetum of anthers		Absolute specificity for spermidine	- Feruloyl-CoA - Caffeoyl/coumaroyl-CoA - Cinnamoyl-CoA	NM 127464 (At2g19070)	Grienenberger et al. (2009)
PHT (EC 2.3.1.138)	*Nicotiana tabacum*	≈50 kDa		Putrescine cadaverine	- Caffeoyl-CoA - Feruloyl/cinnamoyl-CoA		Meurer-Grimes et al. (1989), Negrel (1989), Negrel et al. (1991) and Negrel et al. (1992)
HCBT (EC 2.3.1.44)	*Dianthus caryophyllus*	Monomer 53 kDa	Arylamines	High specificity for anthranilate	- Cinnamoyl/*p*-coumaroyl-CoA - Benzoyl-CoA	Z84383	Yang et al. (1997)
HHT	*Avena sativa*	48 kDa		5-Hydroxyanthranilate	- Feruloyl-CoA - Avenalumoyl-CoA - Cinnamoyl-CoA	AB076980 AB076981 AB076982	Ishihara et al. (1998) and Yang et al. (2004)
THT (EC 2.3.1.110)	*Nicotiana tabacum*	Homodimer subunits of ≈24 kDa		Tyramine ($K_m \approx 24$ μM) octopamine	- Feruloyl-CoA ($K_m \approx 1$ μM)	AJ005062AJ131767	Negrel and Martin (1984), Negrel and Javelle (1997) and Farmer et al. (1999)
THT	*Solanum tuberosum*	Heterodimer (≈49 kDa, gel filtration) subunits of ≈25 kDa		Tyramine ($K_m \approx 22$ μM) octopamine	- Cinnamoyl/feruloyl-CoA($K_m \approx 36$ μM) - *p*-Coumaroyl-CoA	AJ131768	Hohlfeld et al. (1995) and Hohlfeld et al. (1996)
THT	*Solanum tuberosum*	Dimer (63 kDa, gel filtration) subunits of ≈30 kDa		Tyramine ($K_m \approx 40$ μM) octopamine	- Cinnamoyl-CoA ($K_m \approx 60$ μM) - Feruloyl-CoA	AB061243	Schmidt et al. (1999)
THT	*Papaver somniferum*	Homodimer subunits of 25 kDa mature roots		Tyramine ($K_m \approx 76$ μM)	- Cinnamoyl-CoA ($K_m \approx 2$ μM) - Sinapoyl-CoA - *p*-Coumaroyl-CoA		Yu and Facchini (1999)
THT	*Solanum lycopersicum*	≈27 kDa		Tyramine ($K_m \approx 4$ μM) noradrenaline dopamine, octopamine	- Cinnamoyl-CoA - *p*-Coumaroyl-CoA (the only acyl donor tested)	AY081907 AY081905 AY081908 AY081906	Von Roepenack-Lahaye et al. (2003)
THT	*Capsicum annuum*	28 kDa young stems and roots		Tyramine ($K_m \approx 40$ μM) but not serotonin	- Feruloyl/*p*-coumaroyl-CoA ($K_m \approx 20$ μM)	AY819700	Kang et al. (2006)
CASHT	*Capsicum annuum*	28 kDa flowers		Serotonin ($K_m \approx 73$ μM) tyramine	- Caffeoyl-CoA - Feruloyl/*p*-coumaroylCoA	AF329463	Back et al. (2001), Jang et al. (2004) and Kang et al. (2006)
THT	*Zea mays*	40 kDa (gel filtration)		Tyramine, trypcamine ($K_m \approx 59$ μM) dopamine phenethylamine	- Feruloyl-CoA ($K_m \approx 5$ μM) - Sinapoyl-CoA - *p*-CoumaroylCoA		Ishihara et al. (2000)

Table 1 (continued)

Enzyme	Plant source	Size of monomer/ highest or tissue specific expression	Major acyl acceptors (in decreasing order of catalytic efficiency)	Major acyl donors (in decreasing order of catalytic efficiency)	NCBI Genbank Nucleotide Accession No.	References
THT	*Triticum aestivum*	Seedling, roots	Tyramine	– Sinapoyl-CoA – Feruloyl-CoA		Louis and Negrel (1991)
THT	*Hordeum vulgare*	Germinating barley roots	Tyramine phenylethylamine	– Sinapoyl-CoA – Feruloyl-CoA – *p*-CoumaroylCoA		Louis and Negrel (1991)

ACT, agmatine:*N*-coumaroyl transferase; CASHT, serotonin:*N*-hydroxycinnamoyl transferase; HCBT, hydroxycinnamoyl/benzoyl-CoA:anthranilate:hydroxycinnamoyl/benzoyl transferase; HHT, hydroxyanthranilate:hydroxycinnamoyl transferase; PHT, putrescine:*N*-hydroxycinnamoyl transferase; SCT, spermidine:*N*-dicoumaroyl transferase; SDT, spermidine:*N*-di-sinapoyl transferase; SHT, spermidine:*N*-hydroxycinnamoyl transferase; SpmHT, spermine:*N*-hydroxycinnamoyl transferase; THT, tyramine:*N*-hydroxycinnamoyl transferase.

erated evolution leading to the apparition of enzymes dedicated to the *meta*-hydroxylation of tri-*p*-coumaroylspermidine and, for CYP98A8, also to the *meta*-hydroxylation of tri-feruloylspermidine for the formation of N^1,N^5-di(hydroxyferuloyl)-N^{10}-sinapoyl spermidine, a major pollen coat constituent (Matsuno et al., 2009). These enzymes thus seem to be acting downstream of the SHT transferases mentioned above (Fig. 4). However, the function of these phenolamide hydroxylases does not appear restricted to the tapetum during pollen development. CYP98A9 gene expression was also detected in vascular tissues and root stele and tip (Matsuno et al., 2009). Expression of CYP98A8 was detected in seeds (Fellenberg et al., 2008). It is thus possible that pathways involving phenolamide hydroxylation are involved in other aspects of plant development and are active downstream of other polyamine hydroxycinnamoyl transferases such as those described by Luo et al. (2009). Interestingly, Matsuno et al. (2009) pointed the fact that, while the canonical *Arabidopsis* CYP98A3 preferred substrate is the shikimate ester of *p*-coumaric acid, it is also able to hydroxylate the tri-*p*-coumaroyl amide of spermidine, though with a lower efficiency.

3.6.3. Further decoration of the resulting phenolamides

Final products accumulated by the plant usually include feruloyl derivatives and sometimes sinapoylated compounds. Fellenberg et al. (2008) have recently demonstrated that AtTSM1, an *O*-methyltransferase belonging to the *A. thaliana CCoAMT* gene family, was exclusively expressed in the tapetum, where it was catalyzing the terminal methylation of tri-(5-hydroxyferuloyl)spermidine into N^1,N^5-di(hydroxyferuloyl)-N^{10}-sinapoyl spermidine. Further investigations also suggested that CCoAMT1, the methyltransferase involved in caffeoyl-CoA methylation in lignin biosynthesis and also associated with flavonoid and sinapoylmalate biosynthesis (Do et al., 2007), catalyzes methylation of tricaffeoylspermidine into tri-feruloylspermidine (Fellenberg et al., 2009). While hydroxycinnamoylspermidines are apparently not the preferred substrates for AtTSM1 and CCoAMT1 in vitro, it is interesting to note that their roles as phenolamide methylases seem to be validated *in vivo* (Fellenberg et al., 2008).

Enzymes involved in sugar conjugation of phenolamides have not been described yet. Glycosylated derivatives have however been identified during seed (Luo et al., 2009) and flower (Fellenberg et al., 2009) development. Their abundance was higher in *cyp98a8* mutants as compared to wild-type. Glycoside formation was also detected upon SCT overpression in *Arabidopsis* leaves (Luo et al., 2009). Glycosylation thus seems to occur upon intracellular accumulation of hydroxycinnamoylspermidines.

3.6.4. Regulatory genes

The different branches of the phenolamide pathways are obviously differentially regulated with, for example, basic and neutral phenolamides being under separate transcriptional control (see for example Matsuda et al., 2005; Kaur et al., 2010). However, information on their regulatory cascades is still scarce. It is well established that accumulation of caffeoylputrescine in Solanaceae is under the control of the jasmonate signalling pathway (Chen et al., 2006; Tebayashi et al., 2007; Paschold et al., 2007) and, recently, the first regulatory elements involved in this JA-response have been identified. Two homologous R2R3MYB transcription factors, *Nt*MYBJS1 from *N. tabaccum* and *Na*MYB8 from *N. attenuata*, were found to control *p*-coumaroyl- caffeoyl- and feruloylputrescine accumulation in BY-2 cell cultures and intact plants respectively (Gális et al., 2006; Kaur et al., 2010). Overexpression and antisense constructs of *Nt*MYBJS1 showed that it specifically controls the JA-response of a subset of genes in the phenylpropanoid and polyamine pathways. The MYB DNA-binding domain was able to specifically bind PAL-A and PAL-B promoters. Consistently,

*Nt*MYBJS1 overexpression led to the accumulation of hydrocinnamoylputrescines (Gális et al., 2006). The role of *Na*MYB8 was analyzed by Kaur et al. (2010) in the more ecologically relevant context of *N. attenuata*. They showed that its expression was activated by mechanical wounding, and amplified by the simultaneous application of *M. sexta* oral secretions. Caterpillar feeding-induced local and systemic accumulation of caffeoylputrescine and dicaffeoylspermidine was suppressed in *Na*MYB8 inverted-repeat silenced plants. *Na*MYB8 silencing also reduced transcriptional activation of a large set of genes related to the phenylpropanoid and polyamine pathways. As mentioned above, suppression of phenolamide production in the silenced plants was associated with an improved performance of generalist and specialist herbivores. The same study also demonstrated that *Na*MYB8 silencing suppresses the constitutive accumulation of caffeoylputrescine in young leaves and reproductive tissues. Thus *Na*MYB8 also controls caffeoylputrescine synthesis during plant development.

