

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

IGOR CESARINO



**“CLONAGEM, EXPRESSÃO E CARACTERIZAÇÃO DE UMA
FLAVINA MONOOXIGENASE DE *Coffea arabica*”**

Este exemplar corresponde à redação final da tese defendida pelo(a) candidato (a) <i>IGOR CESARINO</i>
e aprovado pela Comissão Julgadora.

Dissertação apresentada ao Instituto de Biologia para obtenção do Título de Mestre em Biologia Vegetal.

Orientador: Prof. Dr. Paulo Mazzafera

Campinas, 2009

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

C337c	<p>Cesarino, Igor Clonagem, expressão e caracterização de uma flavina monooxigenase de <i>Coffea arabica</i> / Igor Cesarino. – Campinas, SP: [s.n.], 2009.</p> <p>Orientador: Paulo Mazzafera. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.</p> <p>1. <i>Coffea arabica</i>. 2. Flavina monooxigenase. 3. Caracterização funcional. I. Mazzafera, Paulo. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.</p> <p>(scs/ib)</p>
--------------	--

Título em inglês: Cloning, expression and characterization of flavin-containing monooxygenase from *Coffea arabica*.

Palavras-chave em inglês: *Coffea arabica*; Flavin-containing monooxygenase; Functional characterization.

Área de concentração: Biologia Vegetal.

Titulação: Mestre em Biologia Vegetal.

Banca examinadora: Paulo Mazzafera, Marcelo Carnier Dornelas, Jörg Kobarg.

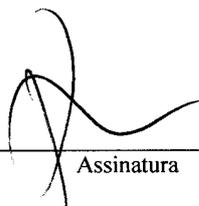
Data da defesa: 19/02/2009.

Programa de Pós-Graduação: Biologia Vegetal.

Campinas, 19 de fevereiro de 2009.

BANCA EXAMINADORA

Prof. Dr. Paulo Mazzafera (orientador)



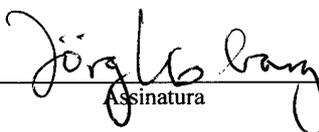
Assinatura

Prof. Dr. Marcelo Carnier Dornelas



Assinatura

Prof. Dr. Jörg Kobarg



Assinatura

Dr. Milton Massao Shimizu

Assinatura

Profa. Dra. Anita Jocelyne Marsaioli

Assinatura

AGRADECIMENTOS

Aos meus pais, pela presença constante em minha formação, pelo apoio em minhas decisões e pelo afeto.

Ao professor Paulo Mazzafera, meu orientador, pelas lições e ensinamentos ao longo de todos estes anos de orientação e amizade.

Ao Programa de Pós-Graduação em Biologia Vegetal.

Ao Departamento de Fisiologia Vegetal, pela infra-estrutura.

Ao CNPq, pelo apoio financeiro.

À professora Anete Pereira de Souza, do LAGM/CBMEG, e seus alunos, pela ajuda experimental em praticamente todo o trabalho.

Ao Dr. Milton Massao Shimizu, pela ajuda em meus primeiros passos.

Ao Dr. Marcelo Prudêncio Giovanini, pelos conhecimentos passados e por aumentar meu interesse pela pesquisa científica.

Ao Dr. Marcelo J. Murai, pela ajuda ao recomeçar todo o projeto.

À Dra. Alexandra Christine Helena Frankland Sawaya, pela colaboração nas análises em HPLC e o empréstimo da única coluna capaz de separar triptamina.

A todos do Depto de Fisiologia Vegetal, alunos, técnicos e professores, por me ajudarem a crescer como pessoa e como pesquisador.

Aos membros da banca examinadora, por aceitarem o convite e pelas valiosas sugestões ao meu trabalho.

À Juliana, pelo carinho, pela paciência e pelo companheirismo em absolutamente todos os momentos.

A todos os meus amigos e pessoas que, de qualquer forma, colaboraram para a realização deste trabalho.

Índice

Resumo.....	vi
Summary.....	vii
1. Introdução	1
<i>FMO: definições e distribuição</i>	<i>1</i>
<i>Mecanismo de ação.....</i>	<i>2</i>
<i>FMOs em plantas</i>	<i>4</i>
<i>Estrutura e motivos conservados</i>	<i>5</i>
<i>A família YUCCA</i>	<i>6</i>
2. Objetivo Geral.....	13
<i>Objetivos Específicos</i>	<i>13</i>
3. Material e Métodos	14
<i>Material vegetal</i>	<i>14</i>
<i>Aquisição da seqüência gênica e análises com ferramentas de bioinformática.....</i>	<i>14</i>
<i>Alinhamentos e análise filogenética.....</i>	<i>15</i>
<i>Isolamento de RNA e análises de expressão</i>	<i>16</i>
<i>Produção da proteína CaFMO8 recombinante</i>	<i>17</i>
4. Resultados.....	26
<i>Determinação da ORF e da seqüência traduzida.....</i>	<i>26</i>
<i>CaFMO8 apresenta expressão preferencial em raízes, folhas e flores</i>	<i>31</i>
<i>Produção de CaFMO8 recombinante em E. coli.....</i>	<i>32</i>
5. Discussão.....	41
6. Conclusões	49
7. Referências Bibliográficas.....	50

Resumo

Uma grande quantidade de genes que codificam flavina monooxigenases (Flavin-containing monooxygenases – FMOs) é encontrada em genomas vegetais, embora poucas funções biológicas tenham sido relacionadas com esse grupo de enzimas em plantas. Um importante papel desempenhado por FMOs é a conversão de triptamina em N-hidroxil triptamina, reação catalisada pelas proteínas YUCCA de *Arabidopsis thaliana* e que constitui o passo limitante da via de síntese de auxina a partir de triptofano. Proteínas similares às YUCCA foram descobertas e caracterizadas em outras espécies vegetais, como OsYUCCA em arroz, FLOOZY em petúnia, ToFZY em tomate e SPI1 em milho, todas comprovadamente envolvidas na produção do hormônio citado. Análises da proteína recombinante CaFMO8 de *Coffea arabica* revelou características similares às YUCCA, sugerindo que esta proteína de café é a primeira YUCCA-like descrita para esta espécie e, inclusive, para a família Rubiaceae. CaFMO8 apresenta os mesmos motivos protéicos conservados entre FMOs vegetais, e particularmente entre proteínas YUCCA-like. O padrão de expressão espacial do gene que codifica CaFMO8 indica possível relação com o desenvolvimento de raízes, folhas e flores de café. Apesar de grandes semelhanças com as proteínas YUCCA, a atividade de N-hidroxilação da triptamina não foi comprovada para CaFMO8 recombinante *in vitro*. Uma análise minuciosa a respeito da funcionalidade de CaFMO8 produzida em *E. coli* deve ser feita antes de descartar a participação desta proteína na síntese de auxina.

Summary

A large number of genes coding flavin-containing monooxygenases (FMOs) is found in plant genomes, although only few biological functions have been related with these enzymes in plants. An important role for FMOs is the conversion of tryptamine in N-hydroxyl tryptamine, catalyzed by the YUCCA protein family in *Arabidopsis thaliana*. These proteins perform the rate-limiting step in tryptophan-dependent auxin biosynthesis. Similar YUCCA proteins were discovered and characterized in other plant species, like OsYUCCA in rice, FLOOZY in petunia, ToFZY in tomato and SPI1 in maize. All of them are shown to be involved in auxin synthesis. Analysis of the recombinant CaFMO8 from *Coffea arabica* showed features similar to YUCCA proteins, suggesting that CaFMO8 is the first described YUCCA-like protein from coffee and, indeed, from the entire Rubiaceae family. CaFMO8 has the same conserved motifs found in other plant FMOs and particularly conserved in YUCCA-like proteins. The spatial expression pattern from the CaFMO8 coding gene suggests a probable role in the development of roots, leaves and flowers. Although very similar to YUCCA proteins, the CaFMO8-mediated conversion of tryptamine in N-hydroxyl tryptamine has not been confirmed *in vitro*. A further analysis of CaFMO8 functionality should address the relation of CaFMO8 to auxin production.

1. Introdução

FMO: definições e distribuição

Flavina monooxigenases (Flavin-containing monooxygenases - FMOs) foram descobertas na década de 1960 em microsossomos hepáticos de mamíferos, com função de converter *N,N'*-dimetilanilina em seu respectivo N-óxido, mediante utilização de NADPH e oxigênio molecular (Ziegler & Pettit, 1964). Características das FMOs, como a localização, o requerimento de co-fator e a atividade sobre xenobióticos, as tornam muito semelhantes a outro sistema enzimático, as proteínas citocromos P450 ou CYPs (Krueger & Williams, 2005). Embora ambos catalisem a *N*- ou *S*-oxigenação dependente de NADPH de compostos contendo hetero-átomo, estes dois sistemas de monooxigenases microsossomais são distintos, diferindo nos mecanismos catalíticos, regulação, especificidade de substratos, entre outros (Cashman, 2005). Ao contrário das CYPs, que são inibidas por uma significativa variedade de compostos, as FMOs raramente são inibidas. Inclusive, há exemplos de metabólitos reativos produzidos pela atividade de FMOs que inibem as CYPs, embora não inibam as próprias FMOs (Cashman *et al.*, 1995).

Em mamíferos, as FMOs estão relacionadas com a remoção de produtos naturais indesejados provenientes da alimentação, os chamados xenobióticos. De maneira geral, as FMOs convertem substratos nucleofílicos lipofílicos em metabólitos mais polares, oxigenados e prontos para serem excretados (Cashman & Zhang, 2006). Acredita-se que essas enzimas evoluíram como um sistema de proteção de mamíferos ao “ataque” de compostos químicos nucleofílicos/lipofílicos de origem vegetal, como alguns metabólitos secundários (Ziegler, 1990). Os mamíferos apresentam cinco genes *FMO* funcionais, nomeados *FMO1* a *FMO5* (Lawton *et al.*, 1994). Um sexto gene foi descoberto em humanos, mas pesquisas posteriores mostraram se tratar de um pseudogene (Hines *et al.*, 2002).

A levedura *Saccharomyces cerevisiae* apresenta somente um gene cujo produto é homólogo à FMOs (Suh *et al.*, 1996). Um trabalho com a FMO de levedura (yFMO) recombinante mostrou sua função na manutenção do potencial de oxido-redução intracelular, sendo que yFMO estaria envolvida na formação de pontes dissulfeto em proteínas que se encontram no retículo endoplasmático (proteínas que serão exportadas) e na regulação dos níveis celulares de thiol/dissulfeto pela oxidação da glutatona (Suh *et al.*, 2000). Os autores sugerem que este seria o papel ancestral das FMOs e que, em mamíferos, a função de oxigenação de xenobióticos teria surgido depois, com uma família multigênica derivada desta monooxigenase original. Outros pesquisadores também sugerem que a origem da família gênica das FMOs em mamíferos ocorreu a partir da duplicação de um gene ancestral ou via uma série de eventos independentes de duplicação gênica (Hernandez *et al.*, 2004).

Choi *et al.* (2003) caracterizaram uma FMO proveniente da bactéria *Methylophaga* sp. strain SK1, responsável pela produção do pigmento azul índigo, a mais antiga tinta conhecida pela humanidade, tornando-se a primeira FMO de origem procariota a ser caracterizada.

Comparadas a outros organismos, as plantas apresentam um número muito maior de genes *FMO*. Em *Arabidopsis thaliana*, 29 genes estão anotados como *FMO*-like. Isto sugere um uso muito mais amplo deste sistema enzimático por parte dos vegetais em relação a outros organismos (Schlaich, 2007). De fato, pesquisas dos últimos sete anos revelaram funções específicas de FMOs em plantas, como a síntese de auxina (Zhao *et al.*, 2001), metabolismo de glucosinolatos (Hansen *et al.*, 2007) e defesa contra patógenos (Koch *et al.*, 2006; Mishina & Zeier, 2006).

Mecanismo de ação

FMOs são flavoenzimas que apresentam uma molécula de FAD (flavin adenine dinucleotide) por monômero (Kubo *et al.*, 1997). Seu mecanismo de ação difere do

mecanismo de outras monooxigenases, não havendo necessidade de ligação do substrato para a redução do grupo prostético FAD pelo co-fator NADPH (Eswaramoorthy *et al.*, 2007).

No primeiro passo do ciclo catalítico, FAD sofre uma dupla redução pelo NADPH, tornando-se capaz de reagir com o oxigênio molecular e formando o intermediário 4 α -hydroperoxyFAD. Este intermediário é responsável pela inserção de um átomo de oxigênio no substrato, tornando-o mais polar e gerando o segundo intermediário 4 α -hydroxyFAD. Portanto, quando o substrato se liga à FMO, a enzima já está na forma ativa, pronta para efetuar a oxigenação. A perda do segundo átomo de oxigênio pela flavina na forma de uma molécula de água regenera o co-fator oxidado. A característica mais marcante do ciclo catalítico destas enzimas é sua capacidade de estabilizar o essencial intermediário 4 α -hydroperoxyFAD até a inserção do substrato no seu sítio ativo. A presença do NADP⁺ é crucial para a estabilização do intermediário e, conseqüentemente, NADP⁺ permanece ligado ao complexo enzimático durante todo o ciclo catalítico, sendo o último produto a ser liberado (Alfieri *et al.*, 2008).

Este mecanismo catalítico foi descrito para FMOs de mamíferos (Cashman, 2005; Krueger & Williams, 2005). Eswaramoorthy *et al.* (2007) descreveram o mecanismo catalítico de uma FMO de *Schizosaccharomyces pombe* após a determinação de sua estrutura cristalizada. Eles observaram que o co-fator (NADPH) e o substrato (metimazole, no caso) interagem com o grupo prostético (FAD) alternadamente, permitindo que a enzima catalise os substratos continuamente. Eles sugeriram ainda que a redução do FAD ocorre antes da ligação do substrato, mas a reação com o oxigênio molecular aconteceria apenas na presença do substrato. Os resultados obtidos sugerem que o substrato deve substituir o NADP⁺, ligando-se no mesmo sítio de ligação que a nicotinamida do co-fator. Pelo fato deste processo ter sido demonstrado com apenas um substrato, em apenas uma estrutura protéica, generalizações

podem ser prematuras. Alfieri *et al.* (2008) determinaram a estrutura cristalizada da mesma FMO bacteriana caracterizada por Choi *et al.* (2003). Neste estudo, além de sugerirem uma segunda função para o NADPH, a de estabilização do intermediário 4 α -hydroperoxyFAD, eles tentam explicar os resultados adquiridos por Eswaramoorthy *et al.* (2007). O mecanismo proposto envolve uma mudança de conformação da proteína para a aceitação do substrato, ao invés da liberação do co-fator, como propuseram Eswaramoorthy *et al.* (2007). Assim, haveria o deslocamento da alça em que se encontra o motivo de ligação ao NADPH, somado ao estiramento de resíduos presentes no final da α -hélice, onde se encontra o motivo de ligação ao FAD. Finalmente, eles ainda citam outras monooxigenases que utilizam o mesmo mecanismo de estabilização de intermediários, como as Baeyer-Villiger monooxigenases (Malito *et al.*, 2004).

FMOs em plantas

A primeira caracterização funcional de uma FMO de origem vegetal foi realizada por Zhao *et al.* (2001), quase quarenta anos após a descoberta dessas enzimas em animais. Estes pesquisadores descobriram o papel de uma FMO, nomeada YUCCA, na via de síntese de auxina dependente de triptofano em *A. thaliana*. Nos anos seguintes, outros genes *YUCCA*-like foram descobertos em *Arabidopsis*, formando uma família gênica com 11 membros (*YUC1* a *YUC11*). Sequências homólogas a genes *YUCCA* foram descobertas em *Petunia x hybrida* (Tobeña-Santamaria *et al.*, 2002), *Oryza sativa* (Yamamoto *et al.*, 2007), *Zea mays* (Gallavotti *et al.*, 2008), entre outras.

Em 2006, três laboratórios identificaram independentemente uma FMO envolvida na defesa contra patógenos em *Arabidopsis* (Bartsch *et al.*, 2006; Koch *et al.*, 2006; Mishina & Zeier, 2006). Por meio da técnica “activation tagging” (Weigel *et al.*, 2000), foi isolado um mutante com alta resistência a bactérias virulentas, bem como avirulentas, e patógenos

oomicetos. A superexpressão de *FMO1* em transformantes estáveis restabeleceu este fenótipo. Além disto, a infecção por patógenos ativou a transcrição de *FMO1*, enquanto o mutante *fmo1* apresentou alta suscetibilidade a microorganismos virulentos (Koch *et al.*, 2006).

Glucosinolatos (GSLs) são metabólitos secundários derivados de aminoácidos que apresentam variadas funções biológicas e ocorrem em plantas pertencentes à ordem Brassicales. Hansen *et al.* (2007) identificaram FMO_{GS-OX1} como a enzima da biossíntese de GSLs alifáticos em *Arabidopsis*, que catalisa a *S*-oxigenação de GSLs methylthioalkyl em GSLs methylsulfinylalkyl, conhecidos por suas propriedades anticancerígenas. Em 2008, o mesmo grupo de pesquisadores identificou outros quatro genes FMO_{GS-OX} de *Arabidopsis* também relacionados com a síntese de glucosinolatos anticancerígenos (Li *et al.*, 2008).

