Universidade de Trás-os-Montes e Alto Douro

Screening of heterologous cytochrome P450 monooxygenases for key catalytic activities involved in the biosynthesis of the anticancer alkaloids from *Catharanthus roseus*

Dissertação de Mestrado em

Genética Molecular Comparativa e Tecnológica

Diogo Francisco Maurício Coelho

Orientadora: Professora Doutora Mariana Sottomayor

Co-orientadora: Professora Doutora Raquel Chaves

Vila Real, 2015

Universidade de Trás-os-Montes e Alto Douro

Screening of heterologous cytochrome P450 monooxygenases for key catalytic activities involved in the biosynthesis of the anticancer alkaloids from *Catharanthus roseus*

Dissertação de Mestrado em

Genética Molecular Comparativa e Tecnológica

Diogo Francisco Maurício Coelho

Orientador: Professora Doutora Mariana Sottomayor

Coorientador: Professora Doutora Raquel Chaves

Composição do Júri:

_____

_____

_____

Vila Real, 2015

Declaração

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

“A mente que se abre a uma nova ideia jamais

voltará ao seu tamanho original.”

Albert Einstein

Aos meus pais e irmã,

A toda a minha família e amigos,

A todos aqueles que sempre

acreditaram em mim,

Agradecimentos

Agora que chegou ao fim mais uma grande etapa do meu percurso académico não posso deixar de agradecer a todas as pessoas que me acompanharam durantes estes dois anos de mestrado, estando sempre disponíveis para me ajudar, tendo sido essenciais no meu sucesso ao longo destes dois anos.

À Universidade de Trás-os-Montes e Alto Douro, na pessoa do magnífico Reitor PROFESSOR DOUTOR ANTÓNIO FONTAÍNHAS FERNANDES, instituição que me acolheu durante a licenciatura e mestrado, e que me forneceu todos os meios que me possibilitaram a conclusão dos mesmos.

À direção de curso, na pessoa da PROFESSORA DOUTORA PAULA FILOMENA MARTINS LOPES, por todos os esforços reunidos para manter a qualidade do mestrado e pela enorme disponibilidade em resolver os problemas que foram surgindo ao longo destes dois anos.

Ao BNP, por me ter acolhido para elaborar a minha dissertação de mestrado, por ter confiado no meu trabalho e por me ter proporcionado uma enorme experiência a nível de trabalho de equipa em investigação.

À PROFESSORA DOUTORA MARIANA SOTTOMAYOR, minha orientadora, agradeço por toda a ajuda, e atenção que me disponibilizou ao longo deste projeto. Agradeço todos os conhecimentos e ideais que me passou durante este ano, tornando-me uma pessoa mais completa, tanto a nível académico como a nível pessoal. O seu gosto pela investigação e o entusiasmo que demonstra pelo trabalho são sem dúvida uma grande motivação para mim.

À PROFESSORA DOUTORA RAQUEL CHAVES, minha co-orientadora, agradeço por todo o apoio que me deu durante ao longo da minha dissertação e por ter estado sempre presente. Seria ingrato agradecer apenas pelo apoio que me deu neste ano à pessoa que foi uma verdadeira inspiração para mim durante a licenciatura e o mestrado. Tenho a agradecer todas as palavras de incentivo e todos os conselhos, contribuíram sem dúvida para o meu crescimento enquanto aluno e pessoa.

À SARA, a minha "scientific adviser", agradeço todo o apoio que me deu ao longo deste ano, agradeço todos os conhecimentos que me transmitiu, agradeço as muitas horas passadas no laboratório a ajudar-me e agradeço toda a paciência que teve comigo quando eu interrompia o seu trabalho para ir ver espectros esquisitos ou alcaloides que não estavam a dissolver. A sua ajuda foi imprescindível na realização desta dissertação.

Ao PESSOAL DO LABORAÓRIO, agradeço por toda a disponibilidade que demonstraram, interrompendo muitas vezes o seu trabalho para me ajudar ou dar alguma explicação. Agradeço pelo ótimo ambiente de trabalho que me proporcionaram durante este estágio e pelos conhecimentos que me transmitiram. Não podia deixar de agradecer a ótima companhia aos almoços e lanches, bons momentos de descontração que davam aquela força para voltar ao trabalho.

À MALTA DE S.VERISSIMO, minhas colegas de casa, minhas amigas do coração, agradeço por tudo! MIKI, isto tinha sido muito mais difícil sem a tua boa "vibe". MAGDA, agradeço-te todo o apoio, conversas, desabafos e grandes momentos passados cá em casa. CATHY, a minha vizinha de quarto, agradeço toda a ajuda e motivação que me deste. A tua tranquilidade mascarada de stresse foi inspiradora. MAF, meu irmão ou meu filho, obrigado pelas saídas, pelas maluquices, pelos batuques, pelos abraços, pelo apoio e por seres uma grande pessoa. SORAIA, a minha jocki, primeiro tenho a agradecer por me teres mantido alimentado e por fazeres-me marmitas mesmo de madrugada. Tenho a agradecer a companhia nas viagens, nos almoços e nos lanches. Tenho a agradecer a paciência que tens para me aturar mas mais importante, tenho a agradecer todo o apoio que me deste, sem ele tudo teria sido muito mais difícil. 3,2,1 ESTAMOS JUNTOS!

Ao PESSOAL DE ALMEIRIM, Vasco, Francisco, Ruben, David, Magalhães, Neves, Edu, Zé, Marta, Vanessa, Rodrigo e Marques, é difícil encontrar palavras para agradecer a quem nos é tão importante. Mais importante do apoio que me deram foi o facto de terem acreditado sempre em mim. Cada jantarada, cada café e cada "futebolada" são ótimos momentos de descontração e foram grandes motivações para me agarrar ao trabalho com unhas e dentes.

Aos EPIC, tenho a agradecer todos os momentos passados juntos, férias no Algarve, jantares cá em casa e idas aos kebabs. A vossa amizade para mim é mesmo uma fonte de inspiração.

Aos MEUS PAIS e IRMÃ, os principais responsáveis pelo meu sucesso. Tenho que agradecer por serem todos os dias a minha maior inspiração. Agradeço por todo o carinho, compreensão e apoio que me dão a cada segundo e por toda a confiança depositada em mim. Agradeço por terem investido na minha educação e formação e por estarem lá sempre que preciso. Sem vocês nada disto teria sido possível. Aos MEUS AVÓS, agradeço por tudo o que fizeram e fazem por mim, por cuidarem sempre de mim, por me ouvirem e por confiarem em mim. À restante FAIMÍLIA, agradeço por todo o apoio, carinho e confiança que me transmitem.

O Meu Muito Obrigado!

Resumo

Catharanthus roseus (L.) G. Don é comumente denominada vinca de Madagáscar e é uma planta medicinal que tem sido ao longo dos anos objeto de um estudo intensivo. As propriedades medicinais desta planta começaram empiricamente a ser exploradas na antiguidade, mas foi a partir dos anos cinquenta que esta planta recebeu uma maior atenção por parte de diversos grupos de investigação, resultado do isolamento de dois alcaloides anticancerígenos a partir das suas folhas, vinblastina e vincristina. Para além destes alcaloides, até à data, sabe-se que *C. roseus* é uma "fábrica de compostos químicos", com a capacidade de produzir cerca de 130 alcaloides indolicos terpenoides (AIT), incluindo outros com distintas propriedades farmacológicas, tais como atividade hipotensora.

A vinblastina e vincristina são acumulados nas folhas de *C. roseus* em baixas quantidades, tornando a sua produção industrial um processo dispendioso. Portanto, a via de AIT tem sido estudada intensivamente ao longo dos anos, e as novas tecnologias "omicas" têm sido preponderantes nos recentes avanços. No entanto esta via é muito complexa, apresentando uma grande compartimentação a nível intra e intercelular, uma grande quantidade de compostos químicos e etapas bastante reguladas, o que continua a desafiar os investigadores. Consequentemente, os esforços para aumentar a produção destes alcaloides não têm sido bem sucedidos até à data.

Algumas das reações não caracterizadas na biossíntese dos AIT são potencialmente mediadas por citocromos P450 (CIPs). O CIP BM3, um CIP de *Bacillus megaterium* é uma enzima com grande atividade que sofreu uma extensiva diversificação genética para aceitar uma ampla gama de substratos, uma característica comum com os CIPs humanos envolvidos no metabolismo de drogas. Portanto, o objetivo deste estudo foi verificar se os CIPs BM3 e humanos selecionados têm capacidade de executar os passos não caracterizados/limitantes na via dos AIT, levando a formação dos alcaloides anticancerígenos de *C. roseus*.

Neste trabalho, a expressão de 10 variantes recombinantes do CIP BM3 em *E. coli* e 8 CIPs humanos recombinantes em levedura foi tentada. Uma quantidade significativa de proteína funcional foi conseguida para 6 variantes do CIP BM3 e a sua atividade de conversão de diversos AIT foi avaliada. AIT testados incluem tabersonina, α-3',4'-anidrovinblastina e vinblastina, substituída pela vindolina, com o objetivo de investigar se os BM3 têm potencial para serem utilizados na engenharia da via dos AIT. Apesar das condições de ensaio exigirem mais otimizações, e dos produtos necessitarem de caracterização, a variante 2571

correspondendo a BM3 9-10A F87V TS – 8C7 demonstrou resultados promissores. Outras variantes demonstraram sinais de afinidade para com α-3',4'-anidrovinblastina. Os CIPs humanos revelaram uma expressão mais difícil e esse processo requer otimizações. Finalmente, foi possível concluir que as variantes do CIP BM3 são expressas em células de planta e apresentam localização citoplasmática em células de folha de *C. roseus*. Esta informação é relevante para o planeamento de futuras estratégias de engenharia de *C. roseus* utilizando o CIP BM3. Globalmente, este estudo forneceu informações importantes e experiência para a futura implementação de estratégias baseadas nos CIPs para o melhoramento ou diversificação da via dos AIT de *C. roseus* através de engenharia metabólica.

Palavras-chave: *Catharantus roseus*; Tabersonina; Vindolina; α-3',4'-anidrovinblastina; Citocromos P450 humanos; Variantes do Citocromo P450 BM3.

Abstract

Catharantus roseus (L.) G. Don is commonly known as Madagascar periwinkle and is a medicinal plant that has been the subject of intense study. The medicinal properties of this plant were empirically exploited since ancient times but it was from the nineteen-fifties that this plant started receiving a growing attention from several research groups, as a result of the isolation of two anticancer alkaloids from its leaves, vinblastine and vincristine. In addition to these alkaloids, to date it is known that *C. roseus* is a chemical factory with the ability to produce more than 130 terpenoid indole alkaloids (TIAs), including others with further pharmacological properties, such as hypotensive activity.

The anticancer vincristine and vinblastine are accumulated by *C. roseus* in extremely low levels, making their industrial production an expensive process. Therefore, the TIA pathway has been intensively studied over the years, and the new "omic" technologies have been instrumental in fast recent advances. However, this is a very complex pathway, displaying a high intra and intercellular compartmentalization, a large range of chemical compounds and a tight regulation that continue to challenge researchers. Therefore, efforts to increase production of these alkaloids have not been successful so far.

Some of the key uncharacterized bottleneck reactions of TIA biosynthesis are potentially mediated by cytochromes P450 (CYPs). CYP BM3 from *Bacillus megaterium* is a highly active enzyme that has undergone extensive genetic diversification to accept a wide range of substrates, a feature common with human CYPs involved in drug metabolism. Therefore, the goal of this study was to screen the potential of selected CYP BM3 and human CYPs to function as heterologous surrogates for uncharacterized and/or bottleneck steps of the TIA pathway leading to the anticancer alkaloids of *C. roseus*.

Here, the overexpression of 10 recombinant CYP BM3 variants in *E. coli* and of 8 recombinant human CYPS in yeast was attempted. Significant amounts of correctly folded protein were successfully generated for 6 CYP BM3 variants and their conversion activity towards several target TIAs was evaluated. TIAs tested included tabersonine, α-3',4'-anhydrovinblastine and the vinblastine surrogate vindoline, in order to investigate if the BM3 variants have potential to be used for the engineering of the TIA pathway. Although the assay conditions need further optimization and the reaction products need characterization, the variant 2571 corresponding to BM3 9-10A F87V TS – 8C7 showed promising results. Several other variants also showed signs of potential affinity to α-3',4'-anhydrovinblastine. Human CYPs

revealed to be more difficult to be expressed with success and further optimization of the process is needed. Finally, it was possible to conclude that the CYP BM3 variants may be successfully expressed in plant cells and they assume a cytosolic localization in *C. roseus* leaf cells. This is relevant information for the design of future engineering strategies of *C. roseus* using CYP BM3s. Overall, this study generated important information and expertise for the future implementation of CYP based strategies for the improvement or diversification of the valuable TIA pathway from *C. roseus* through metabolic engineering.

Keywords: *Catharantus roseus*; Tabersonine; Vindoline; α-3',4'-anhydrovinblastine; Human Cytochrome P450; Cytochrome P450 BM3 variants.

Index

Figure Index

Table Index

Abbreviations

35S – Cauliflower Mosaic virus promoter

APS – Ammonium persulfate

AVLB – α-3',4'-anhydrovinblastine

BC – Before Christ

BP – Base pairs

cDNA – Complementary DNA

CFP – Cyan fluorescent protein

CO – Carbon monoxide

CPR – Cytochrome P450 reductase

CR – Cathenamine reductase

CrMPK3 – *C. roseus* mitogen-activated protein kinase

CrPrx1 – *Catharanthus roseus* class III peroxidase

CYP – Cytochrome P450

D4H – Deacetoxyvindoline-4-hydroxylase

dALA – δ-aminolevulinic acid

DAT – Deacetylvindoline-4-O-acetyltransferase

DMSO – Dimethylsulfoxide

DNA – Deoxyribonucleic acid

dNTPs – Deoxyribonucleotide triphosphates of purine bases (adenine, guanine), and pyrimidine bases (uracil, thymine, cytosine)

e.g. – *exempli gratia*

EDTA – Ethylenediaminetetraacetic acid

FAD – Flavin adenine dinucleotide

FMN – Flavin mononucleotide

i.e – *id est*

IPAP – Internal phloem associated parenchyma

IPTG – Isopropyl β-D-1-thiogalactopyranoside

kat, nkat – Katal, Nanokatal

kDa, Da – kiloDalton, Dalton

LB – Luria Bertrani

MOPS – 3-(N-morpholino)propanesulfonic acid

NADH – Nicotinamide adenine dinucleotide

NADPH – Nicotinamide adenine dinucleotide phosphate

nos ter – NOS terminator from the nopaline synthase gene from *A.tumefaciens*

OD – Optical density

ON – Over night

ORCA3 – Jasmonate-responsive transcriptional regulator

PAGE – Polyacrylamide gel electrophoresis

PCR – Polymerase chain reaction

pDNA – Plasmid DNA

PEG – Polyethylene glycol

rpm – Rotations per minute

RT – Room temperature

SDS – Sodium dodecyl sulfate

SGD – Strictosidine - β -D-glucosidase

SNP – Single nucleotide polymorphism

SNV – Single nucleotide variations

STR – Strictosidine synthase

T16H – Tabersonine 16-hydroxilase

T3O – Tabersonine 3-oxygenase

TAB – Tabersonine

TAE – Tris-acetate-EDTA

TB – Terrific broth

TDC – Tryptophan decarboxylase

TE – Tris-EDTA

TEK – Tris-EDTA-potassium

TEMED – Tetramethylethylenediamine

TES B – Tris-EDTA-sorbitol

TIA – Terpenoid indole alkaloid

Tm – Melting temperature

Tris – 2-Amino-2-hydroxymethyl-propane-1,3-diol

U – Enzyme unit

UV – Ultra-violet

v/v – volume/volume

VIND – Vindoline

VLB – Vinblastine

w/v – weight/volume

x g– G force

ε – Molar extinction coefficient

1.INTRODUCTION

1.1. *Catharanthus roseus*

Catharantus roseus (L.) G. Don, commonly known as Madagascar periwinkle **(Figure 1)**, is one of the most investigated medicinal plants (Verpoorte 1997).

This plant is popularly called Madagascar periwinkle because it is believed to be indigenous of central Madagascar. During the epoch of discoveries, this species was spread into many parts of the world by sailors, who carried the plants on board and chewed the leaves to fight the sensations of hunger and fatigue (Mishra and Kumar 2000, Heijden, Jacobs *et al.* 2004). It is believed to have been introduced into India by Portuguese missionaries in the middle of the eighteenth century (Mishra and Kumar 2000). *C. roseus* now has a pantropical distribution, being naturalized in continental Africa, the America's, Asia, Australia and Southern Europe and on some islands in the Pacific Ocean (Plaizier 1981, Van Bergen and Snoeijer 1996).

Concerning the nomenclature and taxonomy of *C. roseus*, there was initially a lot of ambiguity in the botanical literature regarding the correct nomenclature of the genus and, for several years, this species got several names, including *Vinca rosea*. In 1835, George Don assigned the name *Catharanthus* to the genus in his *General system of gardening and botany* (4: 95). The name *Catharanthus* is derived from the Greek words katharos (pure) and anthos (flower), referring to the neatness, beauty, and elegance of the flower (Shukla and Khanuja 2013).

The genus *Catharanthus* belongs to the family Apocynaceae and is closely related to the genera *Vinca* and *Amsonia*. This genus comprises eight species of small shrubs and herbs, seven of which are predominantly endemic of Madagascar and one is originated from India (Heijden, Jacobs *et al.* 2004). The chromosome number of all species of the genus *Catharanthus* is $2n = 16$, with a genome size of 1500 mega base pairs (Nejat, Valdiani *et al*. 2015).

C. roseus is a fast-growing, erect, evergreen, everblooming, perennial herb or small shrub, growing optimally in tropical and subtropical climates and being highly heat and drought tolerant (Shukla and Khanuja 2013).

One of the reasons why this plant is well known is its beauty and is therefore much appreciated as an ornamental plant. In the 1980's, the interest in developing *Catharanthus* cultivars as an ornamental has given rise to the breeding of plants with more colors, bigger flowers and smaller size. Today, over 100 cultivars are commercially available (Heijden, Jacobs *et al.* 2004).

But *C. roseus* is mostly known for their medicinal properties, which were already applied even in 50 BC (Virmani, Srivastava *et al.* 1978). Taylor and Farnsworth in 1975 were the first to consolidate all the information regarding the phytochemical, chemical, and pharmacological research on *C. roseus* alkaloids, which are the basis of the medicinal properties of the plant (Verpoorte 1997). Nowadays, the great economic importance of *C. roseus* arises from its highly valuable leaf anticancer alkaloids vincristine and vinblastine, its antihypertensive root alkaloid ajmalicine and its sedative alkaloid serpentine (Moreno, van der Heijden *et al.* 1995, Jaleel, Sankar *et al.* 2008). All these compounds are terpenoid indole alkaloids and their pharmaceutical interest associated to the low levels in the plant stimulated intense research generating a tremendous amount of literature, and making of *C. roseus* one of the most investigated medicinal plants (El-Sayed and Verpoorte 2007).

Figure 1. *Catharanthus roseus* (L.) G. Don, also known as Madagascar periwinkle.

C. roseus was initially investigated due to its use to treat diabetes. Two independent groups: one at the University of Western Ontario (Noble, Beer and Cutts) and the other at Eli Lilly and Company, Indianapolis (Svoboda's group) investigated its putative hypoglycemic activity and found no compounds responsible for such properties (Shukla and Khanuja 2013). However this investigation allowed Noble, Beer *et al.* (1958) to observe a depletion of white cells in rats treated with *C. roseus* extracts that led to the isolation of the first known *C. roseus* alkaloid initially named vincaleukoblastine, in view of its origin and its effect on leukoblasts (Noble, Beer *et al.* 1959). Later, the name was shortened to the more accessible vinblastine **(Figure 2A)** (Sottomayor and Barceló 2006). Almost at the same time, Svoboda, Neuss *et al.* (1959) in a general screening program for medicinal activities detected an antitumor activity in

C. roseus extracts, which enabled to isolate vincristine **(Figure 2 B)**, later shown to be a dimeric alkaloid closely related in chemical structure to vinblastine.

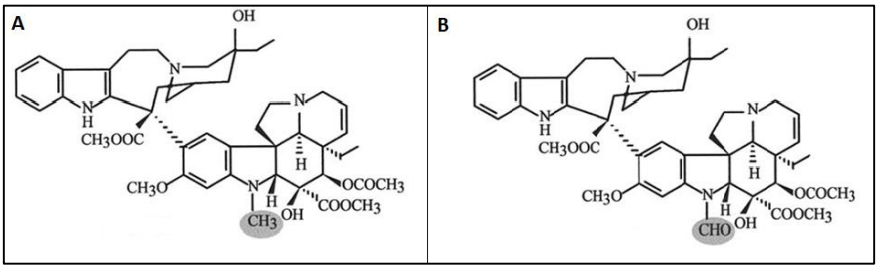

Figure 2. Chemical structure of the two anticancer alkaloids from *C. roseus*. **A)** Vinblastine. **B)** Vincristine. The shaded areas indicate the structural differences between the two compounds. Adapted from Sottomayor, Cardoso *et al.* (2004).

In modern medicine, vinblastine and vincristine are important agents used in cancer therapy. They may be applied alone or in combination with other anticancer agents, integrating therapeutic programs of different cancer types (Sottomayor and Barceló 2006). The therapeutic ability of these alkaloids lies in the fact that they interfere with the polymerization of microtubules, arresting tumor cells during mitosis (Almagro, Fernández-Pérez *et al.* 2015). *C. roseus* was shown to act as a chemical factory, producing more than 130 different terpenoid indole alkaloids, with several of them depicting further pharmacological properties that potentially could be applied in other areas of medicine (Almagro, Fernández-Pérez *et al.* 2015, Pan, Mustafa *et al.* 2015). Those pharmacological properties were reviewed by Almagro, Fernández-Pérez *et al.* (2015) and are summarized in **(Table 1).**

Table 1. Pharmacological properties of *C. roseus* alkaloids.

Alkaloid	Pharmacological Properties
Yohimbine	Antiviral, Antimicrobial, Important role on erectile dysfunction treatment
Catharoseumine	Antimicrobial
Vinblastine and Vincristine	Antimicrobial, Antitumor
Vindoline, Vindolidine, Vindolicine and Vindolinine	Antidiabetic
Ajmalicine and Serpentine	Anti-Neuro-Inflammatory, Hypotensive
α-3',4'-anhydrovinblastine	Antitumor
Vindolicine	Antioxidant

1.2. Terpenoid indole alkaloids

Alkaloids are one of the major groups of plant secondary metabolites and have been critical in the formulation of modern drugs. They comprise pharmacologically active, nitrogen-containing heterocyclic compounds and their general classification is based on their chemical structures. In the plant, alkaloids are thought to perform a defense role, operating as a defense against herbivorous pests, and some have been shown to be lethal against certain fungi and bacteria (Bennett and Wallsgrove 1994, Wink 1998).

Terpenoid indole alkaloids (TIAs) are characterized by the presence of an indole moiety derived from tryptophan and a monoterpenoid moiety derived from secologanin. They constitute one of the most well characterized classes of alkaloids and represent a major alkaloid group in the plant kingdom, including more than 3000 identified members to date (Shukla and Khanuja 2013, Pan, Mustafa *et al.* 2015).

TIAs are found mainly in plant species belonging to the families Apocynaceae, Loganiaceae, and Rubiaceae (Heijden, Jacobs *et al.* 2004).

Some TIAs have been applied in medicine due to their therapeutic properties like the antimalarial alkaloid quinine from *Cinchona officinalis*, the antineoplastic camptothecine from *Camptotheca acuminate* and the antihypertensive and tranquilizer reserpine from *Rauvolfia* species (Kutchan 1995).

1.2.1. Biosynthesis of terpenoid indole alkaloids

The biosynthesis of all TIAs in *C. roseus* involves the production of the central precursor strictosidine, which arises from the coupling between tryptamine and secologanin under the catalysis of strictosidine synthase (STR) (Rischer, Orešič *et al.* 2006). Tryptamine contributes with the indole moiety in the molecules of TIAs and is synthetized by decarboxylation of the amino acid tryptophan by tryptophan decarboxylase (TDC) (De Luca, Marineau *et al.* 1989). Secologanin is involved in the formation of the monoterpene seco-iridoid moiety and results from the conversion of two isopentenyl diphosphate units through several reactions that compose the iridoid pathway **(Figure 3)** (Oudin, Courtois *et al.* 2007).

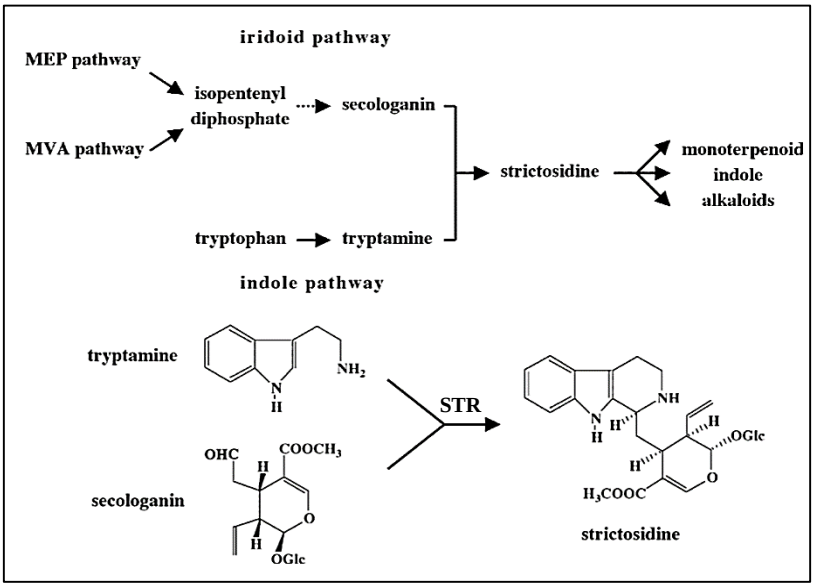

Figure 3. Synthesis of the central precursor strictosidine through the coupling reaction between tryptamine and secologanin performed by strictosidine synthase (STR). MEP – Methylerythritol 4-phosphate pathway. MVA – Mevalonate pathway. Adapted from (Oudin, Courtois *et al.* 2007).

Once synthesized, strictosidine experiences a range of reactions and rearrangements, ultimately resulting in the formation of hundreds of different TIAs (Brown, Clastre *et al.* 2015). These reactions begin with the removal of the glucose moiety of strictosidine by strictosidine-β -D-glucosidase (SGD) leading to the formation of an unstable aglycone **(Figure 4 ,arrow 1)** (Guirimand, Courdavault *et al.* 2010). Subsequently, the aglycone may be converted into cathenamine and reduced to ajmalicine by the enzyme cathenamine reductase (CR) **(Figure 4, arrows 3 and 4)**. The putative oxidation of ajmalicine by class III peroxidases then leads to the

formation of serpentine (Sottomayor, Cardoso *et al.* 2004). Another two branches of the pathway include the conversion of the aglycone into catharanthine and alternatively into tabersonine through several uncharacterized steps **(Figure 4, arrows 2 and 6)**. Tabersonine is transformed into vindoline by a sequence of seven steps comprising different characterized enzymatic reactions **(Figure 4, arrow 5)** (O'Connor and Maresh 2006, Kellner, Geu-Flores *et al.* 2015, Qu, Easson *et al.* 2015).

The formation of the anticancer alkaloids vinblastine and vincristine represents the last stage of this pathway. First, the monomeric alkaloids catharanthine and vindoline are coupled into α-3′,4′-anhydrovinblastine (AVLB) in a reaction that is thought to be mediated by a basic class III peroxidase, CrPrx1 **(Figure 4, arrow 7)** (Sottomayor, Lopez-Serrano *et al.* 1998) (Costa, Hilliou *et al.* 2008). AVLB is then converted into vinblastine and vincristine through bottleneck uncharacterized reactions **(Figure 4, arrow 8 and 9)** (Zhu, Wang *et al.* 2015).

So far, five reactions of the TIA pathway were shown to be catalyzed by cytochromes P450 (CYPs), which belong to an enzymatic family that will be presented in detail in chapter 1.3 of the introduction. CYP 76B6 was shown to catalyze the hydroxylation of geraniol into 10-hydroxigeraniol, in the beginning of the iridoid pathway **(Figure 3)** (Collu, Unver *et al.* 2001). (Salim, Yu *et al.* 2013) after recombinant expression and characterization of CYP 72A224 in yeast observed that this enzyme displays 7-deoxyloganic acid 7-hydroxylase activity, catalyzing the conversion of 7-deoxyloganic acid into loganic acid, also in the iridoid pathway **(Figure 3)**. CYP 72A1 has secologanin synthase activity and is responsible for the conversion of loganin into secologanin, closing the iridoid pathway **(Figure 3)** (Irmler, Schröder *et al.* 2000). *C. roseus* CYPs 71D12 and 71D351, are both tabersonine 16-hydroxilases (T16Hs) and perform the conversion of tabersonine into 16-hydroxytabersonine, an enzymatic step in the final pathway leading to vindoline **(Figure 4, arrow 5)** (Besseau, Kellner *et al.* 2013). Applying an enzyme assay and analyzing several characteristics of the reaction, St-Pierre and De Luca (1995) initially concluded that the conversion of tabersonine into 16-hidroxytabersonine could be performed by a CYP (T16H) . Schröder, Unterbusch *et al.* (1999) reported the cloning of the T16H cDNA from cell suspension cultures and the expression of the enzyme in *Escherichia coli* allowing their functional identification. Through analysis of the sequence, the researchers named it CYP71D12 in accordance with the rules of CYP nomenclature. Later, a second cytochrome P450 (CYP71D351) also displaying T16H activity was isolated from young leaves., showing the existence of two isoenzymes with different regulation for the hydroxylation of tabersonine (Besseau, Kellner *et al.* 2013). Finally, very

recently, Kellner, Geu-Flores *et al.* (2015) and Qu, Easson *et al.* (2015) characterized CYP71D1 and CYP71D1V2 as being a tabersonine 3-oxygenase (T3O) responsible for another step in the biosynthesis of vindoline from tabersonine.

It is important to note that the conversion of AVLB into vinblastine and then into vincristine may involve uncharacterized CYPs, and that these reactions are most likely the major bottleneck in the biosynthesis of these compounds. In fact, AVLB has been shown to be often present in high levels in *C. roseus*, with vinblastine and vincristine being almost undetectable (Sottomayor, Lopez-Serrano *et al.* 1998, Costa, Hilliou *et al.* 2008, Carqueijeiro, Noronha *et al.* 2013).

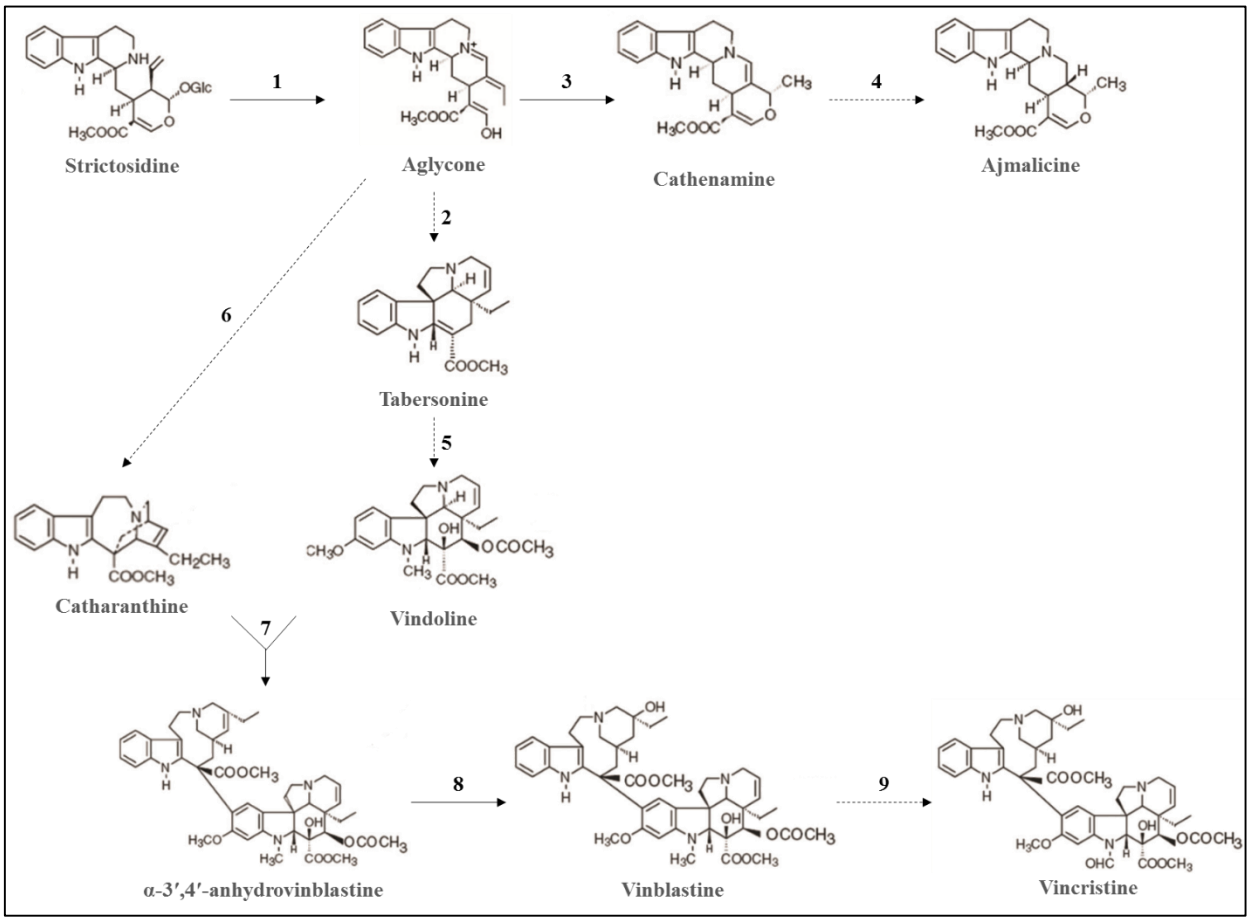

Figure 4. A simplified overview of the terpenoid indole alkaloids pathway, beginning on strictosidine. The reactions are indicated by the numbers. The dashed arrows represent multi step reactions and filled arrows represent single step reactions. Adapted from (Zhu, Wang *et al.* 2015).

