

UNIVERSIDADE FEDERAL DO PARANA

ANELIS MARIA MARIN

CARACTERIZAÇÃO ESTRUTURAL E FUNCIONAL DA REGIÃO DO GENE *nodD*
DE *Herbaspirillum seropedicae*

CURITIBA
2009

ANELIS MARIA MARIN

CARACTERIZAÇÃO ESTRUTURAL E FUNCIONAL DA REGIÃO DO GENE *nodD*
DE *Herbaspirillum seropedicae*

Dissertação apresentada ao Curso de Pós-Graduação em Bioquímica, Área de Concentração em Biologia Molecular, Departamento de Bioquímica e Biologia Molecular, Setor de Ciências Biológicas, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Ciências.

Orientadora: Prof. Dra. Rose Adele Monteiro
Co-Orientadora: Prof. Dra. Roseli Wassem

CURITIBA
2009

AGRADECIMENTOS

Às minhas orientadoras, Prof. Dra. Rose Adele Monteiro e Prof. Dra. Roseli Wassem. Por todos esses anos de orientação, dedicação, atenção, carinho, amizade e por toda a paciência que sempre tiveram.

Ao Prof. Dr. Emanuel Maltempi de Souza, que igualmente me orienta. Por toda a dedicação, atenção, e por todas as sugestões no decorrer do trabalho.

Ao Prof. Dr. Fábio de Oliveira Pedrosa, pela oportunidade em trabalhar neste laboratório.

Ao Programa GENOPAR pelas sequencias.

Ao Prof. Dr. Leonardo Magalhães Cruz, pelo auxílio com o seqüenciamento completo da região estudada.

A Prof. Dra. Maria Berenice Steffens, pela orientação no estágio em docência; a Prof. Dra. Leda Satie Chubatsu e a Prof. Dra. Liu Un Rigo, pelo apoio.

Ao Valter Antônio de Baura, por toda a ajuda, pela amizade, e pelas idéias em diversos experimentos.

A Roseli Prado e a Julieta Pie por todo o carinho, preocupação e atenção.

À galera do anexo: Maria Augusta, Michelle Tadra, Arnaldo, Hellison, Marcelo Müller, Stefania, Eduardo, Viviane, Thalita, Ana Carolina, Bruno e Marcelo Batista, pela ajuda e pelos momentos de descontração.

Aos demais membros do Núcleo de Fixação de Nitrogênio, e em especial à Ana Cláudia, ao Marco Antonio, ao Marco Aurélio, e ao Giovanni Pisa por toda a ajuda e amizade.

À Dona Marilza por toda a paciência.

Ao Grupo de Estrutura de Carboidratos, deste departamento, em especial ao Prof. Dr. Guilherme L. Sasaki e ao Pós-doutorando Lauro M. de Souza por toda a ajuda com o HPLC.

Ao Grupo de Oxidações Biológicas, deste departamento, em especial a Prof. Dra. Glaucia Regina Martinez e ao aluno de mestrado Paulo Worfel por conceder alguns flavonóides para testes.

A minha turma de mestrado, por todos os momentos de descontração, e pela discussão das aulas.

As minhas eternas amigas, Maria Augusta e Carolina Heyse. Por tudo o que passamos juntas (e quanta coisa, hein?), por todo apoio, carinho, amor e amizade. Aos meus amigos, Leandro Roberto, Arquimedes Paixão por toda a amizade e carinho

Aos amigos de sempre Tuca, Carol, Ana Cláudia, Mônia, Rinaldo Oliveira, Danillo Gardenal, Danillo Carneiro, Alison Batista, Roger de Souza, Juliana Cassolato e Graciele Viccini que moram no meu coração. Por todos os momentos de diversão, conselhos, amizade e carinho.

A minha família que é a minha base; minha mãe Evanira Marin, meu pai Santo Marin (*in memoriam*), meu irmão Carlos Alberto, minha cunhada Marinês, e a minha pequena princesa Laura. Tudo o que tenho e o que sou devo a eles. Pela paciência e compreensão em todos os momentos (que não foram poucos) em que não pude estar presente.

A Deus, que sempre me guiou, que nunca me deixa desistir, e que sempre me acalma.

Ao Programa de Pós-Graduação em Bioquímica.

Ao CNPq pelo suporte financeiro.

Eu prefiro na chuva caminhar,
que em dias frios em casa me esconder

Martin Luther King

Resumo

Herbaspirillum seropedicae é uma bactéria diazotrófica que tem a capacidade de se associar endofiticamente à raízes de plantas sem que ocorra a formação de nódulos radiculares. A anotação do genoma pelo Programa GENOPAR identificou a presença do gene *nodD*, sendo este o único gene de nodulação encontrado em *H. seropedicae*. Este gene codifica para a proteína NodD, que é a principal proteína regulatória do processo de nodulação em rizóbios, e pertence a família LysR de reguladores de transcrição. Em *H. seropedicae* esta proteína é denominada HsNodD e ativa a transcrição de um operon a montante e divergente a *nodD*, que contém 10 *orfs*, e foi chamado operon *fde* (Flavonoid Degradation). Uma vez que este operon ainda não tem a sua função descrita em outros microrganismos, o presente estudo teve como objetivo principal esclarecer sua função. Estudos de bioinformática mostraram que as proteínas codificadas por este operon contêm domínios conservados relacionados com a quebra de anéis fenólicos. Uma estirpe mutante na 1ª *orf* do operon *fde*, denominada Orf1Km foi obtida e testada quanto ao seu crescimento em meio contendo fontes alternativas de carbono contendo anéis fenólicos, como naringenina, quercetina, ácido salicílico, salicilato de sódio, ácido benzóico, resorcinol, tirosina, fenilalanina e triptofano. Também foram testadas as estirpes selvagem (SmR1) e a estirpe mutante *nodD*⁻ (DR2), mas nenhuma diferença de crescimento foi observada não sendo possível a determinação do fenótipo da estirpe Orf1Km. A região intergênica entre *nodD* e *fdeA* foi clonada em ambas as orientações a jusante do gene repórter *lacZ*. Ensaio de β-galactosidase mostraram que a proteína HsNodD não ativou a expressão da fusão *nodD::lacZ*, e mesmo na presença de naringenina não houve alteração., indicando que HsNodD não se autoregula, como acontece com outras NodDs já descritas. Em relação à expressão da fusão *orf1::lacZ* a HsNodD ativa a sua expressão, e na presença de naringenina, apigenina e crisina ocorre um aumento na expressão desta fusão. As estirpes SmR1, DR2 e Orf1Km foram cultivadas em meio contendo naringenina por 30 horas, e a cada 6 horas foram coletadas amostras e estas foram injetadas em HPLC. Os resultados obtidos mostraram que a estirpe SmR1 com 24 horas de cultivo degrada a naringenina, porém as estirpes DR2 e Orf1Km não, indicando que estes genes estão envolvidos na degradação de naringenina. Assim, os resultados obtidos indicam que a naringenina é uma molécula sinalizadora, que desencadeia a sua própria degradação. HsNodD é sensível a naringenina, e na presença deste composto aumenta a ativação do operon *fde*, que por sua vez está envolvido na degradação de naringenina em *H. seropedicae*.

Palavras-chave: *Herbaspirillum seropedicae*; NodD; flavonóides; degradação.

Abstract

Herbaspirillum seropedicae is a nitrogen fixing bacterium found in association with plants of the Poales Order. Unlike rhizobium and leguminosae interaction, where bacteria are found in nodules, *H. seropedicae* are endophytic, colonizing intercellular spaces and vascular tissues of Poales plants. GENOPAR Program identified a *nodD* gene, which is the unique nodulation gene found in the *H. seropedicae* genome. This gene codifies the NodD protein, that is the main regulatory protein in the nodulation process, belonging to the LysR transcriptional regulators family. In *H. seropedicae* this protein is called HsNodD and activates the transcription of an operon, which contain 10 *orfs* and was called *fde* (Flavonoid Degradation) operon. Bioinformatic studies showed that the proteins codified by this operon contain conserved domains involved with phenolics ring degradation. A mutant strain in the first *orf* of the *fde* operon was obtained, and named Orf1Km. The ability of the wild strain (SmR1), *nodD*⁻ strain (DR2) and *orf1*⁻ strain (Orf1Km) to growth in medium containing alternative carbon source, like, naringenin, quercetin, salicylic acid, benzoic acid, sodium salicylate, resorcinol, tyrosine, phenylalanine, tryptofane were tested, but no growth differences were observed. The intergenic region of *nodD* and *fdeA* was cloned in both orientations upstream the *lacZ* gene reporter. β -galactosidase assays showed that the HsNodD protein did not activate the *nodD::lacZ* expression in the presence or in the absence of naringenin. The HsNodD activates *orf1::lacZ* expression, and in the presence of naringenin, apigenin and chrysin occurred an increased in the expression. The SmR1, DR2 e Orf1Km strains were cultivated in medium containig naringenin for 30 hours, and aliquots were removed at each 6 hours and analysed by HPLC. The results showed that SmR1 strain was able to degradate naringenin with 24 hours, but DR2 and Orf1Km strains did not, indicating that *orf1* and *nodD* are involved in the degradation of naringenin. The results obtained showed that naringenin is a signal molecule, which induces its own degradation. HsNod is sensible to naringenin, and in the presence of this compound increased the *fde* operon transcription activation, whose the function seems to degrade naringenin in *H. seropedicae*.

Keywords: *Herbaspirillum seropedicae*; NodD; flavonoids, degradation.

LISTA DE FIGURAS

FIGURA 1 – MOTIVOS CONSERVADOS ENCONTRADOS NA PROTEÍNA HsNodD.....	16
FIGURA 2 – MOTIVO HÉLICE-VOLTA-HÉLICE DA PROTEÍNA HsNodD	17
FIGURA 3 – ALINHAMENTO DOS PROMOTORES DEPENDENTES DE FATOR σ^{70} ENCONTRADOS NA REGIÃO PROMOTORA DOS GENES <i>nodD</i> E <i>orf1</i> DE <i>H. seropedicae</i> A SEQUÊNCIA CONSENSO DE <i>E. coli</i>	17
FIGURA 4 – MODELO PROPOSTO PARA A ATIVAÇÃO DA TRANSCRIÇÃO DE GENES <i>nod</i>	24
FIGURA 5 – ESTRUTURA BÁSICA DE UM FATOR NOD	28
FIGURA 6 – ESTRUTURA BÁSICA DE UM FLAVONÓIDE	29
FIGURA 7 – VIA DE BIOSÍNTESE DOS FLAVONÓIDES.	30
FIGURA 8 – ESTRUTURAS BÁSICAS DE CADA CLASSE DE FLAVONÓIDES.....	32
FIGURA 9 – VIA DE DEGRADAÇÃO DA QUERCITINA EM <i>P. putida</i> PML2.....	34
FIGURA 10 – REAÇÃO DE OXIDAÇÃO QUE OCORRE NO INTERIOR DA CÉLULA NA PRESENÇA DO COMPOSTO CLORETO DE 2,3,5-TRIFENIL TETRAZÓLIO.....	50
FIGURA 11 – ESQUEMA DA DISPOSIÇÃO DO GENE <i>nodD</i> E DO OPERON <i>fde</i> DE <i>H. seropedicae</i>	57
FIGURA 12 – COMPARAÇÃO ENTRE A REGIÃO DO GENE <i>nodD</i> DE <i>H. seropedicae</i> E A REGIÃO SEMELHANTE ENCONTRADA EM <i>B. japonicum</i>	61
FIGURA 13 – PEPTÍDEO SINAL ENCONTRADO NA PROTEÍNA FdeA.	63
FIGURA 14 – CONSTRUÇÃO DO PLASMÍDEO pGAP.....	64
FIGURA 15 – SEQUÊNCIA DE NUCLEOTÍDEOS COMPLETA DO GENE <i>fdeF</i>	65
FIGURA 16 – CONSTRUÇÃO DO PLASMÍDEO pORF1.....	66
FIGURA 17 – HIBRIDIZAÇÃO DO DNA GENÔMICO DAS ESTIRPES SmR1 E AMM1 DE <i>H. seropedicae</i> COM O PRODUTO DE PCR DO PLASMÍDEO pORF1.....	67

FIGURA 18 – ESQUEMA DA RECOMBINAÇÃO HOMÓLOGA SIMPLES OCORRIDA PARA ORIGINAR A ESTIRPE MUTANTE AMM1 DE <i>H. seropedicae</i>	68
FIGURA 19 – CROMATOGRAMAS REFERENTES AOS PADRÕES	73
FIGURA 20 – CROMATOGRAMAS REFERENTES AOS CONTROLES NO TEMPO 0 HORAS.....	74
FIGURA 21 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 0 HORAS.	74
FIGURA 22 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 6 HORAS	75
FIGURA 23 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 12 HORAS.....	76
FIGURA 24 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 24 HORAS.....	76
FIGURA 25 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 30 HORAS.....	78
FIGURA 26 – REGIÃO CLONADA EM pMP220 E EM pPW452, ORIGINANDO AS FUSOES <i>nodD::lacZ</i> (pSU11) E <i>fdeA::lacZ</i> (pSUI).	79
FIGURA 27 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn CONTROLE (SEM FLAVONÓIDES)	80
FIGURA 28 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO <i>fdeA::lacZ</i>	82
FIGURA 29 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO <i>nodD::lacZ</i>	83
FIGURA 30 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO <i>nodABC::lacZ</i> DE <i>Rhizobium</i> sp.....	84
FIGURA 31 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE, UTILIZANDO O TRANSCONJUGANTE SmR1 <i>fdeA::lacZ</i> , VARIANDO A CONCENTRAÇÃO DE APIGENINA, LUTEOLINA, KAEMPFEROL E CRISINA NO MEIO.	85
FIGURA 32 - ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE UTILIZANDO 200 μ mol/L DE FLAVONA, CATEQUINA, GENISTEÍNA E DAIDZEÍNA COMO POSSÍVEIS INDUTORES DA FUSÃO <i>fdeA::lacZ</i>	86

LISTA DE TABELAS

TABELA 1 – BACTÉRIAS E PLASMÍDEOS	37
TABELA 2 – ANTIBIÓTICOS UTILIZADOS PARA O CULTIVO DE <i>E. coli</i> E DE <i>H. seropedicae</i>	40
TABELA 3 – COMPOSTOS UTILIZADOS NO SCREENING EM MEIO SÓLIDO	49
TABELA 4 – CONDIÇÕES UTILIZADAS NO EXPERIMENTO DE DETERMINAÇÃO DO FENÓTIPO	50
DA ESTIRPE AMM1 EM MEIO LÍQUIDO; pH 6,9	50
TABELA 5 – COMPOSTOS UTILIZADOS NO DECORRER DESTE TRABALHO.....	53
TABELA 6 - OPERON A MONTANTE E DIVERGENTE AO GENE <i>nodD</i> em <i>H. seropedicae</i>	58
TABELA 7 – COMPARAÇÃO ENTRE <i>H. seropedicae</i> SmR1 E <i>B. japonicum</i> USDA110	62
TABELA 8 – TESTE DE CRESCIMENTO DAS ESTIRPES SmR1, DR2 E AMM1 EM MEIO NFbHPN E NFbHPN-MALATO.....	69
TABELA 9 – LOCALIZAÇÃO POR POÇO DAS CONDIÇÕES TESTADAS.....	71
TABELA 10 – CÁLCULO DO TEMPO DE GERAÇÃO OBTIDO A PARTIR DE UMA CURVA DE CRESCIMENTO DAS ESTIRPES SmR1, DR2 e AMM1.	87

SUMÁRIO

RESUMO.....	5
ABSTRACT.....	6
LISTA DE FIGURAS.....	7
LISTA DE TABELAS.....	10
1 INTRODUÇÃO.....	13
2 REVISÃO BIBLIOGRÁFICA.....	15
2.1 <i>Herbaspirillum seropedicae</i>	15
2.1.1 Proteína NodD e gene <i>nodD</i> de <i>Herbaspirillum seropedicae</i>	16
2.2 FAMÍLIA LysR DE REGULADORES DE TRANSCRIÇÃO	18
2.3 PROTEÍNA NodD DE RIZÓBIOS	23
2.3.1 <i>nod</i> -box	26
2.3.2 Genes <i>nod</i> e sua regulação em rizóbios	27
2.4 FLAVONÓIDES	28
3 OBJETIVOS.....	36
3.1 OBJETIVOS GERAIS	36
3.2 OBJETIVOS ESPECÍFICOS.....	36
4 MATERIAL E MÉTODOS	37
4.1 BACTÉRIAS E PLASMÍDEOS.....	37
4.2 MEIOS DE CULTURA E CONDIÇÕES DE CRESCIMENTO	38
4.2.1 Cultivo de <i>E. coli</i>	38
4.2.2 Cultivo de <i>Herbaspirillum seropedicae</i>	39
4.2.3 Antibióticos utilizados	40
4.3 PURIFICAÇÃO DE DNA EM PEQUENA ESCALA.....	40
4.4 PURIFICAÇÃO DO DNA GENÔMICO DE <i>Herbaspirillum seropedicae</i>	41
4.5 ELETROFORESE DE DNA	42
4.6 DIGESTÃO DE DNA COM ENDONUCLEASES DE RESTRIÇÃO.....	42
4.7 PREPARO DE VETORES	42
4.8 LIGAÇÃO DE DNA AOS VETORES.....	42
4.9 SEQUENCIAMENTO DE DNA	43
4.10 PREPARO DE CÉLULAS ELETROCOMPETENTES E TRANSFORMAÇÃO POR ELETROPORAÇÃO.....	44
4.10.1 Preparo de células eletrocompetentes de <i>E. coli</i>	44
4.10.2 Preparo de células eletrocompetentes de <i>H. seropedicae</i>	44
4.11 CONJUGAÇÃO BACTERIANA.....	45
4.12 ARMAZENAMENTO DAS ESTIRPES BACTERIANAS.....	45
4.13 OBTENÇÃO DA ESTIRPE MUTANTE NO GENE <i>fdeA</i> DE <i>H. seropedicae</i>	46
4.14 HIBRIDIZAÇÃO DO DNA GENÔMICO DAS ESTIRPES MUTANTES PARA CONFIRMAÇÃO DA MUTAÇÃO DE <i>H. seropedicae</i>	46
4.14.1 Transferência do DNA.....	46
4.14.2 Preparo da sonda.....	47

4.14.3 Hibridização	47
4.15 SEQUENCIAMENTO DO GENE <i>fdeF</i> DO OPERON A MONTANTE DO GENE <i>nodD</i> DE <i>H. seropedicae</i>	48
4.16 CARACTERIZAÇÃO DO FENÓTIPO DO MUTANTE AMM1 DE <i>H. seropedicae</i>	48
4.16.1 Crescimento em meio NFbHPN e NFbHPN-malato sólido variando as fontes de carbono	49
4.16.2 Crescimento em meio NFbHPN e NFbHPN-malato líquido variando as fontes de carbono	50
4.16.3 Caracterização do fenótipo da estirpe AMM1 por HPLC	51
4.17 DETERMINAÇÃO DA ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE.....	52
4.18 SOLUBILIZAÇÃO DOS COMPOSTOS	53
4.20 CURVA DE CRESCIMENTO DAS ESTIRPES SMR1, DR2 E AMM1 NA PRESENÇA E NA AUSÊNCIA DE NARINGENINA.....	54
4.21 ANÁLISE DE BIOINFORMÁTICA.....	55
5 RESULTADOS E DISCUSSÃO	56
5.1 Estudo de bioinformática da região a montante do gene <i>nodD</i> de <i>H. seropedicae</i>	56
5.1.1 Sequenciamento da <i>fdeF</i>	63
5.2 OBTENÇÃO E CARACTERIZAÇÃO DA ESTIRPE MUTANTE <i>fdeA</i> ⁻ DE <i>H. seropedicae</i>	65
5.2.1 Construção do mutante <i>fdeA</i> ⁻ de <i>H. seropedicae</i>	65
5.3 DETERMINAÇÃO DO FENÓTIPO DO MUTANTE AMM1	68
5.3.1 Crescimento em meio NFbHPN e NFbHPN-malato sólido variando as fontes de carbono	68
5.3.2 Crescimento em meio NFbHPN e NFbHPN-malato líquido variando as fontes de carbono	70
5.3.5 Análise do fenótipo da estirpe AMM1 em HPLC	72
5.4 REGULÇÃO DA EXPRESSÃO DOS GENES <i>nodD</i> E <i>fdeA</i> DE <i>H. seropedicae</i> E DO OPERON <i>nodABC</i> EM <i>H. seropedicae</i>	78
5.4.1 Efeito da proteína HsNodD sobre a expressão dos genes <i>fdeA</i> e <i>nodD</i> de <i>H. seropedicae</i> e sobre o operon <i>nodABC</i> de <i>Rhizobium</i> sp. NGR234	80
5.4.2 Efeito do flavonóide naringenina sobre a atividade da proteína HsNodD	81
5.4.3 Efeito de diversos flavonóides sobre a expressão do gene <i>fdeA</i> em <i>H. seropedicae</i>	84
5.5 EFEITO DA NARINGENINA SOBRE O CRESCIMENTO DE <i>H. seropedicae</i>	87
6 CONCLUSÕES	89
7 REFERÊNCIAS BIBLIOGRÁFICAS	90
ANEXO	98

1 INTRODUÇÃO

O processo da fixação biológica de nitrogênio, realizado por algumas bactérias, é essencial para a manutenção da vida, e somente estes organismos possuem a capacidade de realizá-lo. Este processo gera amônio (NH_4^+), a partir do nitrogênio atmosférico (N_2), que é então utilizado pelos demais seres vivos na síntese de biomoléculas como, por exemplo, proteínas e ácidos nucléicos. O microrganismo capaz de realizar a redução do nitrogênio atmosférico a amônio é chamado de diazotrofo e pode ser utilizado como inoculante na agricultura, o que colabora para a diminuição da poluição causada pelos adubos nitrogenados, e também diminui o custo para o produtor. Dessa maneira, torna-se importante o estudo dos vários aspectos que envolvem todo o processo da fixação biológica de nitrogênio.

As bactérias diazotróficas podem estabelecer uma relação simbiótica ou associativa com o hospedeiro, sendo que esta relação pode ocorrer através de pelo menos duas formas: a primeira, e melhor estudada, é a interação simbiótica, que é o caso da associação leguminosa-rizóbio; a segunda, é a interação endofítica, que é o caso da associação gramínea-*Herbaspirillum/Azospirillum*. A diferença básica entre os dois tipos de processos é que na interação simbiótica há a formação do nódulo radicular, onde ocorre a fixação de nitrogênio, enquanto que na associação endofítica gramínea *Herbaspirillum/Azospirillum* não há formação de nódulo radicular, sendo as bactérias encontradas nos espaços intercelulares, onde realizam a fixação de nitrogênio.

Em rizóbios, a formação do nódulo radicular depende da expressão dos genes *nod*, cuja transcrição é ativada pela proteína NodD, codificada pelo gene *nodD*. Esta proteína é uma reguladora de transcrição da família LysR, e, em muitos casos, tem sua atividade controlada por interação com flavonóides. Quando ativa, a NodD altera a transcrição dos demais genes *nod* envolvidos no processo de nodulação. Esta alteração, na maioria dos casos, é a indução da transcrição dos demais genes *nod* como, por exemplo, o operon *nodABC*, que é responsável pela síntese do esqueleto central dos fatores Nod, lipoquitooligosacarídeos essenciais no processo de nodulação. Contudo, em alguns casos, a proteína NodD pode reprimir a expressão

gênica, em especial se houver mais de uma cópia deste gene no genoma do microrganismo.

Em *Herbaspirillum seropedicae*, uma bactéria diazotrófica endofítica encontrada em certas plantas da ordem Poales, foi identificado um gene (Programa GENOPAR) semelhante ao gene *nodD* (46% de identidade com o gene *nodD* de *Bradyrhizobium japonicum*). Uma vez que esta bactéria não induz a formação de nódulos na planta com a qual se associa, a função deste gene em *H. seropedicae* permanece desconhecida. Ensaios de interação entre uma estirpe *nodD*⁻ de *H. seropedicae* e milho sugerem que este gene não está envolvido no processo inicial de interação planta-bactéria (MARIN, 2006b). Outros resultados indicam que a proteína NodD está envolvida na regulação de um operon divergente ao gene *nodD*, uma vez que a expressão deste operon, em *E. coli*, é diminuída na presença de NodD (MARIN, 2006a).

Uma vez que a proteína NodD de *H. seropedicae* pode ter a sua atividade regulada por flavonóides e que o operon divergente ao gene *nodD* parece estar envolvido com o metabolismo de compostos fenólicos, este trabalho visa esclarecer a função deste operon.

2 REVISÃO BIBLIOGRÁFICA

2.1 *Herbaspirillum seropedicae*

Herbaspirillum seropedicae é uma bactéria diazotrófica, ou seja, fixadora de nitrogênio, pertencente à classe Beta das Proteobactérias. Possui de 0,6 a 0,7 µm de diâmetro e 1,5 a 5 µm de comprimento, é gram-negativa, vibrióide, e com dois (ocasionalmente, um ou três) flagelos em um ou ambos os pólos. *H. seropedicae* fixa nitrogênio em condições microaeróbicas, tem a capacidade de usar muitos açúcares como fonte de carbono e coloniza, endofiticamente, membros da família Poaceae (BALDANI *et al.*, 1986), como milho (*Zea mays*), arroz (*Oryza sativa*), sorgo (*Sorghum bicolor*) e cana-de-açúcar (*Saccharum officinarum*) (BALDANI *et al.*, 1984; PIMENTEL *et al.*, 1991; BODDEY *et al.*, 1995); Musaceae (*Musa* sp) e Bromeliaceae (*Ananas* sp) (WEBER *et al.*, 1999).

Estudos de microscopia mostraram que *H. seropedicae* Z67 é capaz de penetrar através de rachaduras das raízes laterais, colonizar os espaços intercelulares, aerênquima e xilema, tanto de raízes quanto de partes aéreas de plântulas de arroz (JAMES *et al.*, 2002). Outros estudos mostraram que *H. seropedicae* coloniza vasos do xilema de folhas de cana-de-açúcar (OLIVARES *et al.*, 1997). A partir da inoculação de diversas estirpes de *H. seropedicae* foi possível mostrar o grande potencial biofertilizante desta bactéria, uma vez que na interação arroz-*Herbaspirillum* a estirpe Z94 contribuiu com até 54% do nitrogênio total acumulado pela planta e as outras estirpes também contribuíram com aproximadamente 30% do nitrogênio (BALDANI *et al.*, 1995). Além disto, *H. seropedicae* também é capaz de produzir fitohormônios que estimulam o crescimento vegetal (OLIVARES *et al.*, 1997).

O genoma de *H. seropedicae* foi seqüenciado pelo Programa GENOPAR, e agora os genes envolvidos nos processos de fixação de nitrogênio e de colonização estão sendo estudados para entender o processo de interação planta-bactéria.

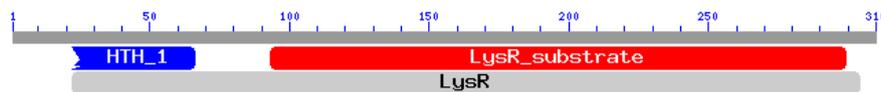
Apesar de *H. seropedicae* não induzir a formação de nódulos radiculares, o Programa GENOPAR encontrou no genoma deste microrganismo uma seqüência similar ao gene *nodD* de rizóbios. Este gene codifica a proteína NodD, que é chamada de HsNodD. Análise da seqüência desta proteína revela que a HsNodD possui dois

motivos estruturais característicos de proteínas da família LysR de reguladores transcricionais. O motivo hélice-volta-hélice, formado na região N-terminal é responsável pela ligação da proteína ao DNA, enquanto a região C-terminal está relacionada com a ligação da proteína a um possível indutor (SECCON, 2004).

2.1.1 Proteína NodD e gene *nodD* de *Herbaspirillum seropedicae*

A proteína HsNodD contém 309 resíduos de aminoácidos, e foi anotada pelo Programa GENOPAR como sendo uma proteína reguladora de transcrição com similaridade de 66% à proteína NodD de *Bradyrhizobium japonicum* USDA110. Quando submetida ao programa blastx, a similaridade é maior (74%) a NodD de *Burkholderia vietnamiensis*.

Através de análises de bioinformática da seqüência de aminoácidos da proteína HsNodD, foi encontrado dois motivos estruturais comuns de proteínas da família LysR de reguladores transcricionais. O motivo hélice-volta-hélice de ligação ao DNA encontra-se na região N-terminal, e o outro motivo, de ligação a indutores, encontra-se na região C-terminal (Figura 1).



<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>

FONTE: SECCON, 2005

FIGURA 1 – MOTIVOS CONSERVADOS ENCONTRADOS NA PROTEÍNA HsNodD. O motivo HTH (Helix-Turn-Helix) na região N-terminal encontra-se em azul, e o motivo LysR, de ligação a indutores na região C-terminal encontra-se em vermelho.

A mesma seqüência de aminoácidos foi submetida ao programa Helix-Turn-Helix Motif Prediction (DODD & EGAN, 1990), que indica onde se encontra o motivo hélice-volta-hélice. Segundo este programa o motivo de ligação ao DNA, se estende do resíduo 21 ao resíduo 42 (figura 2), que como revisto por SCHELL, 1993, é o chamado HTH-clássico.

```

          10          20          30          40          50          60
          |          |          |          |          |          |
MRFNKLDLNLVLDALLTEMSISRAAEKIHLSQSAMSNALARLREYFDDELLIQVGRRM
EPTPRAEVLKDAVHDLRRLRIDGSAALPAFVPAESTREFRISVSDFTLSVLIPRVLARAH
AEGKHIRFALMPQVQDPTRSLDRAEVDLLVLPQEFCTPDHPAEEVFRERHVCVVWRDSAL
AQGELTLERYMASGHVVMVPPGANASSVEAUMARKLGFARRVEVTSFSFASALALVQGTD
RIATVHARLAQLLAPQWPVVIKESPLSLGEMRQMMQWHRYSNDPQIQWLRRVFLESAQE
MDAALPGIC

```

FONTE: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hth.html

FIGURA 2 – MOTIVO HÉLICE-VOLTA-HÉLICE DA PROTEÍNA HsNodD. Em vermelho está a seqüência que o programa encontrou como sendo o motivo hélice-volta-hélice.

Estudos anteriores mostraram que quando superexpressa em *E. coli* a proteína HsNodD é capaz de reprimir a sua própria expressão, assim como a transcrição do operon divergente a ela. Contudo, ensaios de β -galactosidase na presença dos flavonóides naringenina e quercetina mostraram que em relação ao operon divergente à *nodD* estes compostos foram capazes de reverter a repressão causada pela proteína HsNodD, porém, em relação a sua própria expressão, tais compostos não revertem a repressão (MARIN, 2006a).

A proteína HsNodD parece não estar envolvida nos processos iniciais da interação planta-bactéria, uma vez que ensaios de adesão e colonização interna não mostraram diferença quando as estirpes selvagem (SmR1) e a estirpe *nodD*⁻ (DR2) de *H. seropedicae* foram comparadas (MARIN, 2006b).

Na região intergênica entre *nodD* e o 1º gene do operon vizinho (denominado *orf1*) há duas seqüências promotoras dependentes do fator σ^{70} , que tem como característica uma região consenso nas regiões -35 e -10 anterior ao início do gene. A seqüência consenso responsável pela regulação da transcrição do gene *nodD*, do gene *orf1*, bem como a homologia com o promotor -35/-10 de *E. coli* estão na figura 3.

<i>E. coli</i>	TTGACA (N ₁₇) TATAAT
HsnodD	TTGATA (N ₁₇) TTTGAT
Hsorfl	TGGTCA (N ₁₈) TATCAT
	* * * * * * * *

FONTE: MARIN, 2006a

FIGURA 3 – ALINHAMENTO DOS PROMOTORES DEPENDENTES DE FATOR σ^{70} ENCONTRADOS NA REGIÃO PROMOTORA DOS GENES *nodD* E *orf1* DE *H. seropedicae* A SEQÜÊNCIA CONSENSO DE *E. coli*.

* bases conservadas

Como mostrado na figura 3, as seqüências promotoras de *nodD* e *orf1* têm identidade à seqüência de *E. coli*, e como nesta região intergênica não foram encontradas outras seqüências promotoras candidatas, essas, provavelmente, são as reais seqüências promotoras dos genes em questão.

Estudos de footprinting (WASSEM – não publicado) revelaram que há uma seqüência de nucleotídeos similar à seqüência consenso de ligação para proteínas NodD, que é a seqüência que o motivo LysR reconhece, similar ao encontrado em rizóbios.

Uma vez que *Herbaspirillum seropedicae* não induz a formação de nódulos nas plantas com que se associa, mas possui um gene de nodulação, este projeto vem a auxiliar no esclarecimento sobre a presença deste gene no genoma deste microrganismo. Além disto, este projeto tenta esclarecer a função do operon divergente e a montante do gene *nodD* que é regulado pela proteína HsNodD, uma vez que os genes presentes neste operon têm homologia com genes, cujos produtos estão envolvidos na abertura e degradação de compostos aromáticos, que podem estar envolvidos na sinalização entre a planta e a bactéria, durante o estágio inicial do processo de interação planta-bactéria.

2.2 FAMÍLIA LysR DE REGULADORES DE TRANSCRIÇÃO

A família LysR forma um grupo de proteínas que possui mais de 100 representantes e é considerada a maior família de reguladores de transcrição de procariotos, responsáveis pela regulação da expressão de genes pertencentes a diferentes processos metabólicos. LTTRs (LysR Type Transcriptional Regulators) compartilham um ancestral comum dentro dos procariotos e são amplamente distribuídas, principalmente nas classes α e γ de Proteobactérias, sendo poucas encontradas na classe β e nenhuma na δ (revisito por SCHELL, 1993).