A análise de genes *FMO*-like dos três genomas vegetais completamente seqüenciados revela que as FMOs de *A. thaliana*, *O. sativa* e *Populus trichocarpa* se distribuem em três clados: (i) o Clado I (defesa contra patógenos) contém *FMO1* e um pseudogene de *Arabidopsis*, três FMOs de arroz e seis de *Populus*; (ii) no Clado II (síntese de auxina) encontram-se as proteínas YUCCA, juntamente com oito FMOs de arroz e doze de *Populus*; (iii) o Clado III (*S*-oxigenação) é composto de dezesseis FMOs de *Arabidopsis*, sendo uma delas FMO_{GS-OX1} , sete FMOs de arroz e outras quatro de *Populus* (Hansen *et al.*, 2007; Schlaich, 2007). Além das diferenças funcionais, FMOs de clados distintos apresentam peculiaridades em relação aos motivos protéicos que possuem, havendo grande variação de tipos e localização dos mesmos.

Estrutura e motivos conservados

As FMOs se ligam à adenina do FAD por meio de um domínio conservado do tipo “*Rossmann fold*”, constituído pelos resíduos “GXGXXG”, e localizado próximo ao N-terminal da proteína (Eswaramoorthy *et al.*, 2007). Mutações em qualquer um dos resíduos de glicina

foram correlacionadas com atenuação ou mesmo perda total da atividade enzimática (Kubo *et al.*, 1997). O motivo de ligação ao NADPH, que ocorre em uma região mais central da proteína, é idêntico ao motivo para ligação ao FAD (“GXGXXG”), embora bem menos conservado (Schlaich, 2007). Em algumas FMOs esse domínio apresenta somente o primeiro resíduo de glicina conservado, enquanto os demais foram substituídos por outros resíduos (Choi *et al.*, 2003; Eswaramoorthy *et al.*, 2007; Alfieri *et al.*, 2008). Em direção ao C-terminal, encontra-se o motivo “FATGY”, provavelmente relacionado com proteínas que realizem *N*-oxidação. Este motivo é conservado entre FMOs de animais, levedura e bactéria. Em *Arabidopsis*, ele é encontrado principalmente como “LATGY” nas YUCCA e em FMO1. No caso das FMOs pertencentes ao Clado III, um motivo similar é encontrado, “HCTGYK”, que compartilha os três resíduos centrais “TGY” com o motivo “F/LATGY” (Schlaich, 2007). O motivo “HCTGYK” está localizado antes, enquanto o motivo “F/LATGY” encontra-se depois do chamado motivo identificador de FMOs (“*FMO-identifying motif*”). O motivo identificador de FMOs “FXGXXXHXXXY/F” é encontrado em todas as FMOs de plantas e de outros organismos (Suh *et al.*, 1996; Falls *et al.*, 1997; Tobeña-Santamaria *et al.*, 2002; Choi *et al.*, 2003), embora sua posição ao longo da proteína seja variável.

A família YUCCA

A técnica “activation tagging” (Weigel *et al.*, 2000), que utiliza T-DNA para inserir ativadores de superexpressão (promotor constitutivo CaMV 35S, por exemplo) aleatoriamente em uma determinada sequência genômica, mostrou-se uma ferramenta interessante para o descobrimento de novas funções gênicas em plantas. Foi por meio desta técnica que Zhao *et al.* (2001) desvendaram o papel de uma FMO na síntese de auxina em *Arabidopsis*, marcando a primeira caracterização funcional de uma FMO de origem vegetal. Estes pesquisadores identificaram mutantes com características típicas de plantas com alto teor endógeno de

auxina, como hipocótilos longos, cotilédones com forte epinastia, pecíolos alongados, raiz curta com pêlos radiculares numerosos e compridos. A análise da seqüência genômica adjacente à borda direita do T-DNA inserido possibilitou a identificação do gene superexpresso: tratava-se de um gene *FMO*, nomeado *YUCCA*. Este nome se deve às semelhanças fenotípicas compartilhadas entre o mutante estudado e a planta yucca (*Agave* sp.), como folhas curvadas e hábito de crescimento semi-ereto. Buscas realizadas no BLAST identificaram outros genes da família *YUCCA*, totalizando 11 genes que compartilham de 44 a 66% de identidade em termos de seqüência de aminoácidos. As superexpressões de *YUCCA2* (53% de identidade com *YUCCA*) e *YUCCA3* (51% de identidade com *YUCCA*) também resultaram em fenótipos semelhantes ao fenótipo *yucca*, demonstrando que estes genes são parálogos e provavelmente apresentam funções redundantes. Baseados em ensaios bioquímicos *in vitro* com a proteína *YUCCA* recombinante e experimentos com o fornecimento de um análogo tóxico do triptofano, propôs-se que *YUCCA* catalisa a *N*-oxigenação da triptamina em *N*-hidroxil triptamina, reação pertencente à via de síntese de auxina dependente de triptofano.

Posteriormente, outros membros da família gênica *YUC* foram isolados em *Arabidopsis* e todos se mostraram envolvidos na biossíntese de auxina. Após Zhao *et al.* (2001) terem descrito os três primeiros genes *YUC* (*YUC1*, *YUC2* e *YUC3*), Marsch-Martinez *et al.* (2002) também utilizaram “activation tagging” para caracterizar um gene *YUC* com 66% de identidade com *YUCCA*, cuja superexpressão causava o mesmo fenótipo relativo à alta produção de auxina. Este gene foi nomeado *THREAD/YUCCA4* e apresenta baixo nível de expressão em plantas do tipo selvagem. Woodward *et al.* (2005) isolaram o quinto gene *YUC* de *Arabidopsis*, *SUPER1/YUCCA5*, e observou que este gene não apresentava nenhum intron, o que o diferenciava dos quatro genes *YUC* previamente caracterizados. Enquanto a

superexpressão dos outros genes *YUC* resulta em um fenótipo bastante característico, no caso de *SUPER1/YUCCA* este fenótipo é menos intenso: os hipocótilos são mais curtos e cotilédones não exibem forte epinastia. As diferenças de fenótipos encontradas entre os mutantes de superexpressão de *SUPER1/YUCCA5* e dos outros genes *YUCCA* devem refletir papéis redundantes, porém distintos para cada membro da família ao longo do crescimento e desenvolvimento da planta. Outro gene *YUCCA* apresentou fenótipo com certas características distintas do fenótipo padrão dos mutantes *yucca* em *Arabidopsis*. Kim *et al.* (2007) isolaram os mutantes *hypertall*, referentes à superexpressão do gene *YUCCA6*, e observaram que estes não apresentam raízes curtas ou com muitos pêlos radiculares. Além disso, as inflorescências são extremamente altas (por isso o nome *hypertall*) com elevada dominância apical. Estes pesquisadores também analisaram a atividade funcional *in vitro* da proteína *YUCCA6* recombinante e confirmaram a *N*-oxigenação da triptamina em *N*-hidroxil triptamina.

Para entender melhor a importância dos genes *YUCCA* no desenvolvimento de *Arabidopsis*, Cheng *et al.* (2006) produziram diversos mutantes duplos, triplos e quádruplos com baixos níveis de expressão de *YUCCA1*, 2, 4 e 6, em diferentes combinações. Além disso, superexpressaram e reduziram a expressão de cada um desses genes individualmente. A superexpressão individual de cada um desses genes *YUCCA* resultou em altos níveis de auxina endógenos e fenótipos característicos semelhantes. Por outro lado, a inativação individual não resultou em nenhum fenótipo visível, confirmando a redundância de funções. Interessantemente, eles observaram que inativações duplas de genes que pertencem ao mesmo clado (*YUC1* e *YUC4*; *YUC2* e *YUC6*) causam severos defeitos de desenvolvimento, enquanto nenhum fenótipo marcante foi observado com combinações inativadoras entre dois genes de clados diferentes, sugerindo que genes do mesmo clado apresentam funções redundantes em um nível mais específico, como estádios de desenvolvimento e condições de ambiente. Eles

também constataram que os genes *YUCCA* desempenham papel fundamental no desenvolvimento floral e vascular.

O mesmo grupo de pesquisadores utilizou estratégia semelhante para mensurar a importância dos genes *YUCCA* durante a embriogênese em *Arabidopsis* (Cheng *et al.*, 2007). Eles produziram mutantes duplos e quádruplos com baixos níveis de expressão de todas as combinações entre *YUCCA1*, *4*, *10* e *11*. Os mutantes quádruplos *yuc1 yuc4 yuc10 yuc11* apresentam características incomuns, como ausência de hipocótilo ou raízes, plântulas com apenas um cotilédono e pouco tecido vascular. Estes efeitos ocorrem logo na fase de embrião globular. O embrião também não apresenta hipófise, responsável pela posterior formação do centro quiescente do meristema apical radicular, e o hipocótilo não é formado. Uma característica proeminente do embrião do mutante *yuc1 yuc4 yuc10 yuc11* é apresentar a região basal defeituosa e sem formação de raízes primárias. Estes resultados demonstram que a auxina sintetizada pelos genes *YUCCA* é essencial para a embriogênese em *A. thaliana*.

Os estudos envolvendo os genes *YUCCA*-like não se restringem apenas à planta modelo *A. thaliana*. Tobeña-Santamaria *et al.* (2002) identificaram em *Petunia x híbrida* o gene *FLOOZY* (*FZY*), requerido para a iniciação dos órgãos florais e para o desenvolvimento da vascularização em folhas e brácteas. *FLOOZY* codifica uma FMO com alta homologia às *YUCCA* de *Arabidopsis*. Além de alta similaridade em termos de sequência de aminoácidos, esses genes compartilham estrutura gênica similar, sendo o número e a posição dos introns de *FLOOZY* idênticos aos de *YUCCA1*, *2* e *4*. A superexpressão dos genes *YUCCA* em *Arabidopsis* ou *FLOOZY* em petúnia resulta em altos níveis de auxina, folhas epinásticas, pêlos radiculares abundantes e alongados e crescimento independente de auxina em cultura de tecidos. Todos esses resultados indicam que *FLOOZY* e *YUCCA1/YUCCA2* são ortólogos e que as proteínas codificadas por tais genes catalisam reações enzimáticas semelhantes.

Exposito-Rodriguez *et al.* (2007) identificaram o ortólogo de *FLOOZY* em tomate (*Solanum lycopersicon*), ao qual deram o nome de *ToFZY*. A proteína deste gene de tomate compartilha 87, 70 e 69% de identidade com as proteínas de *FLOOZY*, *YUCA* e *YUCI*, respectivamente. O ensaio *in vitro* de *ToFZY* recombinante evidenciou a capacidade de transformar triptamina em N-hidroxil triptamina, sugerindo participação desta proteína na síntese de auxina *in vivo*.

Em monocotiledôneas os estudos referentes aos genes *YUCCA*-like envolvem principalmente o arroz, única planta deste grupo com o genoma completo seqüenciado. Yamamoto *et al.* (2007) identificaram sete genes *YUCCA*-like no genoma de arroz por meio de buscas no tBLASTn, utilizando a seqüência de *YUCCA1* como referência. Para selecionar quais destes genes estavam diretamente relacionados com a síntese de auxina, eles analisaram o padrão de expressão espacial dos sete genes em coleóptile, cuja extremidade é abundante em auxina, e observaram que três genes eram preferencialmente expressos na ponta dos coleóptiles, mas que *OsYUCCA1* era o gene predominante. Por este motivo estes pesquisadores utilizaram um sistema de expressão induzido pelo hormônio esteróide dexametasona (DEX) para observar os efeitos da superexpressão de *OsYUCCA1* no desenvolvimento de plantas transgênicas regeneradas por cultura de tecidos. Na ausência de DEX as plantas transgênicas apresentaram fenótipo normal, enquanto a adição de DEX causou o desenvolvimento abundante de pêlos radiculares. Além disto, as raízes destas plantas não responderam normalmente ao gravitropismo e a quantidade de auxina por miligrama de massa fresca era maior em plantas tratadas com DEX, sugerindo participação direta de *OsYUCCA1* na síntese deste hormônio. A produção de plantas transgênicas expressando o cDNA antisense de *OsYUCCA1* sob controle do promotor ACT1 resultou em plantas anãs com severa inibição da formação e alongamento de raízes. Todas estas informações corroboram a importância de *OsYUCCA1* para o crescimento e desenvolvimento de arroz.

Gallavotti *et al.* (2008) identificaram e caracterizaram um mutante de milho (*sparse inflorescence1 - spi1*) que apresenta defeitos na iniciação de meristemas axilares e órgãos laterais durante o desenvolvimento vegetativo e da inflorescência. O gene *SP11* codifica uma flavina monooxigenase similar às proteínas *YUCCA* de *Arabidopsis*. No entanto, ao contrário de *Arabidopsis*, em que a inativação de ao menos dois genes *YUCCA* é necessária para a observação de um fenótipo, os mutantes *spi1* por si só apresentam defeitos marcantes de desenvolvimento. Análises filogenéticas sugerem que isto pode ocorrer devido à aquisição de um papel específico no desenvolvimento da inflorescência por parte de *SP11*, juntamente com a expansão independente da família *YUCCA* entre monocotiledôneas e eudicotiledôneas.

Embora a auxina seja conhecida há muito tempo como um importante hormônio regulador do crescimento e desenvolvimento vegetal, uma série de características da sua via de biossíntese permanecem obscuras. A descoberta dos genes *YUCCA*, bem como a caracterização funcional de alguns membros desta família, proporcionou um grande avanço no conhecimento acerca dos mecanismos envolvidos na produção da auxina. De fato, um melhor entendimento da biossíntese da auxina e da regulação do crescimento de plantas será conseguido após a integração de dados resultantes da análise de outros genes *YUCCA* em *Arabidopsis* e, especialmente, por meio da caracterização de genes *YUCCA*-like em outras espécies de plantas.

Escolheu-se trabalhar com a espécie *Coffea arabica* porque i) é a principal espécie comercial de café, ii) apresenta um extenso banco de ESTs com informações ainda pouco exploradas, iii) representa a principal espécie estudada pelo grupo de pesquisa e iv) não possui nenhuma FMO caracterizada até o momento. A escolha deste grupo de enzimas justifica-se pelo pouco conhecimento acerca das funções biológicas das FMOs em plantas. Poucos substratos endógenos foram identificados, sendo limitados a um número restrito de espécies

vegetais. Outras caracterizações funcionais envolvendo FMOs de diferentes espécies são necessárias para determinar a real especificidade dessas enzimas por potenciais substratos naturais, o que constitui o principal objetivo do presente trabalho.

2. Objetivo Geral

O objetivo do trabalho foi realizar a caracterização funcional *in vitro* de uma flavina monooxigenase (FMO) recombinante de *Coffea arabica* que apresenta alta similaridade com proteínas YUCCA de *A. thaliana*.

Objetivos Específicos

- Clonagem, expressão e purificação de uma FMO recombinante de *C. arabica*;
- Caracterização funcional *in vitro* da proteína recombinante;
- Determinação do padrão de expressão espacial do gene estudado por RT-PCR semi-quantitativa;
- Análise filogenética de proteínas YUCCA-like.

3. Material e Métodos

Material vegetal

Com exceção das raízes, todo o material vegetal utilizado nos experimentos foi obtido a partir de uma única planta de *Coffea arabica* var. Mundo Novo cultivada na casa de vegetação do Depto de Fisiologia Vegetal da UNICAMP, não tendo sido feito nenhum controle de temperatura ou umidade. As raízes utilizadas nas análises de expressão foram adquiridas de plantas cultivadas em hidroponia, após a germinação de sementes oriundas do indivíduo descrito acima.

Aquisição da seqüência gênica e análises com ferramentas de bioinformática

Foi feita uma busca por palavras-chave (“flavin monooxygenase”) no Banco de ESTs do Projeto Genoma Café (<http://www.lge.ibi.unicamp.br/cafe>) e após a clusterização foram formados 8 contigs, entre os quais apenas um estava completo. O unigene (contig) C8 foi selecionado e submetido a uma série de análises com programas de bioinformática para a determinação das estratégias de clonagem, expressão e purificação da proteína recombinante.