1.2.2. Organization of the alkaloid pathway in *C. roseus*

The enzyme systems involved in TIA biosynthesis in *C. roseus* have an elaborated intra - and intercellular organization, adding further complexity, to the pathway. In fact, the *C. roseus* alkaloid pathway is one of the most complex and elaborated forms of compartmentalization described to date, representing a "beautiful confusing" network with a highly regulated flow of intermediate metabolites between the different intra- and intercellular plant structures (St-Pierre, Besseau *et al.* 2013, de Bernonville, Clastre *et al.* 2014).

The initial steps of the terpenoid pathway occur in plastids, where the synthesis of ther precursor geraniol takes place (Mahroug, Burlat *et al.* 2007). Geraniol is then transported across the plastid membranes and undergoes various reactions by enzymes localized in the cytosol and endoplasmic reticulum to yield secologanin (Saiman 2014). Tryptophan is derived from the shikimate pathway in the plastid and it moves out to the cytosol, where tryptamine is synthetized (De Luca and Cutler 1987). The formation of strictosidine is described as taking place in the vacuole, therefore both secologanin and tryptamine have to be transported into this organelle (Guirimand, Courdavault *et al.* 2010). After formed, strictosidine leaves the vacuole and the known reactions leading to the formation of tabersonine occur in the nucleus and the cytosol. Then, tabersonine is converted into vindoline by several reactions performed by enzymes localized in the ER membranes, the cytosol and in the membrane of the thylakoids (Courdavault, Papon *et al.* 2014). Regarding the synthesis of catharanthine, its location remains unknown (Saiman 2014). Finally, catharanthine and vindoline are transported to the vacuole where is located the enzyme CrPrx1, which putatively mediates their coupling to produce AVLB (Costa, Hilliou *et al.* 2008). It is thought that reactions which lead to formation of vinblastine and vincristine also occur in the vacuole (Saiman 2014).

Beyond the diversity of intracellular localization of the TIA reactions, it is noted that the TIA pathway genes are also expressed in different cell types in the leaves, thus suggesting intercellular translocation of intermediates in the TIA biosynthethic pathway (Saiman 2014).

The genes involved in the first steps of the pathway from geraniol until loganic acid are expressed in the internal phloem associated parenchyma (IPAP) cells of leaves (Miettinen, Dong *et al.* 2014). Further analysis of the location of gene expression, showed that the formation of tryptamine, secologanin and TIA intermediates like tabersonine occurs in epidermal cells (Courdavault, Papon *et al.* 2014). In addition, secretion and accumulation of catharanthine in the leaf wax surface of young leaves suggest that catharanthine biosynthesis in the leaf takes

place in the epidermis cells (Roepke, Salim *et al.* 2010). The last two steps of vindoline biosynthesis occur in specialized cells, laticifer and idioblast cells of aerial tissues (Verma, Mathur *et al.* 2012). The location of the synthesis of vinblastine and vincristine remains unknown. On the one hand, the synthesis of the first dimeric TIA, AVLB, involve the coupling between vindoline and catahranthine, which have been suggested to be physically separated in the plant, hampering researchers to find which location is "home" to the last reactions of this tricky pathway (Courdavault, Papon *et al.* 2014). On the other hand, it has been shown that the levels of AVLB in leaves may be high (Carqueijeiro, Noronha *et al.* 2013), suggesting co-localization of catharanthine and vindoline and a possible location of the last stage reactions in laticifers and idioblasts, known to accumulate alkaloids (Yoder and Mahlberg 1976).

Summarizing, TIA biosynthesis in *C. roseus* is a complex pathway involving at least 30 coordinately regulated enzymatic steps producing at least 35 known intermediates. The diverse steps of this pathway exhibit different subcellular localization and are spread over at least four different cell types (Courdavault, Papon *et al.* 2014).

1.3. Cytochromes P450

Cytochromes P450 (CYPs) represent one of the largest and oldest gene superfamilies coding for enzymes present in the genomes of all biological kingdoms (Degtyarenko and Archakov 1993, Hannemann, Bichet *et al.* 2007). These enzymes are heme *b* containing monooxygenases which were recognized and defined as a distinct class of hemeproteins around 50 years ago (Klingenberg 1958). The heme consistes in a prosthetic group formed by an iron ion coordinated by four nitrogen atoms of a porphyrin ring, and is linked to the apoprotein via a conserved cysteine (Urlacher and Girhard 2012).

The terminology P450 stems in the spectral properties of this b-type heme containing red pigments, which display a typical absorption band at 450 nm when reduced by carbon monoxide, resulting from the establishment of a bond between the heme and carbon monoxide (Omura and Sato 1964).

The current knowledge about the CYP enzymatic mechanism is the result of an intense research developed over the last years in CYP biochemistry. Much of this work was done using the model enzyme CYPcam from *Pseudomonas putida*, probably the best-characterized CYP (Wong 1998, Guengerich 2014).

CYPs catalyze regio- and stereospecic oxidations of non-activated hydrocarbons (C– H) under mild reaction conditions. These oxidations may be briefly described by catalysis of the reductive scission of molecular oxygen, thereby introducing one atom into the substrate, while the second oxygen atom is reduced to water, although the mechanism is much more complex (Urlacher and Girhard 2012). There are some CYPs fleeing the general reaction mechanism, such as CYP152 peroxygenases from *Sphingomona paucimobilis, Bacilus subtilis and Clostridium acetobutylicum*, which utilize an oxygen atom from peroxides instead of molecular oxygen for oxidation reactions (Van Bogaert, Groeneboer *et al.* 2011, Guengerich and Munro 2013).

To scroll through the CYP catalytic cycle and perform the oxidation reactions, it is necessary the transfer of electrons to the CYP heme center (Urlacher and Girhard 2012). Almost all CYPs use electrons derived from the pyridine cofactors NADH or NADPH and, therefore, they must be associated with redox partner proteins which are responsible for the transfer of electrons from the cofactors to the heme center. Based on the types of redox proteins participating in electron transfer, CYPs can be inserted into different classes, with the number of classes rising due to constant growing of gene numbers discovered in the genome projects (Hannemann, Bichet *et al.* 2007).

CYPs catalyze a wide variety of oxidation reactions like hydroxylations, N-, O- and S-dealkylations, sulphoxidations, epoxidations, deaminations, desulphurations, dehalogenations, and N-oxide reductions, with individual CYPs often catalyzing multiple reactions. This makes it impossible to designate the enzyme according to the reaction that it performs (Hannemann, Bichet *et al.* 2007). This problem was solved by the introduction of a systematic nomenclature based on CYP's structural homology (Nelson, Kamataki *et al.* 1993, Nelson, Koymans *et al.* 1996). The nomenclature system for CYPs was first devised in 1987 (Nebert, Adesnik *et al.* 1987) and has undergone some updates, but its basic principle was never broken (Nelson 2006). The nomenclature system refers that CYP genes are identified by the abbreviation CYP followed by a number indicating the family (proteins with more than 40% sequence identity), a letter nominating a subfamily (proteins with more than 55% sequence identity) and a number representing the individual gene within the subfamily, for example, CYP 106A2 (Nelson 2006). A large part of the work on CYP nomenclature was developed by David Nelson, who provides updates to the system and has categorized all the information by creating a repository for CYP nomenclature and sequence data (http://drnelson.uthsc.edu/CytochromeP450.html).

Despite the fact that the sequence conservation among CYPs of different families may be less than 20%, their structural fold are highly conserved and display a common mechanism of oxygen activation with few exceptions. Nonetheless, the most variable regions correspond to the flexible substrate recognition sites (Gotoh 1992, Munro, Girvan *et al.* 2013). This characteristic explains the remarkable variety of chemical reactions catalyzed and the wide number of substrates of CYPs, accepting fatty acids, terpenes, steroids, prostaglandins, polyaromatic and heteroaromatic compounds and a vast number of drugs. The ability of CYPs to catalyze different reactions of a vast number of substrates made these versatile enzymes quite appealing to biotechnology applications as discussed in chapter 1.3.3 (Urlacher and Girhard 2012).

The wide distribution of CYPs among various forms of life including plants and insects was already known in the early years of CYP research but technological limitations made the elucidation of the physiological functions of CYPs a difficult work, specially for nonmammalian CYPs. Fortunately, with advances in molecular biology and genomics, there have been major advances in this field since the end of the last century (Omura 2013).

An impressive functional breadth of CYP activities was observed in different organisms (Danielson 2002). Prokaryotic CYPs display fatty acid and steroid hydroxylating activities in *Bacillus* spp., participate in the biosynthesis of antibiotics and other natural products in *Streptomyces* spp. and display roles in bacterial virulence and novel secondary metabolite production in mycobacteria (McLean, Leys *et al.* 2015). Fungal CYPs participate in a wide variety of physiological reactions that contribute to the fitness and fecundity of fungi in various ecological niches, being introduced in certain metabolic pathways and leading to the production of various compounds. Fungal CYPs facilitate also the pathogenesis by detoxifying the chemical defenses of target host plants (Chen, Lee *et al.* 2014). Plant CYPs are implicated in different pathways: including the synthesis of fatty acids and sterols, of structural components (cell wall), of signaling networks components (gibberllins) and of defense secondary compounds (e.g. terpenoids, alkaloids). Insect CYPs are essentially involved in the metabolism of fatty acids and hormones and in the production of insect defense toxins and pheromones (Schuler 2015). Finally, mammalian CYPs are involved in the synthesis of steroids, metabolism of fatty acids and, play a central role as phase I enzymes in mammalian drug metabolism, being also implicated in the conversion of procarcinogens and promutagens to deleterious genotoxic compounds (Danielson 2002).

Searching in repositories of CYP data, it is verified that on the date of the last update (August 2013) there are 1056 CYP sequences identified in mammals, 3452 in insects, 7446 in plants, 5729 in fungi and 1254 in bacteria, with a total of 21039 CYP sequences identified, including other organisms like *Sulfolobus solfataricus* (http://drnelson.uthsc.edu/CytochromeP450.html). The presence of CYPs in archaea suggests that they are of ancient origin (Wright, Harris *et al.* 1996). Possibly they evolved in response to the toxicity risk posed by rising planetary dioxygen levels and initially played functions related to the lipids that formed the cell membranes of early microbes (Whitehouse, Bell *et al.* 2012).

1.3.1. The cytochrome P450 BM3

The bacterial cytochrome P450 BM3 (CYP102A1, CYP BM3) was discovered in the early seventies by Fulco (1974), who identified it as a soluble medium- to long – chain fatty acid hydroxylase that required only NADPH and oxygen to function (Fulco 1974, Matson, Hare *et al.* 1977, Whitehouse, Bell *et al.* 2012). The BM3 designation comes from the fact that this CYP was the third to be isolated from the soil bacterium *Bacillus megaterium* and the name became widely used (Narhi and Fulco 1986, Whitehouse, Bell *et al.* 2012). Besides having been identified in *B. megaterium* this enzyme has also been found in a number of *Bacillus* spp. as well as in other bacteria (Munro, Leys *et al.* 2002).

The hydroxylation performed by BM3 takes place exclusively at a sub-terminal position of a range of fatty acids, near the ω terminus, a specificity shared by most bacterial CYPs (Narhi and Fulco 1986, Whitehouse, Bell *et al.* 2012). The CYP BM3 activity is different according to the chain length of fatty acids, and it also accepts as substrate fatty amides, alcohols, hydroxylated fatty acids, ω-oxo fatty acids and unsaturated fatty acids (Whitehouse, Bell *et al.* 2012).

CYP BM3 is considered the most catalytically efficient of the known P450 oxidase enzymes with a catalytic rate constant of ~285 s^{-1} reported for the oxidation of arachidonic acid **(Figure 5)** (Noble, Miles *et al.* 1999, McLean, Leys *et al.* 2015).

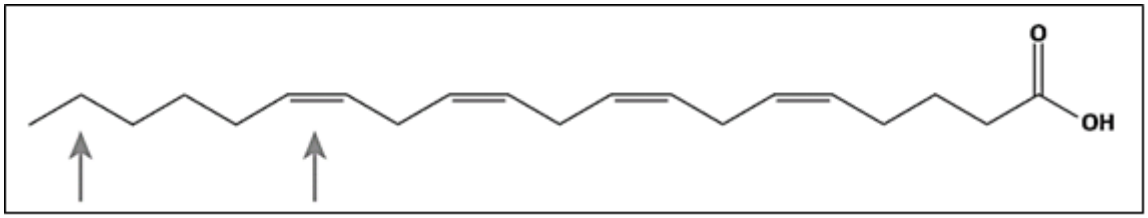

Figure 5. Arachidonic acid is a polyunsaturated fatty acid that is present in membrane phospholipids and is abundant in brain and liver. CYP BM3 hydroxylates AA at the ω-2 carbon and epoxidizes in the ω-5 and 6 positions (arrows). Adapted from (Munro, Girvan *et al.* 2013)

Structural characterization of CYP BM3 revealed an 119500 Da enzyme composed of two major domains, a 55 kDa P450 domain present in the N-terminal portion, containing the heme center responsible for binding to the substrate, was fused to a 66 kDa "reductase" domain present in the C-terminal portion showing reducing activity (Narhi and Fulco 1987). This connection is made through a short protein linker which is fundamental in the electron transfer process as evidenced by Munro, Lindsay *et al.* (1994), after comparing the flow of electrons between isolated domains and in the intact CYP protein. BM3 was the first CYP discovered to exist as a fusion with its CYP reductase (Ruettinger, Wen *et al.* 1989).

Analyzing the "reductase" domain, it was found that it includes two prosthetic flavin groups, FAD (Flavin Adenine Dinucleotide) and FMN (Flavin Mononucleotide), in an equimolar ratio, and may be termed a diflavin domain (Narhi and Fulco 1987, Munro, Leys *et al.* 2002). During the BM3 reaction, NADPH initially reduces the BM3 diflavin "reductase" domain with electrons transferred to the FAD cofactor, which are then transferred one by one from FAD to FMN and further on to the heme iron in the P450 domain (Ost, Clark *et al.* 2003).

The composition of the diflavin domain and the P450 system are similar to those found in mammalian hepatic CYP enzymes, with the exception that the last works by separate membrane-embedded CYP and reductase enzymes. Thus, the diflavin domain may be known as eukaryotic – like diflavin and the CYP BM3 may be used as model in the study of the mammalian CYP enzymes (Warman, Roitel *et al.* 2005).

As mentioned above, Hannemann, Bichet *et al.* (2007) proposed the classification of CYPs in different classes taking account the types of redox proteins participating in electron transfer. These authors classified CYP BM3 as Class VIII, which comprise CYP proteins that are fused to their eukaryotic – like diflavin reductase partner (CYP reductase) in a single polypeptide chain, and therefore are catalytically self-sufficient as monooxygenases. Although

CYP BM3 is the most studied member of this class, CYPs of this type have been discovered in various prokaryotes and lower eukaryotes (Warman, Roitel *et al.* 2005).

Boddupalli, Estabrook *et al.* (1990) have suggested that the higher levels of activity of CYP BM3 than other CYP fatty acid hydroxylases were due to the self-sufficient structure of CYP BM3, since the transfer of electrons is not dependent on the need to find a redox partner.

Despite being extensively studied in what concerns its molecular properties, CYP BM3's physiological functions remain obscure (Whitehouse, Bell *et al.* 2012, McLean, Leys *et al.* 2015). For many years, the natural substrates were assumed to be straight – chain fatty acids, and these have been copiously researched. However, while CYPcam expression can be induced in *Pseudomonas putida* by camphor, the natural substrate, CYP BM3 expression is not induced in *Bacillus megaterium* by saturated straight – chain fatty acids or monounsaturated acids. By contrast, polyunsaturated analogues such as linoleic and arachidonic acids do induce CYP BM3 expression, binding 1000-fold more strongly to the repressor protein, Bm3R1 (Whitehouse, Bell *et al.* 2012). English, Hughes *et al.* (1994) documented that unsaturated fatty acids are toxic to *B. megaterium* if applied exogenously and Fulco (1967) had already given evidence to suggest that they are synthesized only as a transient response to cold, and even then in small quantities. These evidence point to a possible role of the CYP BM3 in detoxification of xenobiotic lipids produced by plants (Palmer, Axen *et al.* 1998, McLean, Leys *et al.* 2015). Not only unsaturated fatty acids, but also branched – chain saturated fatty acids promote CYP BM3 expression. These fatty acids play a key role in the regulation of membrane fluidity and they are believed to be the enzyme's true substrate since they account for almost 90% of the fatty acid content of *B. megaterium* (Whitehouse, Bell *et al.* 2012, McLean, Leys *et al.* 2015).

As a highly active self-sufficient monooxygenase expressed successfully in *Escherichia coli*, CYP BM3 has been the target of enzymatic engineering with applied aims, a subject that will be further specified in chapter 1.3.3 (Jung, Lauchli *et al.* 2011).

1.3.2. Human Cytochromes P450

The history of CYPs really began with studies on the metabolism of drugs, carcinogens and steroids. The early research in these fields were supported by animal models, but the intent was always to understand how human systems processed these xenobiotics (Guengerich 2015).

Early work showed that human drug metabolism was inducible and different among individuals, characteristics that suggested a possible involvement of CYPs (Conney, Miller *et al.* 1956, Rubin and Lieber 1968, Vesell and Page 1968, Guengerich 2015).

With technological advancement it was possible to carry out the purification of human CYPs from liver microsomes and these studies brought evidence that multiple CYPs exist in humans (Wang, Beaune *et al.* 1983). Human studies demonstrated that the metabolism of an individual drug was genetically controlled, which led to the introduction of the hypothesis that the monogenic control of the oxidation of a drug makes a single CYP dominant in its metabolism, a premise that enabled the isolation of different human CYPs (Guengerich 2015).

With purified CYPs, their antibodies and other approaches, it became possible to define the roles of individual CYPs in the metabolism of individual drugs, carcinogens and steroids. The development of molecular biology led to the isolation of cDNA clones for many human CYPs, contributing for their heterologous expression in different systems like yeast and bacteria (Sakaki, Kenji *et al.* 1985, Li and Chiang 1991). This advance was central for expression in sufficient quantities to allow the crystallization of human CYPs, which began to be made in the early nighties by Johnson and his associates (Williams, Cosme *et al.* 2003). At present, the three-dimensional structures of at least 21 human CYPs have been determined (Guengerich 2015). The completion of the three-dimensional structures with *in silico* techniques, such as computer modeling have provided valuable clues about the characteristics of individual CYPs (Rowland, Blaney *et al.* 2006). Recombinant DNA technology was not only important in determining the three dimensional structures of CYPs but also allowed to understand the regulation of these genes and opened doors for research of single nucleotide variations (SNVs), which could sometimes be associated with altered drug or steroid metabolism (Guengerich 2015).

Through the Human Genome Project it was possible to identify the number of human CYPs as 57. Today it is also known that human CYP enzymes have two subcellular locations, in the endoplasmic reticulum (microsomal) and in mitochondria and are expressed in a wide range of tissues like liver, brain, lung, etc., with the liver presenting the greater diversity of expressed CYPs (Guengerich 2015). It was shown that CYPs play a major role in drug

metabolism. CYPs are involved in the metabolic clearance of the 200 most prescribed drugs, with 75% of drugs cleared mainly by CYPs. ~90% of those CYP mediated drug metabolism reactions can be accounted for by a set of five CYPs (1A2, 2C9, 2C19, 2D6 and 3A4) and the largest fraction (~46%) of the reactions is catalyzed by CYP 3A enzymes (Rendic and Guengerich 2014). It should be noted that all these CYPs are microsomal liver CYPs (Guengerich 2015).

Drug clearance is a process comprising two phases (phase I and phase II). The oxidations performed by CYPs are included in phase I, with phase II involving a conjugation step performed by transferases. These two phases lead to the modification of hydrophobic drugs allowing its excretion as water-soluble forms (Meunier, De Visser *et al.* 2004).

In the panoply of drugs metabolized by CYPs are included a number of anticancer drugs, as mentioned by Wahlang, Falkner *et al.* (2015) in their recent review. In the set of anticancer drugs, the *C. roseus* vinblastine and vincristine alkaloids mark their presence as being metabolized by CYP3A4 and CYP3A5 (Yao, Ding *et al.* 2000, Dennison, Kulanthaivel *et al.* 2006).

Defects in several CYPs have been linked to serious human diseases not only due to changes in drug metabolism but also due to the importance of CYPs in the metabolism of endogenous compounds like steroids (Guengerich 2015).

In addition to drugs and endogenous compounds, CYPs are also involved in the metabolism of chemical carcinogens. Metabolism of certain carcinogens such as nitrosamines, furans and vinyl chloride by some CYPs causes the formation of active metabolites which provoke alterations in DNA like its alkylation (Wahlang, Falkner *et al.* 2015).

Regarding the reaction molecular mechanisms of microsomal human CYPs, they follow the standard mechanism of CYPs mentioned above. They are included in class II according to the electron transfer mechanism and types of redox proteins. The monooxygenase system of class II found in the endoplasmic reticulum of eukaryotes contains two integral membrane proteins: the CYP and the NADPH CYP reductase (CPR), the latter containing the prosthetic groups FAD and FMN responsible for the transfer of electrons from NADPH to one of the many CYPs isozymes (Hannemann, Bichet *et al.* 2007).

1.3.3. Cytochromes P450 in biotechnology

No other enzyme group has a range of accepted substrates and catalyzed reaction types, as CYPs, including so many complex molecules, namely alkaloids, and "unusual" reactions. These findings make CYPs outstanding biocatalysts with broad applicability including a tremendous potential for biotechnological use (Bernhardt and Urlacher 2014). Another fact that supports the potential of CYPs in biotechnology is the hydroxylation of nonactivated C–H bonds. This reaction is one of the major challenges in chemistry since it involves expensive and complex chemical catalysts, require harsh reaction conditions, are often not very effective, and most existing chemical catalysts display lack in selectivity. In contrast, oxygenations catalyzed by CYPs in one step are often highly regioselective and stereoselective leading to high-value compounds that are difficult or even impossible to synthesize via traditional chemical routes. In addition, CYPs operate under mild reaction conditions utilizing molecular oxygen, which is abundant, environmentally friendly, and inexpensive. Finally, some P450s are reported to mediate multiple sequential modifications on a single substrate, which is particularly attractive when complex multistep biotechnological processes should be established (Girhard, Bakkes *et al.* 2015).

In industry, CYPs have applications in the development of new important compounds including several chemicals and pharmaceuticals. Moreover, CYPs have a great potential for the development of biosensors, as well as in bioremediation (Girhard, Bakkes *et al.* 2015).

The application of human CYPs in biotechnology started with the production of drug metabolites, since this is of outstanding interest in drug development. In order to characterize the metabolic pathways of a candidate drug in detail, metabolites of this compound are required as reference compounds, but also as candidates to investigate toxicity, biological activity, and drug/drug interactions. However, drug metabolites standards were not commercially available and classical chemical synthesis was difficult and stereochemically demanding. So, the application of human liver CYPs for the production of drug metabolites was an important challenge. Heterologous expression of CYPs together with CYP reductase was essential in this field, since it allowed the large scale expression of CYPs (Schroer, Kittelmann *et al.* 2010). As an example, there is the production of metabolites of the Novartis compound AAG561 by recombinant *E.coli* expressing human CYP3A4 described by Hanlon, Friedberg *et al.* (2007).

Human CYPs also have possible applications in the production of commercial products like indigo, widely used in the dye industry, in the development of biosensors and in

bioremediation, as these enzymes have the ability to metabolize chemical pollutants and herbicides and are therefore implicated in the development of transgenic plants with resistance to these latter compounds (Kumar 2010).

Despite numerous characteristics that make CYPs an optimal investment in biotechnology, there are certain limitations that have hindered the implementation of these enzymes in certain procedures. Low activity, especially of mammalian CYP enzymes, expensive cofactors (NADPH and NADH), substrate or product toxicity, low substrate solubility and stability are examples of limitations that have been target of optimizations in order to implement these enzymes in biotechnological processes (Girhard, Bakkes *et al.* 2015). Some of the optimization strategies tried were the fusion between human CYPs and the reductase systems from other organisms, in order to increase the catalytic rates of human CYPs, and the insertion of mutations in the CYP active site making it able to perform reactions that are not exercised by the wild type enzyme. These changes are known as enzyme engineering and several enzymes have been improved over the years (Girhard, Bakkes *et al.* 2015).

In biotechnological applications, the CYP BM3 has the advantage over the human CYPs of being a self-sufficient monooxygenase with a high catalytic activity (Kumar 2010, Urlacher and Girhard 2012). Obtaining the crystal structure of CYP BM3 and its combination with computational methods were key approaches to the knowledge of several enzymatic aspects of the CYP BM3 structure, including substrate interaction with the CYP active site and key residues, opening doors to the engineering and introduction of CYP BM3 on biotechnology (Whitehouse, Bell *et al.* 2012).

Engineering of the CYP BM3 enzyme has been made in order to change the substrate spectrum of the enzyme. This was done following two general strategies: directed evolution and rational protein design **(Figure 6)** (Girhard, Bakkes *et al.* 2015). Directed evolution does not require prior knowledge of the protein structure and the DNA sequence. This strategy uses random (point) mutagenesis of a whole gene or domain, insertion and deletions (non – recombinant techniques) or recombination of homologous genes (recombinant techniques). Thus, the enzyme will acquire changes as compared to wild type that may be advantageous in adapting to new functions or environments (Arnold 2009). However, directed evolution requires the screening of large variant libraries with thousands of clones and, in most cases the hit rates for new activities are rather low (Girhard, Bakkes *et al.* 2015). Rational protein design require prior knowledge of the protein structure, relying on crystal structures and computational modeling. This strategy involves modifying particular CYP structures directly related to

catalysis of the substrate (site-directed mutagenesis). The disadvantages of this strategy lie in the fact that few crystal structures are available, the number of potential substrate-interacting residues is often quite high, and therefore an exhaustive analysis of possible cooperative effects is required (Girhard, Bakkes *et al.* 2015).

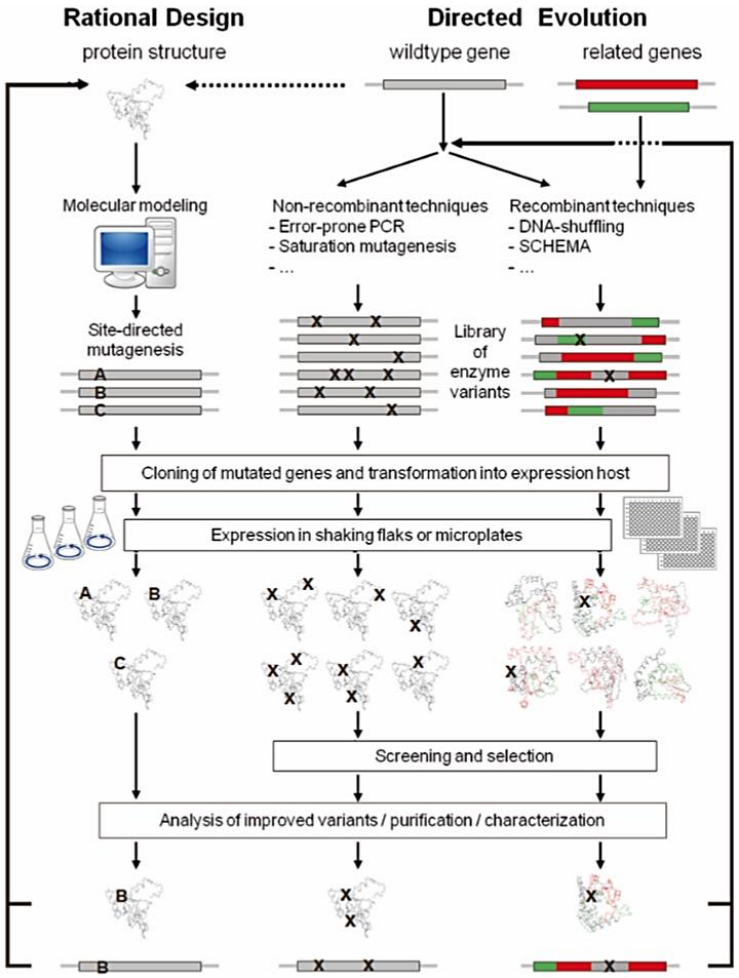

Figure 6. The two general strategies used to perform the engineering of CYPs and the identification of variants with desirable characteristics. Adapted from (Girhard, Bakkes *et al.* 2015).

CYP BM3 has been engineered for oxidation of alkanes, terpenes, heteroaromatics, steroids, alkaloids and other classes of chemical substances, catalyzing hydroxilations, epoxidations and demethylations of such compounds (Whitehouse, Bell *et al.* 2012). In some cases, turnover rates and coupling efficiencies of the CYP BM3 variants were comparable to those measured for the wild type enzyme with fatty acids (Fasan, Chen *et al.* 2007). There are also CYP BM3 variants designed by rational protein design and directed evolution with the

ability to oxidize certain human drugs and production of drug metabolites, reactions that are not performed by the wild type enzyme (Girhard, Bakkes *et al.* 2015).

The increase in the substrate spectrum catalyzed by CYP BM3 through its engineering has paved the way for the introduction of this CYP on various processes. Therefore, synthesis of alcohol, indigo, indirubin, and metabolites of propranolol and buspirone, as well as the degradation of pollutants, such as PAHs and petroleum components (e.g. alkanes) are now possible practical applications of this versatile enzyme (Kumar 2010).

1.4. Biotechnological approaches to the terpenoide indole alkaloid pathway of *C. roseus*

The importance of vincristine and vinblastine as anticancer drugs, allied to the fact that *C. roseus* produces these compounds in extremely low amounts has motivated a great interest in the alkaloids of this plant (Mujib, Ilah *et al.* 2012). The chemical synthesis of these complex compounds is highly difficult and, consequently, industrial production of the anticancer alkaloids relies on laborious and expensive extraction from the plant (Brown, Clastre *et al.* 2015). This poses a serious challenge to researchers and biotechnologists as to improve the yield and reduce the unit cost of these alkaloids (Mujib, Ilah *et al.* 2012).

In biotechnology based alkaloid enrichment programs, the following two approaches have currently been adopted: (1) the use of superior cell lines from various sources and optimization of cultural conditions/factors that control synthesis and (2) metabolic engineering and/or over expression of key enzymes (Mujib, Ilah *et al.* 2012).

The first is an empirical strategy involving a constant optimization of the media and culture conditions, changing many variables that can lead to increased production (or not) of certain alkaloids. This approach can also involve the feeding of *in vitro* cultures with precursors (tryptamine, geraniol) and/or the treatment of cultures with elicitors (UV light, methyl jasmonate) which has proved effective in increasing the level of certain TIAs such as ajmalicine, strictosidine, tabersonine and others (Almagro, Fernández-Pérez *et al.* 2015). Several studies have been developed in order to understand how the feeding of the cultures with precursors or elicitors led to increased accumulation of TIAs in cell cultures, verifying that this accumulation is caused through the overexpression of certain enzymes of the biosynthetic

pathway (Almagro, Fernández-Pérez *et al.* 2015). These studies included combined effect of elicitors as demonstrated by Almagro, Gutierrez *et al.* (2014), who confirmed that the joint action of cyclodextrins and methyl jasmonate leads to overexpression of several TIA enzymes like SLS and STR, resulting in the increased accumulation of ajmalicine and catharanthine.

The second approach is based on rational methods to obtain the increased biotechnological production of TIAs. To proceed with the metabolic engineering approach it is necessary to know in detail every step of the pathway (enzymes/genes, regulation and compartmentalization), wherein the increase of data available thanks to the "omics revolution" has been fundamental (Almagro, Fernández-Pérez *et al.* 2015).

Most of the known genes codifying enzymes from TIA biosynthetic pathway are tightly regulated by specific transcription factors in a coordinate manner together with developmental and environmental factors (Almagro, Fernández-Pérez *et al.* 2015). Several studies have linked overexpression of transcription factors to increased expression levels of some genes involved in TIA biosynthesis (Van Der Fits and Memelink 2001). This may then lead to accumulation of TIAs as confirmed by the study of Tang, Liu *et al.* (2011), in which the overexpression of ORCA3 (transcription factor) in hairy root cultures increased the content of catharanthine up to 2.5-fold in comparison with the control cultures. Another example is CrMPK3, an important component of the intracellular signaling pathway triggered by different stresses in *C.roseus.* The overexpression of CrMPK3 in *C. roseus* leaves raised both the expression of ORCA3 and genes involved in TIA biosynthetic pathway as well as serpentine, vindoline, catharanthine and vincristine levels (Raina, Wankhede *et al.* 2012).

Due to the compartmentalization of the TIA pathway, a constant multi-level transmembrane transport of alkaloid intermediate is required. Thus, identification of the transporters involved and their overexpression may be an effective strategy to affect the regulation of the alkaloid biosynthetic pathway and provide enhancement of alkaloid levels (Mujib, Ilah *et al.* 2012).

The utilization of heterologous hosts for the reconstitution of the TIA pathway has also been under study. As an example, the elegant work of Brown, Clastre *et al.* (2015) demonstrated how strictosidine can be produced *de novo* in a *Saccharomyces cerevisiae* host by the engineering of 14 known TIA pathway genes, along with additional seven genes and three gene deletions that enhance the chanelling of precursors.

