

PATRÍCIA CASTELLEN

ANÁLISE ESTRUTURAL E FUNCIONAL DA
REGIÃO *glnBA* DE *Azospirillum brasilense*

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências – Bioquímica, Curso de Pós-Graduação em Bioquímica, Setor de Ciências Biológicas, Universidade Federal do Paraná.

Orientador: Prof. Dr. Fábio de Oliveira Pedrosa

Co-orientadora: Prof. Dra. Roseli Wassem

Curitiba
2005

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Ao meu pai Alfredo, à minha mãe Maria
Áurea e às minhas irmãs Mireille, Juliane e
Andreza.

AGRADECIMENTOS

Agradeço ao Prof. Fábio pela oportunidade de trabalhar no Núcleo de Fixação de Nitrogênio e pela orientação que me mostrou o que é, na realidade, fazer ciência.

Agradeço à Prof. Roseli Wassem que me ensinou a ver a pesquisa científica como alguém que faz parte dela. Obrigada pelas horas que gastou me ajudando a ver o caminho.

À coordenação do curso de pós-graduação pelo apoio prestado.

Aos colegas do Anexo: Marcelo M., Karen, Daniela, Helisson, Leo, Stefan, Valter, Pércio, Giovani, Ane, Michele, Marco Aurélio, Marco Antônio, Salah, Fernanda. Obrigada pelo empréstimo de material, mesmo involuntário... Não esquecendo as conversas e risadas que aliviaram os dias mais pesados.

A todos os colegas do núcleo: Miló, Lupi, Luíza, Adriana I., Feibis, Marcelão, Carol, Rafael, Gás, Claudião, Magda e Michelle Daniel. Pelas informações, dicas e por termos convivido muito bem.

Aos professores Emanuel, Elaine, Gisele, Cynthia, Leda, Berenice, Rose Adele, Adriana T., Liu, por dividerem seu conhecimento.

À D. Jú e Roseli Prado por nos tratarem como filhos, ajudando e aconselhando.

À Djudju, Giovana, André, Ana C. e Lys por serem meus amigos queridos, que me dão conselhos e procuram me ajudar e me animar, cada um do seu jeitinho. Um agradecimento especial à Ju e ao André por me obrigarem a ir à natação quando a água está a 4°C.

Aos amigos que sabem o que significam pra mim: Jana, Carolzinha, Cynthia, Pry, Ricardo, Erich e Roberta. Aprendi muito com vocês, mesmo estando um pouco distante ultimamente.

À minha família, que é a melhor do mundo! Meu pai e minha mãe que abriram mão de tanta vida para que eu estivesse aqui e para que minhas irmãs também chegassem aonde queriam. Eu amo tanto vocês.

Ao Edson pelo carinho, paciência e compreensão. E por me mostrar que não estava sozinha.

A Deus, Jeová, pela vida. Que é surpreendente, por isso eu a estudo.

SUMÁRIO

| | |
|---|------|
| LISTA DE FIGURAS | VI |
| LISTA DE TABELAS | VII |
| LISTA DE ABREVIATURAS | VIII |
| RESUMO | X |
| 1. INTRODUÇÃO | 1 |
| 1.1 Importância da Fixação Biológica de Nitrogênio | 1 |
| 1.2 Nitrogenase | 2 |
| 1.3 Regulação da Nitrogenase | 3 |
| 1.4 <i>Azospirillum brasilense</i> | 4 |
| 1.5 Regulação da Fixação de Nitrogênio em <i>A. brasilense</i> | 4 |
| 1.6 Controle da Atividade da Nitrogenase | 8 |
| 1.7 Sistemas de Assimilação de Amônio | 9 |
| 1.8 Operon <i>glnBA</i> de <i>A. brasilense</i> | 11 |
| 2. OBJETIVOS | 15 |
| 3. MATERIAIS E MÉTODOS | 16 |
| 3.1 Bactérias e Plasmídeos | 16 |
| 3.2 Seção 1: Seqüenciamento e Análise do Plasmídeo pAB441 | 17 |
| 3.2.1 Condições de Cultivo | 17 |
| 3.2.2 Composição dos Meios de Cultura | 17 |
| 3.2.3 Antibióticos | 18 |
| 3.2.4 Construção de Biblioteca Genômica | 19 |
| 3.2.5 Seqüenciamento | 25 |
| 3.2.6 Análise dos dados de seqüenciamento | 26 |
| 3.3 Seção 2: Análise Fisiológica dos mutantes HM26, HM053 e HM210 de <i>A. brasilense</i> | 27 |
| 3.3.1 Condições de cultivo | 27 |
| 3.3.2 Composição dos meios de cultura | 27 |
| 3.3.3 Antibióticos | 28 |
| 3.3.4 Manipulação de DNA | 29 |
| 3.4 Transferência de plasmídeo por conjugação | 30 |
| 3.5 Dosagem de proteínas | 31 |
| 3.6 Atividade de nitrogenase em meio semi-sólido | 31 |
| 3.7 Seqüenciamento do operon <i>glnBA</i> dos mutantes HM26, HM053 e HM210 | 31 |
| 3.7.1 Amplificação de DNA por PCR | 31 |
| 3.7.2 Seqüenciamento | 36 |
| 3.7.3 Análise das seqüências | 36 |
| 4. RESULTADOS E DISCUSSÃO | 37 |
| 4.1 Seqüenciamento do inserto do plasmídeo pAB441 | 37 |
| 4.2 Análise fisiológica dos mutantes HM26, HM053 e HM210 de <i>A. brasilense</i> | 48 |
| 4.3 Análise das seqüências de aminoácidos da glutamina sintetase dos mutantes HM26, HM053 e HM210 de <i>A. brasilense</i> | 48 |
| 5. CONCLUSÃO | 61 |
| 6. REFERÊNCIAS BIBLIOGRÁFICAS | 80 |

LISTA DE FIGURAS

| | |
|---|----|
| FIGURA 1 REGULAÇÃO DO PROCESSO DE FIXAÇÃO DE NITROGÊNIO EM <i>A. brasilense</i> | 5 |
| FIGURA 2 REGULAÇÃO DA ATIVIDADE DE GS E NtrC EM RESPOSTA À CONCENTRAÇÃO DE AMÔNIO EM <i>E. coli</i> | 10 |
| FIGURA 3 REGIÃO <i>glnBA</i> DE <i>A. brasilense</i> E LOCALIZAÇÃO DOS “PRIMERS” UTILIZADOS | 34 |
| FIGURA 4 - ANEXO 1 SEQÜÊNCIA DO INSERTO DO PLASMÍDEO pAB441..... | 63 |
| FIGURA 5 MAPA FÍSICO E GENÉTICO DO INSERTO DO PLASMÍDEO pAB441.. | 41 |
| FIGURA 6 ESQUEMA DA CLONAGEM DE REGIÕES DO INSERTO DE pAB441 EM pLAFR3.18..... | 51 |
| FIGURA 7 EFEITO DE pL+13 SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES <i>Nif^c</i> DE <i>A. brasilense</i> | 52 |
| FIGURA 8 EFEITO DE pL+7 SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES <i>Nif^c</i> DE <i>A. brasilense</i> | 53 |
| FIGURA 9 EFEITO DE pLAFR <i>glnA</i> SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES <i>Nif^c</i> DE <i>A. brasilense</i> | 54 |
| FIGURA 10 ALINHAMENTO DE AMINOÁCIDOS DA PROTEÍNA GS..... | 58 |

LISTA DE TABELAS

| | |
|--|----|
| TABELA 1 BACTÉRIAS E PLASMÍDEOS..... | 16 |
| TABELA 2 SEQUÊNCIA DOS PRIMERS E SUAS TEMPERATURAS DE ANELAMENTO..... | 36 |
| TABELA 3 SIMILARIDADE DE SEQUÊNCIA ENTRE AS PROTEÍNAS CODIFICADAS PELAS ORF`s DO PLASMÍDEO pAB441 E SEQUÊNCIAS DE PROTEÍNAS DO GENBANK | 42 |
| TABELA 4 ATIVIDADE DE GLUTAMINA SINTETASE NOS MUTANTES Nif ^C DE <i>Azospirillum brasilense</i> | 61 |