A similar study has been carried out by Shinya et al. (2007) and led to the identification of a β-glucan and laminarin responsive elements in tobacco BY-2 cells. This work allowed the identification of another R2R3MYB-type transcription factor termed *Nt*MYBGR1 that targets phenylpropanoid genes for glucan-induced accumulation of caffeoylputrescine and feruloylputrescine and might be representative of plant response to fungal attack.

3.7. Overlooked branching in the phenylpropanoid metabolism?

It gradually appears that phenolamides have to be regarded as metabolic intermediates rather than just final products. They undergo hydroxylation and methylation of the phenolic rings, can be stored as hexose conjugates and remobilized upon demand. High polyamine concentrations accumulate in the plant cells. The low K_m of *N*-hydroxycinnamoyl transferases for their amine acceptors results in the formation of significant pools of amide derivatives where *p*-coumaroyl shikimate/quinate concentrations are usually below detection threshold (Matsuda et al., 2005; our own observations). This might be corrected by a higher affinity of P450 enzymes for ester conjugates than for the corresponding amides, but the large gap in the size of the pools of precursors suggests that the role of phenolamides in relation to the phenylpropanoid pathway has to be reassessed. To evaluate the potential contribution of hydroxycinnamoyl amides in the formation and storage of guayacyl and syringyl units, a more extensive description of the enzymes catalyzing their hydroxycinnamoyl transfer, hydroxylation, methylation and glycosylation/deglycosylation steps will be required. Of particular interest will be the reverse reactions catalyzed by hydroxycinnamoyl transferases to convert the amides into CoA esters. More information is also required about fluxes through the amides and esters branches of conjugates in different plant tissues, and about their storage and transport.

Besides participation to branched pathways, phenolamides have clearly acquired specific functions in plant development and defence. Two of them seem to be shared by several compounds. Those are (1) a contribution to cell-wall cross-linking and reinforcement and (2) a direct toxicity for predators and pathogens, either as built-in or inducible defence. Differences in species-specific metabolites also account for evolution of specific defence or recognition systems. Some amine conjugates such as putrescine, spermidine or tyramine derivatives are present in all Angiosperms and probably beyond. They are expected to have more generic functions in plant structure, signalling, in particular in the reproductive cycle. Mutants isolated so far failed to identify clear effects on plant reproduction resulting from the suppression of the putrescine and spermidine conjugate pathways (Fellenberg et al., 2008, 2009; Matsuno et al., 2009; Grienenberger et al., 2009; Kaur et al., 2010). This might be due to partial functional redundancies or to the fact that the phenotypes were usually evaluated with RNAi or antisense lines or under laboratory conditions. Kaur et al. (2010) interpreted the absence of developmental phenotype as a demonstration that phenolamides are exclusively involved in defence and are accumulated in reproductive tissues to protect them against abiotic or biotic stresses, according to the Optimal Defence Theory. Further studies involving KO suppression mutants tested under field conditions are however required to confirm this hypothesis.

Phenolamides are quantitatively major plant metabolites that are tightly regulated during plant ontogeny and adaptative responses. Their role in plant biology is still surprisingly overlooked but raises an increasing interest. While biological functions of phenolamides are obviously linked to those of their polyamine and phenolic constituents, evolution of species-specific branched pathways is indicative of the acquisition of a diversity of metabolite-specific functions. It should also be considered that phenolamides, in particular via polyamines, offer an opportunity of cross-talk between nitrogen and phenolic metabolisms.

Acknowledgements

J.-E.B is grateful to the French Ministry of Research for the funding of his PhD. DW is grateful to the Human Frontier of Science Program to the Agence Nationale de la Recherche and to the European Commission SmartCell project for financial support.

References

Alcázar, R., Garcia-Martinez, J.L., Cuevas, J.C., Tiburcio, A.F., Altabella, T., 2005. Overexpression of ADC2 in Arabidopsis induces dwarfism and late flowering through GA deficiency. Plant J. 43, 425–436.

Alcázar, R., Marco, F., Cuevas, J.C., Patron, M., Ferrando, A., Carrasco, P., Tiburcio, A.F., Altabella, T., 2006a. Involvement of polyamines in plant response to abiotic stress. Biotechnol. Lett. 28, 1867–1876.

Alcázar, R., Cuevas, J.C., Patron, M., Altabella, T., Tiburcio, A.F., 2006b. Abscisic acid modulates polyamine metabolism under water stress in *Arabidopsis thaliana*. Physiol. Plant. 128, 448–455.

Alcázar, R., Planas, J., Saxena, T., Zarza, X., Bortolotti, C., Cuevas, J., Bitrián, M., Tiburcio, A.F., Altabella, T., 2010. Putrescine accumulation confers drought tolerance in transgenic Arabidopsis plants over-expressing the homologous Arginine decarboxylase 2 gene. Plant Physiol. Biochem. 48, 547–552.

Angelini, R., Tisi, A., Rea, G., Chen, M.M., Botta, M., Federico, R., Cona, A., 2008. Involvement of polyamine oxidase in wound healing. Plant Physiol. 146, 162–177.

Aribaud, M., Martin-Tanguy, J., 1994a. Polyamine metabolism, floral initiation and floral development in chrysanthemum (*Chrysanthemum morifolium* Ramat.). Plant Growth Regul. 15, 23–31.

Aribaud, M., Martin-Tanguy, J., 1994b. Polyamine metabolism in normal and sterile *Chrysanthemum morifolium*. Phytochemistry 37, 927–932.

Back, K., Jang, S.M., Lee, B.C., Schmidt, A., Strack, D., Kim, K.M., 2001. Cloning and characterization of a hydroxycinnamoyl-CoA:tyramine *N*-(hydroxycinnamoyl) transferase induced in response to UV-C and wounding from *Capsicum annuum*. Plant Cell Physiol. 42, 475–481.

Ben Hassine, A., Ghanem, M.E., Bouzid, S., Lutts, S., 2009. Abscisic acid has contrasting effects on salt excretion and polyamine concentrations of an inland and a coastal population of the Mediterranean xero-halophyte species Atriplex halimus. Ann. Bot. 104, 925–936.

Bernards, M.A., Fleming, W.D., Llewellyn, D.B., Priefer, R., Yang, X., Sabatino, A., Plourde, G.L., 1999. Biochemical characterization of the suberization-associated anionic peroxidase of potato. Plant Physiol. 121, 135–146.

Berta, G., Alatamura, M.M., Fusconi, A., Cerruti, F., Capitani, F., Bagni, N., 1997. The plant cell wall is altered by inhibition of polyamine biosynthesis. New Phytol. 137, 569–577.

Berwanger, A., Eyrisch, S., Schuster, I., Helms, V., Bernhardt, R., 2009. Polyamines: naturally occurring small molecule modulators of electrostatic protein–protein interactions. J. Inorg. Biochem. 204, 118–125.

Bienz, S., Bisegger, P., Guggisberg, A., Hesse, M., 2005. Polyamine alkaloids. Nat. Prod. Rep. 22, 647–658.

Bird, C.R., Smith, T.A., 1981. The biosynthesis of coumaroylagmatine in barley seedlings. Phytochemistry 20, 2345–2346.

Bird, C.R., Smith, T.A., 1983. Agmatine coumaroyltransferase from barley seedlings. Phytochemistry 22, 2401–2403.

Bohlmann, J., Lins, T., Martin, W., Eilert, U., 1996. Anthranilate synthase from *Ruta graveolens*. Duplicated AS alpha genes encode tryptophan-sensitive and tryptophan-insensitive isoenzymes specific to amino acid and alkaloid biosynthesis. Plant Physiol. 111, 507–514.

Bokern, M., Witte, L., Wray, V., Nimtz, M., Meurer-Grimes, B., 1995. Trisubstituted hydroxycinnamic acid spermidines from *Quercus dentata* pollen. Phytochemistry 39, 1371–1375.

Bonneau, L., Carré, M., Martin-Tanguy, J., 1994a. Polyamines and related enzymes in rice seeds differing in germination potential. Plant Growth Regul. 15, 75–82.

Bonneau, L., Carré, M., Dreumont, C., Martin-Tanguy, J., 1994b. Polyamine metabolism during seedling development in rice. Plant Growth Regul. 15, 83–92.

Borg-Olivier, O., Monties, B., 1993. Lignin, suberin, phenolic acids and tyramine in the suberized, wound-induced potato periderm. Phytochemisty 32, 601–606.

Bors, W., Langebartels, C., Michel, C., Sandermann, H.Jr., 1989. Polyamines as radical scavengers and protectants against ozone damage. Phytochemistry 28, 1589–1595.

Bouchereau, A., Aziz, A., Larher, F., Martin-Tanguy, J., 1999. Polyamines and environmental challenges: recent development. Plant Sci. 140, 103–125.

Buanafina, M.M., 2009. Feruloylation in grasses: current and future perspectives. Mol. Plant. 2, 861–872.

Burhenne, K., Kristensen, B.K., Rasmussen, S.K., 2003. A new class of *N*-hydroxycinnamoyltransferases: purification cloning, and expression of a barley agmatine coumaroyltransferase (EC 2.3.1.64). J. Biol. Chem. 278, 13919–13927.

Burtin, D., Martin-Tanguy, J., Tepfer, D., 1991. Alpha-dl-difluoromethylornithine, a specific, irreversible inhibitor of putrescine biosynthesis, induces a phenotype in tobacco similar to that ascribed to the root-inducing, left-hand transferred DNA of *Agrobacterium rhizogenes*. Plant Physiol. 95, 461–468.

Cabanne, F., Dalebroux, M.A., Martin-Tanguy, J., Martin, C., 1981. Hydroxycinnamic acid amides and ripening to flower of *Nicotiana tabaccum* var. Xanthi n.c.. Physiol. Plant. 53, 399–404.