In addition to these methods, the metabolic engineering of the TIA pathway could be made through the use of external enzymes to perform certain reactions (Shukla and Khanuja

2013). In this field, Glenn, Runguphan *et al.* (2013) propose the introduction of CYPs, due to the diversity and ability of these enzymes to perform numerous reactions including difficult chemical transformations.

The introduction of CYPs in engineering of numerous biosynthetic pathways for low-cost production of fine chemicals has been a widely used approach (Pateraki, Heskes *et al.* 2015). One example is the introduction of the CYP BM3 previously described, in the biosynthetic pathway of the sesquiterpene lactone endoperoxide artemisinin, a highly important antimalarial pharmaceutical. Since the wild type CYP BM3 did not perform the desired reaction, it was necessary to resort to enzymatic engineering to increase the size of the active site binding pocket and thus allow the acceptance of an unnatural substrate by the enzyme (Dietrich, Yoshikuni *et al.* 2009). The yield obtained with the CYP BM3 mutagenized enzyme was higher than when using the native enzyme (CYP71AV1) expressed in the same host. Here it is possible to observe the high potential of the combination of metabolic engineering with the introduction of an improved heterologous CYP, taking advantage of the great benefits associated with CYP BM3 (Pateraki, Heskes *et al.* 2015).

1.5. The aims of the project

As introduced above, the alkaloids vinblastine and vincristine are of great importance in medicine, having won their space among the most valuable anticancer drugs. These compounds are produced by the medicinal plant *C. roseus* through a quite complex biosynthetic pathway and are present in extremely low levels in the leaves of the plant, making commercial exploitation expensive and laborious. Since chemical synthesis is difficult and expensive, the metabolic engineering of this pathway has been a very attractive approach to increase production yield, but this approach have been unsuccessful so far. An obvious target for the metabolic engineering approach are the last bottleneck uncharacterized biosynthetic steps of vinblastine and vincristine, likely involving CYPs. Due to the characteristics of CYPs discussed above, of their biotechnological potential and previous success cases, the investigation of heterologous CYP enzymes as possible key catalysts of specific reactions of the TIA pathway is a promising strategy for the implementation of successful metabolic engineering approaches aiming to increase the levels of the anticancer vinblastine and vincristine, or to generate new derivatives.

Therefore, the main goal of this study was to perform the screening of selected collections of human and bacterial CYPs for their capacity to perform with high efficiency three reactions of the TIA pathway: the hydroxylation of tabersonine (TAB), already shown to be performed by a CYP, the hydroxylation/hydration of α-3',4'-anhydrovinblastine (AVLB) and the hydroxylation of vinblastine (VLB).

The characterization of the intracellular sorting of the heterologous CYPs in leaf cells of *C. roseus* was also a goal, since this is a crucial factor for the design of future metabolic engineering methodologies involving CYPs in *C. roseus*.

2. MATERIALS AND METHODS

2.1. Biological material

2.1.1. Plant material

Plants of *Catharanthus roseus* (L.) G. Don cv. Little Bright Eye were grown at 25°C in a growth chamber, under a 16 h photoperiod, using white fluorescent light with a maximum intensity of 70 µmol m^{-2} s^{-1}. Seeds were acquired from B&T World Seeds (France). For the isolation of protoplasts, the 2nd and 3rd leaf pairs of adult *C. roseus* plants were used.

2.1.2. Strains and vectors

In this work, the following *Escherichia coli* strains were used: DH5α for the expression of CYP BM3 and TOP10 for plasmid maintenance and replication. CYP BM3 variants were cloned in the expression vector pCWOri **(Table 2)**, which includes a lac operon, a taq promoter and an ampicillin-resistance marker.

For the expression of human CYPs, the *Saccharomyces cerevisiae* strain W(hR) was used. This yeast strain has a human CYP reductase expression cassette inserted in the yeast chromosome VIII in replacement of the original yeast reductase gene (Pompon, Louerat *et al.* 1996). Human CYP genes were cloned in the expression vector pYEDP60 **(Table 2)**, which contains both URA3 and ADE2 as selection markers and a GAL10-CYC1 hybrid promoter.

For the transient expression in C. roseus mesophyll protoplasts, the plant expression vector pMON999 **(Table 2)**, carrying the cyan fluorescent protein sCFP4/mTurquoise (Goedhart, van Weeren *et al.* 2010), was used. This vector has an ampicillin-resistance marker and the expression is controlled by a double 35S promoter. The plasmid map is in **Appendix A.**

Table 2. List and main features of the plasmids used in this work.

Vector	Type	Size (BP)	Promoter	Type of promoter
pCWOri	Bacteria expression vector	5000	*taq*	Inducible
pYEDP60	Yeast expression vector	9265	GAL10-CYC1	Inducible
pMON999	Plant expression vector	4108	e35S	Constitutive

2.1.3. CYP collections

The collection of CYP BM3 variants was developed by the group of Frances H. Arnold: Lewis, Mantovani *et al.* (2010) and Sawayama, Chen *et al.* (2009), and kindly provided by the California Institute of Technology **(Table 3).**

The collection of human CYPs (1A2, 1B1, 3A4, 3A5, 2C8, 2C9, 2C19, 2D6) as well as the *S. cerevisiae* strain W(hR) were kindly provided by Prof. Marie-Agnés Sari from the University Paris Descartes, UMR 8601 CNRS, France.

Table 3. Collection of CYP BM3 variants used in this study.

CYP BM3	Code Number
9-10A F87V TS – 4H9	2566
9-10A F87V TS – 7A1	2567
9-10A F87V TS – 8F11	2568
9-10A F87V TS – F1	2569
9-10A F87V TS – 4H5	2570
9-10A F87V TS – 8C7	2571
9-10A A78F	609
9-10A A82S	615
9-10A F87L	617
9-10A A82I	621

2.2. Culture media

For the growth of *E. coli*, were used Luria-Bertani (LB) medium (1% w/v tryptone, 0.5% w/v yeast extract and 1% w/v NaCl) and Terrific Broth (TB) medium (1.2% w/v tryptone, 2.4% w/v yeast extract, 0.4% v/v glycerol, 0.017 M KH_2PO_4 and 0.072 M K_2HPO_4).

For yeast manipulations were used several media:

- YPDE – YPD medium (1% w/v yeast extract, 2% w/v peptone and 2% w/v dextrose) supplemented with 3% (v/v) ethanol;
- YPDAW – YPD medium supplemented with 0.002% (w/v) adenine and 0.003% (w/v) tryptophan;
- SGI – SI medium (0.7% w/v Yeast Nitrogen Base and 0.1% w/v Casamino acids) supplemented with 2% (w/v) glucose and 0.003% (w/v) tryptophan.

2.3. Preparation of competent *E. coli* cells

Chemically competent *E. coli* cells were prepared according to a protocol adapted from Hanahan *et al.* (1991). A single colony from a LB medium with 1.5% (w/v) agar plate was inoculated in 25 mL of LB medium and incubated overnight (ON) at 37ºC with vigorous shaking. The entire ON culture was used to inoculate 225 mL of SOB medium (LB medium supplemented with 10 mM $MgCl_2$ and 10 mM $MgSO_4$). Cells were grown at 37ºC with vigorous shaking until the culture reached an OD_{600}~0.6. Culture was placed in ice for 10 min and the cells were recovered by centrifugation (3000 rpm for 5 min at 4ºC). The cells were gently resuspended in 100 mL of chilled RF1 solution (30 mM KCH_3COO, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl and 15% v/v glycerol) and kept on ice for 15 min. The cells were pelleted by centrifugation at 3000 rpm for 5 min at 4ºC and gently resuspended in 10 mL of chilled RF2 solution (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl and 15% v/v glycerol). Cells were incubated on ice for 45 min to 1 h. 50 and 100 µL aliquots were prepared in pre-chilled 1.5 mL tubes, frozen in liquid nitrogen and stored at -80ºC.

2.4. Transformation of *E. coli* by heat shock

All *E. coli* transformation procedures were carried out according to a protocol adapted from Sambrook, Fritsch *et al.* (1989). Typically, 50-100 µL of chemically competent cells were used for transformation either with plasmid DNA (pDNA) or a ligation reaction. The cells were incubated in ice for 15 min. Thereafter, the mixture of competent bacteria and DNA was heat-shocked at 42ºC for 1.5 min and placed back on ice for 2 min. It was added 1 mL of LB to the mixture and the cells were left to recover for 30 min at 37ºC. The cell suspension was centrifuged at 4000 rpm for 3 min at room temperature (RT) and 1 mL of supernatant was discarded. The cells were gently resuspended in the remaining volume and were plated onto LB agar plates supplemented with 100 µg/mL ampicillin (Amp100) for selection. Plates were then incubated ON at 37ºC.

2.5. Agarose gel electrophoresis

Unless stated otherwise, all DNA samples were separated in a 1% (w/v) agarose gel electrophoresis in 1x TAE buffer (40 mM Trizma base, 10% v/v acetic acid and 10 mM EDTA) supplemented with 0.5 µg/mL ethidium bromide (Bio-Rad) to allow visualization of the DNA bands under UV light. The electrophoretic run was performed at 90-120 V (PowerPac Basic, Bio-Rad) and as molecular marker the GeneRuler™ 1kb DNA Ladder (0.5 µg/µL; Thermo Scientific) was used. Images of gels were acquired in Gel Doc™ XR (Bio-Rad) system and processed with the Quantity One® (Bio-Rad) software.

2.6. Molecular Biology

2.6.1. Plasmid DNA extraction

All minipreparations (mini-prep) of pDNA were carried out using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and all midipreparations (midi-prep) were carried out using the Plasmid Midi Kit (Qiagen), according to manufacturers' instructions. The success of each extraction was confirmed by agarose gel electrophoresis.

2.6.2. Enzymatic manipulations

All commercial enzymes necessary in this study were purchased from Thermo Scientific and were used with their respective buffers and according to the manufacturer's instructions.

2.7. Screening of CYP BM3 variants for alkaloid conversion activity

2.7.1. Quality control of mini-preps

In order to confirm the quality of the received CYP BM3 mini-preps, a single colony of DH5α transformed with each variant was inoculated in LB medium supplemented with Amp100 and was grown ON at 37ºC at 225 rpm. The pDNA was extracted, run in an agarose gel and tested by a restriction reaction. CYP BM3 gene was cloned between *Bam*HI and *Eco*RI restriction sites and, therefore, the reaction was performed with these enzymes. The restriction products were also analyzed by agarose gel electrophoresis.

2.7.2. Induction of CYP BM3 expression

For the variants 2566 to 2571 **(Table 3)**, the protein expression protocol was performed following an adapted methodology from Lewis, Mantovani *et al.* (2010). A single colony of DH5α transformed with each variant was inoculated in 1 mL of LB medium supplemented with Amp100 and incubated ON at 37ºC at 225 rpm. 50 µL of these ON cultures were used to inoculate 1 mL of TB medium supplemented with Amp100 and then these cultures were incubated for 4 hours at 37ºC at 225 rpm. After this time, the incubation temperature was reduced to 25ºC and 30 min after this decrease it was added isopropyl β-D-1-thiogalactopyranoside (IPTG) and δ-aminolevulinic acid (dALA), both to a final concentration of 1 mM. The cultures were allowed to continue growing for another 24 hours at 25ºC at 225 rpm. Cells were then pelleted for 15 min at 3000 g at 4ºC and stored at -20ºC.

For the variants 609 to 621 **(Table 3)** the protein expression was performed following the procedure of Sawayama, Chen *et al.* (2009) with a few optimizations. A single colony of DH5α transformed with each variant was inoculated in 1 mL of LB medium supplemented with Amp100 and incubated for 24 hours at 30ºC at 225 rpm. 50 µL of these ON cultures were used to inoculate 1 mL of TB medium supplemented with Amp100 and then these cultures were incubated for 5 hours at 30ºC at 225 rpm. After this time, the cultures were placed on ice for 15 min. 1 mM of IPTG and of dALA were added and the cultures were allowed to grow for another 24 hours at 20ºC at 225 rpm. Cells were then pelleted for 15 min at 3000 g at 4ºC and stored at

-20ºC. This optimized protocol was reached after several tests involving changes in the concentration of IPTG and incubation temperature adjustments.

For each variant, the induction was performed in triplicate and it was also included a non-induced control (without the addition of IPTG and dALA).

2.7.3. Cell lysis

The buffer used in cell lysis differed depending on the substrate used in the activity assay: 0.1 M phosphate buffer pH 6.2 was used for assays with vindoline (VIND) and α- 3',4'-anhydrovinblastine (AVLB), and 0.1 M Tris-acetate buffer pH 5.0 was used for assays with tabersonine (TAB).

The pellets from the previous section were allowed to defrost at room temperature. Then, they were resuspended in the lysis buffer composed by the appropriate buffer, 10 mM $MgCl_2$, 0.5 mg/L lysozyme and 8 U/mL DNaseI. The lysis reaction was incubated for 1 hour at 37ºC and finally centrifuged for 15 min at 5000 x g at 4ºC. The supernatants (cell lysates) were used in SDS-PAGE analysis, CO binding assay and alkaloid conversion activity.

2.7.4. Analysis of CYP BM3 expression by SDS-PAGE

The level of expression of CYP BM3 variants was analyzed by vertical electrophoresis on polyacrylamide gel under denaturing conditions performed in a Mini-Protean II (Bio-Rad), using 10% running gel and 5% stacking gel **(Table 4)**. Samples were prepared by adding the loading buffer for SDS-PAGE (50 mM Tris-HCl pH 6.8, 2% v/v SDS, 0.1% w/v bromophenol blue, 10% v/v glycerol and 1% v/v β-mercaptoethanol) and then they were boiled for 5 min immediately before application to gel. The molecular weight marker used was the PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific). Electrophoresis was conducted at 100 V in running buffer for SDS-PAGE (25 mM Trizma base, 192 mM glycine and 0.1% w/v SDS). After the run, gels were stained with a solution of Coomassie Brilliant Blue R (0,125% w/v *Coomassie Brilliant Blue R*, 50% v/v methanol and 10% v/v acetic acid) for 30 min and then destained with a solution of 5% (v/v) methanol and 7.5% (v/v) acetic acid ON. Images were

acquired in a densitometer GS-800™ (Bio-Rad) and processed in Quantity One® (Bio-Rad) software.

Table 4. Composition of polyacrylamide gels used in SDS-PAGE.

Running Gel	Stacking Gel
380 mM TRIS-HCl pH 8.8	125 mM TRIS-HCl pH 6.8
10% (v/v) Bisacrylamide	5.1% (v/v) Bisacrilamide
0.1% (v/v) SDS	0.1% (v/v) SDS
0.1% (v/v) APS	0.1% (v/v) APS
0.4% (v/v) TEMED	0.1% (v/v) TEMED

2.7.5. CO binding assay

CO binding assay is used to measure the amount of folded heme proteins, since they absorb light at 450 nm when CO is bound to the iron ion of the heme domain (Omura and Sato 1964, Schenkman and Jansson 2006). These assays were performed according to the protocol of Omura and Sato (1964) and spectral readings were acquired in a UV-2401PC spectrophotometer (Shimadzu) at 25ºC.

Cell lysates (from section 2.7.3) where diluted 5x in 0.1 M phosphate buffer pH 6.2 in a cuvette. A few milligrams of sodium dithionite were added to the cuvette and the baseline for wavelengths between 400 and 500 nm was established. Then, the solution was bubbled for 20 sec with CO and allowed 2 min to rest. After this time, the spectrum was recorded.

The calculation of the concentration (mM) of folded CYP BM3 was carried out according with de Beer's Law equation $A = \varepsilon \cdot c \cdot l$ considering that A (Absorbance) = $A_{450nm} - A_{490nm}$, l (light path of the cuvette) = 1 cm, and the extinction coefficient $\varepsilon_{450nm–490nm} = 91\ mM^{-1}\ cm^{-1}$ (Schenkman and Jansson 2006).

2.7.6. Alkaloid conversion activities

Alkaloid conversion assays were performed by monitoring the NADPH consumption in real-time by absorbance readings at 340 nm (Rastogi, Khanna *et al.* 2002), acquired in a UV-2401PC spectrophotometer (Shimadzu) at 25ºC. Spectra were recorded between 300 and 400

nm and 5 in 5 minutes over 45 minutes. For each CYP BM3 variant three different reactions were recorded: control (cell lysate not induced, i.e. without CYP BM3) plus alkaloid, cell lysate induced with no alkaloid added, and cell lysate induced plus alkaloid. The buffer used in the assays with TAB was 0.1 M Tris-acetate pH 5.0 and with VIND and AVLB was 0.1 M phosphate buffer pH 6.2. Stock solutions of NADPH were prepared with the respective buffer and alkaloids were prepared in methanol.

The mixtures were performed in a cuvette to a final volume of 350 µL: the buffer (variable volume), 70 µL of cell lysate, and the alkaloid (250 µM VIND, 250 µM TAB or 125 µM AVLB). The baseline was defined for wavelengths between 300 and 400 nm. Finally, it was added 500 µM NADPH and the first spectrum was recorded (time = 0). After 45 min, the reaction was transferred to a 1.5 mL tube and continued at 25ºC until complete 2 hours. During the monitoring, whenever verified a complete consumption of NADPH, an equal amount was again added to the reaction. Before stopping the reaction, 25 µg/mL papaverine was added as an internal control for the calculation of extraction efficiency. To stop the reaction, tubes were frozen in liquid nitrogen and store at -80ºC. Finally, samples were lyophilised and stored at -80ºC.

2.7.7. Alkaloid extraction and HPLC analysis

The alkaloid conversion by CYP BM3 was analyzed through High Performance Liquid Chromatography (HPLC). Lyophilised samples from the previous section were resuspended in 1 mL of methanol and vortexed for 15 min. Then, the mixtures were sonicated for 30 min and centrifuged at 13225 x g for 15 min at RT. The supernatants were collected into glass tubes and completely dried with nitrogen gas. Finally, samples were ressuspended in 300 µL of methanol and analysed by HPLC.

Separations were achieved on a Phenomenex Gemini® C18 column (150 x 4.6 mm; 3 µm particle size) and a guard column with the same characteristics. The mobile phase consisted of methanol (A) and 25 mM ammonium acetate pH 10.0 (B). The gradient was as follows: 40% B at 0 min, 38% B at 6 min, 38% B at 35 min, 25% B at 40 min, 10% B at 45 min, 40% B at 50 min, and 40% B at 55 min. Elution was performed with a flow rate of 0.5 mL/min and the injection volume was 20 µL. Spectral data from all peaks were accumulated in the range of 200–600 nm, and chromatograms were recorded at 260 nm. The HPLC system (Jasco

Corporation, Japan) was composed by a low-pressure quaternary gradient unit (model PU-2089 Plus) with an in-line degasser (model DG-1580-54), a spectrophotometric detector with diode array (model MD-1510) and an automatic injector (model AS-950). The system control and data analysis was processed on a Jasco ChromPass Chromatography Data software (version 1.8.6.1).

Quantification of alkaloids was achieved using calibrations curves obtained with external standards by the host lab, and normalization was performed using papaverine as internal standard. Alkaloid concentrations were used to calculate the rate of alkaloid consumption in katal and katal/mM for CYP BM3 variants and each alkaloid tested.

2.8. Subcellular localization of CYP BM3 in *C. roseus* mesophyll protoplasts

In order to know the localization of the CYP BM3 in plant cells, three variants chosen randomly (617, 2567 and 2569) were cloned in the plant expression plasmids pMON999_sCFPstop and pMON999_sCFP to generate respectively N-terminal and C-terminal fusions with the reporter gene cyan fluorescent protein (CFP), and localization was investigated by PEG-mediated transformation of mesophyll protoplasts.

2.8.1. N-terminal fusions (CYP BM3-CFP)

2.8.1.1. Amplification of *CYP BM3*

CYP BM3 variants genes were amplified from the plasmid pCWOri by polymerase chain reaction (PCR). Primers were designed to amplify the full coding sequence and included the restrictions sites for *Bgl*II and *Xba*I to allow directional cloning **(Table 5)**. *Pfu* DNA polymerase (Thermo Scientific) was used, since this enzyme has capacity exonuclease 3'→5' (proofreading), decreasing the probability of occurrence of mutations in the amplified during the PCR reaction. The conditions used in the PCR reaction are indicated in

Table 6.

The PCR program consisted of one cycle at 95ºC for 3 min, 35 cycles at 95ºC for 30 sec, 55ºC for 30 sec and 72ºC for 6.5 min and a final extension at 72ºC for 7 min. The PCR product was run on agarose gel electrophoresis and the band of interest was recovered from the agarose gel using the GeneJET™ Gel Extraction Kit (Thermo Scientific), according to the manufacturer's instructions.

Table 5. Primers for *CYP BM3* amplification for *CFP* N-terminal fusion. The *CYP BM3* initiation codon is in red and the restriction site of the respective enzyme is underlined.

Name	Sequence (5' – 3')	BP	Tm (°c)
BM3FWBGLII	CCAGATCTATGACAATTAAAGAAATGCCTCAGCC	34	56.2
BM3REVXBAI	AAATCTAGACCCAGCCCACACGTCTTTTGC	30	61.6

Table 6. Composition of PCR reaction.

Reagents	Volumes (µL)
DNA (1:20)	0.5
***Pfu* buffer (10x)**	5
Deoxynucleotide mix (dntps, thermo scientific) (10 mm)	1
Primer forward (10 µm)	2
Primer reverse (10 µm)	2
***Pfu* DNA polymerase**	0.5
H_2O	39

2.8.1.2. Ligation to the plant expression vector

The purified products and pMON999_sCFPstop were digested with the restriction enzymes *Bgl*II and *Xba*I, following the manufacturer's instructions. The reactions were incubated at 37ºC for 3 hours. After 2 hours of incubation, 1U of FastAP Thermosensitive Alkaline Phosphatase (FastAP) was added to the restriction reaction of pMON999_sCFPstop to avoid plasmid re-ligation. Then, the restriction reactions were inactivated through the addition of 20 mM EDTA pH 8.0 and incubation at 80ºC for 20 min. Finally, the restriction product was run on agarose gel electrophoresis and the bands of interest were extracted and purified with GeneJET™ Gel Extraction Kit (Thermo Scientific), according to the manufacturer's instructions.

Ligation reactions were set with T4 DNA ligase (Thermo Scientific) as recommended by the manufacturer, and incubated for 4 hours at RT. Before transforming, the ligations were heat inactivated at 65ºC during 10 min. *E. coli* competent cells were transformed with the ligation reaction product.

2.8.2. C-terminal fusions (CFP-CYP BM3)

2.8.2.1. Modification of pMON999_sCFP

Since *CYP BM3* variants have different sequences, with different point mutations, was performed a restriction analysis in order to know which restriction enzymes could be used to perform the C-terminal fusions with *CFP*: *Bam*HI, *Eco*RI, *Kpn*I and *Sac*I. Since only *Bam*HI and *Eco*RI did not cut *CYP BM3* sequences, the option was to change the original location of the *CFP* in the pMON999_sCFP, in order to insert *CYP BM3* among those restriction sites.

The *CFP* gene was amplified by PCR. Primers were designed to amplify the full coding sequence and included the restrictions sites for *Bgl*II and *Eco*RI to allow directional cloning **(Table 7)**. The conditions used in the PCR reaction are indicated in Table 5. The PCR program consisted of one cycle at 95ºC for 3 min, 35 cycles at 95ºC for 30 sec, 55ºC for 30 sec and 72ºC for 2 min and a final extension at 72ºC for 7 min. The PCR product was run on agarose gel electrophoresis and the band of interest was recovered from the agarose gel using the GeneJET™ Gel Extraction Kit (Thermo Scientific), according to the manufacturer's instructions.

The amplified and purified products and the plasmid pMON999 (without *CFP*) were digested with the restriction enzymes *Bgl*II and *Eco*RI, following the manufacturer's instructions. The restriction and ligation reactions were performed as described in section 2.8.1.2.

Table 7. Primers for *CFP* amplification. The *CFP* initiation codon is in red and the restriction site of the respective enzyme is underlined.

NAME	SEQUENCE (5' – 3')	BP	Tm (°C)
CFPFWBGLII	CCAGATCTATGGTGAGCAAGGGCGAGGAGC	30	66.3
CFPREVECORI	CCGAATTCCTTGTACAGCTCGTCCATGCCG	30	65.3

2.8.2.2. Amplification of *CYP BM3*

The amplification of *CYP BM3* variants was performed as described in section 2.8.1.1, with the exception of the primers **(Table 8)**.

Table 8. Primers for *CYP BM3* amplification for *CFP* C-terminal fusion. The *CYP BM3* initiation and termination codons are in red and the restriction site of the respective enzyme is underlined.

Name	Sequence (5' – 3')	BP	Tm (°C)
BM3FWEcoRI	CCGAATTCATGACAATTAAAGAAATGCCTCAGCC	34	61.1
BM3REVBamHI	TTTGGATCCTTACCCAGCCCACACGTCTTTTGC	33	61.9

2.8.2.3. Ligation to the plant expression vector

The purified products from the previous section and the modified pMON999_sCFP were digested with the restriction enzymes *Eco*RI and *Bam*HI, following the manufacturer's instructions. The restriction and ligation reactions were performed as described in section 2.8.1.2.

2.8.3. Selection of positive clones

The presence of positive clones was screened by colony PCR. Single isolated colonies were picked to transfer a small amount of bacteria into a PCR tube previously filled with 10 µL of H_2O, and to re-plate onto LB-agar supplemented with Amp100. The colonies were numbered to facilitate their identification. The conditions used in the PCR reaction are indicated in Table **9**. Primers used were chosen in a way that it was possible to check for the presence of the inserted DNA. The PCR reactions were set by adding 10 µL of the mix to each tube. The PCR program consisted of one cycle at 95ºC for 5 min, 35 cycles at 95ºC for 30 sec, 50ºC for 30 sec and 72ºC for 5 min and a final extension at 72ºC for 5 min. The positive clones were identified by gel electrophoresis. Positive clones were picked from the second plate to inoculate 5 mL of LB medium supplemented with Amp100. Cultures were grown ON at 37ºC at 225 rpm. Mini-preps were performed as described in 2.6.1. All the constructs were confirmed by sequencing (STAB Vida) and analysis of sequence homology between the predicted sequences and the obtained sequences was performed using the multiple sequence alignment program ClustalW2.

For each construct, one error-free clone was used for midi-prep purification in order to obtain highly pure and concentrated DNA to be used in transformation of *C. roseus* mesophyll protoplasts.

Table 9. Composition of colony PCR reaction.

Reagents	Volumes (µL)
***Dreamtaq* buffer (10x)**	2
dNTPs (10 mm)	0.4
Primer forward (10 µm)	0.4
Primer reverse (10 µm)	0.4
***Dreamtaq* DNA polymerase**	0.1
H_2O	6.7

2.8.4. Isolation of *C. roseus* mesophyll protoplasts

C. roseus mesophyll protoplasts were isolated according to Duarte, Rocha *et al.* (2011). Leaves of the 2nd or 3rd pair were cut into ~1 mm strips, after excising the central vein, and were immediately transferred to a Petri dish, keeping the abaxial face down, with 10 mL of digestion medium freshly prepared: 2% (w/v) cellulase Onozuka R-10 (Duchefa), 0.3% (w/v) macerozyme Onozuka R-10 (Serva) and 0.1% (v/v) pectinase (Sigma) dissolved in MM buffer (0.4 M mannitol, 20 mM MES, 20 mM KCl and 10 mM $CaCl_2.2H_2O$, pH 5.7). The leaf strips were vacuum infiltrated for 15 min and the digestion was further allowed for 3 h, at 25ºC, in the dark. The enzyme solution containing protoplasts was filtered through a 100 µm nylon mesh and the filtrate was centrifuged at 60 x g for 5 min at RT, with acceleration and deceleration set at the minimum. The protoplasts pellet was washed twice in MM buffer. In the last wash, protoplasts were resuspended in a minimum volume of MM buffer and it was determined the protoplast concentration using a haemocytometer under the optical microscope. In the meantime, protoplasts are left to rest for 30 min at RT. After that time, the suspension was once again centrifuged and the pellet was resuspended in an appropriate volume of MMg buffer (0.4 M mannitol, 15 mM $MgCl_2$ and 4 mM MES, pH 5.7) to yield a final protoplast concentration of $5x10^6$ cells/mL.

2.8.5. PEG-mediated transformation of protoplasts

C. roseus leaf protoplasts were transformed following the procedure described by Duarte, Rocha *et al.* (2011). 20 µg of plasmid DNA (midi-prep) were mixed with 100 µL of protoplast suspension in a 2 mL round bottom tube. 110 µL of PEG solution freshly prepared (40% w/v PEG, 0.2 M mannitol and 0.1 M $CaCl_2.2H_2O$) was added drop by drop to this mixture, flicking the tube after every drop. The tubes were left to incubate for 15 min at RT and then 440 µL of W5 solution (154 mM NaCl, 125 mM $CaCl_2.2H_2O$, 5 mM KCl and 2 mM MES, pH 5.7) were slowly added. Protoplasts were recovered by a 600 rpm centrifugation for 2 min, with acceleration and deceleration set at the minimum. The supernatant was removed and the protoplasts were gently resuspended in 100 µL of W5 solution and transferred to a 15 mL falcon tube containing 900 µL of W5. Transformed protoplasts were incubated in the dark at 25ºC, with the tubes lying in a slight slope, for at least 24 hours.

2.8.6. Confocal microscopy

After 24, 48 and 72 hours of incubation, fluorescence inside the protoplasts was observed using a Leica SP2 AOBS SE confocal microscope. Visualization of CFP was performed using an excitation wavelength of 458 nm and an emission wavelength window from 465 to 540 nm. Visualization of chloroplast autofluorescence was performed using an emission wavelength window from 650 to 750 nm.

2.9. Characterization of human CYPs

2.9.1. Transformation of yeast

The yeast strain W(hR) was transformed following the lithium acetate method (Gietz and Woods 2002). The following stock solutions were previously prepared and stored at RT: 10x lithium acetate solution (1M lithium acetate, pH 7.5 adjusted with dilute acetic acid), filter sterilized; 10x TE buffer (100 mM Tris-HCl pH 7.5 and 10 mM EDTA pH 8.0), autoclaved. 50% (w/v) PEG was freshly prepared and filter sterilized.

A single colony was used to inoculate 5 mL of YPDAW medium. Culture grew for ~24 hours at 30ºC at 225 rpm. The OD600 was measured and a suitable amount of the saturated culture was added to 100 mL of fresh YPDAW medium in order to obtain an OD600 of 0.2 (~$2x10^6$ cells/mL). This culture grew in the same conditions until reached an OD600 of 1 (~$2x10^7$ cells/mL). Then, the culture was centrifuged for 5 min at 5000 rpm at RT and the pellet was resuspended in 20 mL of sterile water. This solution was again centrifuged and the pellet was resuspended in 2 mL of sterile water. This solution was transferred to a new tube and centrifuged 5 min at 7000 rpm at RT. The pellet was resuspended in 2 mL of lithium buffer (1 vol of 10x TE buffer, 1 vol of 10X lithium acetate and 8 vol of water) freshly prepared and centrifuged 5 min at 7000 rpm at RT. The pellet was again resuspended in 500 µL of lithium buffer and the cells were incubated ≤1 hour at 30°C. For each transformation a 1.5 mL tube with 50 µg of single stranded carrier DNA was prepared (boiled for 5 min and then chilled for 5 min before use) with 1 µg of pDNA. Finally, 50 µL of yeast suspension and 300 µL of PEG solution (8 vol of 50% PEG, 1 vol of 10x TE buffer and 1 vol of 10x lithium acetate) freshly prepared were added and stirred thoroughly. This mixture was gently shacked for 30 min at 30ºC. Then, was added 40 µL of DMSO and was applied a heat shock for exactly 15 min at 42ºC. The mixture was centrifuged for 10 sec at 7000 rpm at RT and the pellet was resuspended in 1 mL of 1x TE buffer. This step was repeated once. 200 µL of the mixture was plated on selective medium (SGI) and incubated at 30ºC until the transformed yeast colonies appeared.

2.9.2. Selection of positive colonies

The efficiency of transformation of yeast W(hR) with the plasmids of interest was confirmed by colony PCR. Primers were designed to hybridize in the CYP sequence and other primers were complementary to the vector and were described by Abécassis, Pompon *et al.* 2000 **(Table 10).**

10 µL of a 20 mM NaOH solution was added to each PCR tube. An isolated colony of each clone was picked and resuspended in this solution, and re-plated onto SGI-agar. The samples were incubated at 99ºC for 10 min and allowed to cool to RT. The PCR reaction mix was prepared as indicated in **Table 11** and distributed to new PCR tubes. To each tube 1 µL of the respective sample was added. The PCR program consisted in one cycle at 95ºC for 5 min, 35 cycles at 95ºC for 30 sec, 50ºC for 30 sec and 68ºC for 10 min, and a final extension at 72ºC

for 10 min. The positive clones were identified by agarose gel electrophoresis. A positive control with pDNA as template was included.

Table 10. Primers for yeast colony PCR.

CYP	NAME	HYBRIDIZATION SITE	SEQUENCE (5' – 3')	BP	Tm (°C)
1A2	FwGal	Plasmid promotor GAL	CGTGTATATAGCGTGGATGGCCAG	24	59.3
	Rv1A2	CYP sequence	GAAGGATGGGGAAGAAGTCC	20	54.9
1B1	FwGal	Plasmid promotor GAL	CGTGTATATAGCGTGGATGGCCAG	24	59.3
	Rv1B1	CYP sequence	TGTCCAGGATGAAGTTGCTG	20	55.0
2C19	Fw2C19	CYP sequence	GACCTATGTCCTGACTGTGG	20	54.5
	RvPGK	Plasmid terminator PGK	GCACCACCACCAGTAG	16	52.6
3A4and 3A5	FwGal	Plasmid promotor GAL	CGTGTATATAGCGTGGATGGCCAG	24	59.3
	Rv3A	CYP sequence	GGCTCACAGTCTTGCTTACTC	21	55.5

Table 11. Composition of yeast colony PCR reaction.