LISTA DE ABREVIATURAS

μg – micrograma

μL – microlitro

Amp – ampicilina

ADP – Adenosina difosfato

AMP – Adenosina monofosfato

ATP – Adenosina trifosfato

bp – pares de bases

Cm – cloranfenicol

CRP – receptor de AMP cíclico

CTAB – Brometo de N, N'-cetil-trimetil amônio

D. O. – densidade óptica

DNA – ácido desoxirribonucléico

dNTP – Desoxiribonucleosídeo trifosfato

EDA – etilenodiamina

EDTA – Ácido etilenodiamino tetra acético

g – grama

h - hora

kb – quilo pares de bases

mg – miligrama

min – minuto

miniprep – minipreparação de plasmídeos

NAD⁺ - nicotinamida adenina dinucleotídeo

Nal – ácido nalidíxico

Pi – fosfato inorgânico

PCR – Reação em cadeia da polimerase

r – resistência à

RNase - ribonuclease

rpm – rotações por minuto

Sm – estreptomicina

Tc – tetraciclina

Tris – hidroximetil aminometano

q. s. p. – quantidade suficiente para

Nif – fenótipo de capacidade de fixação de nitrogênio

Nif^C – fenótipo de capacidade de fixação de nitrogênio constitutiva

GS – glutamina sintetase

Sistema Ntr – sistema de regulação geral do metabolismo de nitrogênio

GOGAT – glutamato sintase

nm – nanômetro

UV – ultra violeta

V - volts

RESUMO

A fixação biológica de nitrogênio é a redução do nitrogênio atmosférico a amônio, catalisado pela nitrogenase. O metabolismo de nitrogênio é controlado por um complexo sistema regulatório denominado *ntr*, do qual fazem parte as proteínas PII e glutamina sintetase (GS). Em *Azospirillum brasilense*, a proteína GlnB (produto do gene *glnB*) atua como sinalizador dos níveis intracelulares de amônio, controlando a atividade de NifA, a qual é responsável pela ativação dos genes envolvidos na síntese do complexo da nitrogenase. A GS, produto do gene *glnA*, participa do sistema de assimilação do nitrogênio fixado, catalisando a síntese de glutamina, utilizando como substrato glutamato e amônio. O operon *glnBA* de *A. brasilense* foi identificado no plasmídeo pAB441, que foi isolado a partir de uma biblioteca genômica. Em experimentos de complementação genética com estirpes Nif^C de *A. brasilense* (HM26, HM053 e HM210) realizados por Vitorino (2001), observou-se que a presença do plasmídeo pAB441 leva a uma repressão completa da nitrogenase nestes mutantes, tanto na presença quanto na ausência de amônio. Com o objetivo de identificar os genes responsáveis por tal efeito foi construída uma biblioteca genômica do plasmídeo pAB441. Os insertos dos clones foram sequenciados e a análise da seqüência permitiu a identificação de dezessete ORF's na região do inserto do plasmídeo pAB441. Regiões deste inserto, exceto a região central, entre os sítios de restrição *SacI/PstI*, foram sub-clonadas para determinar seu efeito sobre a atividade da nitrogenase dos mutantes HM. Estas clonagens foram feitas em vetor estável em *A. brasilense* e a atividade de nitrogenase foi determinada nos mutantes contendo tais plasmídeos. Uma vez que o fenótipo Nif^C dos mutantes pode ser decorrente de mutações nos genes *glnBA*, este operon foi parcialmente seqüenciado nestas linhagens. Nas linhagens HM26 e HM210 foi detectada uma transversão C→T que levou à troca de uma glicina por cisteína na GS destes mutantes. Na HM26 a troca de aminoácido ocorreu na posição 53, correspondente ao sítio de ligação a amônio da GS enquanto na HM210 a troca deu-se na posição 129, o sítio de coordenação de íons metálicos. Ambas as mutações podem ser responsáveis pela baixa atividade biossintética de GS observada nestes mutantes.

1. INTRODUÇÃO

1.1 Importância da Fixação Biológica de Nitrogênio

A fixação biológica de nitrogênio consiste na redução do dinitrogênio atmosférico (N_2) a amônio (NH_4^+) pelas bactérias diazotróficas. Este processo é de grande importância econômica, pois o nitrogênio é um dos fatores limitantes para o desenvolvimento de plantas de interesse comercial. A produção brasileira de fertilizantes nitrogenados é de 1 milhão de toneladas para uma demanda de 1,7 milhão, sendo a diferença suprida pela importação, o que implica em altos custos na produção de grãos (Ministério da Agricultura, Pecuária e Abastecimento, 2003).

Na cultura de soja o nitrogênio é o nutriente requerido em maior quantidade. Estima-se que para produzir 1000 kg de grãos de soja são necessários 80 kg de nitrogênio. As fontes de nitrogênio disponíveis para a cultura da soja são os fertilizantes nitrogenados químicos e a fixação biológica do nitrogênio. Nesta cultura, bactérias do gênero *Bradyrhizobium* são utilizadas como inoculantes e em alguns casos podem fornecer todo o nitrogênio de que ela necessita (HUNGRIA et al., 2001).

A utilização de microrganismos associativos capazes de fixar nitrogênio em outras culturas é de grande interesse tendo em vista os benefícios econômicos e ambientais dessa técnica. Sabe-se, por exemplo, que o nitrogênio é o nutriente mais requerido, o que mais limita a produtividade de milho e também o que mais onera a cultura. Cerca de 22 kg de nitrogênio são necessários para produzir 1.000 kg de grãos (COELHO et al., dados não publicados – Embrapa Milho). Na lavoura de arroz, a quantidade de nitrogênio recomendada está em torno de 90 kg ha⁻¹ (Embrapa Arroz).

Bactérias fixadoras de nitrogênio dos gêneros *Azotobacter*, *Enterobacter*, *Bacillus*, *Klebsiella*, *Azospirillum*, *Herbaspirillum*, *Gluconacetobacter*, *Burkholderia* e *Azoarcus* são capazes de associar-se a uma grande variedade de gramíneas, aumentando sua produção

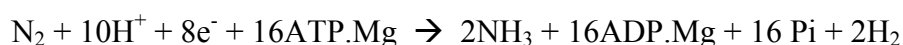
(TARRAND et al., 1978; RENNIE et al., 1982; GILLIS et al., 1989, 1995; BALDANI et al., 1992; REINHOLD-HUREK et al., 1993; DONG et al., 1994). Visto que a associação entre gramíneas e bactérias fixadoras de nitrogênio possibilita uma substancial economia de fertilizantes nitrogenados na agricultura (PEDROSA, 1987), sendo também menos nociva ao ambiente, o estudo destas bactérias é importante para posterior utilização como inoculantes em culturas de interesse comercial.

1. 2 Nitrogenase

A reação de redução do dinitrogênio a amônia é catalisada pelo complexo enzimático nitrogenase, composto por duas proteínas: a dinitrogenase ou proteína-MoFe (proteína ferro-molibdênio) e a dinitrogenase redutase ou proteína-Fe (proteína ferro) (EADY, 1986). A dinitrogenase (MoFe) é um tetrâmero formado por duas subunidades α e duas subunidades β codificadas pelos genes *nifD* e *nifK*, respectivamente. Cada tetrâmero apresenta dois cofatores ferro-molibdênio (FeMoco) que são os sítios de ligação e redução do nitrogênio e dois centros P (revisado por ZHANG et al., 1997). A dinitrogenase redutase é um homodímero γ_2 codificado pelo gene *nifH*, que contém um núcleo 4Fe-4S e sua função é transportar elétrons até a dinitrogenase (DEAN *et al.*, 1993).

Os centros ativos fortemente reduzidos no complexo da nitrogenase são facilmente destruídos por oxidação com O_2 , inativando-a. Em função de sua sensibilidade a O_2 a síntese da nitrogenase é reprimida em altas tensões de oxigênio (ZUMFT e MORTENSON, 1975).

A reação catalisada pela nitrogenase é a seguinte:



Esta reação é inibida por ADP.Mg, mas não por NH₃ ou glutamina, este último, produto imediato da assimilação da NH₃ (HAAKER e KLUGKIST, 1987).

A biossíntese e o funcionamento da nitrogenase requerem, além dos produtos dos genes *nifHDK*, produtos de outros genes responsáveis por sua regulação transcricional e pós-traducional, transporte de elétrons, biossíntese de cofatores e transporte de íons. Condições ambientais favoráveis tais como a ausência de nitrogênio fixado, baixas tensões de oxigênio, presença de molibdênio e temperatura adequada permitem a expressão da capacidade de fixação de nitrogênio pelos diazotrofos (POSTGATE, 1982).

1. 3 Regulação da nitrogenase

O processo de fixação de nitrogênio catalizado pelo complexo da nitrogenase envolve um alto gasto energético para o microorganismo, por isso é altamente regulado tanto na atividade da nitrogenase como na expressão dos genes *nif* (POSTGATE, 1982; FISCHER, 1994).

Uma vez que o oxigênio tem efeito deletério sobre os componentes da nitrogenase é vantajoso para as bactérias reprimir sua transcrição quando os níveis de oxigênio são altos. Além disso, a transcrição se torna desnecessária quando o nível de nitrogênio fixado é suficiente. Sendo assim, ao nível transcricional, a síntese da nitrogenase é regulada em resposta às condições ambientais de oxigênio e amônio (HARTMANN et al., 1986). O grau em que cada estímulo afeta a transcrição é característico de cada espécie de diazotrofo. Em diazotrofos de vida livre a expressão dos genes *nif* é mais sensível aos níveis celulares de amônio do que em organismos simbióticos (MERRICK, 1992).