Calvin, A., Hostettmann, K., Dyatmyko, W., Potterat, O., 1998. Antioxidant and lipophilic constituents of *Tinospora crispa*. Planta Med. 64, 393–396.

Cervelli, M., Di Caro, O., Di Penta, A., Angelini, R., Federico, R., Vitale, A., Mariottini, P., 2004. A novel C-terminal sequence from barley polyamine oxidase is a vacuolar sorting signal. Plant J. 40, 410–418.

Chattopadhayay, M.K., Tiwari, B.S., Chattopadhayay, G., Bose, A., Sengupta, D.N., Ghosh, B., 2002. Protective role of exogenous polyamines on salinity-stressed rice (*Oryza sativa*) plants. Physiol. Plant. 116, 192–199.

Chen, H., Jones, A.D., Howe, G.A., 2006. Constitutive activation of the jasmonate signaling pathway enhances the production of secondary metabolites in tomato. FEBS Lett. 580, 2540–2546.

Clarke, D.D., 1982. The accumulation of cinnamic acid amides in the cell walls of potato tissues as an early response to fungal attack. In: Wood, R.K.S. (Ed.), Active Defense Mechanisms in Plants, vol. 37. NATO Advanced Study Institute Series A: Life Sciences, Plenum Press, NY, pp 321–322.

Clay, N.K., Nelson, T., 2005. Arabidopsis thickvein mutation affects vein thickness and organ vascularization, and resides in a provascular cell-specific spermine synthase involved in vein definition and in polar auxin transport. Plant Physiol. 138, 767–777.

Collins, F.W., 1989. Oat phenolics: avenanthramides, novel substituted *N*-cinnamoylanthranilate alkaloids from oat groats and hulls. J. Agr. Food Chem. 37, 60–66.

Cona, A., Rea, G., Angelini, R., Federico, R., Tavladoraki, P., 2006. Functions of amine oxidases in plant development and defence. Trends Plant Sci. 11, 80–88.

Cowley, T., Walters, D.R., 2002. Polyamine metabolism in barley reacting hypersensitively to the powdery mildew fungus *Blumeria graminis* f. sp. hordei. Plant Cell Environ. 25, 461–468.

Creuss, J.A., Eucuentra, A., Gavalda, E.G., Barcelo, J., 1991. Binding of polyamines to different macromolecules in plants. In: Galston, A.W., Tiburcio, A.F.N. (Eds.), Polyamines as Modulators of Plant Development. Ediciones Peninsular, Madrid, pp. 30–34.

Cuevas, J.C., López-Cobollo, R., Alcázar, R., Zarza, X., Koncz, C., Altabella, T., Salinas, J., Tiburcio, A.F., Ferrando, A., 2008. Putrescine is involved in Arabidopsis freezing tolerance and cold acclimation by regulating abscisic acid levels in response to low temperature. Plant Physiol. 148, 1094–1105.

Daoudi, E.H., Bonnet-Masimbert, M., 1998. Polyamines conjuguées et différenciation florale chez le sapin de Douglas (*Pseudotsuga menziesii* (Mirb.) Franco). Can. J. Bot. 76, 782–790.

D'Auria, J.C., 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. 9, 331–340.

Del Duca, S., Bregoli, A.M., Bergamini, C., Serafini-Fracassini, D., 1997. Compatible and self-incompatible pollination in *Pyrus communis* displays different polyamine levels and transglutaminase activity. Amino Acids 38, 659–667.

Delétang, J., 1974. Présence de caféoylputrescine, de caféoylspermidine et de dicaféoyl spermidine chez *Nicotiana tabacum*. Ann. Tobacco 11, 124–130.

Della Mea, M., De Filippis, F., Genovesi, V., Serafini Fracassini, D., Del Duca, S., 2007a. The acropetal wave of developmental cell death of tobacco corolla is preceded by activation of transglutaminase in different cell compartments. Plant Physiol. 144, 1211–1222.

Della Mea, M., Serafini-Fracassini, D., Del Duca, S., 2007b. Programmed cell death: similarities and differences in animals and plants. A flower paradigm. Amino Acids 33, 395–404.

Do, C.T., Pollet, B., Thévenin, J., Silbout, R., Denoue, D., Barriere, Y., Lapierre, C., Jouanin L., 2007. Both caffeoyl coenzyme A 3-O-methyltransferase and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *A. thaliana*. Planta 226, 1117–1129.

Edreva, A., Yordanov, I., Kardjieva, R., Gesheva, E., 1998. Heat shock responses of bean plants: involvement of free radicals, antioxidants and free radical/active oxygen scavenging systems. Biol. Plant. 41, 185–191.

Edreva, A.M., Velikova, V.B., Tsonev, T.D., 2007. Phenylamides in plants. Russian J. Plant Physiol. 54, 287–301.

Efdi, M., Ohguchi, K., Akao, Y., Nozawa, Y., Koketsu, M., Ishihara, H., 2007. *N*-trans-feruloyltyramine as a melanin biosynthesis inhibitor. Biol. Pharm. Bull. 30, 1972–1974.

Facchini, P.J., Yu, M., Penzes-Yost, C., 1999. Decreased cell wall digestibility in canola transformed with chimeric tyrosine decarboxylase genes from opium poppy. Plant Physiol. 120, 653–664.

Facchini, P.J., Huber-Allanach, K.L., Tari, L.W., 2000. Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. Phytochemistry 54, 121–138.

Facchini, P.J., Hagel, J., Zulak, K.G., 2002. Hydroxycinnamic amide metabolism: physiology and biochemistry. Can. J. Bot. 80, 577–589.

Farmer, M.J., Czernic, P., Michael, A., Negrel, J., 1999. Identification and characterization of cDNA clones encoding hydroxycinnamoyl-CoA:tyramine *N*-hydroxycinnamoyltransferase from tobacco. FEBS J. 263, 686–694.

Fattorusso, E., Lanzotti, V., Taglialatela-Scafati, O., 1999. Antifungal *N*-feruloyl amides from roots of two *Allium* species. Plant Biosyst. 133, 199–203.

Fellenberg, C., Milkowski, C., Hause, B., Lange, P.R., Böttcher, C., Schmidt, J., Vogt, T., 2008. Tapetum-specific location of a cation-dependent *O*-methyltransferase in *Arabidopsis thaliana*. Plant J. 56, 132–145.

Fellenberg, C., Böttcher, C., Vogt, T., 2009. Phenylpropanoid polyamine conjugate biosynthesis in *Arabidopsis thaliana* flower buds. Phytochemistry 70, 1392–1400.

Fixon-Owoo, S., Levasseur, F., Williams, K., Sabado, T.N., Lowe, M., Klose, M., Joffre Mercier, A., Fields, P., Atkinson, J., 2003. Preparation and biological assessment of hydroxycinnamic acid amides of polyamines. Phytochemistry 63, 315–334.

Flores, H.E., 1991. Changes in polyamine metabolism in response to abiotic stress. In: Slocum, R.D., Flores, H.E. (Eds.), Biochemistry and Physiology of Polyamines in Plants. CRC Press. Boca Raton. pp. 213–226.

Fraga, M.F., Berdasco, M., Diego, L.B., Rodríguez, R., Cañal, M.J., 2004. Changes in polyamine concentration associated with aging in *Pinus radiata* and *Prunus persica*. Tree Physiol. 24, 1221–1226.

Fujiwara, T., Maisonneuve, S., Isshiki, M., Mizutani, M., Chen, L., Wong, H.L., Kawasaki, T., Shimamoto, K., 2010. Sekiguchi lesion gene encodes a cytochrome P450 monooxygenase that catalyzes conversion of tryptamine to serotonin in rice. J. Biol. Chem. 285, 11308–11313.

Funayama, S., Yoshida, K., Konno, C., Hikino, H., 1980. Structure of kukoamine A, a hypotensive principle of *Lycium chinense* root barks. Tetrahedron Lett. 21, 1355–1356.

Funayama, S., Zhang, G.R., Nozoe, S., 1995. Kukoamine B, a spermine alkaloid from *Lycium chinense*. Phytochemistry 38, 1529–1531.

Gális, I., Simek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H., Matsuoka, K., 2006. A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. Plant J. 46, 573–592.

Ge, C., Cui, X., Wang, Y., Hu, Y., Fu, Z., Zhang, D., Cheng, Z., Li, J., 2006. *BUD2*, encoding an *S*-adenosylmethionine decarboxylase, is required for Arabidopsis growth and development. Cell Res. 16, 446–456.

Graça, J., 2009. Hydroxycinnamates in suberin formation. Phytochem. Rev. 9, 85–91.

Grandmaison, J., Olah, G.M., Van Calsteren, M.R., Furlan, V., 1993. Characterisation and localisation of plant phenolics likely involved in the pathogen resistance expressed by endomycorrhizal roots. Mycorrhiza 3, 155–164.

Grienenberger, E., Besseau, S., Geoffroy, P., Debayle, D., Heintz, D., Lapierre, C., Pollet, B., Heitz, T., Legrand, M., 2009. A BAHD acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. Plant J. 58, 246–259.

Groppa, M.D., Benavides, M.P., 2008. Polyamines and abiotic stress: recent advances. Amino Acids 34, 35–45.

Groppa, M.D., Tomaro, M.L., Benavides, M.P., 2001. Polyamines as protectors against cadmium or copper-induced oxidative damage in sunflower leaf discs. Plant Sci. 161, 481–488.

Groppa, M.D., Rosales, E.P., Iannone, M.F., Benavides, M.P., 2008. Nitric oxide, polyamines and Cd-induced phytotoxicity in wheat roots. Phytochemistry 69, 2609–2615.

Guillet, G., De Luca, V., 2005. Wound-inducible biosynthesis of phytoalexin hydroxycinnamic acid amides of tyramine in tryptophan and tyrosine decarboxylase transgenic tobacco lines. Plant Physiol. 137, 692–699.