A grande maioria destes reguladores é sensível a indutores, que podem ser substratos, metabólitos intermediários ou produtos finais das vias nas quais os genes alvo estão envolvidos. Como características estruturais principais as LTTRs apresentam: um domínio de ligação ao DNA, chamado de motivo conservado Hélice-

Volta-Hélice (HTH) e um motivo menos conservado, denominado motivo LysR (revisto por SCHELL, 1993). O motivo HTH é encontrado do resíduo 1 a 65. Do resíduo 66 até o final da proteína, encontra-se o motivo LysR que é subdividido em 3 domínios: 2 domínios na região central envolvidos com o reconhecimento e/ou resposta ao indutor (resíduos 95-173 e 196-206) e um domínio envolvido na multimerização (resíduos 227 a 253) (revisto por TROPEL & van der MEER, 2004). Muraoka, *et al.*, 2003 elucidaram a estrutura tri-dimensional de CbnR, uma LTTR ativadora da expressão do operon *cbnABCD* de *Ralstonia eutropha* NH9, responsável pela degradação do composto 3-clorobenzoato. CbnR foi cristalizada na forma de um tetrâmero, sendo que cada monômero contém dois domínios: HTH do resíduo 1-58, e um domínio regulatório do resíduo 88-294, sendo que a ligação entre estes dois domínios é feita por uma hélice que vai dos resíduos 59-87. Apesar de as 4 subunidades serem idênticas, elas assumem conformações diferentes; duas delas (A e P) adotam uma forma compacta, enquanto as outras duas (B e Q) adotam uma forma estendida, e por isso, um tetrâmero pode ser chamado de um dímero de dímeros (MURAOKA *et al.*, 2003). A ligação do efetor e subsequente ativação da transcrição resultam da modificação da dobra no DNA provocada pela ligação da LTTR sob condições não-indutoras (PÉREZ-MARTÍN, *et al.*, 1994). Algumas outras características que as proteínas dessa família apresentam em comum são: tamanho variando de 276 a 324 resíduos; se ligam a seqüências regulatórias do DNA com posição e motivos similares; são divergentemente transcritos de um promotor que é muito próximo ou até mesmo se sobrepõe ao promotor do gene que regula, e essa sobreposição permite controle simultâneo e bidirecional da transcrição, e assim, muitas LTTRs reprimem a sua própria transcrição para manter seus níveis constantes. LTTRs permanecem ligadas à região proximal do promotor do gene alvo, independente da presença de indutores. Estudos mostram que, na ausência de indutores, estas proteínas se ligam perto da região -65, e, a presença do indutor causa interação da proteína com a região -35, que é o sítio de ligação da RNA polimerase, fazendo uma dobra no DNA, e, assim ativando a transcrição (revisto por SCHELL, 1993).

Algumas das LTTRs melhor conhecidas, bem como a função de seus genes alvo são: Nac (resposta a variações nos níveis de nitrogênio), e CysB (envolvida na

biossíntese de cisteína) são reguladas negativamente em relação à sua própria expressão; MleR (enzima maloláctica) e AlsR (síntese de acetoína) parecem não estar envolvidos na sua própria expressão; SpvR (síntese de fatores de virulência), alguns NodDs (regulação dos genes *nod* envolvidos na interação planta-bactéria), e PhcA (reguladores de fatores de virulência) apresentam regulação positiva onde a expressão é aumentada em cerca de 6 vezes na presença das mesmas. Nem todas as LTRs necessitam de indutores, como é o caso de NodD3 (gene envolvido no processo de nodulação em leguminosas) e CatM (genes alvo envolvidos no catabolismo de catecóis), AmpR (genes alvo envolvidos com a β -lactamase) e IciA (genes alvo envolvidos na inibição da iniciação da replicação de DNA *in vitro*) que agem como repressoras (revisto por SCHELL, 1993).

A proteína Nac ativa a expressão de operons envolvidos no catabolismo de fontes alternativas de nitrogênio, como a histidina (genes *hut*), prolina (genes *put*) e urease (genes *ure*). Ensaio *in vitro* com o operon *hutUH* mostram que a presença apenas de Nac é suficiente para que ocorra transcrição, e que esta proteína protege um fragmento de 26 pb, que se encontra a 64 pb a jusante do início de transcrição na região promotora de *hutU* (GOSS & BENDER, 1995).

CysB é um regulador positivo da expressão do regulon de biossíntese de cisteína, e o gene que o codifica *cysB* não está ligado a nenhum dos genes alvo de CysB, perdendo a típica divergência de promotores da família LysR. Ensaio de transcrição *in vitro* com CysB purificada mostram que esta se autoregula negativamente na ausência dos indutores N-acetilserina e O-acetilserina, enquanto que na presença de N-acetilserina CysB estimula a transcrição dos genes *cysJIH*, *cysK*, e *cysP* (OSTROWSKI & KREDICH, 1991). A ligação de CysB ao promotor de *cysK* induz uma dobra no DNA que é relaxada na presença do indutor (MONROE *et al.*, 1990).

CatM em *Acinetobacter calcoaceticus* regula os genes *cat* envolvidos na degradação de catecol (intermediário da degradação de benzoato) via β -cetoadipato (NEIDLE *et al.*, 1989). CatM ativa a expressão de *catA*, cujo produto é a enzima 1,2-catecol dioxigenase, que converte o catecol a *cis-cis* muconato (CCM); e do operon *catBCIJFD*, que codifica enzimas envolvidas na formação de intermediários do ciclo

dos ácidos tricarboxílicos em resposta ao indutor *cis-cis* muconato (ROMERO-ARROYO *et al.*, 1995).

No plasmídeo NAH7 de *Pseudomonas putida* é encontrado o gene *nahR* que codifica para NahR reguladora dos operons *nah* e *sal*, envolvidos no metabolismo de naftaleno e salicilato como fonte de carbono, respectivamente. Para que a transcrição destes operons ocorra é necessária a presença de salicilato e de NahR. O gene *nahR* é divergentemente transcrito em relação ao operon *sal*, e sua expressão é inibida pela proteína NahR (SCHELL, 1993).

SyrM é uma proteína regulatória envolvida na expressão do gene *nodD3* e dos genes *exo* (genes envolvidos com a síntese de polissacarídeos extracelulares), envolvidos com a invasão da planta pelo rizóbio. NodD3 ativa a expressão de *nodC-lacZ* de maneira independente de flavonóides. SyrM afeta a regulação dos genes *nod* e também dos genes *exo*, necessários para a invasão pela raiz da planta, e assim tem um papel central na simbiose (MULLIGAN & LONG, 1989; SWANSON *et. al.*, 1993).

Apesar de quase todas as LTTRs serem citoplasmáticas, a proteína NodD de *Rhizobium leguminosarum* está associada com a membrana citoplasmática, facilitando a interação com seus indutores hidrofóbicos (flavonóides), mais abundantes na membrana. Aparentemente, além das NodDs, somente a NhaR (envolvido com o transporte antiporte $\text{Na}^+ - \text{H}^+$) é encontrada associada à membrana, sendo as demais LTTRs proteínas citoplasmáticas (revisto por SCHELL, 1993).

LTTRs possuem regiões conservadas e a substituição de determinados resíduos de aminoácidos proporcionam um melhor entendimento sobre quais são os aminoácidos envolvidos na ligação ao DNA ou na ligação a indutores. A região de maior identidade entre as LTTRs encontra-se entre os resíduos 6-66 da região N-terminal, sendo que na porção central (resíduos 23-42), idêntica em 40% das LTTRs, está o motivo de ligação ao DNA do tipo hélice-volta-hélice (HTH – Helix-Turn-Helix). Mutações nestes 20 aminoácidos mostram que o HTH encontrado em LTTRs é diferente dos HTH clássicos, principalmente em relação a pouca conservação do resíduo da Gly na posição 9 do HTH e da Pro na posição 13 (posições consideradas a partir do resíduo 23 da LTTR, que no caso do HTH é considerado o 1º resíduo). Contudo, outros aminoácidos são altamente conservados, sendo idênticos em 70%

das LTTRs, como Ala27, Thr(Ser)22, Gln34, Pro35, Ser(Thr)38, Leu44, e Glu45 (posições consideradas a partir do aminoácido 1 da LTTR). Apesar da presença do HTH, outras regiões das LTTRs também são importantes na ligação ao DNA. Em NahR, substituições únicas em aminoácidos situados entre as posições 26 e 56, retirando Ala27 ou a Thr26 acabam com a habilidade desta proteína em se ligar especificamente aos seus genes-alvos e transcrevê-los. Na proteína NodD, a substituição de Ala40 ou Ser42 causa perda da habilidade de ligação ao DNA, e com isso uma diminuição em 15 vezes na ativação da transcrição estimulada por indutores. Assim, as substituições que afetam a ligação de LTTRs ao DNA são encontradas entre os resíduos 23 e 62, que se encontram próximas ou inseridas no motivo HTH (revisto por SCHELL, 1993).

O motivo LysR, como já mencionado acima, possui menor conservação de seqüência, e isso pode ser explicado pela grande variedade de indutores aos quais essas proteínas se ligam. Mesmo entre as mesmas proteínas, encontradas em diferentes microrganismos, como é o caso das proteínas NodDs, é possível identificar uma grande variabilidade nos resíduos de aminoácidos encontrados nesta região, uma vez que estas proteínas reconhecem diferentes indutores (flavonóides). Alguns tipos de substituições podem alterar a conformação da proteína de maneira que esta fique parecida com a proteína selvagem ligada ao indutor. Isso pode ser observado em relação a substituições nas posições 95, 123 ou 154 de *nodD* em *R. leguminosarum* que causam um aumento na ativação da transcrição de genes alvo de aproximadamente 15 vezes, na ausência do indutor. Em NahR, substituições nos aminoácidos 116 ou 154, e em CysB, nos aminoácidos 149 e 165 também resultam em um fenótipo independente de indutores. Talvez seja essa a região (do resíduo 102-154) envolvida com a ligação ao indutor (revisto por SCHELL, 1993).

Dentro do motivo LysR, a região com maior similaridade de seqüência encontra-se entre os resíduos 236-246, e mutações nestes aminoácidos causam perda da capacidade destas proteínas em se ligarem ao DNA ou perda da capacidade de ligação a indutores (revisto por SCHELL, 1993).

2.3 PROTEÍNA NodD DE RIZÓBIOS

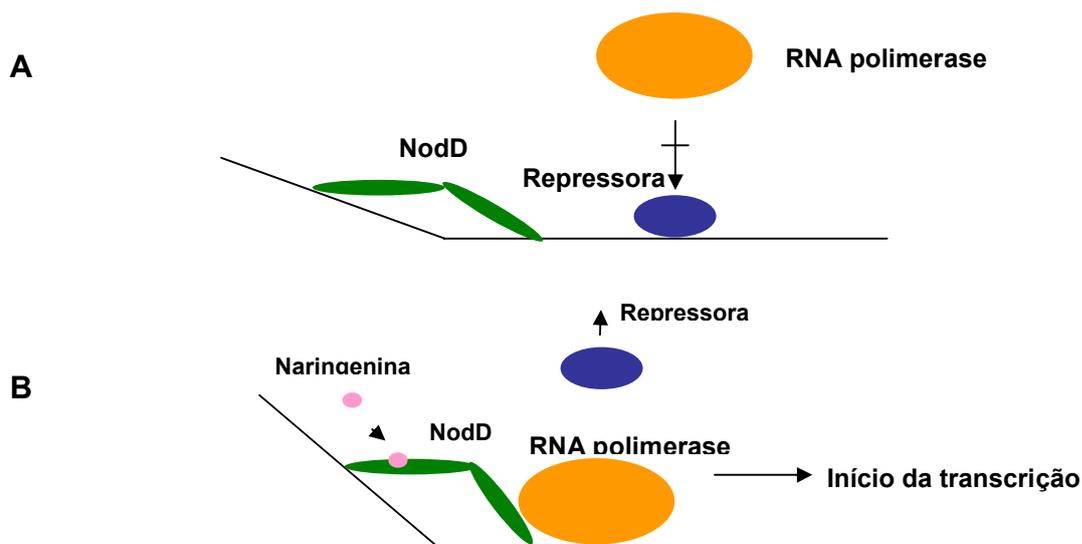
A proteína NodD pertence à família LysR de reguladores de transcrição, e em rizóbios ativa a transcrição de genes envolvidos no processo de nodulação nos estágios iniciais do processo de simbiose com leguminosas. Esta proteína se liga a uma seqüência de DNA chamada *nod*-box encontrada na região promotora, mas geralmente só é capaz de ativar a expressão na presença de flavonóides (FISHER *et al.*, 1988). Apresenta conservação de seqüência na região N-terminal, mas perde similaridade de seqüência na região C-terminal (HENIKOFF *et al.*, 1988; GOETHALS *et al.*, 1992; revisto por SCHLAMMAN *et al.*, 1992a). Estas proteínas possuem características comuns de proteínas LysR. Normalmente tem um tamanho médio de 32 a 36 kDa; possuem um motivo de ligação ao DNA do tipo hélice-volta-hélice (HTH) altamente conservado na região N-terminal (HENNIKOF *et al.*, 1988; revisto por SCHLAMMAN *et al.*, 1992a), perdendo homologia de seqüência na região C-terminal; frequentemente inibem a sua própria transcrição. Além disso, estudos com proteínas NodD mutantes indicam que a ligação ao indutor muda a estrutura tridimensional da proteína, uma vez que alguns tipos de mutações causam a ativação da transcrição dos genes-alvo mesmo na ausência de indutores (revisto por SCHLAMMAN *et al.*, 1992; GOETHALS *et al.*, 1992).

NodD usualmente não se liga a seqüências longas de DNA, e estas geralmente contêm dois subsítios de ligação de proteína. Devido ao ancoramento nos dois subsítios, estas proteínas podem causar uma dobra no DNA-alvo que estaria envolvido com a maneira de regulação das proteínas reguladoras de transcrição da família LysR (FENG *et al.*, 2003).

A proteína NodD de *R. leguminosarum* parece ser tetramérica e se liga aos dois subsítios *nod*-box causando uma dobra no DNA, tanto na presença como na ausência de indutores, impedindo que a RNA polimerase forme um complexo transcricional aberto ativo (Figura 4A). Além disso, na região promotora dos genes *nod* há uma proteína repressora que mantém a expressão destes genes reprimidos. Porém, na presença de naringenina a NodD aumenta a dobra no DNA, permitindo o início da transcrição pela RNA polimerase, uma vez que libera o repressor de seu sítio,

permitindo a formação de um complexo transcricional ativo (Figura 4B). Contudo, a identidade do repressor e o sítio onde se liga na região promotora ainda não são conhecidos (CHEN *et al.*, 2005). Segundo HU *et al.*, 2000, em *R. leguminosarum*, a NodD age como repressora da sua própria transcrição competindo com a RNA polimerase pela ligação na região promotora do gene *nodD*. Assim, a autorregulação negativa ocorre devido ao bloqueio do acesso da RNA polimerase à região promotora de *nodD* pela própria proteína NodD, sendo que esse bloqueio ocorre na ausência de flavonóides. Para que a proteína NodD deixe de ser repressora de sua própria transcrição é necessária a presença de um flavonóide. Em relação à transcrição de outros genes *nod*, em *R. leguminosarum* a NodD ativa a transcrição de tais genes na presença de flavonóides.

Em *R. leguminosarum* foi proposto que a proteína NodD está associada à monocamada interna da membrana citoplasmática, auxiliando na resposta a indutores hidrofóbicos presentes na membrana, como os flavonóides (SCHLAMMAN *et al.*, 1989). Em *R. meliloti* a proteína NodD migra para a membrana citoplasmática apenas na presença de um determinado indutor, permanecendo livre no citoplasma quando o indutor está ausente (SCHLAMMAN *et al.*, 1992). Contudo, ainda não se conhece totalmente o mecanismo de como a proteína NodD responde aos indutores para iniciar a transcrição dos genes *nod* (CHEN *et al.*, 2005).



FONTE: CHEN *et al.*, 2005

FIGURA 4 – MODELO PROPOSTO PARA A ATIVAÇÃO DA TRANSCRIÇÃO DE GENES *nod*.

O número de seqüências de proteínas NodD encontrados em rizóbios varia de espécie para espécie. *R. leguminosarum* bv. *trifolii* possui apenas uma proteína NodD, enquanto *Bradyrhizobium japonicum*, *Rhizobium* sp. NGR234, *R. meliloti* e *R. tropici* possuem de 2 a 4 proteínas NodDs (van RHIJIN *et al.*, 1993). A vantagem em ter mais de uma proteína NodD está na variedade de moléculas indutoras reconhecíveis.

Em *Rhizobium* sp. BR816 foram encontradas 4 proteínas NodD sendo todas funcionais, porém de maneiras diferentes em relação à nodulação. A ativação do promotor *nodABC* na presença de flavonóides é feita por NodD1 e NodD2, enquanto NodD3 e NodD4 não ativaram a expressão de outros genes *nod* na presença dos flavonóides testados. A apigenina foi quem apresentou maior indução da atividade de NodD1 e NodD2, sendo que NodD2 é a que apresenta maior atividade (van RHIJIN *et al.*, 1994).

Em *Rhizobium* sp. NGR234 há duas proteínas NodD. NodD1 é ativada por vários flavonóides, e isso explica a grande variedade de plantas-hospedeiras com que NGR234 é capaz de se associar, uma vez que mutantes *nodD1*⁻perdem completamente a capacidade de induzir a formação de nódulos radiculares. NodD2 está envolvida na repressão do operon *nodABC*, uma vez que na estirpe mutante *nodD2*⁻ há aumento, em cerca de 5 vezes, na produção de fatores Nod, em relação à estirpe selvagem. Isso acarreta na formação de nódulos Fix⁻ e nódulos vazios em *Vigna unguiculata* e *Tephrosia vogelii*, respectivamente. Porém em *Leucaena leucocephala* a mutação em *nodD2* de NGR234 não afeta a nodulação, indicando que esta NodD2 está envolvida com a especificidade de hospedeiro, inibindo a produção de fatores Nod, e, assim, permitindo o estabelecimento do bacteróide (FELLAY *et al.*, 1998).

R. meliloti possui 3 genes *nodD* localizados no seu megaplasmídeo, chamadas *nodD1*, *nodD2*, *nodD3*. Dentre as proteínas codificadas por esses genes, NodD1 e NodD2 ativam a expressão de outros genes *nod* na presença de flavonóides (HONMA & AUSUBEL, 1987). Contudo, NodD3 precisa de SyrM para que se torne ativadora (item 2.1)

2.3.1 *nod*-box

O sítio de ligação específico da proteína NodD, denominado *nod*-box está presente na região promotora dos genes *nod*, e contém, aproximadamente 37 pb. O alinhamento entre os *nod*-box de vários genes *nod* de *R. leguminosarum* biovars *viciae* e *trifolii*, *R. galagae*, e *S. meliloti* mostra a alta conservação destas seqüências (SUOMINEM *et al.*, 1999), sendo que a região central do *nod*-box é a que apresenta maior conservação (GOETHALS *et al.*, 1992). A presença de um *nod*-box sugere a presença de genes regulados por flavonóides na região (SUOMINEM *et al.*, 1999).

A organização dos nucleotídeos de um *nod*-box ocorre em seqüências palindrômicas ATC-N₉-GAT, e foi proposto que cada *nod*-box possui duas seqüências ATC-N₉-GAT, uma distal e outra proximal em relação ao 1º nucleotídeo do gene, separadas por 7 pb. Assim, os 25 pb da região central conservada são encontrados entre o GAT do subsítio distal do *nod*-box até o GAT do subsítio proximal. Cada NodD dimérica se liga a um subsítio do *nod*-box (GOETHALS *et al.*, 1992) o que auxilia na explicação da hipótese de que a proteína NodD se liga ao *nod*-box na forma de um tetrâmero (SCHLAMMAN *et al.*, 1992; GOETHALS *et al.*, 1992; CHEN *et al.*, 2005). Dentro da seqüência ATC-N₉-GAT está a seqüência reconhecida pelo motivo LysR, T-N₁₁-A, que é a seqüência reconhecida pela proteína NodD, isto é, onde a proteína se liga ao DNA (GOETHALS *et al.*, 1992; SCHLAMMAN *et al.*, 1992). No subsítio distal o motivo LysR está localizado entre -75 e -50 pb a montante do início de transcrição, sendo que neste subsítio do *nod*-box de *A. caulinodans* se um T ou A for trocado por G ou T, respectivamente, na seqüência T-N₁₁-A, a proteína NodD perde a capacidade de ativar a transcrição (GOETHALS *et al.*, 1992). O subsítio proximal ATC-N₉-GAT (50 a -25 pb) é sobreposto ao sítio distal (-35pb) da RNA polimerase, e por isso é este sítio nem sempre é tão conservado quanto o subsítio distal (FENG *et al.*, 2003).

De uma maneira geral, a região de um *nod*-box protegida pela proteína NodD é a mesma tanto na presença quanto na ausência de indutores, porém foi observado que em *R. meliloti* AK41 e em *A. caulinodans* a proteína NodD tem maior afinidade pelo *nod*-box na presença de indutores (SCHLAMMAN *et al.*, 1992).

2.3.2 Genes *nod* e sua regulação em rizóbios

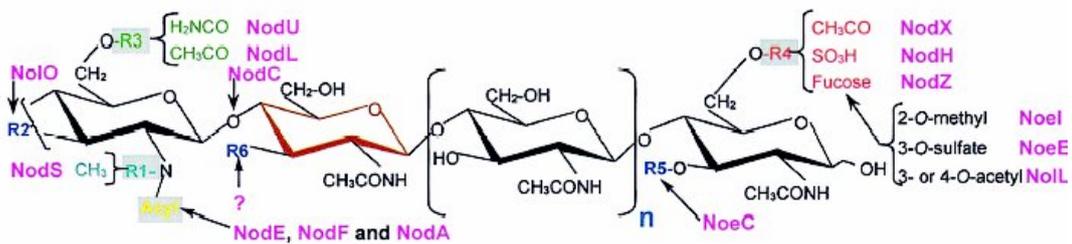
Os genes tipo *nod* podem ser classificados em estruturais (ou comuns), que são os genes *nodABC* e específicos como o gene *nodS* e *nodEF* (DÉNARIÉ *et al.*, 1992). Além dos genes *nod*, os genes *nol* e *noe* também são necessários para a nodulação, estando envolvidos na biossíntese dos Fatores Nod. A expressão destes genes também é regulada pela proteína NodD (FENG *et al.*, 2003; SCHLAMMAN *et al.*, 1992). A expressão destes genes se inicia ainda na rizosfera, antes da entrada do rizóbio na raiz (SCHLAMMAN *et al.*; 1992), e com exceção do gene *nodD* nos casos em que sua expressão é constitutiva, os demais genes de nodulação não são expressos continuamente durante o processo de nodulação (MYLONA *et al.*, 1995).

Os genes comuns *nodABC* são encontrados em todos os rizóbios e, quando delecção, as estirpes mutantes perdem a habilidade de induzir o enrolamento dos pêlos radiculares, em formar o cordão de infecção, e em induzir a divisão das células corticais, que são os processos que originam os nódulos radiculares (DÉNARIÉ *et al.*, 1992). Os produtos dos genes *nodABC* são necessários para a síntese da cadeia central do lipo-chito-oligossacarídeo, e parecem ser funcionalmente semelhantes em várias espécies (ROCHE *et al.*, 1996).

Os fatores Nod têm uma estrutura básica similar, composta por um chito-oligossacarídeo, que é uma cadeia linear de 3 ou 5 moléculas de N-acetilglucosamina com ligações β -1,4, e possui um radical acil ligado no resíduo de açúcar não-redutor (Figura 2). O gene *nodC* codifica a enzima NodC, uma N-acetilglucosamiltransferase, que sintetiza o oligômero de quitina; NodB, produto do gene *nodB*, é uma chito-oligossacarídeo desacetilase que remove a porção N-acetil da N-acetilglucosamina da extremidade não-redutora do oligossacarídeo; e NodA é uma acil-transferase que liga uma cadeia acil ao carbono 2 da extremidade não-redutora do oligossacarídeo (revisto por PERRET *et al.*, 2000). Modificações na estrutura básica são decorrentes da atividade de genes *nod* específicos, cuja presença varia amplamente nos diversos rizóbios (MYLONA *et al.*, 1995).

Os genes *nod* específicos estão envolvidos na especificidade pelo hospedeiro, sendo essenciais para a nodulação de certas plantas, mas não para outras (LONG,

1996). Mutações em genes *nod* específicos não bloqueiam a nodulação, mas tornam demoradas a formação ou diminuem o número de nódulos formados, e ainda podem alterar o número de plantas-hospedeiras da estirpe (DÈNARIÈ *et al.*, 1992). O produto de *nodEF* é responsável pela adição de ácidos graxos N-ligados ao terminal não redutor da porção açúcar. O produto de *nodZ* é responsável pela fucosilação do fator Nod, e o produto do gene *nodH* adiciona um grupamento fosfato ao fator Nod. Estas modificações, dentre outras mostradas na figura 5, na estrutura básica do fator Nod são responsáveis pela especificidade (revisito por PERRET, 2000).



FONTE: PERRET *et al.*, 2000

FIGURA 5 – ESTRUTURA BÁSICA DE UM FATOR NOD. A figura mostra possíveis substituições (R) e os genes responsáveis pela substituição, sendo estas substituições responsáveis pela especificidade; são os genes *nod* específicos.
n- número de N-acetilglucosaminas

2.4 FLAVONÓIDES

Para que ocorra a ativação da expressão dos genes *nod*, via NodD, é necessária a presença de flavonóides presentes em exudatos de raízes de plantas. A natureza química e quantidade de flavonóides liberados pela planta dependem da planta e de seu estágio de desenvolvimento. Os flavonóides são liberados como agliconas ou conjugados glicosídicos de baixa atividade e alta solubilidade, sendo convertidos à forma ativa por glicosidases presentes na bactéria (DÈNARIÉ *et al.*, 1992).

Flavonóides são metabólitos secundários, extremamente diversos, sendo encontrados em plantas vasculares e alguns musgos, e contém mais de 9000 estruturas descritas (MITHÖFER & MARTENS, 2005). São formados por 2 anéis benzênicos (A e B) conectados por um anel pirano contendo um átomo de oxigênio

(figura 6) (WILLIAMS *et al.*, 1986), apresentando um total de 15 átomos de carbono. As diferentes classes de flavonóides se devem ao nível de oxidação do anel C, enquanto que as diferenças entre os compostos dentro de uma mesma classe, se devem a diferenças nas configurações dos substituintes nos anéis A e B (REDDY *et al.*, 2007). Nas plantas, os flavonóides são encontrados como glicosídeos, e ao menos 8 diferentes monossacarídeos podem se ligar aos grupos hidroxila da forma aglicona (que perdeu a porção açúcar) do flavonóide, resultando nos inúmeros tipos de flavonóides descritos. As unidades glicosídicas mais comuns, encontradas nos flavonóides, são a D-glucose e a L-ramnose (WILLIAMS *et al.*, 1986). A principal reação da via de biossíntese de flavonóides (figura 7) é a condensação de uma molécula de *p*-cumaril-CoA a 3 moléculas de malonil-CoA, originando uma chalcona intermediária (tetrahidroxi-chalcona). A partir desta chalcona é que as demais classes de flavonóides se originam (REDDY *et al.*, 2007).

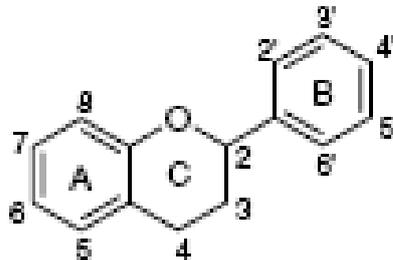
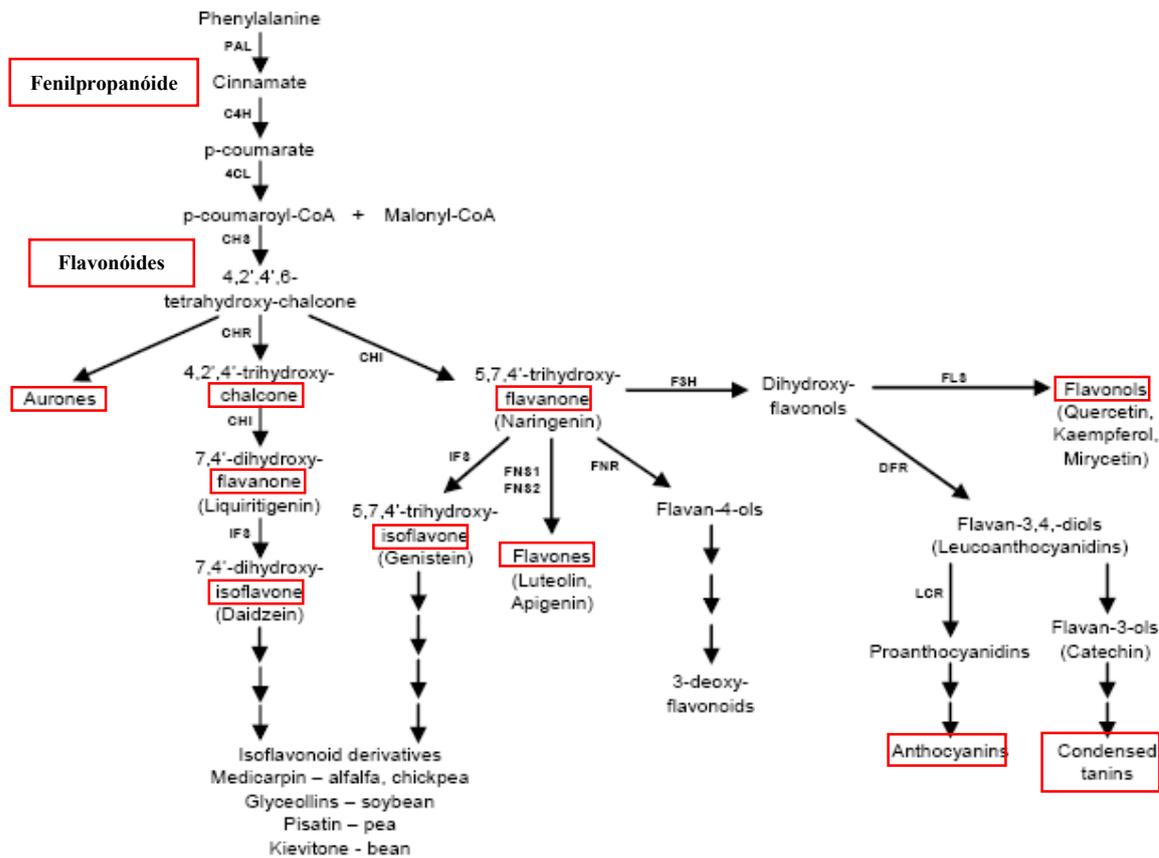


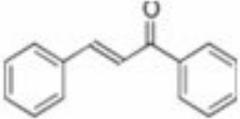
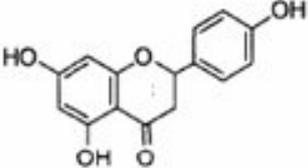
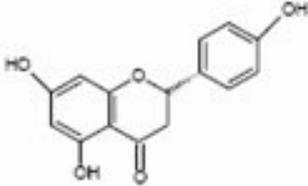
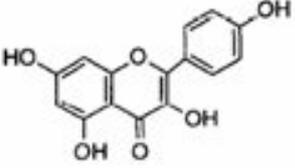
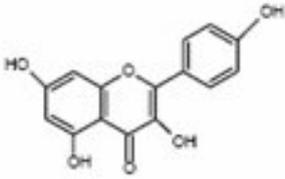
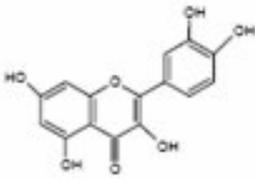
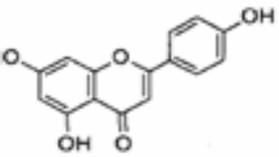
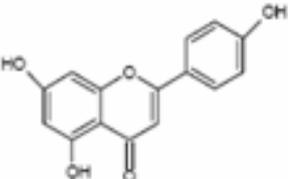
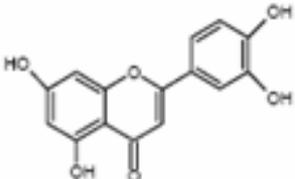
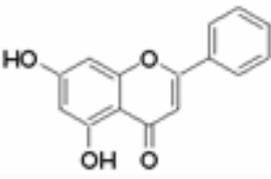
FIGURA 6 – ESTRUTURA BÁSICA DE UM FLAVONÓIDE. Os flavonóides contêm os anéis A, B e C, sendo que os carbonos do anel B levam um apóstrofe para diferenciar dos carbonos dos anéis A e C (WILLIAMS, 1986).



REDDY *et al.*, 2007

FIGURA 7 – VIA DE BIOSÍNTESE DOS FLAVONÓIDES. As principais classes de flavonóides e outros produtos estão nas caixas vermelhas. PAL, fenilalanina amônia-liase; 4CL, 4-cumarato-CoA ligase; C4H, ácido 4-hidroxilase cinâmico; CHI, chalcona isomerase, CHR, chalcona redutase; CHS, chalcona sintase; DFR, dihidroflavonol 4-redutase; F3H, flavona 3-hidroxilase; FLS, flavonol sintase; FNR, flavona 4-redutase; FNS, flavona sintase; IF8, isoflavona sintase; LCR, leucoantocianidina redutase; PAL, fenilalanina amônia liase.

A estrutura química das classes de flavonóides, bem como as estruturas dos flavonóides utilizados neste trabalho estão na figura 8.