Estudos sobre a fixação de nitrogênio em *Klebsiella pneumoniae* geraram um modelo para a regulação transcricional da nitrogenase. Neste modelo, o controle da transcrição do “cluster” de genes *nif* depende de NifA (produto do gene *nifA*), um ativador de transcrição de promotores dependentes de σ^{54} . A transcrição de NifA, por sua vez, está sob controle dos

produtos dos genes *ntrB* e *ntrC*, que fazem parte do sistema *ntr*, o qual controla o metabolismo geral de nitrogênio em enterobactérias (MERRICK, 1992). Em *K. pneumoniae*, o gene *nifA* é co-transcrito com *nifL* que codifica para NifL, uma flavoproteína redox sensível a nitrogênio (HILL, 1996; MERRICK, 1992). NifL age como regulador negativo de NifA aparentemente formando um complexo inibitório que bloqueia a atividade de nucleosídeo trifosfatase de NifA (DIXON, 1998; EYDMANN, 1995).

1.4 *Azospirillum brasilense*

As bactérias do gênero *Azospirillum* são microrganismos diazotróficos, aeróbicos, capazes de se associar endofiticamente à plantas de interesse agrícola como milho, arroz, trigo e sorgo (DÖBEREINER e DAY, 1976; DÖBEREINER, 1991; PEDROSA, 1988).

1.5 Regulação da fixação de nitrogênio em *Azospirillum brasilense*

A expressão dos genes envolvidos na regulação do metabolismo de nitrogênio em *A. brasilense*, como em outros organismos, está sob o controle de fatores ambientais como fonte de nitrogênio e concentração de oxigênio (DOBEREINER e PEDROSA, 1987).

Um modelo de regulação do metabolismo de nitrogênio em *A. brasilense* foi proposto por VANDE BROEK e colaboradores (2000) e modificado por Huergo (2002) e contempla o conhecimento atual (Figura 1).

A regulação da fixação de nitrogênio em *A. brasilense* apresenta algumas diferenças em relação a *K. pneumoniae*. Em *A. brasilense*, *nifA* é expresso constitutivamente, mas seu nível de expressão é menor em presença de altas concentrações de oxigênio e amônio. As proteínas regulatórias NtrB e NtrC não atuam diretamente na transcrição de *nifA*, como indicam estudos da sua expressão em mutantes *ntrBC* de *A. brasilense* (LIANG et al., 1992). Apesar disso, NtrC é necessária para a expressão máxima de *nifA* sob condições de fixação de nitrogênio (LIANG et al., 1993).

REGULAÇÃO DO PROCESSO DE FIXAÇÃO DE NITROGÊNIO EM *A. brasilense*

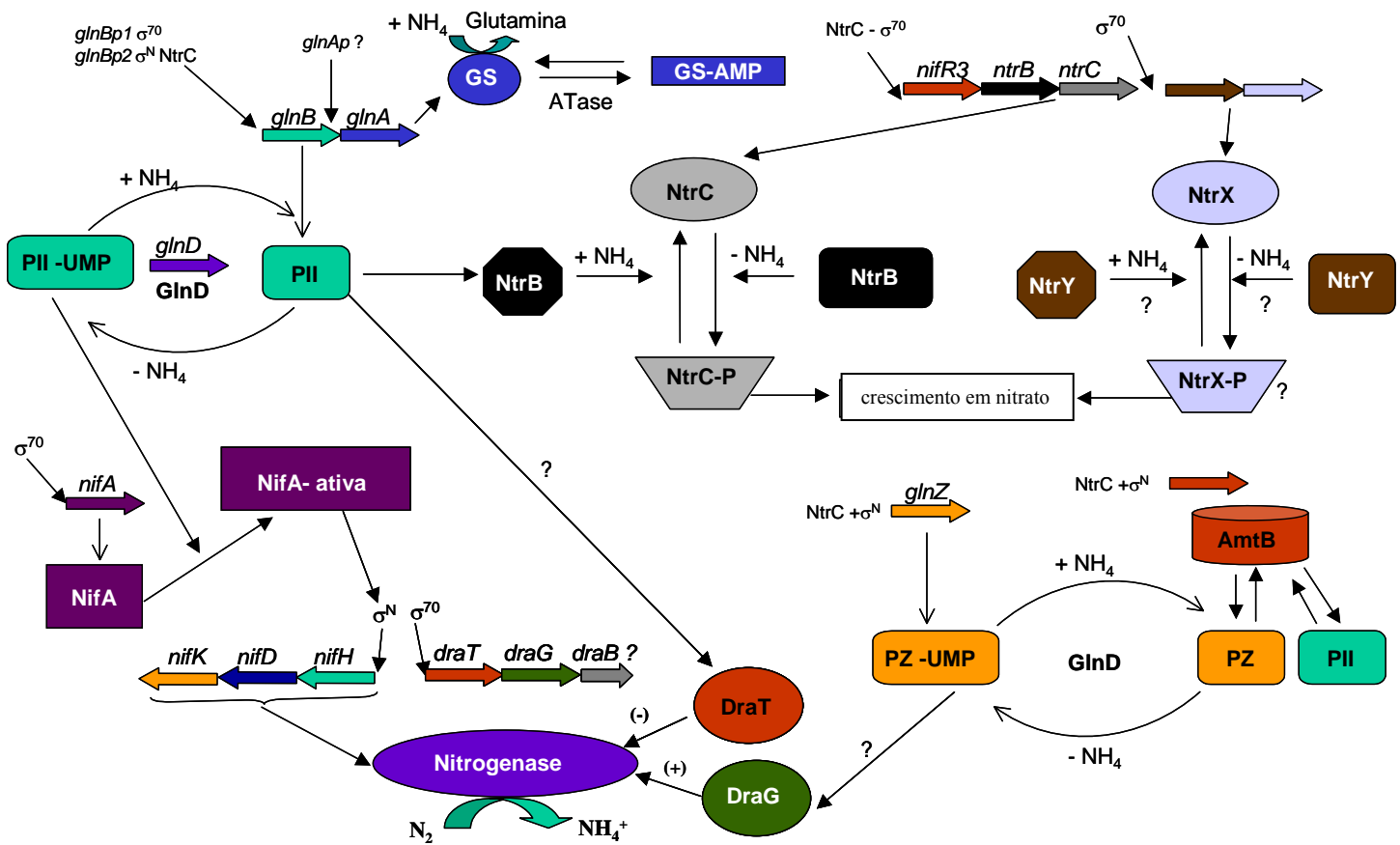


Figura 1: Regulação do sistema Ntr em *A. brasilense*. Adaptado de VANDE BROEK (2000) por HUERGO (2002).

As proteínas NtrX e NtrY, apesar de exercerem funções regulatórias de forma similar a NtrBC, também não foram capazes de ativar a transcrição de *nifA* em *A. brasilense* (ISHIDA et al., 2002). Uma região essencial para a atividade do promotor de *nifA* foi identificada entre os nucleotídeos -67 e -47 do início de transcrição de *nifA*. Um provável sítio reconhecido pela subunidade σ^{70} foi identificado nesta região e pode constituir o promotor de *nifA*. Estudos de expressão utilizando a fusão deste promotor com o gene *lacZ* mostraram que um efeito combinado de altas concentrações de oxigênio e amônio é responsável pela repressão do promotor de *nifA* (FADEL-PICHETH et al., 1999).

Em mutantes *glnB* de *A. brasilense*, a proteína NifA é provavelmente sintetizada na forma inativa, sugerindo o envolvimento de GlnB na regulação da atividade de NifA por amônio (LIANG et al., 1992). A atividade de NifA também é modulada em resposta à concentração de nitrogênio na célula e pela tensão de oxigênio (LIANG, et al., 1992). NifA não é controlada por NifL como em *K. pneumoniae*, mas sim auto-regulada por modificações conformacionais. Através de ensaios utilizando a proteína NifA truncada foi possível determinar que seu domínio N-terminal está envolvido na regulação negativa de sua própria atividade em presença de amônio (ARSÈNE et al., 1996). Uma estirpe mutante em *glnB*, apresentou fenótipo Nif⁻, apesar de não ter a transcrição de *nifA* afetada (LIANG, et al., 1992). Isso pode indicar que a inibição de NifA por amônio é revertida pela ação da proteína GlnB em condições de fixação de nitrogênio (de ZAMAROCZY, 1998). Acredita-se que a forma uridililada da proteína GlnB (GlnB-UMP) seja necessária para a ativação da proteína NifA (STEENHOUDT e VANDERLEYDEN, 2000).