Guo, D.P., Sun, Y.Z., Chen, Z.J., 2003. Involvement of polyamines in cytoplasmic male sterility of stem mustard (*Brassica juncea* var. tsatsai). Plant Growth Regul. 41, 33–40.

Hagel, J.M., Facchini, P.J., 2005. Elevated tyrosine decarboxylase and tyramine hydroxycinnamoyltransferase levels increase wound-induced tyramine-derived hydroxycinnamic acid amide accumulation in transgenic tobacco leaves. Planta 221, 904–914.

Hamana, K., Matsuzaki, S., 1985. Distinct difference in the polyamine compositions of bryophyta and pteridophyta. J. Biochem. 97, 1595–1601.

Han, S.H., Lee, H.H., Lee, I.S., Moon, Y.H., Woo, E.R., 2002. A new phenolic amide from *Lycium chinense* Miller. Arch. Pharm. Res. 225, 433–437.

Han, E.H., Kim, J.Y., Kim, H.G., Choi, J.H., Im, J.H., Woo, E.R., Jeong, H.G., 2010. Dihydro-*N*-caffeoyltyramine down-regulates cyclooxygenase-2 expression by inhibiting the activities of C/EBP and AP-1 transcription factors. Food Chem. Toxicol. 48, 579–586.

Hanfrey, C., Sommer, S., Mayer, M.J., Burtin, D., Michael, A.J., 2001. Arabidopsis polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. Plant J. 27, 551–560.

Hanfrey, C., Elliott, K.A., Franceschetti, M., Mayer, M.J., Illingworth, C., Michael, A.J., 2005. A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. J. Biol. Chem. 280, 39229–39237.

Hanzawa, Y., Takahashi, T., Komeda, Y., 1997. *ACL5*: an Arabidopsis gene required for internodal elongation after flowering. Plant J. 12, 863–874.

Hanzawa, Y., Takahashi, T., Michael, A.J., Burtin, D., Long, D., Pineiro, M., Coupland, G., Komeda, Y., 2000. *ACAULIS5*, an Arabidopsis gene required for stem elongation, encodes a spermine synthase. EMBO J. 19, 4248–4256.

Havelange, A., Lejeune, P., Bernier, G., Kaur-Sawhney, R., Galston, A.W., 1996. Putrescine export from leaves in relation to floral transition in *Sinapis alba*. Physiol. Plant. 96, 59–65.

Hedberg, C., Hesse, M., Werner, C., 1996. Spermine and spermidine hydroxycinnamoyl transferases in *Aphelandra tetragona*. Plant Sci. 113, 149–156.

Hoffmann, L., Maury, S., Martz, F., Geoffroy, P., Legrand, L., 2003. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. J. Biol. Chem. 278, 95–103.

Hohlfeld, H, Schürmann, W., Scheel, D., Strack, D., 1995. Partial purification and characterization of hydroxycinnamoyl-coenzymea:tyramine hydroxycinnamoyl transferase from cell suspension cultures of *Solanum tuberosum*. Plant Physiol. 107, 545–552.

Hohlfeld, H., Scheel, D., Strack, D., 1996. Purification of hydroxycinnamoyl-CoA:tyramine hydroxycinnamoyltransferase from cell-suspension cultures of *Solanum tuberosum* L. Cv.. Datura. Planta 199, 166–168.

Hu, W., Werner, C., Hesse, M., 1998. The photoisomerization behavior of the (4-hydroxycinnamoyl)-spermidines. Helv. Chim. Acta 81, 342–352.

Imai, A., Akiyama, T., Kato, T., Sato, S., Tabata, S., Yamamoto, K.T., Takahashi, T., 2004a. Spermine is not essential for survival of Arabidopsis. FEBS Lett. 556, 148–152.

Imai, A., Matsuyama, T., Hanzawa, Y., Akiyama, T., Tamaoki, M., Saji, H., Shirano, Y., Kato, T. Hayashi, H., Shibata, D., Tabata, S., Komeda, Y., Takahashi, T., 2004b. Spermidine synthase genes are essential for survival of Arabidopsis. Plant Physiol. 135, 1565–1573.

Ioannidis, N.E., Kotzabasis, K., 2007. Effects of polyamines on the functionality of photosynthetic membrane *in vivo* and *in vitro*. Biochim. Biophys. Acta 1767, 1372–1382.

Ishihara, A., Miyagawa, H., Matsukawa, T., Ueno, T., Mayama, S., Iwamura, H., 1998. Induction of hydroxyanthranilate hydroxycinnamoyl transferase activity by oligo-*N*-acetylchitooligosaccharides in oats. Phytochemistry 47, 969–974.

Ishihara, A., Kawata, N., Matsukawa, T., Iwamura, H., 2000. Induction of *N*-hydroxycinnamoyltyramine synthesis and tyramine *N*-hydroxycinnamoyltransferase (THT) activity by wounding in maize leaves. Biosci. Biotechnol. Biochem. 64, 1025–1031.

Jang, S.M., Ishihara, A., Back, K., 2004. Production of coumaroylserotonin and feruloylserotonin in transgenic rice expressing pepper hydroxycinnamoyl-coenzyme A:serotonin N-(hydroxycinnamoyl)transferase. Plant Physiol. 135, 346–355.

Janicka-Russak, M., Kabała, K., Młodzińska, E., Kłobus, G., 2010. The role of polyamines in the regulation of the plasma membrane and the tonoplast proton pumps under salt stress. J. Plant Physiol. 167, 261–269.

Kakehi, J., Kuwashiro, Y., Niitsu, M., Takahashi, T., 2008. Thermospermine is required for stem elongation in *Arabidopsis thaliana*. Plant Cell Physiol. 49, 1342–1349.

Kakkar, R.K., Rai, V.K., 1993. Plant polyamines in flowering and fruit ripening. Phytochemistry 33, 1281–1288.

Kakkar, R.K., Sawhney, V.K., 2002. Polyamine research in plants – a changing perspective. Physiol. Plant. 116, 281–292.

Kamada-Nobusada, T., Hayashi, M., Fukazawa, M., Sakakibara, H., Nishimura, M., 2008. A putative peroxisomal polyamine oxidase, AtPAO4, is involved in polyamine catabolism in *Arabidopsis thaliana*. Plant Cell Physiol. 49, 1272–1282.

Kang, S., Back, K., 2006. Enriched production of N-hydroxycinnamic acid amides and biogenic amines in pepper (*Capsicum annuum*) flowers. Sci. Hortic. 108, 337–341.

Kang, S., Kang, K., Chung, G.C., Choi, D., Ishihara, A., Lee, D.S., Back, K., 2006. Functional analysis of the amine substrate specificity domain of pepper tyramine and serotonin *N*-hydroxycinnamoyltransferases. Plant Physiol. 140, 704–715.

Kang, S., Kang, K., Lee, K., Back, K., 2007. Characterization of tryptamine 5-hydroxylase and serotonin synthesis in rice plants. Plant Cell Rep. 26, 2009–2015.

Kang, K., Park, M., Park, S., Kim, Y.S., Lee, S., Lee, S.G., Back, K., 2009. Production of plant-specific tyramine derivatives by dual expression of tyramine *N*-hydroxycinnamoyltransferase and 4-coumarate:coenzyme A ligase in *Escherichia coli*. Biotechnol. Lett. 31, 1469–1475.

Kasukabe, Y., He, L., Nada, K., Misawa, S., Ihara, I., Tachibana, S., 2004. Overexpression of spermidine synthase enhances tolerance to multiple environmental stresses and up-regulates the expression of various stress-regulated genes in transgenic *Arabidopsis thaliana*. Plant Cell Physiol. 45, 712–722.

Kasukabe, Y., He, L., Watanabe, Y., Otani, M., Shimada, T., Tachibana, S., 2006. Improvement of environmental stress tolerance of sweet potato by production of genes for spermidine synthase. Plant Biotechnol. 23, 75–83.

Kaur, H., Heinzel, N., Schottner, M., Baldwin, I.T., Galis, I., 2010. R2R3-NaMYB8 regulates the accumulation of phenylpropanoid-polyamine conjugates which are essential for local and systemic defense against insect herbivores in *Nicotiana attenuata*. Plant Physiol. 152, 1731–1747.

Keller, H., Hohlfeld, H., Wray, V., Hahlbrock, K., Scheel, D., Strack, D., 1996. Changes in the accumulation of soluble and cell wall-bound phenolics in elicitor-treated cell suspension cultures and fungus-infected leaves of *Solanium tuberosum*. Phytochemistry 42, 389–396.

King, R.R., Calhoun, L., 2005. Characterization of cross-linked hydroxycinnamic acid amides isolated from potato common scab lesions. Phytochemistry 66, 2468–2473.

Klapheck, S., 1983. Polyamines and cinnamoyl-putrescines in normal and sulfur-starved suspension cultures of *Nicotiana tabacum*. Z. Pflanzenphysiol. 112, 275–279.

Klose, M.K., Atkinson, J.K., Mercier, A.J., 2002. Effects of a hydroxy-cinnamoyl conjugate of spermidine on arthropod neuromuscular junctions. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. 187, 945–952.

Knott, J.M., Römer, P., Sumper, M., 2007. Putative spermine synthases from *Thalassiosira pseudonana* and *Arabidopsis thaliana* synthesize thermospermine rather than spermine. FEBS Lett. 581, 3081–3086.

Kristensen, B.K., Burhenne, K., Rasmussen, S.K., 2004. Peroxidases and the metabolism of hydroxycinnamic acid amines in Poaceae. Phytochem. Rev. 3, 127–140.

Kumar, A., Taylor, M.A., Mad Arif, S.A., Davies, H.G., 1996. Potato plants expressing antisense and sense *S*-adenosylmethionine decarboxylase (SAMDC) transgenes show altered levels of polyamines and ethylene: antisense plants display abnormal phenotypes. Plant J. 9, 147–158.