A maioria das análises foi realizada com algoritmos das plataformas Expasy (www.expasy.org) e NCBI (www.ncbi.nlm.nih.gov). Primeiramente, identificou-se a ORF (*Open Reading Frame*) do unigene C8 utilizando concomitantemente os programas *ORF finder*, do NCBI (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), e *Translate*, do Expasy (<http://www.expasy.org/tools/dna.html>). Esta seqüência foi então submetida à comparação com outras seqüências depositadas no NCBI usando-se o módulo blastp, que compara a proteína de interesse com um banco de dados de proteínas depositadas, bem como detecta domínios conservados na seqüência submetida. A identificação de motivos conservados característicos de FMOs foi realizada manualmente com base em Schlaich (2007) e Exposito-Rodriguez *et al.* (2007) e permitiu a classificação de C8 em um dos três clados determinados

para FMOs de plantas. Com a análise do ProtParam (<http://www.expasy.org/tools/protparam.html>) computou-se uma série de parâmetros físico-químicos da proteína teórica, como peso molecular, pI teórico, composição de aminoácidos, coeficiente de extinção, entre outros. A predição da topologia da proteína e identificação de possíveis hélices transmembranares foi realizada com os programas TMHMM (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) e HMMTOP (<http://www.enzim.hu/hmmtop/index.html>). A presença e localização de peptídeo sinal foram preditas com o programa SignalP (<http://www.cbs.dtu.dk/services/SignalP/>). A provável localização subcelular foi determinada com os algoritmos TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) e WoLF PSORT (<http://wolfsort.org/>). Finalmente, o algoritmo PredictProtein (<http://www.predictprotein.org/>) foi usado para uma análise diversificada da seqüência protéica, gerando informações como presença de motivos de acordo com o PROSITE, regiões de baixa complexidade, predição de estrutura secundária, regiões globulares e transmembranares, pontes dissulfeto, entre muitas outras.

Alinhamentos e análise filogenética

As seqüências de aminoácidos referentes a proteínas YUCCA de *A. thaliana* foram adquiridas no banco de dados de proteínas do NCBI (www.ncbi.nlm.nih.gov). Os números de acesso das seqüências utilizadas são: YUCCA1, At4g32540; YUCCA2, At4g13260; YUCCA3, At1g04610; YUCCA4, At5g11320; YUCCA5, At5g43890; YUCCA6, At5g25620; YUCCA7, At2g33230; YUCCA8, At4g04610; YUCCA9, At1g04180; YUCCA10, At1g48910; YUCCA11, At1g21430. A proteína FMO1 (At1g19250) de *A. thaliana*, pertencente a um clado distinto das YUCCA (Schlauch, 2007), foi escolhida como grupo externo. Essas seqüências, juntamente com a seqüência de aminoácidos referente à CaFMO8, foram alinhadas no programa BioEdit 7.0.5.3 (Tom Hall, Ibis Biosciences, Carlsbad, CA,

USA) e salvas no formato FASTA. Este arquivo foi convertido para o formato MEG no programa MEGA 4.0 (Tamura *et al.*, 2007) e uma árvore filogenética foi construída neste próprio programa com base no alinhamento das seqüências protéicas. Foi utilizado o método de Neighbor-joining com bootstrap de 1000 replicações e modelo p-distance.

Após estas análises, realizou-se um alinhamento no BioEdit entre a FMO de café e as duas proteínas YUCCA de *Arabidopsis* mais próximas. O arquivo foi salvo no formato FASTA e utilizado no ESPript 2.2 (<http://esript.ibcp.fr/>), uma ferramenta online que permite melhorar características estéticas de alinhamentos de seqüências. Na opção “execute”, o arquivo FASTA gerado no BioEdit foi utilizado e submetido à análise do ESPript, após a seleção das opções de “output”.

Isolamento de RNA e análises de expressão

Para determinar o padrão de expressão espacial do unigene C8, o RNA total de diferentes partes da planta foi extraído com perclorato de sódio (Rezaian and Krake, 1987). Os materiais vegetais utilizados foram: raízes, folhas jovens, botão floral, fruto estágio F2 (30 dias após a florada) e fruto estágio F4 (45 dias após a florada). A quantificação foi feita pela medição da absorbância a 260 nm e a integridade do RNA foi analisada por meio de gel de agarose 1%, em tampão MOPS 1X e formaldeído 1,85%. DNA contaminante foi eliminado por tratamento com DNAase “TURBO DNA-free” (Ambion, Inc), que utilizou 2 µg de RNA total, sendo o volume recuperado utilizado diretamente na síntese da primeira fita de cDNA pelo kit “SuperScript First-Strand Synthesis System for RT-PCR” (Invitrogen). Actina foi utilizada como gene normalizador (Melo *et al.*, 2006) utilizando-se os primers AcF1 (5'-GACCTCACAGATCACCTCAT-3') e AcR1 (5'-GTAGTCTCGTGGATAACCACG - 3'), gentilmente cedidos pela Dra Miriam Peres Maluf (Instituto Agronômico de Campinas). As reações continham 1,5 mM de MgCl₂, 0,2 µM de cada primer, 1U de Taq DNA polimerase

(Fermentas), 0,2 mM de dNTPs, variando-se a quantidade de cDNA conforme necessário e acertando o volume final da reação para 25 µl. A programação do termociclador para estas reações foi: desnaturação inicial a 94°C por 3 minutos; trinta e cinco ciclos envolvendo desnaturação a 94°C por 30 segundos; anelamento a 53°C por 30 segundos; extensão a 72°C por 1 minuto. Para a normalização, reações foram retiradas após 23, 28 e 30 ciclos, garantindo a comparação da expressão antes da saturação. O resultado foi avaliado em gel de agarose 1% com brometo de etídio 0,5 mg/ml. Para as análises de expressão de C8, foram desenhados primers específicos do unigene C8 usando o programa Primer3 v. 0.4.0 (<http://frodo.wi.mit.edu/>), de modo a gerar um amplicon de 412 bp, apresentando as seguintes seqüências: FMO_Foward 5' – ATTTGGGTTCTTCCGATTCC – 3' e FMO_Reverse 5' – GGCCTTTTGATTCCGTACCT – 3'. As reações foram idênticas àquelas descritas para os primers de actina. A programação do termociclador para a reação com estes primers foi: desnaturação inicial a 94°C por 3 minutos; 30 ou 35 ciclos envolvendo desnaturação a 95°C por 30 segundos; anelamento a 50°C por 30 segundos; extensão a 72°C por 1 minuto. A comparação da expressão foi feita tanto após 30 quanto 35 ciclos, visto que mesmo depois de tantos ciclos, alguns tecidos não apresentaram sinais de saturação.

Produção da proteína CaFMO8 recombinante

Vetores de expressão e desenho de primers: a seleção dos vetores de expressão baseou-se em (i) características da seqüência protéica, principalmente hidrofobicidade e regiões transmembranares; (ii) disponibilidade de vetores no laboratório e (iii) utilização de diferentes caudas de afinidade e solubilidade, visando alternativas durante a expressão e purificação. Assim, dois vetores foram selecionados, pETMBP e pETNusA (ambos gentilmente cedidos pelo Dr. Marcelo J. Murai, Laboratório de Fisiopatologia Médica, UNICAMP), que adicionam respectivamente as caudas 6xHis-MBP (maltose-binding protein) e 6xHis-NusA (N utilization

substance A) à porção N-terminal da proteína recombinante. Em ambos os casos, a cauda adicionada constitui um sistema combinatório de afinidade (6xHis) e aumento de solubilidade (MBP e NusA). A cauda de MBP, além de aumentar a solubilidade também pode ser utilizada como cauda de afinidade. O programa WebCutter (<http://rna.lundberg.gu.se/cutter2/>) foi utilizado para gerar um mapa de restrição da seqüência do ORF C8 e possibilitou a determinação das endonucleases utilizadas para a ligação do fragmento aos vetores de expressão. NcoI e XhoI foram as enzimas de restrição escolhidas para as digestões do fragmento e dos vetores.

Com esses parâmetros estabelecidos, primers específicos para o unigene C8 foram desenhados no programa GeneRunner 3.05 (Hastings Software Inc., Hastings-on-Hudson, NY, USA) e analisados quanto à porcentagem de GC, temperatura de melting e complementaridade entre pares no BioEdit 7.0.5.3 (Tom Hall, Ibis Biosciences, Carlsbad, CA, USA) e Primer3 v. 0.4.0 (<http://frodo.wi.mit.edu/>), bem como no próprio GeneRunner. Em vista da escolha das mesmas enzimas de restrição para ambos os vetores utilizados, um único par de primers que adicionariam os sítios de tais enzimas foi desenhado, com as seguintes seqüências: Foward_NcoI 5'-CCATGGCAACCAAAGAACAAGT-3' e Reverse_XhoI 5'-AAACTCGAGCTATGTTTGGGCATCCAA-3'.

Extração de RNA total e síntese de cDNA: RNA total de folhas jovens *C. arabica* var. Mundo Novo foi extraído pelo método TRIzol (Invitrogen), de acordo com o protocolo do fabricante. A quantificação foi feita pela medição da absorbância a 260 nm (Ultrospec 1000, Pharmacia Biotech). O comprimento de 280 nm foi usado para analisar a quantidade de proteínas e determinar o grau de pureza do ácido nucléico. A integridade do RNA foi analisada por meio de gel de agarose 1% desnaturante, com tampão MOPS 1X e formaldeído 1,85%. Um volume referente a 5 µg de RNA total foi tratado com DNase “TURBO DNA-free” (Ambion, Inc.) e

posteriormente usado como template para a síntese da primeira fita cDNA com o kit “SuperScript First-Strand Synthesis System for RT-PCR” (Invitrogen), de acordo com o protocolo do fabricante. O cDNA total produzido foi diluído dez vezes para a utilização nas RT-PCRs subsequentes. A confirmação da síntese foi feita por RT-PCR utilizando os primers de actina (ver acima), em reações idênticas às descritas previamente, e o resultado analisado em gel de agarose 1% com brometo de etídio. A programação do termociclador para tais reações foi: desnaturação inicial a 95°C por 3 minutos; trinta e cinco ciclos envolvendo desnaturação a 95°C por 30 segundos; anelamento a 55°C por 30 segundos; extensão a 72°C por 1 minuto.

Clonagem e seqüenciamento: Após a validação do cDNA, os fragmentos da região codificadora de C8 foram produzidos por RT-PCR com o kit “Platinum PCR High Fidelity Supermix” (Invitrogen), de acordo com determinações do fabricante e utilizando 0,2 µM de cada primer. A programação do termociclador foi: desnaturação inicial a 94°C por 3 minutos; 35 ciclos envolvendo desnaturação a 94°C por 30 segundos; anelamento a 51°C por 30 segundos; extensão a 68°C por 1 minuto e 30 segundos; extensão final a 68°C por 15 minutos. As reações foram imediatamente purificadas com kit “illustra GFX PCR DNA and Gel Band Purification Kit” (GE Healthcare) e cerca de 50 ng do inserto purificado foi utilizado para a ligação ao vetor de clonagem “pGEM T-easy Vector System” (Promega). Bactérias DH5α termocompetentes (Novagen) foram transformadas com o plasmídeo recombinante, plaqueadas em meio LB semi-sólido contendo 100 µg/ml de ampicilina, 0,5 mM de IPTG e 80 µg/ml de XGal (protocolo “pGEM T-easy Vector System”, Promega), e incubadas a 37°C “overnight”. Colônias brancas foram selecionadas para a minipreparação de plasmídios pelo método de lise alcalina (modificação e adaptação de “PureLink Quick Plasmid Miniprep Kit”, Invitrogen). A presença do inserto foi confirmada por uma reação de restrição com EcoRI (Promega), que

corta o vetor pGEM T-easy em ambas as extremidades, liberando o fragmento inserido. A identidade do fragmento foi confirmada pelo seqüenciamento com primers do vetor (T7 promoter e SP6 promoter), e dois pares distintos de primers internos: o primeiro referente aos primers utilizados para a produção do fragmento a ser clonado, e o segundo referente aos primers usados na análise do padrão de expressão do gene em diferentes partes da planta. As reações foram realizadas com “BigDye Terminator v3.1 Cycle Sequencing Kit” (Applied Biosystems) de acordo com as instruções do fabricante, seguidas de precipitação com etanol 70%, e os produtos separados pelo seqüenciador ABI Prism 377 (Applied Biosystems). Todos estes processos relativos ao seqüenciamento foram realizados no Laboratório de Análise Genética Molecular, do Centro de Biologia Molecular e Engenharia Genética (CBMEG) da UNICAMP.

Expressão de FMO recombinante em Escherichia coli: Fragmentos devidamente validados pelo seqüenciamento foram utilizados para a ligação aos vetores de expressão pETMBP e pETNusA. A digestão dos vetores e do fragmento clonado em pGEM T-easy foi realizada com enzimas FastDigest NcoI e Fast Digest XhoI (Fermentas) a 37°C por 1 hora. Plasmídios e fragmentos digeridos foram purificados do gel de agarose por meio de “illustra GFX PCR DNA and Gel Band Purification kit” e posteriormente ligados com T4 DNA ligase (Promega) a 4°C “overnight”. Bactérias *E. coli* “One Shot Mach1 – T1” (Invitrogen) foram transformadas por choque térmico, plaqueadas em LB semi-sólido contendo kanamicina a 50 ug/ml e incubadas a 37°C “overnight”. Cinco colônias isoladas de cada construção foram escolhidas para minipreparação de plasmídios pelo método de lise alcalina. A identificação de clones positivos foi feita por digestões duplas com NcoI e XhoI que liberam o fragmento clonado, a 37°C por 1 hora. Os plasmídios recombinantes foram introduzidos em bactérias *E. coli* BL21(DE3) “One Shot” (Invitrogen), linhagem própria para a expressão de proteínas. Devido

às dificuldades encontradas para a produção de proteínas recombinantes solúveis, uma alternativa para comparação de resultados foi transformar pETMBP carregando o gene *FMO* em bactérias “Rosetta (DE3) Competent Cells” (Novagen), linhagem suplementada com T-RNAs para códons raramente utilizados por *E. coli*. Em vista desta característica, é necessário adicionar cloranfenicol ao meio de cultura, em concentração final de 34 µg/ml. Como controle para expressão e para os futuros testes enzimáticos, os vetores fechados (sem o gene *FMO* de café) de pETMBP e pETNusA foram introduzidos respectivamente em Rosetta (DE3) e BL21(DE3). Diferentemente de outros vetores, cuja indução da expressão do plasmídeo fechado culmina somente na produção da cauda, pETMBP e pETNusA foram desenvolvidos para expressar a proteína GFP (Green Fluorescent Protein) fusionada à cauda.

Vários procedimentos foram realizados para otimização dos parâmetros de expressão, visando maior produção de proteínas recombinantes solúveis. Os parâmetros analisados foram temperatura de indução (de 37°C até 18°C), tempo de indução (de 3h a 18 horas), concentração do indutor (de 1mM à 0,2 mM), tipo de indutor (IPTG ou lactose) e linhagem bacteriana (BL21 DE3 e Rosetta DE3). Os resultados desses testes não foram adicionados, permanecendo somente os dados referentes aos valores escolhidos para cada parâmetro.

Um teste em pequena escala foi realizado para avaliar a capacidade e eficiência da expressão. Para tal, 500 µl de pré-inóculo foi adicionado a 50 ml de LB líquido contendo 50 µg/ml de kanamicina e, no caso de Rosetta, 34 µg/ml de cloranfenicol, e incubado a 37°C, sob agitação de 200 rpm até OD₆₀₀ atingir 0,6. Após este período, 20 ml de meio foram coletados e centrifugados a 5000 rpm, 4°C por 10 minutos para constituir o tempo zero (controle). Ao restante de meio, aplicou-se o indutor de expressão IPTG (isopropyl-β-D-thiogalactopyranoside) em concentração final de 0,5 mM, seguido de incubação a 25°C, sob agitação de 200 rpm. Após 5 horas, 10 ml de meio foram coletados e centrifugados nas

mesmas condições descritas acima. Os pellets foram ressuspensos em 500 µl de tampão de lise (Tris-HCl 50 mM pH 7,5, NaCl 300 mM) e tratados com lisozima (concentração final: 1 mg/ml) por 30 minutos, na presença de 1 mM do inibidor de proteases PMSF (phenylmethylsulfonyl fluoride). Para garantir o rompimento das células, a solução foi sonicada (Ultrasonic Homogenizer 4710, Cole-Parmer Instruments) utilizando ponteira fina e pulsos de 15 segundos, separados por intervalos de 1 minuto. Este procedimento foi feito devido à considerável instabilidade térmica apresentada por proteínas FMOs (Cashman, 2005). Seguiu-se com a centrifugação a 15000 rpm, 4°C por 20 minutos, para separação da fração solúvel (sobrenadante) e da fração insolúvel (pellet). O resultado do teste foi avaliado por SDS-PAGE de 10% (Laemmli, 1970). As amostras foram desnaturadas pela adição de igual volume de tampão de desnaturação (Tris-HCl 0,5M pH 6,8, glicerol, β-mercaptoetanol 2%, azul de bromofenol 0,001g/ml e SDS 0,02g/ml) e incubação em água fervente por 5 minutos. Ao final da corrida, o gel foi corado em solução de Coomassie Blue 2,7% por 30 minutos e posteriormente descorado em solução de metanol 30% e ácido acético 10%. Após digitalização, o gel foi conservado em solução de metanol 5% e ácido acético 7% em temperatura ambiente.