O gene *glnB*, que codifica para a proteína GlnB, é adjacente ao gene *glnA*. O operon *glnBA* apresenta três diferentes promotores, utilizados seletivamente conforme a disponibilidade de nitrogênio (de ZAMAROCZY et al, 1993) (Figura 1). Em condições de excesso de nitrogênio fixado, *glnBA* é cotranscrito em baixos níveis a partir do promotor

glnBp1 (σ^{70}) e *glnA* é transcrito em níveis mais altos a partir de seu promotor, até então não identificado. A transcrição máxima de *glnBA* ocorre em baixas concentrações de íons amônio e é iniciada a partir do promotor *glnBp2* (σ^{54}). Nestas condições, *glnA* é pouco expresso a partir de *glnAp* e é cotranscrito com *glnB*. Sendo assim, a proteína GlnB de *A. brasilense* é sintetizada em altos níveis sob condições de amônio limitante e em baixos níveis sob condições de excesso de nitrogênio (de ZAMAROCZY *et al.*, 1993). O promotor *glnBp1* é sobreposto a duas seqüências reconhecidas por NtrC, que em situação limitante de amônio, são ocupadas por NtrC-P, favorecendo a transcrição a partir do promotor *glnBp2*. A transcrição de *glnA*, aparentemente, não é ativada por NtrC mas está alterada em um mutante *glnB* (estirpe 7628) sugerindo que a proteína GlnB age na regulação da expressão do gene *glnA* (HUERGO *et al.*, 2003).

Em *A. brasilense*, foi identificada a proteína PZ, com alta similaridade a GlnB (66% de identidade e 82% de similaridade). Apesar de similares, estas duas proteínas desempenham papéis diferentes na regulação da fixação de nitrogênio. A proteína PZ pode estar relacionada com o transporte de íons amônio, visto que um mutante *GlnZ⁻* apresentou uma taxa aumentada na captação de metilamônio e a expressão de *glnZ* neste mutante foi capaz de reverter este fenótipo (de ZAMAROCZY, 1998).

Os genes que codificam para o sistema NtrY/NtrX de *A. brasilense*, apresentam alta homologia com os de *Azorhizobium caulinodans* (ISHIDA *et al.*, 2002). Neste microrganismo, este sistema está envolvido na expressão da proteína NifA e no crescimento em nitrato (PAWLOWSKI *et al.*, 1991). Em *A. brasilense*, foi evidenciada a participação deste sistema apenas no metabolismo de nitrato e não na expressão da proteína NifA (VITORINO *et al.*, 2001; ISHIDA *et al.*, 2000).

1.6 Controle da atividade da nitrogenase

Em *A. brasilense*, a atividade da nitrogenase é regulada por NH_4^+ e oxigênio. A subunidade dinitrogenase redutase é inativada reversivelmente por ADP-ribosilação sob condições incompatíveis com a fixação de nitrogênio (HARTMANN *et al.*, 1986). Este processo envolve as proteínas DRAT e DRAG (produtos dos genes *draT* e *draG*). Na presença de altas concentrações de amônio, DRAT inativa a subunidade dinitrogenase redutase catalizando a transferência de um grupamento ADP-ribosil do NAD^+ para o resíduo Arg-101. Em baixas concentrações de amônio, a inativação é revertida pela ação de DRAG, a qual remove o grupamento ADP-ribosil (ZHANG *et al.*, 1993).

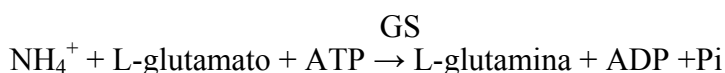
DRAT e DRAG também estão sujeitas a regulação pós-traducional (ZHANG *et al.*, 1994). DRAG é ativa e DRAT é inativa sob condições de fixação de nitrogênio. Quando células de *A. brasilense* são tratadas com NH_4^+ ou transferidas de um meio microaeróbico para um meio anaeróbico, DRAG se torna inativa e DRAT é ativada. Após o consumo do amônio do meio ou o retorno ao meio microaeróbico, DRAG se torna ativa e a atividade da nitrogenase é recuperada (ZHANG *et al.*, 1993). Mutantes *ntrBC* não sofrem perda da atividade da nitrogenase em altos níveis de amônio (PEDROSA e YATES, 1984). Este resultado sugere que o efeito destas mutações provavelmente envolve mudanças na atividade de DRAG (MACHADO *et al.*, 1995 e LIANG *et al.*, 1993).

Uma estirpe GlnZ⁻ apresentou somente uma recuperação parcial (de 20 a 40%) da atividade da nitrogenase após 80 minutos de sua inativação reversível induzida por amônio, enquanto a estirpe selvagem recuperou totalmente a atividade após 10 minutos. Esse resultado mostra que GlnZ pode estar envolvida no controle pós-traducional da atividade da nitrogenase em *A. brasilense* (KLASSEN *et al.*, 2001). Estudos recentes mostram que a proteína GlnZ é importante para a reativação da nitrogenase, provavelmente pela interação

com a proteína DRAG e que a proteína GlnB é fundamental para o desligamento da nitrogenase (KLASSEN et al., 2005 – submetido).

1.7 Sistemas de Assimilação de amônio

A amônia produzida ou captada do ambiente é utilizada na síntese de glutamina e glutamato, que servem como doadores de nitrogênio para as reações biossintéticas (MERRICK e EDWARDS, 1995). A assimilação do amônio pode ocorrer através de duas vias, sendo que a principal envolve duas enzimas: a glutamina sintetase (GS) e a glutamato sintase (GOGAT). Em *A. brasilense*, os genes que codificam para as duas subunidades da glutamato sintase são *gltDB* (WESTBY et al., 1987). A glutamina sintetase é codificada pelo gene *glnA*, o qual faz parte de um operon com *glnB* (BOZOUKLIAN et al., 1986). As reações que compõem esta via são as seguintes:



A glutamato sintase catalisa a aminação redutiva do α -cetogluturato, produzindo glutamato, empregando glutamina como doadora de nitrogênio. A glutamina sintetase catalisa a reação de formação da glutamina a partir de glutamato e NH_4^+ . Ela é uma enzima altamente regulada, tanto a nível transcricional como pós-traducional. Produtos finais do metabolismo da glutamina, como a alanina e a glicina, são inibidores alostéricos de GS. Além da regulação alostérica há a inibição por adenililação, que aumenta ainda mais a sensibilidade da enzima aos inibidores alostéricos (MAGASANIK, 1993). No modelo para bactérias entéricas (Figura 2), baseado em *Klebsiella pneumoniae*, a adenililação e desadenililação da GS são promovidas pela adenilil transferase, que tem sua atividade modulada pela ligação com a proteína GlnB. Esta modulação permite um decréscimo da

**REGULAÇÃO DA ATIVIDADE DE GS E NtrC EM RESPOSTA À
CONCENTRAÇÃO DE AMÔNIO EM *K. pneumoniae***

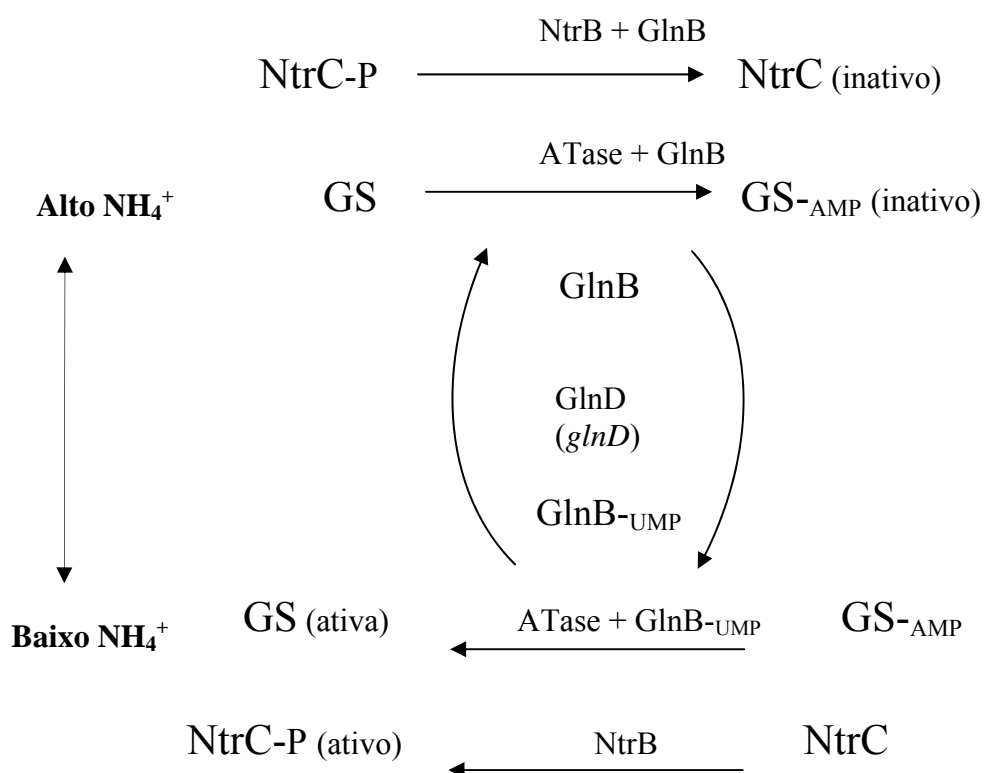


Figura 2: Modelo esquemático da regulação das atividades da glutamina sintetase e sistema Ntr em resposta à pressão de nitrogênio no meio. A Utase (GlnD) catalisa a uridililação e deuridililação da GS. A NtrB catalisa a fosforilação e defosforilação da NtrC (MERRICK & EDWARDS, 1995).

atividade da glutamina sintetase quando os níveis de glutamina são altos, e um aumento dessa atividade quando os níveis de glutamina são baixos e os substratos 2-oxoglutarato e ATP estão disponíveis (MERRICK e EDWARDS, 1995). Diferentemente do que ocorre em bactérias entéricas, em *A. brasilense*, a regulação do nível de adenililação de GS não é mediada por PII (de ZAMAROCZY, 1998).