Kumar, A., Altabella, T., Taylor, M., Tiburcio, A.F., 1997. Recent advances in polyamine research. Trends Plant Sci. 2, 124–130.

Kusano, T., Yamaguchi, K., Berberich, T., Takahashi, Y., 2007. Advances in polyamine research in 2007. J. Plant. Res. 120, 345–350.

Kusano, T., Berberich, T., Tateda, C., Takahashi, Y., 2008. Polyamines: essential factors for growth and survival. Planta 228, 367–381.

Kuznetsov, V.V., Radyukina, N.L., Shevyakova, N.I., 2006. Polyamines and stress: biological role, metabolism, and regulation. Russ. J. Plant Physiol. 53, 583–604.

Langebartels, C., Kerner, K., Leonardi, S., Schraudner, M., Trost, M., Heller, W., JR, Sandermann.H., 1991. Biochemical plant responses to ozone. I, differential induction of polyamine and ethylene biosynthesis in tobacco. Plant Physiol. 95, 882–889.

Lee, D.G., Park, Y., Kim, M.R., Jung, H.J., Seu, Y.B., Hahm, K.S., Woo, E.R., 2004. Anti-fungal effects of phenolic amides isolated from the root bark of *Lycium chinense*. Biotechnol. Lett. 26, 1125–1130.

Lehmann, T., Pollmann, S., 2009. Gene expression and characterization of a stress-induced tyrosine decarboxylase from *Arabidopsis thaliana*. FEBS Lett. 583, 1895–1900.

Leubner-Metzger, G., Amrhein, N., 1992. Hydroxycinnamoylputrescines are not causally involved in the tuberization process in potato plants. Physiol. Plant. 86, 495–501.

Leubner-Metzger, G., Amrhein, N., 1993. The distribution of hydroxycinnamoylputrescines in different organs of *Solanum tuberosum* and other solanaceous species. Phytochemistry 32, 551–556.

Li, R.W., Li, H.S., 1997. Changes of polyamine content in fertility transformation of PGMS rice (in Chinese). Plant Physiol. Commun. 33, 101–104.

Lopez-Meyse, M., Nessler, C.L., 1997. Tryptophan decarboxylase is encoded by two autonomously regulated genes in *Camptotheca acuminata* which are differentially expressed during development and stress. Plant J. 11, 1167–1175.

Louis, V., Negrel, J., 1991. Tyramine hydroxycinnamoyl transferase in the roots of wheat and barley seedlings. Phytochemistry 30, 2519–2522.

Lu, L., Berkey, K.A., Casero, R.A., 1996. RGFGIGS is an amino acid sequence required for acetyl coenzyme A binding and activity of human spermidine/spermine N^1 acetyltransferase. J. Biol. Chem. 271, 18920–18924.

Luo, J., Fuell, C., Parr, A., Hill, L., Bailey, P., Elliott, K., Fairhurst, S.A., Martin, C., Michael, A.J., 2009. A novel polyamine acyltransferase responsible for the accumulation of spermidine conjugates in Arabidopsis seed. Plant Cell 21, 318–333.

Malmberg, R.L., McIndoo, J., 1983. Abnormal floral development of a tobacco mutant with elevated polyamine levels. Nature 305, 623–625.

Marina, M., Maiale, S.J., Rossi, F.R., Romero, M.F., Rivas, E.I., Gárriz, A., Ruiz, O.A., Pieckenstain, F.L., 2008. Apoplastic polyamine oxidation plays different roles in local responses of tobacco to infection by the necrotrophic fungus *Sclerotinia sclerotiorum* and the biotrophic bacterium *Pseudomonas viridiflava*. Plant Physiol. 147, 2164–2178.

Marini, F., Betti, L., Scaramagli, S., Biondi, S., Torrigiani, P., 2001. Polyamine metabolism is upregulated in response to tobacco mosaic virus in hypersensitive, but not in susceptible, tobacco. New Phytol. 149, 301–309.

Martin-Tanguy, J., 1985. The occurrence and possible function of hydroxycinnamoyl acid amines in plants. Plant Growth Regul. 3, 381–399.

Martin-Tanguy, J., 1997. Conjugated polyamines and reproductive development: biochemical, molecular and physiological approaches. Physiol. Plant. 100, 675–688.

Martin-Tanguy, J., Martin, C., Gallet, M., Vernoy, R., 1976. Sur de puissants inhibiteurs de multiplication du virus de la mosaïque du tabac. C.R. Acad. Sci. 282, 2231–2234.

Martin-Tanguy, J., Cabanne, F., Perdrizet, E., Martin, C., 1978. Distribution of hydroxycinnamic acid amides in flowering plants. Phytochemistry 17, 1927–1928.

Martin-Tanguy, J., Perdrizet, E., Prevost, J., Martin, C., 1982. Hydroxycinnamic acid amides in fertile and cytoplasmic male sterile lines of maize. Phytochemistry 21, 1939–1945.

Martin-Tanguy, J., Sun, L.Y., Burtin, D., Vernoy, R., Rossin, N., Tepfer, D., 1996. Attenuation of the phenotype caused by the root-inducing, left-hand, transferred DNA and Its rolA gene (correlations with changes in polyamine metabolism and DNA methylation). Plant Physiol. 111, 259–267.

Mata, R., Morales, I., Pérez, O., Rivero-Cruz, I., Acevedo, L., Enriquez-Mendoza, I., Bye, R., Franzblau, S., Timmermann, B., 2004. Antimycobacterial compounds from *Piper sanctum*. J. Nat. Prod. 67, 1961–1968.

Matsuda, F., Morino, K., Ano, R., Kuzawa, M., Wakasa, K., Miyagawa, H., 2005. Metabolic flux analysis of the phenylpropanoid pathway in elicitor-treated potato tuber tissue. Plant Cell Physiol. 46, 454–466.

Matsuda, F., Yonekura-Sakakibara, K., Niida, R., Kuromori, T., Shinozaki, K., Saito, K., 2009. MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. Plant J. 57, 555–577.

Matsuno, M., Compagnon, V., Schoch, G., Schmitt, M., Debayle, D., Bassard, J.E., Pollet, B., Hehn, A., Heintz, D., Ullmann, P., Lapierre, C., Bernier, F., Ehlting, J., Werck-Reichhart, D., 2009. Retroposition and positive selection led to evolution of a novel phenolic pathway for pollen development. Science 225, 1688–1692.

Mayama, S., Tani, T., Matsuura, Y., Ueno, T., Fukami, H., 1981. The production of phytoalexins by oat (*Avena sativa*) in response to crown rust, *Puccinia coronata* f. sp. avenae. Physiol. Plant Pathol. 19, 217–226.

Mayama, S., Matsuura, Y., Iida, H., Tani, T., 1982. The role of avenalumin in the resistance of oat crown rust, *Puccina coronata* f. sp. Avenae. Physiol. Plant Pathol. 20, 189–199.

McLusky, S.R., Bennett, M.H., Beale, M.H., Lewis, M.J., Gaskin, P., Mansfield, J.W., 1999. Cell wall alterations and localized accumulation of feruloyl-3′-methoxytyramine in onion epidermis at sites of attempted penetration by *Botrytis allii* are associated with actin polarisation, peroxidase activity and suppression of flavonoid biosynthesis. Plant J. 17, 5523–5534.

Messiaen, J., Van Cutsem, P., 1999. Polyamines and pectins. II. Modulation of pectic-signal transduction. Planta 208, 247–256.

Meurer, B., Wray, V., Grotjahn, L., Wiermann, R., Strack, D., 1986. Hydroxycinnamic acid spermidine amides from pollen of *Corylus avellana* L. Phytochemistry 25, 433–435.

Meurer, B., Wiermann, R., Strack, D., 1988. Phenylpropanoid patterns in Fagales pollen and their phylogenetic relevance. Phytochemistry 27, 823–828.

Meurer-Grimes, B., 1995. New evidence for the systematic significance of acylated spermidines and flavonoids in pollen of the higher Hamamelidae. Brittonia 47, 130–142.

Meurer-Grimes, B., Berlin, J., Strack, D., 1989. Hydroxycinnamoyl-CoA: putrescine hydroxycinnamoyltransferase in tobacco cell cultures with high and low levels of caffeoylputrescine. Plant Physiol. 89, 488–492.

Minguet, E.G., Vera-Sirera, F., Marina, A., Carbonell, J., Blázquez, M.A., 2008. Evolutionary diversification in polyamine biosynthesis. Mol. Biol. Evol. 25, 2119–2128.

Mitsuya, Y., Takahashi, Y., Berberich, T., Miyazaki, A., Matsumura, H., Takahashi, H., Terauchi, R., Kusano, T., 2009. Spermine signaling plays a significant role in the defence response of *Arabidopsis thaliana* to cucumber mosaic virus. J. Plant Physiol. 166, 626–643.

Miyagawa, H., Ishihara, A., Nishimoto, T., Ueno, T., Mayama, S., 1995. Induction of avenanthramides in oat leaves inoculated with crown rust fungus, *Puccinia coronata* f. sp. Avenae. Biosci. Biotech. Biochem. 59, 2305–2306.

Morant, M., Schoch, G.A., Ullmann, P., Ertunç, T., Little, D., Olsen, C.E., Petersen, M., Negrel, J., Werck-Reichhart, D., 2007. Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. Plant Mol. Biol. 63, 1–19.

Moreau, R.A., Nuñez, A., Singh, V., 2001. Diferuloylputrescine and *p*-coumaroyl-feruloylputrescine, abundant polyamine conjugates in lipid extracts of maize kernels. Lipids 36, 839–844.

Moschou, P.N., Paschalidis, K.A., Roubelakis-Angelakis, K.A., 2008a. Plant polyamine catabolism: the state of the art. Plant Signal Behav. 3, 1061–1066.