Reagents	Volumes (µL)
Betaine (5 m)	3.2
***Dreamtaq* buffer (10x)**	1
dNTPs (10 mm)	0.2
Primer forward (100 µm)	0.2
Primer reverse (100 µm)	0.2
***Dreamtaq* DNA polymerase**	0.1
H_2O	4.1

2.9.3. Protein expression

The induction of human CYPs expression was performed according to the High Density Procedure described by Pompon, Louerat *et al.* (1996). Briefly, a positive colony was toothpicked and inoculated in 30 mL of SGI. The culture was grown at 30ºC at 225 rpm until the stationary phase (~48 hours). The culture was diluted 1:50 in 250 mL YPGE medium and cells were grown at 28ºC at 160 rpm until an OD_{600} of ~6-7 (~24 hours). Induction was started by the addition of 10% of 200 g/L galactose sterile solution and continued for ~15 hours. After this time, is advisable to start the microsomes isolation immediately.

2.9.4. Isolation of microsomes

The isolation of microsomes containing human CYPs was performed following the Mechanical Procedure described by Pompon, Louerat *et al.* (1996), with some optimizations.

All the solutions necessary to this protocol were previously prepared and sterilized: TE buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA pH 8.0), TEK buffer (0.1 M KCl in TE), TES B (0.6 M sorbitol in TE), TEG (20% v/v glycerol in TE). TE and TES B should be chilled for the procedure.

The cells were centrifuged for 4 min at 7500 x g at RT. The pellet was resuspended in 25 mL of TEK and the suspension was allowed to rest for 5 min at RT. After this time, it was centrifuged for 4 min at 6000 x g at RT and the pellet was resuspended in 2.5 mL of TES B. Zirconia beads with 0.5 mm of diameter (BIOSPEC) were gently added until skimming the top of the cell suspension. Then, yeast cell walls were mechanically disrupted by hand shaking for 15 min (occasionally the tube should be placed on ice to cold down). Cell disruption was monitored under an optical microscope and performed until a significant portion of cells were lysed. From this step forward the protocol was proceed on ice. 5 mL of TES B was added to the cell lysate to wash the beads and the supernatant above the top of the beads was recovered. This step was repeated three times. The supernatants were pooled and centrifuged at 27000 x g for 10 min at 4ºC. The supernatant was recovered and centrifugation was repeated. The supernatant was diluted twice with TES B and microsomes were precipitated through the addition of a PEG solution (0.1 g/mL PEG-4000 and 0.15 M NaCl) freshly prepared. The suspension was allowed to rest for at least 15 min on ice. The microsomes were recovered by centrifugation at 12000 x g for 10 min at 4ºC. Pellet was transferred to a potter to facilitate homogenization and was resuspended in 2 mL of TEG. Finally, aliquots were prepared in pre-chilled 1.5 mL tubes, frozen in liquid nitrogen and stored at -80ºC.

2.9.5. CO binding assay

CO binding assays were performed as described in section 2.7.5, with the difference that microsomes were diluted 40 and 80x in the cuvette using the TE buffer used for isolation of microsomes.

3. RESULTS AND DISCUSSION

3.1. Selection of CYP BM3 variants and human CYPs

To search for CYP surrogates capable of performing the bottleneck uncharacterized oxidation steps in the biosynthesis of VLB and VCR, and the conversions of tabersonine with higher efficiency, CYP BM3 variant libraries and human CYPs were considered.

The bacterial CYP BM3 is one of the most active CYPs known possibly due to the fact that its hydroxylase, reductase and electron-transfer domains are all in one contiguous polypeptide chain (Lewis, Mantovani *et al.* 2010). CYP BM3 is soluble, readily over-expressed in a variety of heterologous hosts, and requires only atmospheric oxygen and a supply of NADPH for activity. These properties led several groups to expand the substrate scope of BM3 by extensive genetic diversification methodologies with the goal of creating efficient enzymes for a variety of catalytic transformations (Lewis, Mantovani *et al.* 2010). In this study, the catalytic profiles of the CYP BM3 variants reported by Sawayama, Chen *et al.* (2009) and Lewis, Mantovani *et al.* (2010) were screened to select variants potentially accepting TIAs as substrates.

Sawayama, Chen *et al.* (2009) evaluated 120 CYP BM3 variants for their capacity to accept as substrate the drug astemizole, which is a nitrogen-containing heterocyclic compound, similarly to alkaloids, and has a molecular weight in the range of monomeric TIAs. Among the 15 variants shown to accept astemizole as substrate, the 4 capable of performing hydroxylations of the drug with a high % of conversion were selected for this study **(Table 12)**. The selected BM3 variants are all derived from the alkane-hydroxylating 9-10A variant, which is 13 mutations away from the wild type. These variants were constructed by using error-prone PCR and targeted mutagenesis of active site residues.

The other CYP BM3 variants selected for this study are the ones developed by Lewis, Mantovani *et al.* (2010) to accept opiate alkaloids and steroid derivatives as substrates. The authors hypothesized that the shape and volume of these substrates exceeded the capacity of the active site of CYP BM3 variants originated from error prone PCR and concluded that sufficient expansion of the active site to obtain significant improvements in activity would require more extensive mutation. Therefore, they used extensive replacement of bulky active site residues with alanine. Based on computational models of these compounds docked in the enzyme active site and enzymatic engineering, these authors were able to identify active sites residues likely to clash with large substrates and create a combinatorial library. The BM3 variant 9-10A F87V TS, a thermostable form of an enzyme (9-10AF87V) previously found to

have activity against a-1,2,3,4-tetramethoxymethyl xylose was selected as a parent for library creation. A library with 256 members was generated, of which 6 were shown to depict a significant activity towards alkaloids and were therefore selected in the present study **(Table 12)**.

A particular set of CYPs comprises the primary enzymes responsible for drug metabolism in humans (Guengerich 2015). These CYPs accept a broad range of substrates and, since most drugs either come directly from plants or have been designed from plant products, it has been suggested that they evolved initially to detoxify defense plant secondary metabolites, when animal species colonized land and started using plants as primary food source (Lewis 2001). Therefore, it is predictable that some of the human drug-metabolism CYPs accept alkaloids as substrates, since these compounds are thought precisely to play a defense role against herbivores. This is precisely the case of CYP 3A4 and 3A5 which have been confirmed as vinblastine and vincristine metabolizers (Yao, Ding *et al.* 2000, Dennison, Kulanthaivel *et al.* 2006). The major drug-metabolizing CYPs come from families CYP1, CYP2 and CYP3, which include the 8 CYPs selected for this study due to their particularly broad substrate profile **(Table 12).**

Table 12. CYP BM3 variants and human CYPs selected for evaluation of TIA conversion activity.

CYP		Substrates	Reference
BM3 9-10A A78F	**609**	Astemizole and verapamil	(Sawayama, Chen *et al.* 2009)
BM3 9-10A A82S	**615**	Astemizole	(Sawayama, Chen *et al.* 2009)
BM3 9-10A F87L	**617**	Astemizole and verapamil	(Sawayama, Chen *et al.* 2009)
BM3 9-10A A82I	**621**	Astemizole	(Sawayama, Chen *et al.* 2009)
BM3 9-10A F87V TS – 4H9	**2566**	Alkaloids and thioglycosides	(Lewis, Mantovani *et al.* 2010)
BM3 9-10A F87V TS – 7A1	**2567**	Alkaloids and thioglycosides	(Lewis, Mantovani *et al.* 2010)
BM3 9-10A F87V TS – 8F11	**2568**	Alkaloids and steroids	(Lewis, Mantovani *et al.* 2010)
BM3 9-10A F87V TS – F1	**2569**	Steroids	(Lewis, Mantovani *et al.* 2010)
BM3 9-10A F87V TS – 4H5	**2570**	Alkaloids	(Lewis, Mantovani *et al.* 2010)
BM3 9-10A F87V TS – 8C7	**2571**	Alkaloids and thioglycosides	(Lewis, Mantovani *et al.* 2010)
CYP1A2		Alkaloids, arachidonic acid and several drugs	(Stiborová, Frei *et al.* 2001); (Ward, Morocho *et al.* 2004)
CYP1B1		Arachidonic acid, retinoids and steroids	(Choudhary, Jansson *et al.* 2004); (Jang, Kim *et al.* 2012)
CYP2C8		Arachidonic acid and several drugs	(Fisher, Lickteig *et al.* 2009); (Kot and Daniel 2008)
CYP2C9		Monoterpenoids, urea and several drugs	(Jiang, He *et al.* 2006); (Jeong, Nguyen *et al.* 2009); (Kringen, Haug *et al.* 2010)
CYP2C19		Steroids and several drugs	(Korhonen, Turpeinen *et al.* 2008); (Laine, Auriola *et al.* 2009)
CYP2D6		Alkaloids, coumarin and several drugs	(Fisher, Lickteig *et al.* 2009); (Cook Sangar, Anandatheerthavarada *et al.* 2009); (Laine, Auriola *et al.* 2009)
CYP3A4		Alkaloids and several drugs	(Lalovic, Phillips *et al.* 2004); (Laine, Auriola *et al.* 2009)
CYP3A5		Alkaloids and steroids	(Aoyama, Yamano *et al.* 1989); (Lalovic, Phillips *et al.* 2004)

3.2. Expression of the CYP BM3 variants in *E.coli* DH5α

3.2.1. Transformation of *E.Coli* with the plasmids pCWOri harboring the CYP BM3 variant genes

In order to express the CYP BM3 variants it was necessary to transform *E. coli* DH5α bacteria with the plasmids pCWOri containing the cloned genes of the variants of CYP BM3. This bacterial strain was chosen to express the CYP BM3 variants since it is the strain used in numerous studies involving the expression of this and other CYPs (Sawayama, Chen *et al.* 2009, Zelasko, Palaria *et al.* 2013). After transformation and selection, the pDNA was extracted in order to determine whether the competent bacteria were correctly transformed with the pCWOri harboring the CYP BM3 variant genes (pCWOri + CYP BM3). The pDNA was analyzed by agarose gel electrophoresis **(Figure 7)** and showed a band for each variant located above the 6000 BP, higher than the ~ 5000 BP reported for pCWOri (Cirino 2004).

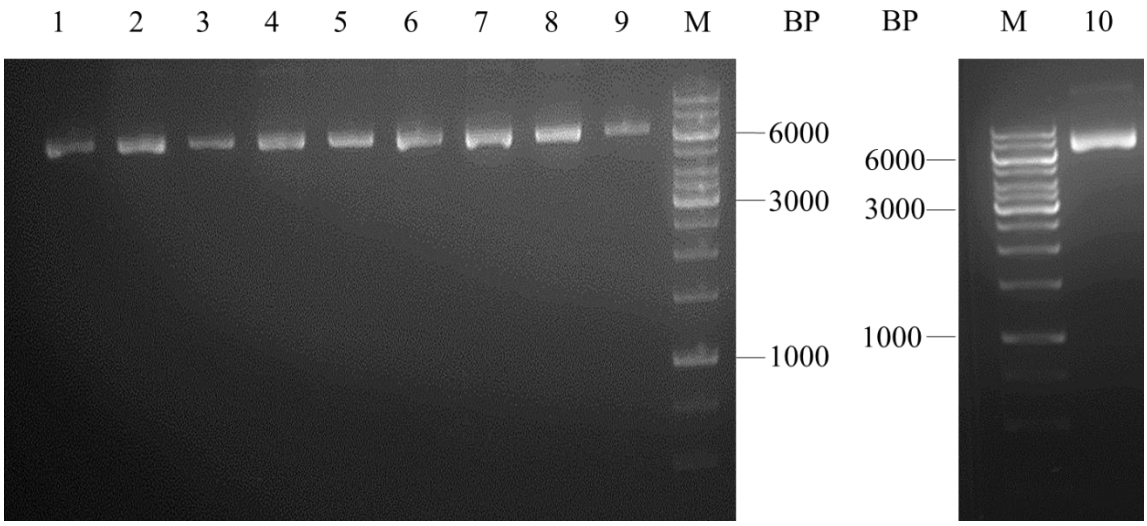

Figure 7. Agarose gel electrophoresis of pDNA extracted from *E.coli* DH5α transformed with different BM3 variants. Lanes 1 to 10 correspond respectively to variants 609, 615, 621, 2566, 2567, 2568, 2569, 2570, 2571 and 617.

In order to confirm if the genes of the different variants of CYP BM3 were indeed correctly cloned in the plasmid and that the bacteria DH5α were therefore transformed with the correct pDNA, a restriction reaction was performed. As previously described by Cirino (2004), the insertion of the CYP BM3 variants in pCWOri is flanked by the restriction sites *BamHI* and

EcoRI **(Figure 8)**. Thus, the pDNA showed in **Figure 7** was submitted to a restriction reaction with the enzymes *BamHI* and *EcoRI* and the reaction products were analyzed by agarose gel electrophoresis **(Figure 9)**.

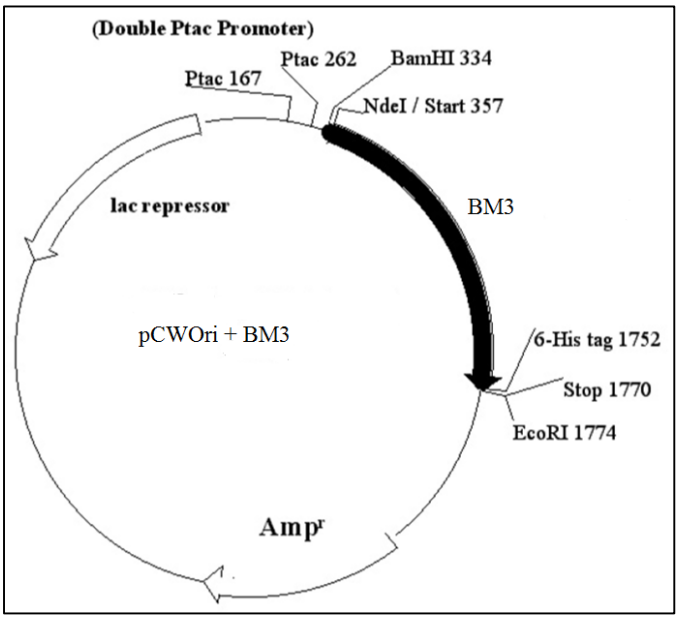

Figure 8. Plasmid map of pCWOri. The CYP BM3 genes are flanked by the restriction sites *BamHI* and *EcoRI*. Adapted from Cirino (2004).

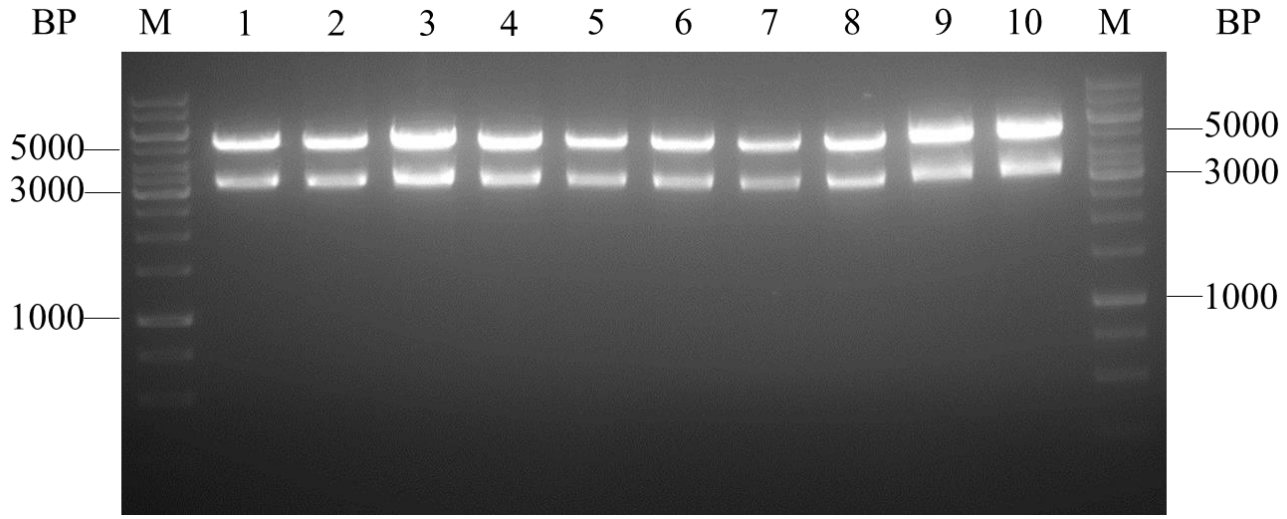

Figure 9. Restriction analysis of the pDNA from *E.coli* DH5α transformed with different BM3 variants. Lanes 1 to 10 correspond respectively to variants 609, 615, 617, 621, 2566, 2567, 2568, 2569, 2570 and 2571.

All pDNAs analysed showed the presence of two fragments: one with approximately 3000 BP likely corresponding to each CYP BM3 (Ruettinger, Wen *et al.* 1989), and another with approximately 5000 BP likely corresponding to the linearized pCWOri. Complementing the information of **Figure 7** with the information of the restriction reaction it appears that the *E. coli* DH5α are properly transformed with the correct constructs from the 10 variants of CYP BM3. The fact that the bands shown in **Figure 7** have approximately 7000 BP, below the predicted 8000 BP, may be explained by the fact that pDNA in a circular form migrate more rapidly than pDNA in a linear form (Hintermann, Fischer *et al.* 1981).

The presence of the two bands with the expected size rules out all possible three hypotheses for negative clones: i) total absence of bands when a clone is not correctly transformed; ii) appearance of two fragments with different sizes of the aforementioned if the clone was transformed with the plasmid harboring an incorrect gene (e.g. due to unexpected recombination); iii) presence of only one fragment corresponding to the linearized plasmid if the clone was transformed with the plasmid without any gene.

3.2.2. Induction of expression and isolation of variants of CYP BM3

After confirmation of the correct transformation of *E. coli* DH5α, work proceeded to the induction of protein expression. The plasmid pCWOri contains two *tac* promoter cassettes, contains a strong *trpA* transcription terminator sequence and the *lacI*^{q} gene encoding the Lac repressor molecule that prevents transcription from the *tac* promoters prior to addition of inducing agents (Barnes, Arlotto *et al.* 1991). The *tac* promotors can be derepressed with isopropyl beta-D-thiogalactoside (IPTG) (De Boer, Comstock *et al.* 1983). This compound was therefore used for induction of the expression of the CYP BM3 variant proteins.

Induction of protein expression was performed at 25°C for 24 hours, through the addition of IPTG and δ-aminolevulinic acid (dALA). dALA is a precursor of heme, a key prosthetic group of CYPs, and its addition aims to achieve a higher concentration of folded active CYP protein (Otey, Bandara *et al.* 2006). Induction of protein expression was performed in triplicate and included a non-induced control (without addition of IPTG and dALA) for each variant of CYP BM3. After the 24 hours induction, samples were collected and the respective protein extracts were analyzed by SDS-PAGE **(Figure 10)**. The overexpression of a protein of

approximately 130 kDa was observed for the variants 2566 to 2571, while variants 609 to 621 were similar to the non-induced control.

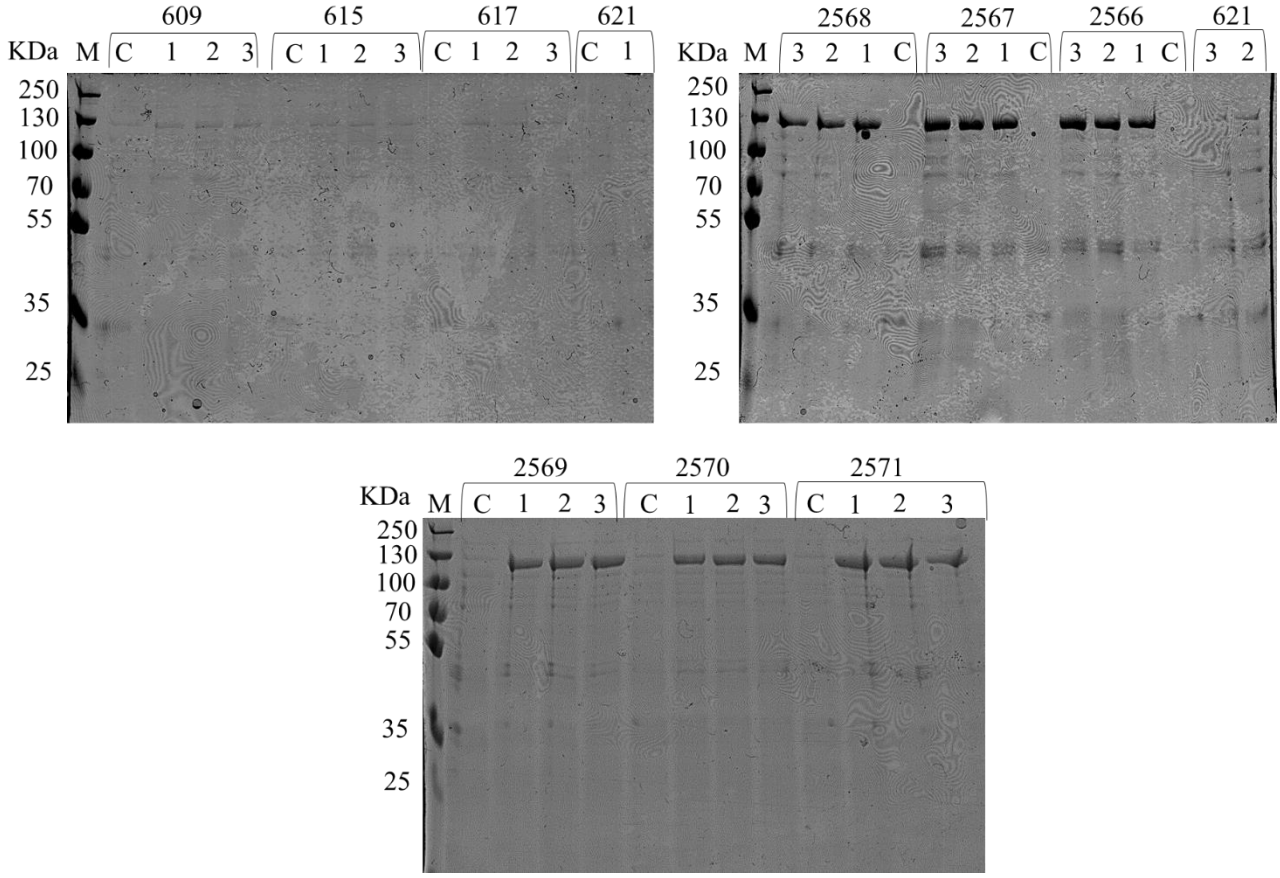

Figure 10. Analysis of the induction of the expression of CYP BM3 variants in DH5α by SDS-PAGE. For each variant, induction of protein expression was done in triplicate (1, 2 and 3) in parallel with a non-induced control (C). Each lane was loaded with10 μL of sample.

The molecular weight of CYP BM3 calculated from its aminoacid sequence is 117,781 Daltons. Narhi and Fulco (1986), who first isolated and characterized CYP BM3, showed that the enzyme was composed by a single polypeptide with a molecular weight estimated to be 119,000 ± 5,000 Daltons. Later, the same authors documented that this polypeptide was composed by two domains, a catalytic domain and a reductase domain (Narhi and Fulco 1987). Therefore, the overexpressed protein identified by SDS-PAGE should correspond to the variants of CYP BM3, since the enzymatic engineering does not alter significantly the molecular weight of the enzyme.

This induction procedure was not effective to express variants 609-621, and hence it has undergone several optimizations. The first optimization involved the modification of the final

concentration of IPTG to 4.5 mM, maintaining the remaining procedure. Another optimization tested included the decrease in temperature to 20 ºC and the incubation of the bacteria 15 min on ice prior to the addition of IPTG, maintaining the final concentration of IPTG in 1 mM. This was done since Weickert, Doherty *et al.* (1996) in their review about optimization of heterologous protein production in *Escherichia coli*, describe that lower temperatures may improve the overexpression in *E.coli*. In both optimization attempts, the variant 2566 of CYP BM3 was also included, to assess if a further improvement in the expression was possible.

As shown in **Figure 11**, the increase in the final concentration of IPTG to 4.5 mM did not improve the expression of the variants 609 to 621 and actually decreased expression for the variant 2566 in comparison with the previous conditions. This slight decrease in protein expression is possibly explained by the fact that IPTG may cause toxicity to the cells as described by Hannig and Makrides (1998). Thus, an increase in the final concentration of IPTG may hinder both cell division as protein synthesis, reflecting in a decrease in expression of the variant 2566.

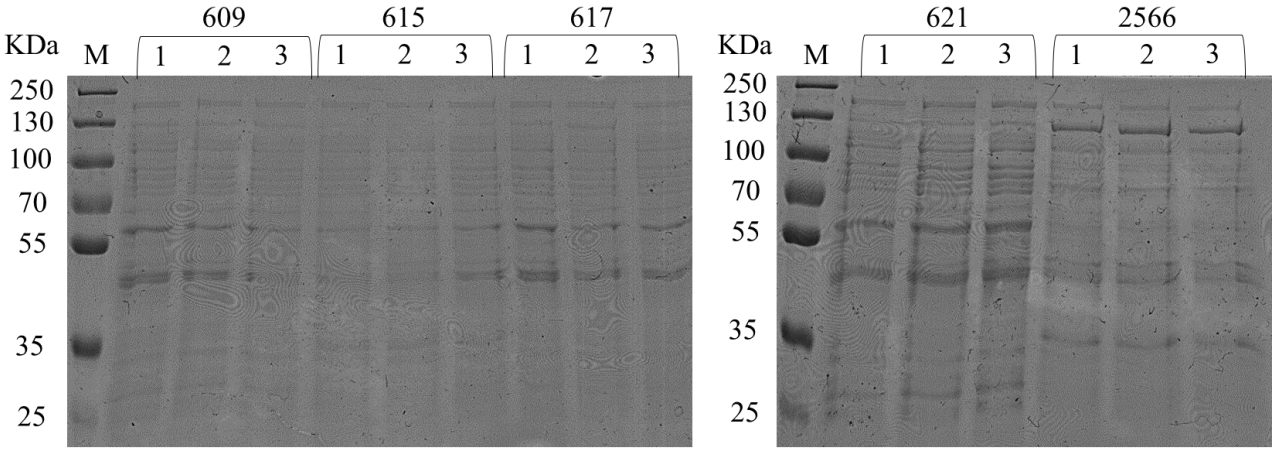

Figure 11. Analysis of the effect of increasing the final concentration of IPTG in the induction of expression of CYP BM3 variants in DH5α by SDS-PAGE. For each variant, induction of protein expression was done in triplicate (1, 2 and 3). Each lane was loaded with10 μL of sample.

The **Figure 12 A** demonstrates that a lower temperature is slightly effective in increasing the expression of the BM3 variants, since it is observable a modest increase of recombinant protein expression at 20 °C compared to the expression at 25 °C, accompanied by a decrease in the expression of secondary proteins. However, **Figure 12 B** reveals high levels of secondary proteins for variants 615, 617 and 2566 at 20 ºC. Li, Anumanthan *et al.* (2007)

refer that lowering the temperature reduces proteolytic degradation in the culture medium, what may lead to interference with the proteolytic degradation of *E.coli* proteins, resulting in a greater appearance of these proteins in the extracts.

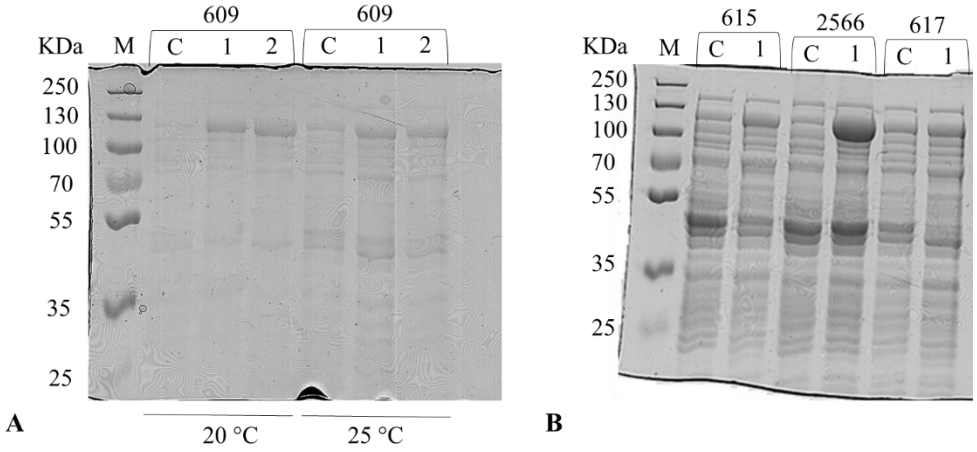

Figure 12. Analysis of the effect of lowering the temperature in the induction of expression of CYP BM3 variants in DH5α by SDS-PAGE. In A, after 15 min on ice and the addition of IPTG, variant 609 was expressed at two different temperatures (20 °C and 25 °C). The experiment was performed in duplicate (1 and 2) together with non-induced controls (C). In B, after 15 min on ice and the addition of IPTG, variants 615, 617 and 2566 were expressed at 20 ºC. The experiment was performed for single samples together with non-induced controls (C). Each lane was loaded with10 µL of sample.

Since it was not possible to obtain a significant amount of recombinant protein for variants 609 to 621, it was decided to proceed with only variants 2566 to 2571. The temperature decrease was shown to reduce expression efficiency for variant 2566, and therefore variants 2566 to 2571 were produced in large scale using the original protocol without any changes.

In addition to the analysis of the expression of CYP BM3 variants, it was also necessary to verify if the expressed variants acquire their native conformation, before proceeding for activity assays. This was done using the CO binding assay, which uses UV/Vis spectroscopy to identify the presence of intact holocytochrome P450. This methodology is based in carbon monoxide's ability to bind to the heme domain of CYP generating a complex that exhibits a characteristic spectrum between 448 and 452 nm, considerably distinct from the usual Soret absorption peaks of hemeproteins. The ability to display this characteristic spectrum is achieved only when the CYP reaches its full maturity and is functional conformation, with its heme domain incorporated correctly. The CO Binding assay not only allows to verify if the enzyme is correctly folded but also allows to calculate the concentration of protein correctly folded (Schenkman and Jansson 2006).

The CO binding assay was performed following the procedure of Omura and Sato (1964) and the concentration of folded enzyme was calculated according with Beer's Law equation as described in chapter 2.7.5. After the previous dilution of the protein extract in phosphate buffer, the sample was reduced, subjected to bubbling with CO and the absorption spectra was recorded **(Figure 13 to Figure 15)**.

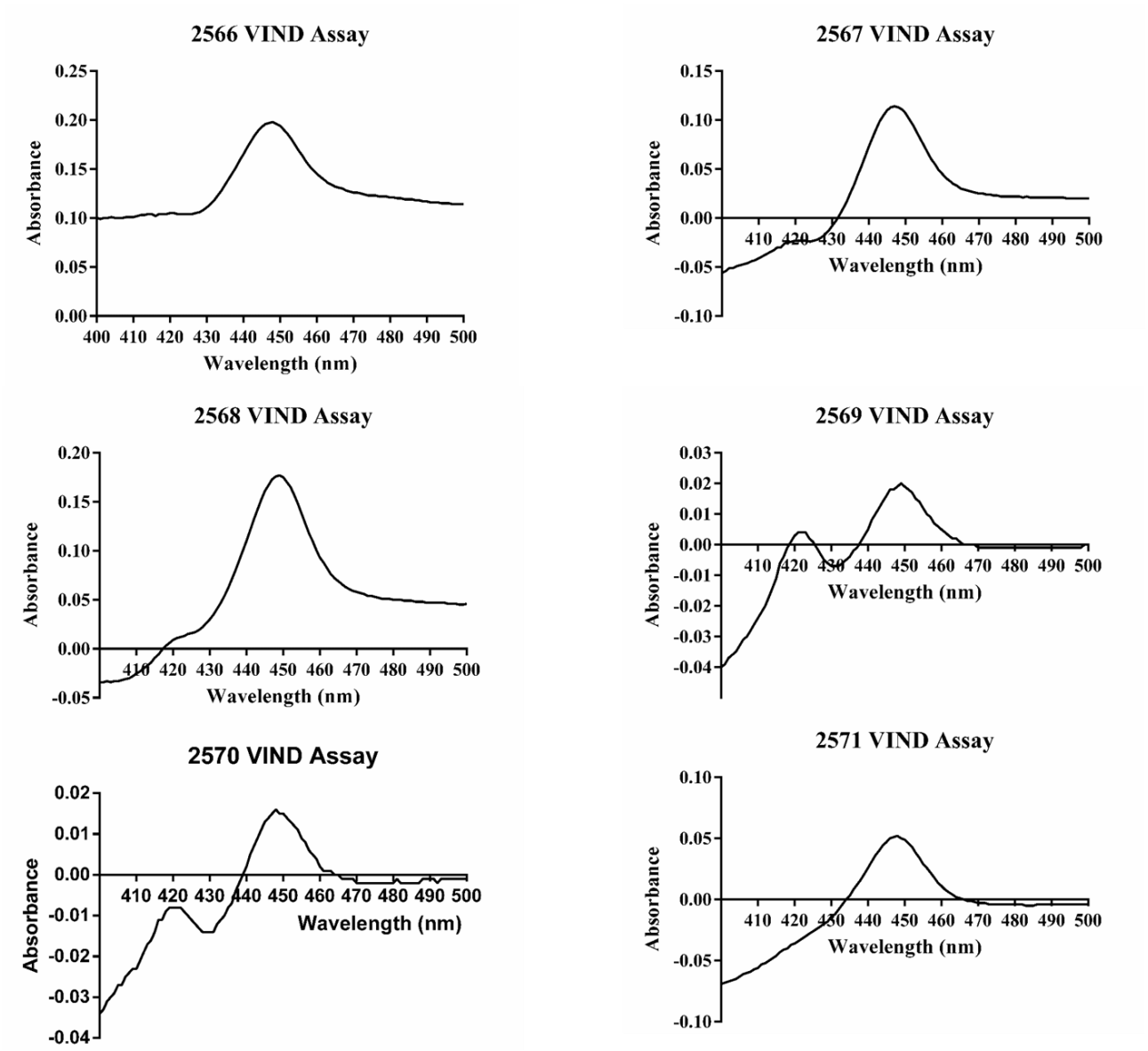

Figure 13. Absorption spectra of CO-CYP BM3 variants later used in the assays with vindoline. The variants shown were chosen to proceed to the following tests. Samples were diluted 5 times.