Classe	Estrutura	Exemplo	
Chalcona			
Flavonona		 Naringenina	
Flavonol		 Kaempferol	 Quercetina
Flavona		 Apigenina	 Luteolina
	 Crisina		

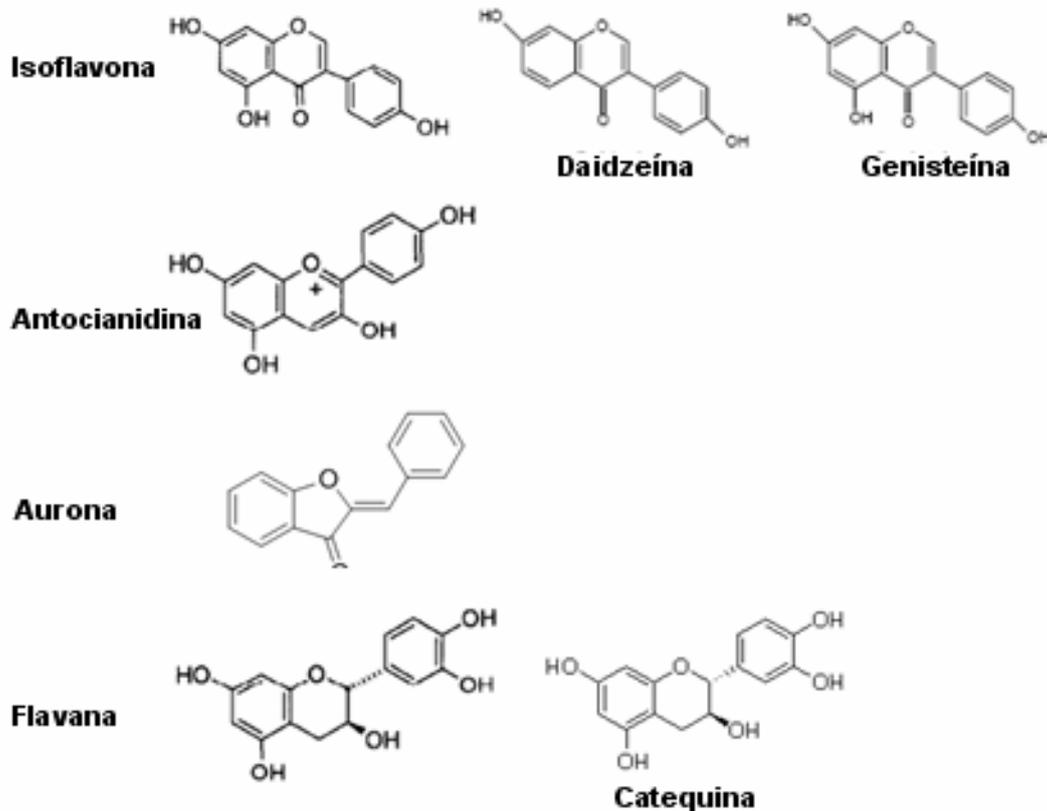


FIGURA 8 – ESTRUTURAS BÁSICAS DE CADA CLASSE DE FLAVONÓIDES. Dentro de cada classe são mostradas as estruturas dos compostos que foram utilizados no decorrer deste trabalho.

As diversas proteínas NodD de diferentes rizóbios reconhecem flavonóides específicos, com um estrutura distinta. Proteínas NodD de rizóbios que possuem pouca variedade de hospedeiros, como *R. meliloti*, *R. leguminosarum*, e *R. trifolii* respondem a poucos flavonóides. Por outro lado proteínas NodD de rizóbios com grande variedade de hospedeiros, como é o caso de *Rhizobium* sp. NGR234, respondem a vários flavonóides (BASSAM *et al.*, 1988).

Usualmente, as flavonas e flavononas são os indutores de NodD de rizóbios, porém a NodD de *Bradyrhizobium* spp. é geralmente ativada pelas isoflavonas. Contudo, as plantas também podem liberar flavonóides que agem como anti-indutores (SCHLAMMAN *et al.*, 1992). Geralmente estes anti-indutores apresentam estruturas similares a dos indutores e a inibição pode ser superada com um aumento na concentração do indutor (DÈNARIÈ *et al.*, 1992). Uma característica relevante encontrada em flavonóides indutores de genes *nod* é o radical hidroxil em dois dos

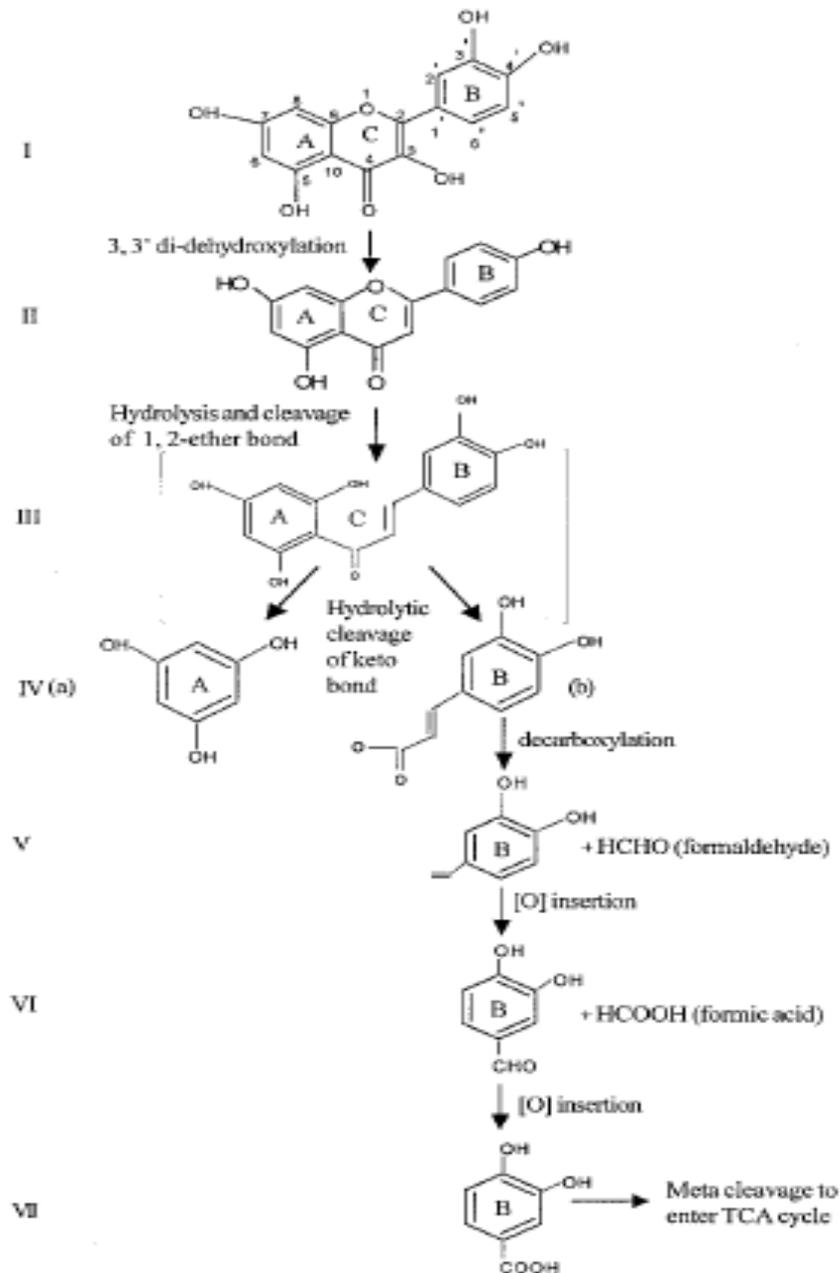
carbonos que formam a estrutura básica de um flavonóide (C-7 e C-4) (BRENCIC & WINANS, 2005). Peters & Long, 1988, observaram que a luteolina é o flavonóide que apresenta maior indução em relação aos genes *nod* em *R. meliloti*. Ao adicionar certos flavonóides na presença de luteolina, esta indução é diminuída, ou seja, estes flavonóides agem como anti-indutores da luteolina. Analisando estruturalmente estes compostos, foi proposto que qualquer modificação na estrutura da luteolina causa essa diminuição da atividade. A adição de grupos hidroxila na posição 2' (morina) ou 3' (quercetina), a remoção de grupos hidroxila na posição 3' (apigenina) e/ou 4' (crisina), ou saturação da ligação entre C2-C3 (eriodictiol) diminuem ou eliminam por completo a atividade da molécula. A produção de compostos indutores ou de anti-indutores de genes *nod* está intimamente relacionada com o estágio de desenvolvimento da raiz, sendo que evidências apontam para o fato de que a produção de indutores está localizada na parte de alongação da raiz que é a parte mais sujeita à infecção pelos rizóbios (REDDY *et al.*, 2007).

A naringenina (flavonona) é, preferencialmente, o indutor dos genes *nod* de *R. leguminosarum*, sendo encontrada na membrana citoplasmática (membrana interna) onde a proteína NodD está localizada (DÈNARIÈ *et al.*, 1992). Apesar de sua hidrofobicidade natural, estudos mostraram que a naringenina tem maior afinidade pela membrana citoplasmática interna em um pH de aproximadamente 5,7 (RECOURT *et al.*, 1989).

Em várias espécies de rizóbios foi observado que o catabolismo de flavonóides se inicia pela quebra de ligações no anel C, liberando os anéis A e B, que, então, formam outros compostos. O mesmo foi observado na bactéria *Pseudomonas putida*, e esta cliva o flavonóide no anel A (RAO & COOPER, 1994), mostrando, assim, que a bactéria é responsável pela especificidade em relação ao mecanismo de clivagem. O mecanismo de clivagem pode ser visto como o oposto do mecanismo de biossíntese realizado pelas plantas, onde ocorre a condensação de 3 moléculas de malonil-CoA (compõem anel A) com 1 molécula de *p*-cumaril-CoA (componentes dos anéis B e C), para formar a estrutura básica dos flavonóides (RAO & COOPER, 1994).

P. putida PML2 foi descrita como uma estirpe capaz de utilizar vários flavonóides, como flavonóis, flavononas, flavonas e isoflavononas como fonte de

carbono, sendo a quercetina (flavanol) um dos mais rapidamente metabolizados por este microrganismo. A via proposta para o catabolismo da quercetina em *P. putida* é mostrada na figura 9 (PILLAI & SWARUP, 2002).



PILLAI & SWARUP, 2002

FIGURA 9 – VIA DE DEGRADAÇÃO DA QUERCETINA EM *P. putida* PML2. (I) Quercetina (3,3',4',5,7-pentahidroxi flavona); (II) Naringenina (4',5,7-trihidroxi flavonona); (III) intermediário instável; (IVa) Phloroglucinol; (IVb) 3,4-dihidroxi ácido cinâmico; (V) 3,4-dihidroxi estireno; (VI) aldeído protocatecuico; (VII) ácido protocatecuico. Todos os compostos foram encontrados na estirpe selvagem PML2, mas não na estirpe mutante Flav 1-9.

Os resultados sugerem que a quercetina é dehidroxilada à naringenina antes de ser convertida em compostos menores. A via acima apresentada parece o oposto da via de biossíntese realizada pelas plantas, onde a hidroxilação da naringenina origina a quercetina (PILLAI & SWARUP, 2002). O ácido protocatecuíco originado pela degradação da quercetina, entra na via meta de clivagem de anéis aromáticos, e por várias clivagens origina os intermediários succinil-CoA e acetil-CoA que entram na via dos ácidos tricarbóxicos (HARWOOD & PARALES, 1996).

Alguns trabalhos mostram que os flavonóides podem agir como reguladores de crescimento em algumas espécies de rizóbios (REDDY *et al.*, 2007). Em relação ao crescimento de *R. meliloti* foi observado que os flavonóides contendo hidroxilas nas posições C5 e C7 (ex. luteolina e quercetina) aumentam a taxa de crescimento deste microrganismo, e este efeito provavelmente não envolve a proteína NodD, uma vez que mesmo uma estirpe, de *R. meliloti*, mutada em todos os genes *nodD* mantém a resposta ao crescimento na presença de luteolina e quercetina (HARTWIG *et al.*, 1991; REDDY *et al.*, 2007). Contudo, estudos mostram que o flavonóide que induz o crescimento pode não ser o mesmo que induz a ativação dos genes *nod* via NodD, como é o caso da quercetina que aumenta o crescimento mas não induz a expressão da fusão *nodC::lacZ* por NodD1 em *R. meliloti* (HARTWIG *et al.*, 1991; PETERS & LONG, 1988). Em *Bradyrhizobium japonicum* a luteolina é o flavonóide que aumenta a taxa de crescimento (REDDY *et al.*, 2007).

3 OBJETIVOS

3.1 OBJETIVOS GERAIS

Este projeto tem como objetivo estudar a função do gene *nodD* encontrado no genoma de *H. seropedicae*, assim como determinar a função do operon localizado a montante deste gene, cuja expressão é regulada pela proteína HsNodD.

3.2 OBJETIVOS ESPECÍFICOS

- Obter uma estirpe mutante de *H. seropedicae*, no gene *orf1*, primeiro gene do operon a montante e com transcrição divergente ao gene *nodD*;
- Finalizar o sequenciamento do operon entre os contigs 186 e 232, dentro da *orf6* do operon em estudo;
- Caracterizar através de ferramentas de bioinformática todo o operon a montante do gene *nodD*;
- Determinar quais compostos fenólicos são degradados por *H. seropedicae*.

4 MATERIAL E MÉTODOS

4.1 BACTÉRIAS E PLASMÍDEOS

As bactérias e plasmídeos utilizados neste trabalho estão descritos na tabela 1.

TABELA 1 – BACTÉRIAS E PLASMÍDEOS

Estirpes	Características	Referência
<i>Herbaspirillum</i>		
<i>seropedicae</i>		
SmR1	Estirpe selvagem, Sm ^R	MACHADO <i>et al.</i> , 1996
DR2	Inativação cromossomal do gene <i>nodD</i> através da dupla recombinação homóloga e inserção do cassete de canamicina, Sm ^R , Km ^R	CORAL, 2005
AMM1	Inativação cromossomal do gene <i>fdeA</i> através da simples recombinação homóloga e inserção do gene de resistência a canamicina, Sm ^R , Km ^R	Este trabalho
<i>Escherichia coli</i>		
S17.1	Sm ^R	SIMON <i>et al.</i> , 1983
DH10B	Sm ^R	INVITROGEN
Plasmídeos		
pGEM-T Easy	Ap ^R	PROMEGA
pCR 2.1	Ap ^R	INVITROGEN
pORF1	Ap ^R , fragmento de 512 pb da região codificadora do gene <i>orf1</i> (<i>fdeA</i>) do operon a montante do gene <i>nodD</i> de <i>H. seropedicae</i> SmR1, clonado em pCR 2.1	Este trabalho
pGAP	Ap ^R , fragmento de 900 pb da região codificadora da <i>orf6</i> do operon a montante do gene	Este trabalho

PMP-NB8	<i>nodD</i> de <i>H. seropedicae</i> SmR1, clonado em pGEM-T Easy Tc ^R , fragmento de 300 pb <i>KpnI/XbaI</i> , contendo a região promotora do operon <i>nodABC</i> de <i>Rhizobium</i> sp. NGR234, clonado em pMP220	FELLAY <i>et al.</i> , 1998
pSU11	Tc ^R , fragmento de 820 pb <i>EcoRI/PstI</i> , contendo a região promotora do gene <i>nodD</i> de <i>H. seropedicae</i> clonado em pPW452	KARP, 2004
pSUI	Tc ^R , fragmento de 820 pb <i>EcoRI/PstI</i> , contendo a região promotora do gene <i>orf1 (fdeA)</i> de <i>H. seropedicae</i> clonado em pMP220	MARIN, 2006a
HS23-RN-00-000-094-C11	Amp ^R , <i>orf6 (fdeF)</i> do operon a montante do gene <i>nodD</i> de <i>H. seropedicae</i> . Clone que liga os contigs 186 e 232.	Programa GENOPAR
HS25-EG-00-000-146-G01	Amp ^R , <i>orf6 (fdeF)</i> do operon a montante do gene <i>nodD</i> de <i>H. seropedicae</i> . Contig 186.	Programa GENOPAR

FONTE: O autor (2008)

4.2 MEIOS DE CULTURA E CONDIÇÕES DE CRESCIMENTO

4.2.1 Cultivo de *E. coli*

As estirpes de *E. coli* foram cultivadas em meio LB a 37°C (SAMBROOK *et al.*, 1989), sob agitação à 150 rpm. A composição do meio LB está descrita abaixo:

Triptona	10 g/L
Extrato de levedura	5 g/L
NaCl	10 g/L
Água destilada q.s.q.	1 L

O meio LA, utilizado para o cultivo das estirpes de *E. coli* em meio sólido foi obtido pela adição de ágar (15 g/L) ao meio líquido LB.

Para o preparo de células eletrocompetentes o meio utilizado foi o SOB (HANAHAN *et al.*, 1983), cuja composição está descrita abaixo:

Bacto triptona	20 g/L
Extrato de levedura	5 g/L
Cloreto de sódio	0,5 g/L
Cloreto de potássio	0,186 g/L

O meio SOC foi obtido pela adição de 20 mmol/L de glucose, e 5 mL/L de cloreto de magnésio, e 2,4 g/L de sulfato de magnésio ao meio SOB.

4.2.2 Cultivo de *Herbaspirillum seropedicae*

Para o cultivo das estirpes de *H. seropedicae* foi utilizado o meio NFb-malato (KLASSEN *et al.*, 1997) acrescido de 50 mL/L de mistura de fosfatos (K_2HPO_4 17,8 g/L; KH_2PO_4 159,5 g/L; pH 5,8), que foi denominado NFbHP-malato. Foi utilizado 20 mmol/L de cloreto de amônio como fonte de nitrogênio (onde não foi utilizado fonte de nitrogênio está indicado) e o meio foi denominado NFbHPN-malato. A cultura foi incubada a 30°C em agitador rotatório a 120 rpm. A composição de sais do meio NFb-malato segue abaixo:

$MgSO_4 \cdot 7H_2O$	2×10^{-1} g/L
NaCl	1×10^{-1} g/L
$CaCl_2$	2×10^{-2} g/L
Ácido nitrilo-triacético (NTA)	$5,6 \times 10^{-2}$ g/L
$FeSO_4 \cdot 7H_2O$	2×10^{-2} g/L
Biotina	1×10^{-4} g/L
Solução de microelementos	10 mL/L
Malato de potássio	5 g/L

Solução de microelementos:

$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	1 g/L
$\text{MnSO}_4 \cdot \text{H}_2\text{O}$	1,175 g/L
H_3BO_3	1,4 g/L
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	4×10^{-2} g/L
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	$1,2 \times 10^{-1}$ g/L

Os meios NFbHPN-malato sólido e semi-sólido foram obtidos após adição de 15g/L e 1,75 g/L de ágar, respectivamente, ao meio líquido.

Para determinar a capacidade de crescimento das estirpes de *H. seropedicae* em fontes alternativas de carbono foi utilizado o meio NFbHP, mas sem adição de malato de potássio e a fonte de carbono está indicada.

4.2.3 Antibióticos utilizados

A tabela 2 abaixo mostra os antibióticos utilizados e suas concentrações para o cultivo de *E. coli* e *H. seropedicae*.

TABELA 2 – ANTIBIÓTICOS UTILIZADOS PARA O CULTIVO DE *E. coli* E DE *H. seropedicae*

antibiótico	<i>E. coli</i>	<i>H. seropedicae</i>
Estreptomicina (Sm)	10 µg/mL	80 µg/mL
Tetraciclina (Tc)	10 µg/mL	10 µg/mL
Canamicina (Km)	50 µg/mL	500 µg/mL
Ampicilina (Ap)	250 µg/mL	*

* este antibiótico não foi utilizado para o crescimento desta bactéria

4.3 PURIFICAÇÃO DE DNA EM PEQUENA ESCALA

A extração de plasmídeos de células de *E. coli* foi realizado segundo o método de lise alcalina (SAMBROOK *et al.*, 1989), com modificações.

Estirpes de *E. coli*, contendo os plasmídeos de interesse foram crescidas por aproximadamente 16 horas em meio LB, contendo os antibióticos adequados. As culturas crescidas foram transferidas para tubos plásticos de 1,5 mL, centrifugadas a 13.400 rpm (centrífuga mini-spin EPPENDORF) por 2 minutos e o precipitado ressuspense em 200 µl de tampão GET (Tris.HCl 25 mmol/L, pH 8,0; glucose 50 mmol/L e EDTA 10 mmol/L). A lise das células ocorreu após a adição de 200 µl de solução de lise (NaOH 0,2 mol/L e SDS 1% (v/v)). Em seguida foram adicionados 200 µl de acetado de potássio 3 mol/L e ácido fórmico 1,8 mol/L pH 5,3, e a mistura foi centrifugada por 10 minutos. Ao sobrenadante foi adicionado 200 µl de fenol:clorofórmio:álcool isoamílico (25:24:1), a amostra foi agitada, e centrifugada por 10 minutos. O sobrenadante foi coletado e o DNA plasmidial foi precipitado pela adição de 1 volume de isopropanol, e centrifugado por 20 minutos. O precipitado foi lavado com 500 µl de etanol 80% e centrifugado por 5 minutos. O sobrenadante foi dispensado, e o precipitado foi seco à vácuo e dissolvido em 20 µl de água ultra pura estéril.

Os plasmídeos purificados foram analisados por eletroforese em gel de ágar 1% corado com brometo de etídeo e visualizados sob luz ultravioleta (SAMBROOK *et al.*, 1989).

4.4 PURIFICAÇÃO DO DNA GENÔMICO DE *Herbaspirillum seropedicae*

A purificação do DNA genômico foi realizada a partir de 10 mL de cultura, crescida em meio NFb malato líquido, contendo os antibióticos necessários. Dez mL de cultura foram centrifugados a 10.000 rpm por 10 minutos, o precipitado lavado em 10 mL de tampão GET (Tris.HCl 25 mmol/L, pH 8,0; glucose 50 mmol/L e EDTA 10 mmol/L), e centrifugado a 10.000 rpm por 10 minutos. As células foram ressuspensas em 500 µl de GET, e foram lisadas pela adição de SDS 1% e incubação a 55°C por 30 minutos. A degradação das proteínas foi feita pela adição de pronase e incubação a 37°C por aproximadamente 16 horas. A extração de proteínas foi realizada uma vez com fenol:clorofórmio:álcool isoamílico (25:24:1), e uma vez com clorofórmio:álcool isoamílico (24:1). Após a extração, o DNA foi precipitado com a adição de um volume

de isopropanol, lavado com etanol 80%, seco à vácuo e ressuspenso em 100 µl de água ultra pura estéril. O DNA foi analisado por eletroforese em gel de agarose 1%.

4.5 ELETROFORESE DE DNA

A eletroforese de DNA foi realizada em gel de ágar ou agarose em cuba horizontal (SAMBROOK *et al.*, 1989). Para a corrida eletroforética, foi utilizado o tampão TBE 1X, e o gel foi tratado com brometo de etídeo (0,5 µg/mL) e visualizado sob luz ultravioleta a 312 nm em transluminador EC₃ System - UVP BioImaging Systems (UVP, Inc. Upland, CAIUSA).

4.6 DIGESTÃO DE DNA COM ENDONUCLEASES DE RESTRIÇÃO

A digestão de DNA foi feita de acordo com o fabricante da enzima de restrição Invitrogen, GE Healthcare e Fermentas.

4.7 PREPARO DE VETORES

Os vetores foram incubados para a digestão com endonucleases de interesse, e após o período de 3 horas as enzimas foram inativadas por aquecimento ou pela adição de 40% (v/v) de fenol:clorofórmio:álcool isoamílico (25:24:1), dependendo da recomendação do fabricante. Após este tratamento o DNA foi precipitado com a adição de 1 volume de isopropanol, lavado com etanol, seco à vácuo, e dissolvido em água.

4.8 LIGAÇÃO DE DNA AOS VETORES

Para a ligação, foram utilizados de 5 a 20 ng do DNA vetor digerido (item 4.7), e 3 a 5x mais de DNA inserto em relação ao DNA vetor. Ao sistema foi adicionado

tampão de ligação (1x) e 0,5 a 1U de enzima T4 DNA ligase. O sistema foi incubado a 16°C por aproximadamente 24 horas.

A ligação dos produtos de PCR aos vetores pCR 2.1 e pGEM-T Easy foi feita de acordo com a indicação dos fabricantes. O fabricante do pCR 2.1 recomenda que seja feito um sistema de 20 µL, contendo 2 µL do vetor (5ng/ µL), 3U de T4 DNA ligase e 1 µL de tampão de ligase (presentes no kit), além de 10 ng de DNA inserto. O fabricante do pGEM-T easy recomenda um sistema de 10 µL, dos quais 0,2 µL de vetor (10 ng), 4U de T4 DNA ligase, 5 µL de tampão de ligase (presentes no kit), e 10 ng de DNA inserto (produto de PCR).

4.9 SEQUENCIAMENTO DE DNA

O DNA seqüenciado foi purificado como descrito no item 4.3. e tratado com 0,1 mg/µL de RNase a 37°C por 3 horas. A RNase foi inativada e extraída pelo tratamento com 0,4 volume de fenol:clorofórmio:álcool isoamílico (25:24:1). O DNA foi precipitado após a adição de 2/3 (v/v) de acetado de amônio (7,5 mol/L) e 2,5 volumes (v/v) de etanol absoluto, seco e ressuspenso em 10 µL de água estéril.

Para a reação de sequenciamento foram utilizados de 200 a 500 ng de DNA, 5 pmol do oligonucleotídeo (reverso ou universal), 3 µL de ET terminator (DYEnamic™ ET Dye Terminator Cycle Sequencing Kit, Amersham Biosciences) e 7,5 µL de água estéril para 10 µL de volume final de reação. Os parâmetros do programa, utilizado no termocilador, foram: 95°C por 1 minuto; e 35 ciclos de 94°C por 20 segundos, e 60°C por 1 minuto e 30 segundos.

Ao produto da reação de sequenciamento foram adicionados 12,5 µL de água estéril, 2 µL de acetato de amônio 7,5 mol/L, 65 µL de etanol 96%, e posteriormente, lavado com 150 µL de etanol 80%, e seco. A amostra foi dissolvida em 4 µL de Formamide Loading Dye (Applied Biosystems) e desnaturado a 96°C e submetida a eletroforese em seqüenciador automático (ABI-PRISM 377 – Applied Biosystems).

4.10 PREPARO DE CÉLULAS ELETROCOMPETENTES E TRANSFORMAÇÃO POR ELETROPORAÇÃO

4.10.1 Preparo de células eletrocompetentes de *E. coli*

Para o preparo de células eletrocompetentes, as estirpes S17.1 e DH10B de *E. coli* foram crescidas em meio SOB a 37°C sob agitação de 150rpm até atingir uma D.O.₆₀₀ de 0,6. A cultura foi então resfriada em banho de gelo por 30 minutos, lavada por 2 vezes com água ultra pura estéril gelada, intercalada por centrifugação a 5000 rpm por 5 min. Em seguida as células foram lavadas em glicerol 15%, centrifugadas a 5000 rpm por 5 min, e ressuspensas em glicerol 15%.

A eletroporação foi realizada com 35 µL de células onde foram adicionados 50 ng de DNA plasmidial, e no caso de ligações 1 µL da reação. A mistura de células e DNA foi transferida para cubas de eletroporação e submetida a um choque elétrico de 330-400 µF, 200 Ω e 4 KV/cm por cerca de 6-10 ms (eletroporador Cell-porator – Gibco BRL). Após o choque, as células foram ressuspensas em meio SOC sem antibiótico e incubadas a 37°C por 30 minutos a 1 hora. Após recuperação as células foram plaqueadas em meio LA contendo os antibióticos adequados. Colônias isoladas foram visualizadas nas placas após incubação a 37°C.

4.10.2 Preparo de células eletrocompetentes de *H. seropedicae*

Para o preparo de células eletrocompetentes de *H. seropedicae*, as estirpes foram cultivadas em meio NFb malato suplementado com 20mmol/L de NH₄Cl e, apenas 1/10 do volume da mistura de fosfatos para facilitar a remoção de sais durante as lavagens. A cultura foi crescida até atingir uma D.O.₆₀₀ de 0,6, e, então resfriada em banho de gelo por 30 minutos, lavada por 2 vezes em água ultra pura estéril gelada intercaladas por centrifugação a 5000 rpm por 5 min. Em seguida as células foram lavadas 1X em glicerol 15%, centrifugadas a 5000 rpm, e ressuspensas em de glicerol 15%.

O procedimento de eletroporação foi realizado da mesma maneira descrita no item 4.10.1. Após eletroporadas, as células foram ressuspensas em meio NFb malato sem antibióticos, para a recuperação. Após 3 a 4 horas de recuperação, a cultura foi plaqueada em meio NFb malato sólido contendo os antibióticos seletivos. Colônias foram visualizadas nas placas após incubação a 30°C por 48 horas.

4.11 CONJUGAÇÃO BACTERIANA

A estirpe S17.1 de *E. coli* (doadora) e as estirpes SmR1, DR2 e AMM1 de *H. seropedicae* (receptoras) foram cultivadas em meio LB a 37°C por 3 horas (*E. coli*), e em meio NFb malato a 30°C por 6 horas (*H. seropedicae*), ambos sem antibiótico. Após o tempo de crescimento, 100 µL das culturas de S17.1 foram misturadas à 1000 µL das culturas de *H. seropedicae*, incubadas a 30°C por 30 minutos, e centrifugadas por 30 segundos à 13.400 rpm. O precipitado de células foi ressuspensão e transferida como uma gota para meio sólido, contendo uma relação de 3:1 entre o meio NFb malato sólido e LA. As placas foram incubadas a 30°C por 24 horas.

Após este período de incubação, a biomassa de células foi coletada e ressuspensa em 1 mL de meio NFb malato líquido, 100 µL desta suspensão foram plaqueados em meio NFb malato sólido, contendo os antibióticos adequados. Estas placas foram incubadas a 30°C por 48 horas.

4.12 ARMAZENAMENTO DAS ESTIRPES BACTERIANAS

As estirpes de *E. coli* foram armazenadas em glicerol 50% a -20°C, enquanto que as estirpes de *H. seropedicae* foram armazenadas em meio NFb malato sólido e semi-sólido, e repicadas periodicamente.

4.13 OBTENÇÃO DA ESTIRPE MUTANTE NO GENE *fdeA* DE *H. seropedicae*

A mutagênese de *fdeA* foi realizada através de inativação cromossomal por recombinação com uma cópia truncada do gene, presente em plasmídeo suicida. A cópia truncada foi obtida por PCR, amplificando-se um fragmento de 512 pb da *orf1*. Os oligonucleotídeos *orf1for* e *orf1rev* (5' TTCTTCCAACGGTCTCTA 3' (*orf1for*) $T_m = 52^\circ\text{C}$, e 5' CGGTTACGCCAAACTTCC 3' (*orf1rev*) $T_m = 56^\circ\text{C}$) foram utilizados para a amplificação, utilizando-se os seguintes parâmetros: desnaturação inicial a 95°C por 5 minutos; seguido por 30 ciclos de 94°C por 45 segundos, 55°C por 1 minuto, e 72°C por 1 minuto; seguidos de uma última extensão a 72°C por 2 minutos.

O produto de PCR obtido foi ligado ao vetor comercial pCR 2.1 (item 4.8) e a mistura de ligação inserida por transformação em células DH10B eletrocompetentes (item 4.10.1).

Após a confirmação da clonagem através de restrição, o plasmídeo, então denominado pORF1, foi transformado em células eletrocompetentes da estirpe SmR1 (item 4.10.2), e a seleção feita com a adição do antibiótico canamicina ($500\ \mu\text{g}/\text{mL}$) ao meio NFb malato sólido. Algumas colônias resistentes à canamicina foram submetidas à PCR para a confirmação da inserção e de sua orientação no genoma. O DNA genômico de algumas estirpes mutantes foi extraído, digerido e submetido à hibridização para a confirmação da mutação.

4.14 HIBRIDIZAÇÃO DO DNA GENÔMICO DAS ESTIRPES MUTANTES PARA CONFIRMAÇÃO DA MUTAÇÃO DE *H. seropedicae*

4.14.1 Transferência do DNA

O DNA cromossomal da estirpe SmR1 e dos prováveis mutantes foi extraído, digerido com as enzimas adequadas, e os fragmentos separados por eletroforese em gel de agarose 0,7%. Antes da transferência o gel foi tratado com 4 soluções: solução de depurinação (HCl $125\ \text{mmol}/\text{L}$) por 7 minutos; solução de desnaturação (NaOH $0,4\ \text{mol}/\text{L}$, NaCl $1,5\ \text{mol}/\text{L}$) por 30 minutos; solução de neutralização (Tris-HCl pH 7,5 $0,7$

mol/L, NaCl 1,5 mol/L) por 20 minutos; e com solução de transferência (SSC 20X) por 10 minutos. O gel foi lavado com água destilada entre os tratamentos descritos acima. A transferência foi feita por capilaridade, através do sistema de Southern Blotting (SAMBROOK *et. al.*, 1989), para uma membrana de náilon (Hibond-N⁺, Amersham Biosciences). Após a transferência a membrana foi seca em câmara de fluxo laminar e o DNA fixado por luz ultravioleta (312 nm) em transiluminador por 3 minutos.

4.14.2 Preparo da sonda

O DNA utilizado como sonda foi um fragmento de 512 pb do gene *orf1*, obtido por PCR utilizando como molde o plasmídeo pORF1. Cerca de 50 ng desse DNA foi fervido por 5 minutos, resfriado em gelo, e marcado com α [³²P]dCTP (GE Healthcare). Os demais componentes do sistema foram: tampão OLB (1X), 1 μ L de BSA (10 mg/mL), 1 μ L de Klenow (1-2 unidades), água em q.s.p. para 25 μ L, e 2 μ L de α [³²P]dCTP. O sistema foi incubado por 24 horas a temperatura ambiente.

4.14.3 Hibridização

A membrana foi pré-hibridizada em tampão de hibridização (Na₂HPO₄ 0,5 mmol/L pH 8,0; SDS 2%) por 1 hora a 65°C. Em seguida foi adicionado 20 μ g/mL de DNA de timo desnaturado por fervura e o sistema mantido nas mesmas condições por mais 3 horas. À sonda marcada foram adicionados 100 μ L de água e a mistura foi fervida por 5 minutos, sendo então resfriada em gelo. Esta foi transferida para o sistema contendo a membrana pré-hibridizada e incubada por mais 24 horas nas mesmas condições. Após a incubação, a membrana foi lavada 3 vezes em solução de alta stringência (SSC 0,1X; SDS 0,1%). As lavagens foram feitas por 15 minutos cada, a 65°C sob agitação.