Moschou, P.N., Paschalidis, K.A., Delis, I.D., Andriopoulou, A.H., Lagiotis, G.D., Yakoumakis, D.I., Roubelakis-Angelakis, K.A., 2008b. Spermidine exodus and oxidation in the apoplast induced by abiotic stress is responsible for H_2O_2 signatures that direct tolerance responses in tobacco. Plant Cell 20, 1708–1724.

Muñiz, L., Minguet, E.G., Singh, S.K., Pesquet, E., Vera-Sirera, F., Moreau-Courtois, C.L., Carbonell, J., Blázquez, M.A., Tuominen, H., 2008. ACAULIS5 controls Arabidopsis xylem specification through the prevention of premature cell death. Development 135, 2573–2582.

Muroi, A., Ishihara, A., Tanaka, C., Ishizuka, A., Takabayashi, J., Miyoshi, H., Nishioka, T., 2009. Accumulation of hydroxycinnamic acid amides induced by pathogen infection and identification of agmatine coumaroyltransferase in *Arabidopsis thaliana*. Planta 230, 17–527.

Negrel, J., 1989. The biosynthesis of cinnamoylputrescines in callus tissue cultures of *Nicotiana tabacum*. Phytochemistry 28, 447–481.

Negrel, J., Javelle, F., 1997. Purification, characterization and partial amino acid sequencing of hydroxy cinnamoyl-CoA: tyramine *N*-(hydroxycinnamoyl) transferase from tobacco cell-suspension cultures. Eur. J. Biochem. 247, 1127–1135.

Negrel, J., Jeandet, P., 1987. Metabolism of tyramine and feruloyltyramine in TMV inoculated leaves of *Nicotiana tabacum*. Phytochemistry 26, 2185–2190.

Negrel, J., Martin, C., 1984. The biosynthesis of feruloyltyramine in *Nicotiana tabacum*. Phytochemistry 23, 2797–2801.

Negrel, J., Javelle, F., Paynot, M., 1991. Separation of putrescine and spermidine hydroxycinnamoyl transferases extracted from tobacco callus. Phytochemistry 30, 1089–1092.

Negrel, J., Paynot, M., Javelle, F., 1992. Purification and properties of putrescine hydroxycinnamoyl transferase from tobacco (*Nicotiana tabacum*) cell suspensions. Plant Physiol. 98, 1264–1269.

Negrel, J., Pollet, B., Lapierre, C., 1996. Ether-linked ferulic acid amides in natural and wound periderm of potato tuber. Phytochemistry 43, 1195–1199.

Newman, M.A., von Roepenack-Lahaye, E., Parr, A., Daniels, M.J., Dow, J.M., 2001. Induction of hydroxycinnamoyl-tyramine conjugates in pepper by *Xanthomonas campestris*, a plant defense response activated by hrp gene-dependent and hrp gene-independent mechanisms. Mol. Plant Microbe Interact. 14, 85–92.

Nezbedova, L., Hesse, M., Drandarov, K., Bigler, L., Werner, C., 2001. Phenol oxidative coupling in the biogenesis of the macrocyclic spermine alkaloids aphelandrine and orantine in *Aphelandra* sp.. Planta 213, 411–417.

Niemi, K., Sutela, S., Häggman, H., Scagel, C., Vuosku, J., Jokela, A., Sarjala, T., 2006. Changes in polyamine content and localization of *Pinus sylvestris ADC* and *Suillus variegatus ODC* mRNA transcripts during the formation of mycorrhizal interaction in an *in vitro* cultivation system. J Exp Bot 57, 2795–2804.

Niggeweg, R., Michael, A.J., Martin, C., 2004. Engineering plants with increased levels of the antioxidant chlorogenis acid. Nature Biotech. 22, 746–754.

Oh, S.I., Park, J., Yoon, S., Kim, Y., Park, S., Ryu, M., Nam, M.J., Ok, S.H., Kim, J.K., Shin, J.S., Kim, K.N., 2008. The Arabidopsis calcium sensor calcineurin B-like 3 inhibits the 5′-methylthioadenosine nucleosidase in a calcium-dependent manner. Plant Physiol. 148, 1883–1896.

Okazaki, Y., Isobe, T., Iwata, Y., Matsukawa, T., Matsuda, F., Miyagawa, H., Ishihara, A., Nishioka, T., Iwamura, H., 2004. Metabolism of avenanthramide phytoalexins in oats. Plant J. 39, 60–72.

Panicot, M., Minguet, E.G., Ferrando, A., Alcázar, R., Blázquez, M.A., Carbonell, J., Altabella, T., Koncz, C., Tiburcio, A.F., 2002. A polyamine metabolon involving aminopropyl transferase complexes in Arabidopsis. Plant Cell 14, 2539–2551.

Park, M.H., 2006. The post-translational synthesis of a polyamine-dependent amino acid hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). J. Biochem. (Tokyo) 139, 161–169.

Park, J., 2007. Caffedymine from cocoa has COX inhibitory activity suppressing the expression of a platelet activation marker, P-selection. J. Agric. Food. Chem. 55, 2171–2175.

Park, J.B., 2009. Isolation and characterization of *N*-feruloyltyramine as the P-selectin expression suppressor from garlic (*Allium sativum*). J. Agric. Food. Chem. 57, 8868–8872.

Park, J.B., Schoene, N., 2002. Synthesis and characterization of *N*-coumaroyltyramine as a potent phytochemical which arrests human transformed cells via inhibiting protein tyrosine kinases. Biochem. Biophys. Res. Commun. 292, 1104–1110.

Park, S., Kang, K., Kim, Y.S., Back, K., 2009a. Endosperm-specific expression of tyramine *N*-hydroxycinnamoyltransferase and tyrosine decarboxylase from a single self-processing polypeptide produces high levels of tyramine derivatives in rice seeds. Biotechnol. Lett. 31, 911–915.

Park, S., Kang, K., Lee, K., Choi, D., Kim, Y.S., Back, K., 2009b. Induction of serotonin biosynthesis is uncoupled from the coordinated induction of tryptophan biosynthesis in pepper fruits (*Capsicum annuum*) upon pathogen infection. Planta 230, 1197–1206.

Parr, A.J., Mellon, F.A., Colquhoun, I.J., Davies, H.V., 2005. Dihydrocaffeoyl polyamines (Kukoamine and Allies) in potato (*Solanum tuberosum*) tubers detected during metabolite profiling. J. Agric. Food. Chem. 53, 5461–5466.

Paschalidis, K.A., Moschou, P.N., Aziz, A., Toumi, I., Roubelakis-Angelakis, K.A., 2009. Polyamines in grapevine: an update, second ed., Grapevine Molecular Physiology & Biotechnology, Springer Science, 2007–2028.

Paschold, A., Halitschke, R., Baldwin, I.T., 2007. Co(i)-ordinating defenses: NaCOI1 mediates herbivore-induced resistance in *Nicotiana attenuata* and reveals the role of herbivore movement in avoiding defenses. Plant J. 51, 79–91.

Passardi, F., Penel, C., Dunand, C., 2004. Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci. 9, 534–540.

Paynot, M., Perennec, P., Martin, C., Martin-Tanguy, J., Vernoy, R., Javelle, F., 1983. Photopériodisme, tubérisation et phénolamides. CR Acad. Sci. Paris 297, 87–90.

Pegg, A.E., 2008. Spermidine/spermine-N^1-acetyltransferase: a key metabolic regulator. Am. J. Physiol. Endocrinol. Metab. 294, E995–E1010.

Peipp, H., Maier, W., Schmidt, J., Wray, V., Strack, D., 1997. Arbuscular mycorrhizal fungus-induced changes in the accumulation of secondary compounds in barley roots. Phytochemistry 581, 587.

Petersen, M., Abdullah, Y., Benner, J., Eberle, D., Gehlen, K., Hücherig, S., Janiak, V., Kim, K.H., Sander, M., Weitzel, C., Wolters, S., 2009. Evolution of rosmarinic acid biosynthesis. Phytochemistry 70, 1663–1679.

Ponasik, J.A., Strickland, C., Faerman, C., Savvides, S., Karplus, P.A., Ganem, B., 1995. Kukoamine A and other hydrophobic acylpolyamines: potent and sélective inhibitors of *Crithidia fasciculata* trypanothione reductase. Biochem. J. 311, 371–375.

Ponchet, M., Martin-Tanguy, J., Marais, A., Martin, C., 1980. Hydroxycinnamoyl acid amides and aromatic amines in the inflorescences of some araceae species. Phytochemistry 21, 2865–2869.

Rastogi, R., Sawhney, V.K., 1990. Polyamines and flower development in the male sterile stamenless-2 mutant of tomato (*Lycopersicon esculentum* mill.). I, level of plyamines and their biosynthesis in normal and mutant flowers. Plant Physiol. 93, 439–445.

Rogoza, L.N., Salakhutdinov, N.F., Tolstikov, G.A., 2005. Plant alkaloids of the polymethyleneamine series. Russ. Chem. Rev. 381, 396.

Schmidt, A., Scheel, D., Strack, D., 1998. Elicitor-stimulated biosynthesis of hydroxycinnamoyltyramines in cell suspension cultures of *Solanum tuberosum*. Planta 205, 51–55.

Schmidt, A., Grimm, R., Schmidt, J., Scheel, D., Strack, D., Rosahl, S., 1999. Cloning and expression of a potato cDNA encoding hydroxycinnamoyl-CoA:tyramine *N*-(hydroxycinnamoyl) transferase. J. Biol. Chem. 274, 4273–4283.

Sen, A., Bergvinson, D., Miller, S.S., Atkinson, J., Fulcher, R.G., Arnason, J.T., 1994. Distribution and microchemical detection of phenolic acids, flavonoids, and phenolic acid amides in maize kernels. J. Agr. Food Chem. 42, 1879–1883.

Serafini-Fracassini, D., Del Duca, S., 2008. Transglutaminases: widespread cross-linking enzymes in plants. Ann. Bot. 102, 145–152.