Figure 13 exhibits the absorption spectra of each CYP BM3 variant used subsequently in the assay with vindoline. The test with the alkaloids will be discussed in more detail in the chapter 3.4. All spectra show a peak between 448 and 452 nm, characteristic of CYPs, and are

similar with the spectra obtained by Omura and Sato (1964) in their studies involving the early characterization of CYPs, as well as with the spectra obtained by van der Hoeven and Coon (1974) in their research on the purification of CYPs from rabbit liver microsomes. In the process of characterization of CYP BM3, Narhi and Fulco (1986) described the absorption spectra of the reduced-CO, observing a peak at 450 nm with a small shoulder at 420 nm. This description is in agreement with all the spectrum shown in **Figure 13** and seemingly these isolated variants exhibit a correct conformation and are suitable for use in the alkaloid assays.

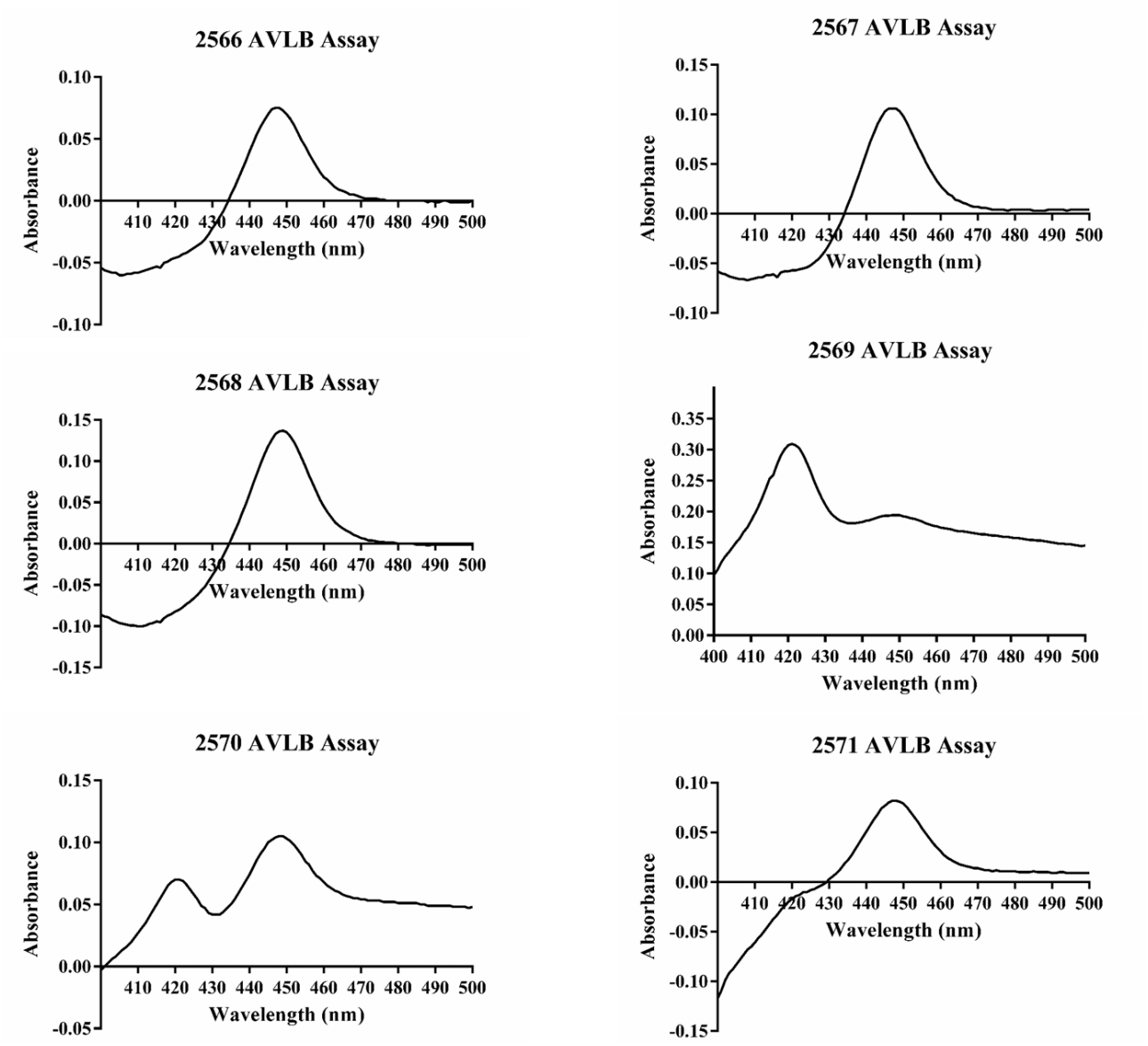

Figure 14. Absorption spectra of CO-CYP BM3 variants later used in the assays with α-3',4'-anhydrovinblastine. The variants shown were chosen to proceed to the following tests. Samples were diluted 5 times.

Figure 14 exhibits the absorption spectra of each CYP BM3 variant subsequently used in the assay with α-3',4'-anhydrovinblastine. Virtually all spectra are in agreement with what was presented and discussed in **Figure 13** as one might expect since the process of induction and isolation of CYPs was exactly the same. However, for variant 2569 it is possible to observe a peak at 420 nm well above the peak at 450 nm. Narhi and Fulco (1986) indicate that this peak at 420 nm is the result of a conformational change or unfolding of CYP BM3. These authors also indicate that the increased peak at 420 nm is accompanied with a decrease of the peak at 450 nm, which explains the reduced size of the peak at 450 nm observed for the 2569 variant. Interestingly, this variant also shows the highest peak at 420 nm in the previous induction experiment shown in **Figure 13**. The peak at 420 nm is also detectable for the 2570 variant in the two figures. This may indicate that due to the enzymatic engineering that both variants had experienced, these two proteins are more sensitive to the process of induction of protein expression and isolation, thus presenting a greater level of deterioration when compared with other variants. Nonetheless, these variants were used for the tests with alkaloids as they all present a peak at 450 nm indicating the presence of correctly folded enzyme.

Figure 15 exhibits the absorption spectra of each CYP BM3 variant which were subsequently used in the assay with tabersonine. In this third round of expression induction, only the variants 2568 and 2571 exhibited the characteristic spectra of functional CYPs. The variant 2569 continues to show a large peak at 420 nm that represents a large proportion of unfolded enzyme in this sample, with almost no detectable peak at ~450 nm. Variant 2567, although not displaying a peak at 420 nm, shows a very small peak at 450 nm. Finally, the spectra of variants 2566 and 2570 are not characteristic of CYPs. As previously reported, the binding of CO to the heme domain is responsible for the characteristic spectrum of CYPs, therefore the fact that these variants do not show a characteristic spectrum may be indicative that either conformation is not correct or induction was not effective and there is no CYP present. Note that the isolation process of the BM3 variants for the test with tabersonine was conducted in Tris-acetate buffer at pH 5 due to solubility requirements of the alkaloid that will be deepened in the chapter 3.4. However, studies conducted with CYP BM3 often use buffers with a higher pH range. Brühlmann, Fourage *et al.* (2014) in their study about the terminal oxidation of palmitic acid by engineered CYP BM3 used phosphate buffer at pH 7, and Lewis, Mantovani *et al.* (2010) used phosphate buffer at pH 8 for the isolation of variants 2566 to 2571 and for the subsequently assays. One of the possible causes not only for the appearance of non-characteristic spectra of CYPs in the variants 2566 and 2570, but also for the apparent loss of

quality of the enzymes displayed by the spectra of remaining variants, may be the interference that low pH (5) has in the process of isolation of the CYPs and in obtaining the CYPs with the correct conformation.

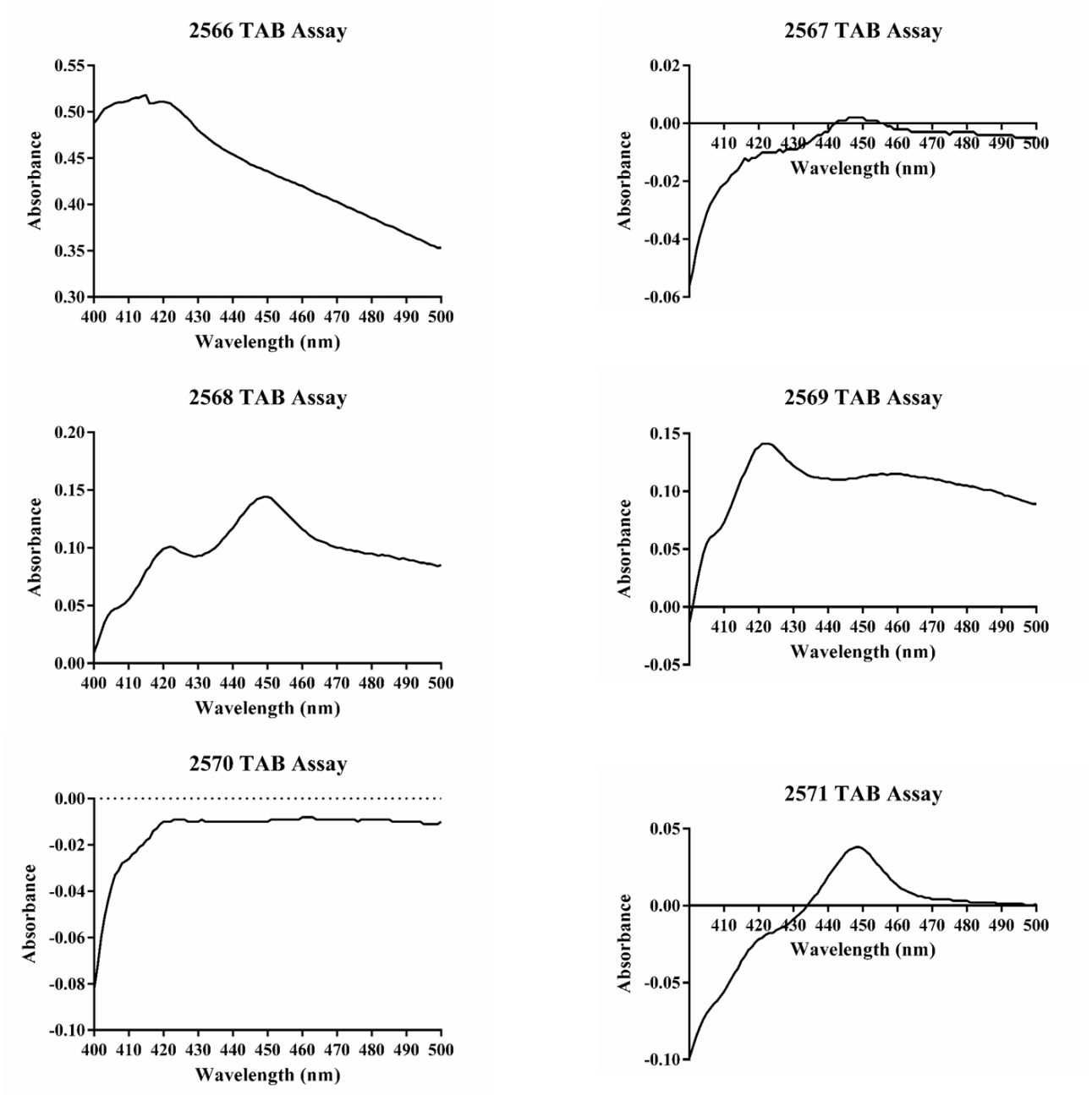

Figure 15. Absorption spectra of CO-CYP BM3 variants later used in the assays with tabersonine. The variants shown were chosen to proceed to the following tests. Samples were diluted 5 times.

3.3. Expression of human CYPs in yeast

3.3.1. Transformation of yeast with the plasmids pYEDP60 harboring the human CYP genes

In order to express the human CYPs it was necessary to transform the *Saccharomyces cerevisiae* strain W(hR) with the plasmids pYEDP60 containing the cloned genes of the human CYPs. In contrast with CYP BM3, which is a self-sufficient monooxygenase, the human CYP genes only encode the catalytic domain, and therefore for its proper functioning it is necessary to co-express a cytochrome P450 reductase (CPR). The yeast strain W(hR) provides the best option since it has the human CPR expression cassette inserted in the yeast chromosome VIII in replacement of the original yeast reductase, reducing incompatibility problems (Pompon, Louerat *et al.* 1996).

Before starting the transformation of yeast with the human CYP constructs, it was necessary to perform the transformation of *E.coli* TOP 10 and the extraction and characterization of pDNA in order to verify if the constructs were in good conditions. This was possible since the pYEDP60 can be replicated in *E. coli* at high copy numbers (Pompon, Louerat *et al.* 1996). The pDNA was analyzed by agarose gel electrophoresis and is represented in **Figure 16**.

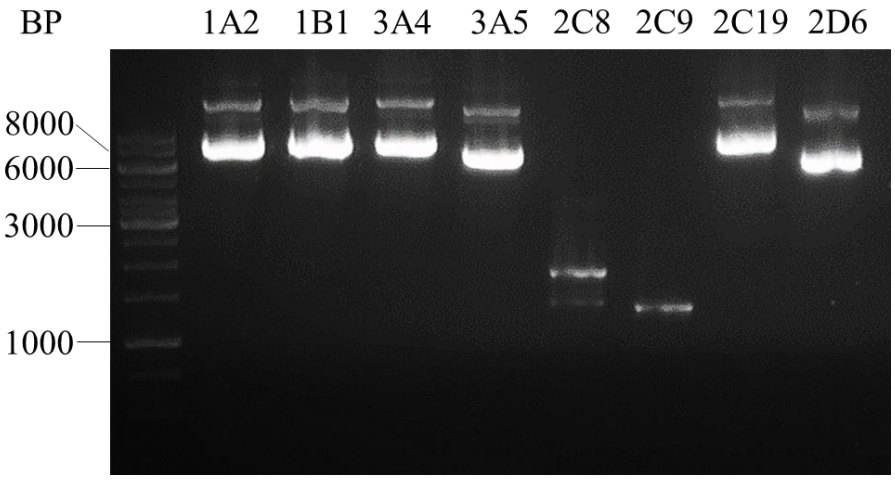

Figure 16. Agarose gel electrophoresis of pDNA extracted from *E.coli* TOP 10 transformed with human CYPS. All human CYPs are represented.

It was observed that the pDNA for CYPs 2C8 and 2C9 did not display the correct migration pattern that could be observed for the remaining CYPs. Transformation of TOP 10 and extraction of pDNA was repeated for these CYPs and the presence of this non-characteristic pattern of migration was observed in all attempts. The cause of this migration pattern may be loss of DNA quality, since the process of transformation and extraction of pDNA was the same for all CYPs. Given that the constructs of CYP 2C8 and 2C9 did not seem to be in proper condition, yeast transformation was not performed with these two CYPs.

After transformation of yeast, colony PCR was performed in order to verify if the competent yeast was correctly transformed with the constructs (pYEDP60 + human CYPs). Colony PCR is a rapid and sensitive method for DNA amplification that allows the selection of yeast transformants (Ward 1992, Mirhendi, Diba *et al.* 2007). For colony PCR, a primer that hybridized in the vector (in the GAL promoter or in the PGK terminator) and a primer that hybridized within the sequence of CYP were used. For each CYP, a PCR using pDNA with the same primers as used as a positive control.

The basic principle of colony PCR is that only if yeasts are effectively transformed with the constructs will primers anneal and the PCR reaction occur, resulting in the appearance of a specific band when the PCR product is analyzed by agarose gel electrophoresis. This band will have to necessarily appear in the positive control. In the case yeast is not correctly transformed with the construct, the primers will not be able to hybridize in appropriate locations and will not lead to the appearance of the expected band. The colony PCR was performed and the PCR product was analyzed by agarose gel electrophoresis **(Figure 17)**. Analyzes of **Figure 17 A** allows to conclude that apparently yeasts have been properly transformed with the constructs harboring the genes of CYP 3A5 and 1A2. The primers designed for these constructs should lead to the formation of fragments with 1500 and 1600 BP respectively. These fragments are represented by the bands marked with the black arrows, and are present in both the colony PCR and the positive control. This allows to conclude that the yeasts were apparently transformed with these constructs. The lanes of 2C19 and 1B1 should present bands with 1500 and 1600 BP respectively. However, colony PCR for these constructs did not result in the formation of the expected fragments and therefore the yeasts were not properly transformed with these constructs. The results of colony PCR for the construct with CYP 3A4 were the ones raising more doubts. Although a band with 1500 BP is observed, this band in the positive control is quite faint, raising the possibility of nonspecific amplification.

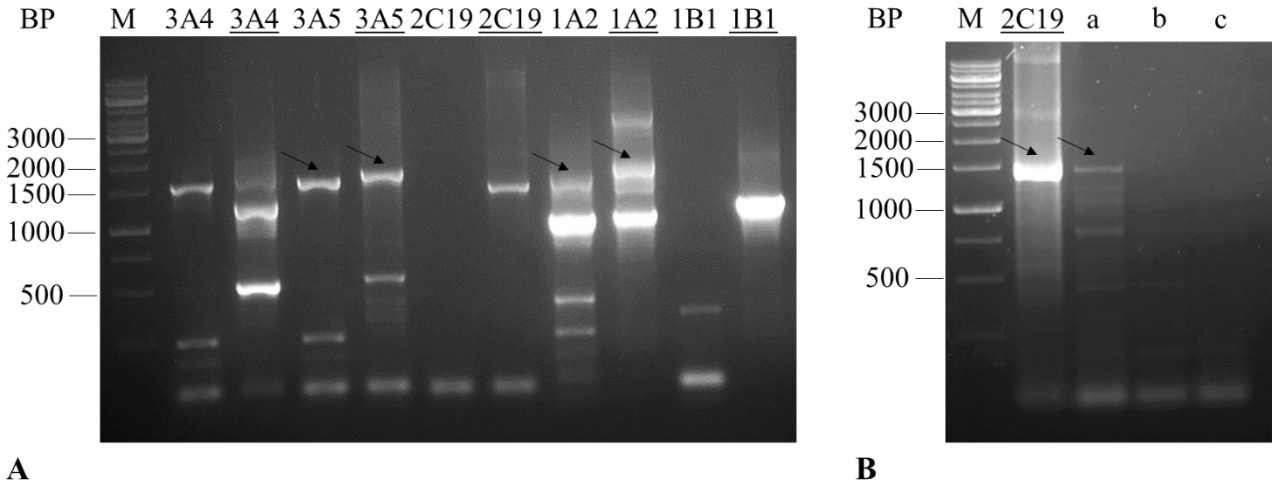

Figure 17. Analysis of the yeast colony PCR product by agarose gel electrophoresis. A presents the analysis of the products of colony PCR for CYPs 3A4, 3A5, 2C19, 1A2 and 1B1. B presents the analysis of the products of colony PCR for CYP 2C19 using three different yeast colonies (a, b and c). Underlined names indicate the positive controls for each CYP. The black arrows indicate the expected bands.

Figure 17 B indicates that apparently yeasts have been properly transformed with the construct harboring the gene of CYP 2C19. This colony PCR was repeated using more than one colony, and it was possible to identify a colony (a) where the yeasts have been properly transformed with the constructs containing the CYP 2C19 gene. Although the band is faint, it's normal since it comes from yeast colony PCR. In yeast colony PCR, burst of yeasts is carried out, which eventually "contaminates" the PCR reaction with various components of yeast, including DNA. This may hinder the PCR reaction and increase non-specific amplifications, which explains the appearance of several bands or faint bands, being therefore very important to perform parallel positive controls.

After design of new primers, yeast transformation and colony PCR were repeated for constructs 3A4 and 1B1, with minor changes in the protocol, but yeast transformation with these constructs was never successful. After transformation of yeast with the construct 2D6, there were no formation of colonies in minimal medium, indicating that the transformation had not occurred. This transformation was also repeated several times without success.

3.3.2. Induction of expression and isolation of human CYPs

Induction of expression was performed only for human CYPs for which transformation of yeast was indeed confirmed, that is for CYP 3A5, 1A2 and 2C19. The plasmid pYEDP60 contains the galactose-inducible GAL10-CYC1 promoter and the PGK terminator, together with two markers (URA3 and ADE2). The markers allow the yeasts which were transformed with the plasmids to survive on minimal medium (without adenine and uracil) (Pompon, Louerat *et al.* 1996). This selection did not allow the appearance of transformants with the construct CYP2D6.

Taking into account the characteristics of the plasmid, the induction of expression is performed through addition of galactose, which also contributes for an higher activity of the human CPR (Pompon, Louerat *et al.* 1996). The induction of protein expression was performed according to a high density procedure which is best adapted for all applications requiring large amounts (up to micromoles) of the recombinant microsomal CYPs at a high specific content. The induction is triggered by the addition of galactose and occurs during 15 hours. After the 15 hours induction, samples were collected for subsequent microsomal isolation.

It was not possible to perform the analysis of microsomal extracts by SDS-PAGE, since the gel only showed an aggregate of proteins (data not shown), therefore the correct expression and conformation of human CYPs was directly analyzed using the CO binding assay. The microsomes had to be diluted 40 to 80 times in TEG to be possible to perform a spectrum measurement. After carrying out the CO binding assay, the absorption spectra was recorded and is presented in **Figure 18**. Even diluting 80 times the sample of CYP 2C19, it was not possible to measure the absorption spectra for this CYP and, therefore, only the absorption spectra of CYPs 1A2 and 3A5 are presented **(Figure 18)**. Analyzing **Figure 18**, it is possible to verify that no spectrum shows the characteristic CYP peak between 448 and 452 nm and they show a large peak around or near 420 nm, indicating the virtual absence of folded CYP enzyme.

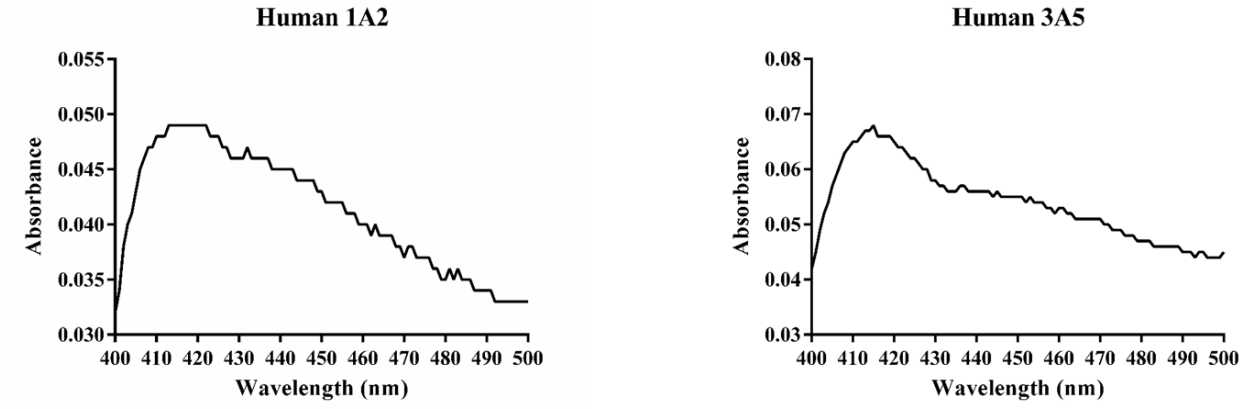

Figure 18. Absorption spectra of human CO-CYPs 1A2 and 3A5. The sample of 1A2 was diluted 80 times and the sample of 3A5 was diluted 40 times.

Saito, Hanioka *et al.* (2005) show that the reduced CO absorption spectra for wild type CYP 1A2 exhibits the normal characteristics of a reduced-CO-CYP absorption spectra. However, in this study, the authors also analyzed naturally occurring CYP 1A2 variants found in a Japanese population, derived from genetic polymorphisms, and differing from wild type by means of single amino acid substitutions. The reduced CO absorption spectra of some of these variants contain a hardly detectable peak at 450 nm and a prominent peak at 420 nm. According with the authors, this finding suggests marked reductions in functional hemeprotein formation, which coincide with the dramatic reductions in the activities of the variant enzymes. Through the evidence discussed by Saito, Hanioka *et al.* (2005), the absorption spectra of CYP 1A2 shown in **Figure 18** suggests that CYP 1A2 does not have the correct conformation and probably will not be properly functional. Problems in the process of induction of CYP expression or isolation may be responsible for the reductions in functional hemeprotein formation.

Sandhu, Guo *et al.* (1994), submitted CYP 1A2 to the action of nonionic detergents (treated CYP 1A2) and recorded the reduced CO absorption spectra of the untreated and treated CYP 1A2. While the untreated CYP exhibited the characteristic spectrum of CYPs, the treated CYP exhibited a prominent peak in the region of 420 nm, which results from the destruction of CYP native conformation caused by non-ionic detergents. Based on the spectrum shown by CYP 1 A2 it may be concluded that this CYP is not in the appropriate conditions to be applied in the following assays.

Lee, van der Heiden *et al.* (2007) studied the activity of a variant of CYP 3A5 in the metabolism of nifedipine, an antihypertensive drug. This variant was found in a white European and results from a single nucleotide polymorphism (SNP) (g.3775A>G) in exon 2, which leads to a Tyr53Cys substitution. The absorption spectra of the wild type CYP displayed the characteristic reduced-CO absorption spectra of CYPs with a major peak at 450 nm and a minor peak at 420 nm. However, the absorption spectra of the variant exhibit a major peak at 420 nm and authors attributed the appearance of this peak to a large amount of denatured CYP, resulting from a probable instability of this enzymatic variant. This instability was corroborated by the low activity of the enzyme in the metabolism of nifedipine. Furthermore, the authors stated that the increased peak at 420 nm and decreased catalytic activity against nifedipine by the variant protein is likely to result from conformational changes. Therefore, the absorption spectra of CYP 3A5 shown in **Figure 18** suggests that there is a large amount of denatured enzyme which is not in its correct conformation. The reasons for this may be the same as those discussed for CYP 1A2 and, therefore, CYP 3A5 could not be used also for further assays. The dilution of 40 and 80 times may also have been excessive, leading to difficulties in the detection of CYPs, but smaller dilutions were tested and were not successful in obtaining a readable absorption spectrum, due to the high concentration of the microsomal extracts.

3.4. Searching for alkaloid conversion activities

3.4.1. Monitoring of NADPH consumption during BM3 reactions in the presence of alkaloid substrates

The CYP BM3 variants 2566, 2567, 2568, 2569, 2570 and 2571 were incubated with vindoline, tabersonine and α-3',4'-anhydrovinblastine during two hours in order to investigate their capacity to accept these alkaloids as substrates. Vindoline was used as a substitute to evaluate the capacity of the CYPs to oxidize vinblastine (VLB), since it is much cheaper and corresponds to the VLB moiety that suffers oxidation during the conversion to vincristine VCR. During the first 45 minutes of reaction the consumption of NADPH was monitored as a means to detect the occurrence of a catalytic reaction. The principle of this procedure lies in the fact the NADPH absorption spectra shows a peak at 340 nm that is absent from $NADP^+$. Therefore, NADPH consumption causes the gradual decrease of the absorbance peak at 340 nm.

This test was performed with a non-induced control (without CYP BM3) incubated with alkaloids, with the CYP BM3 in an alkaloid-free reaction medium and with the CYP BM3 incubated with the alkaloid. Monitoring the NADPH consumption will allow to check if the BM3 reductase activity is functional, and to verify if the presence of the alkaloid leads to an increase of NADPH consumption indicating that it is accepted as substrate by the BM3 variant. The consecutive absorption spectra of NADPH were measured for the different reactions (e.g. **Figure 19 A**) and used to build plots representing the decrease of absorbance at 340 nm over time **(Figure 20 and Figure 21).** However, in the assays with TAB it was not possible to detect the characteristic absorption spectra of NADPH, possibly due to an incompatibility in solution between TAB and NADPH (e.g. **Figure 19 B**) and consequently the monitoring of the consumption of NADPH was not realized for the incubations with TAB.

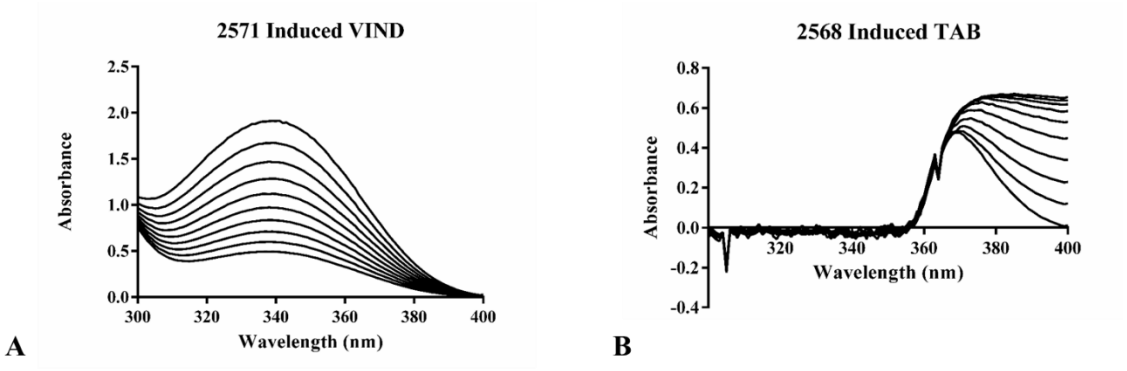

Figure 19. Consecutive absorption spectra recorded during incubation of BM3 variants with alkaloids. A presents the NADPH absorption spectra measured for incubation of variant 2571 with VIND. B presents the NADPH absorption spectra measured for incubation of variant 2568 with TAB.

Figure 20 shows the changes overtime of absorbance at 340 nm representative of the levels of NADPH, for the incubations with VIND. In all the tests it was possible to observe consumption of NADPH translated by decreasing lines. The non-induced control (red lines) in all variants displays a slower decrease of absorbance at 340 nm then the samples with BM3 variant in the absence (black lines) and presence of VIND (green lines), which indicates that it has the lowest rate of NADPH consumption. This low consumption rate was expected since this incubation does not contain CYP and consequently does not contain reductase, which is a good indication that the reductase is functional. The comparison between the incubation of BM3 variants without substrate and in the presence of VIND does not show an increase in NADPH consumption when the alkaloid is present, as would be expected if the alkaloid was a substrate of the CYP, as this should lead to increased CYP activity and consequently to a higher rate of consumption of NADPH. This may mean that VIND is not a substrate for these BM3 variants, or that the extract has compounds that may function either as electron acceptors of the reductase domain of BM3 variants or as substrates. The variants 2566 and 2568 exhibited elevated NADPH consumption rates and thus there was a need to add more NADPH during the incubation period in order to guarantee that the reaction could proceed. The variants 2570 and 2571 also exhibited significant NADPH consumption rates. Interestingly, 2569 is the variant which exhibits a lower NADPH consumption rate. This result corroborates the results of the CO binding assay discussed above, which indicated that a fraction of this CYP BM3 in the protein extracts is not in its proper conformation predicting lower activity.