A membrana foi seca na capela e colocada em contato com placa sensível à radiação (Molecular Dynamics – GE Healthcare) e as imagens foram capturadas em PhosphorImager Storm (Molecular Dynamics – GE Healthcare).

4.15 SEQUENCIAMENTO DO GENE *fdeF* DO OPERON A MONTANTE DO GENE *nodD* DE *H. seropedicae*

Para o sequenciamento completo de *fdeF* foi construído um par de oligonucleotídeos específicos para esta região, bem como foram encontrados plasmídeos da biblioteca de *H. seropedicae* que contém esta região.

A seqüência dos oligonucleotídeos é: 5' CTGTTCTATCTCTTTGCC 3' (gapfor) e 5' CGTGCTGTTGGTAGGTGA 3'(gaprev) e permitiu a amplificação da região não seqüenciada. O produto de PCR foi ligado em vetor pGEM-T easy, como descrito no item 4.8, e transformado na estirpe DH10B eletrocompetente. A confirmação dos possíveis clones foi feita através da digestão com as enzimas de restrição *PstI* e *SphI*. Após esta confirmação, foi realizada a reação de sequenciamento (item 4.9), e para verificar se a região foi completamente seqüenciada foi utilizado o programa PHRED-PHRAP-CONSED.

A amplificação utilizando o par de oligonucleotídeos gapfor e gaprev seguiu os seguintes parâmetros: desnaturação inicial a 95°C por 5 minutos; 30 ciclos de 94°C por 45 segundos, 55°C por 1 minuto, e 72°C por 1 minuto; e uma última extensão a 72°C por 2 minutos.

Os produtos de PCR, de ambos os pares de oligonucleotídeos, foram analisados em gel de agarose 1%, como descrito no item 4.5, antes da ligação aos vetores.

4.16 CARACTERIZAÇÃO DO FENÓTIPO DO MUTANTE AMM1 DE *H. seropedicae*

A caracterização do fenótipo do mutante foi feita testando a viabilidade das estirpes SmR1, DR2 e AMM1 frente a diversos compostos fenólicos, uma vez que o produto do gene *fdeA*, assim como o restante do operon, parece estar envolvido na via de degradação destes compostos. A escolha dos compostos foi feita com base na literatura e na análise da provável função dos genes do operon em estudo.

4.16.1 Crescimento em meio NFbHPN e NFbHPN-malato sólido variando as fontes de carbono

O crescimento e viabilidade das estirpes SmR1, DR2 e AMM1 foi testado frente a diversos compostos fenólicos, em meio NFbHPN e NFbHPN-malato sólido. As estirpes SmR1, DR2 e AMM1 foram crescidas em meio NFbHPN-malato por 18 horas, quando então a DO_{600nm} das culturas foi determinada e ajustada para 1,7. Diluições 10^{-5} , 10^{-6} , 10^{-7} foram plaqueadas em triplicatas de 5 μ L, em meio NFbHPN-malato e NFbHPN contendo as diferentes fontes de carbono. As placas foram incubadas por 48 horas a 30°C, e as UFC foram contadas. Os compostos e as concentrações utilizadas estão descritos na tabela 3.

TABELA 3 – COMPOSTOS UTILIZADOS NO SCREENING EM MEIO SÓLIDO

Compostos	Concentração final (mmol/L)
Malato	37 mmol/L
Salicilato de sódio	0,12
	0,31
	0,62
	1,2
	2,0
	3,1
	5,0
	6,2
Ácido benzóico	0,8
	5,0
Resorcinol	5,0
Naringenina	1,0
	2,0
	3,0
	4,0
	5,0
Quercetina	1,0
	2,0
	3,0
Quercetina	4,0
	5,0
Tirosina	1,0
Triptofano	1,0
Fenilalanina	1,0
Etanol	500 μ L (equivalente ao volume máximo utilizado dos compostos)

4.16.2 Crescimento em meio NFbHPN e NFbHPN-malato líquido variando as fontes de carbono

Para a observação de crescimento em meio líquido, foi utilizado a metodologia descrita em BOCHNER *et al.*, 2008. Esta metodologia é baseada na oxidação do NADH a NAD^+ , com mudança na coloração do meio de incolor para púrpura na devido a presença do composto tetrazólio violeta. Se o microrganismo não está utilizando o composto testado como fonte de carbono, não há crescimento, logo não há formação de NADH e o meio permanece incolor. Neste trabalho foi utilizado o cloreto de 2,3,5-trifenil tetrazólio (figura 10), que é reduzido, passando de incolor à púrpura, da mesma maneira que o tetrazólio violeta.

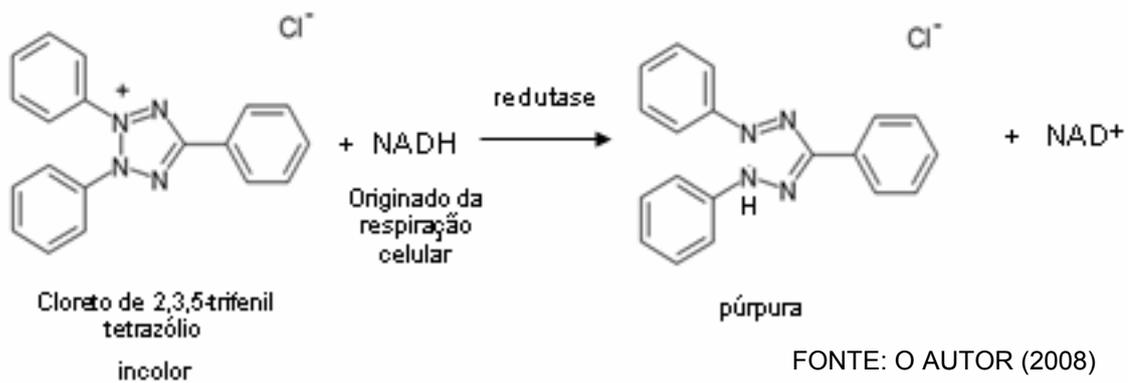


FIGURA 10 – REAÇÃO DE OXIDAÇÃO QUE OCORRE NO INTERIOR DA CÉLULA NA PRESENÇA DO COMPOSTO CLORETO DE 2,3,5-TRIFENIL TETRAZÓLIO.

As fontes de carbono utilizadas neste ensaio estão descritas na tabela 4.

TABELA 4 – CONDIÇÕES UTILIZADAS NO EXPERIMENTO DE DETERMINAÇÃO DO FENÓTIPO DA ESTIRPE AMM1 EM MEIO LÍQUIDO; pH 6,9

Condição	Concentração final da fonte de carbono (mmol/L)
NFbHPN	-
NFbHPN-malato	37 mmol/L
NFbHPN glucose	50 mmol/L
NFbHPN sacarose	50 mmol/L
NFbHPN trealose	50 mmol/L
NFbHPN manose	50 mmol/L
NFbHPN ramnose	50 mmol/L
NFbHPN ácido benzóico	1 mmol/L
NFbHPN salicilato de sódio	1 mmol/L

NFbHPN Tirosina (Tyr como fonte de carbono)	1 mmol/L
NFbHP malato+tirosina (Tyr como fonte de nitrogênio)	37mM malato; 1mM Tyr
NFbHPN fenilalanina (Phe como fonte de carbono)	1mM
NFbHP malato+fenilalanina (Phe como fonte de nitrogênio)	37mM malato; 1mM Phe
NFbHPN fenol	1mM
NFbHPN malato+fenol	37mM malato; 1mM fenol
NFbHPN naringenina	0,2mM
NFbHPN resorcinol	1mM
NFbHPN etanol	32 µL (equivalente ao volume de naringenina utilizado)

As estirpes SmR1, DR2 e AMM1 foram riscadas em meio NFbHPN malato sólido, e todas as colônias crescidas na placa após 2 dias de incubação a 30°C foram coletadas com o auxílio de um palito estéril, e ressuspensas em 5 mL de meio NFb líquido. A DO_{600nm} das culturas foi determinada, ajustadas para 1,0 em 1 mL de meio, e então 10µL foram inoculados em cada poço do bloco, que foi incubado a 30°C por 2 dias, sob agitação a 120 rpm. Após este período de incubação, foi adicionado 0,1% de cloreto de 2,3,5 trifenil tetrazólio, e uma nova incubação de 2 horas foi feita para que a coloração púrpura fosse visualizada.

4.16.3 Caracterização do fenótipo da estirpe AMM1 por HPLC

Para a análise da capacidade de *H. seropedicae* em degradar naringenina utilizando o HPLC, as estirpes SmR1, DR2 e AMM1 foram cultivadas em meio NFbHPN e NFbHPN-malato contendo 2 mmol/L de naringenina. Para cada estirpe o crescimento foi feito em duplicata biológica, ou seja, foram utilizadas 2 colônias diferentes de cada estirpe. A DO_{600nm} das células do pré-inóculo, que foi feito em meio NFbHPN-malato, foram ajustadas para 1,0, em 1 mL, e as células lavadas através de uma centrifugação por 2 minutos a 13.400 rpm, e ressuspensas em 1 mL da mistura de sais do meio NFb. A partir desta suspensão foram feitos os inóculos, com uma DO_{600nm} inicial de 0,01 nas condições acima mencionadas. O controle foi feito em meio NFbHPN ambos contendo apenas 2 mmol/L de naringenina, e sem inóculo. As culturas foram incubadas a 30°C, 120 rpm por 30 horas, e foram coletadas amostras

de 1 mL no tempo (horas) 0, 6, 12, 24 e 30. As amostras foram centrifugadas por 10 minutos a 13.400 rpm e o sobrenadante filtrado, sendo então adicionado como padrão interno o flavonóide crisina, na concentração final de 10 µg/mL.

Foram injetados 5 µL de cada amostra em HPLC Shimadzu (LC10A), e a detecção realizada utilizando o detector SPA-UV/Vis, em 280 nm. A temperatura foi mantida a 60°C, e a separação realizada com um fluxo de 1,2 mL.min⁻¹ em coluna Synergy Fusion-RP (Phenomenex), de 150 X 2,0 (4 µM de tamanho de partícula). Duas fases móveis foram utilizadas. Solvente A: 1% de ácido acético em água (v/v), e solvente B: 1% de ácido acético em acetonitrila (v/v).

Foi utilizado um gradiente linear: inicial com 10% B, 10-40% B em 6 minutos, que foi mantido em 40% até os 7 minutos, retornando para 10% B em 9 minutos, e mantido por mais 5 minutos em 10% B para a estabilização.

4.17 DETERMINAÇÃO DA ATIVIDADE ESPECÍFICA DE β-GALACTOSIDASE

A determinação da atividade de β-galactosidase foi feita segundo MILLER, 1972.

As estirpes SmR1, DR2 e AMM1 de *H. seropedicae* contendo as fusões pNB8 (em AMM1 este plasmídeo não foi inserido), pSU11 e pSUI foram inoculadas em meio NFbHPN-malato e incubadas por 16 horas. Após este período, as células foram centrifugadas, ressuspensas nas condições a serem testadas (a partir de uma DO_{600nm} de 0,2), e incubadas a 30°C por 6 ou 18 horas (conforme indicado nas legendas das figuras em resultados) a 120 rpm, quando então foi realizado o ensaio de β-galactosidase.

O sistema de reação continha 100 µl de cultura, 900 µl de tampão Z (SDS 0,27% e β-mercaptoetanol 0,39% em Na₂HPO₄.7H₂O 60 mmol/L, NaH₂PO₄.H₂O 40 mmol/L, KCl 10 mmol/L, MgSO₄.7H₂O 1 mmol/L, pH 7,0), e 25 µl de clorofórmio. Esta mistura foi agitada por 10 segundos, e incubada a 30°C por 5 minutos. Foram adicionados 200 µl do o-nitrofenil-β-D-galactopiranosídeo – ONPG (4 mg/mL) – em cada tubo de ensaio, incubados por 30 minutos, ou até que fosse observada a coloração amarela. A reação foi interrompida após adição de 500 µl de Na₂CO₃ 1M. Os tubos de reação

foram centrifugados por 2 minutos a 13.400 rpm (centrífuga mini-spin Eppendorf) para que o clorofórmio e as células lisadas ficassem no fundo do tubo e não interferissem na leitura da absorbância. Após centrifugação, 200 µL da reação foram transferidos para placa de 96 poços e a absorbância foi determinada a 415 nm e 550 nm em leitor de microplacas (ELX-800 BioTek Instruments, Inc.).

A atividade específica de β-galactosidase (WASSEM *et al.*, 2002; SCHWAB *et al.*, 2007) foi obtida em nanomoles de ONP formado por minuto por miligrama de proteína total (nmol ONP/min/mg proteína).

Para a determinação da concentração de proteínas o método utilizado foi o descrito por BRADFORD (1976), utilizando solução de albumina de soro bovino (BSA) como padrão. A lise das células foi feita com a adição de 1 volume de NaOH 0,2 mol/L e incubação por no mínimo 1 hora, a temperatura ambiente.

Os flavonóides utilizados como possíveis indutores foram: naringenina, quercetina, apigenina, crisina, luteolina, kaempferol, flavona, catequina, genisteína e daidzeína. Alguns compostos intermediários da via de degradação de compostos aromáticos, como ácido benzóico, ácido salicílico, salicilato de sódio e resorcinol, e os aminoácidos aromáticos tirosina, fenilalanina e triptofano também foram testados como possíveis indutores. Porém, só houve resposta com os flavonóides e, assim sendo, neste trabalho serão mostrados apenas os resultados referentes a estes compostos.

4.18 SOLUBILIZAÇÃO DOS COMPOSTOS

Todos os compostos e solventes utilizados neste trabalho estão descritos na tabela 5:

TABELA 5 – COMPOSTOS UTILIZADOS NO DECORRER DESTE TRABALHO

Composto	Solvente	Concentração do estoque
Naringenina	Etanol absoluto	125 mmol/L
Quercetina	DMSO	200 mmol/L
Resorcinol	Etanol absoluto	500 mmol/L
Fenol	Água	500 mmol/L
Salicilato de sódio	Água	500 mmol/L

Ácido salicílico	Etanol absoluto	500 mmol/L
Acido benzóico	Etanol absoluto	500 mmol/L
Tirosina	Água+HCl concentrado (37%)	125 mmol/L
Fenilalanina	Água	125 mmol/L
Triptofano	Água+HCl concentrado (37%)	125 mmol/L
Malato	Água	20%
Glucose	Água	20%
Ramnose	Água	20%
Manose	Água	20%
Trealose	Água	500 mmol/L
Sacarose	Água	10%

4.20 CURVA DE CRESCIMENTO DAS ESTIRPES SMR1, DR2 E AMM1 NA PRESENÇA E NA AUSÊNCIA DE NARINGENINA

Para a realização das curvas de crescimento as estirpes SmR1, DR2 e AMM1 foram cultivadas em meio NFbHPN-malato por aproximadamente 12 horas. As células foram então centrifugadas e lavadas uma vez em meio NFb, e a DO_{600nm} ajustada para 0,1 em 50 mL de meio líquido, em erlenmeyers de 250 mL respeitando a proporção de 1:5 (meio:ar). As condições testadas foram: controle (NFbHPN-malato), controle etanol (NFbHPN-malato + etanol), e naringenina (NFbHPN-malato + 0,2 mmol/L de naringenina), sendo que foi feito apenas uma repetição para cada condição. O volume de etanol utilizado foi o mesmo volume de naringenina utilizado. A DO_{600nm} foi determinada no tempo zero e a cada hora, retirando-se 1 mL de cultura. O espectrofotômetro utilizado para as leituras de absorbância foi o Beckman.

Para o cálculo do tempo de geração foi utilizado o valor de inclinação da reta, que é o coeficiente angular da equação da reta, sendo que o gráfico foi construído a partir do \log_{10} das DO_{600nm} com os pontos que mais se ajustaram à reta. A taxa de crescimento (μ) é igual a constante $2,303 \cdot \text{inclinação}$, e o tempo de geração foi calculado a partir da seguinte fórmula: tempo de geração (g) = \log_2 / μ .

4.21 ANÁLISE DE BIOINFORMÁTICA

As ferramentas de bioinformática utilizadas neste trabalho foram: FramePlot (ISHIKAWA & HOTTA, 1999), Blast (ALTSCHUL, 1997), String (von MERING, 2007), e SignalP (BENDTSEN, 2004).

O programa FramePlot é utilizado para identificar ORFs (Open Read Frames) em genomas de procariotos ricos em G+C, o que resulta em uma alta distribuição dos nucleotídeos G e C na 3^a letra dos códons, permitindo uma predição do códon de início de cada região codificadora (ORF). A seqüência de nucleotídeos da região a montante do gene *nodD*, presente no contig 186, bem como a seqüência de nucleotídeos do contig 232 foram submetidos a este programa para identificar quantas ORFs fazem parte do operon. Após esta etapa, as seqüências foram submetidas ao blastx que traduz a seqüência de nucleotídeos em aminoácidos e compara com o banco de dados de proteínas já depositadas no NCBI (National Center for Biotechnology Information).

O String é um programa que, a partir da seqüência de aminoácidos submetida, faz a comparação com as proteínas encontradas em microrganismos que já tenham o genoma seqüenciado, cujas seqüências estejam depositadas no banco de dados. Após esta comparação ele fornece os microrganismos onde a proteína é encontrada assim como as proteínas vizinhas.

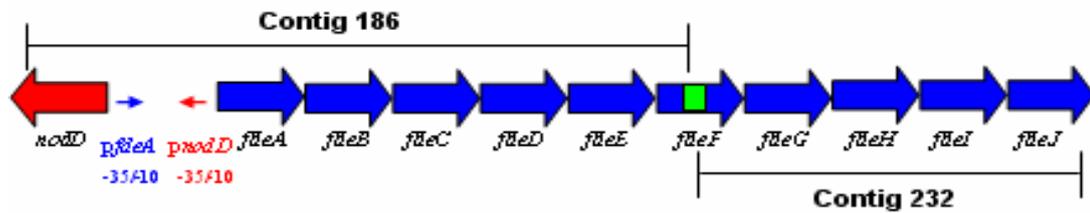
Para procurar por possíveis peptídeos sinais foi utilizado o programa SignalP, que prediz um sítio de clivagem na região N-terminal das proteínas.

5 RESULTADOS E DISCUSSÃO

5.1 Estudo de bioinformática da região a montante do gene *nodD* de *H. seropedicae*

Em 2006, Marin e colaboradores utilizaram o plasmídeo pVAN1, que contém o gene *nodD* de *H. seropedicae* SmR1 clonado no vetor pDK7 para avaliar a expressão da proteína HsNodD em *E. coli*. Também foi utilizada a fusão pSU11 (KARP, 2004), que contém a região promotora do gene *nodD* de *H. seropedicae* fusionado ao gene *lacZ*. Nesta mesma região intergênica se encontra o promotor do gene *fdeA* (1ª *orf* do operon em estudo, como será descrito logo abaixo) porém em orientação oposta em relação ao promotor de *nodD*, que também foi clonado fusionado ao gene *lacZ*, originando o plasmídeo pSUI (MARIN, 2006). Assim, temos a mesma região intergênica clonada em ambas as orientações para estudar a possível regulação da proteína HsNodD na transcrição do gene *nodD* assim como do operon encontrado a montante deste gene. Para isto foi utilizada a estirpe MC1061 (*lac*⁻) de *Escherichia coli*, sendo observado que em *E. coli* a proteína HsNodD reprime a sua própria transcrição assim como a do gene *orf1*, porém na presença de flavonóides a repressão, em relação ao gene *orf1*, é revertida, ocorrendo a indução da expressão do mesmo. Em relação à expressão de *nodD*, a presença de flavonóides não alterou a repressão de HsNodD sobre a sua própria expressão, observada na ausência de flavonóides (controle).

Através de análises de bioinformática, foi possível observar que a *orf1* é a primeira *orf* de um operon que contém 10 *orfs* (figura 12), uma vez que a única região promotora encontrada está na região intergênica entre os genes *nodD* e *orf1*, e que após as 10 *orfs* há uma grande região, aparentemente, não codificadora, observado no programa FramePlot. Uma vez que a principal hipótese é que este operon seja responsável pela degradação de flavonóides, ele será chamado de operon *fde*, e os genes nomeados *fdeA*, *fdeB*, *fdeC*, *fdeD*, *fdeE*, *fdeF*, *fdeG*, *fdeH*, *fdeI* e *fdeJ*. Durante a análise de bioinformática foi encontrada uma região não seqüenciada dentro da *fdeF* (quadrado verde – figura 11).



FONTE: O AUTOR 2008

FIGURA 11 – ESQUEMA DA DISPOSIÇÃO DO GENE *nodD* E DO OPERON *fde* DE *H. seropedicae*. A região em azul mostra os genes encontrados no operon, sendo que a região de *fdeA* até a região não seqüenciada de *fdeF* (em verde) encontra-se no contig 186, e a região do final da *fdeF* até a *fdeJ* encontra-se no operon 232 (anotação de 18/08/2003). A seta azul indica a posição do promotor dependente de σ^{70} do operon, e a seta vermelha indica o promotor σ^{70} de *nodD*.

Uma vez que o contig 186 (Programa GENOPAR – anotação de 18/08/2003) termina na região inicial do gene *fdeF*, foi necessário encontrar o restante deste gene dentre os demais contigs montados durante o sequenciamento do genoma de *H. seropedicae*. Para isto, foi realizada uma busca por plasmídeos na biblioteca genômica cuja seqüência estivesse na ponta do contig 186. Assim, foi encontrado o plasmídeo HS23-RN-00-000-094-C11 que possui o gene *fdeF* e anela nos contigs 186 e 232, que contém o restante do operon. Este gene foi completamente seqüenciado e a estratégia utilizada está descrita no item 5.1.1. O gene *fdeA* é o primeiro gene de um operon que contém 10 *orfs*, cujos produtos podem estar envolvidos na metabolização de compostos fenólicos. Antes do códon de início de transcrição de *fdeA*, foi identificado um sítio de ligação para ribossomo, que fica 6 pb a montante do início de tradução. Além disso, foi encontrada uma região promotora -35/-10, presente na região intergênica entre *nodD* e *fdeA*. Esta foi a única região que apresentou promotor. Entre o final e começo das demais *orfs* não foram identificadas seqüências promotoras, o que indica que as 10 *orfs* em questão formam um operon, sendo co-transcritas (ANEXO – operon todo).

Para tentar encontrar a possível função do operon em estudo, as seqüências das proteínas codificadas pelos genes presentes no operon foram comparadas com a seqüência de outras proteínas encontradas no banco de dados com o objetivo de encontrar domínios estruturais conservados de proteínas encontradas em outros organismos. A tabela 6 mostra o que foi encontrado por esta análise.

TABELA 6 - OPERON A MONTANTE E DIVERGENTE AO GENE *nodD* em *H. seropedicae*

Gene	Tamanho (pb)	Hit de maior similaridade no blastx (%)	Tamanho do polipeptídeo	Domínios
<i>fdeA</i>	939 pb	77% proteína hipotética conservada de <i>Burkholderia</i> sp. H160	312 aa	COG4313 Proteína envolvida na via meta de degradação de fenol
<i>fdeB</i>	1161 pb	88% com Bll6428 de <i>Bradyrhizobium japonicum</i> USDA110	386 aa	superfamília das esterases e lípases; COG1073
<i>fdeC</i>	927 pb	81% glioxalase/ proteína de resistência a bleomicina/dioxigenase de <i>Burkholderia ambifaria</i> MC40-6	309 aa	pfam 00913 – glioxalase/proteína de resistência à bleomicina/ superfamília das dioxigenases
<i>fdeD</i>	339 pb	71% proteína Rieske (2Fe-2S) de <i>Burkholderia vietnamiensis</i> G4	113 aa	cd03467 – domínio Rieske
<i>fdeE</i>	1131 pb	75% monooxigenase ligadora de FAD de <i>Xanthobacter autotrophicus</i> Py2	376 aa	pRK05868 pRK06847 proteína hipotética
<i>fdeF</i>	1116 pb	76% superfamília MSF-1 de <i>Bradyrhizobium</i> sp. BTAi1	372 aa	Superfamília dos transportadores MFS-1
<i>fdeG</i>	813 pb	96% provável ciclase de <i>Comamonas testosteroni</i> KF-1	270 aa	pfam04199 família de prováveis ciclases
<i>fdeH</i>	552 pb	86% com o domínio cupina 2 de <i>Burkholderia ambifaria</i> MC40-6	198 aa	pfam07883 representa o domínio em barril conservado da superfamília cupina
<i>fdeI</i>	985 pb	78% com fumaril-acetoacetato hidrolase de <i>Burkholderia ambifaria</i> MC40-6.	344 aa	pfam01557 família fumarilacetoacetato hidrolase COG0179
<i>fdeJ</i>	963 pb	65% proteína hipotética PP_3206 de <i>Pseudomonas putida</i> KT2440	328 aa	MhpD – via do catecol pfam01063 família 3-beta hidroxiesteroide dehidrogenase/isomerase COG0451 WcaG epimerase

A proteína FdeA encontra-se agrupada ao COG4313, que é um grupo de proteínas ortólogas que atuam na via meta de degradação de fenol. Essa proteína também é encontrada em outros microrganismos como *Bradyrhizobium japonicum* USDA110 (número de acesso: NC_004463.1), *Yersinia frederiksenii* ATCC 33641 (número de acesso: NZ_AALE01000026.1), *Magnetospirillum magnetotacticum* MS-1 (número de acesso: NZ_AAAP01001463.1) e *Ralstonia solanacearum* MolK2 (número de acesso: NW_002196568.1), entre outros, que possuem seqüências depositadas no banco NCBI. Porém ainda não há trabalhos que mostrem a real função desta proteína (NCBI – Conserved Domains).

A FdeB pertence à superfamília das esterases/lipases e ao COG1073. Esterases e lipases agem em ésteres carboxílicos, cujo sítio catalítico contém 3 resíduos de aminoácidos (tríade catalítica): serina, glutamato ou aspartato, e histidina. Estes resíduos são responsáveis pelo ataque nucleofílico ao carbono da carbonila da ligação éster. O COG1073 contém proteínas da superfamília de hidrolases α/β , de função ainda desconhecida (NCBI – Conserved Domains).

A FdeC contém um domínio conservado da família das glioxalases, sendo que este domínio também é encontrado em proteínas de resistência à bleomicina e em dioxigenases, que são enzimas importantes na clivagem do anel de compostos aromáticos. Há dois grupos de dioxigenases que clivam anéis aromáticos, sendo classificadas de acordo com o modo de cisão do anel. Intradiol dioxigenases são enzimas que contêm Fe(III) para clivar o anel aromático entre dois carbonos hidroxilados adjacentes (clivagem orto) (OHLENDORF *et al.*; 1988), enquanto extradiol dioxigenases usam Fe(II) para clivar o anel entre um carbono hidroxilado adjacente a um não-hidroxilado (clivagem meta) (ASTURIAS *et al.*, 1994). Os íons de ferro férrico (Fe^{+3}) e ferroso (Fe^{+2}) não são ligados a grupamentos heme.

A FdeD apresenta um domínio Rieske, contendo um cluster 2Fe-2S comumente encontrado em proteínas Rieske oxigenases que não contém grupamentos heme, como as naftaleno e bifenil dioxigenases. O cluster 2Fe-2S está envolvido na transferência de elétrons e está ligado a dois resíduos de cisteína e dois resíduos de histidina presentes em uma seqüência conservada, denominada motivo Rieske (NCBI – Conserved Domains).

A FdeE possui um domínio pertencente à família das piridinas nucleotídeo-dissulfito oxidorreduções, que inclui as oxidorreduções de classe I e II, NADH oxidases e NADH peroxidases. Através de estudos de fingerprinting foram encontrados sítios de ligação pra ADP, FAD e NAD ou NADP. A região N-terminal possui um sítio de ligação pra ADP que está envolvido com a ligação ao FAD; na região central encontra-se um sítio de ligação para NAD ou NADP; e na região C-terminal há um sítio de ligação ao grupamento flavina do FAD (EGGINK *et al.*, 1990). As flavoproteínas FAD, pertencentes a esta família de oxidorreduções (glutaciona redutase, tripanotiona redutase, lipoamida desidrogenase, mercúrio redutase, e

tioredoxinas redutase, entre outras) compartilham similaridade de seqüência com outras flavoproteínas oxidoredutases, em particular com as redutases ferredoxina NAD⁺ envolvidas no metabolismo oxidativo de vários hidrocarbonetos (ferredoxina-NAD⁺ redutase encontrada nas enzimas benzeno 1,2-dioxigenase, tolueno 1,2-dioxigenase, clorobenzeno dioxigenase, bifenil dioxigenase) (EMBL-EBI IPR001327).

A FdeF possui um domínio MFS-1 (Major Facilitator Superfamily) também conhecida como família uniporte/simporte/antiporte, e as proteínas que possuem tal domínio são capazes de transportar pequenos solutos em resposta a um gradiente quimiosmótico. Há várias famílias de MFS descritas na literatura, sendo uma delas relevante para este trabalho que é a família de simporte de ácidos aromáticos:H⁺. Alguns dos substratos deste transporte são: muconato, benzoato; 4-hidroxibenzoato; 2,4-diclorofenoxiacetate, protocatecuato, e 3-hidroxipropionato (PAO *et al.*, 1998).

A FdeG possui um domínio da superfamília das ciclases, sendo encontradas principalmente em proteínas envolvidas na biossíntese de antibióticos, catalizando a formação de anéis aromáticos (NCBI – Conserved Domains).

A FdeH possui um domínio conservado pertencente à superfamília das cupinas 2, que tem como característica a presença da estrutura secundária do tipo barril β. Uma das enzimas que contém esse tipo de estrutura é a quercetina 2,3-dioxigenase (também conhecida como quercetinase ou flavonol 2,4-dioxigenase; EC 1.13.11.24), que é expressa por *Aspergillus flavus* após o crescimento deste fungo em meio contendo rutina (forma glicosilada da quercetina). Esta enzima, oxidativamente, cliva o anel heterocíclico da quercetina (anel C) liberando monóxido de carbono e o ácido carboxílico 2-protocatecuilfloroglucinol, e contém um íon cobre (Cu⁺) em seu sítio ativo (OKA & SIMPSON, 1971).

A Fdel possui o domínio pertencente à família fumarilacetoacetato hidrolase (FAA), que também é encontrado na enzima bifuncional HHDD (ácido 2-hidroxihept-2,4-diene-1,7-diólico) isomerase/OPET (ácido 5-oxo-pent-3-eno-1,2,5-tricarboxílico) decarboxilase. A FAA é a última enzima a atuar no catabolismo da tirosina e da fenilalanina, clivando a molécula de fumarilacetoacetato em fumarato e acetoacetato, que entram no ciclo dos ácidos tricarboxílicos. Um exemplo de enzima bifuncional contendo este domínio é a HpcE, responsável pela degradação do 4-

hidroxifenilacetato, que é um produto do catabolismo da tirosina e da fenilalanina (IPR002529 Fumarylacetoacetase).

A FdeJ contém um domínio S-adenosil-L-homocisteína hidrolase, que catalisa a hidrólise do S-adenosil-L-homocisteína (AdoHyc) para formar adenosina (Ado) e homocisteína (Hcy), reação que faz parte da via de degradação da metionina. Em *Bradyrhizobium japonicum* USDA110 esta proteína contém estes mesmos domínios e codifica para UDP-glucose 4-epimerase. Esta enzima interconverte UDP-glucose e UDP-galactose que são precursores das moléculas de glucose e galactose, encontradas em exopolissacarídeos (IPR005886 UDP-glucose 4-epimerase).

Devido à presença de proteínas cuja estrutura sugere um envolvimento na degradação de compostos aromáticos, este operon parece estar envolvido no metabolismo de tais compostos, como benzoato, salicilato, aminoácidos aromáticos, resorcinol, fenol, e flavonóides.

Utilizando o programa String – Protein and their interactions (von MERING *et al.*, 2007), foi possível observar que em *Bradyrhizobium japonicum* USDA110 há um operon semelhante a este de *H. seropedicae* SmR1 (figura 12). A partir de uma seqüência polipeptídica já depositada no banco de dados, o programa String compara a seqüência submetida com a seqüência depositada, procurando pela sua vizinhança em outros organismos que possuem genomas sequenciados.

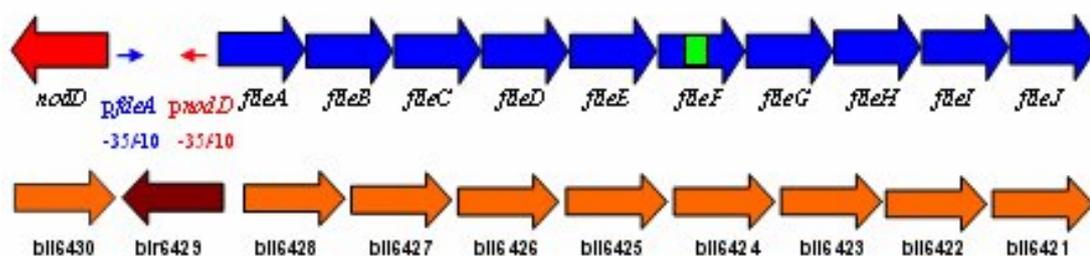


FIGURA 12 – COMPARAÇÃO ENTRE A REGIÃO DO GENE *nodD* DE *H. seropedicae* E A REGIÃO SEMELHANTE ENCONTRADA EM *B. japonicum*.

Os 2 microrganismos possuem esta região genômica bastante semelhante, porém ainda não há indícios da função deste operon, uma vez que ele está descrito estruturalmente, mas não funcionalmente. Várias α -Proteobactérias, além de *B. japonicum* USDA110, como *R. leguminosarum* bv. *viciae* 3841, *Sinorhizobium meliloti*

1021, *Agrobacterium tumefaciens* str. C58, *Gluconobacter oxydans* 621H, *Magnetospirillum magneticum* AMB-1, algumas β -Proteobactérias como *Burkholderia* sp., *Azoarcus* sp., *Bordetella* sp., e *Ralstonia* sp., e algumas γ -Proteobactérias, como *Pseudomonas* sp., e *Marinomonas* sp. também possuem parte deste operon, lembrando que o único microrganismo que contém este operon inteiro é o *B. japonicum* USDA110. Em rizóbios a ciclase codificada pela *fdeG* de *H. seropedicae* SmR1 é encontrada fora do operon ou depois de *bll6421*.