A segunda via de assimilação de amônio envolve a enzima glutamato desidrogenase (GDH):



A glutamato desidrogenase catalisa a reação reversível da aminação do α -cetoglutarato pela amônia formando glutamato e utilizando as coenzimas NADH ou NADPH. Em microrganismos, a GDH NADH-dependente parece estar envolvida em funções catabólicas enquanto a NADPH-dependente em funções anabólicas (BENDER et al., 1977).

A primeira via (GS – GOGAT) é a principal responsável pela assimilação de amônio em condições limitantes deste composto. A segunda via (GDH) funciona quando a concentração de amônio é alta, pois a GDH possui uma afinidade relativamente baixa para NH_4^+ (MERRICK e EDWARDS, 1995). Em *A. brasilense* a via GS-GOGAT parece ser a via predominante de assimilação de amônio independente da fonte de nitrogênio utilizada para crescimento (WESTBY et al., 1987).

1.8 Operon *glnBA* de *A. brasilense*

Duas estirpes de *A. brasilense* excretoras de amônio apresentaram uma baixa atividade de glutamina sintetase (GS). Uma delas é incapaz de fixar nitrogênio (Nif⁻; 7029) e a outra apresenta atividade de nitrogenase independente do nível de nitrogênio fixado do meio (Nif^c; 7028). A estirpe 7029 apresenta a troca de um resíduo conservado de ácido aspártico na posição 52 da GS por um resíduo de asparagina (van DOMMELEN et al.,

2003). Esta mutação se encontra no sítio proposto para ligação de íons amônio à GS (LIAW et al., 1995), não permitindo a geração de espécies ativas de NH_3 , necessárias na reação biossintética da GS. A GS da estirpe 7028 apresenta um resíduo de cisteína na posição 322 onde, na selvagem, é encontrado um resíduo de arginina. Esta mutação se encontra no sítio de ligação a glutamato (LIAW et al., 1993). Ambos os mutantes tiveram seus fenótipos restaurados pelo gene *glnA* selvagem de *A. brasilense*, inclusive recuperando sua capacidade de fixar nitrogênio (7029) e a regulação da nitrogenase por NH_4^+ (7028). Estes resultados confirmam o papel central de GS na regulação da fixação de nitrogênio (van DOMMELEN et al., 2003).

Mutantes espontâneos de *A. brasilense* estirpe Sp7, resistentes a etilenodiamina, excretoras de amônio e capazes de fixar nitrogênio constitutivamente, foram isoladas por MACHADO (1988). Quatro dessas estirpes foram caracterizadas fisiologicamente: HM14, HM26, HM053 e HM210 e apresentam atividade normal de glutamato sintase e da glutamato desidrogenase. O sistema de regulação da atividade da nitrogenase estava parcialmente afetado, visto que o aumento da concentração de amônio no meio implicava apenas em diminuição da atividade da nitrogenase e esta era insensível ao desligamento (“switch off”) induzido por NH_4^+ . Estes mutantes foram classificados em três grupos quanto à atividade de GS: o mutante HM053 que apresentava baixos níveis de GS altamente adenililada em todas as condições testadas; o mutante HM14 que apresentava baixos níveis de GS regulada por amônio como no selvagem; os mutantes HM26 e HM210 que apresentavam níveis normais de GS, porém adenililada constitutivamente (MACHADO, 1988).

A caracterização genética mais detalhada destes mutantes foi feita por VITORINO (2001), através de estudos de complementação onde foi possível verificar que os fenótipos apresentados pelos mutantes HM são corrigidos por diferentes genes. Assim, os plasmídeos

pMAK4 (*nifA* de *A. brasilense*, sob o controle de seu próprio promotor), pLHR31 (*nifR3*), pLHB1 (*ntrB* de *A. brasilense*), pLHC1 (*ntrC* de *A. brasilense*), pLHBC1 (*ntrB* e *ntrC* de *A. brasilense*) e pLHR3BC1 (*nifR3ntrBntrC* de *A. brasilense*), cujos genes estavam sob o controle de um promotor constitutivo (*plac*), não foram capazes de alterar o fenótipo Nif^c destes mutantes (VITORINO, 2001).

O mutante HM053, é provavelmente, um mutante na região promotora do operon *glnBA* visto que o gene *glnB* expresso constitutivamente (plasmídeo pJC1) restaura o controle da atividade de GS neste mutante. O plasmídeo pAB441 (que contém entre outros o operon *glnBA* de *A. brasilense*) restaurou o controle da atividade e parcialmente os níveis de GS deste mutante. A presença do gene *glnB* expresso constitutivamente (pJC1) também restaurou o controle da atividade de GS na estirpe HM210, o que sugere que esta seja mutante em *glnB*.

O plasmídeo pHM9 (*nifR3ntrBCXY* sob o controle de seus próprios promotores) foi capaz de complementar a capacidade de crescimento em nitrato e a regulação da atividade da nitrogenase apenas no mutante HM26. Entretanto, quando os plasmídeos pLHR3BC1 (*nifR3ntrBntrC* sob o promotor *plac*) e pL46 (*ntrXY* sob seu próprio promotor) estão presentes neste mutante, não são capazes de restaurar o controle da atividade da nitrogenase por NH₄⁺. Esse resultado sugere a existência de interação entre estes genes e seus produtos e que esta interação seja responsável pela complementação do fenótipo do mutante HM26.

O plasmídeo pL46 (*ntrXY*) restaurou o controle e a regulação por amônio da atividade da nitrogenase no mutante HM14. Este e outros resultados sugerem que a linhagem HM14 apresenta alteração funcional no sistema NtrY/X (VITORINO et al., 2001).

Uma observação surpreendente foi a de que transconjugantes das estirpes HM26, HM053 e HM210 contendo o plasmídeo pAB441, embora tivessem o crescimento dependente de nitrato restaurado, perderam completamente a atividade da nitrogenase tanto

na ausência como na presença de amônio. É relevante citar que a atividade da nitrogenase do mutante HM14 e da estirpe selvagem não foi alterada em presença do plasmídeo pAB441 (VITORINO, 2001). Foi então levantada a hipótese de que a perda da atividade da nitrogenase destes transconjugantes decorre de um efeito conjunto de suas mutações e outros genes presentes no inserto de 20kb de DNA de *A. brasilense* no pAB441, em adição aos genes *glnB* e *glnA*.

2. OBJETIVOS

2.1 Objetivo geral

Determinar os fatores genéticos envolvidos na repressão da nitrogenase nos mutantes HM053, HM26 e HM210 pelo plasmídeo pAB441.

2.2 Objetivos específicos:

- Construir uma biblioteca aleatória do plasmídeo pAB441;
- Seqüenciar os clones obtidos visando obter a seqüência completa do inserto de DNA de *A. brasilense* no pAB441;
- Analisar as seqüências para a identificação de possíveis genes envolvidos na regulação da fixação de nitrogênio e metabolismo nitrogenado;
- Identificar os genes responsáveis pela ausência da atividade da nitrogenase nos mutantes HM26, HM053 e HM210;
- Realizar ensaios fisiológicos que avaliem o envolvimento destes genes na regulação da fixação de nitrogênio em *A. brasilense*;
- Seqüenciar o operon *glnBA* dos mutantes HM e determinar as interações de seus produtos com os genes do pAB441.

3. MATERIAIS E MÉTODOS

Para atingir os objetivos propostos foram utilizadas duas principais estratégias: seqüenciamento do inserto do plasmídeo pAB441 para identificação dos genes presentes e testes de complementação dos fenótipos para identificação das mutações presentes nos mutantes HM. Os métodos utilizados neste trabalho serão descritos separadamente em duas seções. A primeira se refere à metodologia empregada para o seqüenciamento e análise do inserto do plasmídeo pAB441, enquanto a segunda se refere à análise fisiológica dos mutantes HM26, HM053 e HM210 de *A. brasilense*.