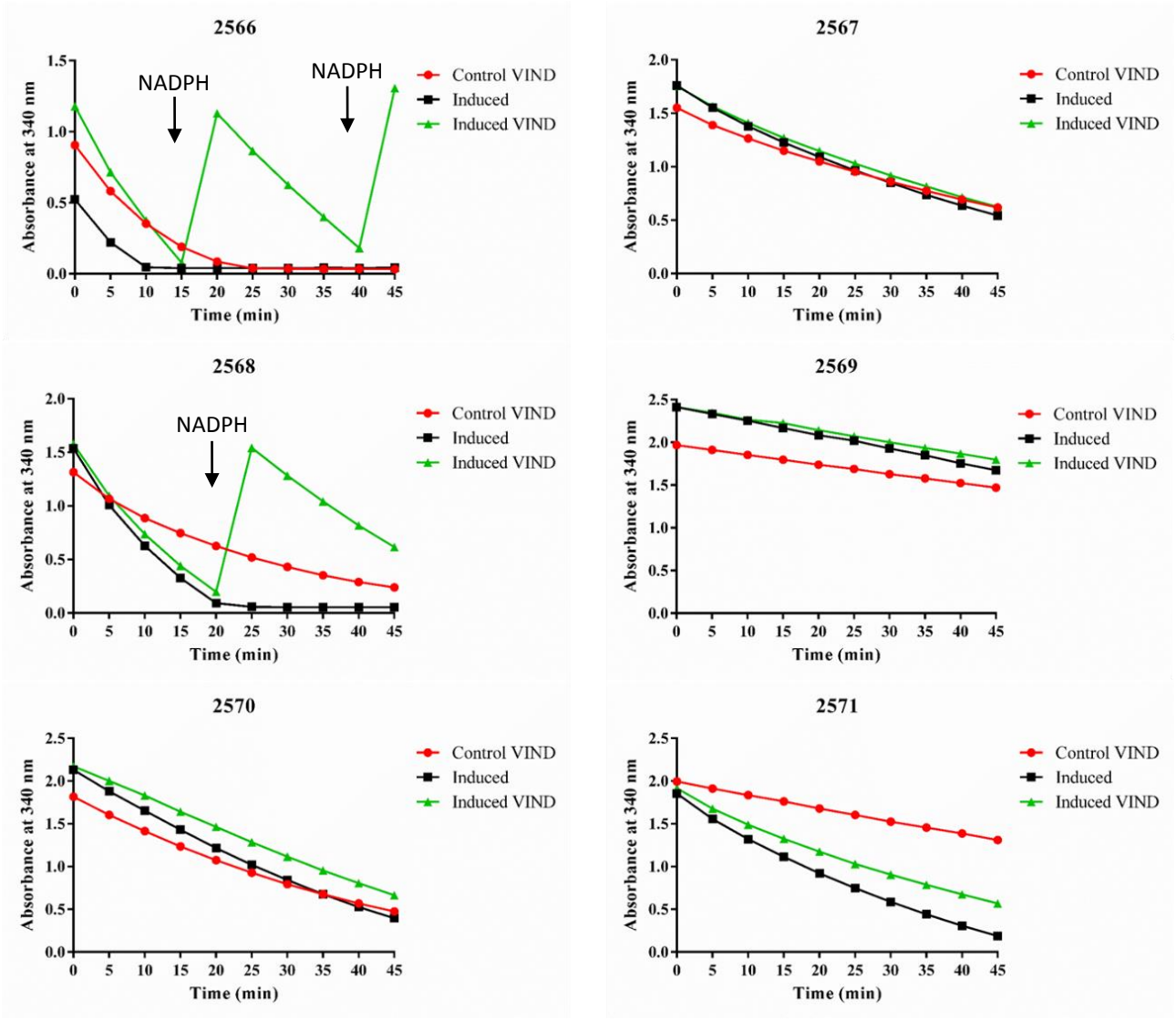

Figure 20. Time-course of changes in absorbance at 340 nm representative of the presence of NADPH during incubation of BM3 variants with vindoline. The variants 2566 to 2571 are represented. For 2566 and 2568 further NADPH was added at 15+40 and 20 min respectively. Control VIND - incubation of non-induced control and vindoline (red line); Induced – incubation of induced samples in the absence of alkaloid (black line); Induced VIND - incubation of induced samples with vindoline (green line).

Figure 21 shows the changes overtime of absorbance at 340 nm representative of the levels of NADPH, for the incubations with AVLB. Although the initial values of NADPH differ a lot, the NADPH consumption rate of non-induced control is always the lowest, as expected due to the facts already discussed above for **Figure 20**. The comparison between the incubation of BM3 variants without substrate and in the presence of AVLB does not show an increase in NADPH consumption when the alkaloid is present, on the contrary, the presence of AVLB seems to somewhat slow down the rate of NADPH consumption. Moreover, in the presence of

AVLB, the consumption rate of NADPH begins slowly at the start of reaction and increases through time, particularly for the variants 2566, 2567 and 2570, contrary to the rate of NADPH consumptions in incubations with VIND, which is practically constant throughout the reaction time that was monitored. The variant 2569 shows again the lowest rate of consumption of NADPH, in accordance with the results of CO binding of this variant, previously discussed.

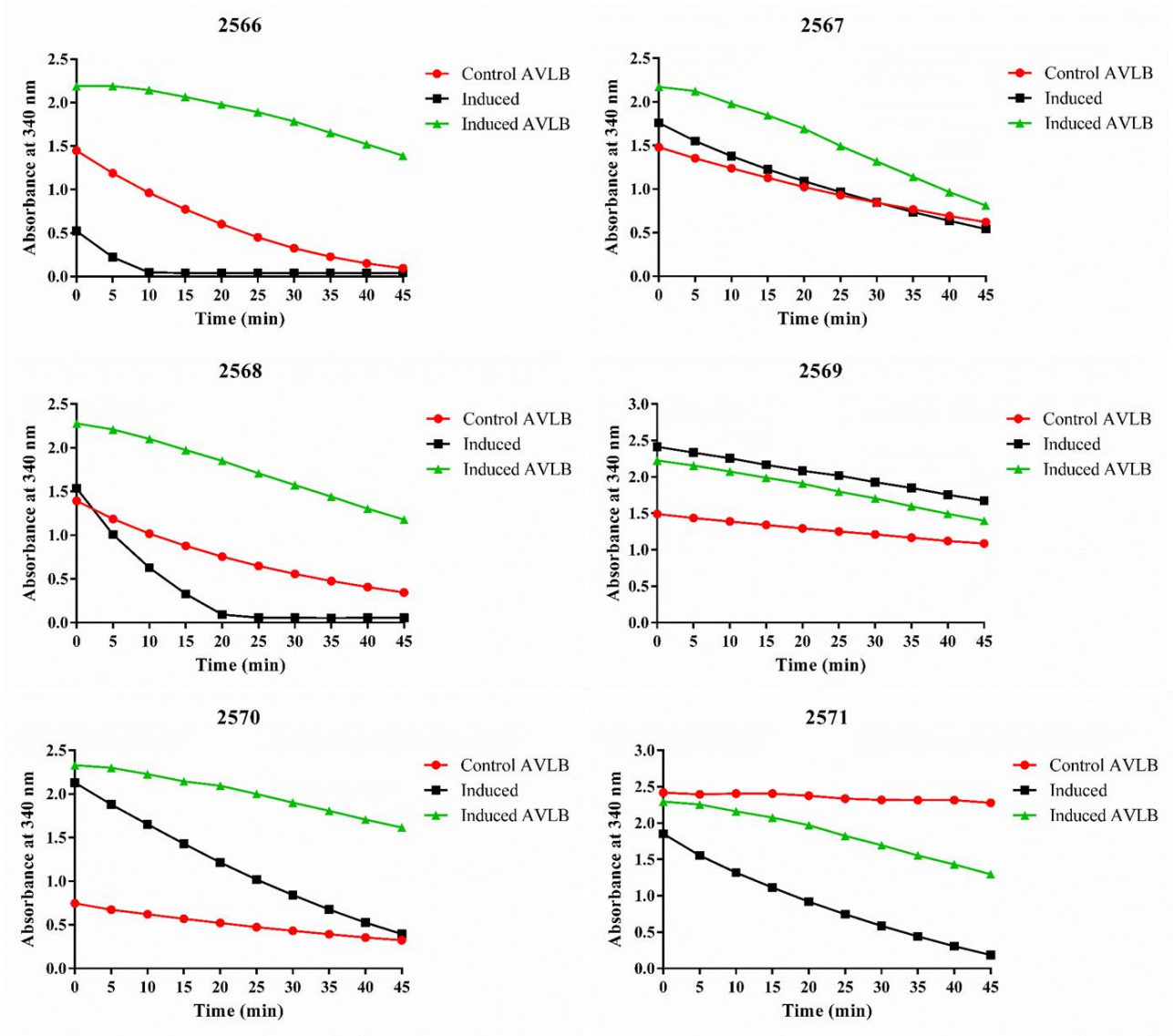

Figure 21. Time-course of changes in absorbance at 340 nm representative of the presence of NADPH during incubation of BM3 variants with α-3',4'-anhydrovinblastine. The variants 2566 to 2571 are represented. Control VIND - incubation of non-induced control and vindoline (red line); Induced – incubation of induced samples in the absence of alkaloid (black line); Induced VIND - incubation of induced samples with vindoline (green line).

Over the years various studies have been conducted in order to confirm which compounds CYP BM3 could metabolize, particularly in the range of fatty acids. Some of these studies were analyzed by Whitehouse, Bell *et al.* (2012) in a review. These studies have linked the consumption of NADPH with the CYP BM3 enzyme activity in incubations with different fatty acids. However, some of these studies, including that conducted by Cryle and De Voss (2008) allowed to verify that NADPH consumption is not fully coupled to product formation when saturated fatty acids are oxidized *in vitro*. It seems that, even in the absence of substrate binding to the monooxygenase domain of BM3 the reductase domain can still function, oxidise NADPH and channel the electrons possibly to a soluble acceptor instead of transferring the electrons to the monooxygenase domain. In this case, it is likely that the presence of a substrate for the monooxygenase will change the conformation of the enzyme to couple the two domains, potentially resulting in slower reduction of NADPH rather than faster, since electron transfer will proceed slower than if directly from the reductase domain to a soluble acceptor. If this is true, the results obtained for AVLB may indicate that AVLB is a substrate for the BM3 variants. Moreover, this would explain the increase of NADPH consumption rate during the reactions in the presence of AVLB, since the consumption of the substrate will decrease its concentration leading to a progressive increase of BM3 enzyme molecules not bound to substrate, which will start reducing NADPH in the faster uncoupled mode.

3.4.2. Alkaloid conversion activity of BM3 variants

After monitoring the consumption of NADPH during 45 min, the reaction was left at 25 ºC until it completed two hours of incubation. Then, the reactions were properly stored and posteriorly prepared in order to be analyzed by High Performance Liquid Chromatography (HPLC) (Tikhomiroff and Jolicoeur 2002). The HPLC data was analyzed and the concentration of each of the alkaloids tested was calculated. The difference between the concentration of added alkaloid and the concentration obtained after incubation corresponds to the amount of alkaloid consumed during the incubations. In **Figure 21** it is possible to observe that for BM3 variants 2566, 2567 and 2568 there was no conversion of VIND. For variants 2569, 2570 and 2571 there was a significant VIND consumption, but only for variant 2571 it was possible to observe a higher activity of VIND conversion relative to the non-induced control. When the conversion rate was calculated as specific activity for BM3 variant 2571 the value obtained was

1.88 kat mol $^{-1}$. The fact that no conversion of VIND occurred in the presence of variants 2566, 2567 and 2568 may be due to the fact that the active site of these enzymes is incompatible with the molecule of VIND, and therefore there is no binding between enzyme and the molecule of VIND. The conversion activity observed by the non-induced controls in assays with the variants 2569, 2570 and 2571 leads to believe that there is some compound or enzyme in these protein extracts that is leading to changes in the molecule of VIND. However, VIND conversion did not change in the presence of variants 2569 and 2570 indicating that they also don't accept VIND as substrate. On the other hand, the activity exhibited in the presence of the variant 2571 is significantly higher than that exhibited by its control, strongly suggesting that it accepts VIND as substrate.

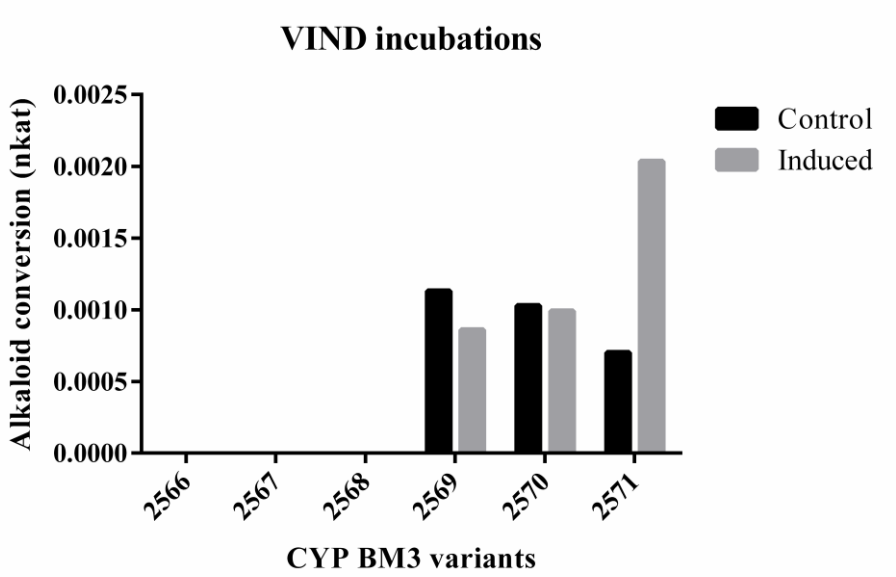

Figure 22. Consumption rates of VIND in the presence of CYP BM3 variants.

It should still be considered that the low activity exhibited by the variants 2569 and 2570 may be due to the low pH at which the reaction was performed. As was described in chapter 3.2.2, the reactions with CYP BM3 are usually performed at higher pH, however, vindoline showed precipitation at pH above 6.2 and thus the reactions had to be performed at this pH, which may not be the most suitable for the proper functioning of CYP BM3, hindering the enzymatic activity of these variants.

There was no TAB conversion activity detected for any of the six variants of CYP BM3. This may result from denaturation of the CYP BM3 variants as a consequence of the low pH at

which these variants were isolated, rather than absence of activity. TAB precipitates at pHs above 5, which forced the incubations and isolations of variants to be performed at pH 5, again potentially interfering with activity.

The entire set of assays for the different BM3 variants was also conducted with AVLB but, unfortunately, when the extracts were run in the HPLC it became evident that the AVLB standard used was actually corrupted, since the peak of AVLB was substituted two peaks neither of which seemed to be AVLB, although one of them was probably a dimer with a close structure judging from the absorption peaks **(Appendix B)**. Due to lack of time, these assays will be have to be repeated outside the timeframe of this study.

3.5. Subcellular localization of CYP BM3 variants in *C. roseus* cells

In order to evaluate the potential of using BM3 variants as enzyme surrogates for *in vivo* TIA biosynthetic reactions it is essential to know where they will localize when expressed in cells of *C. roseus* leaves. This was investigated by fusion protein (FP-tagging) using cyan fluorescent protein (CFP) and transient expression in *C. roseus* mesophyll protoplasts. Both N-terminal fusions using the plasmid pMON999_sCFPstop and C-terminal fusions using the plasmid pMON999_sCFP were generated, to be sure that the position of the fusion did no affect subcellular sorting.

3.5.1. Generation of CYP BM3-CFP N-terminal fusions

The genes of the CYP BM3 variants 617, 2567 and 2569 were randomly selected as representative for subcellular localization studies and were subcloned in the plant expression vector pMON999_sCFPstop, to generate N-terminal fusions with sCFP (a bright CFP).

Figure 23 represents the strategy used to generate the CYP BM3-CFP N-terminal fusions, which can be summarized in 3 steps: 1) amplification of the CYP BM3 variant genes; 2) restriction reaction of amplified fragments and pMON999; 3) ligation of amplified fragments to the digested pMON999_sCFPstop.

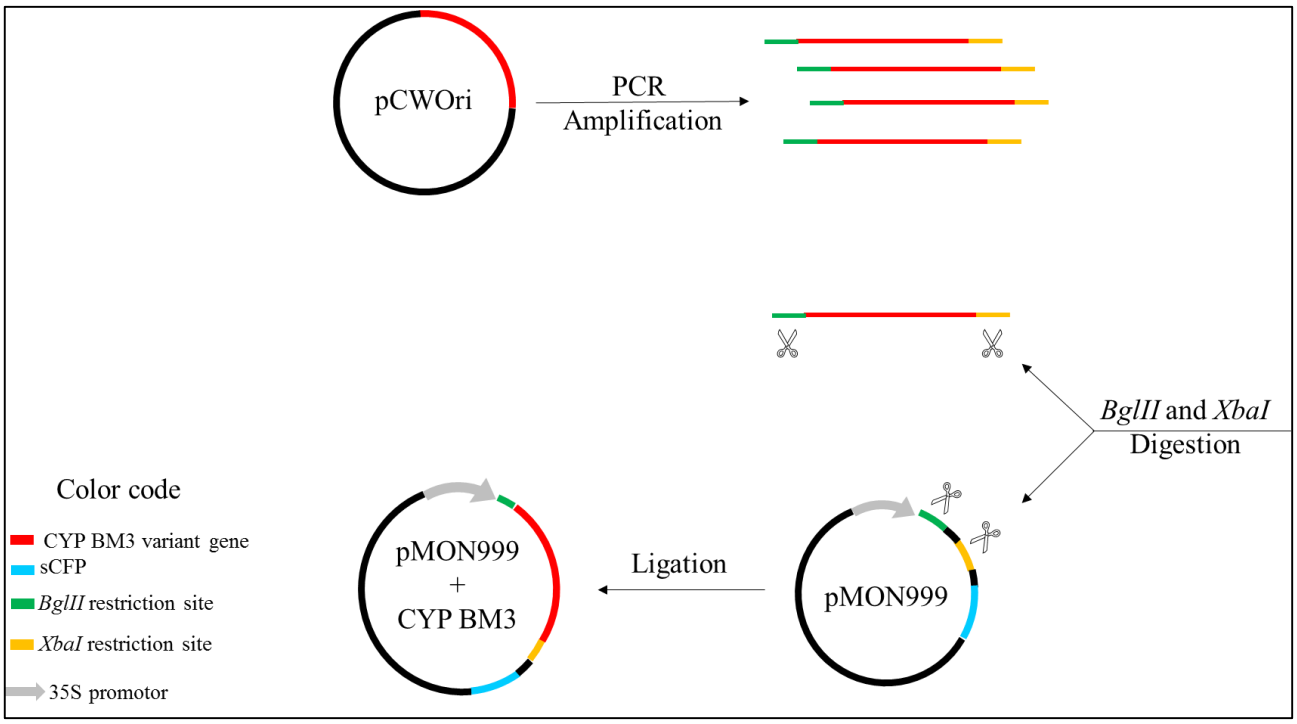

Figure 23. Schematic representation of the subcloning strategy followed to generate CYP BM3-CFP N-terminal fusions. The process starts with the amplification of the BM3 CYP gene of each chosen variant, followed by restriction digestion of the amplified fragments and pMON999-sCFPstop, and terminates with ligation of amplified fragments to the pMON999, to generate the construct *35S::CYP BM3-CFP*. Scissors symbolize the restriction enzymes.

In step 1, the CYP BM3 variant genes were amplified from the plasmid pCWOri by PCR using primers designed to hybridize in the regions flanking the sequence of CYP BM3, thereby allowing the amplification of the complete sequence. The primers further contained at their 5 'ends extensions including the restriction sites of *BglII* and *XbaI*, needed for ligation with correspondent restriction sites in pMON999_sCFPstop to generate the N-terminal fusion. The PCR product was run on agarose gel electrophoresis, the bands corresponding to the amplified fragments were excised from the gel and purified. The purification product was run on agarose gel electrophoresis showing the expected band of 3000 BP corresponding to the CYP BM3 variant genes **(Figure 24)**.

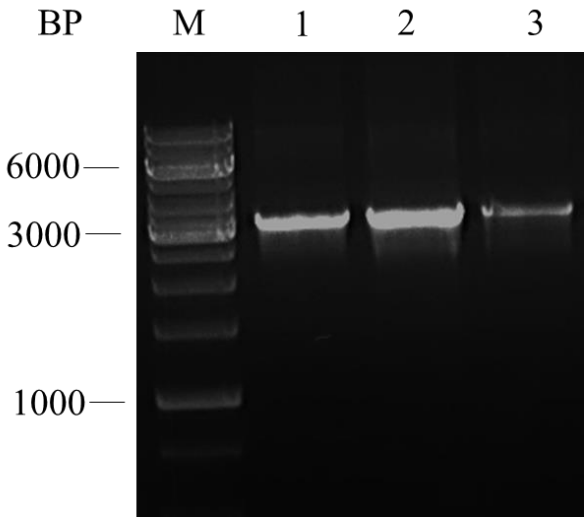

Figure 24. Agarose gel electrophoresis of the PCR-amplification products from CYP BM3 variant genes. 1, 2 and 3 correspond respectively to the variants 2569, 2567 and 617.

In step 2, the purification product shown in **Figure 24** and pMON999_sCFPstop were subjected to restriction digestion with *BglII* and *XbaI*. The restriction reaction linearizes the pMON999_sCFPstop, creating sticky ends in the restriction sites of the respective enzymes. Since the amplified fragment contains restriction sites of the same enzymes at its ends, due to the primers used in step 1, the restriction digestion will form complementary sticky ends in the amplified fragments. The restriction products were run on agarose gel electrophoresis and the bands corresponding to the BM3 fragments and pMON999_sCFPstop were excised from the gel and purified **(Figure 25)**. The linearized plasmid showed a size of ~5000 BP as expected, since the original pMON999 has 4108 BP and the added sequence of CFP described by Goedhart, van Weeren *et al.* (2010) contains 720 BP.

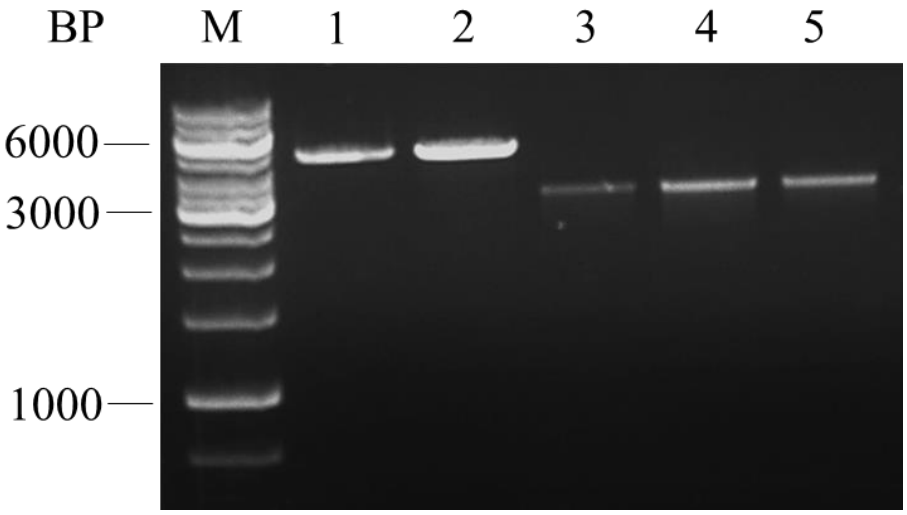

Figure 25. Agarose gel electrophoresis of digested pMON999 and BM3 PCR-amplified products. 1 and 2 correspond to pMON999. 3, 4 and 5 correspond respectively to the variants 2569, 2567 and 617.

In step 3, the amplified fragments were ligated with pMON999_sCFPstop, *E.coli* TOP10 were transformed with the ligation product and colony PCR was performed in order to screen for the presence of positive clones containing the fusion CYP BM3-CFP, more specifically the construct *35S::CYP BM3-CFP*. In the colony PCR M13 primers were used and PCR was also performed using an empty pMON999_sCFPstop as template to act as negative control. The colony PCR products were analyzed by agarose gel electrophoresis **(Figure 26)**.

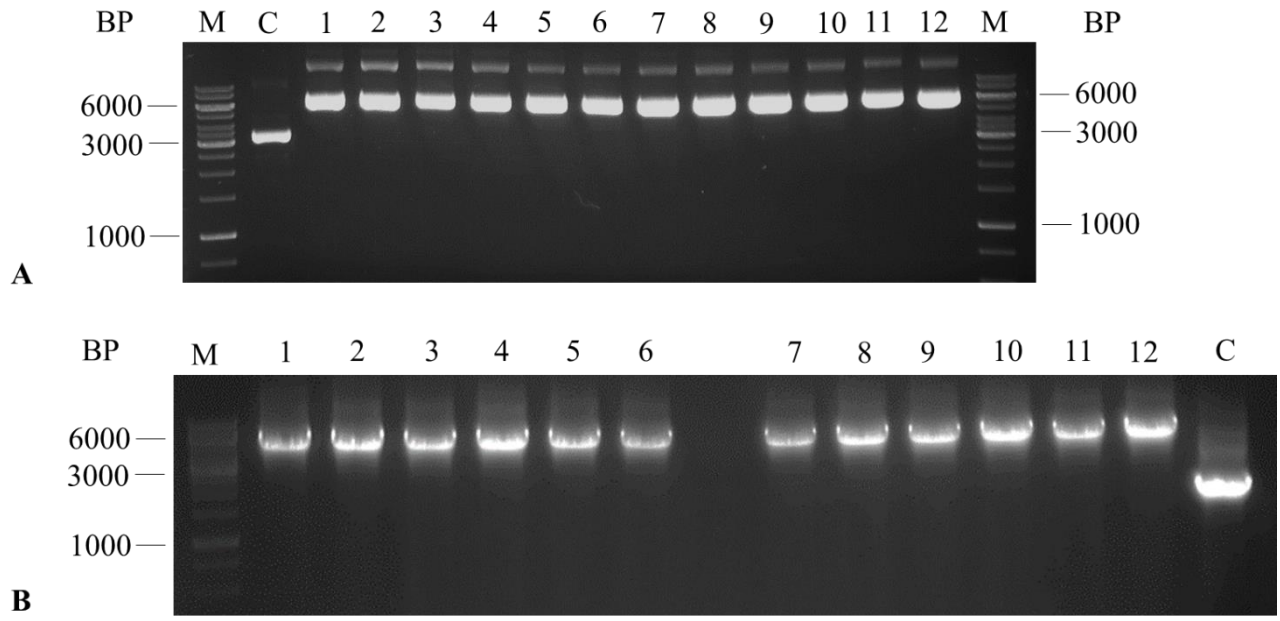

Figure 26. Agarose gel electrophoresis of the colony PCR amplification products for E. coli transformed with BM3 – pMON999 ligations. A presents the results of colony PCR of 12 colonies from TOP 10 transformed with pMON999 containing the CYP BM3 variant 617. B, lanes 1 to 6 present the results of colony PCR of 6 colonies from TOP 10 transformed with pMON999 containing the variant 2567, lanes 7 to 12 present the results of colony PCR of 6 colonies from TOP 10 transformed with pMON999 containing the variant 2569. Lane C in both A and B presents the negative control. All colonies were positive.

Figure 26 shows that the amplification product in all tested colonies had a size larger than the negative control, indicating that apparently all the colonies are positive and therefore contained the construct *35S::CYP BM3-CFP*. The inserted nucleotide sequences were confirmed by sequencing and for each construct, and one error free clone was used for midi-prep purification of plasmid DNA to be used in transient expression assays with *C. roseus* mesophyll cells.

3.5.2. Modification of the plasmid pMON999_sCFP

To generate CFP-CYP BM3 C-terminal fusions using pMON999_sCFP a similar strategy was used to N-terminal fusions adjusting the selection of restriction sites. The first time the subcloning strategy was tried, one of the selected restriction enzymes cut the BM3 fragment in the middle. This indicated that the sequence changes that were introduced in the CYP BM3 variants had formed *de novo* restriction sites not present in the wild type BM3 (the only accessible sequence for *in silico* prediction of restriction sites). Therefore, empirical restriction

tests were performed using the restriction enzymes useful for cloning in pMON999_sCFP, *BamHI, SacI, KpnI,* and *EcoRI,* for all the CYP BM3 variant genes, either using as template PCR amplification products or directly the pCWOri clones. After extensive analysis, it was concluded that *EcoRI* and *BamHI* did not cut the BM3 variant genes and were suitable for cloning. However, this led to the emergence of a problem: since the CFP sequence had been cloned using the plasmid *EcoRI* and *SacI* restriction sites (Appendix A), it became located between *EcoRI* and *BamHI* sites, meaning that using these two enzymes would remove the CFP gene. To overcome this problem it was necessary to modify the pMON999_sCFP, more specifically to change the location of the sequence of CFP in the plasmid. The modification process followed was quite similar to the process of the generation of CYP BM3-CFP N-terminal fusions and its schematic representation is shown in **Figure 27**.

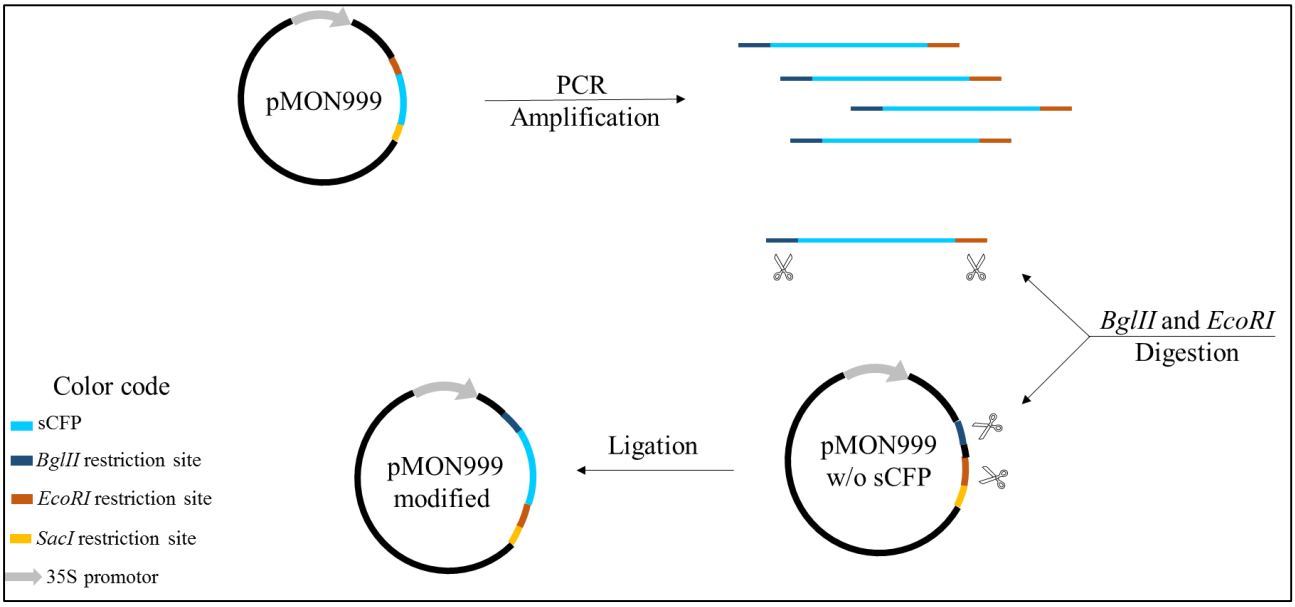

Figure 27. Schematic representation of the modification of pMON999_sCFP. The process starts with the amplification of the CFP gene, followed by restriction digestion of the amplified sCFP and of pMON999 without CFP, and terminates with ligation of the digested amplified *sCFP* with pMON999, to generate the modified pMON999_sCFP. Scissors symbolize the restriction enzymes.

The CFP sequence located between the *EcoRI* and *SacI* restriction sites was removed and was re-introduced between *BglII* and *EcoRI* restriction sites, freeing the downstream *EcoRI* - *BamHI* sites for subcloning of the variant genes in a position leading to the generation of the desired CFP C-terminal fusions.

Figure 28 A reveals the PCR amplification product of *sCFP* as a ~750 BP fragment, in accordant with the 720 BP reported for this gene (Goedhart, van Weeren *et al.* 2010). **Figure**

28 B shows that the restriction reaction led to the linearization of the vector, corresponding to the band shown in lane 2 with ~4000 BP, which is the size of pMON999 without *sCFP*. Lane 1 presents only a fragment with ~750 BP, indicating that the restriction digestion did not lead to cutting within *sCFP*. **Figure 28 C** presents the colony PCR of 9 colonies of TOP 10 transformed with the product of the ligation reaction and by comparison with the negative control it can be concluded that colonies 1 to 8 are all transformed with pMON999 correctly modified. The correct insertion was confirmed by sequencing.

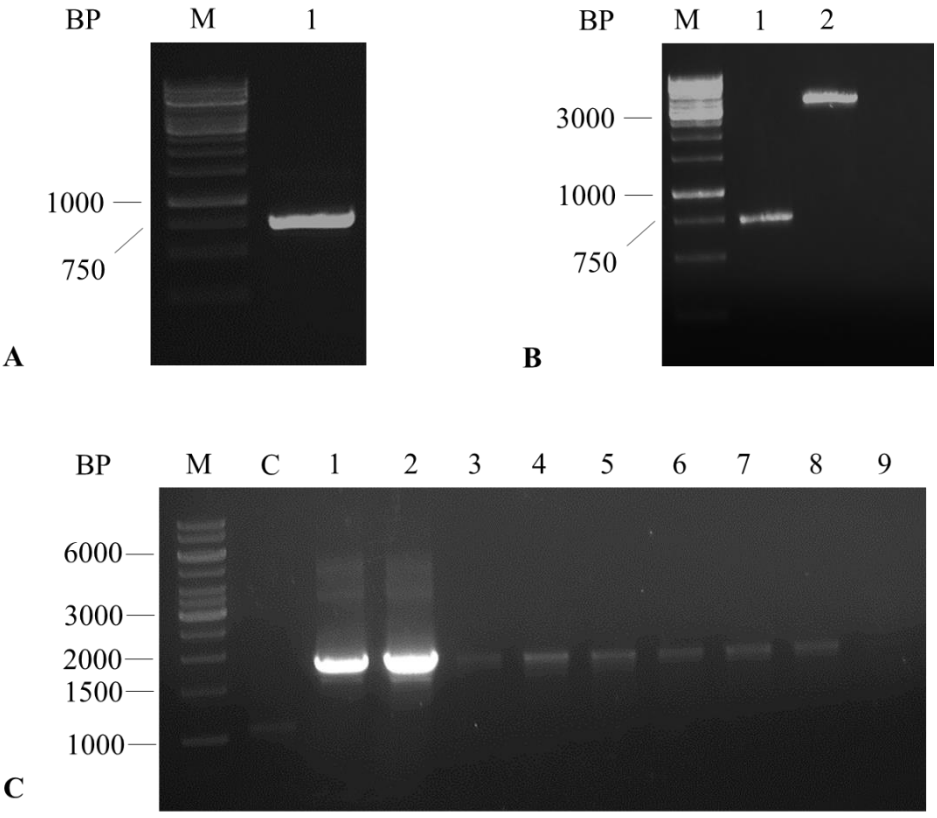

Figure 28. Agarose gel electrophoresis analysis of the DNA elements generated during the modification process of pMON999-sCFP. A presents the purified product from PCR amplification of *sCFP*. B shows the purified products from restriction digestion of *sCFP* (lane 1) and pMON999 (lane2). C presents the colony PCR products for 9 colonies of TOP 10 transformed with ligated pMON999 + *sCFP*. Lane C corresponds to the negative control.