A tabela 7 mostra a equivalência entre os genes de *H. seropedicae* e *B. japonicum*. Contudo, a seqüência da *fdeG* de *H. seropedicae* não foi encontrada nesta região do genoma de *B. japonicum*.

TABELA 7 – COMPARAÇÃO ENTRE *H. seropedicae* SmR1 E *B. japonicum* USDA110

<i>B. japonicum</i>	<i>H. seropedicae</i>
<i>blr6429</i>	<i>nodD</i>
<i>bll6430</i>	<i>fdeA</i>
<i>bll6428</i>	<i>fdeB</i>
<i>bll6427</i>	<i>fdeC</i>
<i>bll6426</i>	<i>fdeD</i>
<i>bll6425</i>	<i>fdeF</i>
<i>bll6424</i>	<i>fdeG</i>
<i>bll6423</i>	<i>fdeH</i>
<i>bll6422</i>	<i>fdeI</i>
<i>bll6421</i>	<i>fdeJ</i>

Além de programas que comparam seqüências, também foi utilizado o programa SignalP 3.0 (BENDTSEN *et al.*, 2004), que prediz a presença e a localização de um peptídeo sinal, através de um sítio de clivagem na cadeia polipeptídica. Assim, foi possível identificar que a FdeA apresenta um peptídeo sinal que se estende do 1º ao 28º resíduo da região N-terminal. Os 28 primeiros aminoácidos são: VKRGFVPACRLATCALLTMAAGPAAYA (figura 13). A presença deste peptídeo sinal sugere que esta proteína seja transportada. Nenhuma outra ORF deste operon apresentou score alto para a presença de peptídeo sinal.

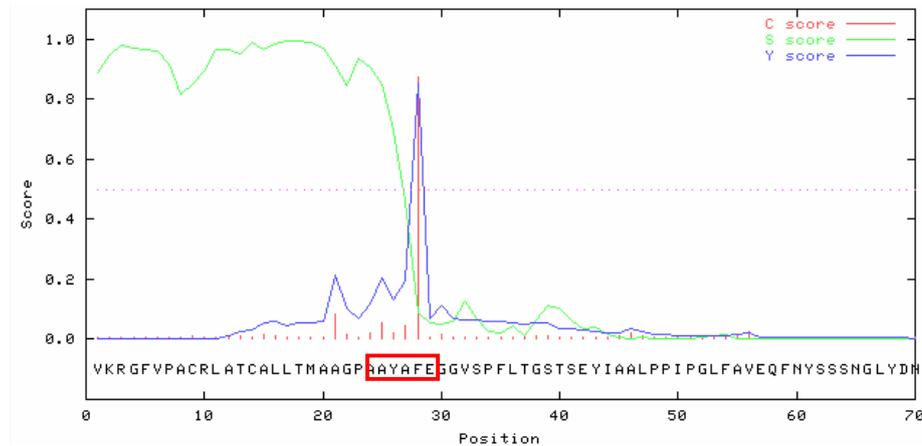


FIGURA 13 – PEPTÍDEO SINAL ENCONTRADO NA PROTEÍNA FdeA. Gráfico obtido após submeter a seqüência de aminoácidos da proteína FdeA ao programa SignalP 3.0, mostrando a predição para o sítio de clivagem AYA-FE (região destacada em vermelho).

5.1.1 Sequenciamento da *fdeF*

O algoritmo blastx (ALTSCHUL *et al.*, 1997) foi utilizado para determinar o número aproximado de nucleotídeos não seqüenciados. Para tal, a seqüência de nucleotídeos da *fdeF*, a jusante (*contig* 186 da anotação de 23/08/2003 do genoma de *H. seropedicae*) e a montante (*contig* 232 da anotação de 23/08/2003 do genoma de *H. seropedicae*) da região não seqüenciada, foram submetidos ao blastx, separadamente. O alinhamento, utilizando o blastx, da região do *contig* 186 mostrou que a seqüência de aminoácidos termina no aminoácido de número 195, e quando a região inicial do *contig* 232 foi submetida à mesma ferramenta mostrou que a seqüência de aminoácidos começa no aminoácido 271, ou seja, há 76 aminoácidos faltando nessa região. Logo, são 228 nucleotídeos não seqüenciados. No caso do oligonucleotídeo que anela no *contig* 232, não foi possível planejá-lo muito próximo da extremidade, uma vez que a seqüência desta região não apresenta alta qualidade. Assim, o fragmento esperado, obtido pela amplificação com o par de oligonucleotídeos é de 847 nucleotídeos.

O produto de PCR obtido foi clonado no vetor pGEM-T Easy, confirmado por restrição e denominado pGAP (figura 14). Os plasmídeos que contém as regiões a montante e a jusante à região não seqüenciada da *fdeF* também foram seqüenciados.

Os plasmídeos escolhidos foram: HS23-RN-00-000-094-C11, e HS25-EG-00-000-146-G01.

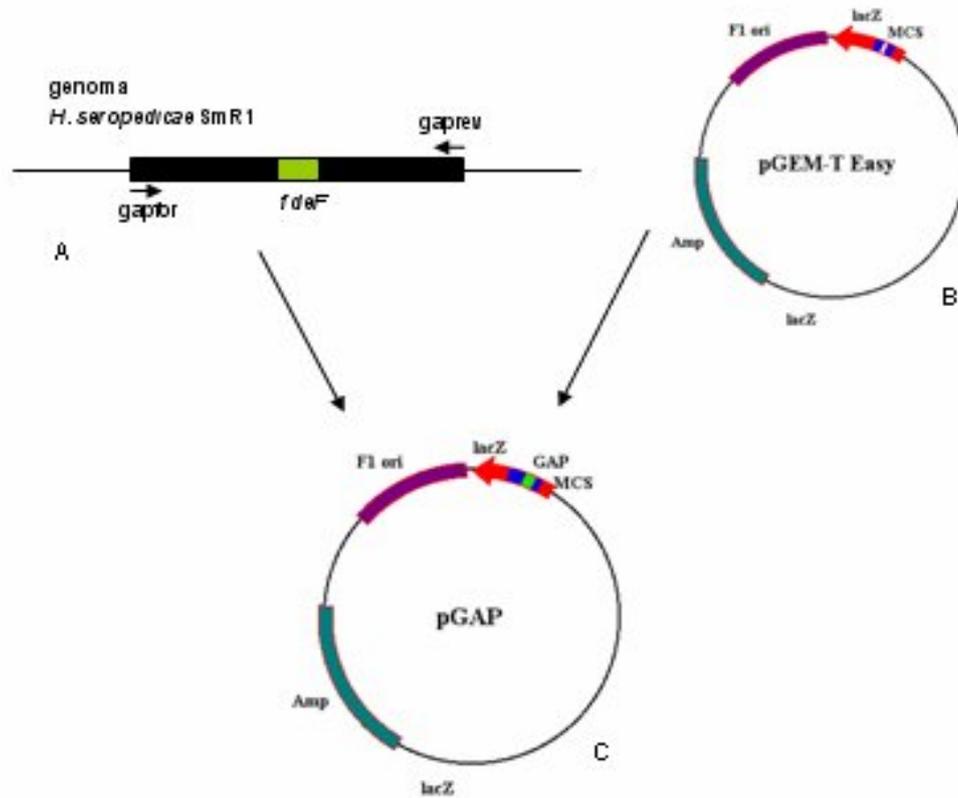


FIGURA 14 – CONSTRUÇÃO DO PLASMÍDEO pGAP. A) gene *fdeF* de *H. seropedicae*, onde se situa o GAP, e onde os oligonucleotídeos *gapfor* e *gaprev* se anelam; B) vetor pGEM-T Easy; C) plasmídeo pGAP.

As seqüências obtidas foram submetidas ao pacote de programas PHRED-PHRAP-CONSED que faz a análise quanto à qualidade das bases, montagem das seqüências e exibição das seqüências em interface linux. Após a montagem foi obtido um único *contig* para esta região, ou seja, os contigs 186 e 232, agora, fazem parte de um único contig, e o gene *fdeF* está completamente seqüenciado (figura 15).

```

ATCGCGCTGGGCGTCTTCGCCATCTGTTTCTCGGCAGTGCGCCGCGCTGGTCGCTCACTTGGTGGCA
GTACGCTGTTCTATCTTTGCCGGCGTGATGGCGGTGGGTGCCTGGCCAGCGCTTGGCTTTCCGCG
TGCCGAGACCGATGGCTGCCGCGCGTGTCCACCGTGCTCCACCGGTGCCGCGTCAGGTCTGGTACGGC
ATCGCCGGATCGCTGCATGGGGTGGTGCAGTCGATGACCTTCAGTTCCTCGAACGCGTCGGCAGTG
ATCGTGGCTTCGGCTGCAAGCCGTCACCGGCGTCATCGCCTGGGTTTCGTCAATCTCTGCCCGC
TCCATGGCGCTCTGTGAAAAGCGCTGCCGGCCCGCAACGTGCTCATGGCCGCCGCTGGTGCAG
GCGCTGCTGGTGCTGGTGATCATGTCCTCGGCGCGCTTCGCACCTATGCCGCCGCCGCGCGTCGGTATTTG
CCGCGGTGATGATCTCACGCATACTTTGCTTCGGTCTCATGGCCGGCTCGACCGAGTGGCCGC
CATGGCCCACGCCCCATGCTGATGACCGGCGCCATCGGCCCATTCTCGGCGCACGCTGGTA
CAGGGCTGGGCTATGGCAGCTGGGACCAGTGGCGTGGCGTGGCCTGACGGCATCGTTCTGTTTCT
CGCGGCTGCCGCCCCCCGCGCTTCAACCTCAGGAGCCCTGGCATGACCGCATCCTTGCAAGCACGT
CGCCGCATCGTTCGTCGACTTGTCCATCTATCTCGAAAACGACGTGCTCCGATCCGCCCCACTGG
CGCCAAGATCACTACCACACGCACCTTGCCCGAATCATGGCATGCTGCCGGCACCCG
GCCCGAAGACTATCCCGACGGCAAGCCCGCCGCTGAATGGGTCA

```

FIGURA 15 – SEQUÊNCIA DE NUCLEOTÍDEOS COMPLETA DO GENE *fdeF* (em verde). Os oligonucleotídeos utilizados para a construção do pGAP estão marcados em rosa.

5.2 OBTENÇÃO E CARACTERIZAÇÃO DA ESTIRPE MUTANTE *fdeA*⁻ DE *H. seropedicae*

5.2.1 Construção do mutante *fdeA*⁻ de *H. seropedicae*

Para a obtenção da estirpe mutante *fdeA*⁻ de *H. seropedicae* foi construído o plasmídeo chamado pORF1 (figura 16), a partir da amplificação de um fragmento central de 512pb do gene *fdeA*. Este plasmídeo foi transferido para a estirpe SmR1 através de eletroporação e colônias resistentes a canamicina foram obtidas, resultantes da integração do plasmídeo no genoma de *H. seropedicae* SmR1 por recombinação homóloga simples, e este mutante deve apresentar alteração na expressão de todo o operon, sendo esperado que a via toda deixe de funcionar na estirpe mutante.

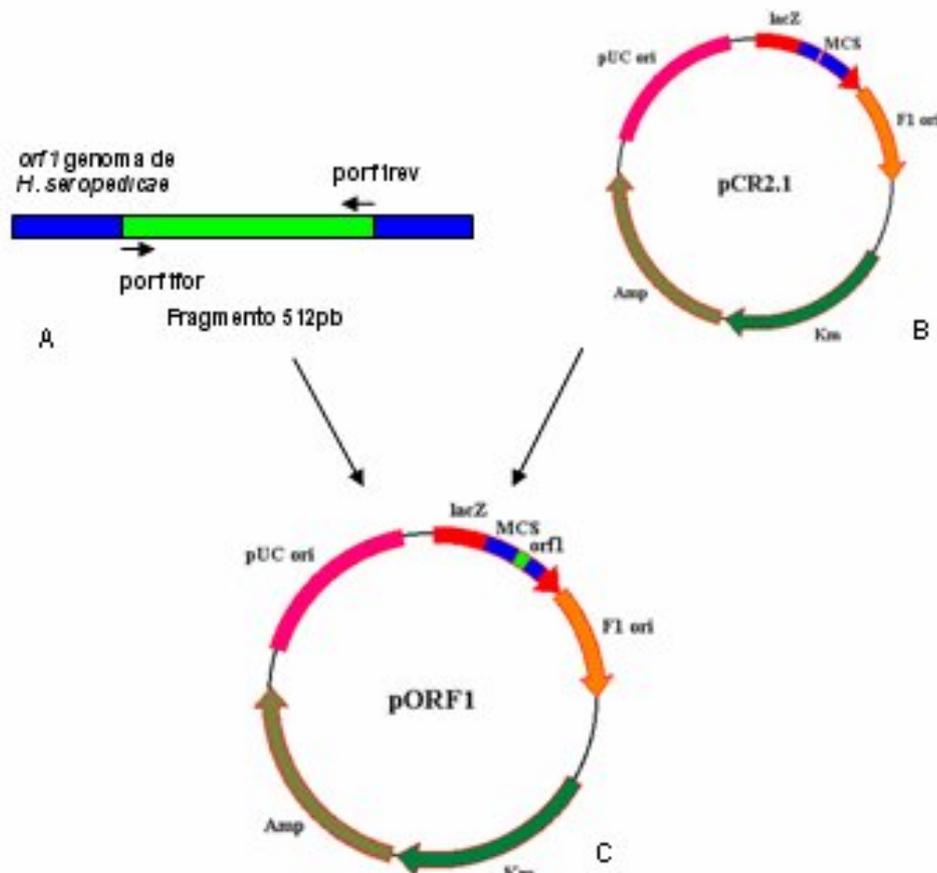


FIGURA 16 – CONSTRUÇÃO DO PLASMÍDEO pORF1. A) DNA genômico da estirpe selvagem SmR1, mostrando o fragmento de 512pb da *orf1* obtido da amplificação por PCR; B) Vektor pCR2.1; C) plasmídeo pORF1 obtido pela ligação do produto de PCR de 512pb ao vetor pCR2.1.

Para a confirmação da inserção do plasmídeo no genoma de *H. seropedicae* foi realizada a hibridização do DNA genômico de 5 possíveis mutantes e também da estirpe SmR1. O DNA genômico foi extraído e submetido à digestão com a enzima *SmaI*, que libera o gene todo, e hibridizado com o produto de PCR do plasmídeo pORF1 para a confirmação da mutação. Os perfis de hibridização da estirpe selvagem e das estirpes mutantes podem ser observados na figura 17. Quando o DNA genômico da estirpe SmR1 foi hibridizado ao produto de PCR do plasmídeo pORF1, foi possível observar uma banda de hibridização de aproximadamente 2 kb, que se refere ao gene *fdeA* selvagem. Por outro lado, a hibridização do DNA genômico da estirpe AMM1 ao produto de PCR apresentou uma banda de aproximadamente 6,4 kb, referente à inserção de todo o plasmídeo pORF1 no genoma.

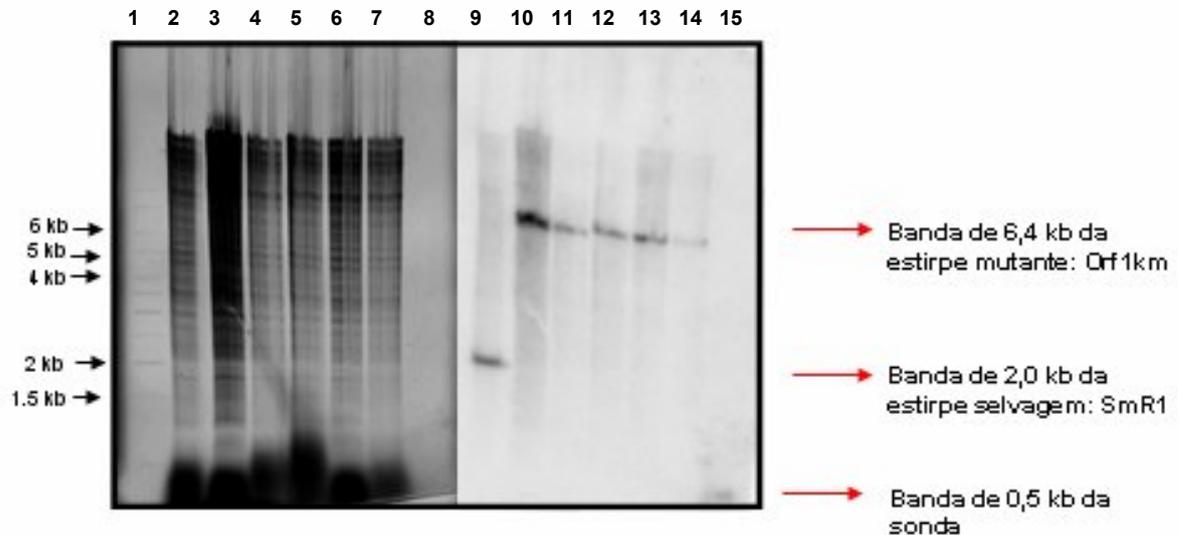


FIGURA 17 – HIBRIDIZAÇÃO DO DNA GENÔMICO DAS ESTIRPES SmR1 E AMM1 DE *H. seropedicae* COM O PRODUTO DE PCR DO PLASMÍDEO pORF1. As linhas de 1 a 8 mostram a imagem do gel de eletroforese de DNA em agarose 0,7% tratado com brometo de etídeo e processado pela inversão de cores, e as linhas 9 a mostram a autoradiografia da hibridização do DNA genômico com a sonda (produto de PCR do pORF1), obtidas pelo PhosphorImager Storm 820 (Molecular Dynamics). A linha 1 mostra o marcador de peso molecular 1kb ladder diluído (2ng). As linhas 2 e 9 mostram o DNA genômico da estirpe SmR1, e a sua hibridização, respectivamente, mostrando um fragmento de aproximadamente 2 kb. As linhas 3 a 7 mostram o DNA genômico das estirpes AMM1, e suas hibridização nas linhas 10 a 14, respectivamente, mostrando um fragmento de aproximadamente 5,9 kb. A linha 8 e a linha 15 referem-se a sonda, e sua autoradiografia, respectivamente, com tamanho de aproximadamente 0,5 kb.

A estirpe mutante obtida foi chamada de AMM1, e um esquema da recombinação ocorrida está na figura 18.

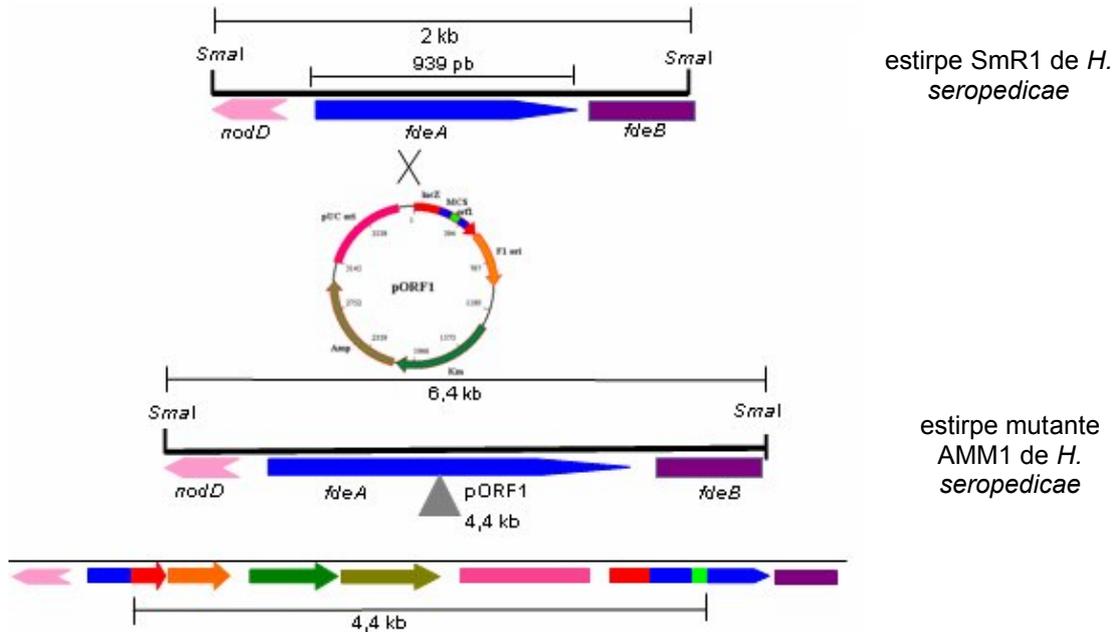


FIGURA 18 – ESQUEMA DA RECOMBINAÇÃO HOMÓLOGA SIMPLES OCORRIDA PARA ORIGINAR A ESTIRPE MUTANTE AMM1 DE *H. seropedicae*. O plasmídeo pORF1 tem tamanho de 4,4 kb

5.3 DETERMINAÇÃO DO FENÓTIPO DO MUTANTE AMM1

5.3.1 Crescimento em meio NFbHPN e NFbHPN-malato sólido variando as fontes de carbono

Uma vez que o operon *fde* de *H. seropedicae* parece estar envolvido no catabolismo de compostos aromáticos, o crescimento das estirpes SmR1, DR2 e AMM1 foi testado frente a diferentes compostos aromáticos, em especial os fenólicos (tabela 8). O meio escolhido para este teste foi o NFb que é o meio preferencial para o crescimento de *H. seropedicae*. O experimento foi realizado neste meio na presença de malato e do composto a ser testado, para determinar uma concentração tóxica para o microrganismo, bem como apenas na presença do composto para observar se as estirpes testadas conseguiriam catabolizar tais moléculas, utilizando-as como fontes alternativas de carbono. Se a mutação em *fdeA* estivesse relacionada com o catabolismo de tais compostos, seria esperado que a estirpe AMM1 não crescesse na presença destes compostos. A quantidade de malato utilizado para os experimentos foi de 37 mmol/L, que é a concentração padrão utilizada. Salicilato de sódio e ácido

benzóico foram utilizados por serem compostos já descritos na literatura como sendo fontes alternativas de carbono para algumas bactérias, como é o caso de *Streptomyces* sp. WA46 (ISHYIAMA *et al.*, 2004) e *Pseudomonas putida* (MORENO & ROJO, 2008). Os flavonóides naringenina e quercetina foram utilizados, pois além de serem compostos fenólicos, são compostos que foram encontrados como tendo a capacidade de induzir a expressão do operon em estudo (MARIN, 2006a), e também há estudos na literatura que mostram que certas bactérias tem a capacidade de clivar estes compostos, utilizando-os como fontes de carbono. Pillai & Swarup, 2002 observaram que *P. putida* PML2 é capaz de utilizar quercetina como fonte de carbono, e Rao & Cooper, 1994, observaram que *Rhizobium leguminosarum* bv. phaseoli é capaz de catabolizar quercetina, e *R. leguminosarum* bv. viciae é capaz de utilizar naringenina como fonte de carbono. Visando isto, uma das hipóteses é que os flavonóides sirvam como sinal para a sua própria degradação, uma vez que, como será visto posteriormente, a naringenina induz a expressão do operon em estudo. Resorcinol é uma molécula que foi encontrada como sendo resultante da clivagem da 7,4'-dihidroxi-flavona por *R. leguminosarum* bv. trifolii, e seria o resultado da clivagem do anel C, liberando os anéis A (resorcinol) e B (ácido p-hidroxibenzóico) (RAO & COOPER, 1994). Os aminoácidos aromáticos foram testados, pois também podem ser utilizados como fontes de carbono.

TABELA 8 – TESTE DE CRESCIMENTO DAS ESTIRPES SmR1, DR2 E AMM1 EM MEIO NFbHPN E NFbHPN-MALATO. (-) significa que não houve crescimento e (+) significa que houve crescimento.

Composto [final]	NFbHPN	NFbHPN- malato
Salicilato de sódio 6,2mM	-	-
Salicilato de sódio 5mM	-	+
Salicilato de sódio 3,1mM	-	+
Salicilato de sódio 2mM	-	+
Salicilato de sódio 1,2mM	-	+
Salicilato de sódio 0,62mM	-	+
Salicilato de sódio 0,31mM	-	+
Salicilato de sódio 0,12mM	-	+
Ácido benzóico 0,8mM	-	-
Ácido benzóico 5mM	-	-
Resorcinol 5mM	+	+
Naringenina 1mM	+	+
Naringenina 2mM	+	+

Naringenina 3mM	+	+
Naringenina 4mM	+	+
Naringenina 5mM	-	+
Quercetina 1mM	+	+
Quercetina 2mM	+	+
Quercetina 3mM	+	+
Quercetina 4mM	+	+
Quercetina 5mM	-	+
Tirosina 1mM	-	+
Triptofano 1mM	-	+
Fenilalanina 1mM	-	+

Como mostra a tabela 8, os únicos compostos utilizados como fonte de carbono foram o resorcinol, a naringenina e a quercetina, porém a estirpe mutante AMM1 também foi capaz de utilizá-las. O crescimento em meio contendo malato foi feito para observar uma possível dose tóxica dos compostos testados, sendo que o salicilato de sódio na concentração de 6,2 mmol/L e o ácido benzóico parece ser tóxicos, uma vez que nesta concentração não houve crescimento de nenhuma das 3 estirpes testadas.

5.3.2 Crescimento em meio NFbHPN e NFbHPN-malato líquido variando as fontes de carbono

Após o teste em meio sólido (item 5.3.1) foi realizado também um teste semelhante em meio líquido, utilizando a metodologia de BOCHNER *et al.*, 2008, que é um método simples, uma vez que a detecção do crescimento é feita através da presença da coloração púrpura (item 4.16.2). Como controles positivos foram testados os monossacarídeos glucose, manose e ramnose, e o ácido carboxílico malato. Como controles negativos foram utilizados os dissacarídeos sacarose e trealose, uma vez que já é conhecida a incapacidade de *H. seropedicae* utilizar dissacarídeos como fonte de carbono. Os compostos aromáticos testados com esta técnica foram naringenina, resorcinol, ácido benzóico, salicilato de sódio e fenol. O etanol foi utilizado também como controle uma vez que a naringenina e o ácido benzóico são diluídos em etanol. O volume utilizado foi de 1,6 µL, que é o mesmo volume da solução de naringenina utilizada.

Na presença dos monossacarídeos houve crescimento das 3 estirpes e das 2 colônias da cada estirpe, assim como na presença tabela 9.

H. seropedicae foi capaz de crescer em meio contendo ácido benzóico, naringenina, resorcinol e etanol como fontes de carbono, como é possível observar na tabela 9.

TABELA 9 – LOCALIZAÇÃO POR POÇO DAS CONDIÇÕES TESTADAS. (+) significa crescimento e (-) significa ausência de crescimento.

Posição	Condição	SmR1	DR2	AMM1
A1/C5	NFbHPN sem bactéria	-	-	-
A2/C6	NFbHPN com bactéria	-	-	-
A3/C7	NFbHPN-malato sem bactéria	-	-	-
A4/C8	NFbHPN-malato com bactéria	+	+	+
B1/D5	NFbHPN Glucose	+	+	+
B2/D6	NFbHPN Glucose sem bactéria	-	-	-
B3/D7	NFbHPN Sacarose	-	-	-
B4/D8	NFbHPN Sacarose sem bactéria	-	-	-
C1/E5	NFbHPN Manose	+	+	+
C2/E6	NFbHPN Manose sem bactéria	-	-	-
C3/E7	NFbHPN Trealose	-	-	-
C4/E8	NFbHPN Trealose sem bactéria	-	-	-
D1/F5	NFbHPN Ramnose	+	+	+
D2/F6	NFbHPN Ramnose sem bactéria	-	-	-
D3/F7	NFbHPN Acido benzóico	+	+	+
D4/F8	NFbHPN Acido benzóico sem bactéria	-	-	-
E1/G5	NFbHPN Salicilato de sódio	-	-	-
E2/G6	NFbHPN Salicilato de sódio sem bactéria	-	-	-
E3/G7	NFbHPN Tirosina (fonte de C)	-	-	-
E4/G8	NFbHPN Tirosina (fonte de C) sem bactéria	-	-	-
F1/H5	NFbHP Tirosina (fonte de N)	-	-	-
F2/H6	NFbHP Tirosina (fonte de N) sem bactéria	-	-	-
F3/H7	NFbHPN Fenilalanina (fonte de C)	-	-	-
F4/H8	NFbHPN Fenilalanina (fonte de C) sem bactéria	-	-	-
G1/A9	NFbHP Fenilalanina (fonte de N)	+	+	+
G2/A10	NFbHP Fenilalanina (fonte de N) sem bactéria	-	-	-
G3/A11	NFbHPN 0,5mmol/L fenol + malato	+	+	+
G4/A12	NFbHPN 0,5mmol/L fenol + malato sem bactéria	-	-	-
H1/B9	NFbHPN 1 mmolL fenol + malato	+	+	+

H2/B10	NFbHPN 1 mmolL fenol + malato sem bactéria	-	-	-
H3/B11	NFbHPN 1 mmolL fenol	-	-	-
H4/B12	NFbHPN 1 mmolL fenol sem bactéria	-	-	-
A5/C9	NFbHPN Naringenina em pó	+	+	+
A6/C10	NFbHPN Naringenina em pó sem bactéria	-	-	-
A7/C11	NFbHPN Naringenina solução	+	+	+
A8/C12	NFbHPN Naringenina solução sem bactéria	-	-	-
B5/D9	NFbHPN Etanol	+	+	+
B6/D10	NFbHPN Etanol sem bactéria	-	-	-
B7/D11	NFbHPN Resorcinol	+	+	+
B8/D12	NFbHPN Resorcinol sem bactéria	-	-	-

Como mostra a tabela 9, *H. seropedicae* é capaz de utilizar resorcinol, etanol e naringenina como fontes alternativas de carbono, que foi o observado no crescimento em placa, mas como a estirpe mutante AMM1 também cresceu não foi possível determinar o seu fenótipo utilizando este teste.

5.3.5 Análise do fenótipo da estirpe AMM1 em HPLC

Com a finalidade de verificar se *H. seropedicae* SmR1 é capaz de degradar naringenina e as estirpes mutantes *nodD*⁻ e *fdeA*⁻ apresentam um fenótipo diferente, foram feitas análises em HPLC, onde foram testadas 2 colônias diferentes das estirpes SmR1, DR2 e AMM1, que foram cultivadas a 30°C, sob agitação a 120 rpm por 30 horas. Foi coletado 1 mL de cultura nos tempos (horas) 0, 6, 12, 24 e 30. As duplicatas biológicas foram reprodutivas, e por isso será mostrado apenas um cromatograma de cada estirpe, condição e tempo. Os padrões de naringenina 1 mg/mL e crisina 1 µg/mL foram injetados no HPLC pra conhecer o tempo de retenção destes compostos na coluna, e os cromatogramas referentes a estes padrões estão na figura 19.

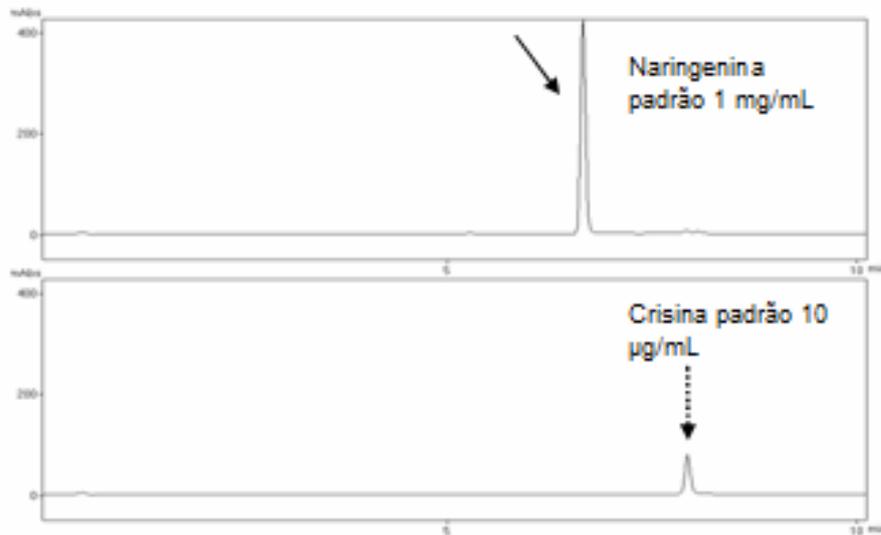


FIGURA 19 – CROMATOGRAMAS REFERENTES AOS PADRÕES. A naringenina e a crisina foram pesadas e diluídas em metanol, filtradas e 5 µL destas soluções foram injetadas separadamente. —→ Naringenina;→ Crisina.

O tempo total de corrida desta coluna foi de 14 minutos, mas depois dos 10 minutos não aparecem mais picos, e por isso as figuras apresentadas acabam em 10 minutos. Como mostra a figura 19, a naringenina é eluída da coluna por volta de 6 minutos de corrida, enquanto que a crisina é eluída em aproximadamente 8 minutos, sendo possível distinguir bem os 2 picos, e sendo assim, a crisina pode ser utilizada como padrão interno de corrida.

Os controles foram feitos em meio NFbHPN-malato e NFbHPN contendo 2 mmol/L de naringenina, e podem ser vistos na figura 20.