Um segundo teste foi realizado para verificar como a solubilidade da proteína recombinante poderia ser influenciada por (i) pH do tampão de lise, (ii) força iônica, (iii) detergentes e (iv) aditivos. Todos os procedimentos de indução, expressão e rompimento celular, bem como a resolução por eletroforese foram idênticos aos descritos para o primeiro teste de expressão. No entanto, dez tampões de lise distintos foram preparados (Current Protocols in Protein Science, 2007), cujas composições podem ser observadas na tabela 1.

Tabela 1: Composição dos tampões de lise.

	Bis-Tris pH 6,8 (mM)	Tris-HCl pH 8,0 (mM)	NaCl (mM)	Arginina (M)	Glicina (M)	Triton X-100 (%)
1	20	-	150	-	-	-
2	20	-	150	0,1	-	-
3	20	-	150	-	0,1	-
4	20	-	500	-	-	0,1
5	20	-	500	-	-	-
6	-	20	150	-	-	-
7	-	20	150	0,1	-	-
8	-	20	150	-	0,1	-
9	-	20	500	-	-	0,1
10	-	20	500	-	-	-

As induções posteriores levaram em consideração os resultados obtidos no teste de solubilidade, sendo a alteração da força iônica de 300 mM para 500 mM de NaCl o único parâmetro modificado.

Purificação da FMO recombinante: Um volume de 2 ml de pré-inóculo foi adicionado à 500 ml de meio LB líquido contendo 50 µg/ml de kanamicina e, no caso de MBP-6xHis:FMO/GFP em Rosetta (DE3), 34 µg/ml de cloranfenicol, e incubado a 37°C, 300 rpm até a OD₆₀₀ atingir 0,6. Posteriormente, a expressão foi induzida com 0,4 mM de IPTG, e o meio foi incubado a 25°C, 300 rpm por 4 horas. Ao final deste período, o meio de cultura foi centrifugado a 5000 rpm, 4°C por 10 minutos e o pellet ressuspendido em 10 ml de tampão de lise contendo Tris-HCl 20 mM pH 8,0, NaCl 500 mM, lisozima 1 mg/ml e PMSF 1 mM. Após incubação em gelo por 30 minutos, a mistura foi sonicada e centrifugada a 15000 rpm, 4°C por 25 minutos, para separação da fração de proteínas solúveis.

Cromatografia por afinidade foi utilizada para purificar a proteína recombinante fusionada com ambas as caudas utilizadas no trabalho. No caso das construções com 6xHis-NusA, utilizou-se IMAC (Immobilized Metal Affinity Chromatography) em resina de níquel Ni-NTA (Qiagen), cujo princípio de funcionamento baseia-se na afinidade dos resíduos de

histidina pelo níquel imobilizado na resina, enquanto as construções 6xHis-MBP utilizaram a afinidade da MBP pela resina de amilose (New England). Desse modo, cerca de 1 ml da respectiva resina foi empacotada em colunas específicas para purificação por gravidade (BioRad) e equilibrada com 10 volumes de coluna (CV) de tampão A (Tris-HCl 20 mM pH 8,0 e NaCl 500 mM). Os extratos totais de proteínas solúveis foram carregados nas respectivas colunas e o *flow through* coletado. Para eliminação do excesso de proteínas contaminantes, as colunas foram lavadas com 20 CV de tampão A, sendo todo esse volume também coletado para análise em eletroforese. A eluição das proteínas recombinantes variou de acordo com a construção, seguindo a tabela 2.

Tabela 2: Série de Eluições para cada cauda

6xHis-MBP:(FMO/GFP)		6xHis-NusA:(FMO/GFP)	
1)	1/5CV x 5 Tampão A + 10 mM Maltose	1)	10CV Tampão A + 5 mM imidazole
2)	1/5CV x 6 Tampão A + 10 mM Maltose	2)	8CV Tampão A + 20 mM imidazole
3)	1/5CV x 4 Tampão A + 10 mM Maltose	3)	2CV Tampão A + 50 mM imidazole
		4)	2CV Tampão A + 100 mM imidazole
		5)	2CV Tampão A + 250 mM imidazole

Os vetores fechados, que expressam GFP, foram utilizados como controle e as proteínas 6xHis-MBP:GFP e 6xHis-NusA:GFP foram expressas e purificadas pelo mesmo procedimento descrito para as FMOs recombinantes fusionadas. O resultado do teste foi avaliado em SDS-PAGE de 10%, utilizando as mesmas especificações descritas para o teste de expressão.

Análise da atividade enzimática da FMO recombinante: Os ensaios enzimáticos *in vitro* envolvendo a FMO recombinante de *C. arabica* foram realizados com base em estudos previamente realizados de caracterizações funcionais de proteínas YUCCA-like em *Arabidopsis* (Zhao *et al.*, 2001; Kim *et al.*, 2007) e tomate (Exposito-Rodriguez *et al.*, 2007). Para tal, utilizou-se a mesma metodologia de expressão e rompimento celular descrita previamente, mudando-se apenas o tampão de lise de Tris-HCl 20 mM pH 8,0 e NaCl 500 mM para fosfato de potássio 100 mM pH 8,0 e NaCl 500 mM.

Semelhante à Exposito-Rodriguez *et al.* (2007), o extrato total de proteínas solúveis foi utilizado para os ensaios enzimáticos. No entanto, ao invés da diálise, optou-se por passar o extrato por uma coluna PD-10 (Amersham) para a retirada do sal, procedendo a eluição com tampão fosfato 100 mM pH 8,0. A quantificação de proteínas foi feita com reagente pronto da BioRad (Bradford, 1976), com leitura no espectrofotômetro (Ultrospec 1000, Pharmacia Biotech) no comprimento de onda de 595 nm. Reações contendo 50 μ M FAD, 100 μ M NADPH, 2 mM triptamina e 300 μ g de proteínas totais foram completadas com tampão fosfato de potássio 100 mM pH 8,0 para um volume final de 500 μ l. Para o controle, as FMOs fusionadas foram substituídas pelas GFPs produzidas pelos vetores fechados. Foram realizadas cinco repetições para cada construção. Após incubação a 37°C por 3 horas, as reações foram analisadas diretamente em HPLC Shimadzu com detector de UV (diode array detector SPD-M10A), com coluna de fase reversa Varian Inertsil ODS-2 (5 μ , 250 mm x 4,6 mm) e fase móvel composta por ácido acético:acetonitrila (5:5 v/v). O fluxo foi regulado para 1,0 ml/min e absorvância em 280 nm. O padrão de triptamina 100 mM foi preparado em água com triptamina-HCl (Sigma-Aldrich), sendo utilizado para calibração do sistema e quantificação da triptamina de cada reação. As análises estatísticas foram realizadas no programa Sisvar 5.0 por meio do teste de Tuckey ($p \leq 0,05$).

4. Resultados

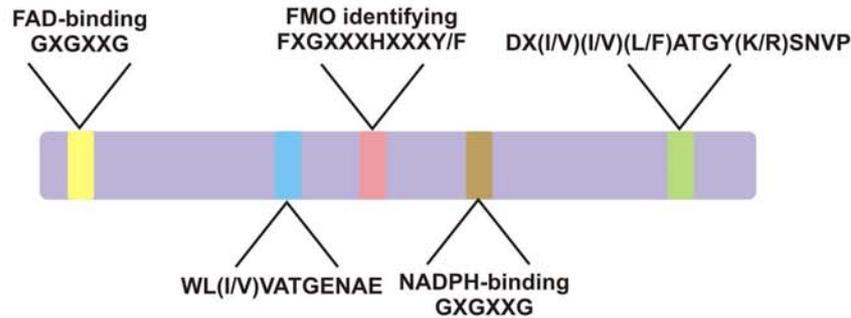
Determinação da ORF e da seqüência traduzida

A seqüência do contig C8 obtida do banco de dados do Projeto EST-Genoma Café, apresenta 1505 bp, das quais 1125 bp são referentes à região codificadora da putativa proteína. As análises da ORF de C8 mostraram que esta seqüência codifica uma proteína de 374 aminoácidos (figura 1A), que apresenta alta similaridade com FMOs da família YUCCA de *Arabidopsis thaliana*, gerando valores de e-value no Blastp de 2e-109, 8e-99 e 3e-93 em relação a YUCCA10 (At1g48910), YUCCA11 (At1g21430) e YUCCA2 (At4g13260), respectivamente. Similar a YUCCA, a FMO de café (*CaFMO8*) apresenta os motivos de ligação ao FAD (GAGPSG) e NADPH (GSGNSG), inclusive em posições semelhantes ao longo da seqüência de aminoácidos (figura 1B). O motivo identificador de FMOs (FXGXXXHXXXY) e o motivo “FATGY” também estão presentes. No caso deste último, (Exposito-Rodriguez *et al.*, 2007) propuseram uma definição mais apurada para seqüências pertencentes à família YUCCA: DX(I/V)(I/V)(L/F)ATGY(K/R)SNVP. Além disto, eles descreveram um motivo particularmente conservado em proteínas YUCCA e sugeriram se tratar de um motivo exclusivo de FMOs de plantas: WL(I/V)VATGENAE. Embora com certa variação (DAVIFATGFKRSTN e CLVVATGETSD – os sublinhados indicam as diferenças), foi possível identificar estes dois motivos na seqüência de CaFMO8, reforçando sua proximidade com proteínas da família YUCCA. Esses resultados permitiram classificar a FMO de café como uma proteína pertencente ao clado II de FMOs vegetais (figura 1C) de acordo com Schlaich (2007), sendo provavelmente a primeira proteína YUCCA-like descrita e caracterizada em *C. arabica*.

A) >CaFM08

MATKEQVVIVV**GAGPS**GLATAACLNNLSIPNIVLEREDCFASLWKKYSYDRLLHLHLAKQ
 FCQLPLKFPPTTYPTVPRDQFLRYLDDYVSHFNICPLYQRSVESARYDEAAEAWIVKA
 RNLGSSDSEEMEEYSSK**CLVVATGETSD**AFIPQLEGLNT**YLGEVI****HSTRY**KNGKSYENK
 NVLVV**SGNS****MEI**AFDLSNYGAKTSIAVRSPLHILSRGMVYIGPVLLKYPSLNTVDWL
 VLMLSKLWYGGDLRSYGIKRPEEGPFTMKVKYGYKYPVIDVGTCCQIKSGEIQVLPAVAS
 LGGNDVVFEDGKSYPF**DAVIFATGFKRSTN**KWLQGADDLLTDDGFAKPAFPNNWKGTKG
 LYCAGLAGRGLYGAALDAQT

B)



C)

	FAD-binding	NADPH-binding	?	FMO identifying
	GXXXXG	GXXXXG	FATGY	FXGXXX_HX_XXFY
CLADO III	Atlg12130 (15) V I GAGAA GLV (212) V I GYQSS GSD (281) VHC TGYK (320) PGLS F I G LPS_MA_LQ F FMFE			
	Atlg12160 (14) V I GLGAA GLV (209) V I GSQAS GND (277) VHC TGYK (316) PGI S F I G LPF_MG_LQ F FMFE			
	Atlg12140 (15) V I GAGAA GLV (210) V I GSQAS GAD (281) VHC TGYK (320) PGLS F I G LPW_MT_LQ F FMFE			
CLADO I	FMO1 (15) I I GAGVS GLA (213) V I GFKKSA ID (360) I LATGYD (180) GPEM FQCKVM HS_MDV CKLE			
	pFMO1 (9) I I GAGVS GLA (205) V I GFKKSA ID (329) I LATGYD (173) GPEI FQCKVL HS_MDV SKLQ			
CLADO II	YUCCA (23) I I GAGFS GLA (187) VV GCNS SGME (325) I FATGYK (159) GRKK FSGK I V HA_S EYK SGE			
	YUCCA4 (19) I V GAGPS GLA (181) VV GCNS SGME (319) I LATGYK (154) GLKK F T G PVV_ HTSAVYK SGS			
	YUCCA2 (30) I V SGPS GLA (198) VV GCNS SGME (336) I LATGYK (172) G I P D F G G P I L_ HTSYK SGE			
	YUCCA6 (34) I V GAGPS GLA (202) VV GCNS SGME (340) I LATGYK (174) GMDK F A A A G V V K H T C H Y K TGG			
	YUCCA3 (39) I V GAGPS GLA (210) VV GCNS SGME (348) I LATGYR (184) GLED F G G F V L_ HAGDYK SGG			
	YUCCA7 (34) I V GAGPS GLA (205) VV GCNS SGME (343) I LATGYR (179) GLED F G G D V L_ HAGDYK SGE			
	YUCCA5 (27) I V GAGPS GLA (197) VV GCNS SGME (335) V LATGYR (170) LTTE F E G E V I_ HSCEYK SGE			
	YUCCA9 (27) I V GAGPS GLA (194) VV GCNS SGME (332) V LATGYR (167) LMTE F D G E V I_ HACEYK SGE			
	YUCCA8 (27) I V GAGPS GLA (197) VV GCNS SGME (335) V LATGYR (170) GLSE F S G E V I_ HACDYJ SGE			
	YUCCA10 (7) I V GAGPA GLA (175) VV GCNS SGME (309) V FATGYK (148) G I D T F G G E I M_ HSSEYK SGR			
	YUCCA11 (11) I I GAGPA GLA (178) VV GCNS SGME (311) V FATGYK (151) LVES F Q G K Y L_ HSSEYK NGE			
CaFM08 (12) V V GAGPS GLA (183) VV GSNS SGME (316) I FATGFK (158) GLNTY L G E V I HSTRYK NGK				

Figura 1. Seqüência completa de aminoácidos de CaFM08. Os motivos conservados da família estão destacados com as mesmas cores utilizadas na representação do item B (A). Localização relativa dos motivos conservados nas proteínas YUCCA, pertencentes ao Clado II de FMOs vegetais (B) e alinhamento dos motivos entre FMOs dos três clados distintos, mostrando os resíduos conservados e indicando onde ocorrem as variações (C).

O alinhamento de CaFMO8 com as proteínas YUCCA de *A. thaliana* mostrou que elas compartilham de 43 a 53% de identidade em termos de seqüência de aminoácidos. A análise filogenética realizada com o alinhamento das seqüências protéicas indicou que a YUCCA de café pertence ao clado YUCCA10/YUCCA11 (figura 2), o grupo mais divergente de proteínas desta família (Exposito-Rodriguez *et al.*, 2007).

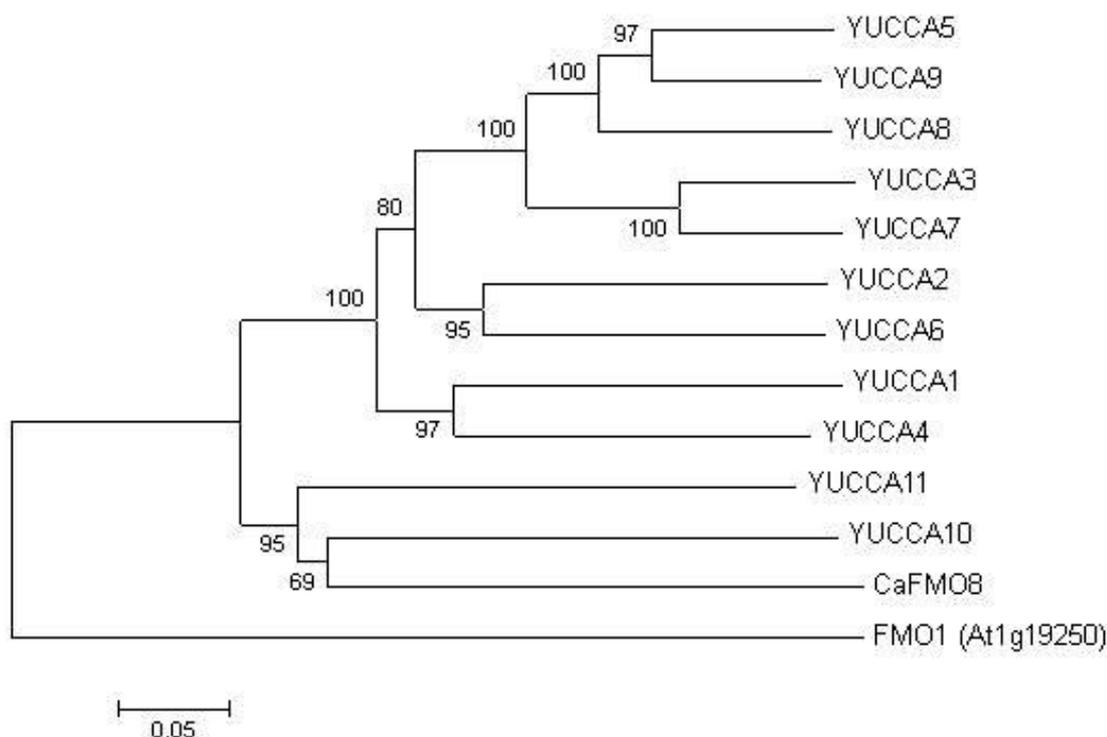


Figura 2. Análise filogenética das proteínas YUCCA de *Arabidopsis* e a FMO de café. A árvore foi construída pelo método de Neighbor-joining, em relação ao alinhamento das seqüências protéicas. A proteína FMO1 (At1g19250) de *A. thaliana* foi utilizada como grupo externo. Os números nos nós dos ramos indicam os valores de bootstrap (1000 repetições).