Serafini-Fracassini, D., Del Duca, S., Beninati, S., 1995. Plant transglutaminases. Phytochemistry 40, 355–365.

Sfakianaki, M., Sfichi, L., Kotzabasis, K., 2006. The involvement of LHCII-associated polyamines in the response of the photosynthetic apparatus to low temperature. J. Photochem. Photobiol. B: Biol. 84, 181–188.

Shevyakova, N.I., Rakitin, V.Y., Stetsenko, L.A., Aronova, E.E., Kuznetsov, V.V., 2006. Oxidative stress and fluctuations of free and conjugated polyamines in the halophyte *Mesembryanthemum crystallinum* L. under NaCl salinity. Plant Growth Regul. 50, 69–78.

Shinya, T., Gális, I., Narisawa, T., Sasaki, M., Fukuda, H., Matsuoka, H., Saito, M., Matsuoka, K., 2007. Comprehensive analysis of glucan elicitor-regulated gene expression in tobacco BY-2 cells reveals a novel MYB transcription factor involved in the regulation of phenylpropanoid metabolism. Plant Cell Physiol. 48, 1404–1413.

Smith, T.A., 1977. Phenethylamine and related compounds in plants. Phytochemistry 16, 9–18.

Son, S., Lewis, B.A., 2002. Free radical scavenging and antioxidative activity of caffeic acid amide and ester analogues: structure-activity relationship. J. Agric. Food. Chem. 50, 468–472.

Stoessl, A., Unwin, C.H., 1970. The antifungal factors in barley. V. Antifungal activity of the hordatines. Can. J. Bot. 48, 465–470.

Strack, D., Eilert, U., Wray, V., Wolff, J., Jaggy, H., 1990. Tricoumaroylspermidine in flowers of Rosaceae. Phytochemistry 29, 2893–2896.

Suzuki, T., Holden, I., Casida, J.E., 1981. Diphenyl ether herbicides remarkably elevate the content in *Spinacia oleracea* of (E)-3-(4-hydroxy-3-methoxyphenyl)-*N*-[2-(4-hydroxy-3-methoxyphenyl)ethyl]-2-propenamide. J. Agric. Food. Chem. 29, 992–995.

Takahashi, T., Kakehi, J.I., 2009. Polyamines: ubiquitous polycations with unique roles in growth and stress responses. Ann. Bot. 105, 1–6.

Takahashi, Y., Berberich, T., Miyazaki, A., Seo, S., Ohashi, Y., Kusano, T., 2003. Spermine signaling in tobacco: activation of mitogen-activated protein kinases by spermine is mediated through mitochondril dysfunction. Plant J. 36, 820–829.

Takahashi, Y., Uehara, Y., Berberich, T., Ito, A., Saitoh, H., Miyazaki, A., Terauchi, R., Kusano, T., 2004. A subset of hypersensitive response marker genes, including HSR203J, is the downstream target of a spermine signal transduction pathway in tobacco. Plant J. 40, 586–595.

Tanaka, E., Tanaka, C., Morii, N., Kuwaharaa, Y., Tsuda, M., 2003. Phenylpropanoid amides of serotonin accumulate in witches' broom diseased bamboo. Phytochemistry 64, 965–969.

Tarenghi, E., Martin-Tanguy, J., 1995. Polyamides, floral induction and floral development of strawberry (*Fragaria ananassa* duch.). Plant Growth Regul. 17, 157–165.

Tassoni, A., Fornalè, S., Bagni, N., 2003. Putative ornithine decarboxylase activity in *Arabidopsis thaliana*: inhibition and intracellular localisation. Plant Physiol. Biochem. 41, 871–875.

Tebayashi, S.I., Ishihara, A., Tsuda, M., Iwamura, H., 2000. Induction of dovamide by jasmonic acid in red clover. Phytochemistry 54, 387–392.

Tebayashi, S., Horibata, Y., Mikagi, E., Kashiwagi, T., Mekuria, D.B., Dekebo, A., Ishihara, A., Kim, C.S., 2007. Induction of resistance against the leafminer, *Liriomyza trifolii*, by jasmonic acid in sweet pepper. Biosci. Biotechnol. Biochem. 71, 1521–1526.

Tian, C.E., Liang, C.Y., Hwang, Y.W., Liu, H.X., 1998. Metabolism of polyamines during the development of panicles in cytoplasmicmale sterile line and its maintainer line in *Oryza sativa* L. (in Chinese with English abstract). Acta Phytophysiol. Sinica 24, 333–339.

Toumi, I., Moschou, P.N., Paschalidis, K.A., Bouamama, B., Ben Salem-Fnayou, A., Ghorbel, A.W., Mliki, A., Roubelakis-Angelakis, K.A., 2010. Abscisic acid signals reorientation of polyamine metabolism to orchestrate stress responses via the polyamine exodus pathway in grapevine. J. Plant Physiol. 167, 516–525.

Tun, N.N., Santa-Catarina, C., Begum, T., Silveira, V., Handro, W., Floh, E.I.S., Scherer, G.F.E., 2006. Polyamines induce rapid biosynthesis of nitric oxide (NO) in *Arabidopsis thaliana* seedlings. Plant Cell Physiol. 47, 346–354.

Uehara, Y., Takahashi, Y., Berberich, T., Miyazaki, A., Takahashi, H., Matsui, K., Ohme-Takagi, M., Saitoh, H., Terauchi, R., Kusano, T., 2005. Tobacco ZFT1, a transcriptional repressor with a Cys2/His2 type zinc finger motif that functions in spermine-signaling pathway. Plant Mol. Biol. 59, 435–448.

Urano, K., Yoshiba, Y., Nanjo, T., Igarashi, Y., Seki, M., Sekiguchi, F., Yamaguchi-Shinozaki, K., Shinozaki, K., 2003. Characterization of Arabidopsis genes involved in biosynthesis of polyamines in abiotic stress responses and developmental stages. Plant, Cell Environ. 26, 1917–1926.

Urano, K., Yoshiba, Y., Nanjo, T., Ito, T., Yamaguchi-Shinozaki, K., Shinozaki, K., 2004. Arabidopsis stress-inducible gene for AtADC2 is required for accumulation of putrescine in salt tolerance. Biochem. Biophys. Res. Commun. 313, 369–375.

Urano, K., Hobo, T., Shinozaki, K., 2005. Arabidopsis ADC genes involved in polyamine biosynthesis are essential for seed development. FEBS Lett. 579, 1557–1564.

Vallée, J.C., Vansuyt, G., Négrel, J., Perdrizet, E., Prévost, J., 1983. Mise en évidence damines liées à des structures cellulaires chez *Nicotiana tabacum* et *Lycopersicum esculentum*. Physiol. Plant. 57, 143–148.

Velikova, V.B., Edreva, A.M., Tsonev, T.D., Jones, H.G., 2007. Singlet oxygen quenching by phenylamides and their parent compounds. Z. Naturforsch. C. 62, 833–838.

Verma, S., Mishra, S.N., 2005. Putrescine alleviation of growth in salt stressed *Brassica juncea* by inducing antioxidative defense system. J. Plant Physiol. 162, 669–677.

Villegas, M., Brodelius, P.E., Kylin, A., 1990. Elicitor-induced hydroxycinnamoyl-CoA:tyramine hydroxycinnamoyltransferase in plant cell suspension cultures. Physiol. Plant. 78, 414–420.

Vogt, T., 2010. Phenylpropanoid biosynthesis. Mol. Plant 3, 2–20.

Von Roepenack-Lahaye, E., Newman, M.A., Schornack, S., Hammond-Kosack, K.E., Lahaye, T., Jones, J.D.G., Daniels, M.J., Dow, J.M., 2003. *p*-Coumaroylnoradrenaline, a novel plant metabolite implicated in tomato defense against pathogens. J. Biol. Chem. 278, 43373–43383.

Von Röpenack, E., Parr, A., Schulze-Lefert, P., 1998. Structural analyses and dynamics of soluble and cell wall-bound phenolics in a broad spectrum resistance to the powdery mildew fungus in barley. J. Biol. Chem. 273, 9013–9022.

Walters, D.R., 2003a. Polyamines and plant disease. Phytochemistry 64, 97–107.

Walters, D.R., 2003b. Resistance to plant pathogens: possible roles for free polyamines and catabolism. New Phytol. 159, 109–115.

Walters, D., Meurer-Grimes, B., Rovira, I., 2001. Antifungal activity of three spermidine conjugates. FEMS Microbiol. Lett. 201, 255–258.

Walters, D., Cowley, T., Mitchell, A., 2002. Methyl jasmonate alters polyamine metabolism and induces systemic protection against powdery mildew infection in barley seedlings. J. Exp. Bot. 53, 747–756.

Wang, Y., Yun, B.W., Kwon, E.J., Hong, J.K., Yoon, J., Loake, G.J., 2006. *S*-nitrosylation: an emerging redox-based post-translational modification in plants. J Exp Bot 57, 1777–1784.

Werner, C., Hu, W.Q., Lorenziriatsch, A., Hesse, M., 1995. Dicoumaroylspermidines and tri-coumaroylspermidines in anthers of different species of the genus *Aphelandra*. Phytochemistry 40, 461–465.

Wise, M.L., Doehlert, D.C., McMullen, M.S., 2008. Association of avenanthramide concentration in oat (*Avena sativa* L.) grain with crown rust incidence and genetic resistance. Cereal Chem. 85, 639–641.

Yamaguchi, K., Takahashi, Y., Berberich, T., Imai, A., Takahashi, T., Michael, A.J., Kusano, T., 2007. A protective role for the polyamine spermine against drought stress in Arabidopsis. Biochem. Biophys. Res. Commun. 352, 486–490.