3.1 Bactérias e Plasmídeos

As bactérias e plasmídeos utilizados neste trabalho estão listadas na tabela abaixo.

| Estirpe ou plasmídeo | Propriedades | Referência ou Fonte |
|-----------------------------|--|--------------------------------|
| Estirpe | | |
| <i>E. coli</i> | | |
| TOP10 | Sm ^r F ⁻ [<i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) <i>lacZ</i> ΔM15] | www.invitrogen.com |
| S17.1 | Rp4-2 Tc ^r ::Mu Km ^r ::Tn7 Tra ⁺ Sm ^r recA ⁺ | SIMON et al., 1983 |
| <i>A. brasilense</i> | | |
| FP2 | Estirpe selvagem, SP7 Nif ⁺ , Nal ^r , Sm ^r | PEDROSA E YATES, 1984 |
| HM14, HM26, HM053, HM210 | FP2, EDA ^r , Nif ^c , Nal ^r , Sm ^r , Amt ⁻ | MACHADO, 1988 |
| Plasmídeos | | |
| pUC18 | <i>lacI</i> , <i>lacZ</i> , Amp ^r , ori pBR322 | YANISCH-PERRON et al., 1985 |
| pTZ18R | pTZ19R com sítio de policlonagem invertido, Amp ^r | MEAD et al., 1986 |
| pRK2013 | Tra ⁺ Km ^r , plasmídeo mobilizador | FIGURSKI & HELINSKI, 1979 |
| pAB441 | 20Kb do genoma de <i>A. brasilense</i> em vetor pVK100, Tc ^r | BOZOUKLIAN & ELMERICH, 1986 |

| | | |
|------------|---|-----------------------------|
| pLAFR3.18 | derivado do vetor pLAFR3 com sítio de policlonagem do vetor pTZ18R, Tc ^r , Cm ^r | SOUZA et al., 1995 |
| pTZ+7 | fragmento <i>SacI/HindIII</i> de 7Kb do pAB441 subclonado em pTZ18 | VITORINO, 2001 |
| pTZ+13 | fragmento <i>HindIII/PstI</i> de 5Kb do pAB441 subclonado em pTZ18 | Este trabalho |
| pLAFRglnA2 | fragmento <i>EcoRI</i> de 1,6 kb contendo GS de <i>A. brasilense</i> clonado em pLAFR3.18 | HUERGO, comunicação pessoal |
| pL+7 | fragmento <i>SacI/HindIII</i> de 7Kb do pAB441 subclonado em pLAFR3.18 | Este trabalho |
| pL+13 | fragmento <i>HindIII/PstI</i> de 5Kb do pAB441 subclonado em pLAFR3.18 | Este trabalho |

3.2 Seção 1: Seqüenciamento e análise do inserto do plasmídeo pAB441

3.2.1 Condições de cultivo

As estirpes de *E. coli* foram cultivadas em meio líquido Luria-Broth (LB) ou Terrific Broth (TB) (SAMBROOK *et al*, 1989) a 37°C, incubadas em agitador rotatório a 200 rpm. Para crescimento em meio sólido foram utilizados MacConkey (MILLER, 1992) ou LA (SAMBROOK *et al*, 1989), em estufa a 37°C.

As culturas de *E. coli* foram mantidas em glicerol 50% e estocadas a -20°C.

3.2.2 Composição dos meios de cultura

| Luria Broth (LB) | Gramas/litro |
|---------------------|--------------|
| Extrato de levedura | 5,0 |
| Cloreto de sódio | 10,0 |
| Triptona | 10,0 |

O meio sólido foi obtido pela adição de ágar na concentração de 15 g/L ao meio líquido e foi denominado LA.

| Terrific Broth (TB) | Gramas/litro |
|---------------------|--------------|
| Extrato de Levedura | 24,0 |
| Bactotripton | 12,0 |
| Glicerol | 4,0 |

| Mac Conkey | Gramas/litro |
|------------------|--------------|
| Peptona | 20,0 |
| Polipeptona | 3,0 |
| Lactose | 10,0 |
| Sais biliares | 1,5 |
| Cloreto de sódio | 5,0 |
| Ágar | 15,0 |
| Vermelho neutro | 0,03 |
| Cristal violeta | 0,001 |

3.2.3 Antibióticos

Os antibióticos usados foram:

| | Concentração final |
|---------------------|--------------------|
| Ampicilina (Amp) | 250 µg/mL |
| Estreptomicina (Sm) | 10 µg/mL |
| Tetraciclina (Tc) | 10 µg/mL |

3.2.4 Construção de biblioteca genômica

3.2.4.1 Purificação de pAB441 em larga escala

A purificação de DNA plasmidial em grande escala foi feita em gradiente de cloreto de céσιο (SAMBROOK *et al.*, 1989). Células de 100 mL de uma cultura de *E. coli* contendo o plasmídeo pAB441 foram coletadas por centrifugação a 5.000 rpm por 10 minutos e ressuspensas em 40mL de GET. Em seguida, foram adicionados 80mL de solução de lise e a solução incubada em gelo por 5 minutos. As proteínas foram precipitadas com a adição de 60 mL de acetato de potássio 3 mol/L pH5,5 e incubação por 10 minutos em gelo. O DNA plasmidial foi precipitado com 100 mL de isopropanol absoluto, lavado em 100mL de etanol 80%, seco a vácuo e dissolvido em 9 mL de água milli-Q estéril. Mais tarde as amostras foram analisadas em gel de ágar. Em seguida, foi adicionada água até a solução atingir uma massa de 10,5 g e foram adicionados 11,05g de CsCl₂ à preparação. Após a dissolução foram adicionados 10 mg de brometo de etídeo. A solução foi centrifugada a 2.000 rpm por 15 minutos e o sobrenadante transferido para tubos de ultracentrífuga. A amostra foi centrifugada por 48 horas em ultracentrífuga Beckman L65B (rotor Ti 60) a 55.000 rpm a 15°C. As bandas formadas no gradiente foram visualizadas sob luz ultravioleta e coletadas com o auxílio de agulha e seringa. Dois volumes de água foram adicionados ao volume coletado das bandas (aproximadamente 2 mL) e posteriormente, foram adicionados seis volumes de etanol absoluto. A solução foi centrifugada a 3.000 rpm por 20 minutos. O DNA foi ressuspense em 300µL, tratado com RNase e extraído com 300µL de fenol:clorofórmio. Após isso, foi precipitado com 600µL de etanol absoluto, lavado com etanol 70% e redissolvido em 500µL de água.

3.2.4.2 Nebulização

O plasmídeo pAB441 foi nebulizado seguindo-se o seguinte protocolo. No interior de um nebulizador (TOPO Shotgun Subcloning Kit - Invitrogen) foi adicionada uma solução contendo 25µg de DNA, 0,3M de acetato de sódio e glicerol 30% em um volume final de

1mL. Esta solução foi nebulizada sob 1,96kPa de nitrogênio gasoso durante 15 segundos. Após isso, a solução de DNA foi coletada e o resultado da nebulização foi verificado através de eletroforese em gel de agarose. Os passos de nebulização e verificação foram repetidos até serem obtidos fragmentos de DNA com cerca de 1kb. Após isso a solução foi precipitada com etanol absoluto 3 volumes por aproximadamente 2 horas a 4°C. O DNA foi coletado por centrifugação a 13.000 rpm por 20 min a 4°C, lavado com etanol 70% e seco.

3.2.4.3 Eletroforese e purificação de fragmentos de DNA

A eletroforese de DNA foi feita em gel de ágar horizontal (SAMBROOK, 1989).

O tampão de corrida utilizado foi o TBE 1X concentrado (Tris-base 89 mmol/l, ácido bórico 89 mmol/L e EDTA 2 mmol/L pH 8,0) e a solução de aplicação de amostra em gel de agarose foi FSUDS [azul de bromofenol 0,25% (p/v), xileno cianol 0, 25% (p/v), Ficoll 25% (p/v), SDS 0,5% (p/v), Tris-HCl 10 mmol/L pH 8,0 e EDTA 1 mmol/L pH 8,0].

O DNA foi visualizado após tratamento do gel com solução de brometo de etídeo (0,5 µg/mL) em transiluminador de luz ultravioleta (312 nm) e o perfil eletroforético foi registrado fotograficamente utilizando o sistema de vídeo BioChemi System (UVP BioImaging Systems) acoplada a equipamento de impressão.

Para purificação de fragmentos de DNA, foi utilizado gel de agarose de baixo ponto de fusão e tampão TAE 1X (Tris-acetato 40 mmol/L e EDTA 1mmol/L pH 8,0). As bandas de interesse foram coradas em solução de 0,5 µg/mL de brometo de etídeo por 20 min e cortadas sob iluminação luz ultravioleta. A extração de DNA do gel foi feita pela transferência da fatia de gel para um tubo de 1,5mL, a esta foi acrescentada solução de NaCl 5 M até uma concentração final de 250 mM e incubada a 65°C até a fusão do gel. Depois disso, foi adicionado ao tubo 1 volume de fenol tamponado, a solução foi misturada vigorosamente e centrifugada a 13.400 rpm por 10 min. A fase aquosa foi transferida para novo tubo e tratado com 1 volume de fenol:clorofórmio seguindo o procedimento descrito acima. A fase aquosa foi novamente transferida para um tubo e precipitada com 0,5 volume

de acetato de amônio e 2 volumes de iso-propanol absoluto, centrifugado, lavado com etanol 80%, seco e ressuspensão em 20 µL de água milli-Q.