Já o alinhamento de CaFMO8 com estas duas proteínas YUCCA mais próximas permitiu observar o alto grau de identidade (53% e 51% com YUCCA10 e YUCCA11, respectivamente) entre as seqüências analisadas, bem como identificar a presença e posição dos motivos conservados para a família YUCCA (figura 3).

```

1         10         20         30         40         50         60         70
YUCCA10 METVVVIVGAGPAGLATSVCLNQHSIPNVILEKEDIYASLWKRRAYDRLKLHLAKEFCQLPFMDHGRREVP
YUCCA11 IKILVLIIGAGPAGLATSACLNRLNIPNIVVERDVCSASLWKRRSYDRLKLHLAKQFCQLPHMDFPSNTP
CaFM08  KEQVIVIVGAGPAGLATAACLNNLSIPNIVLEREDCFASLWKRYSYDRLHLHLAKQFCQLPLKRFPTTYP

80         90         100        110        120        130        140
YUCCA10 FMSKELFVNYLDAYVARFDINFRYNRTVKSTFDESKWRVVAENTVTGETEVYWSEFLVVATGENGDGN
YUCCA11 TFVSKLGFINYLDEYATRFNVNFRYNRNVKSAYFKDGQWIVKVVNKTTALIEVYSAKFMVAATGENGEV
CaFM08  TYVPRDQFLRYLDDYVSHFNICDLYQRSVESARYDEAAWIVKARNLGSSEMEVYSSKCLVVATGETSDAF

150        160        170        180        190        200        210
YUCCA10 IPMVEGIDTFGCEIMHSSEYKSCRDFKDKNVLVVGGCNSGMEISFDLCNFGANTTLIRTPRHVVTKEVI
YUCCA11 IPEIFGVESPOCKYLHSSEYKNCEKFAGKDVLVVGCCNSGMEIAYDLSKCNANVSTVVESQVHVLTRCIV
CaFM08  IPQLECLNTYLGEVIHSTRYKNGKSYENKNVLVVGSCNSGMEIAFDLSNYGAKTSIAVRSPLHILSRGMV

220        230        240        250        260        270        280
YUCCA10 HIGMTLLKYAPVAMVDTLVTTMAKILYGDLSKYGLFRPKQGPFATKLFTGKAPVIDVGTVEKIRDGEIOV
YUCCA11 RIGMSLLRFFPVKLVDRLCLLAELRFRNTSRYGLVRPNNGPLNKLITGRSATIDVGCVGEIKSGKIOV
CaFM08  YIGPVLLKYFSLNTVDWLVLMLSKLWYGDLSRYGIKRPEGPFTMKVKYCKYPVIDVGTCQKIKSGEIOV

290        300        310        320        330        340        350
YUCCA10 INGIGSINGKTLTFENGHKQDEFDAIVFATGYKSSVCNWLEDYEYVMKKDGFFKAPMPKHWKGEKNLYCAG
YUCCA11 VTSIKRIEGKTVEFIDGNTKNVDSIVFATGYKSSVSKWLEVDDGLFNENGMEREFPDHWKGKNGLYSAG
CaFM08  LPAVASLGCNDVVFEDCKSYPFDAVIFATGFKRSTNKWLQGADDLLTDDCFAKPAFPNNWKCTKGLYCAG

360
YUCCA10 FSRKCIAGGAEDAMS
YUCCA11 FGKQGLAGISRDARN
CaFM08  LAGRGLYGAALDAQT

```

Figura 3. Alinhamento das seqüências protéicas de YUCCA10, YUCCA11 e CaFM08. Análises filogenéticas sugerem que estas proteínas pertencem ao mesmo clado, dentro da família YUCCA. Os resíduos conservados aparecem sombreados, enquanto os motivos característicos da família estão destacados com as cores da figura 1B. YUCCA10 e YUCCA11 compartilham 53 e 51% de identidade com CaFM08, respectivamente.

Informações relevantes sobre a seqüência deduzida de aminoácidos foram adquiridas com programas de bioinformática, especialmente oriundos da plataforma Expsy. O ProtParam informou que o peso molecular esperado seria 41,3 kDa, o pI teórico 6,95 e coeficiente de extinção 65780. Estes dados são importantes para a identificação da proteína expressa no gel de eletroforese, determinação da influência do pH na extração da proteína, e quantificação. No entanto, como a proteína expressa estaria fusionada à cauda, uma nova análise no ProtParam foi realizada para cada construção, e pode ser vista na tabela 3.

Tabela 3: Características físico-químicas das proteínas recombinantes de acordo com o ProtParam.

Construção	Peso Molecular	pI	Coefficiente de Extinção
<i>6xHis-MBP:FMO</i>	84,4 kDa	5,88	133620
<i>6xHis-MBP:GFP</i>	70,2 kDa	5,58	89730
<i>6xHis-NusA:FMO</i>	98,7 kDa	4,93	98210
<i>6xHis-NusA:GFP</i>	84,4 kDa	4,84	54320

Ambos os programas TMHMM e HMMTOP, que fazem a predição da topologia da proteína, identificaram significativa probabilidade da presença de uma hélice transmembranar na porção N-terminal da FMO submetida. De acordo com Graslund *et al.* (2008), recomenda-se eliminar regiões preditas como transmembranares, devido ao significativo aumento da hidrofobicidade da proteína, causando problemas para expressar a proteína recombinante na fração solúvel. No entanto, optou-se por manter esta porção, pois ela contém o motivo de ligação ao FAD (GXGXXG), seqüência de aminoácidos indispensável para a funcionalidade das FMOs. Na mesma região, o SignalP identificou grande probabilidade da presença de um peptídeo sinal, com possível sítio de clivagem entre os resíduos 21 e 22. Este resultado, bem como a predição da topologia, foi considerado um tanto incomum, visto que indicaria a retirada do motivo de ligação ao FAD da proteína expressa. Mutações nos conservados resíduos de glicina deste motivo foram relacionadas com atenuação ou mesmo perda total da atividade de FMOs (Kubo *et al.*, 1997; Bottoms *et al.*, 2002). Em vista dessas informações da literatura, algumas predições realizadas *in silico* não foram consideradas para a continuidade do trabalho.

A predição da localização subcelular foi determinada pelo WoLF PSORT como citoplasmática e pelo TargetP como proteína excretada, embora ambos tenham apresentado baixa confiabilidade. Poucos estudos relataram a localização subcelular de proteínas YUCCA. Kim *et al.* (2007) produziram plantas transgênicas com expressão constitutiva de YUCCA6 fusionada a GFP, para observar onde esta proteína YUCCA se localizava no interior da célula.

Os resultados obtidos não foram muito conclusivos, e os autores sugeriram que YUCCA6 deve funcionar no citoplasma ou em um compartimento endo-membranar não identificado.

CaFMO8 apresenta expressão preferencial em raízes, folhas e flores

Os níveis de transcrição de *CaFMO8* em diferentes órgãos da planta foi investigado por RT-PCR semi-quantitativa. De modo geral, *CaFMO8* é altamente expresso em raízes, folhas e especialmente em flores imaturas, mas é pouco expresso em frutos (nos dois estádios verificados: F2 e F4), como pode ser observado na figura 4. Estes resultados diferem dos resultados obtidos para *YUC10* e *YUC11* em *Arabidopsis* (Cheng *et al.*, 2007). Apesar de não investigarem o padrão de expressão destes dois genes por RT-PCR, esses autores analisaram os dados de um microarray público (Zimmermann *et al.*, 2004) e observaram *YUC10* é altamente expresso em siliquis, o tipo de fruto produzido por *Arabidopsis*. Posteriormente, eles analisaram o padrão de expressão dos dois genes por hibridização *in situ* durante a embriogênese, e descobriram que eles seguem o padrão de expressão de *YUC1* e *YUC4*. Desse modo, embora as análises filogenéticas indiquem que *CaFMO8* pertence ao clado *YUCCA10/YUCCA11*, aparentemente este gene não segue o padrão de expressão observado na planta modelo *A. thaliana*.

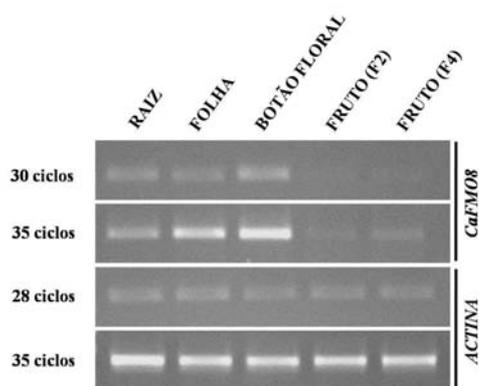


Figura 4. Análises de expressão de *CaFMO8* em diferentes órgãos da planta por RT-PCR semi-quantitativa. A comparação foi feita com 30 e 35 ciclos, visto que uma menor quantidade de ciclos não permitiu visualização de banda no gel. A normalização dos cDNAs totais foi feita com actina (porção inferior).

Produção de CaFMO8 recombinante em E. coli

RNA total e validação do cDNA: RNA total foi extraído de folhas jovens de *C. arabica* e a separação em gel de agarose desnaturante (figura 5A) permitiu observar a integridade do RNA, garantindo uma boa qualidade de template para a síntese do DNA complementar. Após o tratamento do RNA total com DNase promoveu-se a produção da primeira fita de DNA complementar. Esta síntese foi confirmada por meio de RT-PCR utilizando primers de actina. O resultado foi verificado em gel de agarose 1% com brometo de etídio (figura 5B).

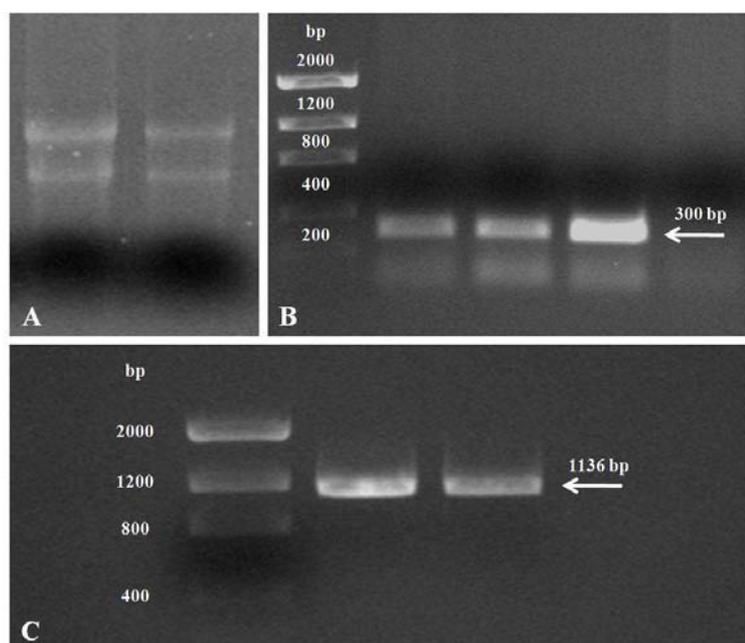


Figura 5. RNA extraído de folhas jovens de *C. arabica* (A), RT-PCR com primers de actina para validação da síntese de cDNA, realizada em triplicata (B) e RT-PCR de alta fidelidade para produção dos fragmentos de CaFMO8 a serem clonados (C). Em destaque, a banda de 1136 bp referente ao gene amplificado.

Clonagem em pGEM T-easy: os fragmentos da região codificadora de *CaFMO8* foram produzidos com alta fidelidade por RT-PCR (figura 5C) e purificados para a clonagem direta em pGEM T-easy, em reação catalisada pela T4 DNA ligase. Bactérias DH5 α foram transformadas com os plasmídios recombinantes para a produção de grande número de cópias do inserto. Após a seleção de colônias brancas e a extração dos plasmídios por miniprep, a

presença do inserto foi confirmada pela reação de restrição com EcoRI, que corta o plasmídio em ambas as extremidades do sítio múltiplo de clonagem, liberando o fragmento clonado (figura 6). As reações enzimáticas referentes aos clones positivos apresentam uma banda de cerca de 1130 bp, referente ao inserto liberado, e outra banda referente ao plasmídio linearizado, com cerca de 3000 bp. Os clones positivos foram submetidos ao seqüenciamento para confirmação da identidade da seqüência clonada. A figura 7 mostra características da cobertura do seqüenciamento, como o sítio múltiplo de clonagem do pGEM T-easy, os primers utilizados e a posição em que estes primers pareiam com a seqüência.

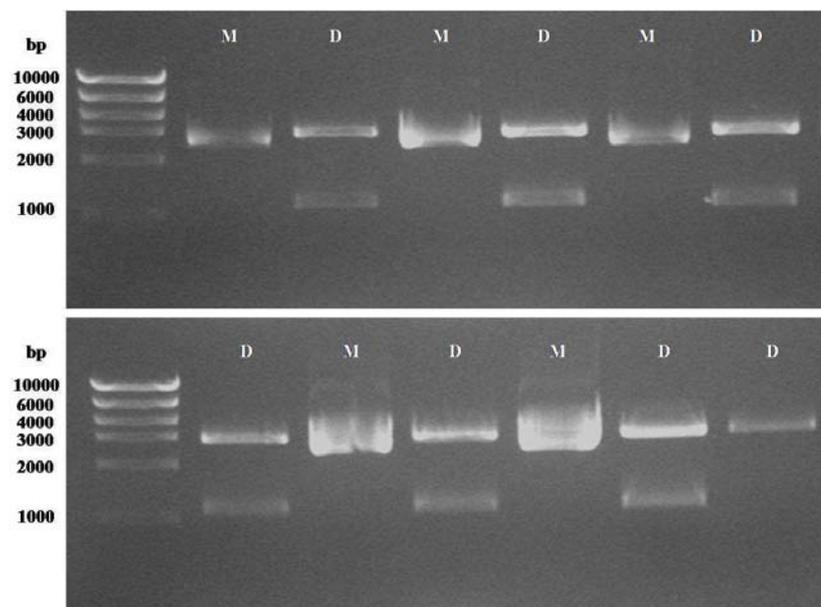


Figura 6. Seleção de clones positivos em pGEM T-easy. A digestão do plasmídio extraído por miniprep com EcoRI libera o fragmento clonado, que apresenta exatamente o tamanho esperado de cerca de 1100 bp. A banda referente ao plasmídio linearizado também apresenta o tamanho esperado de 3000 bp. Legenda: M, miniprep; D, digestão.

A)

5' ... TG **TAATACGACTCACTATAGGG** CGAATTGGGCCCCGACGTCGCATGCTCCCGCCGCCATGGCGCCCGGGGAATTCGATT **ATGCAACCA**
AGAACAAGTGGTAATAGTAGTGGGAGCTGGGCGTCGGGGCTTGCAACTGCAGCATGCTTGAACAACCTTCCATTCCAAACCTAGTCCTTGAG
AGAGAAGATTGTTTTGCTTCCTGTGGAAGAAATATTCCTACGACCGCCTTCATCTTCACTGGCTAAACAGTTCTGTCAACTTCCTCTCAAGCC
GTTCCCGACCACCTACCCGACTTACGTGCCAGAGACCAGTCTTTCGCTTATTTAGACGACTACGTCCTCCATTCAACATTTGTCCCTGTACC
AAAGATCGGTGGAGTCGGCGCGGTACGATGAAGCCGCCGAAGCATGGATTGTCAAGGCGAGGA **ATTGGGTTCTCCGATTCC** GAGGAGATGGAG
GAGTATCCAGCAAGTGCTTGGTGGTGGCCACGGGAGAAACCTCCGATGCATTCATCCCGCAGCTTGAAGGTCTCAACACTTACCTGGCGAGGT
CATCCATAGCACCCGATATAAAAACGGCAAGTCCTATGAAAACAAGAATGTTTTGGTTGTTGGGTCTGGAAATCTGGTATGGAAATTCATTTG
ATCTATCAAACCTACGGTGCAGAACTTCCATCGCCGTCCTAGCCCGCTTACATTTCTGTCAAGGGGAATGGTATACATAGGACCGTTTTGTTG
AAATATTTTTCTCTGAATACGGTTGACTGGCTGGTGTGATGCTGAGCAAGCTGTGGTATGGAGGAGACTTGTC **GGTACCGAATCAAAGGCC**
AGAGGAAGTCCCTTCACCATGAAAGTTAAGTATGGCAAATACCCCGTCATTGACGTTGGAACCTGTCAGAAAATCAAATCCGGCGAGATTCAGG
TACTGCCGCGAGTAGCAAGTCTTGGAGGCAATGATGTGGTGTGAGGATGGCAAGTCTTACCCATTTGATGCTGTTATTTTTGCAACTGGCCTT
AAAAGATCGACGAATAAGTGGCTCCAGGGGGCTGATGATCTCTGACTGACGACGATTGCAAACCTGCCTTCCCTAACCAACTGGAAGGGTAC
CAAGGGATTGACTGCGCTGGTTGGCAGGAAGAAGATTGTATGGAGCTG **CTTTGGATGCCCAAACA** TAGATCACTAGTGAATTCGGCGCCGCT
CGAGTTCGACCATATGGGAGAGCTCCAACCGCTTGGATGCATAGCTTGAGT **ATTCTATAGTGCACCTAAAT** ... 3'

B)

	Primers	Início
Vetor	T7 Promoter (direto)	-78 (bases do vetor)
	SP6 Promoter (reverso)	1125 + 98 (bases do vetor)
Internos 1	Foward_NcoI (direto)	1 →
	Reverse_XhoI (reverso)	←1125
Internos 2	FMO_Foward (direto)	359 →
	FMO_Reverse (reverso)	←770

Figura 7. Cobertura do seqüenciamento do fragmento clonado em pGEM T-easy. Localização relativa do pareamento dos primers na seqüência (A) e lista de primers usados no seqüenciamento, bem como o número do nucleotídeo de início de cada oligo (B). As trincas destacadas em azul indicam os nucleotídeos inicial e final.