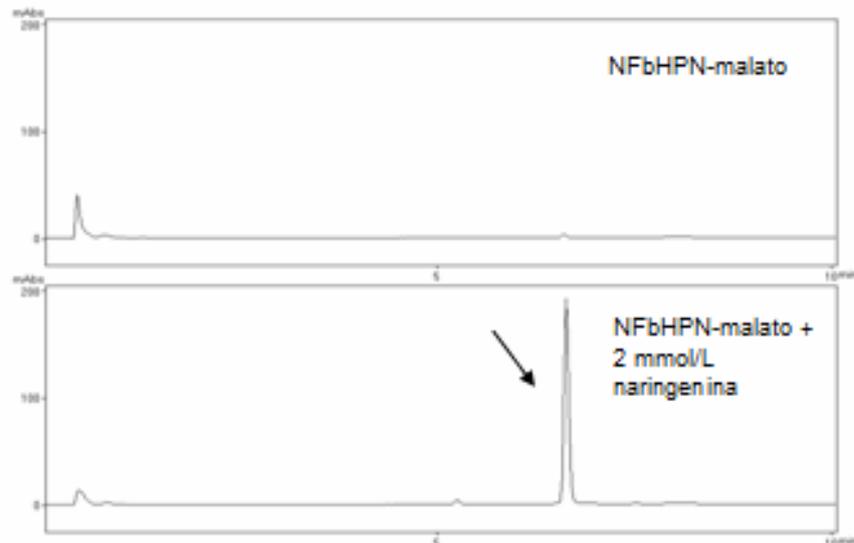


FIGURA 20 – CROMATOGRAMAS REFERENTES AOS CONTROLES NO TEMPO 0 HORAS. NFbHPN-malato sem adição de naringenina, e NFbHPN-malato contendo 2 mmol/L de naringenina. As amostras foram filtradas antes da injeção, sendo que o volume injetado foi de 5 μ L. —→ Naringenina.

Como mostra a figura 20, o meio NFbHPN-malato não absorve em 280nm, confirmando que o pico visto no meio NFbHPN-malato contendo 2 mmol/L de naringenina é realmente a naringenina e não algum outro composto do próprio meio de cultura. Além disto, a naringenina é eluída no mesmo tempo que mostra a injeção da solução padrão (figura 19).

O ponto inicial da curva foi em 0 hora (figura 21) e apresenta o pico de naringenina e o pico de crisina conforme o padrão.

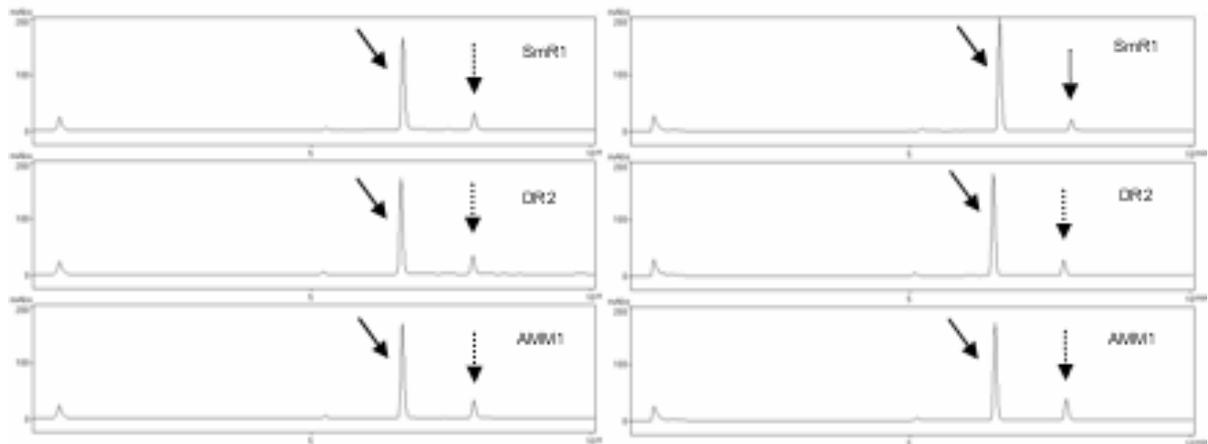


FIGURA 21 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 0 HORAS. Os cromatogramas da esquerda são das estirpes SmR1, DR2 e AMM1 cultivadas em meio NFbHPN contendo 2 mmol/L de naringenina, e os cromatogramas da direita são dessas mesmas

estirpes cultivadas em meio NFbHPN-malato contendo 2 mmol/L de naringenina. 1 mL das culturas foi centrifugado a 13400 rpm, filtrado, e adicionado a solução padrão de crisina, que ficou com concentração final de 10 µg/mL. —▶ Naringenina;▶ Crisina.

Como é possível observar na figura 21 tanto os cromatogramas das culturas em meio NFbHPN contendo naringenina quanto os cromatogramas das culturas em meio NFbHPN-malato contendo naringenina mostram o mesmo perfil. A naringenina é eluída por volta dos 6 minutos e a intensidade do pico está próxima aos 200 mAbs, mostrando que a naringenina está presente nas culturas de maneira semelhante.

Após 6 horas de cultivo o perfil não se alterou como mostra a figura 22.

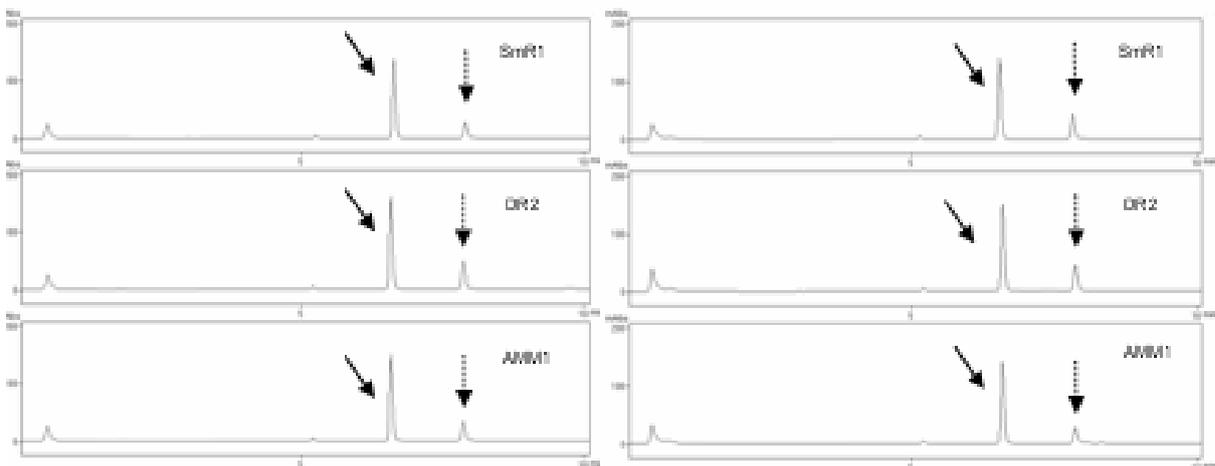


FIGURA 22 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 6 HORAS. Os cromatogramas da esquerda são das estirpes SmR1, DR2 e AMM1 cultivadas em meio NFbHPN contendo 2 mmol/L de naringenina, e os cromatogramas da direita são dessas mesmas estirpes cultivadas em meio NFbHPN-malato contendo 2 mmol/L de naringenina. 1 mL das culturas foi centrifugado a 13400 rpm, filtrado, e adicionado a solução padrão de crisina, que ficou com concentração final de 10 µg/mL. —▶ Naringenina;▶ Crisina.

Até as 12 horas de cultivo também não houve alteração no perfil cromatográfico seja em relação às estirpes ou em relação aos meios utilizados (figura 23).

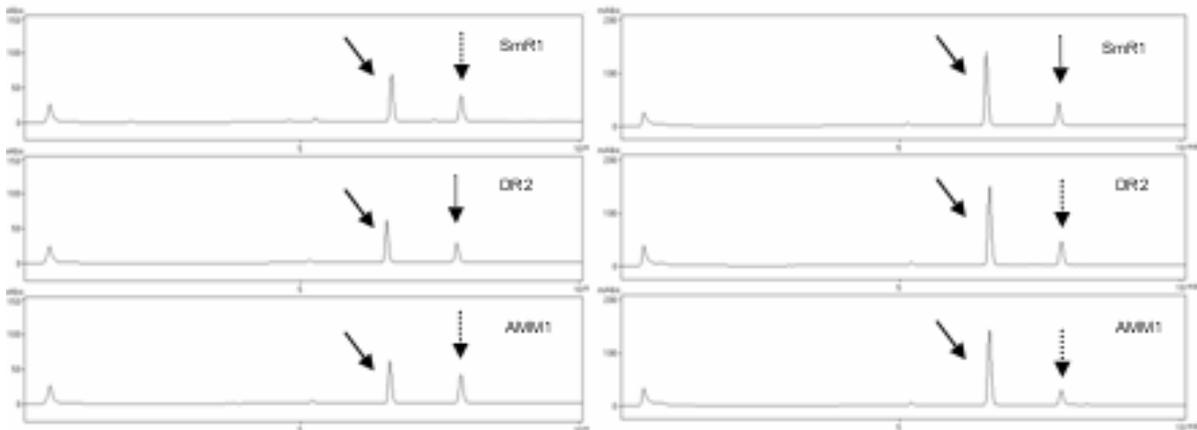


FIGURA 23 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 12 HORAS. Os cromatogramas da esquerda são das estirpes SmR1, DR2 e AMM1 cultivadas em meio NFbHPN contendo 2 mmol/L de naringenina, e os cromatogramas da direita são dessas mesmas estirpes cultivadas em meio NFbHPN-malato contendo 2 mmol/L de naringenina. 1 mL das culturas foi centrifugado a 13400 rpm, filtrado, e adicionado a solução padrão de crisina, que ficou com concentração final de 10 µg/mL. —> Naringenina;> Crisina.

Com 24 horas de cultivo a estirpe SmR1 mostrou um perfil diferente, tanto em meio NFbHPN contendo naringenina, quanto em meio NFbHPN-malato contendo naringenina (figura 24).

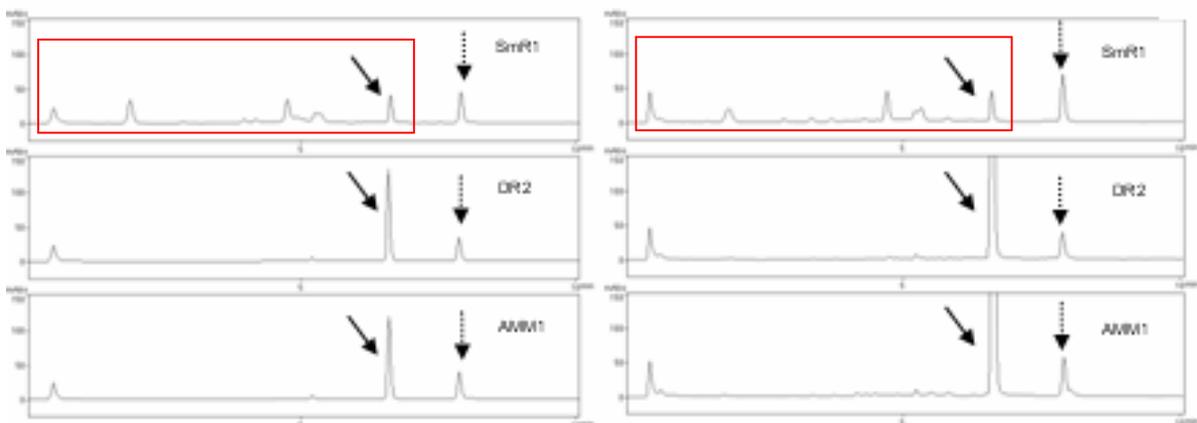


FIGURA 24 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 24 HORAS. Os cromatogramas da esquerda são das estirpes SmR1, DR2 e AMM1 cultivadas em meio NFbHPN contendo 2 mmol/L de naringenina, e os cromatogramas da direita são dessas mesmas estirpes cultivadas em meio NFbHPN-malato contendo 2 mmol/L de naringenina. 1 mL das culturas foi centrifugado a 13400 rpm, filtrado, e adicionado a solução padrão de crisina, que ficou com concentração final de 10 µg/mL. —> Naringenina;> Crisina. Os picos novos, orinados da degradação da naringenina estão nas caixas vermelhas.

Após 24 horas de cultivo é possível observar que a intensidade do pico de naringenina diminuiu e novos picos, eluídos anteriormente à naringenina, apareceram

na injeção referente à estirpe SmR1, tanto em meio com malato e naringenina quanto em meio contendo apenas naringenina. Isto mostra que a naringenina é degradada pela estirpe selvagem SmR1 e que está sendo metabolizada de maneira semelhante mesmo quando há malato no meio de cultura, ou seja, ela é degradada independente da sua utilização como fonte alternativa de carbono. Assim como *Pseudomonas putida* PML2 (PILLAI & SWARUP, 2002) e *Rhizobium leguminosarum* bv. phaseoli, que clivam quercetina, *R. leguminosarum* bv. viciae que cliva naringenina, *R. leguminosarum* bv. trifolii que cliva 7,4'-dihidroxi-flavona, *R. meliloti* que cliva luteolina, *R. fredii*, que cliva genesteína e *Rhizobium* sp. NGR234 que cliva daidzeína e apigenina (RAO & COOPER, 1994), *H. seropedicae* SmR1 também tem a capacidade de clivar flavonóides, sendo o único testado até o momento a naringenina. Contudo, o mecanismo de clivagem realizado por *H. seropedicae* ainda não foi esclarecido, uma vez que os produtos da clivagem da naringenina não foram identificados.

As estirpes DR2 e AMM1 parecem não terem clivado a naringenina, indicando que a proteína HsNodD, que é a proteína ativadora da expressão da via, assim como o operon *fde* estão envolvidos no metabolismo deste composto. Porém no crescimento em placa (item 5.3.1) as 3 estirpes testadas cresceram na presença deste composto. O que pode ter acontecido é que as estirpes mutantes ainda clivem a naringenina, mas em menor quantidade e por isso não foi possível observar uma mudança nos cromatogramas destas estirpes.

O último ponto de coleta foi com 30 horas de cultivo, e os cromatogramas obtidos apresentam um perfil semelhante ao observado em 24 horas, porém em 30 horas houve o consumo quase completo da naringenina (figura 25).

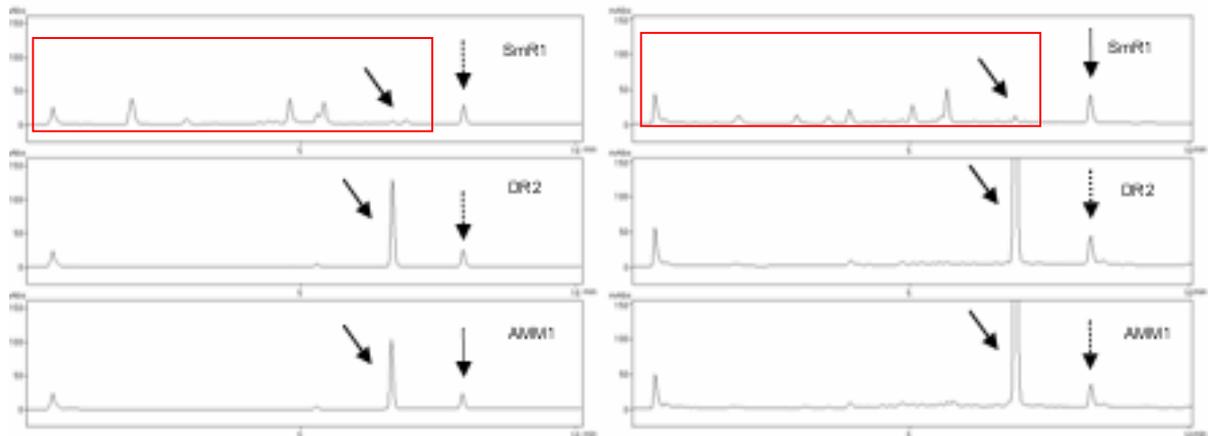


FIGURA 25 – CROMATOGRAMAS REFERENTES À INJEÇÃO DAS AMOSTRAS NO TEMPO 30 HORAS. Os cromatogramas da esquerda são das estirpes SmR1, DR2 e AMM1 cultivadas em meio NFbHPN contendo 2 mmol/L de naringenina, e os cromatogramas da direita são dessas mesmas estirpes cultivadas em meio NFbHPN-malato contendo 2 mmol/L de naringenina. 1 mL das culturas foi centrifugado a 13400 rpm, filtrado, e adicionado a solução padrão de crisina, que ficou com concentração final de 10 µg/mL. —> Naringenina;> Crisina. Os picos novos, originados da degradação da naringenina estão nas caixas vermelhas.

Com 30 horas de cultivo o pico de naringenina desapareceu quase por completo, ou seja, ainda neste tempo a estirpe SmR1 continua degradando esta molécula.

Após este experimento foi possível concluir que a via mutada neste trabalho (estirpe AMM1) está envolvida na degradação de naringenina, pois a estirpe mutante não foi capaz de degradá-la. Contudo, o mecanismo de clivagem ainda não pode ser proposto.

5.4 REGULAÇÃO DA EXPRESSÃO DOS GENES *nodD* E *fdeA* DE *H. seropedicae* E DO OPERON *nodABC* EM *H. seropedicae*

Para avaliar a regulação da expressão dos genes *nodD* e *fdeA* de *H. seropedicae* as fusões pSU11 (*nodD::lacZ*) e pSUI (*fdeA::lacZ*), que contêm o mesmo fragmento da região intergênica entre esses 2 genes clonados em ambas as direções (figura 26; item 4.1), foram conjugadas nas estirpes SmR1, DR2 e AMM1 (item 4.1) de *H. seropedicae*, e testados quanto ao seus níveis de expressão normais e na presença de alguns compostos.

```

GGTGAGGAACGGCGAGACGCCGCCCTCGAAGGCATAGGCTGCAGGCCCG PstI
CGGCCATGGTCAGCAGCGCGCAAGTTGCCAGCCGGCATGCCGGCACGAAT
CCTCTCTTACGCTCTGTCTCCTTGTCCTCTACCCCTGCCTGGCGGGGTGTG
GCCGCTATCGGC GCGTTGTCAATGGCCAGGCTCAACA GCGGC GCATC CAGT
GCGAGTGC GCGCGCCGGCTTGC GCGCG CAGCC CTTTC TGGTGGGGCCATGC
AGCGGATGAGTGAAGCTTAAGAGTGA GCTTGATA TTGATCAAATGGATTG -35/-10 nodD
TTTTGATTCATGATATGGACGGCATCAATACATTGACCACC CCATCCGTC -35/-10 fdeA
GTCTATC CGAGGGTCCGGCAAGAACAA GCACAACAAC GGAGACCAGCGCCA
TGCCTTTCAACAAGCTCGACCTCAATCTTCTGGTCCCTGGATGC ACTG
CTCAGGAGATGAGCATCAGCGCGCGCGAAAAGATCCATCTGAGCCA
GTCGGCCATGAGCAATGCCCTGGCGCGGCTGCGCGAGTATTTTCGATGATG fdeA
AATTGCTGATCCAGGTGGGCCGGCGCATGGAGCCCA GCGCGCGCGCGAG nodD
GTGCTCAAGGATGCGGTGCATGATGTGCTGCGGCGTATCGATGGCTCCAT região intergênica
CGCGGCGCTGCCGGCCTTTCGTGCCGGCCGAGTCCACGCGCGAGTTTCGCA
TCTCGGTTTTCGGACTTTACGCTCTCCGTCCTCATCCCGCGGTGCTGGCG
CGCGCGCACGCCGAGGGCAAGCACATCCGCTTTGCCCTGATGCCCGAGGT
GCAAGACCGGACCCGCTCGCTGGATCGGGCCGAGGTGGACCTGCTGCTCT
TGCCGCAAGGAATTCTGCACGCCCGATCATCTGCCGAAGAGGTCTTCCGC EcoRI

```

FIGURA 26 – REGIÃO CLONADA EM pMP220 E EM pPW452, ORIGINANDO AS FUSOES *nodD::lacZ* (pSU11) E *fdeA::lacZ* (pSUI). A região de estudo foi digerida com as enzimas *Pst*I e *Eco*RI, liberando um fragmento de 0,82 kb. Em roxo esta a sequência de nucleotídeos do gene *fdeA*, e também a região promotora -35/-10 dependente do fator σ^{70} . Em vermelho esta a sequência de nucleotídeos do gene *nodD*, e também a região promotora -35/-10 dependente do fator σ^{70} . Na região em azul estão os nucleotídeos da região intergenica entre os genes *nodD* e *fdeA*.

Uma vez que o principal alvo de regulação das proteínas NodD em rizóbios é o promotor do operon *nodABC*, a fusão pNB8 (*nodABC::lacZ* de *Rhizobium* sp. NGR234; FELLAY *et al.*, 1998) também foi conjugada nas estirpes SmR1 e DR2 de *H. seropedicae* testadas neste trabalho. Esta fusão foi utilizada como controle para observar uma possível regulação deste promotor pela HsNodD, devido à sua semelhança com NodDs de rizóbios. Além disto, a HsNodD reconhece o *nod*-box do operon *nodABC* de *Rhizobium* sp. NGR234, mostrado em ensaios *in vitro* (WASSEM, comunicação pessoal).

Neste trabalho foram testados alguns compostos fenólicos como possíveis moduladores da expressão desses genes. Os compostos utilizados foram: os flavonóides naringenina, quercetina, apigenina, luteolina, kaempferol, crisina, genisteína, daidzeína e catequina, e alguns intermediários da via de degradação de compostos fenólicos, como ácido benzóico, ácido salicílico, salicilato de sódio e resorcinol, e os aminoácidos aromáticos tirosina, fenilalanina e triptofano. Contudo, os únicos compostos que apresentaram resposta como indutores foram a naringenina, apigenina e crisina. Quercetina, kaempferol e luteolina não ativaram a expressão do operon *fde*. Os resultados referentes aos intermediários do metabolismo de

compostos fenólicos não serão mostrados, pois não houve diferença em relação ao controle. Serão mostrados apenas os resultados obtidos em relação aos flavonóides.

5.4.1 Efeito da proteína HsNodD sobre a expressão dos genes *fdeA* e *nodD* de *H. seropedicae* e sobre o operon *nodABC* de *Rhizobium* sp. NGR234

Para conhecer o nível basal de expressão dos promotores testados, foram realizados 3 ensaios de β -galactosidase na ausência de qualquer flavonóide. Cada um destes ensaios foi realizado com 3 colônias de cada transconjugante em duplicata de cultura, e o padrão de expressão pode ser visto na figura 27.

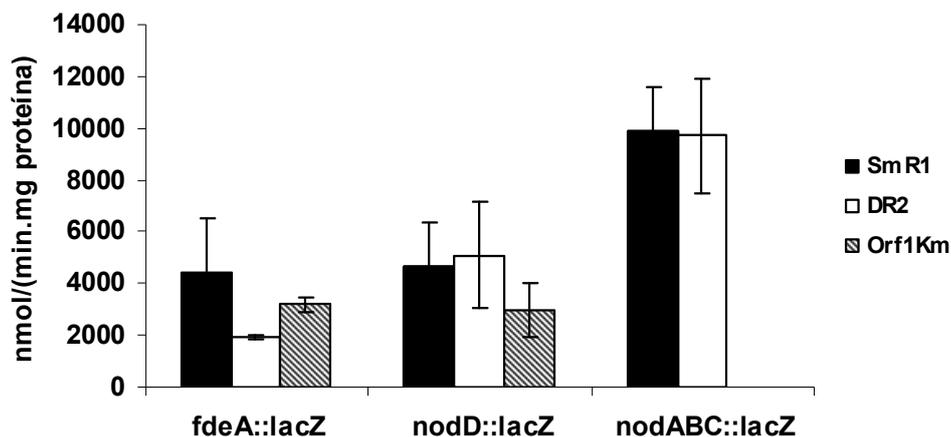


FIGURA 27 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn CONTROLE (SEM FLAVONÓIDES). A fusão pNB8 (*nodABC::lacZ*) não foi conjugada na estirpe AMM1. Para os transconjugantes com as fusões pSU11 (*nodD::lacZ*) e pSUI (*fdeA::lacZ*) foram realizados 3 ensaios independentes, sendo que para cada um deles foram utilizadas 3 colônias diferentes de cada transconjugante, em duplicata de cultura. Para os transconjugantes com a fusão pNB8 foi realizado apenas 1 ensaio, que continha 5 colônias diferentes em duplicata de cultura. Para todos os ensaios e todos os transconjugantes, o tempo de crescimento das culturas foi de 6 horas, a 30°C e 120 rpm, assim como o modo de inoculação foi o mesmo.

Em relação a expressão de *fdeA* a diferença observada entre as estirpes SmR1 e DR2 (figura 27) não será considerada devido ao desvio padrão, ou seja não há diferença de expressão. O mesmo foi observado em relação às fusões *nodD::lacZ* e *nodABC::lacZ* a proteína HsNodD parece não regular a expressão destes promotores, uma vez que não há diferença entre as estirpes. A fusão *nodABC::lacZ* é expressa tanto na estirpe selvagem quanto na estirpe mutante DR2, mostrando que o operon

nodABC de *Rhizobium* sp. NGR234 é expresso por *H. seropedicae* e, provavelmente, a sua expressão independe de HsNodD.

5.4.2 Efeito do flavonóide naringenina sobre a atividade da proteína HsNodD

A proteína NodD de rizóbios tem sua atividade modificada na presença de determinados flavonóides, devido às mudanças conformacionais em sua estrutura tridimensional ocasionadas pela ligação do flavonóide. Marin e colaboradores (2005a) mostraram que o gene *fdeA* de *H. seropedicae*, em *E. coli*, tem a sua expressão regulada pela proteína HsNodD em resposta aos flavonóides naringenina e quercetina. Assim, neste trabalho estes compostos foram utilizados para determinar o seu efeito sobre a atividade da proteína HsNodD frente às fusões *nodD::lacZ* e *fdeA::lacZ* de *H. seropedicae* e *nodABC::lacZ* de *Rhizobium* sp. NGR234, em *H. seropedicae* estirpes SmR1, DR2 e AMM1.

A concentração de naringenina utilizada para os ensaios foi de 200 µmol/L. Tendo os transconjugantes (SmR1 pSUI, DR2 pSUI e AMM1 pSUI), foram realizados 3 ensaios de β-galactosidase, sendo que em cada um foram utilizadas as culturas de 3 colônias de cada transconjugante em duplicata de cultura. O controle foi feito com etanol (figura 28).

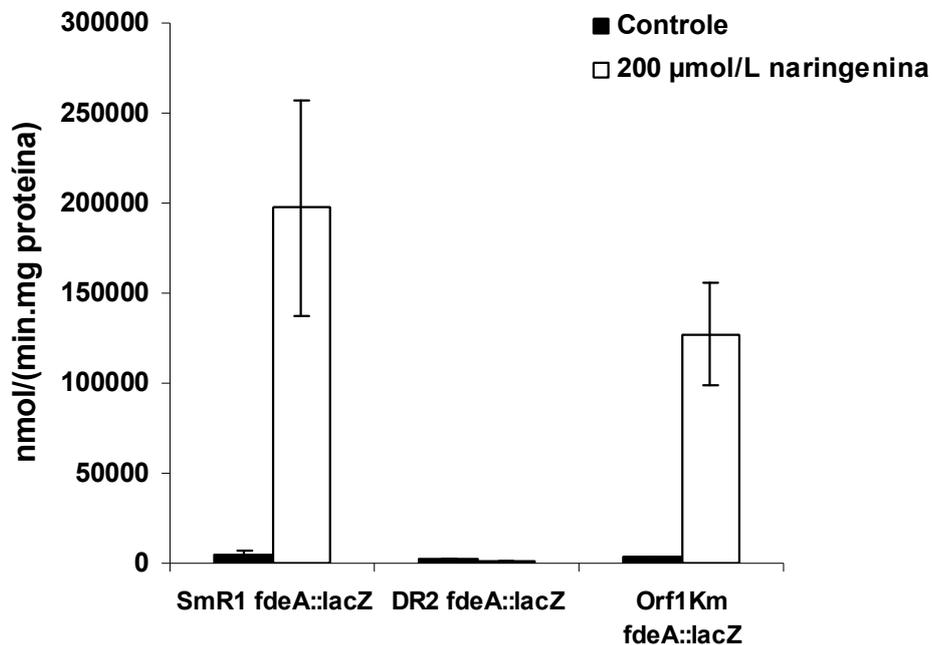


FIGURA 28 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO *fdeA::lacZ*. Média de 3 ensaios utilizando a fusão *fdeA::lacZ*, sendo que para cada ensaio foram utilizadas 3 colônias de cada transconjugante em duplicata de cultura. Para os ensaios, os transconjugantes foram cultivados em meio NFbHPN-malato contendo 0,2 mmol/L de naringenina e como controle os transconjugantes foram cultivados em meio NFbHPN-malato contendo etanol (mesmo volume da solução de naringenina utilizada). O crescimento foi feito durante 6 horas, a 30°C e 120 rpm, antes dos ensaios.

A figura 28 mostra que a expressão de *fdeA* é aumentada cerca de 50 vezes na presença de naringenina, tanto na estirpe selvagem SmR1 quanto na estirpe mutante AMM1, indicando que a naringenina é indutora da expressão deste operon. Além disto, também pode ser visto que a expressão deste operon é dependente da proteína HsNodD, uma vez que na estirpe *nodD*⁻ (DR2) não há aumento de expressão na presença de naringenina. Contudo, este resultado contraria o que foi observado em *E. coli* onde a HsNodD era repressora e para que a repressão deixasse de ocorrer fazia-se necessária à presença de naringenina ou quercetina.

A fusão pSU11 (*nodD::lacZ*) também teve a sua expressão testada na presença de naringenina, e o resultado está na figura 29.

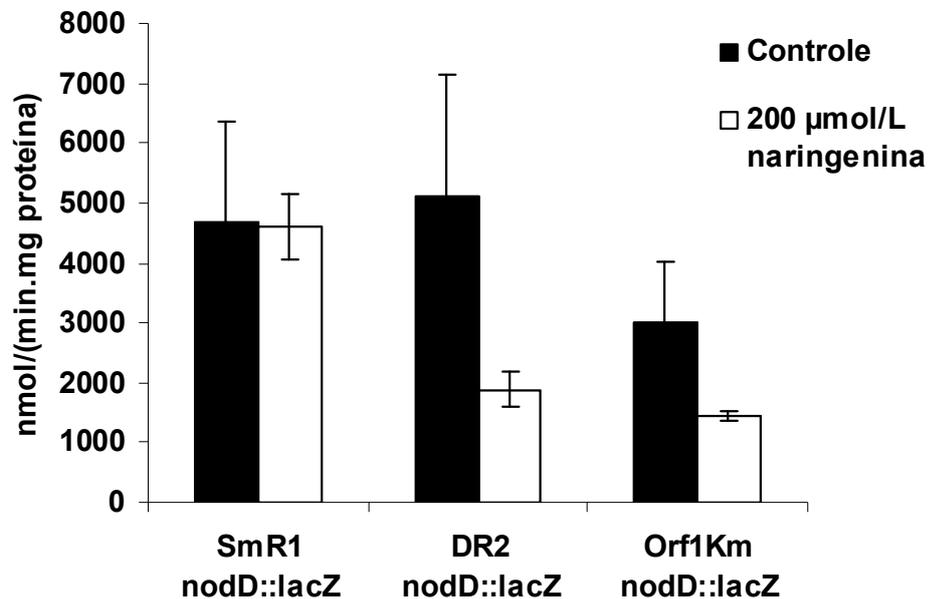


FIGURA 29 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO *nodD::lacZ*. Média de 3 ensaios utilizando a fusão *nodD::lacZ*, sendo que para cada ensaio foram utilizadas 3 colônias de cada transconjugante em duplicata de cultura. Para os ensaios, os transconjugantes foram cultivados em meio NFbHPN-malato contendo 0,2 mmol/L de naringenina e como controle os transconjugantes foram cultivados em meio NFbHPN-malato contendo etanol (mesmo volume da solução de naringenina utilizada. O crescimento foi feito durante 6 horas, a 30°C e 120 rpm, antes dos ensaios.

A figura 29 mostra que a expressão do gene *nodD* responde a naringenina somente nas estirpes mutantes DR2 e AMM1. Este efeito não é observado na estirpe selvagem SmR1, indicando que as mutações nos genes *nodD* e *fdeA* podem estar relacionadas com um sistema que responde à naringenina e que regula a expressão do gene *nodD*. Estes resultados também indicam que a HsNodD não está envolvida na expressão do gene *nodD* de *H. seropedicae*, mesmo na presença de naringenina.

A expressão da fusão *nodABC::lacZ* foi utilizada como controle, e como mostra a figura 30 este operon é expresso em ambas as estirpes SmR1 e DR2, indicando novamente que a proteína HsNodD não está envolvida neste processo. No entanto, a presença de naringenina inibe a expressão deste operon sugerindo que uma outra proteína de *H. seropedicae* que responde a naringenina pode ser a ativadora da expressão desses genes.

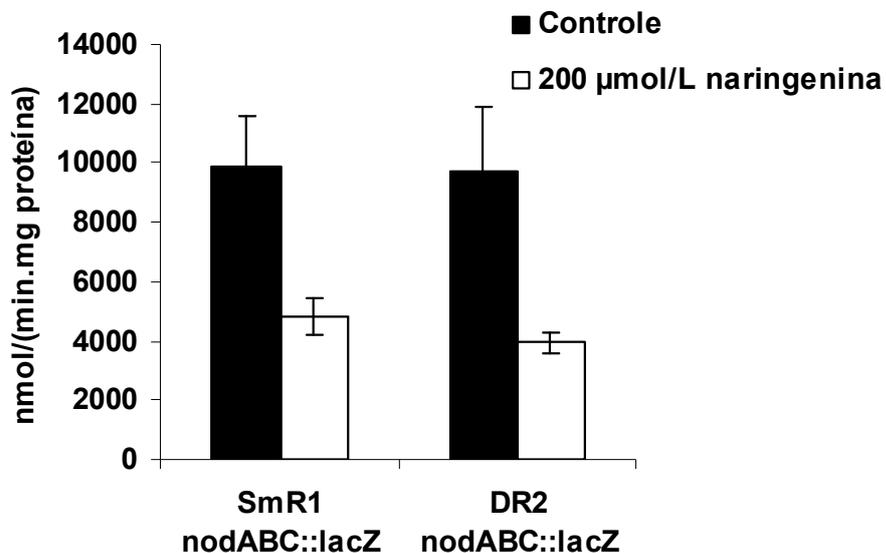


FIGURA 30 – ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE EM nmol/min.mg ptn DA FUSÃO *nodABC::lacZ* DE *Rhizobium* sp. NGR234. Média de 1 ensaio que foi realizado com 5 colônias de cada transconjugante em duplicata de cultura. Para os ensaios, os transconjugantes foram cultivados em meio NFbHPN-malato contendo 0,2 mmol/L de naringenina e como controle os transconjugantes foram cultivados em meio NFbHPN-malato contendo etanol (mesmo volume da solução de naringenina utilizada). O crescimento foi feito durante 6 horas, a 30°C e 120 rpm, antes dos ensaios.