Yamakawa, H., Kamada, H., Satoh, M., Ohashi, Y., 1998. Spermine is a salicylate-independent endogenous inducer for both tobacco acidic pathogenesis-related proteins and resistance against tobacco mosaic virus infection. Plant Physiol. 118, 1213–1222.

Yamasaki, H., Cohen, M.F., 2006. NO signal at the crossroads: polyamine-induced nitric oxide synthesis in plants. Trends Plant Sci. 11, 522–524.

Yamazaki, Y., Sudo, H., Yamazaki, M., Aimi, N., Saito, K., 2003. Camptothecin biosynthetic genes in hairy roots of *Ophiorrhiza pumila*: cloning, characterization and differential expression in tissues and by stress compounds. Plant Cell Physiol. 44, 395–403.

Yang, Q., Reinhard, K., Schiltz, E., Matern, U., 1997. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:anthranilate *N*-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, *Dianthus caryophyllus* L. Plant Mol. Biol. 35, 777–789.

Yang, Q., Trinh, H.X., Imai, S., Ishihara, A., Zhang, L., Nakayashiki, H., Tosa, Y., Mayama, S., 2004. Analysis of the involvement of hydroxyanthranilate hydroxycinnamoyltransferase and caffeoyl-CoA 3-*O*-methyltransferase in phytoalexin biosynthesis in oat. Mol. Plant Microbe Interact. 17, 81–89.

Yoda, H., Yamaguchi, Y., Sano, H., 2003. Induction of hypersensitive cell death by hydrogen peroxide produced through polyamine degradation in tobacco plants. Plant Physiol. 132, 1973–1981.

Yoda, H., Hiroi, Y., Sano, H., 2006. Polyamine oxidase is one of the key elements for oxidative burst to induce programmed cell death in tobacco cultured cells. Plant Physiol. 142, 193–206.

Yoda, H., Fujimura, K., Takahashi, H., Munemura, I., Uchimiya, H., Sano, H., 2009. Polyamines as a common source of hydrogen peroxide in host- and non-host hypersensitive response during pathogen infection. Plant Mol. Biol. 70, 103–112.

Youhnovski, N., Bigler, L., Werner, C., Hesse, M., 1998. On-line coupling of high-performance liquid chromatography to atmospheric pressure chemical ionization mass spectrometry (HPLC/APCI-MS and MS/MS). The pollen analysis of *Hippeastrum* x *hortorum* (Amaryllidaceae). Helv. Chim. Acta 81, 1654–1671.

Youhnovski, N., Werner, C., Hesse, M., 2001. N,N′,N″-triferuloylspermidine, a new UV absorbing polyamine derivative from pollen of *Hippeastrum x hortorum*. Z. Naturforsch. 56, 526–530.

Yu, M., Facchini, P.J., 1999. Purification, characterization, and immunolocalization of hydroxycinnamoyl-CoA:tyramine *N*-(hydroxycinnamoyl)transferase from opium poppy. Planta 209, 33–44.

Zamble, A., Sahpaz, S., Hennebelle, T., Carato, P., Bailleul, F., 2006. N^1,N^5,N^{10}-tris(4-hydroxycinnamoyl)spermidines from Microdesmis keayana roots. Chem. Biodivers. 3, 982–989.

Posters and publications

Publications:

Phenolamides: bridging polyamines to the phenolic metabolism. (2010) Jean-Etienne Bassard, Pascaline Ullmann, François Bernier and Danièle Werck-Reichhart. Phytochemistry. 71(16): 1808-1824.

Evolution of a novel phenolic pathway for pollen development. (2009) Michiyo Matsuno, Vincent Compagnon, Guillaume Schoch, Martine Schmitt, Delphine Debayle, Jean-Etienne Bassard, Brigitte Pollet, Alain Hehn, Dimitri Heintz, Pascaline Ullmann, Catherine Lapierre, François Bernier, Jurgen Ehlting, Danièle Werck-Reichhart . Science. 325(5948): 1688-1692.

Publication in writting:
Protein-protein and membrane-protein interactions in the phenylpropanoid pathway. Jean-Etienne Bassard, Jonas Borch, Hui Duan, Peter Roepstorff, Steven Sligar, Danièle Werck-Reichhart.

Communications by poster:

Investigations on PROTEIN-MEMBRANE and PROTEIN-PROTEIN associations in the formation of METABOLONS in plant secondary metabolism using LSCM and BIA-anchored NANOdiscs™ Jean-Etienne BASSARD, Frédéric D. DUVAL, Jonas BORCH, Hui DUAN, Peter ROEPSTORFF, Steve SLIGAR, Danièle WERCK-REICHHART. FEBS Advanced course: cytochromes P450 Systems, KRANJSKA-GORA, Slovenia, 23-26 september 2008.

Investigations on PROTEIN-MEMBRANE and PROTEIN-PROTEIN associations in the formation of METABOLONS in phenylpropanoids metabolism using confocal microscopy and BIA-anchored lipid Nanodiscs™. Jean-Etienne BASSARD, Frédéric D. DUVAL, Jonas BORCH, Hui DUAN, Peter ROEPSTORFF, Steve SLIGAR, Danièle WERCK-REICHHART. 8ième colloque national de la SFBV, STRASBOURG, France, 8-10 july 2009.

Investigations on PROTEIN-MEMBRANE and PROTEIN-PROTEIN associations in phenylpropanoids metabolism. Jean-Etienne BASSARD, Jonas BORCH, Hui DUAN, Frédéric D. DUVAL, Peter ROEPSTORFF, Steve SLIGAR, Danièle WERCK-REICHHART. 8° PlantGEM meeting, LISBON, Portugal, 7-10 october 2009.

Les métabolons sont des complexes supramoléculaires constitués d'éléments structuraux et d'enzymes consécutives d'une voie de biosynthèse. De telles structures offrent de nombreux avantages pour l'organisation et l'efficacité des voies métaboliques et du métabolisme dans son ensemble. La formation de métabolons dans la voie de biosynthèse des phénylpropanoïdes est envisagée depuis de longues années. L'association/dissociation de ces derniers assurerait la canalisation des flux métaboliques entre les différentes branches constituant cette voie de biosynthèse. Deux cytochromes P450 membranaires, CYP73A5 et CYP98A3, ainsi que deux enzymes solubles, la *p*-coumaroyl :CoA ligase 1 (4CL-1) et l'hydroxycinnamoyl-CoA :shikimate hydroxycinnamoyl transférase (HCT), sont des éléments clé dans la formation du métabolon spécifique de la voie de biosynthèse de la lignine. Ce métabolon a été étudié *in vitro* par reconstitution sur nanodisques lipidiques, ainsi que par imagerie confocale sur des cellules épidermiques de *Nicotiana benthamiana* exprimant transitoirement des protéines de fusion fluorescentes. Ce travail a révélé des propriétés non décrites des enzymes cibles, comme le mouvement rapide des P450 ancrés au réticulum endoplasmique (RE), ou l'interaction de 4CL-1 et HCT avec les membranes artificielles ou cellulaires. *In vivo*, une relocalisation vers le RE des enzymes solubles a été observée en présence des P450s. Des interactions protéine-protéine ont été mises en évidence entre protéines membranaires et solubles, et furent accrues lors de la co-expression de l'ensemble des partenaires du métabolon. Ce travail a également permis de démontrer une homo- et une hétéro-oligomérisation des CYP73A5 et CYP98A3. CYP98A3 semble jouer un rôle déterminant pour la formation du métabolon. Ce travail souligne le caractère extrêmement dynamique du métabolon.

En parallèle, une analyse fonctionnelle des familles CYP98 et CYP73 a été réalisée chez *N. tabacum*. Une caractérisation fonctionnelle précise des divers membres de ces familles n'a pas pu être réalisée en raison de leur faible expression en système recombinant. En revanche, les patrons d'expression des différents membres de ces familles chez le tabac sain ou stressé indiquent une spécialisation fonctionnelle des différents paralogues.

Mots clef: cytochromes P450, CYP73A, CYP98A, phénylpropanoïdes, métabolon, interactions protéine-protéine, *Arabidopsis thaliana*, *Nicotiana tabacum*.

Metabolons are supramolecular complexes of sequential metabolic enzymes and cellular structural elements. This organization of metabolic pathways at the molecular level is expected to have several advantages on the metabolism efficiency. The existence of metabolons in the phenylpropanoid pathway was proposed many years ago. Metabolons association and dissociation was proposed to coordinate fluxes in the different branches of this complex pathway. Two membrane-anchored cytochromes P450, CYP73A5 and CYP98A3, and two soluble enzymes, the *p*-coumaroyl:CoA ligase 1 (4CL-1) and the hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), are expected to play essential roles in the lignin branch metabolon. The formation of this lignin metabolon has been investigated by *in vitro* reconstruction on lipid nanodiscs, and by confocal microscopy on *N. benthamiana* epidermal cells after transient expression of fluorescent fusion proteins. This work revealed unexpected features of the target enzymes, including fast movement of the P450 enzymes with the plant endoplasmic reticulum (ER) and membrane binding of 4CL and HCT, independent from the presence of the P450 proteins. It also indicated membrane relocalization of soluble enzymes *in vivo* in the presence of their partner proteins and demonstrated direct protein/protein interactions that were enhanced when CYP73A5, CYP98A3 and their two soluble partners were co-expressed in the same cells. This work also revealed P450 homo- and hetero-oligomerization and suggests that CYP98A3 plays an essential role in the formation of the lignin metabolon. These data underscore a very dynamic model for plant metabolon.

In parallel, functional investigations on the members of the CYP98 and CYP73 families in *Nicotiana tabacum* have been carried out. Enzyme functions could not be precisely described due to their low expression in the recombinant system. The different expression patterns of the paralogues in healthy or stressed tobacco plants were, however, clearly indicative of their subfunctionalization.

Key words: cytochromes P450, CYP73A, CYP98A, phenylpropanoids, metabolon, protein-protein interactions, *Arabidopsis thaliana*, *Nicotiana tabacum*.