Expressão de CaFMO8 recombinante em *E. coli*: os fragmentos clonados em pGEM T-easy foram digeridos com as enzimas de restrição NcoI e XhoI e ligados aos vetores pETMBP e pETNusA previamente linearizados com as mesmas endonucleases. Após a transformação de bactérias com os plasmídios recombinantes, algumas colônias foram selecionadas para identificação de clones positivos. Esta análise foi feita por digestões duplas com NcoI e XhoI que liberam a seqüência presente no sítio de clonagem do vetor (figura 8), revelando uma banda referente ao inserto (1130 bp) e outra referente ao vetor linearizado, com cerca de 7200 bp para pETMBP e 7550 bp para pETNusA. Vetores contendo o fragmento foram então introduzidos em bactérias próprias para a expressão, e os testes de expressão e solubilidade foram realizados para conferir as características da proteína.

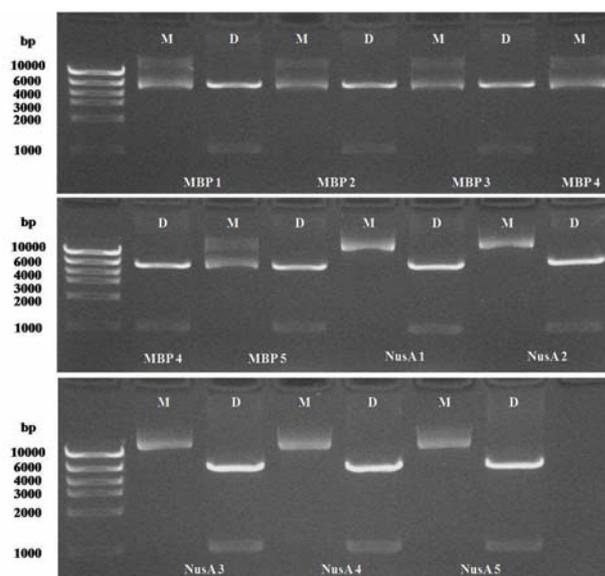


Figura 8. Seleção de clones positivos em pETMBP e pETNusA. Cinco colônias de cada construção foram escolhidas para a minipreparação de plasmídios. A dupla digestão com NcoI e XhoI libera o fragmento clonado em ambos os vetores, gerando uma banda de cerca de 1100 bp referente ao inserto e outra de cerca de 7000 bp para pETMBP e 7500 bp para pETNusA. Como observado na figura, todas as colônias selecionadas para o teste foram positivas. Legenda: M, miniprep; D, digestão.

Um teste em pequena escala mostrou que a proteína foi expressa com peso molecular esperado (figura 9), com ambas as caudas utilizadas.

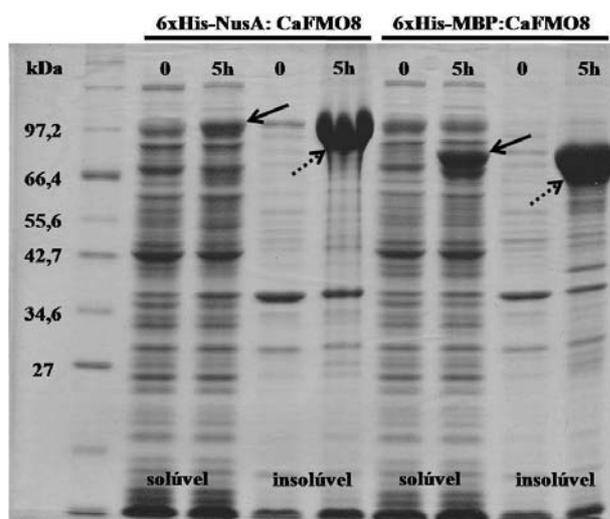


Figura 9. SDS-PAGE de 10% do teste de expressão de CaFM08 fusionada a NusA e MBP após 5 horas de indução com 0,5 mM de IPTG. Nota-se que a maior parte da proteína foi expressa na fração insolúvel (seta pontilhada), enquanto pequena quantidade pode ser identificada referente à proteína solúvel (seta sólida).

Ainda que seja possível observar CaFMO8 recombinante na fração solúvel, a maior parte da proteína expressa apresentou-se na forma insolúvel, tanto fusionada com MBP quanto com NusA. Em vista deste problema, um teste de solubilidade alterando-se diversas variáveis foi realizado, em busca de estratégias que aumentassem a quantidade de proteína solúvel. As variáveis alteradas individualmente no tampão de lise foram pH, força iônica (quantidade de NaCl), aditivos que impedem a agregação de proteínas (arginina e glicina) e detergente (Triton X-100). Resultados significativos foram conseguidos com a utilização de tampão Tris-HCl pH 8,0, maior força iônica e adição de Triton X-100 (figura 10).

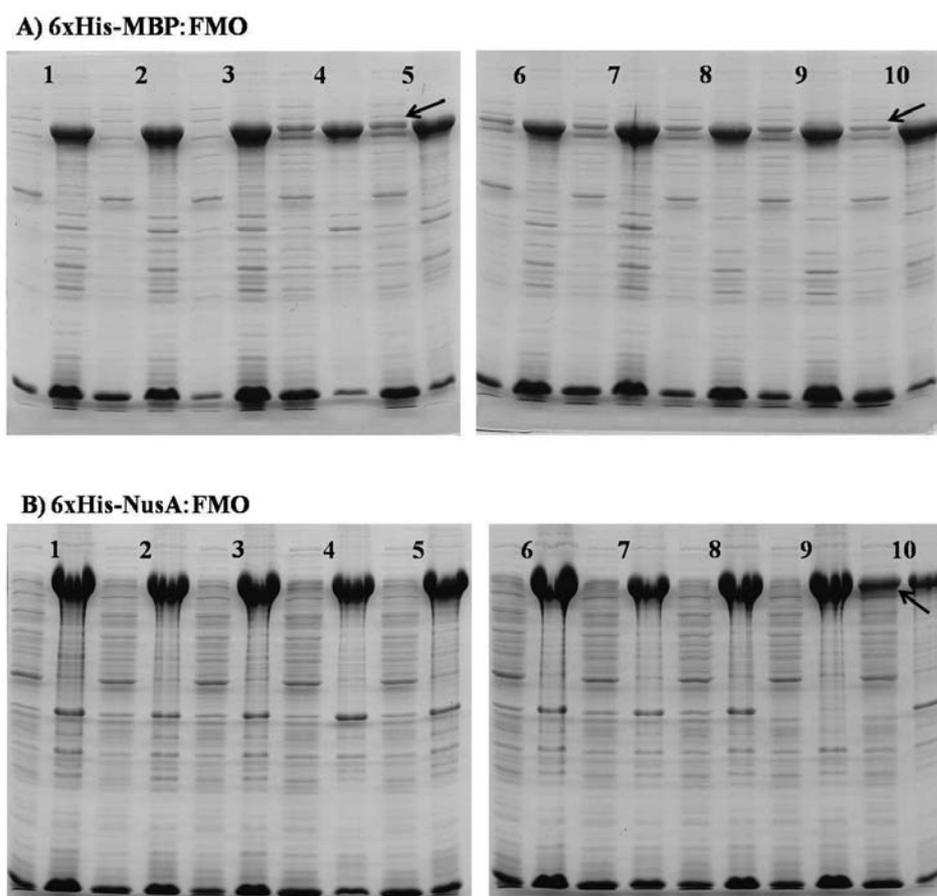


Figura 10. SDS-PAGE de 10% do teste de solubilidade da FMO fusionada às caudas MBP (A) e NusA (B). Para cada número referente a um tampão de lise diferente, o primeiro poço refere-se à fração solúvel, enquanto o segundo refere-se à insolúvel. Os números observados correspondem aos tampões discriminados na tabela 1, página 23. A seta indica a proteína solúvel.

No caso de FMO fusionada à cauda MBP (figura 10A), todos os tampões compostos por Tris-HCl pH 8,0 geraram quantidades de proteína solúvel semelhantes, independentemente da adição de um composto adicional. No entanto, quando o tampão Bis-Tris pH 6,8 foi usado, identifica-se proteína solúvel somente nos tampões de lise com Triton X-100 a 0,1% e com NaCl 500 mM. Já no caso da FMO ligada à cauda NusA (figura 10B), nenhum tampão contendo Bis-Tris pH 6,8 gerou quantidade significativa de proteína solúvel. Por outro lado, o tampão contendo Tris-HCl pH 8,0 e 500 mM de NaCl resultou em grande quantidade de FMO solúvel, constituindo o melhor resultado dentre todos os adquiridos com o teste.

Os dados obtidos no teste de solubilidade permitiram inferir algumas propriedades da proteína expressa. A utilização proveitosa de Triton X-100, um detergente não-iônico, sugere que a proteína realizou interações hidrofóbicas, possivelmente formando agregados e migrando para a fração insolúvel. No entanto, devido ao objetivo de se produzir proteínas cataliticamente funcionais, optou-se por não utilizar Triton X-100 no processo, para se evitar qualquer perda de atividade por parte da enzima recombinante. Nenhum efeito positivo foi observado com o uso de arginina, um comprovado aditivo que suprime agregação de proteínas e solubiliza pellets insolúveis (Arakawa *et al.*, 2007). Acredita-se que a arginina causa leve desestruturação em proteínas, pela perturbação local de ambientes aromáticos, embora não resulte na perda da conformação estrutural ou na inativação de enzimas (Ishibashi *et al.*, 2005).

Resultados positivos também foram obtidos com a utilização de alta força iônica (500 mM NaCl), condição que aumenta a estabilidade e solubilidade de proteínas. De fato, uma significativa variedade de proteínas precipita se a concentração de sal é reduzida para níveis fisiológicos (no caso de proteínas animais), principalmente quando elas se tornam mais puras e concentradas (Graslund *et al.*, 2008).

Após estas observações, o tampão de lise composto de Tris-HCl pH 8,0 e NaCl 500 mM foi adotado para todas as extrações futuras.

Purificação parcial de CaFMO8: as proteínas recombinantes fusionadas à MBP e NusA foram purificadas por cromatografia de afinidade, embora o princípio de funcionamento varie de acordo com a cauda fusionada. Para a purificação de 6xHis-NusA:FMO e 6xHis-NusA:GFP utilizou-se IMAC em resina de níquel imobilizado em ácido nitrotriacético (Ni-NTA, Qiagen). Quantidades crescentes de imidazol, um competidor pelo níquel da resina, foram usadas para eluir as proteínas fusionadas. Embora pequena, a quantidade de 6xHis-NusA:FMO purificada foi relativamente satisfatória (figura 11A).

O grande problema na purificação desta construção foi a grande quantidade de contaminantes que eluiu juntamente com a proteína de interesse. Por outro lado, 6xHis-MBP:FMO e 6xHis-MBP:GFP foram purificadas em resina de amilose (New England), utilizando volumes diferentes de uma mesma solução de maltose 10 mM para a eluição das proteínas recombinantes. No caso de 6xHis-MBP:FMO, o grau de pureza foi maior em relação à 6xHis-NusA:FMO, mas a quantidade de proteína recuperada foi bem inferior.

Uma breve comparação com os controles, 6xHis-NusA:GFP e 6xHis-MBP:GFP (figura 11C), demonstra a gritante diferença na eficiência de produção de proteínas solúveis entre as GFPs e as FMOs fusionadas. A quantidade de proteína solúvel purificada foi o fator limitante para a realização dos ensaios enzimáticos.

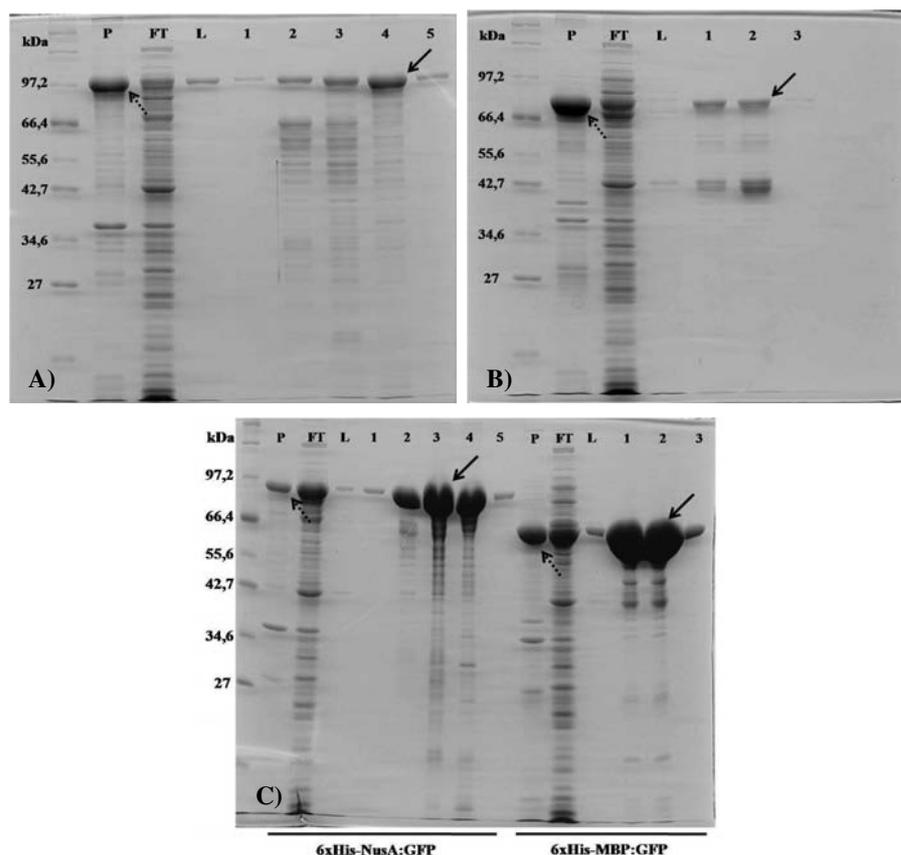


Figura 11. SDS-PAGE de 10% referente à purificação das proteínas fusionadas por cromatografia de afinidade. Purificação de 6xHis-NusA:FMO por IMAC em resina de níquel, sendo que a seta sólida indica a banda referente à 6xHis-NusA:FMO, equivalente a um peso molecular de 98,7 kDa (A). Purificação de 6xHis-MBP:FMO em resina de amilose, sendo que a seta sólida indica a tênue banda referente à FMO recombinante, equivalente a 84,4 kDa (B). Purificação de 6xHis-NusA:GFP e 6xHis-MBP:GFP em resinas de níquel e amilose respectivamente, sendo que as setas sólidas indicam as proteínas solúveis eluídas como descrito acima, com peso molecular de 84,4 kDa (Nusa/GFP) e 70,2 kDa (MBP/GFP) (C). As setas pontilhadas representam, em todos os casos, a expressão da proteína na fração insolúvel. Legenda: P, pellet; FT, flow through; L, lavagem; 1 a 5, eluições.

CaFMO8 não apresentou atividade de oxidação de triptamina in vitro: vários estudos que caracterizaram proteínas YUCCA e YUCCA-like recombinantes demonstraram a capacidade destas enzimas em converter a triptamina, um composto derivado do triptofano, em N-hidroxil triptamina, reação pertencente à via de síntese de auxina dependente de triptofano (Zhao *et al.*, 2001; Exposito-Rodriguez *et al.*, 2007; Kim *et al.*, 2007). Para os testes *in vitro*, estes autores utilizaram a proteína fusionada a MBP, expressa em bactéria e purificada por cromatografia de

afinidade (Zhao *et al.*, 2001; Kim *et al.*, 2007) ou o extrato de proteínas solúveis diretamente (Exposito-Rodriguez *et al.*, 2007). Assim, optou-se por utilizar somente a segunda opção, pela dificuldade de se conseguir quantidades significativas de proteína solúvel purificada. Além deste problema, as FMOs recombinantes parcialmente purificadas precipitaram na diálise, mesmo utilizando um tampão de diálise com 500 mM de NaCl.