5.4.3 Efeito de diversos flavonóides sobre a expressão do gene *fdeA* em *H. seropedicae*

Uma vez que o promotor que apresentou indução na presença de naringenina foi o promotor do gene *fdeA*, a fusão contendo este promotor (pSUI) foi utilizada testando outros flavonóides que também ativassem a sua expressão na presença da proteína HsNodD em *H. seropedicae*. Foram testados flavonóides de algumas classes, como as flavonas apigenina, luteolina e crisina, os flavonóis kaempferol e quercetina (figura 31), e as isoflavonas genisteína e daidzeína, e a flavana catequina (figura 32). A naringenina foi utilizada como controle positivo do ensaio, pois já é conhecida a resposta deste flavonóide na expressão do promotor *fdeA* (figura 28).

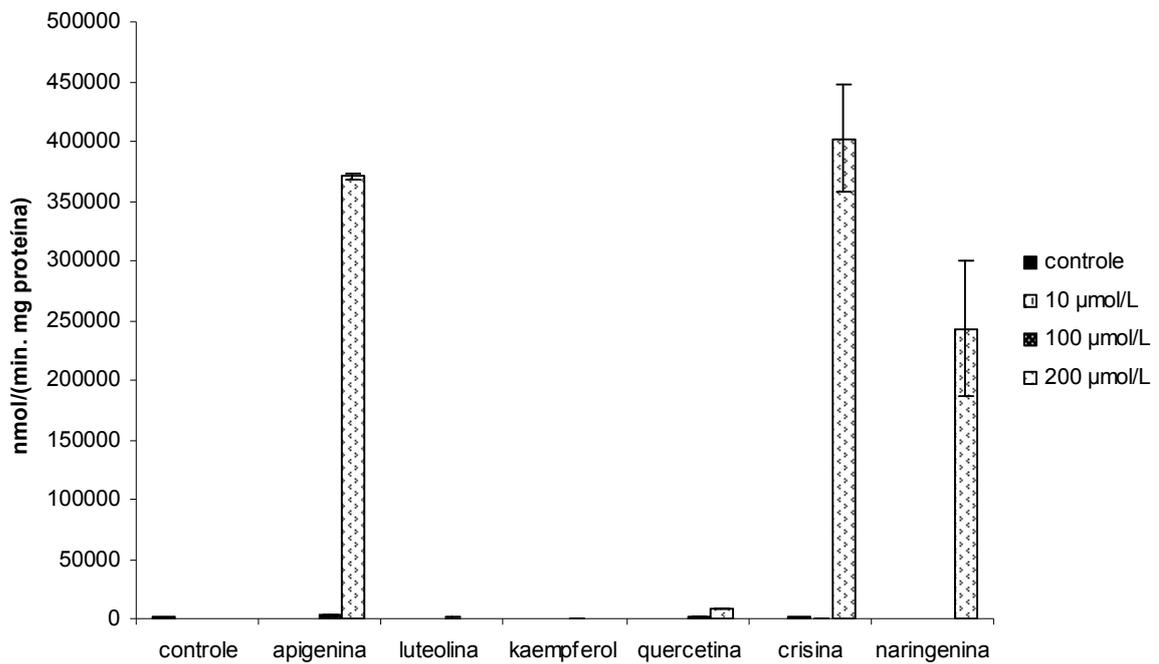


FIGURA 31 - ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE, UTILIZANDO O TRANSCONJUGANTE SmR1 *fdeA::lacZ*, VARIANDO A CONCENTRAÇÃO DE APIGENINA, LUTEOLINA, KAEMPFEROL E CRISINA NO MEIO. A naringenina foi utilizada como controle positivo. Foram realizados 2 ensaios com o transconjugante, que foi deixado crescer por 18 horas, a 30°C e 120 rpm, com os flavonóides no meio de cultura, para que então fosse realizado o ensaio de β -galactosidase. O valor controle mostrado é referente à cultura controle sem a adição de qualquer composto, contudo, também foram realizados ensaios a partir de culturas crescidas com a adição de etanol, metanol e isopropanol (solventes onde se encontram diluídos os flavonóides) e não houve diferença em relação ao controle mostrado aqui.

Dentre os flavonóides testados, os que apresentam atividade indutora são a apigenina e a crisina, indicando que a proteína HsNodD responde a estes flavonóides. Assim, temos que a proteína HsNodD é ativadora da expressão do operon *fde* de *H. seropedicae* SmR1, na presença desses flavonóides, ou seja, HsNodD tem sua atividade aumentada na presença destes compostos, ativando ainda mais a expressão do operon.

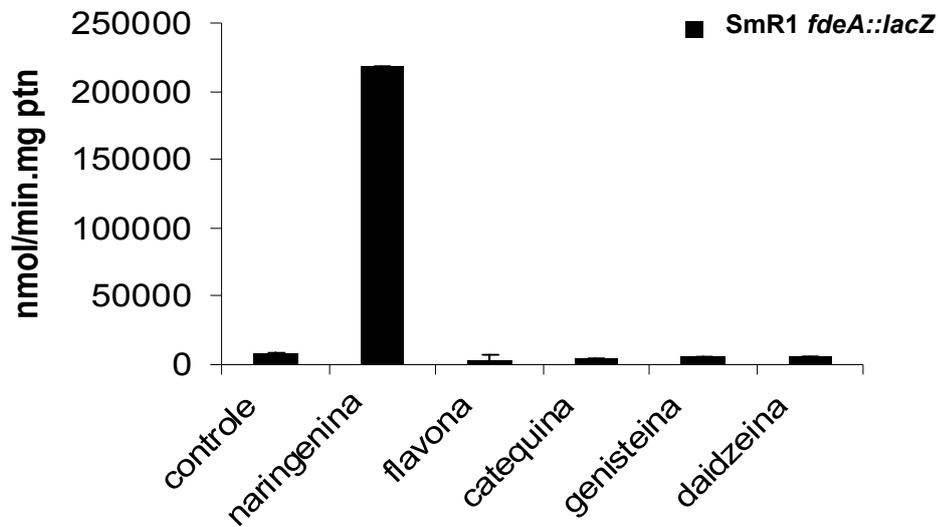


FIGURA 32 - ATIVIDADE ESPECÍFICA DE β -GALACTOSIDASE UTILIZANDO 200 μ mol/L DE FLAVONA, CATEQUINA, GENISTEÍNA E DAIDZEÍNA COMO POSSÍVEIS INDUTORES DA FUSÃO *fdeA::lacZ*. A naringenina foi utilizada como controle positivo. Foi realizado 1 ensaio nestas condições com o transconjugante, que foi deixado crescer por 18 horas, a 30°C e 120 rpm, com os flavonóides no meio de cultura, para que então fosse realizado o ensaio de β -galactosidase. Foi feito um controle com DMSO, uma vez que a flavona e a catequina têm como solventes este composto, porém não houve diferença de indução em relação ao controle (sem flavonóides), e por isso não está mostrado no gráfico.

Como é possível observar na figura 32, nenhum dos flavonóides testados neste experimento mostrou indução sobre a atividade da proteína HsNodD em relação à ativação da fusão *fdeA::lacZ*.

Quando certos flavonóides foram adicionados ao meio de cultura, a ativação da expressão da fusão *fdeA::lacZ* pela proteína HsNodD tornou-se evidente, pois os níveis de expressão aumentaram mais de 200 vezes quando a naringenina era o indutor, e mais de 400 vezes quando a apigenina ou crisina estavam presentes no meio de cultura. Estes resultados indicam que esta proteína é dependente de uma molécula sinal, neste caso um flavonóide, que se ligue a ela e torne-a mais ativa diante da regulação do promotor *fdeA*. Contudo, outros flavonóides como quercetina, kaempferol, luteolina, catequina, daidzeína, genisteína e flavona não foram capazes de aumentar a ativação causada pela proteína HsNodD na fusão pSUI. Isto indica que há alguma mudança na estrutura básica do flavonóide que esta proteína é capaz de reconhecer, como já descrito por Peters & Long, 1988 para *Rhizobium meliloti*, onde

qualquer alteração na estrutura da luteolina leva à diminuição dos níveis de expressão dos genes *nod*.

5.5 EFEITO DA NARINGENINA SOBRE O CRESCIMENTO DE *H. seropedicae*

Alguns trabalhos (REDDY *et al.* 2007; HARTWIG *et al.*, 1991; PETERS & LONG, 1988) com rizóbios mostram que certos flavonóides exercem efeito sobre o crescimento destes microrganismos, e por essa razão foi realizada uma curva de crescimento com as estirpes SmR1, DR2, e AMM1 na ausência e na presença de naringenina. O controle foi feito com e sem etanol (solvente da solução de naringenina estoque utilizada), e com 0,2 mmol/L de naringenina. A DO_{600nm} inicial das culturas foi de 0,1. O tempo de geração calculado está na tabela 10.

TABELA 10 – Cálculo do tempo de geração obtido a partir de uma curva de crescimento das estirpes SmR1, DR2 e AMM1.

Condição	SmR1	DR2	AMM1
NFbHPN-malato	2,12	1,59	1,43
NFbHPN-malato + 0,2 mmol/L naringenina	1,23	0,9	0,9
NFBHPN-malato + etanol	2,07	1,63	1,15

As estirpes mutantes DR2 e AMM1 mostraram um tempo de geração menor do que a estirpe selvagem SmR1 em meio NFbHPN-malato. A presença de etanol não alterou o tempo de geração das estirpes testadas em relação à condição preferencial que é na presença apenas de malato. Contudo, quando a naringenina foi adicionada ao meio de cultura foi possível observar uma diminuição no tempo de geração das 3 estirpes, indicando que este composto interfere no crescimento de *H. seropedicae*, mas que esta alteração não está ligada às mutações, uma vez que a estirpe SmR1 também teve o seu tempo de geração diminuído na presença de naringenina.

A curva de crescimento na presença de naringenina mostrou que este composto atua como um regulador positivo do crescimento em *H. seropedicae*, uma vez que na presença deste composto houve diminuição no tempo de geração da estirpe SmR1 assim como das estirpes mutantes DR2 e AMM1, de maneira semelhante à observada para *R. meliloti* (HARTWIG *et al.*, 1991) e *Bradyrhizobium japonicum* (REDDY *et al.*,

2007), onde a luteolina é o flavonóide que aumenta a taxa de crescimento dos microrganismos citados.

Assim temos que *H. seropedicae* é capaz de crescer na presença de naringenina, e, apesar da naringenina diminuir o tempo de geração, isto foi observado também na estirpe selvagem, indicando que esta diminuição não está ligada às mutações nos genes *nodD* e *fdeA*. Além disto, a naringenina é a molécula que quando presente no meio de cultura torna a proteína HsNodD ativa em relação a expressão do operon *fde*, como foi possível observar nos ensaios de β -galactosidase (figura 28). Logo, a presença de naringenina é sentida pela HsNodD, que por sua vez ativa a transcrição do operon *fde*, cuja a principal possível função atribuída a ele seria a degradação de flavonóides. Esta degradação foi observada nos ensaios em HPLC, onde o perfil cromatográfico da estirpe SmR1 em comparação com as estirpes DR2 e AMM1 após 24 horas de cultivo apresentou diferenças. Com 24 horas de cultivo, a estirpe SmR1 mostrou diminuição da intensidade do pico de naringenina e o surgimento de novos picos que foram eluídos em um tempo anterior à naringenina, enquanto que nos cromatogramas das estirpes DR2 e AMM1 foi observado o pico de naringenina com alta intensidade, e nenhum outro pico anterior ao da naringenina, indicando que estas estirpes não têm a capacidade de metabolizar este composto, e estas mutações estão relacionadas com a degradação de flavonóides, sendo a naringenina, até o momento, o único testado. A mutação na proteína regulatória da via (estirpe DR2) faz com que esta não seja expressa e assim não ocorra a degradação da naringenina. Logo, a naringenina é a molécula sinal para a sua própria degradação, uma vez que HsNodD é sensível a ela e ativa a expressão do operon *fde* que, por resultados de HPLC, se mostrou a via responsável pela degradação da naringenina.

6 CONCLUSÕES

- ✓ A mutação no gene *orf1 (fdeA)* de *H. seropedicae* foi obtida, originando a estirpe mutante AMM1;
- ✓ Resultados obtidos por HPLC mostram que a mutação em AMM1 altera a capacidade de *H. seropedicae* em metabolizar naringenina, que deixa de ser degradada, indicando que esta é a via responsável pela degradação de flavonóides em *H. seropedicae*;
- ✓ HsNodD não autoregula a sua expressão;
- ✓ HsNodD é ativadora da expressão do operon *fde* ;
- ✓ Na presença dos flavonóides naringenina, crisina e apigenina há maior ativação da transcrição do promotor *fdeA* na presença da proteína HsNodD, indicando que estes compostos se ligam a ela;
- ✓ Naringenina não interfere no crescimento de *H. seropedicae*.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ASTURIAS, J.A.; ELTIS, L.D.; PRUCHA, M.; TIMMIS, K.N. Analysis of Three 2,3-Dihydroxybiphenyl 1,2-Dioxygenases Found in *Rhodococcus globerulus* P6. IDENTIFICATION OF A NEW FAMILY OF EXTRADIOL DIOXYGENASES. **The Journal of Biological Chemistry**. v.269, n.10, p.7807-7815, 1994.

ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W., and LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**. 25:3389-3402, 1997.

BALDANI, J.I.; BALDANI, V.L.D.; SAMPAIO, M.J.A.M.; DOBEREINER, J. A fourth *Azospirillum* species from cereal roots. **Anais da Academia Brasileira Ciências**, v. 56, p. 365, 1984.

BALDANI, J.I.; OLIVARES, F.L.; DOBEREINER, J. Selection of *Herbaspirillum* spp. strains associated with rice seedlings amended with ¹⁵N-labeled fertilizer. In: **International Symposium on Sustainable Agriculture for the Tropics – The Role of Biological Nitrogen Fixation**. BOODEY, R.M.; DE RESENDE, A.S. Eds., EMBRAPA, P. 202-203. 1995.

BASSAM, B. J., M. A. DJORDJEVIC, J. W. REDMOND, M. BATLEY, B. G. ROLFE Identification of a *nodD*-dependent locus in the *Rhizobium* strain NGR234 activated by phenolic factors secreted by soybeans and other legumes. **Molecular Plant-Microbe Interaction** v. 1, p. 161–168, 1988.

BENDTSEN, J.D.; NIELSEN, H.; von HEIJNE, G.; BRUNAK, S. Improved prediction of signal peptides: SignalP 3.0. **Journal of Molecular Biology** v. 340, p. 783-795, 2004.

BOCHNER, B.R.; GADZINSKI, P.; PANOMITROS, E. Phenotype MicroArrays for High-Throughput Phenotypic Testing and Assay of Gene Function. **Genome Research**. v.11, p.1246-1255, 2008.

BODDEY, R.M.; OLIVEIRA, O. C.; URQUIAGA, S.; REIS, V.M.; OLIVARES, F.L.; BALDANI, V.L.D.; DOBEREINER, J. Biological nitrogen fixation associated with sugar

cane and rice: contributions and prospects for improvement. **Plant and Soil**. V.174.P. 195-209. 1995.

BRADFORD, M.M, 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. 1976

BRENCIC, A.; WINANS, S.C. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. **Microbiology and Molecular Biology Reviews**. v. 69, p. 155-194, 2005.

CHEN, X.C.; FENG, J.; H, B.H.; LI, F.Q.; LI, Q.; HONG, G.F. Modulating DNA bending affects NodD-mediated transcriptional control in *Rhizobium leguminosarum*. **Nucleic Acids Research**. v. 33, p. 2540-2548, 2005.

CORAL, J. Obtenção de mutante do gene *nodD* de *Herbaspirillum seropedicae*. **Relatório de Iniciação Científica**. 2005.

DÉNARIÉ, J.; DEBELLÉ, F.; ROSENBERG, C. Signaling and host range variation in nodulation. **Annual Review Microbiology** v. 46, p. 497-531, 1992.

DODD, I. B., and J. B. EGAN., Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. **Nucleic Acids Research** v, 18, p. 5019-5026, 1990.

EGGINK, E.; ENGEL, H.; VRIEND, G.; TERPSTRA, P.; WITHOLT, B. Rubredoxin reductase of *Pseudomonas oleovorans*. Structural relationship to other flavoprotein oxidoreductases based on one NAD and two FAD fingerprints. **Journal of Molecular Biology**. v.212 (1), p.135-142, 1990.

FELLAY, R.; HANIN, M.; MONTORZI, G.; FREY, J.; FREIBERG, C.; GOLINOWSKI, W.; STAEHELIN, C.; BROUGHTON, W.J.; JABBOURI, S. *nodD2* of *Rhizobium* sp. NGR234 is involved in the repression of *nodABC* operon. **Molecular Microbiology**. v. 27, p. 1039-1050, 1998.

FENG, J.; LI, Q.; HU, H.L.; CHEN, X.C.; HONG, G.F. Inactivation of the *nod* box distal half-site allows tetrameric NodD to activate *nodA* transcription in an inducer-independent manner. **Nucleic Acids Research**. v. 31, n. 12, p. 3143-3156, 2003.

GOETHALS, K.; VAN MONTAGU, M.; HOLSTERS, M. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. **Proceedings of the National Academy of Sciences** v. 89, p. 1646-1650, 1992.

GOSS, THOMAS J. & BENDER, ROBERT A. The Nitrogen Assimilation Control Protein, NAC, Is a DNA Binding Transcription Activator in *Klebsiella aerogenes*. **Journal of Bacteriology**. v.177, n.12, p.3546-3555, 1995.

HANANAN, D. Studies on transformation of *Escherichia coli* with plasmids. **Journal of Molecular Biology**. v. 166 n. 04, p. 930-937, 1983.

HARTWIG, U.A.; JOSEPH, C.M.; PHILIPS, D.A. Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*. **Plant Physiology**. v.95, p.797-803, 1991.

HARWOOD, C.S & PARALES, R.E. The β -keto adipate pathway and the biology of self-identity. **Annual Review Microbiology**. v.50, p.553-590, 1996.

HENIKOFF, S.; HAUGHN, G.W.; CALVO, J.M.; WALLACE, J.C. A large family of bacterial activator proteins **Proceedings of the National Academy of Sciences**. v. 85, p. 6602-6606, 1988.

HONMA, M. & AUSUBEL, F.M. *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. **Proceedings of the National Academy of Sciences**. v.84, p.8558-8562, 1987.

HU, H.; LIU S.; YANG, Y.; CHANG, W.; HONG, G. In *Rhizobium leguminosarum*, NodD represses its own transcription by competing with RNA polymerase for binding sites. **Nucleic Acids Research**. v. 28, n. 14, p. 2784-2793, 2000.

ISHIYAMA, D.; VUJAKLIJA, D.; DAVIS, J. Novel Pathway of Salicylate Degradation by *Streptomyces* sp. Strain WA46. **Applied and Environmental Microbiology**. v.70, n.3, p.1297-1306, 2004.

ISHIKAWA J.; HOTTA, K. FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. **FEMS Microbiology Letters**. v.174, p.251-253, 1999.

JAMES, E.K.; GYANESHWAR, P.; MATHAN, N.; BARRAQUIO, W.L.; REDDY, P.M.; IANNETTA, P.P.M.; OLIVARES, F.L.; LADHA, J.K. Infection and colonization of rice seedlings by the plant growth-promoting bacterium *Herbaspirillum seropedicae* Z67. **The American Phytopathological Society** v. 15, n. 09, p 894-906, 2002.

KLASSEN, G.; PEDROSA, F.O.; SOUZA, E.M.; FUNAYAMA, S.; RIGO, L.U. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMR1. **Canadian Journal of Microbiology**. v.43, n.9, p.887-891, 1997.

LONG, S.R. *Rhizobium* Symbiosis: Nod Factors in Perspective. **The Plant Cell**. v. 8, p. 1885-1898, 1996.

MACHADO IM, YATES MG, MACHADO HB, SOUZA EM, PEDROSA FO. Cloning and sequencing of the nitrogenase structural genes nifHDK of *Herbaspirillum seropedicae*. **Brazilian Journal of Medical and Biological Research**. v.29, n.12, p. 1599-1602, 1996.

MARIN, A.M. Identificação e caracterização do gene *nodD* de *Herbaspirillum seropedicae*. **Monografia**. 2006a.

MARIN, A.M. Obtenção de estirpe *nodD*- de *Herbaspirillum seropedicae* marcada com a proteína fluorescente dsRED (RAM4). **Relatório de Iniciação Científica**. 2006b.

MARTENS, S. AND MITHÖFER, A. Flavones and flavone synthases. **Phytochemistry**. v. 66, p. 2399-2407, 2005.

MONROE, R.S.; OSTROWSKI, J.; HRYNIEWICZ, M.M.; KREDICH, N.M. *In vitro* interactions of CysB protein with the *cysK* and *cysJIH* promoter regions of *Salmonella typhimurium*. **Journal of Bacteriology**. v.172, n.12, p.6919-6929, 1990.

MORENO, R.; ROJO, F. The Target for the *Pseudomonas putida* Crc Global Regulator in the Benzoate Degradation Pathway Is the BenR Transcriptional Regulator. **Journal of Bacteriology**. v.190, n.5, p.1539-1545, 2008.

MULLIGAN, J.T.; LONG, S.R. A Family of Activator Genes Regulates Expression of *Rhizobium meliloti* Nodulation Genes. **Genetics**. v.122, p.7-18, 1989.

MURAOKA, S.; OKUMURA, R.; OGAWA, N.; TAKAMASA NONAKA, T.; KIYOTAKA MIYASHITA, K.; SENDA, T. Crystal Structure of a Full-length LysR-type Transcriptional Regulator, CbnR: Unusual Combination of Two Subunit Forms and Molecular Bases for Causing and Changing DNA Bend. **Journal of Molecular Biology**. v.328, p.555-566, 2003.

MYLONA, P.; PAWLOWSKI, K.; BISSELING, T. Symbiotic Nitrogen Fixation. **The Plant Cell**. v. 7, p. 869-885, 1995.

NEIDLE, E.L.; HARTNETT, C.; ORNSTON, N. Characterization of *Acinetobacter calcoaceticus* catM, a Repressor Gene Homologous in Sequence to Transcriptional Activator Genes. **Journal of Bacteriology**. v.171, n.10, p.5410-5421, 1989.

OKA, T.; SIMPSON, F.J. Quercetinase, a dioxygenase containing cooper. **Biochemical and Biophysical Research Communications**. v.43, n.1, p.1-5, 1971.

OHLENDORF, D.H.; LIPSCOMB, J.D.; WEBER, P.C. Structure and assembly of protocatechuate 3,4-dioxygenase. **Nature**. v.336, 1988.

OLIVARES, F.L.; JAMES, E.K.; BALDANI, J.I.; DÖBEREINER, J. Infection of mitted strip disease-susceptible and resistant sugar cane varieties by the endophytic diazotroph *Herbaspirillum*. **New Phytologist**. Cambridge, v. 135, p. 723-737, 1997.

OSTROWSKI, J. & KREDICH, N.M. Negative autoregulation of *cysB* in *Salmonella typhimurium*: in vitro interactions of CysB protein with the *cysB* promoter. **Journal of Bacteriology**. v.173, n.7, p.2212-2218, 1991.

PAO, S.S.; PAULSEN, I.T.; SAYER, JR, M.H. Major Facilitator Superfamily. **Microbiology and Molecular Biology Reviews**. v.62, n.1, p.1-34, 1998.

PÉREZ-MARTÍN, J.; ROJO, F.; de LORENZO, V. Promoters Responsive to DNA Bending: a Common Theme in Prokaryotic Gene Expression. **Microbiological Reviews**. v.58, n.2, p.268-290, 1994.

PERRET, X.; STAEHELIN, C.; BROUGHTON, W.J. Molecular basis of symbiotic promiscuity. **Microbiology and Molecular Biology Reviews**. v. 64, p. 180-201, 2000.

PETERS, N.K. & LONG, S.R. Alfalfa Root Exudates and Compounds which Promote or Inhibit Induction of *Rhizobium meliloti* Nodulation Genes. **Plant Physiology**. v.88, p.396-400, 1988.

PILLAI, B.V.S. & SWARUP, S. Elucidation of the flavonoid catabolism pathway in *Pseudomonas putida* PML2 by comparative metabolic profiling. **Applied and Environmental Microbiology**. v. 68, n. 01, p. 143-151, 2002.

PIMENTEL, J. P.; OLIVARES, F.; PITARD, R. M.; URQUIAGA, S.; AKIBA, F.; DOBEREINER, J. Dinitrogen fixation and infection of grass leaves by *Pseudomonas rubrisubalbicans* and *Herbaspirillum seropedicae*. **Plant. Soil** v. 137, p. 61-65, 1991.

RAO, J.R. AND COOPER, J.E. Rhizobia catabolize *nod* gene-inducing flavonoids via C-ring fission mechanism. **Journal of Bacteriology**. v. 176, n. 17, p. 5409-5413, 1994.

RECOURT, K.; van BRUSSEL, A.A.N.; DRIESSEN, A.J.M.; LUGTENBERG, B.J.J. Accumulation of a *nod* gene inducer, the flavonoid naringenin, in the cytoplasmic membrane of *Rhizobium leguminosarum* biovar *viviae* is caused by the pH-dependent hydrophobicity of naringenin. **Journal of Bacteriology**. v. 171, p. 4370-4377, 1989.

REDDY, P.M .; RENDÓN-ANAYA, M.; del RÍO, M.D.S.; KHANDUAL, S. Flavonoids as signaling molecules and regulators of root nodule development. **Dynamic Soil, Dynamic Plant**. 2007.

ROCHE, P.; MAILLET, F.; PLAZANET, C.; DEBELLÉ, F.; FERRO, M.; TRUCHET, G.; PROMÉ, J.C.; DÉNARIÉ, J. The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. **Proceedings of the National Academy of Sciences**. USA. v. 93, p. 15305-15310, 1996.

ROMERO-ARROYO, C.E.; SCHELL, M.A.; GAINES III, G.L.; NEIDLE, E.L. *catM* Encodes a LysR-Type Transcriptional Activator Regulating Catechol Degradation in *Acinetobacter calcoaceticus*. **Journal of Bacteriology**. v.177, n.20, p.5891-5898, 1995.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. **Molecular cloning a laboratory manual**. 2 ed. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press, 1989.

SHELL, M.A. Molecular biology of the LysR family of transcriptional regulators. **Annual Review Microbiology**. v. 47, n. 597-626, 1993.

SCHLAMMAN, H.R.M.; OKKER, R.J.H.; LUGTENBERG, B.J.J. Regulation of nodulation gene expression by NodD in rhizobia. **Journal of Bacteriology**. v. 174, n. 6, p. 5177-5182, 1992a.

SCHWAB, S. Caracterização parcial dos elementos em cis responsáveis pela regulação da expressão do operon *glnAntrBC* de *Herbaspirillum seropedicae*. **Dissertação**. 2002.

SECCON, D.M. Clonagem, superexpressão e purificação da proteína NodD de *Herbaspirillum seropedicae*. **Relatório de Iniciação Científica**. 2004.

SIMON, R.; PRIEFER, U.; PÜHLER A. A Broad Host Range Mobilization System for *In Vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. **BioTechnology**. v.1, p.784-791, 1983.

SUOMINEN, L.; PAULIN, L.; SAANO, A.; SAREN, A.M.; TAS, E.; LINDSTRÖM, K. Identification of nodulation promoter (*nod*-box) regions of *Rhizobium galegae*. **FEMS Microbiology Letters**. v. 177, p. 217-223, 1999.

SWANSON, J.A.; MULLIGAN, J.T.; LONG, S.H. Regulation of *sym* and *nodD3* in *Rhizobium meliloti*. **Genetics**. v.134, p.435-444, 1993.

TROPEL, D. & van der MEER, J.R. Bacterial Transcriptional Regulators for Degradation Pathways of Aromatic Compounds. **Microbiology and Molecular Biology Reviews**. v.68, n.3, p.474-500, 2004.

van RHIJIN, P.J.S.; FEYS, B.; VERRETH, C.; VANDERLEYDEN, J. Multiple copies of *nodD* in *Rhizobium tropici* CIAT899 and BR816. **Journal of Bacteriology**. v.175, n.2, p.438-447, 1993.

van RHIJIN, P.J.S.; DESAIR, J.; VLASSAK, K.; VANDERLEYDEN, J. The NodD Proteins of *Rhizobium* sp. Strain BR816 Differ in Their Interactions with Coinducers and in Their Activities for Nodulation of Different Host Plants. **Applied and Environmental Microbiology**. v.60, n.10, p.3615-3626, 1994.

von MERING, C.; JENSEN, L.J.; KUHN, M.; CHAFFRON, S.; DOERKS, T.; KRÜGER, B.; SNEL, B.; BORK, P. STRING 7--recent developments in the integration and prediction of protein interactions. **Nucleic Acids Research**. v.35, D.358-352, 2007.

WASSEM, R.; PEDROSA, F.O.; YATES, M.G.; REGO, F.G.; CHUBATSU, L.S.; RIGO, L.U.; SOUZA, E.M. Control of autogenous activation of *Herbaspirillum seropedicae* *nifA* promoter by the IHF protein. **FEMS Microbiology Letters**. V. 212, n. 02, p. 177-182, 2002.

WEBER, O.B.; BALDANI, V.L.D.; TEIXEIRA, K.R.S; KIRCHHOF, G.; BALDANI, J.I.; DÖBEREINER, J. Isolation and characterization of diazotrophic bacteria in banana and pineapple plants. **Plant Soil**. V. 210, p. 103-113, 1999.

WILLIAMS CA, HARBORNE JB. IN: HARBORNE JB, editor. The Flavonoids. Advances in research since 1986. London7 Chapman & Hall; 1994. p. 337- 85, 1986.