3.5.3. Generation of CFP-CYP BM3 C-terminal fusions

The process followed for the generation of CFP-CYP BM3 C-terminal fusion was exactly the same as the one used for the N-terminal fusions, changing only the restriction enzymes for *EcoRI* and *BamHI*. The process is schematically shown in **Figure 29**.

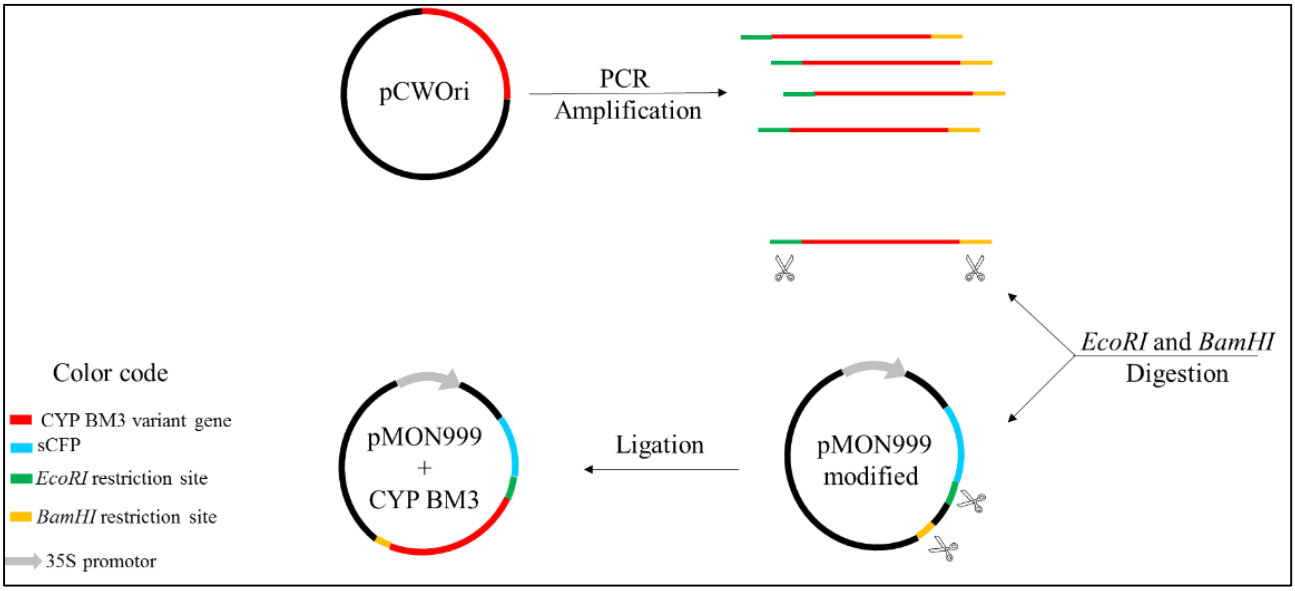

Figure 29. Schematic representation of the subcloning strategy followed to generate CFP-CYP BM3 C-terminal fusions. The process starts with the amplification of the BM3 CYP gene of each chosen variant, followed by restriction digestion of the amplified fragments and the modified pMON999-sCFP, and terminates with ligation of amplified fragments to the pMON999, to generate the construct *35S::CFP-CYP BM3*. Scissors symbolize the restriction enzymes.

Figure 30 A reveled the amplification of a fragment with ~3000 BP for the three variants of CYP BM3 as expected. The **Figure 30 B** shows that restriction digestion led to the linearization of the vector, corresponding to the band shown in lane 4 with ~5000 BP, which is the size of pMON999 with the *sCFP*. The lanes 1 to 3 present only a fragment with ~3000 BP, further conforming that *EcoRI* and *BamHI* indeed do not cut the sequence of these three variants of CYP BM3. Finally, **Figure 30 C** shows that all the colonies are actually transformed with the construct *35S::CFP-CYP BM3*. The inserted nucleotide sequences were confirmed by sequencing and, for each construct, one error free clone was used for midi-prep purification of plasmid DNA to be used for transient expression assays in *C. roseus* cells.

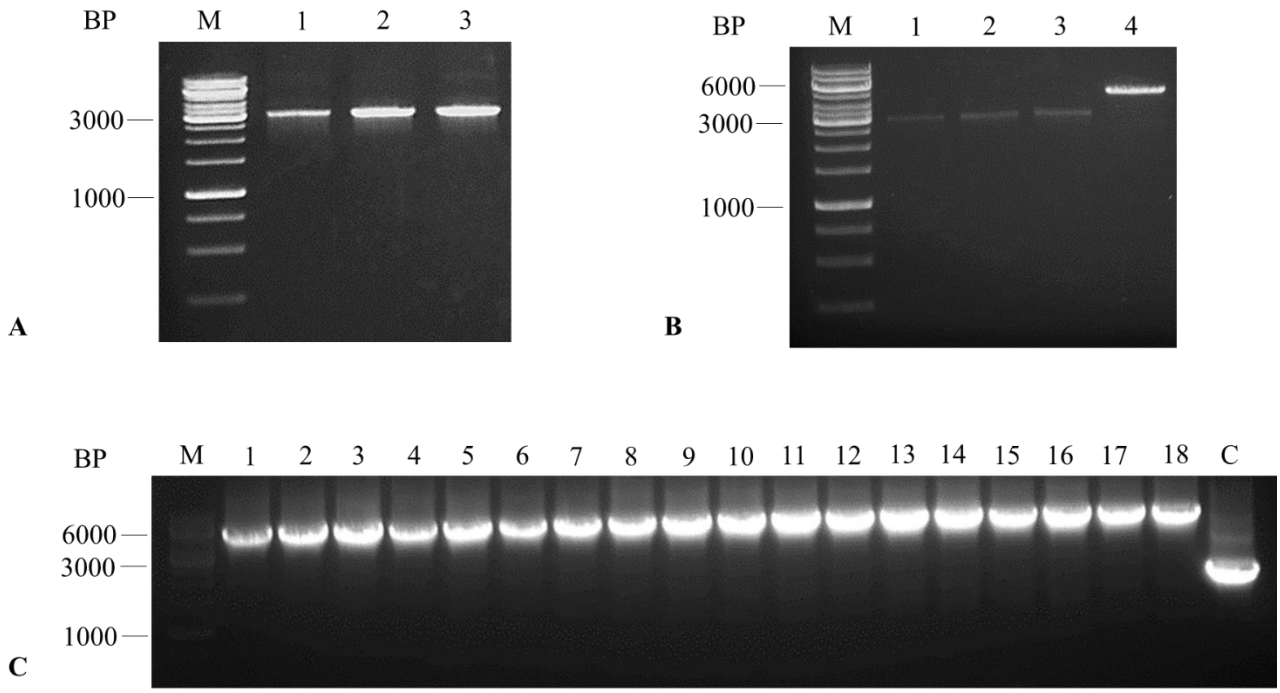

Figure 30. Agarose gel electrophoresis analysis of the DNA elements generated during the several steps followed for the generation of CFP-CYP BM3 C-terminal fusions. A presents the purified product from PCR amplification of the BM3 variant genes: lanes 1, 2 and 3 correspond respectively to variants 2569, 2567 and 617. B shows the purified products from restriction digestion: lanes 1, 2, 3 and 4 correspond respectively to variants 2569, 2567, 617 and to modified pMON999_sCFP. C presents the colony PCR products of 18 colonies of TOP 10 transformed with ligated pMON999-sCFP with the gene variants: colonies 1 to 6 correspond to variant 2569, 7 to 12 correspond to variant 2567 and colonies 13 to 18 correspond to variant 617. Lane C corresponds to the negative control.

3.5.4. Transient expression of CYP BM3-CFP fusions in *C. roseus* leaf protoplasts

The subcellular sorting and final localization of the CYP BM3 variants in *C. roseus* cells was investigated by transformation of *C. roseus* leaf protoplasts with the CYP BM3-CFP fusions generated the precious chapters. As control, the protoplasts were also transformed with empty pMON999_sCFPstop, which leads to expression of CFP (*35S::sCFP*), a cytosolic protein.

The localization of fluorescence was monitored from 24 to 72 hours after transformation by confocal microscopy. The cytosolic CFP appeared after 24 hours and the fluorescence signal started to decline from 48 to 72 hours, without changing the fluorescence location. The images shown below were recorded 48 hours after transformation, corresponding to the best intensity of the fluorescence signal and a large proportion of the protoplasts being transformed and healthy.

Figure 31 shows the presence of a strong blue fluorescence signal homogeneously spread in the cytosol and the nucleoplasm upon transformation of cells with pMON999_sCFPstop, as expected, since CFP is known to be a cytosolic protein.

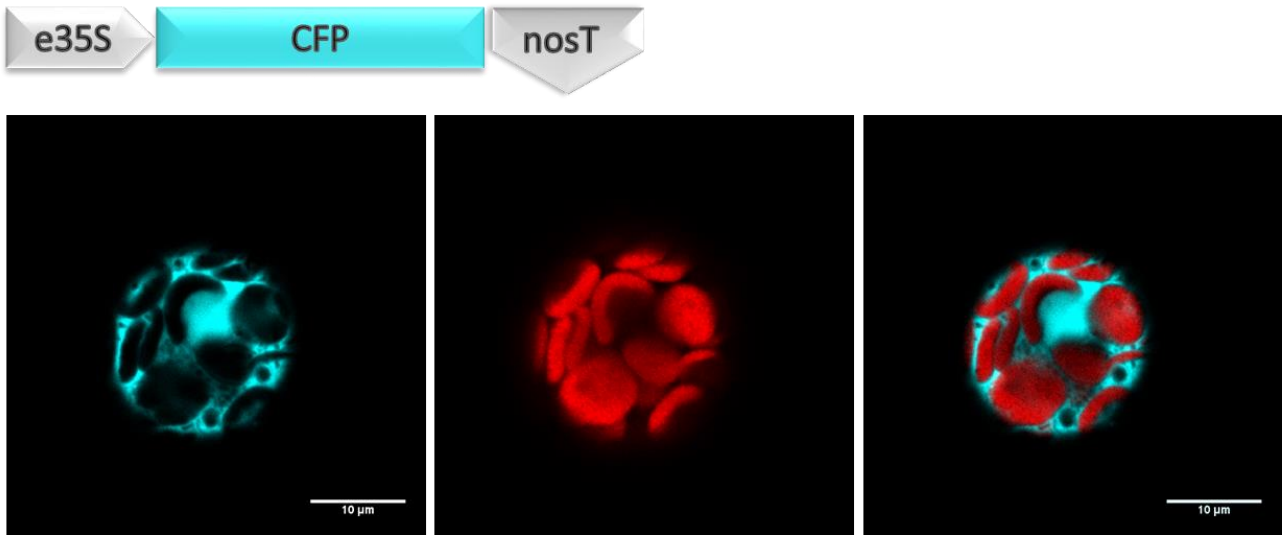

Figure 31. Transient expression of sCFP in *C. roseus* leaf protoplasts. The schematic representation of the construct used for transformation is depicted right above the set of confocal images. Left – CFP channel. Middle – red channel showing chloroplast autofluorescence. Right – merged images. Scale bars, 10 µm.

When the protoplasts were transformed with the CYP BM3-CFP N-terminal fusions **(Figure 32)**, no differences were observed in the fluorescence pattern relative to the expression of sCFP **(Figure 31)**, indicating that all CYP BM3 are sorted to the cytosol in plant cells.

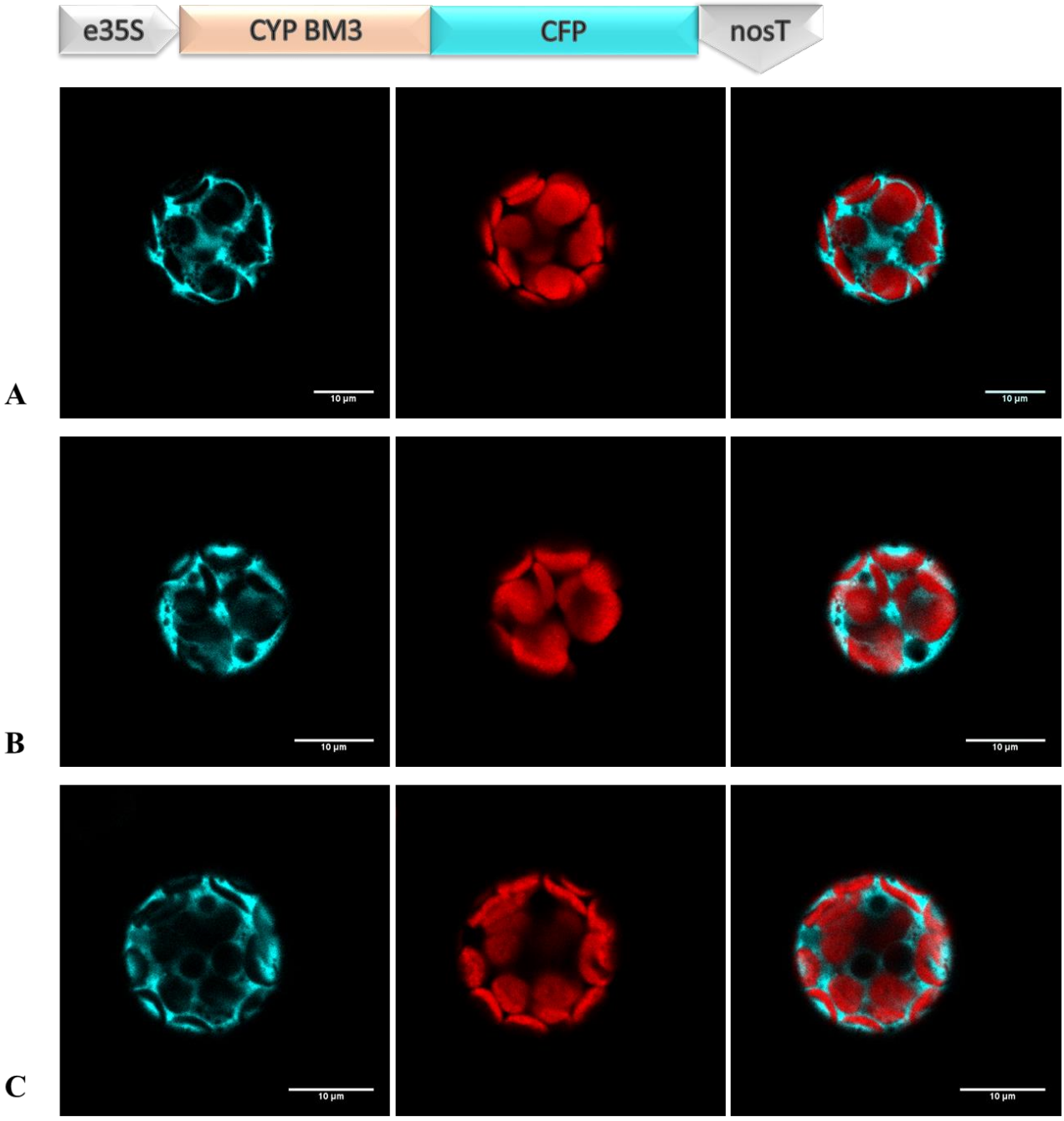

Figure 32. Transient expression of CYP BM3-sCFP N-terminal fusions in *C. roseus* leaf protoplasts. The schematic representation of the construct used for transformation is depicted right above the set of confocal images. A – CFP N-terminal fusion with variant 617, B - CFP N-terminal fusion with variant 2567, C - CFP N-terminal fusion with variant 2569. Left – CFP channel. Middle – red channel showing chloroplast autofluorescence. Right – merged images. Scale bars, 10 µm.

When the protoplasts were transformed with the CFP-CYP BM3 C-terminal fusions (**Figure 33**), no differences were observed again in the observed fluorescence pattern relative to the expression of sCFP (**Figure 31**), reinforcing the premise that the CYP BM3 variants exhibit a cytosolic localization when expressed in plant cells. These results are in agreement with the description that CYP BM3 is a soluble protein in its organism of origin, *Bacillus megaterium* (Narhi and Fulco 1986, Boddupalli, Pramanik *et al.* 1992).

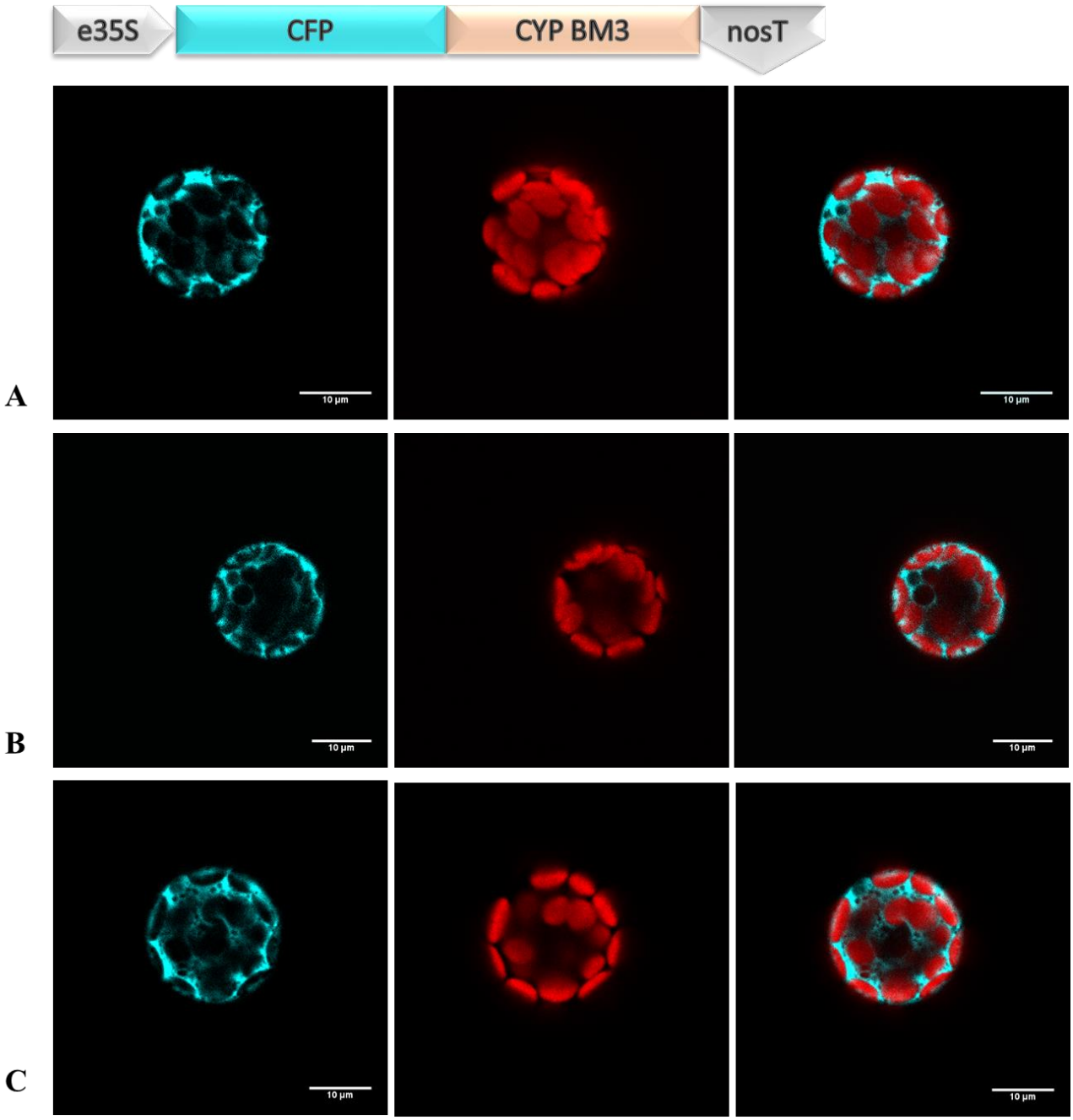

Figure 33. Transient expression of sCFP-CYP BM3 C-terminal fusions in *C. roseus* leaf protoplasts. The schematic representation of the general construct used for transformation is depicted right above the set of confocal images. A – CFP N-terminal fusion with variant 617, B - CFP N-terminal fusion with variant 2567, C - CFP N-terminal fusion with variant 2569. Left – CFP channel. Middle – red channel showing chloroplast autofluorescence. Right – merged images. Scale bars, 10 µm.

The identification of the location shown by CYP BM3 in the cell of *C. roseus* is of special importance. As previously introduced, the different TIAs exhibit a compartmentalized subcellular localization, with the substrates of interest being likely formed in the cytosol in the case of tabersonine, and possibly in the vacuole in the case of AVLB and VLB (Costa *et al.* 2008, Carqueijeiro *et al.* 2013). Therefore, for direct metabolic engineering of the TIA biosynthetic pathway using CYP BM3, the target reaction should by cytosolic. If intended that the CYP accesses to TIAs with other subcellular locations such as the vacuole, there will be the need to add to the sequence of the CYP BM3 the relevant sorting signal sequences.

4. CONCLUSION AND FUTURE DIRECTIONS

In this project, we have expressed in *E. coli* and successfully isolated ten variants of CYP BM3. Six of these variants were properly folded and were used to assay activity against critical TIA substrates from the biosynthetic pathway of the low level anticancer alkaloids from *C. roseus*, vinblastine and vincristine. The successful expression of the remaining variants was not achieved with the desired efficiency, and therefore its protocol requires further optimization. The success in producing recombinant CYPs is not only important for this work but will also be extremely useful in any future work involving tests with CYPs.

The six variants of CYP BM3 successfully isolated were incubated with vindoline, tabersonine and α-3',4'-anhydrovinblastine with the aim to identify variants which exhibit the ability to convert any of these alkaloids, with vindoline being used as a surrogate of VLB. Before each test, the structural integrity of the variants was evaluated, demonstrating satisfying results when the protein extracts were prepared at pH 6.2. This result strengthens the fact that these CYPs were expressed and isolated correctly. Nevertheless, when the protein extracts were prepared at pH 5, as necessary for assays with tabersonine, the CYPs showed a non-correct conformation. The test with tabersonine thus requires optimization. On the other hand, since one of the locations of the targeted reactions is the acidic vacuole, it would be advantageous to further evolve the BM3 variants to become correctly folded at acid pH.

The preliminary results from the incubation assays showed that the variant 2571 may be promising in the conversion of vindoline. However, the incubation assays also require further optimizations, and additional studies are needed to understand the exact nature of the catalyzed reaction, and if and how this variant could be applied in the engineering of the TIA biosynthetic pathway. One of the possibilities is that the variant may mimic one of the *in planta* reactions but presenting higher efficiency as is typical of BM3, and would therefore be a candidate for replacement of the endogenous enzyme. The other possibility is that new derivatives are produced, which should be evaluated for novel or improved medicinal activities.

The subcellular localization of CYP BM3 variants in *C. roseus* cells revealed that this CYP localizes in the cytosol of plant cells. It may become important in the future to transform *C. roseus* cell cultures or plants with CYP BM3 variants in order to engineer the TIA pathway for improved efficiency or new products. To achieve that goal it is important to have the previous knowledge concerning the subcellular sorting of the BM3 proteins in plant cells. This will enable to design the correct recombinant strategies to direct the CYP BM3 for the target subcellular location, in order to perform the desired reaction.

In this study we have also expressed human CYPs in yeast cells. However, the transformation of yeast with the genes of human CYPs was not totally effective, and was only achieved for two CYPs. In addition, when the structural integrity of these CYPs was analyzed, a high degree of denaturation was observed, and therefore it was not possible to proceed to incubation with the alkaloids. However, due to the capacity of certain human CYPs to metabolize alkaloids (vinblastine and vincristine), it is important to proceed to the proper isolation of these CYPs and to carry out tests with tabersonine, vindoline and α-3',4'-anhydrovinblastine.

In order to investigate if any of the heterologous CYP leads to the formation of novel conversion products of TIAs, it is necessary to develop large scale reactions, enabling the detection of the products by HPLC, and/or to use the more sensitive MS instead of UV detection coupled to the HPLC. These should also be made using purified CYP BM3 variants, in order to avoid the use of large amounts of protein extracts, since they come from bacterial lysates and may introduce adverse contaminations in the reaction media. The products formed should be further characterized and may have interesting and novel properties. The CYP BM3 variants could then be applied as biocatalysts for the production of these new compounds.

It would also be interesting to carry out incubations with other CYPs, in view of the large number of variants that have been created via enzymatic engineering. This analysis could lead to the identification of CYPs that have higher conversion rates of these compounds. Above all, the search for an efficient CYP variant capable of performing the bottleneck conversion of AVLB into VLB should proceed.

5. REFERENCES

Almagro, L., F. Fernández-Pérez and M. A. Pedreño (2015). "Indole Alkaloids from *Catharanthus roseus*: Bioproduction and Their Effect on Human Health." Molecules **20**(2): 2973-3000.

Almagro, L., J. Gutierrez, M. A. Pedreño and M. Sottomayor (2014). "Synergistic and additive influence of cyclodextrins and methyl jasmonate on the expression of the terpenoid indole alkaloid pathway genes and metabolites in *Catharanthus roseus* cell cultures." Plant Cell, Tissue and Organ Culture (PCTOC) **119**(3): 543-551.

Aoyama, T., S. Yamano, D. Waxman, D. Lapenson, U. Meyer, V. Fischer, R. Tyndale, T. Inaba, W. Kalow and H. Gelboin (1989). "Cytochrome P-450 hPCN3, a novel cytochrome P-450 IIIA gene product that is differentially expressed in adult human liver. cDNA and deduced amino acid sequence and distinct specificities of cDNA-expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and cyclosporine." Journal of Biological Chemistry **264**(18): 10388-10395.

Arnold, F. H. (2009). How proteins adapt: lessons from directed evolution. Cold Spring Harbor symposia on quantitative biology, Cold Spring Harbor Laboratory Press.

Barnes, H. J., M. P. Arlotto and M. R. Waterman (1991). "Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*." Proceedings of the National Academy of Sciences **88**(13): 5597-5601.

Bennett, R. N. and R. M. Wallsgrove (1994). "Tansley Review No. 72. Secondary metabolites in plant defence mechanisms." New Phytologist: 617-633.

Bernhardt, R. and V. B. Urlacher (2014). "Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations." Applied microbiology and biotechnology **98**(14): 6185-6203.

Besseau, S., F. Kellner, A. Lanoue, A. M. Thamm, V. Salim, B. Schneider, F. Geu-Flores, R. Höfer, G. Guirimand and A. Guihur (2013). "A pair of tabersonine 16-hydroxylases initiates the synthesis of vindoline in an organ-dependent manner in *Catharanthus roseus*." Plant physiology **163**(4): 1792-1803.

Boddupalli, S. S., R. W. Estabrook and J. A. Peterson (1990). "Fatty acid monooxygenation by cytochrome P-450BM-3." Journal of Biological Chemistry **265**(8): 4233-4239.

Boddupalli, S. S., B. C. Pramanik, C. A. Slaughter, R. W. Estabrook and J. A. Peterson (1992). "Fatty acid monooxygenation by P450 BM-3: product identification and proposed mechanisms for the sequential hydroxylation reactions." Archives of biochemistry and biophysics **292**(1): 20-28.

Brown, S., M. Clastre, V. Courdavault and S. E. O'Connor (2015). "De novo production of the plant-derived alkaloid strictosidine in yeast." Proceedings of the National Academy of Sciences **112**(11): 3205-3210.

Brühlmann, F., L. Fourage, C. Ullmann, O. P. Haefliger, N. Jeckelmann, C. Dubois and D. Wahler (2014). "Engineering cytochrome P450 BM3 of *Bacillus megaterium* for terminal oxidation of palmitic acid." Journal of biotechnology **184**: 17-26.

Carqueijeiro, I., H. Noronha, P. Duarte, H. Gerós and M. Sottomayor (2013). "Vacuolar transport of the medicinal alkaloids from *Catharanthus roseus* is mediated by a proton-driven antiport." Plant physiology **162**(3): 1486-1496.

Chen, W., M.-K. Lee, C. Jefcoate, S.-C. Kim, F. Chen and J.-H. Yu (2014). "Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin." Genome biology and evolution **6**(7): 1620-1634.

Choudhary, D., I. Jansson, I. Stoilov, M. Sarfarazi and J. B. Schenkman (2004). "Metabolism of retinoids and arachidonic acid by human and mouse cytochrome P450 1b1." Drug Metabolism and Disposition **32**(8): 840-847.

Cirino, P. C. (2004). Laboratory evolution of cytochrome P450 peroxygenase activity (Doctoral dissertation), California Institute of Technology.

Collu, G., N. Unver, A. M. Peltenburg-Looman, R. van der Heijden, R. Verpoorte and J. Memelink (2001). "Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis." Federation of European Biochemical Societies letters **508**(2): 215-220.

Conney, A., E. C. Miller and J. Miller (1956). "The metabolism of methylated aminoazo dyes V. Evidence for induction of enzyme synthesis in the rat by 3-methylcholanthrene." Cancer Research **16**(5): 450-459.

Cook Sangar, M., H. K. Anandatheerthavarada, W. Tang, S. K. Prabu, M. V. Martin, M. Dostalek, F. P. Guengerich and N. G. Avadhani (2009). "Human liver mitochondrial cytochrome P450 2D6–individual variations and implications in drug metabolism." Federation of European Biochemical Societies journal **276**(13): 3440-3453.

Costa, M. M. R., F. Hilliou, P. Duarte, L. G. Pereira, I. Almeida, M. Leech, J. Memelink, A. R. Barceló and M. Sottomayor (2008). "Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*." Plant Physiology **146**(2): 403-417.

Courdavault, V., N. Papon, M. Clastre, N. Giglioli-Guivarc'h, B. St-Pierre and V. Burlat (2014). "A look inside an alkaloid multisite plant: the *Catharanthus* logistics." Current opinion in plant biology **19**: 43-50.

Cryle, M. J. and J. J. De Voss (2008). "The Role of the Conserved Threonine in P450BM3 Oxygen Activation: Substrate-Determined Hydroxylation Activity of the Thr268Ala Mutant." ChemBioChem **9**(2): 261-266.

Danielson, P. (2002). "The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans." Current drug metabolism **3**(6): 561-597.

de Bernonville, T. D., M. Clastre, S. Besseau, A. Oudin, V. Burlat, G. Glévarec, A. Lanoue, N. Papon, N. Giglioli-Guivarc'h and B. St-Pierre (2014). "Phytochemical genomics of the Madagascar periwinkle: Unravelling the last twists of the alkaloid engine." Phytochemistry **113**: 9-23.

De Boer, H., L. Comstock and M. Vasser (1983). "The tac promoter: a functional hybrid derived from the trp and lac promoters." Proceedings of the National Academy of Sciences **80**(1): 21-25.

De Luca, V. and A. J. Cutler (1987). "Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*." Plant physiology **85**(4): 1099-1102.

De Luca, V., C. Marineau and N. Brisson (1989). "Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases." Proceedings of the National Academy of Sciences **86**(8): 2582-2586.

Degtyarenko, K. N. and A. I. Archakov (1993). "Molecular evolution of P450 superfamily and P450-containing monooxygenase systems." Federation of European Biochemical Societies letters **332**(1): 1-8.

Dennison, J. B., P. Kulanthaivel, R. J. Barbuch, J. L. Renbarger, W. J. Ehlhardt and S. D. Hall (2006). "Selective metabolism of vincristine in vitro by CYP3A5." Drug metabolism and disposition **34**(8): 1317-1327.

Dietrich, J. A., Y. Yoshikuni, K. J. Fisher, F. X. Woolard, D. Ockey, D. J. McPhee, N. S. Renninger, M. C. Chang, D. Baker and J. D. Keasling (2009). "A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450BM3." ACS chemical biology **4**(4): 261-267.

Duarte, P., A. S. Rocha, D. Ribeiro, F. Lima, F. Hilliou, G. Henriques, I. Amorim and M. Sottomayor (2011). Cloning and characterization of a candidate gene from the medicinal plant *Catharanthus roseus* through transient expression in mesophyll protoplasts. Molecular Cloning - Selected Applications in Medicine and Biology. G. G. Brown, InTech 291-308.

El-Sayed, M. and R. Verpoorte (2007). "*Catharanthus* terpenoid indole alkaloids: biosynthesis and regulation." Phytochemistry Reviews **6**(2-3): 277-305.

English, N., V. Hughes and C. R. Wolf (1994). "Common pathways of cytochrome P450 gene regulation by peroxisome proliferators and barbiturates in *Bacillus megaterium* ATCC14581." Journal of Biological Chemistry **269**(43): 26836-26841.

Fasan, R., M. M. Chen, N. C. Crook and F. H. Arnold (2007). "Engineered alkane-hydroxylating cytochrome P450BM3 exhibiting nativelike catalytic properties." Angewandte Chemie **119**(44): 8566-8570.

Fisher, C. D., A. J. Lickteig, L. M. Augustine, J. Ranger-Moore, J. P. Jackson, S. S. Ferguson and N. J. Cherrington (2009). "Hepatic cytochrome P450 enzyme alterations in humans with

progressive stages of nonalcoholic fatty liver disease." Drug Metabolism and Disposition **37**(10): 2087-2094.

Fulco, A. J. (1967). "The effect of temperature on the formation of Δ 5-unsaturated fatty acids by bacilli." Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism **144**(3): 701-703.

Fulco, A. J. (1974). "Metabolic alterations of fatty acids." Annual review of biochemistry **43**(1): 215-241.

Gietz, R. D. and R. A. Woods (2002). "Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method." Methods in enzymology **350**: 87-96.