O ensaio foi realizado com o extrato de proteínas solúveis totais, dessalinizado em colunas PD-10, na presença de FAD, NADPH e triptamina, em reação incubada a 37°C por 3 horas. Embora se tenha observado previamente que uma alta força iônica garantia maior solubilidade da proteína expressa, decidiu-se eliminar o sal da solução por dois motivos (i) possível interferência do sal na atividade enzimática da FMO e (ii) referências da literatura em que os autores não utilizaram sal no tampão da reação enzimática. Após a incubação, as reações foram analisadas diretamente em HPLC para a quantificação da triptamina restante em cada reação. Os resultados não evidenciaram nenhuma atividade de oxidação deste composto por parte das FMOs recombinantes, visto que as quantidades de triptamina presentes nos controles foram estatisticamente iguais àsquelas presentes nas reações com CaFMO8 fusionada tanto com MBP quanto com NusA (figura 12).

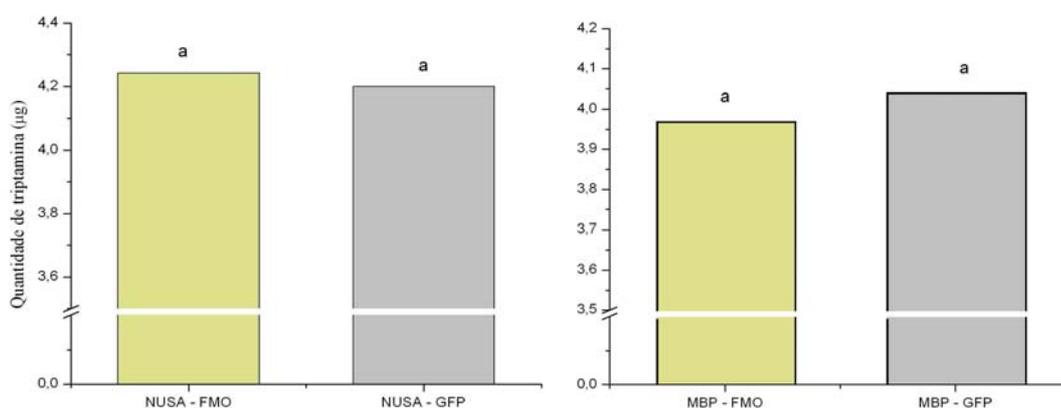


Figura 12. Comparação das quantidades de triptamina restantes ao final das reações *in vitro*. Nenhuma diferença estatisticamente relevante foi encontrada entre as soluções protéicas contendo FMOs recombinantes em relação aos controles com GFP. Não foi possível inferir o papel biológico *in vitro* da proteína CaFMO8 pelos ensaios montados.

5. Discussão

Enquanto outros organismos eucariotos e procariotos apresentam poucos genes FMOs funcionais, as plantas dispõem de um número significativamente maior, indicando uma utilização muito mais ampla dessa família de proteínas. Todavia, poucas funções biológicas envolvendo FMOs foram descritas em plantas, sendo a participação de proteínas FMO-like na biossíntese de auxina o caso mais estudado. Ao menos duas rotas de síntese de ácido indol-3-acético (IAA) a partir de triptofano são ativas em *Arabidopsis* (Cohen *et al.*, 2003), cuja redundância acaba por formar o mesmo intermediário indol-3-acetaldoxime (IAOx). A conversão direta de triptofano em IAOx é catalisada por dois citocromos P450, CYP79B2 e CYP79B3 (Zhao *et al.*, 2002). Em 2001, Zhao e colaboradores identificaram o gene *YUCCA*, que codifica uma FMO responsável pela N-hidroxilação de triptamina em N-hidroxil triptamina, passo limitante na via de síntese de auxina via triptofano. Por sua vez, N-hidroxil triptamina é transformada em IAOx, ponto de convergência entre as duas rotas redundantes citadas acima. A identificação de outros genes *YUCCA* em *Arabidopsis*, bem como a descoberta de genes homólogos em outras espécies, permitiu estender a idéia de redundância e conservação desta via.

No presente trabalho, foi feita a identificação e caracterização parcial do gene *CaFMO8* de *C. arabica*, o primeiro gene *YUCCA*-like descrito em café e provavelmente na família Rubiaceae. *CaFMO8* codifica uma proteína de 374 aminoácidos, com alta identidade em relação a proteínas *YUCCA* e *YUCCA*-like. Além disso, *CaFMO8* apresenta os mesmos motivos conservados característicos da família *YUCCA* (Exposito-Rodriguez *et al.*, 2007; Schlaich, 2007). Embora variações nestes motivos tenham sido observados em proteínas de clados mais divergentes (Exposito-Rodriguez *et al.*, 2007), *CaFMO8* possui diferenças não encontradas em outras proteínas *YUCCA*-like previamente caracterizadas. O motivo

identificador de FMOs (FXGXXXHXXXY/F), altamente conservado entre FMOs de todos os reinos, é significativamente diferente em CaFMO8 (YXGXXXHXXXY), sendo inclusive um fato nunca antes observado para uma FMO, visto que este motivo, como o próprio nome diz, é identificador deste tipo de enzimas. As conseqüências dessa alteração não são conhecidas. Choi *et al.* (2003) verificaram a importância funcional deste motivo na FMO de *Methylophaga* sp. responsável pela produção de índigo. Por meio de mutações sítio-dirigidas, estes pesquisadores substituíram individualmente cada um dos resíduos formadores do motivo por alanina, e observaram sua influência na síntese de índigo. Seus resultados mostraram que todas as substituições aboliram a produção do composto citado, indicando que a seqüência identificadora de FMOs é importante para a atividade da enzima bacteriana. Embora esses dados sejam altamente significativos, seria prematuro afirmar que CaFMO8 não apresenta atividade funcional. Os resíduos de fenilalanina (F) e tirosina (Y) são estruturalmente semelhantes, sendo ambos aromáticos, embora o primeiro seja um aminoácido apolar, enquanto o segundo é polar. Ainda sim, a substituição de um resíduo pelo outro não seria uma mudança dramática ao ponto de alterar definitivamente a interação da proteína com o substrato. Esta hipótese é suportada pelo fato de o último resíduo deste próprio motivo apresentar a mesma variação F/Y, sem a perda da atividade funcional.

Outra variação apresentada por CaFMO8 é em relação ao motivo “(F/L)ATGY”, ou mais atualmente “DX(I/V)(I/V)(L/F)ATGY(K/R)SNVP”. Dentre as proteínas YUCCA, somente YUCCA1, YUCCA10 e YUCCA11 apresentam “FATGY”, enquanto todas as outras YUCCA possuem “LATGY”. De acordo com sua proximidade com YUCCA10/YUCCA11, CaFMO8 também dispõe de fenilalanina (F) como o resíduo inicial deste motivo. No entanto, o motivo desta proteína de café é “FATGF”, com a substituição da tirosina (Y) por uma fenilalanina (F). Novamente, esta diferença não ocorre em nenhuma outra proteína YUCCA-

like descrita na literatura. Exposito-Rodriguez *et al.* (2007), que ampliaram o motivo em questão para “DX(I/V)(I/V)(L/F)ATGY(K/R)SNVP”, descreveram um novo motivo com potencial para identificação de novas proteínas YUCCA-like (WL(I/V)VATGENAE). Ainda sim, eles afirmaram que variações nestes motivos devem ser esperadas para membros do clado YUCCA10/YUCCA11, ao qual pertence CaFMO8, considerado o mais divergente dentre os clados de proteínas YUCCA. De fato, outros estudos sugerem o caráter mais ancestral deste clado. Gallavotti *et al.* (2008) caracterizaram o mutante *spi1* de milho, deficiente em um gene YUCCA-like, que apresenta sérios problemas de desenvolvimento. Neste trabalho, os autores realizaram uma análise filogenética bayesiana de 63 genes *YUC*-like de plantas, enraizados por dois genes de fungos. Baseados na posição do clado de genes *YUC* de musgos (*Physcomitrella patens*) eles consideraram os genes *AtYUC10* e *AtYUC11* mais distantes de outros genes *YUC* de eudicotiledôneas e monocotiledôneas, sugerindo inclusive uma divergência anterior à origem das plantas terrestres. Em vista destes dados, pode-se supor que as diferenças peculiares apresentadas por CaFMO8 podem ser fruto de um maior tempo de divergência e, desse modo, um maior tempo para diferenciação ao longo do tempo evolutivo.

Com o intuito de rastrear outros possíveis casos em que estas variações ocorreram, um alinhamento com outras proteínas YUCCA-like foi realizado. Além das proteínas YUCCA de *Arabidopsis* (YUCCA1 a 11) e OsYUCCA (OsYUCCA1 a 7) de arroz, outras seqüências foram selecionadas com base em Gallavotti *et al.* (2008) e no resultado do blastp realizado para CaFMO8. A variação “YXGXXXHXXXY” foi encontrada somente em OsYUCCA5 (ABA99096), cuja análise filogenética (Gallavotti *et al.*, 2008) sugere ser um produto de duplicação gênica dentro de monocots, não apresentando um ortólogo em eudicots. Não há dados na literatura sobre a caracterização funcional da proteína OsYUCCA5, sendo o padrão de expressão espacial do gene, avaliado por Yamamoto *et al.* (2007), o único dado encontrado

referente a esta proteína. Já a variação “FATGF” foi encontrada em três outras proteínas analisadas. PtXIV000597 de *Populus*, que de acordo com a análise de Gallavotti *et al.* (2008) é aparentemente o ortólogo de *YUC10*, e duas outras proteínas de *Vitis vinifera*, CAO22421.1 e CAO22420.1, que foram obtidas com o blastp de CaFMO8 e apresentam e-value de 4e-142 e 9e-140, respectivamente. Esses resultados comprovam que (i) não houve erros no seqüenciamento de *CaFMO8* e (ii) variações podem ocorrer entre proteínas YUCCA-like, principalmente do clado YUCCA10/YUCCA11. As duas proteínas de *Vitis* e a proteína de *Populus*, compartilham 68, 66 e 70% de identidade com CaFMO8, respectivamente. No entanto, como essas proteínas não foram caracterizadas funcionalmente, nenhuma dica em relação à função de CaFMO8 pôde ser conseguida.

O padrão de expressão espacial de *CaFMO8* foi analisado por RT-PCR semi-quantitativa, utilizando actina como normalizador. Observou-se que este gene é preferencialmente expresso em raízes, folhas e, em especial, flores imaturas. Uma banda bastante fraca foi observada para frutos, em ambas as fases analisadas. Cabe ressaltar que o unigene C8 que identificou *CaFMO8* é formado por 17 reads, sendo que sete são de bibliotecas de folhas e mais 6 de uma biblioteca feita de um pool de tecidos, mas principalmente de folhas de plantas submetidas a estresse hídrico.

Este padrão de expressão gênica não corresponde ao observado para *YUC10* de *Arabidopsis*. Embora nenhum dado de expressão por RT-PCR tenha sido publicado a respeito de *YUC10* e *YUC11*, Cheng *et al.* (2007) identificaram com o GENEVESTIGATOR (Zimmermann *et al.*, 2004) que *YUC10*, mas não *YUC11*, é altamente expresso em siliquis, o fruto produzido por plantas de *Arabidopsis*. Com análises de hibridização *in situ* eles também mostraram que estes genes apresentam expressão semelhante à *YUC1* e *YUC4* durante a embriogênese, sendo significativamente expressos no terço superior do embrião globular e na

região apical dos cotilédones no embrião em fase de coração. Os autores consideram os genes *YUC10/YUC11*, bem como *YUC1* e *YUC4*, fundamentais para a embriogênese em *Arabidopsis*. A partir dessa observação, somada ao fato das análises filogenéticas indicarem proximidade de *CaFMO8* com *YUCCA10* e *YUCCA11*, optou-se por verificar o padrão de expressão do gene de trabalho em diferentes partes da planta, com atenção especial para frutos. Assim, duas fases do fruto de café foram selecionadas, uma fase inicial (F2) e outra mais próxima do fruto maduro (F4). No entanto, o nível de transcritos de *CaFMO8* em frutos foi extremamente baixo, quando comparado aos dados de raiz, folhas e flores em desenvolvimento. Talvez *CaFMO8* não seja o gene ortólogo de *YUC10* em café, mas somente um gene pertencente ao mesmo clado, oriundo de duplicações gênicas e com divergência suficiente para apresentar um padrão de expressão diferente. Nem sempre genes ortólogos apresentam necessariamente os mesmos padrões de expressão gênica. *OsYUCCA1* de arroz e *SP11* de milho são considerados genes ortólogos, embora apresentem padrões de expressão espaciais e mesmo características funcionais distintas (Gallavotti *et al.*, 2008). Como os membros da família *YUCCA* apresentam diferentes padrões de expressão (Cheng *et al.*, 2006), sendo inclusive expressos em regiões extremamente limitadas (Zhao, 2008), as plantas devem utilizar vários genes da família para produzir auxina onde e quando ela for necessária. As proteínas *YUCCA* apresentariam redundância funcional em termos bioquímicos, mas seriam importantes em estágios de desenvolvimento específicos ou em condições ambientais específicas (Kim *et al.*, 2007).

A busca por outros genes *YUC*-like em café não se limita ao entendimento da biossíntese de auxina em plantas, mas também como diferentes espécies desenvolveram estratégias distintas para a produção deste hormônio. Uma discussão interessante que vêm acontecendo na literatura trata da complexidade da redundância de genes *YUC*-like.

Arabidopsis apresenta 11 genes *YUC* em seu genoma (Zhao, 2008). No entanto, nenhum fenótipo aparente é observado quando um único gene *YUC* é inativado (Cheng *et al.*, 2006). A inativação de ao menos dois genes *YUC* é necessária para a observação de um fenótipo característico. Ademais, caso os dois genes inativados pertençam ao mesmo clado dentro da família *YUCCA*, um fenótipo mais dramático é observado, sugerindo uma redundância em nível mais profunda entre genes de um mesmo clado (Cheng *et al.*, 2006). A inativação de *FLOOZY*, o ortólogo de *YUC1* em *Petunia x hybrida*, causa sérios defeitos no desenvolvimento floral e vascular da planta (Tobena-Santamaria *et al.*, 2002). De maneira similar, plantas de arroz antisense para *OsYUCCA1* apresentam fenótipo anormal associado a deficiência de auxina (Yamamoto *et al.*, 2007). Em milho, a perda de função de um único gene *SPII* causa alterações dramáticas no desenvolvimento de órgãos laterais e principalmente em inflorescências (Gallavotti *et al.*, 2008). Esses fatos sugerem que a redundância gênica referente aos genes *YUC* é muito mais complexa em *Arabidopsis* do que em outras plantas estudadas. Além desta diferença, a síntese de IAOx a partir de triptofano nem sempre é ativa em outras plantas. Yamamoto *et al.* (2007) relataram a impossibilidade de se identificar genes homólogos a *CYP79B2* e *CYP79B3* em arroz, sugerindo que esta via não é ativa nesta monocotiledônea e que a síntese de IAA seria diferente do que ocorre em *Arabidopsis*. Portanto, genes *YUC*-like precisam ser identificados e caracterizados em outras plantas para que a síntese de auxina seja entendida de uma forma mais universal, identificando possíveis variações entre diferentes famílias ou grupos taxonômicos.

A capacidade de metabolizar triptamina em N-hidroxil triptamina *in vitro* não pôde ser comprovada para CaFMO8 recombinante fusionada. Uma série de questões pode ser formulada a esse respeito. Primeiramente, uma baixíssima quantidade de proteína solúvel foi conseguida, embora uma série de metodologias e alterações de protocolo tenha sido realizada.

Nos experimentos de caracterização funcional *in vitro* de YUCCA (Zhao *et al.*, 2001), os autores utilizaram uma quantidade bastante alta de proteína recombinante para as reações (350 µg de proteína purificada), fato que não pôde ser realizado para CaFMO8 recombinante. Além disto, os autores utilizaram um sistema de regeneração de NADPH, com o uso de glicose-6-fosfato desidrogenase, que permite a atividade contínua da enzima analisada. No caso de CaFMO8, utilizou-se 0,1 mM de NADPH, a mesma concentração usada por Kim *et al.* (2007), que caracterizaram YUCCA6 de *Arabidopsis*.