ANEXO

Seqüência de nucleotídeos e de aminoácidos das *orfs* encontradas no operon *fde* de *H. seropedicae*

nodD

		← <i>fdeA</i>		
1	GTCTGTCTCCTTGTCTCCTACCCTGCCTGGCGGGGTGTGGCCGCTATCGGCGCGTTGTC		59	
60	ATGGCCAGGCTCAACAGCGGCATCCAGTGCGAGTGCGCCGCCGCTTGCCGCGCCAG		118	
119	CCCTTTCTGGTGGGGCCATGCAGCGGATGAGTGAAGCTTAAGAGTGAGC	-35 <i>pnodD</i>	177	
		-10 <i>pnodD</i> -10 <i>pfdeA</i> -35 <i>pfdeA</i>		
178	TCAAATGGATTGT	TTGATTCATGATA	236	
237	GTCGTCTATCCGAGGGTCGGCAAGAACAAGCACAACAACGGAGACCAGCGCC	ATG CGT	295	
1		M R	2	
296	TTC AAC AAG CTC GAC CTC AAT CTT CTG GTC GCC CTG GAT GCA CTG		340	
3	F N K L D L N L L V A L D A L		15	
316	CTC ACG GAG ATG AGC ATC AGC CGC GCC GCC GAA AAG ATC CAT CTG		360	
16	L T E M S I S R A A E K I H L		30	
361	AGC CAG TCG GCC ATG AGC AAT GCC CTG GCG CGG CTG CGC GAG TAT		405	
31	S Q S A M S N A L A R L R E Y		45	
406	TTC GAT GAT GAA TTG CTG ATC CAG GTG GGC CGG CGC ATG GAG CCC		450	
46	F D D E L L I Q V G R R M E P		60	
451	ACG CCG CGC GCC GAG GTG CTC AAG GAT GCG GTG CAT GAT GTG CTG		495	
61	T P R A E V L K D A V H D V L		75	
496	CGG CGT ATC GAT GGC TCC ATC GCG GCG CTG CCG GCC TTC GTG CCG		550	
76	R R I D G S I A A L P A F V P		90	
551	GCC GAG TCC ACG CGC GAG TTT CGC ATC TCG GTT TCG GAC TTT ACG		595	
91	A E S T R E F R I S V S D F T		105	
596	CTC TCC GTC CTC ATC CCC CGG GTG CTG GCG CGC GCG CAC GCC GAG		640	
106	L S V L I P R V L A R A H A E		120	
641	GGC AAG CAC ATC CGC TTT GCC CTG ATG CCG CAG GTG CAA GAC CCG		685	
121	G K H I R F A L M P Q V Q D P		135	
686	ACC CGC TCG CTG GAT CGG GCC GAG GTG GAC CTG CTG GTC TTG CCG		730	
136	T R S L D R A E V D L L V L P		150	
731	CAG GAA TTC TGC ACG CCC GAT CAT CCT GCC GAA GAG GTC TTC CGC		775	
151	Q E F C T P D H P A E E V F R		165	
776	GAA CGG CAT GTC TGC GTG GTC TGG CGC GAC AGT GCG CTG GCG CAA		820	
166	E R H V C V V W R D S A L A Q		180	
821	GGC GAG CTG ACG CTG GAA CGC TAC ATG GCC TCA GGC CAT GTG GTG		865	
181	G E L T L E R Y M A S G H V V		195	
866	ATG GTG CCG CCT GGG GCC AAT GCG TCG TCG GTG GAG GCG TGG ATG		930	
196	M V P P G A N A S S V E A W M		310	

745	GAT	ATC	TCT	TTG	GAG	ACC	GGC	TCC	TAC	AGC	CCG	ACC	CGG	CCC	AGC	789
153	D	I	S	L	E	T	G	S	Y	S	P	T	R	P	S	167
790	GTG	GCG	GTC	GGC	TAC	ACC	TCC	TGG	CAG	CCG	GTG	CTG	GCC	GTG	CGT	834
168	V	A	V	G	Y	T	S	W	Q	P	V	L	A	V	R	182
835	CAC	AAC	AAC	CCC	GAC	GGA	CTC	GAC	CTG	GGC	ATC	TCC	AAC	CGC	CTG	879
183	H	N	N	P	D	G	L	D	L	G	I	S	N	R	L	197
880	CTC	ATC	AAC	GAC	CGC	AAC	AGC	AGC	ACC	AAC	TAT	CGT	TCC	GGC	ACG	924
198	L	I	N	D	R	N	S	S	T	N	Y	R	S	G	T	212
925	GCC	TAT	GTG	GCC	GAT	TTC	ATC	GGC	GGC	TGG	AAC	GTC	GGC	AAA	TGG	969
213	A	Y	V	A	D	F	I	G	G	W	N	V	G	K	W	257
970	AAG	TTT	GGC	GTA	ACC	GGT	TCT	TAC	CTT	AAC	CAG	TTC	ACC	GAT	GAC	1014
258	K	F	G	V	T	G	S	Y	L	N	Q	F	T	D	D	272
1015	CGG	CAG	AAC	GGC	GCC	GAC	ATC	ACC	GGC	AAC	CGC	GCC	CGC	AGC	CTG	1059
273	R	Q	N	G	A	D	I	T	G	N	R	A	R	S	L	287
1060	TCC	CTG	GGA	CCG	ACC	GTG	GCC	TAC	AAC	GCC	GGT	CCC	TTC	AGC	ATC	1104
288	S	L	G	P	T	V	A	Y	N	A	G	P	F	S	I	302
1105	AAC	ATG	AAT	TAC	CAG	CGC	GGC	CTG	TAC	GCC	GCC	AAT	ACC	GCC	AAG	1149
303	N	M	N	Y	Q	R	G	L	Y	A	A	N	T	A	K	317
1150	AGC	GAC	GCC	ATC	TGG	ATC	AAC	TTC	GCC	ATG	CCG	CTG	TGG	ATG	GGC	1194
318	S	D	A	I	W	I	N	F	A	M	P	L	W	M	G	332
1195	GGG	GGC	CAT	TGA	GCAAGCGCTCGTCATCATAGGAGTCACACCATG										1239	
333	G	G	H	*											335	

fdeB

1	GTCACACC	ATG	TTC	CGT	TAC	TTC	CCG	ACG	AAC	TAT	GTC	TGG	AAT	CTT	47	
1		M	F	R	Y	F	P	T	N	Y	V	W	N	L	13	
48	TCC	GTT	GAC	CTC	GCC	ATC	GAG	ATG	GGC	GCG	CGC	ATC	GGT	GAA	ATC	92
14	S	V	D	L	A	I	E	M	G	A	R	I	G	E	I	28
93	GAA	GCC	ATG	TGC	GCG	CCC	TTG	CAG	GAG	GCC	GCC	AGG	CAG	CCT	GAC	137
29	E	A	M	C	A	P	L	Q	E	A	A	R	Q	P	D	43
138	GCC	GAA	GGC	ACG	GCG	GCC	TTT	CGC	GCC	ACC	TGG	GCG	GAC	ATG	GCT	182
44	A	E	G	T	A	A	F	R	A	T	W	A	D	M	A	58
183	GAA	AAG	CTG	TGC	AGC	CTG	GCC	GCC	GAG	GAT	GAG	ATG	CGC	GGC	CGC	227
59	E	K	L	C	S	L	A	A	E	D	E	M	R	G	R	73
228	CTG	CTG	TCG	GCG	GGC	GAG	AAA	TAC	AAG	CGC	GCC	GCC	AGC	TAC	CTG	272
74	L	L	S	A	G	E	K	Y	K	R	A	A	S	Y	L	88
273	CTG	ACC	TGC	GAA	CGC	CTG	CAG	GGC	CAT	GGC	GCA	CCG	GGC	CGT	CTG	317
89	L	T	C	E	R	L	Q	G	H	G	A	P	G	R	L	103
318	GAA	CTC	TAC	AAG	CGC	TTC	CTG	GAA	GTG	TTC	CAG	CGC	GGT	ATC	GCC	362
104	E	L	Y	K	R	F	L	E	V	F	Q	R	G	I	A	118
363	CTG	GCG	GGC	GAG	AAC	TGC	GAA	CGG	GTC	GAG	ATT	GCC	TAC	GAA	GGC	407
119	L	A	G	E	N	C	E	R	V	E	I	A	Y	E	G	133

408	CGG GTC ATC TCG GGT CTC TAT ACC CGG GCC AGG AAC GTG CAG GGG	452
134	R V I S G L Y T R A R N V Q G	148
453	CCG GCC CCG GTC CTG GTG CAG CTC AAT GGA CTG GAC TCG ACC AAA	497
149	P A P V L V Q L N G L D S T K	163
498	GAG ATG AAA TAC CTG GTA GGC CTG CCC GGC TGG CTG GCC GAG CGC	542
164	E M K Y L V G L P G W L A E R	178
543	GGT GTC GCG TCG CTG GTG ATC GAC CAG CCT GGC ACC GGC GAA GCC	587
179	G V A S L V I D Q P G T G E A	193
588	CTG CGC CTG CAC GGC TTG ACG GCA CGC TTC GAT GCC GAG CAC TGG	632
194	L R L H G L T A R F D A E H W	208
633	GCG CGT CAT GTG GTG GAC TGG CTG GAG CAG CGC GAG GAC GTC GAT	677
209	A R H V V D W L E Q R E D V D	223
678	CCC ACC CGC ATC GGC TGT GAA GGG GTC TCG CTG GGC GGG TAC TAC	722
224	P T R I G C E G V S L G G Y Y	238
723	TGC CCC CGT GTG GTG GCC ATG GAG CCG CGC TTT GCC TGC GGC GTG	767
239	C P R V V A M E P R F A C G V	253
768	GTG TGG GGC GCC AAT CAC GAT TGG CGC GAT GTG CAG AAA CGG CGT	812
254	V W G A N H D W R D V Q K R R	268
813	CTG GAG AAG GAG GGC GAT TTC CCT GTG CCG CAT TAC TGG CAG CAC	857
269	L E K E G D F P V P H Y W Q H	283
858	GTC TGC TGG GTC TGG GGC GCC AAG GAT ATC GAC GAC TTC ATG CGC	902
284	V C W V W G A K D I D D F M R	298
903	ATC GCC GAG GAC GTC CAC CTG GAC GGC GTG GTC GAG AAG ATC CGC	947
299	I A E D V H L D G V V E K I R	313
948	GTG CCT TTC CTG GTC ACC CAT GGC GAG AAG GAC AGC CAG ATT CCG	992
314	V P F L V T H G E K D S Q I P	328
993	CTC AAG TGG GCG CAT CGC ACC TAC GAA CAA CTG GTC AAC AGC CCC	1037
329	L K W A H R T Y E Q L V N S P	343
1038	AAG CGC GAA TTG AAG GTC TTC ACC GAC CGC GAA GGC GGC GTG CAG	1082
344	K R E L K V F T D R E G G V Q	358
1083	CAC GCC AGC TTC GAC AAC TCC ATC AAC GCC GGC CAC TAC ATC GCC	1127
359	H A S F D N S I N A G H Y I A	373
1128	GAC TGG GTC GCC GAA ACC CTG GGC GGC CGC ACC AAG CGC TGA ACA	1172
374	D W V A E T L G G R T K R *	387
1173	TCGACTGCCAACCTAAGGAATACGCATGAA	1202

fdeC

1	CTGAACATCGACTGCCAACCTAAGGAATACGC	53
1	<u>ATG</u> AAC ATC ATT GGA CCT GAC	7
	M N I I G P D	
54	GCC CTG GTA TTT GGC GTG GAT GAT CTG CCC GCC TGC CGC CAG TAC	99
8	A L V F G V D D L P A C R Q Y	22

100	CTG CTG GAC TAT GGC CTC AAG GAC GCC GGC GGC GAC CGC TTC GAA	144
23	L L D Y G L K D A G G D R F E	37
145	GCC CTG GAC GGC ACA GCC GTC GTG CTG CGC GCC AAG GAC GAC GCC	189
38	A L D G T A V V L R A K D D A	52
190	AGT CTT CCG CCA GCC ATG GGC ACT GCC AGC CTG CTG CGC GAG ACC	234
53	S L P P A M G T A S L L R E T	67
235	GTC TAT GGC GTG GCC GAC AGC GCC ACG CTG GAC GCC ATC GAG ACC	279
68	V Y G V A D S A T L D A I E T	82
280	GAA CTG CGC CGC GAC CGC GAA GTC AGC CGC CGC GAT GGC GTC GTA	324
83	E L R R D R E V S R R D G V V	97
316	CGC AGC GTC GAT GAC ATG GGC TTT GCG CTG GCC TTC CAG CTC ACC	369
98	R S V D D M G F A L A F Q L T	112
370	GTG CGC CGC CCC CTG ACG CTG CCC GCT GAA CGC GTC AAT GGC CCT	414
113	V R R P L T L P A E R V N G P	127
415	GGC CAC GCC CAG CGC GCG CCC AAC GAA CTG GGT CTG CCG CCG GAA	459
128	G H A Q R A P N E L G L P P E	142
460	TTG CCG GCC TTG CCG CGC ACG CTC TCG CAT GTG GTG TAT TTC GTG	504
143	L P A L P R T L S H V V Y F V	157
505	CCG GAC GCG GCC AAG GCC GAG GCC TTC TAT CAT CGC CTG GGC TTC	585
158	P D A A K A E A F Y H R L G F	172
586	GTC TGC ACC GAC CGC TTC ACC GGC GTC GGC CCC TTC CTG CGC CCT	629
173	V C T D R F T G V G P F L R P	187
630	GCA GGC ACC CTG GAC CAT CAC ACC CTG TTC ATG ATC CAG ACC CCG	674
188	A G T L D H H T L F M I Q T P	202
675	CCC TTC ATG AAA GGC TGC GAG CAT TTC ACC TTC CAT ATG GGC GGT	719
203	P F M K G C E H F T F H M G G	217
720	CCG ACC GAA CTG CTG CTG GCC GGC ACC CGC TTC GTG GAA AAG GGC	764
218	P T E L L L A G T R F V E K G	232
765	TAC CAG AGC TTC TGG GGA CCC GGC CGC CAT CGT TTC GGC AGC AAC	809
233	Y Q S F W G P G R H R F G S N	247
810	TGG TTC TGG TAT TTC AAT TCC CCG CTG GGC TGC CAC GTC GAG TAC	854
248	W F W Y F N S P L G C H V E Y	262
855	GAC GCC GAC ATG GAC CTG CAC GAC CAA AGC TGG GCC GCG CGC GAA	899
263	D A D M D L H D Q S W A A R E	277
900	GTT CCC ATG GGC GCC GAT GCC TCG CAG CTG TTC CTG TTC CAG TAC	944
278	V P M G A D A S Q L F L F Q Y	292
945	CGC GAG AAG TGG GCG CCG AGC GGC CCG CCG CCG CCC GGC GCG GCG	989
293	R E K W A P S G P P P P G A A	307
990	CAC TGA GTACGAGGACGAATACAAAGACGAATACGACAGCGGAGACAGCCATGT	1043
308	H *	309

fdeD

1	CTGAGTACGAGGACGAATACAAAGACGAATACGACAGCGGAGACAGCC	<u>ATG</u>	TAC	CTG	57
1		M	Y	L	3
58	TGC CAC ATC ACG CAA CTG CCG GAG AAC GGC GCA CGC GGT TTC GAT				102
4	C H I T Q L P E N G A R G F D				18
103	ACC GAA TCG GCA GGG CAG GCC ACC ATC TTC GTC TTG CGC CGG GGC				147
19	T E S A G Q A T I F V L R R G				33
148	GAC CAG GTC CGT GCC TGG CGC GAC AGC TGT CCG CAC CAC GGC ACG				192
34	D Q V R A W R D S C P H H G T				48
193	CCG CTG CCG TGG CGG CGC GAC GCC TAT TTC GAC GGC ACC GGC CAG				237
49	P L P W R R D A Y F D G T G Q				63
238	CAC CTG GTC TGT GCC GCC CAT GGC GCG CTG TTC GAT CCC CTC ACA				282
64	H L V C A A H G A L F D P L T				78
283	GGC GTC TGT ACG CTC GGT CCC TGC CAG GGC GAC ACG CTC ACG GCA				327
79	G V C T L G P C Q G D T L T A				93
328	GTA CCG CTA CGC ATC GAT GAC GAC GGC GGC CTG CAT GTG GAC GCC				372
94	V P L R I D D D G G L H V D A				108
373	GGC GAT CCG CTT TAA CTCTTTCAAGAAAAACATCCAGGAGACTCACCATGA				422
109	G D P L *				112

fdeE

1	TTAACTCTTTCAAGAAAAACATCCAGGAGACTCACC	<u>ATG</u>	ACA	CTC	GCA	GCC	CAA	54
1		M	T	L	A	A	Q	6
55	CGC ATC CTC ATC ATC GGC GGC GGT TTT TCC GGC ATG TCG GCC GCC							99
7	R I L I I G G G F S G M S A A							21
100	ATC GAC CTG CGC CGC CGC GGC GCC CAG GTC GAT CTG GTG GAG CTC							144
22	I D L R R R G A Q V D L V E L							36
145	GAT GCC CAG TGG CGC AAC TAT GGC GCT GGC ATC AGC CTG GGG CCG							189
37	D A Q W R N Y G A G I S L G P							51
190	GCC ACC CTG CGC GCC TTG AAG CAG CTG GGT GTG CTG GAG GCC TTC							234
52	A T L R A L K Q L G V L E A F							66
235	CTG CGC GAA GGG GCG GCC GGC GAT GGT GTG CGC CTG TGC CTG CCG							279
67	L R E G A A G D G V R L C L P							81
271	CAT GGT CCG CAA GTG GCC GAA CTG CCT ACA CCG CGC CTG GCC AGT							324
91	H G P Q V A E L P T P R L A S							96
325	CCC GAT GTG CCG GGC GGC GGC GCC ATC CTG CGT CCG GTG CTG GCG							369
97	P D V P G G G A I L R P V L A							111
370	CGC ATC CTG GCC GAC GCC ACC CGC GCG GCC GGC GCC GAT GTA CGC							414
112	R I L A D A T R A A G A D V R							126
415	CTG GGT TGC ACC TTC AGC GCA GTG CGC GAA GTC GGG CAC AAG GTC							459
127	L G C T F S A V R E V G H K V							141

460 GAA GTG GAC TTC ACC GAT GGC CAG ACC CGC CGC TAC GAC CTG GTG 504
 142 E V D F T D G Q T R R Y D L V 156
 505 ATC GGC GCC GAT GGT CTC TAT TCC AAG CTG CGC ATG CAT CTC TTC 549
 157 I G A D G L Y S K L R M H L F 171
 550 CCC CAT GCC CCC AAG CCG CGC TAC AGC GGT CAG GCG GTG TGG CGC 594
 172 P H A P K P R Y S G Q A V W R 186
 595 GCC GTA CTG CCG CGT CCG CAA GAG ATC GTC ACC TGC ACC ATG TGG 639
 187 A V L P R P Q E I V T C T M W 201
 640 ATG GGG CCG CGC ATC AAG CCT GGC GTC AAT CCG GTG TCC AAG GAT 684
 202 M G P R I K P G V N P V S K D 216
 685 GAG ATG TAC CTG TTT GTC ACC GAA CCG CGC CCG GTC AAC GAG CAC 729
 217 E M Y L F V T E P R P V N E H 231
 730 GTC GAT CCG ACC ACC TTT GTC TCG CAC CTG CGC GGT TTG CTG GAA 774
 241 V D P T T F V S H L R G L L E 246
 775 GAG TTC AGC GCC CCC GTG CTC AAG ACC ATC CGC GAG CAG CTC GAT 819
 247 E F S A P V L K T I R E Q L D 261
 820 GAC AAT GCA CGC ATC ATC TTC CGT CCG CTG GAA GGC CTG CTG TTG 864
 262 D N A R I I F R P L E G L L L 276
 865 CCA CGT CCG TGG TAC CAG GGA CGC GTG GTG CTG ATC GGC GAC GCC 909
 277 P R P W Y Q G R V V L I G D A 291
 910 GTC CAT GCC ACC ACG CCA CAC CTG GCC TCG GGC GCT TGC ATC GGC 954
 292 V H A T T P H L A S G A C I G 306
 955 ATC GAG GAT GCG CTG GTG CTG GCC GAT GAG CTG GAA CGC CAC GCC 999
 307 I E D A L V L A D E L E R H G 321
 1000 ACG GTG CCG CAA GCC CTG GCG GCC TTC GAG GAG CGG CGC TGG GAG 1044
 322 T V P Q A L A A F E E R R W E 336
 1045 CGC TGC CGC ATG GTG GTG GAA AAC TCC GCG CGC CTG GGT GAG ATC 1089
 337 R C R M V V E N S A R L G E I 351
 1090 GAG GTT GAA GGC GGC GAC AAG GAC GAG CAT TCG CGC ATC ATG CGC 1134
 352 E V E G G D K D E H S R I M R 366
 1126 GCC TCG CAT GCG GCG CTG GTG CAA CCC ATC TGA ATGCGCAATGCAGAA 1179
 367 A S H A A L V Q P I * 376
 1180 GAGCACGTAAAAATATAAATCATTAACAGTTCTTGACCTGCGCACAACGCGAGGAGAC 1238
 1239 GACATGATCACGAGCTATTCCAAGCGCGCCTGGCTGCGCTGATGC 1284

fdeF

1 TCTGAATGCGCAATGCAGAAGAGCACGTAAAAATATAAATCATTAACAGTTCTTGACC 59
 60 TGCGCACAACGCGAGGAGACGACATGATCACGAGCTATTCCAAGCGCGCCTGGCTGCG 118
 119 CTG ATG CTG GCG CAT TGC GCC GGC ATG ATT GAC CTG GTG GCC CTG 163
 1 M L A H C A G M I D L V A L 14
 164 CCG GTG TGG GTC GGC ACG CTG GTG GCG TAT TAC CAT TTC GAT GCG 208
 15 P V W V G T L V A Y Y H F D A 29

209	CGC CAG GCC GGC CTG CTG GCC TCG CTC TTT CTC GGT GGC GCC GTG	253
30	R Q A G L L A S L F L G G A V	44
254	CTG GCC AGC GCC ATG CTG GCG CCG CGC TTC GAT CCC GAG ACG GCG	298
45	L A S A M L A P R F D P E T A	59
299	CGC TGG AAG GCC GCC ATC GGC TTC GCC CTG TCG GCG CTC TTG TTC	343
60	R W K A A I G F A L S A L L F	74
344	ACC GTG GCG GCG CAG ACC CAT CAC TAC CTC CTG CTG GCC CTG TTG	388
75	T V A A Q T H H Y L L L A L L	89
389	CAT GCC GTC TGC GGC ATC ACG ACC GGG CTG TCC TTG AGC GTC ACG	433
90	H A V C G I T T G L S L S V T	104
434	CAT GGC ACC ATC GCC CGT GGC AGC AAT CCG CAT CGC GGC TTT GCC	478
105	H G T I A R G S N P H R G F A	119
479	CTG GTC GGG ATC GCG CTG GGC GTC TTC GCC ATC CTG TTT CTC GGC	523
120	L V G I A L G V F A I L F L G	134
524	AGT GCG CCG CCG CTG GTC GCT CAC CTT GGT GGC AGT ACG CTG TTC	568
135	S A P P L V A H L G G S T L F	149
569	TAT CTC TTT GCC GGC GTG ATG GCG GTG GGT GCG CTG GCC AGC GCC	613
150	Y L F A G V M A V G A L A S A	164
614	TTG GCC TTT CCG CGT GCC GAG ACC GAT GGC CTG CCG CGC GTG TCC	658
165	L A F P R A E T D G L P R V S	179
659	ACC GTG CGT CCA CCG GTG CCG CGT CAG GTC TGG TAC GGC ATC GCC	703
180	T V R P P V P R Q V W Y G I A	194
704	GGG ATC GCC TGC ATG GGG CTG GTG CAG TCG ATG ACC TTC AGC TTC	748
195	G I A C M G L V Q S M T F S F	209
749	CTC GAA CGC GTC GGC AGT GAT CGT GGC TTC GGC CTG CAA GCC GTC	793
210	L E R V G S D R G F G L Q A V	224
794	ACC GGC GTG CTC ATC GCC CTG GGT TTC GTC AAT CTC CTG CCC GCT	838
225	T G V L I A L G F V N L L P A	239
839	CCA CTG GCG GCT CTG CTG GAA AAG CGC CTG CCG GCC CGC AAC GTG	883
240	P L A A L L E K R L P A R N V	254
884	CTC ATG GCC GGC CCG CTG GTG CAG GCG CTG CTG GTG CTG GTG ATC	928
255	L M A G P L V Q A L L V L V I	269
929	ATG TCC TCG GCG CGC TTC GCA CCC TAT GCC GCC GCC GCG TCG GTA	973
270	M S S A R F A P Y A A A A S V	284
974	TTT GCC GCG GTG ATG ATC TTC ACG CAT ACC TTT GCC TTC GGT CTC	1018
285	F A A V M I F T H T F A F G L	299
1019	ATG GCC CGG CTC GAC CGC AGT GGC CGC GCC ATG GCC GCC ACG CCC	1063
300	M A R L D R S G R A M A A T P	314
1064	GCC ATG CTG ATG ACC GGC GCG GCC ATC GGC CCC ATT CTC GGC GGC	1108
315	A M L M T G A A I G P I L G G	329
1109	ACG CTG GTA CAG GGC TGG GGC TAT GGC AGC CTG GGA CCA GTG GCG	1153
330	T L V Q G W G Y G S L G P V A	344

1154 CTG GCG CTG GGC CTG ACG GCA TCG TTC TGT TTC TCG CGG CTG CCG 1198
 345 L A L G L T A S F C F S R L P 359

1199 CGC GCC GCC GCC GTT CAA CCT CAG GAG GCC CTG GCA TGA CCGCATC 1244
 360 R A A A V Q P Q E A L A * 371

1245 CTTGCAAGCACGTCGCCGGCATCGCTTCGTGCGACTTGTCCATCTATCTCGAAAACGACG 1303
 1304 TGCTCTCCGATCCGCGGCCACTGGCGCCCAAGATCACCTACCAACAGCACGCCGACACC 1362
 1363 TTGCCCGAATTCATGGCCATGCTGCCCGGCACCCGGCCCCGAAGACTATCCCGACGGCGA 1421
 1422 AGCCGCCCGCCGCTGAATGGGTACCCCTGACCA 1453

fdeG

1 CATGACCGCATCCTTGAAGCACGTCGCCGGCATCGCTTCGTGCGACTTGTCCATCTATCT 59
 118 CGAAAACGACGTGCTCTCCGATCCGCCGCCACTGGCGCCCAAGATCACCTACCAACAGCA 177
 178 CGCCGACACCTTGCCCGAATTCATGGCCATGCTGCCCGGCACCCGGCCCCGAAGACTATCC 236
 237 CGACGGCGAAGCCGCCCGCTGAATGGGTACCCCTGACC ACG CAC AAC GGC ACC 291
 1 T T H N G T 6

292 CAT CTC GAT GCG CCC TGG CAT TTC CAT TCC ACC CAG GAC GCC AAG 336
 7 H L D A P W H F H S T Q D A K 21

337 AAC GGC GGC GCA CGC CCC TCC ATC ACC ATC GAT GAA GTC CCG CTG 381
 22 N G G A R P S I T I D E V P L 36

382 GAG TGG TGC TTC CAG CCC GGC GTG AAA CTG GAC TTC CGC CAC TTC 426
 37 E W C F Q P G V K L D F R H F 51

427 CCC GAT GGC TAC GTC GCC ACT GCC GCC GAC GTG GAA GCC GAA TTG 471
 52 P D G Y V A T A A D V E A E L 66

472 AAG CGC ATC GGC CAT GTA CTG GAG CCA CTC GAT ATC GTC GTG GTC 516
 67 K R I G H V L E P L D I V V V 81

517 AAC ACC CGC GCC GGC TCC CGC TAT GGT CAC CGC GAC TAT CTC GGC 561
 82 N T R A G S R Y G H R D Y L G 96

562 GCC GGT TGC GGC ATG GGT TAT GAA GCG ACC ATG TAC CTG CTC GAA 606
 97 A G C G M G Y E A T M Y L L E 111

607 CGC GGC GTA CGC CTG ACC GGC ACC GAC GCC TGG AGC TGG GAT GCA 651
 112 R G V R L T G T D A W S W D A 126

652 CCC TTT TCC TAT ACC GCC CAG CGT GTG GCC GAG ACC GGC AAC AAG 696
 127 P F S Y T A Q R V A E T G N K 141

697 GCT CTG ATC TGG GAA GGC CAC AAG GCC GGC CGC GAC ATC GGC TAC 741
 142 A L I W E G H K A G R D I G Y 156

742 TGC CAT CTG GAA AAA CTG CAC AAC CTC GAA GCC CTG CCA GCG CAT 786
 157 C H L E K L H N L E A L P A H 171

787 GGC TTC ATG ATC AGC TGC TTC CCG CAC AAG ATC CGC GGC GCC AGC 831
 172 G F M I S C F P H K I R G A S 186

832 GCC GGA TGG ACG CGC GCC GTG GCG ATC TTC GAC GCC GCC TGA ACC 876
 187 A G W T R A V A I F D A A * 199

877 TGCTTGCCACTGGAGTCATCATGAGCT 903

238	GCC ATC GCA TTC GAC CCG CGC CAA TGC ATC GCC CCG CTG CCG CGC	282
764	A I A F D P R Q C I A P L P R	78
283	AGC CCG CAA TGG CTG GAT GCC TCG GCC TTC CTC AAT CAT GGG CGG	327
79	S P Q W L D A S A F L N H G R	93
328	CTG ATG GAA GAG GCG TTC AAG ACG CCA CCG ATC CCG CAT TTC GAC	372
94	L M E E A F K T P P I P H F D	108
373	ACC GTG CCC GTG ATG TAC CAG GGC GCC AGC GAT GAT TTC CTC GGC	417
109	T V P V M Y Q G A S D D F L G	123
418	CCC TAT CAG GAC GTG GCG CTG CCC TCG GAA GAA GAT GGC ATC GAC	462
124	P Y Q D V A L P S E E D G I D	138
463	TTC GAA GGC GAA TTC GGT GTG ATC GTC GGC CCC GTC CCC ATG GGC	507
139	F E G E F G V I V G P V P M G	153
508	GTG CAG GCC AGC CAA GCC TTG CAG GCA GTG CGC CTG CTG GTG CAG	552
154	V Q A S Q A L Q A V R L L V Q	168
541	ATC AAT GAC TGG AGC CTG CGC GCC CTG GGA CCG CAT GAG ATG AAG	597
181	I N D W S L R A L G P H E M K	183
598	ACT GGC TTC GGC TTC CTG CAG GCC AAG CCT TCC ACC GCC TTC GCT	642
184	T G F G F L Q A K P S T A F A	198
643	CCC ATG GCC GTC ACC CCC GAC GAG CTC GGC CCC GCA TGG CGC GAC	687
199	P M A V T P D E L G P A W R D	213
688	GGC CGC GTG CAG ATG CGC CTG CAG GTA CAG TGG AAC GGC CAG CCG	732
214	G R V Q M R L Q V Q W N G Q P	228
733	TTT GGC CAT CCG CAC GGC GGC GAG ATG AAT TTT TCC TTT GGC GAA	777
229	F G H P H G G E M N F S F G E	243
778	CTG ATC GCC CAC GCC GCC CGC AGT CGC CGG CTC ACC GCT GGC ACC	822
244	L I A H A A R S R R L T A G T	258
823	GTG ATC GGT TCG GGC ACG GTG TCG AAC CAG TCG CGT GCG GCA GGC	867
259	V I G S G T V S N Q S R A A G	273
868	TCG GCC TGC ATC GCC GAA CGC CGC GTC ATC GAA AAG ATC GAC CAC	912
274	S A C I A E R R V I E K I D H	288
913	GGC GAC ATC CGC ACC GGC TTC ATG CGC TTT GGC GAC GAA GTC AGC	957
289	G D I R T G F M R F G D E V S	303
958	ATG CAG GCC TGC TTT GAC GAT GGC CGC AGC GGG CCC TTC GGC CTG	1002
304	M Q A C F D D G R S G P F G L	318
1003	CTG CAA CAA CGC GTG GTG CCA GCA CGA TGA	1032
319	L Q Q R V V P A R *	327

fdeJ

1	GTGCCAGCACG <u>ATG</u> AAG ATC CTC GTC ACC GGC GCA GGC GGC TTC ATC	47
1	M K I L V T G A G G F I	12
48	GGC GCC ACC CTG GTT GCG CTG CTG CGG CAG CGC GGG CAA GCA GGC	92
13	G A T L V A L L R Q R G Q A G	27

93	GGG	CGG	CGC	ATC	AGC	AGC	CTG	GCC	GTG	CTG	GAC	CGG	CAA	CTG	GGC	137
28	G	R	R	I	S	S	L	A	V	L	D	R	Q	L	G	42
138	GAG	ACT	GCG	CCA	GAC	ATG	GTG	CGC	ATC	GAA	GGC	AAT	CTG	CAA	GAT	182
43	E	T	A	P	D	M	V	R	I	E	G	N	L	Q	D	57
183	GCG	CAG	GTG	CAA	GCA	CAG	ATC	GTC	GCC	TTC	GGC	GCC	GAC	CTG	TGC	227
58	A	Q	V	Q	A	Q	I	V	A	F	G	A	D	L	C	72
228	TTT	CAT	CTG	GCC	GCG	CTG	CCC	GGC	GGC	GCC	GCC	GAG	GCC	GAT	CCC	272
73	F	H	L	A	A	L	P	G	G	A	A	E	A	D	P	87
273	GCG	CTG	AGC	CGG	CGC	GTG	AAC	CTG	GAC	GCC	ACG	TTG	GAC	CTG	TTC	317
88	A	L	S	R	R	V	N	L	D	A	T	L	D	L	F	102
313	GAA	CTG	CTG	GCC	CGG	CGC	GAC	CCG	GCC	ACA	CCG	CCC	CCA	GTG	GTG	362
103	E	L	L	A	R	R	D	P	A	T	P	P	P	V	V	117
363	GTC	TAC	GCC	AGC	ACC	ATT	GCC	GTC	TAT	GGC	CTG	GAC	CTG	CCC	GAT	407
118	V	Y	A	S	T	I	A	V	Y	G	L	D	L	P	D	132
408	CCG	GTG	ACC	CCG	GCC	ACA	CCG	CCC	CGC	CCG	GCA	TTG	ATC	TAT	GGC	452
133	P	V	T	P	A	T	P	P	R	P	A	L	I	Y	G	147
453	GCG	CAC	AAG	CTG	GCT	TGC	GAA	ATC	CTG	CTG	GCC	GAC	TAC	ACC	CGG	497
148	A	H	K	L	A	C	E	I	L	L	A	D	Y	T	R	162
498	CGC	GGC	CTG	CTC	GAT	GGA	CGT	TCC	TTG	CGC	CTG	CCC	GGC	ATC	GTG	542
163	R	G	L	L	D	G	R	S	L	R	L	P	G	I	V	177
543	GCG	CGT	CCG	CGA	GCG	GCA	TCG	GGG	CTG	GTG	TCA	GCC	TTC	ATG	AGC	587
178	A	R	P	R	A	A	S	G	L	V	S	A	F	M	S	192
588	GAG	CTG	TTG	CAT	GCA	CTG	GCG	GCG	GGC	GAG	CTC	TTC	ACC	TGC	CCC	632
193	E	L	L	H	A	L	A	A	G	E	L	F	T	C	P	207
633	GTG	GGG	CCG	CAG	GCG	ACG	GCC	TGG	TGG	ATG	TCG	GCG	CAG	CGC	TGT	677
208	V	G	P	Q	A	T	A	W	W	M	S	A	Q	R	C	222
678	GCC	GAG	AAC	CTG	CTG	CAC	GCA	GCC	TCG	ATG	GAC	CCG	ACC	CGC	GAC	722
223	A	E	N	L	L	H	A	A	S	M	D	P	T	R	D	237
723	CCC	GAA	GGC	GCA	GCG	GCG	CGG	GTC	TGG	CCT	TTG	CCA	GTG	CTG	CGG	767
238	P	E	G	A	A	A	R	V	W	P	L	P	V	L	R	252
768	CTG	TCG	ATC	GCG	CAA	GTG	GTC	GCC	ACG	CTC	TCG	GCC	TTG	TAT	GGC	812
253	L	S	I	A	Q	V	V	A	T	L	S	A	L	Y	G	267
813	GTG	GAT	GGA	CAG	GCG	CTG	GTG	CGC	TAT	GCA	CCG	CAG	CCG	CAG	GTG	857
268	V	D	G	Q	A	L	V	R	Y	A	P	Q	P	Q	V	282
858	GAA	GCG	GTC	TTT	GGC	CGC	TAT	CCC	AGG	CTG	GAT	GAC	CGG	GCT	GCG	902
283	E	A	V	F	G	R	Y	P	R	L	D	D	R	A	A	297
903	CGG	CAG	CTT	GGG	CTG	CGT	GAT	GAT	GGG	TCG	GTG	CAG	GAG	ATG	GTG	947
298	R	Q	L	G	L	R	D	D	G	S	V	Q	E	M	V	312
948	GAG	CGG	GCG	CTC	ATA	TCC	CGG	GGT	TGA	GCACTGCCGACT						986
313	E	R	A	L	I	S	R	G	*							320