Girhard, M., P. J. Bakkes, O. Mahmoud and V. B. Urlacher (2015). P450 Biotechnology. Cytochrome P450. P. R. O. d. Montellano, Springer International Publishing. **1:** 451-520.

Glenn, W. S., W. Runguphan and S. E. O'Connor (2013). "Recent progress in the metabolic engineering of alkaloids in plant systems." Current opinion in biotechnology **24**(2): 354-365.

Goedhart, J., L. van Weeren, M. A. Hink, N. O. Vischer, K. Jalink and T. W. Gadella (2010). "Bright cyan fluorescent protein variants identified by fluorescence lifetime screening." Nature methods **7**(2): 137-139.

Gotoh, O. (1992). "Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences." Journal of Biological Chemistry **267**(1): 83-90.

Guengerich, F. P. (2015). Human cytochrome P450 enzymes. Cytochrome P450. P. R. O. d. Montellano, Springer International Publishing. **2:** 523-785.

Guengerich, F. P. and A. W. Munro (2013). "Unusual cytochrome P450 enzymes and reactions." Journal of Biological Chemistry **288**(24): 17065-17073.

Guengerich, P. F. (2014). Fifty Years of Progress in Drug Metabolism and Toxicology: What Do We Still Need to Know About Cytochrome P450 Enzymes? Fifty Years of Cytochrome P450 Research. H. Yamazaki, Springer Japan**:** 17-41.

Guirimand, G., V. Courdavault, A. Lanoue, S. Mahroug, A. Guihur, N. Blanc, N. Giglioli-Guivarc'h, B. St-Pierre and V. Burlat (2010). "Strictosidine activation in Apocynaceae: towards a." BMC plant biology **10**(1): 182.

Hanlon, S. P., T. Friedberg, C. R. Wolf, O. Ghisalba and M. Kittelmann (2007). "Recombinant yeast and bacteria that express human P450s: bioreactors for drug discovery, development, and biotechnology." Modern biooxidation: Enzymes, reactions and applications: 233-252.

Hannemann, F., A. Bichet, K. M. Ewen and R. Bernhardt (2007). "Cytochrome P450 systems—biological variations of electron transport chains." Biochimica et Biophysica Acta (BBA)-General Subjects **1770**(3): 330-344.

Hannig, G. and S. C. Makrides (1998). "Strategies for optimizing heterologous protein expression in *Escherichia coli*." Trends in biotechnology **16**(2): 54-60.

Heijden, R. v. d., D. I. Jacobs, W. Snoeijer, D. Hallard and R. Verpoorte (2004). "The *Catharanthus* alkaloids: pharmacognosy and biotechnology." Current medicinal chemistry **11**(5): 607-628.

Hintermann, G., H.-M. Fischer, R. Crameri and R. Hütter (1981). "Simple procedure for distinguishing CCC, OC, and L forms of plasmid DNA by agarose gel electrophoresis." Plasmid **5**(3): 371-373.

Irmler, S., G. Schröder, B. St-Pierre, N. P. Crouch, M. Hotze, J. Schmidt, D. Strack, U. Matern and J. Schröder (2000). "Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase." The Plant Journal **24**(6): 797-804.

Jaleel, C. A., B. Sankar, R. Sridharan and R. Panneerselvam (2008). "Soil salinity alters growth, chlorophyll content, and secondary metabolite accumulation in *Catharanthus roseus*." Turkish Journal of Biology **32**: 79-83.

Jang, H.-H., S.-Y. Kim, J.-Y. Kang, S. H. Park, S. H. Ryu, T. Ahn and C.-H. Yun (2012). "Increase of human CYP1B1 activities by acidic phospholipids and kinetic deuterium isotope effects on CYP1B1 substrate oxidation." Journal of biochemistry **152**(5): 433-442.

Jeong, S., P. D. Nguyen and Z. Desta (2009). "Comprehensive in vitro analysis of voriconazole inhibition of eight cytochrome P450 (CYP) enzymes: major effect on CYPs 2B6, 2C9, 2C19, and 3A." Antimicrobial agents and chemotherapy **53**(2): 541-551.

Jiang, Y., X. He and P. R. Ortiz de Montellano (2006). "Radical intermediates in the catalytic oxidation of hydrocarbons by bacterial and human cytochrome P450 enzymes." Biochemistry **45**(2): 533-542.

Jung, S. T., R. Lauchli and F. H. Arnold (2011). "Cytochrome P450: taming a wild type enzyme." Current opinion in biotechnology **22**(6): 809-817.

Kellner, F., F. Geu-Flores, N. H. Sherden, S. Brown, E. Foureau, V. Courdavault and S. E. O'Connor (2015). "Discovery of a P450-catalyzed step in vindoline biosynthesis: A link between the aspidosperma and eburnamine alkaloids." Chemical Communications **51**(36): 7626-7628.

Klingenberg, M. (1958). "Pigments of rat liver microsomes." Archives of biochemistry and biophysics **75**(2): 376-386.

Korhonen, T., M. Turpeinen, A. Tolonen, K. Laine and O. Pelkonen (2008). "Identification of the human cytochrome P450 enzymes involved in the *in vitro* biotransformation of lynestrenol and norethindrone." The Journal of steroid biochemistry and molecular biology **110**(1): 56-66.

Kot, M. and W. A. Daniel (2008). "The relative contribution of human cytochrome P450 isoforms to the four caffeine oxidation pathways: an in vitro comparative study with cDNA-expressed P450s including CYP2C isoforms." Biochemical pharmacology **76**(4): 543-551.

Kringen, M. K., K. B. F. Haug, R. M. Grimholt, C. Stormo, S. Narum, M. S. Opdal, J. T. Fosen, A. P. Piehler, P. W. Johansen and I. Seljeflot (2010). "Genetic variation of VKORC1 and CYP4F2 genes related to warfarin maintenance dose in patients with myocardial infarction." BioMed Research International **2011**.

Kumar, S. (2010). "Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation." Expert opinion on drug metabolism & toxicology **6**(2): 115-131.

Kutchan, T. M. (1995). "Alkaloid Biosynthesis [mdash] The Basis for Metabolic Engineering of Medicinal Plants." The Plant Cell **7**(7): 1059.

Laine, J., S. Auriola, M. Pasanen and R. Juvonen (2009). "Acetaminophen bioactivation by human cytochrome P450 enzymes and animal microsomes." Xenobiotica **39**(1): 11-21.

Lalovic, B., B. Phillips, L. L. Risler, W. Howald and D. D. Shen (2004). "Quantitative contribution of CYP2D6 and CYP3A to oxycodone metabolism in human liver and intestinal microsomes." Drug Metabolism and Disposition **32**(4): 447-454.

Lee, S.-J., I. P. van der Heiden, J. A. Goldstein and R. H. van Schaik (2007). "A new CYP3A5 variant, CYP3A5* 11, is shown to be defective in nifedipine metabolism in a recombinant cDNA expression system." Drug metabolism and disposition **35**(1): 67-71.

Lewis, D. (2001). Guide to Cytochromes P450 Structure and Function, London and New York: Taylor & Francis.

Lewis, J. C., S. M. Mantovani, Y. Fu, C. D. Snow, R. S. Komor, C. H. Wong and F. H. Arnold (2010). "Combinatorial alanine substitution enables rapid optimization of cytochrome P450BM3 for selective hydroxylation of large substrates." ChemBioChem **11**(18): 2502-2505.

Li, P., A. Anumanthan, X.-G. Gao, K. Ilangovan, V. V. Suzara, N. Düzgüneş and V. Renugopalakrishnan (2007). "Expression of recombinant proteins in *Pichia pastoris*." Applied biochemistry and biotechnology **142**(2): 105-124.

Li, Y. C. and J. Chiang (1991). "The expression of a catalytically active cholesterol 7 alpha-hydroxylase cytochrome P450 in *Escherichia coli*." Journal of Biological Chemistry **266**(29): 19186-19191.

Mahroug, S., V. Burlat and B. St-Pierre (2007). "Cellular and sub-cellular organisation of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*." Phytochemistry Reviews **6**(2-3): 363-381.

Matson, R. S., R. S. Hare and A. J. Fulco (1977). "Characteristics of a cytochrome P-450-dependent fatty acid ω-2 hydroxylase from *Bacillus megaterium*." Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism **487**(3): 487-494.

McLean, K. J., D. Leys and A. W. Munro (2015). Microbial Cytochromes P450. Cytochrome P450. P. R. O. d. Montellano, Springer International Publishing. **1:** 261-407.

Meunier, B., S. P. De Visser and S. Shaik (2004). "Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes." Chemical reviews **104**(9): 3947-3980.

Miettinen, K., L. Dong, N. Navrot, T. Schneider, V. Burlat, J. Pollier, L. Woittiez, S. van der Krol, R. Lugan and T. Ilc (2014). "The seco-iridoid pathway from *Catharanthus roseus*." Nature communications **5**.

Mirhendi, H., K. Diba, A. Rezaei, N. Jalalizand, L. Hosseinpur and H. Khodadadi (2007). "Colony PCR is a rapid and sensitive method for DNA amplification in yeasts." Iranian Journal of Public Health **36**(1): 40-44.

Mishra, P. and S. Kumar (2000). "Emergence of periwinkle *Catharanthus roseus* as a model system for molecular biology of alkaloids: phytochemistry, pharmacology, plant biology and *in vivo* and *in vitro* cultivation." Journal of Medicinal and Aromatic Plant Sciences **22**(2/3): 306-337.

Moreno, P. R., R. van der Heijden and R. Verpoorte (1995). "Cell and tissue cultures of *Catharanthus roseus*: a literature survey." Plant cell, tissue and organ culture **42**(1): 1-25.

Mujib, A., A. Ilah, J. Aslam, S. Fatima, Z. H. Siddiqui and M. Maqsood (2012). "*Catharanthus roseus* alkaloids: application of biotechnology for improving yield." Plant Growth Regulation **68**(2): 111-127.

Munro, A. W., H. M. Girvan, A. E. Mason, A. J. Dunford and K. J. McLean (2013). "What makes a P450 tick?" Trends in biochemical sciences **38**(3): 140-150.

Munro, A. W., D. G. Leys, K. J. McLean, K. R. Marshall, T. W. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek and C. C. Moser (2002). "P450 BM3: the very model of a modern flavocytochrome." Trends in biochemical sciences **27**(5): 250-257.

Munro, A. W., J. G. Lindsay, J. R. Coggins, S. M. Kelly and N. C. Price (1994). "Structural and enzymological analysis of the interaction of isolated domains of cytochrome P-450 BM3." Federation of European Biochemical Societies letters **343**(1): 70-74.

Narhi, L. O. and A. J. Fulco (1986). "Characterization of a catalytically self-sufficient 119,000-dalton cytochrome P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*." Journal of Biological Chemistry **261**(16): 7160-7169.

Narhi, L. O. and A. J. Fulco (1987). "Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbiturates in *Bacillus megaterium*." Journal of Biological Chemistry **262**(14): 6683-6690.

Nebert, D., M. Adesnik, M. Coon, R. Estabrook, F. Gonzalez, F. Guengerich, I. Gunsalus, E. Johnson, B. Kemper and W. Levin (1987). "The P450 gene superfamily: recommended nomenclature." Dna **6**(1): 1-11.

Nelson, D. R. (2006). Cytochrome P450 Nomenclature, 2004. Cytochrome P450 Protocols. I. R. Phillips, Shephard, Elizabeth A., Humana Press**:** 1-10.

Nelson, D. R., T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus and O. Gotoh (1993). "The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature." DNA and cell biology **12**(1): 1-51.

Nelson, D. R., L. Koymans, T. Kamataki, J. J. Stegeman, R. Feyereisen, D. J. Waxman, M. R. Waterman, O. Gotoh, M. J. Coon and R. W. Estabrook (1996). "P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature." Pharmacogenetics and Genomics **6**(1): 1-42.

Noble, M., C. Miles, S. Chapman, D. Lysek, A. Mackay, G. Reid, R. Hanzlik and A. Munro (1999). "Roles of key active-site residues in flavocytochrome P450 BM3." Biochemical Journal **339**: 371-379.

Noble, R., C. Beer and J. Cutts (1958). "Role of chance observation in chemotherapy." Annals of the New York Academy of Science **76**: 882-894.

Noble, R., C. Beer and J. Cutts (1959). "Further biological activities of vincaleukoblastine—an alkaloid isolated from *Vinca rosea* (L.)." Biochemical Pharmacology **1**(4): 347-348.

O'Connor, S. E. and J. J. Maresh (2006). "Chemistry and biology of monoterpene indole alkaloid biosynthesis." Natural product reports **23**(4): 532-547.

Omura, T. (2013). "Contribution of cytochrome P450 to the diversification of eukaryotic organisms." Biotechnology and applied biochemistry **60**(1): 4-8.

Omura, T. and R. Sato (1964). "The carbon monoxide-binding pigment of liver microsomes." Journal of Biological Chemistry **239**(7): 183-189.

Ost, T. W., J. Clark, C. G. Mowat, C. S. Miles, M. D. Walkinshaw, G. A. Reid, S. K. Chapman and S. Daff (2003). "Oxygen activation and electron transfer in flavocytochrome P450 BM3." Journal of the American Chemical Society **125**(49): 15010-15020.

Otey, C. R., G. Bandara, J. Lalonde, K. Takahashi and F. H. Arnold (2006). "Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450." Biotechnology and bioengineering **93**(3): 494-499.

Oudin, A., M. Courtois, M. Rideau and M. Clastre (2007). "The iridoid pathway in *Catharanthus roseus* alkaloid biosynthesis." Phytochemistry Reviews **6**(2-3): 259-276.

Palmer, C. N., E. Axen, V. Hughes and C. R. Wolf (1998). "The repressor protein, Bm3R1, mediates an adaptive response to toxic fatty acids in *Bacillus megaterium*." Journal of Biological Chemistry **273**(29): 18109-18116.

Pan, Q., N. R. Mustafa, K. Tang, Y. H. Choi and R. Verpoorte (2015). "Monoterpenoid indole alkaloids biosynthesis and its regulation in *Catharanthus roseus*: a literature review from genes to metabolites." Phytochemistry Reviews: 1-30.

Pateraki, I., A. M. Heskes and B. Hamberger (2015). "Cytochromes P450 for Terpene Functionalisation and Metabolic Engineering." Berlin; Heidelberg: Springer.

Plaizier, A. (1981). "A revision of *Catharanthus roseus* (L.) G. Don (Apocynaceae)." Mededelingen Landbouwhogeschool,vol.81,no. 9,pp.1–12.

Pompon, D., B. Louerat, A. Bronine and P. Urban (1996). "Yeast expression of animal and plant P450s in optimized redox environments." Methods in enzymology(272): 51-64.

Qu, Y., M. L. Easson, J. Froese, R. Simionescu, T. Hudlicky and V. De Luca (2015). "Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast." Proceedings of the National Academy of Sciences **112**(19): 6224-6229.

Raina, S. K., D. P. Wankhede, M. Jaggi, P. Singh, S. K. Jalmi, B. Raghuram, A. H. Sheikh and A. K. Sinha (2012). "CrMPK3, a mitogen activated protein kinase from *Catharanthus roseus* and its possible role in stress induced biosynthesis of monoterpenoid indole alkaloids." BMC plant biology **12**(1): 134.

Rastogi, S., S. K. Khanna and M. Das (2002). "A spectrophotometric method to monitor the catalytic activity of microsomal cytochrome P-450 IIB 1/2: Comparison with fluorometric assay." Indian Journal of Biochemistry and Biophysics **39**(3): 191-196.

Rendic, S. and F. P. Guengerich (2014). "Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals." Chemical research in toxicology **28**(1): 38-42.

Rischer, H., M. Orešič, T. Seppänen-Laakso, M. Katajamaa, F. Lammertyn, W. Ardiles-Diaz, M. C. Van Montagu, D. Inzé, K.-M. Oksman-Caldentey and A. Goossens (2006). "Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells." Proceedings of the National Academy of Sciences **103**(14): 5614-5619.

Roepke, J., V. Salim, M. Wu, A. M. Thamm, J. Murata, K. Ploss, W. Boland and V. De Luca (2010). "Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle." Proceedings of the National Academy of Sciences **107**(34): 15287-15292.

Rowland, P., F. E. Blaney, M. G. Smyth, J. J. Jones, V. R. Leydon, A. K. Oxbrow, C. J. Lewis, M. G. Tennant, S. Modi and D. S. Eggleston (2006). "Crystal structure of human cytochrome P450 2D6." Journal of Biological Chemistry **281**(11): 7614-7622.

Rubin, E. and C. Lieber (1968). "Alcohol, other drugs, and the liver." Annals of internal medicine **69**(5): 1063-1067.

Ruettinger, R. T., L.-P. Wen and A. Fulco (1989). "Coding nucleotide, 5'regulatory, and deduced amino acid sequences of P-450BM-3, a single peptide cytochrome P-450: NADPH-P-450 reductase from *Bacillus megaterium*." Journal of Biological Chemistry **264**(19): 10987-10995.

Saiman, M. Z. (2014). "Terpenoids and terpenoid indole alkaloids in *Catharanthus roseus* cell suspension cultures." Institute of Biology (IBL), Faculty of Science, Leiden University.

Saito, Y., N. Hanioka, K. Maekawa, T. Isobe, Y. Tsuneto, R. Nakamura, A. Soyama, S. Ozawa, T. Tanaka-Kagawa and H. Jinno (2005). "Functional analysis of three CYP1A2 variants found in a Japanese population." Drug metabolism and disposition **33**(12): 1905-1910.

Sakaki, T., O. Kenji, M. Miyoshi and H. Ohkawa (1985). "Characterization of rat cytochrome P-450MC synthesized in *Saccharomyces cerevisiae*." Journal of biochemistry **98**(1): 167-175.

Salim, V., F. Yu, J. Altarejos and V. Luca (2013). "Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis." The Plant Journal **76**(5): 754-765.

Sambrook, J., E. Fritsch and T. Maniatis (1989). Molecular cloning: A laboratory manual+ Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sandhu, P., Z. Guo, T. Baba, M. V. Martin, R. H. Tukey and F. P. Guengerich (1994). "Expression of modified human cytochrome P450 1A2 in *Escherichia coli*: stabilization, purification, spectral characterization, and catalytic activities of the enzyme." Archives of biochemistry and biophysics **309**(1): 168-177.

Sawayama, A. M., M. M. Chen, P. Kulanthaivel, M. S. Kuo, H. Hemmerle and F. H. Arnold (2009). "A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds." Chemistry-A European Journal **15**(43): 11723-11729.

Schenkman, J. B. and I. Jansson (2006). Spectral analyses of cytochromes P450. Cytochrome P450 Protocols. I. R. Phillips, Shephard, Elizabeth A., Humana Press**:** 11-18.

Schröder, G., E. Unterbusch, M. Kaltenbach, J. Schmidt, D. Strack, V. De Luca and J. Schröder (1999). "Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase." Federation of European Biochemical Societies letters **458**(2): 97-102.

Schroer, K., M. Kittelmann and S. Lütz (2010). "Recombinant human cytochrome P450 monooxygenases for drug metabolite synthesis." Biotechnology and bioengineering **106**(5): 699-706.

Schuler, M. A. (2015). P450s in Plants, Insects, and Their Fungal Pathogens. Cytochrome P450. P. R. O. d. Montellano, Springer International Publishing. **1:** 409-449.

Shukla, A. K. and S. Khanuja (2013). "*Catharanthus roseus*: The Metabolome That Represents a Unique Reservoir of Medicinally Important Alkaloids Under Precise Genomic Regulation, OMICS Applications in Crop Science." OMICS Applications in Crop Science: 325.

Sottomayor, M. and A. R. Barceló (2006). "The Vinca alkaloids: from biosynthesis and accumulation in plant cells, to uptake, activity and metabolism in animal cells." Studies in Natural Products Chemistry **33**: 813-857.

Sottomayor, M., I. L. Cardoso, L. Pereira and A. R. Barceló (2004). "Peroxidase and the biosynthesis of terpenoid indole alkaloids in the medicinal plant *Catharanthus roseus* (L.) G. Don." Phytochemistry Reviews **3**(1-2): 159-171.

Sottomayor, M., M. Lopez-Serrano, F. DiCosmo and A. R. Barceló (1998). "Purification and characterization of α-3′, 4′-anhydrovinblastine synthase (peroxidase-like) from *Catharanthus roseus* (L.) G. Don." Federation of European Biochemical Societies letters **428**(3): 299-303.

St-Pierre, B., S. Besseau, M. Clastre, V. Courdavault, M. Courtois, J. Creche, E. Ducos, T. D. de Bernonville, C. Dutilleul and G. Glevarec (2013). "Deciphering the evolution, cell biology and regulation of monoterpene indole alkaloids." Advances in Botanical. Research **68**: 73-109.

St-Pierre, B. and V. De Luca (1995). "A cytochrome P-450 monooxygenase catalyzes the first step in the conversion of tabersonine to vindoline in *Catharanthus roseus*." Plant physiology **109**(1): 131-139.

Stiborová, M., E. Frei, M. Wiessler and H. H. Schmeiser (2001). "Human enzymes involved in the metabolic activation of carcinogenic aristolochic acids: evidence for reductive activation by cytochromes P450 1A1 and 1A2." Chemical research in toxicology **14**(8): 1128-1137.

Svoboda, G. H., N. Neuss and M. Gorman (1959). "Alkaloids of *Vinca rosea* Linn.(*Catharanthus roseus* G. Don.) V. Preparation and characterization of alkaloids." Journal of the American Pharmaceutical Association **48**(11): 659-666.

Tang, K., D. Liu, Y. Wang, L. Cui, W. Ren and X. Sun (2011). "Overexpression of transcriptional factor ORCA3 increases the accumulation of catharanthine and vindoline in *Catharanthus roseus* hairy roots." Russian Journal of Plant Physiology **58**(3): 415-422.

Tikhomiroff, C. and M. Jolicoeur (2002). "Screening of *Catharanthus roseus* secondary metabolites by high-performance liquid chromatography." Journal of chromatography A **955**(1): 87-93.

Urlacher, V. B. and M. Girhard (2012). "Cytochrome P450 monooxygenases: an update on perspectives for synthetic application." Trends in biotechnology **30**(1): 26-36.

Van Bergen, M. A. and W. Snoeijer (1996). *Catharanthus roseus*, the Madagascar periwinkle, a review of its cultivars. Revision of Catharanthus G. Don, series of revisions of Apocynaceae XLI. M. A. Van Bergen, Wageningen Agricultural University; Distribution, Backhuys Publishers.

Van Bogaert, I. N., S. Groeneboer, K. Saerens and W. Soetaert (2011). "The role of cytochrome P450 monooxygenases in microbial fatty acid metabolism." Federation of European Biochemical Societies journal **278**(2): 206-221.

Van Der Fits, L. and J. Memelink (2001). "The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element." The Plant Journal **25**(1): 43-53.

van der Hoeven, T. A. and M. J. Coon (1974). "Preparation and properties of partially purified cytochrome P-450 and reduced nicotinamide adenine dinucleotide phosphate-cytochrome P-450 reductase from rabbit liver microsomes." Journal of Biological Chemistry **249**(19): 6302-6310.

Verma, P., A. K. Mathur, A. Srivastava and A. Mathur (2012). "Emerging trends in research on spatial and temporal organization of terpenoid indole alkaloid pathway in *Catharanthus roseus*: a literature update." Protoplasma **249**(2): 255-268.

Verpoorte, R., Van der Heijden, R., Moreno, PRH (1997). Biosynthesis of terpenoid indole alkaloids in *Catharanthus roseus* cells. The Alkaloids: Chemistry and Pharmacology G. A. Cordell, ACADEMIC PRESS. **49:** 221 - 299.

Vesell, E. S. and J. G. Page (1968). "Genetic control of drug levels in man: antipyrine." Science **161**(3836): 72-73.

Virmani, O., G. Srivastava and P. Singh (1978). "*Catharanthus roseus*--the tropical periwinkle." Indian drugs.

Wahlang, B., K. C. Falkner, M. C. Cave and R. A. Prough (2015). "Role of Cytochrome P450 Monooxygenase in Carcinogen and Chemotherapeutic Drug Metabolism." Advances in Pharmacology.

Wang, P. P., P. Beaune, L. S. Kaminsky, G. A. Dannan, F. F. Kadlubar, D. Larrey and F. P. Guengerich (1983). "Purification and characterization of six cytochrome P-450 isozymes from human liver microsomes." Biochemistry **22**(23): 5375-5383.

Ward, A. C. (1992). "Rapid analysis of yeast transformants using colony-PCR." Biotechniques **13**(3): 350-350.

Ward, B. A., A. Morocho, A. Kandil, R. E. Galinsky, D. A. Flockhart and Z. Desta (2004). "Characterization of human cytochrome P450 enzymes catalyzing domperidone N-dealkylation and hydroxylation in vitro." British journal of clinical pharmacology **58**(3): 277-287.

Warman, A., O. Roitel, R. Neeli, H. Girvan, H. Seward, S. Murray, K. McLean, M. Joyce, H. Toogood and R. Holt (2005). "Flavocytochrome P450 BM3: an update on structure and mechanism of a biotechnologically important enzyme." Biochemical Society Transactions **33**(4): 747-753.

Weickert, M. J., D. H. Doherty, E. A. Best and P. O. Olins (1996). "Optimization of heterologous protein production in *Escherichia coli*." Current opinion in biotechnology **7**(5): 494-499.

Whitehouse, C. J., S. G. Bell and L.-L. Wong (2012). "P450 BM3 (CYP102A1): connecting the dots." Chemical Society Reviews **41**(3): 1218-1260.

Williams, P. A., J. Cosme, A. Ward, H. C. Angove, D. M. Vinković and H. Jhoti (2003). "Crystal structure of human cytochrome P450 2C9 with bound warfarin." Nature **424**(6947): 464-468.

Wink, M. (1998). Chemical ecology of alkaloids. Alkaloids: Biochemistry, Ecology, and Medicinal Applications. M. W. Margaret F. Roberts, Springer**:** 265-300.

Wong, L.-L. (1998). "Cytochrome P450 monooxygenases." Current opinion in chemical biology **2**(2): 263-268.

Wright, R. L., K. Harris, B. Solow, R. H. White and P. J. Kennelly (1996). "Cloning of a potential cytochrome P450 from the archaeon *Sulfolobus solfataricus*." Federation of European Biochemical Societies letters **384**(3): 235-239.

Yao, D., S. Ding, B. Burchell, C. R. Wolf and T. Friedberg (2000). "Detoxication of vinca alkaloids by human P450 CYP3A4-mediated metabolism: implications for the development of drug resistance." Journal of Pharmacology and Experimental Therapeutics **294**(1): 387-395.

Yoder, L. R. and P. G. Mahlberg (1976). "Reactions of alkaloid and histochemical indicators in laticifers and specialized parenchyma cells of *Catharanthus roseus* (Apocynaceae)." American Journal of Botany: 1167-1173.

Zelasko, S., A. Palaria and A. Das (2013). "Optimizations to achieve high-level expression of cytochrome P450 proteins using *Escherichia coli* expression systems." Protein expression and purification **92**(1): 77-87.

Zhu, J., M. Wang, W. Wen and R. Yu (2015). "Biosynthesis and regulation of terpenoid indole alkaloids in *Catharanthus roseus*." Pharmacognosy reviews **9**(17): 24.

6. APPENDIXES

Appendix A. Map of the plasmid pMON999

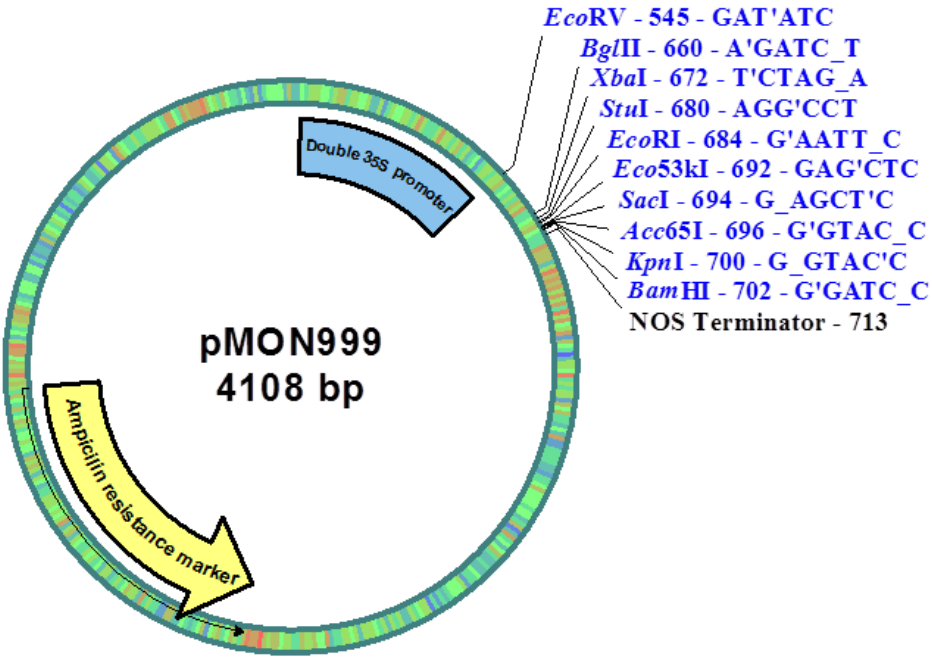

Figure 34.Map of the plasmid pMON999. This map includes: multi cloning site, the ampicillin resistance marker, the 35S promotor and NOS Terminator.

Appendix B. Representative chromatogram and spectra of the degradation products of α-3',4'-anhydrovinblastine used as standard of that compound

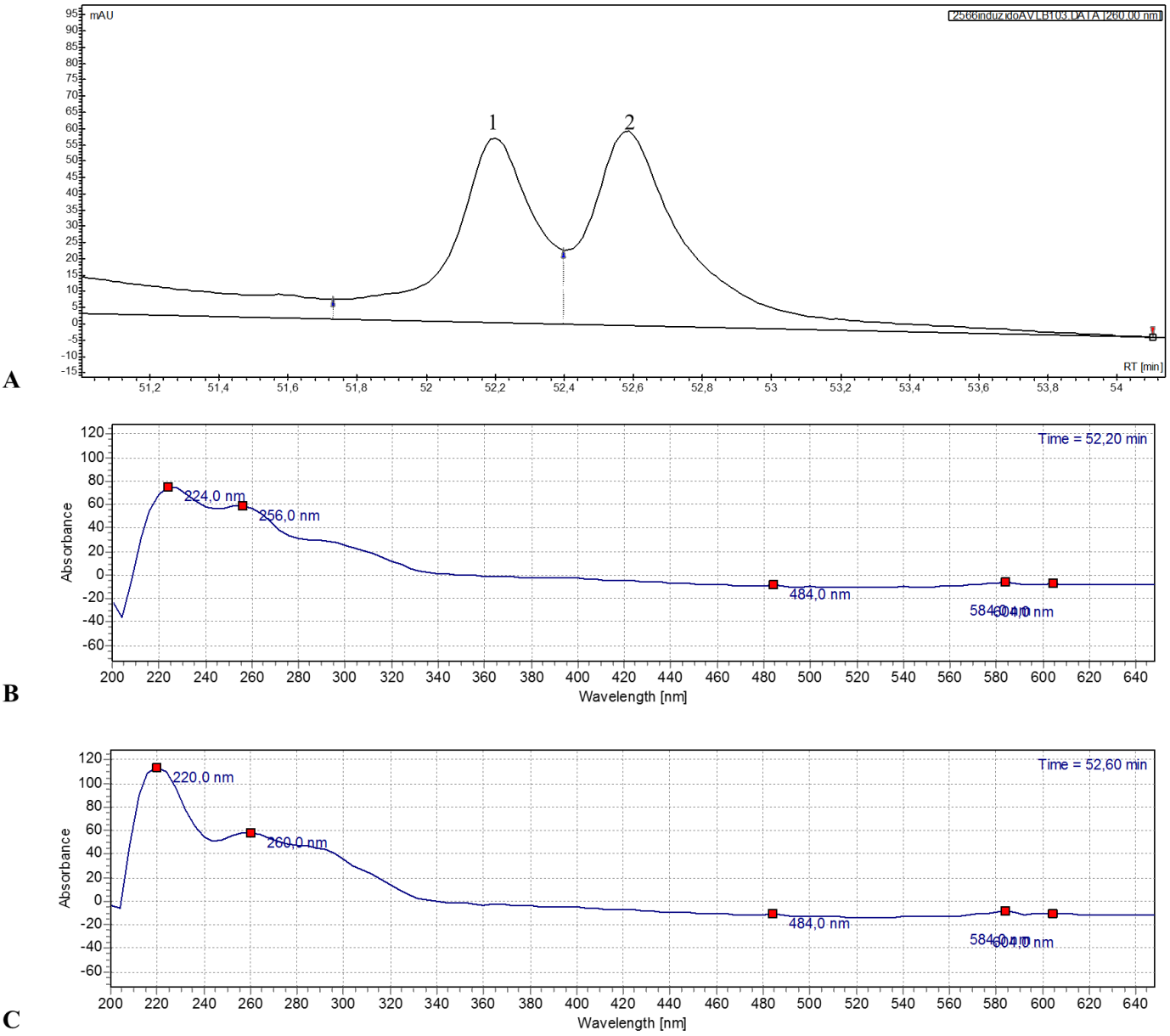

Figure 35. **Representative chromatogram and spectra for degradation of α-3',4'-anhydrovinblastine.** In A is represented the chromatogram at retention time RT of the degraded product composed of two peaks (1 and 2). The chromatogram of α- 3',4'-anhydrovinblastine is characterized by a single peak at 52,4. In B is represented the absorption spectrum corresponding to peak 1. In C is represented the absorption spectrum corresponding to peak 2, with absorption peaks typical of TIA dimers.