3.2.4.4 Preparo de DNA com pontas não-coesivas (“blunt”)

Os fragmentos de DNA a serem ligados em terminais cegos foram tratados com a enzima T4 DNA polimerase, fragmento Klenow da DNA polimerase e T4 Polinucleotídeo quinase, conforme descrito por Sambrook e colaboradores (1989). A reação foi feita utilizando-se o tampão recomendado pelo fabricante (33 mmol/L Tris-acetato pH 7,9; 10 mmol/L de acetato de magnésio; 0,5 mmol/L de ditioneitol e 66 mmol/L de acetato de potássio), 0,3% de BSA (soroalbumina bovina), 0,1 mmol/L de dNTP's, 2mmol/L de ATP, 1U de T4 DNA polimerase, 1U de Klenow, 1U de polinucleotídeo quinase e 30µg DNA. A mistura foi incubada a 12°C durante 40 minutos, em seguida a 37°C durante 45 minutos e a 10 minutos a 75°C. Depois disso, a solução de DNA foi precipitada com 2 volumes de etanol absoluto, lavada com etanol 70%, seca, ressuspensa e aplicada em gel de baixo ponto de fusão para posterior purificação.

3.2.4.5 Digestão dos vetores

O vetor pUC18 foi linearizado com a enzima de restrição *Sma*I (New England Biolabs). Após a reação de 3h a 30°C em presença do tampão NEB 4, a enzima foi desnaturada por calor a 70° durante 10 min. O DNA foi extraído e precipitado com etanol absoluto, lavado, seco e dissolvido em água milli-Q estéril.

3.2.4.6 Ligação de DNA

O vetor linearizado foi ligado ao inserto numa proporção de 1:5 (mol:mol) utilizando tampão de ligação e 0,8 U T4 DNA ligase. A incubação foi feita a 18°C durante a noite.

3.2.4.7 Transformação por eletroporação

3.2.4.7.1 Preparo de células competentes

As células de *E. coli* foram preparadas utilizando-se um método baseado em Dower *et al.*, 1988 e Hanahan, 1983. Um mililitro de uma cultura saturada foi inoculado em 100 mL

de meio LB e crescida até uma D.O.₆₀₀ de 0,5-0,8. As células foram coletadas por centrifugação (5000 rpm por 5 min a 4°C) e lavadas duas vezes com água estéril gelada (100 e 50 mL). Em seguida, foram ressuspensas em glicerol 20%, centrifugadas novamente (10.000 rpm por 5 min) e ressuspensas em volume final de 0,5 mL de glicerol 20%, separadas em alíquotas de 30 µL e estocadas a -70°C.

3.2.4.7.2 Eletroporação

No momento do uso, uma alíquota de 30µL de células competentes foi descongelada em banho de gelo e, após acrescida de 1,0µL da solução de ligação, foi transferida para a cubeta de eletroporação previamente resfriada. A cubeta foi ajustada à câmara de eletroporação contendo gelo e submetida a um pulso elétrico de 10 kV/cm (330µF, 200Ω, 4kV). A seguir, as células foram transferidas para frascos estéreis de 10mL de meio Terrific broth (TB) e incubadas a 37°C, 130rpm por 30min. Após incubação, alíquotas de 200µL de cultura foram plaqueadas em meio LA ou MacConkey contendo os antibióticos apropriados para seleção dos transformantes. As placas foram incubadas a 37°C durante 18 horas. As colônias obtidas foram coletadas e seu DNA purificado para verificar a presença de plasmídeos conforme item 3.5.1.1.

3.2.4.8 Mini-preparação de plasmídeos em microplacas

Em um bloco estéril de 96 poços, foi adicionado a cada poço 1mL de meio Terrific Broth contendo 250 µg/mL de ampicilina. As colônias foram inoculadas com o auxílio de um repicador a partir de uma placa de LA inoculada no dia anterior e incubada a 37°C por cerca de 16 horas. Os blocos foram selados com adesivo, um pequeno orifício foi feito sobre cada poço para permitir a aeração e foram incubados a 37°C a 200rpm, durante 16 horas. As culturas foram centrifugadas a 4000 rpm por 7 minutos. O sobrenadante foi descartado e ao pellet foram adicionados 180µL de GET para lavar as células que foram sedimentadas por nova centrifugação a 4000 rpm por 7 minutos. O sobrenadante foi descartado e os blocos mantidos invertidos sobre papel absorvente.

A cada poço foi adicionado 80µL de GET contendo 0,5 µL de RNase (20mg/mL) e os blocos foram homogeneizados em vortex até a completa suspensão das células. As células foram transferidas para uma microplaca com poços de 250µL e lisadas com 80µL de solução de lise. As placas foram seladas, a solução homogeneizada por inversão e incubadas por 5 minutos à temperatura ambiente. Em seguida foram adicionados 80µL de acetato de potássio 3M, pH 5,2. As placas foram seladas, homogeneizadas por inversão e incubadas por 10 minutos à temperatura ambiente. Então, as placas foram centrifugadas por alguns segundos (4000 rpm) e posteriormente incubadas, sem o adesivo, em estufa a 90°C por 30 minutos. Após este período, as placas foram incubadas em gelo por 15 minutos e centrifugadas a 4000rpm por 10 minutos a 4°C. O sobrenadante foi filtrado em uma placa Millipore (MAGV N22) fixada, com fita adesiva, no topo de uma microplaca de polipropileno nova de fundo em V com poços de 250µL. O conjunto foi centrifugado a 4000 rpm por 5 minutos a 20°C.

A precipitação do DNA foi realizada adicionando-se 80µL de isopropanol ao filtrado. A placa foi homogeneizada por inversão e centrifugada por 45 minutos a 4000 rpm, 20°C. O sobrenadante foi descartado e o precipitado foi lavado com 150µL de etanol 70% e seco em estufa à 37°C por uma hora. O DNA foi dissolvido em 30µL de água milli-Q estéril. As placas foram armazenadas à -20°C.

3.2.4.9 Mutagênese com transposons

Para completar o seqüenciamento do plasmídeo pAB441 o transposon EZ::TN TET 1 foi integrado aleatoriamente nos plasmídeos pTZ+7 e pTZ+13. Como a seqüência das pontas do transposon é conhecida, oligonucleotídeos iniciadores podem ser utilizados para seqüenciar a partir das duas extremidades do TET-1. A reação de inserção foi feita seguindo-se o protocolo do EZ::TN™ Insertion Kit (Epicentre). A reação foi feita em volume total de 10µL, contendo 0,2µg de plasmídeo, 5fmol de EZ::TN <TET-1>

Transposon, 0,1U de tranposase, 1µL de tampão de reação e água estéril. Os clones obtidos foram purificados, analisados e utilizados como molde em reações de seqüenciamento, utilizando oligonucleotídeos iniciadores complementares às extremidades do transposon.

3.2.4.10 Minipreparação de plasmídeos

Preparações em pequena escala de plasmídeos para análise de restrição foram feitas pelo método de lise alcalina (SAMBROOK *et al.*, 1989) descrito abaixo.

A bactéria *E. coli* contendo o plasmídeo de interesse foi crescida em 5mL de meio LB contendo os antibióticos apropriados até uma D.O. de aproximadamente 2,0. A cultura foi transferida para tubos plásticos de 1,5 mL, as células coletadas por centrifugação (13.400 rpm, 1 minuto), ressuspensas em 100µL de solução GET (Tris-HCl 25 mmol/L, pH 8,0; glicose 50 mmol/L e EDTA 10 mmol/L) e lisadas com 150 µL de solução de lise (NaOH 0,2 mol/L e SDS 1%). As proteínas, o DNA cromossomal desnaturado, os restos celulares e o SDS foram precipitados com a adição de 150 µL de acetato de potássio 3 mol/L pH 5,5 e incubação por 10 minutos em gelo. As amostras foram, então, extraídas com 70 µL de fenol-clorofórmio-álcool isoamílico (25:24:1), centrifugadas a 13.400 rpm e o sobrenadante transferido para novo tubo. O DNA plasmidial foi precipitado com 1mL de etanol absoluto, lavado com 1 mL de etanol 80%, seco a vácuo e dissolvido em 20 µL de água milli-Q estéril. Mais tarde as amostras foram tratadas com RNase na concentração final de 500 µg/mL e analisadas em gel de ágar.

3.2.4.11 Digestão de DNA com enzimas de restrição

As condições utilizadas para a digestão de DNA com enzimas de restrição foram aquelas especificadas pelo fabricante.

Usualmente 1U das diferentes enzimas foi utilizada para digestão de 0,1 a 0,5 µg de DNA em um volume de 20µL durante 3 horas na presença do tampão adequado a 37°C.

3.2.5 Seqüenciamento

3.2.5.1 Purificação de DNA plasmidial

A purificação de DNA plasmidial para posterior seqüenciamento foi realizada através da lise alcalina conforme descrita no item 3.2.4.10. Após dissolvido em água, o DNA foi tratado com RNase (concentração final de 500µg/mL) por 1 hora a 37°C. Foi feita, em seguida, uma extração com 0,3 volume de fenol-clorofórmio-álcool isoamílico e uma extração com clorofórmio-álcool isoamílico. O DNA foi precipitado com 0,66 volume de acetato de amônio (7,5 M, pH 8,0) e 3 volumes de etanol absoluto, lavado com etanol 80%, seco e ressuspensão em 20µL de água.