Um fato a ser discutido é se a proteína CaFMO8 recombinante fusionada, produzida pela *E. coli*, estaria funcionalmente ativa. A identificação de atividade enzimática em FMOs fusionadas foi realizada com sucesso por Brunelle *et al.* (1997), Zhao *et al.* (2001) e Kim *et al.* (2007). Além destes estudos, que comprovaram que a presença de uma cauda de afinidade, ou mesmo de solubilidade, não interfere na atividade da proteína recombinante, há estudos que sugerem que caudas como MBP e NusA apresentam uma função de chaperona, promovendo o *fold*ing correto das proteínas a que estão fusionadas (Nallamsetty and Waugh, 2006). Embora esses dados suportem a funcionalidade de 6xHis-MBP:CaFMO8 e 6xHis-NusA:CaFMO8, experimentos que comprovem o estado funcional ativo desta proteína recombinante são necessários, antes de se chegar a conclusões definitivas. Novamente a discussão acerca das variações nos motivos conservados de CaFMO8 tem espaço. Será que essas substituições levaram à perda da atividade da enzima, sendo ela agora incapaz de metabolizar triptamina?

Como observado no teste de solubilidade, a quantidade de CaFMO8 recombinante solúvel apresentou um aumento considerável quando o detergente não-aniônico Triton X-100 foi utilizado. De fato, esse detergente não causa a desnaturação e, conseqüentemente, perda da conformação estrutural por parte das proteínas, ao contrário de SDS (QIAexpressionist, QIAGEN). Brunelle *et al.* (1997) analisaram o efeito de diferentes concentrações de Triton X-

100 na atividade enzimática de MBP:FMO3 de humanos, em ensaios *in vitro*. A máxima atividade desta enzima foi observada na presença de 0,015% de Triton X-100, não só confirmando a manutenção da atividade enzimática, mas também sugerindo vantagens na utilização do detergente. Embora cogitado, o uso de Triton X-100 foi desconsiderado devido a possíveis problemas quanto à quantificação em HPLC, visto que o uso de detergentes pode danificar as colunas utilizadas neste aparelho.

Finalmente, a utilização de extrato de proteínas solúveis totais também pode explicar o insucesso em se detectar atividade para triptamina. Brunelle *et al.* (1997) observaram um aumento significativo de atividade específica quando MBP:FMO3 foi purificada. Além disto, a retirada do sal do tampão pela utilização de colunas PD-10 pode ter causado desestabilização de CaFMO8, visto que o teste de solubilidade indicou a necessidade de grande quantidade de sal para maior solubilidade das enzimas recombinantes. No entanto, é difícil mensurar o quanto essa grande quantidade de sal influenciaria na atividade da enzima. Portanto, mais testes devem ser realizados, alternado-se outros parâmetros e adotando outras metodologias para analisar a conversão de triptamina em N-hidroxil triptamina.

6. Conclusões

Apesar de não conseguir provar a capacidade de CaFMO8 recombinante em converter triptamina em N-hidroxil triptamina *in vitro*, uma grande quantidade de informações a respeito do gene e, em especial, da proteína foram adquiridas. O gene *CaFMO8* apresenta expressão preferencial em raízes, folhas e principalmente flores em desenvolvimento, porém com baixos níveis de transcritos em frutos. A proteína codificada possui 374 aminoácidos e apresenta todos os motivos conservados descritos para a família YUCCA, apesar das variações já esperadas para o clado a que pertence. A análise filogenética indica que CaFMO8 pertence ao clado YUCCA10/YUCCA11, o grupo mais divergente de proteínas YUCCA-like, tendo surgido provavelmente antes do aparecimento das plantas terrestres. A proteína recombinante produzida em *E. coli* foi expressa em sua maior parte na fração insolúvel, mesmo encontrando-se fusionada à caudas com função de aumentar a solubilidade. Uma bateria de testes mostrou que a quantidade de proteína solúvel aumenta com a utilização de um pH mais básico, detergente e alta força iônica, sugerindo características hidrofóbicas e provável formação de agregados protéicos, o que justificaria sua expressão na fração insolúvel. A purificação de CaFMO8 fusionada foi pouco eficiente, gerando pequena quantidade de proteína e baixo grau de pureza.

A comprovação da funcionalidade da proteína expressa é de vital importância para as conclusões finais do trabalho. A metodologia para isso já foi adquirida na literatura (Eswaramoorthy *et al.*, 2007). Se a proteína não estiver sendo produzida na forma ativa, algumas medidas serão tomadas, como a clivagem da cauda, mudanças nos parâmetros de expressão e melhorias na purificação. Caso contrário, os ensaios com triptamina serão repetidos, a fim de comprovar a atividade de CaFMO8 na síntese de auxina dependente de triptofano.

7. Referências Bibliográficas

- Alfieri A, Malito E, Orru R, Fraaije MW, Mattevi A** (2008) Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 6572-6577
- Arakawa T, Tsurnoto K, Nagase K, Ejima D** (2007) The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography. *Protein Expression and Purification* **54**: 110-116
- Bartsch M, Gobbato E, Bednarek P, Debey S, Schultze JL, Bautor J, Parker JE** (2006) Salicylic acid-independent *ENHANCED DISEASE SUSCEPTIBILITY1* signaling in Arabidopsis immunity and cell death is regulated by the monooxygenase FMO1 and the nudix hydrolase NUDT7. *Plant Cell* **18**: 1038-1051
- Bottoms CA, Smith PE, Tanner JJ** (2002) A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Science* **11**: 2125-2137
- Bradford MM** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254
- Brunelle A, Bi YA, Lin J, Russell B, Luy L, Berkman C, Cashman J** (1997) Characterization of two human flavin-containing monooxygenase (form 3) enzymes expressed in *Escherichia coli* as maltose binding protein fusions. *Drug Metabolism and Disposition* **25**: 1001-1007
- Cashman JR** (2005) Some distinctions between flavin-containing and cytochrome P450 monooxygenases. *Biochemical and Biophysical Research Communications* **338**: 599-604
- Cashman JR, Park SB, Berkman CE, Cashman LE** (1995) Role of hepatic FLAVIN-CONTAINING MONOOXYGENASE-3 in drug and chemical metabolism in adult human. *Chemico-Biological Interactions* **96**: 33-46
- Cashman JR, Zhang J** (2006) Human flavin-containing monooxygenases. *Annual Review of Pharmacology and Toxicology* **46**: 65-100
- Cheng Y, Dai X, Zhao Y** (2006) Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in *Arabidopsis*. *Genes & Development* **20**: 1790-1799

- Cheng YF, Dai XH, Zhao YD** (2007) Auxin synthesized by the YUCCA flavin Monooxygenases is essential for embryogenesis and leaf formation in *Arabidopsis*. *Plant Cell* **19**: 2430-2439
- Choi HS, Kim JK, Cho EH, Kim YC, Kim JI, Kim SW** (2003) A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. *Biochemical and Biophysical Research Communications* **306**: 930-936
- Cohen JD, Slovin JP, Hendrickson AM** (2003) Two genetically discrete pathways convert tryptophan to auxin: more redundancy in auxin biosynthesis. *Trends in Plant Science* **8**: 197-199
- Current Protocols in Protein Science (2007). In JE Coligan, BN Dunn, DW Speicher, PT Wingfield, eds. John Wiley & Sons
- Eswaramoorthy S, Bonanno JB, Burley SK, Swaminathan S** (2007) Mechanism of action of a flavin-containing monooxygenase (vol 103, pg 9832, 2006). *Proceedings of the National Academy of Sciences of the United States of America* **104**: 14543-14543
- Exposito-Rodriguez M, Borges AA, Borges-Perez A, Hernandez M, Perez JA** (2007) Cloning and biochemical characterization of *ToFZY*, a tomato gene encoding a flavin monooxygenase involved in a tryptophan-dependent auxin biosynthesis pathway. *Journal of Plant Growth Regulation* **26**: 329-340
- Falls JG, Cherrington NJ, Clements KM, Philpot RM, Levi PE, Rose RL, Hodgson E** (1997) Molecular cloning, sequencing, and expression in *Escherichia coli* of mouse flavin-containing monooxygenase 3 (FMO3): Comparison with the human isoform. *Archives of Biochemistry and Biophysics* **347**: 9-18
- Gallavotti A, Barazesh S, Malcomber S, Hall D, Jackson D, Schmidt RJ, McSteen P** (2008) *sparse inflorescence1* encodes a monocot-specific YUCCA-like gene required for vegetative and reproductive development in maize. *Proceedings of the National Academy of Sciences* **105**: 15196-15201
- Graslund S, Nordlund P, Weigelt J, Bray J, Hallberg BM, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, Dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, Prilusky J, Sussman JL, Stevens RC, Lesley SA, Wilson IA,**

- Joachimiak A, Collart F, Dementieva I, Donnelly MI, Eschenfeldt WH, Kim Y, Stols L, Wu R, Zhou M, Burley SK, Emtage JS, Sauder JM, Thompson D, Bain K, Luz J, Gheyi T, Zhang F, Atwell S, Almo SC, Bonanno JB, Fiser A, Swaminathan S, Studier FW, Chance MR, Sali A, Acton TB, Xiao R, Zhao L, Ma LC, Hunt JF, Tong L, Cunningham K, Inouye M, Anderson S, Janjua H, Shastri R, Ho CK, Wang DY, Wang H, Jiang M, Montelione GT, Stuart DI, Owens RJ, Daenke S, Schutz A, Heinemann U, Yokoyama S, Bussow K, Gunsalus KC, Struct Genomics C, Architecture Fonction M, Berkeley Struct Genomics C, China Struct Genomics C, Integrated Ctr Struct F, Israel Struct Proteomics C, Joint Ctr Struct G, Midwest Ctr Struct G, New York Struct Genomi XRC, Consortium NESG, Oxford Prot Prod F, Prot Sample Prod F, Max Delbruck Ctr Mol M, Proteomics RSG, Complexes S** (2008) Protein production and purification. *Nature Methods* **5**: 135-146
- Hansen BG, Kliebenstein DJ, Halkier BA** (2007) Identification of a flavin-monooxygenase as the *S*-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*. *Plant Journal* **50**: 902-910
- Hernandez D, Janmohamed A, Chandan P, Phillips IR, Shephard EA** (2004) Organization and evolution of the flavin-containing monooxygenase genes of human and mouse: identification of novel gene and pseudogene clusters. *Pharmacogenetics* **14**: 117-130
- Hines RN, Hopp KA, Franco J, Saeian K, Begun FP** (2002) Alternative processing of the human *FMO6* gene renders transcripts incapable of encoding a functional flavin-containing monooxygenase. *Molecular Pharmacology* **62**: 320-325
- Ishibashi M, Tsumoto K, Tokunaga M, Ejima D, Kita Y, Arakawa T** (2005) Is arginine a protein-denaturant? *Protein Expression and Purification* **42**: 1-6
- Kim JI, Sharkhuu A, Jin JB, Li P, Jeong JC, Baek D, Lee SY, Blakeslee JJ, Murphy AS, Bohnert HJ, Hasegawa PM, Yun DJ, Bressan RA** (2007) *yucca6*, a dominant mutation in *Arabidopsis*, affects auxin accumulation and auxin-related phenotypes. *Plant Physiology* **145**: 722-735
- Koch M, Vorwerk S, Masur C, Sharifi-Sirchi G, Olivieri N, Schlaich NL** (2006) A role for a flavin-containing mono-oxygenase in resistance against microbial pathogens in *Arabidopsis*. *Plant Journal* **47**: 629-639

- Krueger SK, Williams DE** (2005) Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacology & Therapeutics* **106**: 357-387
- Kubo A, Itoh S, Itoh K, Kamataki T** (1997) Determination of FAD-binding domain in flavin-containing monooxygenase 1 (FMO1). *Archives of Biochemistry and Biophysics* **345**: 271-277
- Laemmli UK** (1970) Cleavage of structural protein during the assembly of head of bacteriophage T₄. *Nature* **227**: 680-684
- Lawton MP, Cashman JR, Cresteil T, Dolphin CT, Elfarra AA, Hines RN, Hodgson E, Kimura T, Ozols J, Phillips IR, Philpot RM, Poulsen LL, Rettie AE, Shephard EA, Williams DE, Ziegler DM** (1994) A nomenclature for the mammalian flavin-containing monooxygenase gene family based on amino-acid sequence identities. *Archives of Biochemistry and Biophysics* **308**: 254-257
- Li LC, Qin GJ, Tsuge T, Hou XH, Ding MY, Aoyama T, Oka A, Chen ZL, Gu HY, Zhao YD, Qu LJ** (2008) *SPOROXYTELESS* modulates *YUCCA* expression to regulate the development of lateral organs in *Arabidopsis*. *New Phytologist* **179**: 751-764
- Malito E, Alfieri A, Fraaije MW, Mattevi A** (2004) Crystal structure of a Baeyer-Villiger monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 13157-13162
- Marsch-Martinez N, Greco R, Van Arkel G, Herrera-Estrella L, Pereira A** (2002) Activation tagging using the En-I maize transposon system in *Arabidopsis*. *Plant Physiology* **129**: 1544-1556
- Melo GA, Shimizu MM, Mazzafera P** (2006) Polyphenoloxidase activity in coffee leaves and its role in resistance against the coffee leaf miner and coffee leaf rust. *Phytochemistry* **67**: 277-285
- Mishina TE, Zeier J** (2006) The *Arabidopsis* flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance. *Plant Physiology* **141**: 1666-1675
- Nallamsetty S, Waugh DS** (2006) Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expression and Purification* **45**: 175-182

- Rezaian MA, Krake LR** (1987) Nucleic-acid extraction and virus detection in grapevine. *Journal of Virological Methods* **17**: 277-285
- Schlauch NL** (2007) Flavin-containing monooxygenases in plants: looking beyond detox. *Trends in Plant Science* **12**: 412-418
- Suh JK, Poulsen LL, Ziegler DM, Robertus JD** (1996) Molecular cloning and kinetic characterization of a flavin-containing monooxygenase from *Saccharomyces cerevisiae*. *Archives of Biochemistry and Biophysics* **336**: 268-274
- Suh JK, Poulsen LL, Ziegler DM, Robertus JD** (2000) Redox regulation of yeast flavin-containing monooxygenase. *Archives of Biochemistry and Biophysics* **381**: 317-322
- Tamura K, Dudley J, Nei M, Kumar S** (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599
- Tobeña-Santamaria R, Bliet M, Ljung K, Sandberg G, Mol JNM, Souer E, Koes R** (2002) FLOOZY of petunia is a flavin mono-oxygenase-like protein required for the specification of leaf and flower architecture. *Genes & Development* **16**: 753-763
- Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia YJ, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J** (2000) Activation tagging in *Arabidopsis*. *Plant Physiology* **122**: 1003-1013
- Woodward C, Bemis SM, Hill EJ, Sawa S, Koshiba T, Torii KU** (2005) Interaction of auxin and ERECTA in elaborating Arabidopsis inflorescence architecture revealed by the activation tagging of a new member of the YUCCA family putative flavin monooxygenases. *Plant Physiology* **139**: 192-203
- Yamamoto Y, Kamiya N, Morinaka Y, Matsuoka M, Sazuka T** (2007) Auxin biosynthesis by the *YUCCA* genes in rice. *Plant Physiology* **143**: 1362-1371
- Zhao YD** (2008) The role of local biosynthesis of auxin and cytokinin in plant development. *Current Opinion in Plant Biology* **11**: 16-22
- Zhao YD, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J** (2001) A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science* **291**: 306-309

- Zhao YD, Hull AK, Gupta NR, Goss KA, Alonso J, Ecker JR, Normanly J, Chory J, Celenza JL** (2002) Trp-dependent auxin biosynthesis in Arabidopsis: involvement of cytochrome P450s CYP79B2 and CYP79B3. *Genes & Development* **16**: 3100-3112
- Ziegler DM** (1990) Flavin-containing monooxygenases - enzymes adapted for multisubstrate specificity. *Trends in Pharmacological Sciences* **11**: 321-324
- Ziegler DM, Pettit FH** (1964) Formation of intermediate *N*-oxide in oxidative demethylation of *N, N*-dimethylaniline catalyzed by liver microsomes *Biochemical and Biophysical Research Communications* **15**: 188-193
- Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W** (2004) GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiology* **136**: 2621-2632

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação/tese de Mestrado/Doutorado intitulada “Clonagem, expressão e caracterização de uma flavina monooxigenase de *Coffea arabica*”:

() não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

(X) está inserido no **Projeto CIBio/IB/UNICAMP** (Protocolo nº 2006/02), intitulado “Expressão de genes do metabolismo secundário em plantas”;

() tem autorização da **Comissão de Ética em Experimentação Animal/IB/UNICAMP** (Protocolo nº _____);

() tem autorização do **Comitê de Ética para Pesquisa com Seres Humanos/FCM/UNICAMP** (Protocolo nº _____);

() tem autorização de comissão de bioética ou biossegurança externa à UNICAMP.
Especificar: _____

Igor Cesarino

Aluno:

Eliane R. F. Martins

p/ Orientador:

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

Nome:

Função:

Helena Coutinho F. de Oliveira

Profa. Dra. HELENA COUTINHO F. DE OLIVEIRA
Presidente
Comissão Interna de Biossegurança
CIBio/IB - UNICAMP