3.2.5.2 Reação de Seqüenciamento

O seqüenciamento é baseado no método de terminação da cadeia por dideoxiribonucleotídeos (SANGER *et al.*, 1977). Neste trabalho foi utilizado entre 0,1 e 0,5 µg do DNA molde purificado, 5,0 pmol do oligonucleotídeo específico para a reação e 4,0 µL da mistura para seqüenciamento “DYEnamic ET sequence premix terminator” (DYEnamic ET DYE Terminator Kit: Thermo Sequenase II DNA polymerase, DYEnamic ET DYE terminators - Amersham Biosciences), num volume final de 10µL. A reação foi feita utilizando-se um termociclador onde, após um ciclo de desnaturação a 95 °C por 2 minutos foram realizados 35 ciclos a 96°C por 2 minutos foram realizados 35 ciclos a 96°C por 30 segundos e 60°C por 1’30”.

Após a reação de seqüenciamento o produto da reação foi purificado com 2µL de acetato de sódio 3M pH5,2 e 50µL de etanol absoluto, lavado com etanol 70%, seco e submetido à eletroforese no Seqüenciador Automático de DNA ABI377 (Perkin-Elmer).

3.2.5.3 Reação de seqüenciamento em placa de 96 poços

O seqüenciamento em placa foi realizado pelo método de terminação de cadeia utilizando dideoxiribonucleotídeos fluorescentes. As reações de seqüenciamento e purificação foram realizadas segundo o protocolo do kit DYEnamic ET Dye Terminator Cycle

Sequencing (Amersham Biosciences) e lidas em seqüenciador automático MegaBACE 1000 (Amersham Biosciences).

Os oligonucleotídeos iniciadores utilizados foram os apropriados para cada reação. Os parâmetros para a reação de seqüenciamento foram os mesmos do item anterior.

3.2.6 Análise dos dados de seqüenciamento

3.2.6.1 Análise e alinhamento das seqüências

Para o processamento e alinhamento das seqüências foram utilizados os programas PHRED/PHRAP/CONSED (EWING et al., 1998; EWING & GREEN, 1998; GORDON et al., 1998).

O programa PHRED reconhece as seqüências de nucleotídeos provenientes de dados brutos do seqüenciador e atribui valores de qualidade às bases da seqüência. Este processo é denominado identificação de bases ou “base calling” (EWING *et al.*, 1998).

Regiões de vetores de clonagem são filtradas pelo programa Cross-match, que compara a seqüência analisada com arquivos de seqüências de vetores (EWING *et al.*, 1998).

O programa PHRAP (Phragment Assembly Program) é o programa responsável pela montagem dos fragmentos de DNA seqüenciados em regiões contíguas (EWING & GREEN, 1998). Este programa permite a construção da seqüência do contíguo através de um mosaico de partes das seqüências de alta qualidade e a utilização de informações da qualidade dos dados computados internamente (EWING e GREEN, 1998). A visualização e edição das seqüências geradas após a montagem foram realizadas através do programa CONSED (GORDON et al., 1998).

3.2.6.2 Anotação da seqüência

Analisada a seqüência, a etapa seguinte consiste em identificar quais proteínas são codificadas. As ORF's, ou prováveis regiões codificadoras, presentes na seqüência foram identificadas utilizando o programa FramePlot (ISHIKAWA e HOTTA, 1999). Todas as

regiões que provavelmente codificam para proteínas foram comparadas com o banco de dados GeneBank utilizando o programa BLASTX (Basic Local Alignment Search Tool) (ALTSCHUL et al., 1997).

As análises de organização genômica foram feitas utilizando o programa STRING (von MERING et al., 2003) (<http://string.embl.de>).

As análises de domínios estruturais de proteínas foram feitas através do programa SMART (LETUNIC et al., 2004) (<http://smart.embl-heidelberg.de>).

3.3 Seção 2: Análise fisiológica dos mutantes HM26, HM053 e HM210 de *A. brasilense*

3.3.1 Condições de cultivo

O meio utilizado para o cultivo das estirpes FP2 e dos mutantes de *A. brasilense* foi o NFbHP (MACHADO et al., 1991).

3.3.2 Composição dos meios de cultura

A composição e o preparo estão descritos abaixo:

| | |
|--|-----------|
| MgSO ₄ .7H ₂ O | 0,2 g/L |
| NaCl | 0,1 g/L |
| Lactato de sódio | 5,0 g/L |
| CaCl ₂ | 20,0 g/L |
| Ácido nitrilo triacético | 56,0 mg/L |
| FeSO ₄ .7H ₂ O | 20,0 mg/L |
| Biotina | 0,1 mg/L |
| Na ₂ MoO ₄ .H ₂ O | 2,0 mg/L |
| MnSO ₄ .H ₂ O | 2,4 mg/L |
| H ₃ BO ₃ | 2,8 mg/L |
| CuSO ₄ .7H ₂ O | 0,08 mg/L |
| ZnSO ₄ .7H ₂ O | 0,24 mg/L |

| | |
|----------------------|---------|
| Água destilada q.s.p | 1000 mL |
|----------------------|---------|

O meio foi autoclavado por 20 minutos à 120 °C e 1 atmosfera de pressão.

Solução de fosfatos 20x:

| | |
|---------------------------------|----------|
| KH ₂ PO ₄ | 80,0 g/L |
|---------------------------------|----------|

| | |
|---------------------------------|-----------|
| K ₂ HPO ₄ | 120,0 g/L |
|---------------------------------|-----------|

| | |
|----------------------|---------|
| Água destilada q.s.p | 1000 mL |
|----------------------|---------|

O pH do meio foi ajustado para 6,8 através da adição de 50 mL/L da solução de fosfatos 20X. Como fontes de nitrogênio foram utilizadas soluções de glutamato de sódio 1 mol/L e de cloreto de amônio 1 mol/L. Estas soluções foram esterilizadas separadamente e adicionadas ao meio NFbHP no momento do uso em diferentes concentrações conforme o experimento. O meio acrescido destas fontes de nitrogênio foi denominado NFbHPG ou NFbHPN, respectivamente.

Nos ensaios de determinação de atividade da nitrogenase foi utilizado o meio NFbHP semi-sólido acrescido de 0,5 mM de glutamato ou 10 mM de cloreto de amônio. O meio NFbHP semi-sólido foi preparado adicionando-se 0,17% (m/v) de ágar ao meio líquido e o meio sólido foi preparado adicionando-se 1,5% (m/v) de ágar ao meio líquido.

3.3.3 Antibióticos

Os antibióticos usados foram:

| | Concentração final |
|------------------------|--------------------|
| Ácido Nalidíxico (Nal) | 20 µg/mL |
| Estreptomicina (Sm) | 80 µg/mL |
| Canamicina (Km) | 50 µg/mL |
| Tetraciclina (Tc) | 10 µg/mL |

3.3.4 Manipulação de DNA

3.3.4.1 Isolamento de plasmídeos de *A. brasilense*

O isolamento de plasmídeos de alta massa molecular das estirpes de *A. brasilense* foi feito conforme descrito abaixo (KADO & LIU,1981).

As estirpes transconjugantes e controles (FP2 e mutantes) de *A. brasilense* foram cultivadas em 10mL de meio NFbHPN até D.O.₆₀₀ de aproximadamente 1,0. Uma alíquota de 1,0 mL das culturas foi transferida para tubo de 1,5mL e centrifugada a 13.400 rpm durante 1 minuto, o sobrenadante foi descartado e as células ressuspensas em 100µL de tampão T₁₀E₁ (Tris-HCl 10mmol/L, EDTA 1mmol/L). Foram adicionados 200µL de solução de lise (Tris base 50mmol/L, SDS 3%, NaOH 82 mmol/L), homogeneizados por inversão suave e mantidas em repouso por 15 minutos. Após a adição de 800µL de fenol:clorofórmio:álcool isoamílico (25:24:1) a solução foi homogeneizada suavemente por inversão durante 5 minutos. Após centrifugação por 10 minutos, a fase aquosa foi transferida para tubo novo e o DNA foi precipitado com 2 volumes de etanol absoluto. O DNA foi coletado por centrifugação a 13.400 rpm por 20 minutos, lavado com etanol 70%, seco e ressuspensão em 15µL de água contendo 100 µg/mL de RNase. A solução de DNA resultante foi submetida a eletroforese em gel de agarose 0,8% em tampão TBE a 60 V por 6 horas para visualização dos plasmídeos.

3.3.4.2 Extração de DNA cromossomal

Células de uma cultura de *A. brasilense* (1,5 mL) D.O.₆₀₀= 1,5 foram coletadas por centrifugação a 13.400 rpm por 1 minuto, e ressuspensas em 500µL de tampão GET (5mmol/L de glucose, 25mmol/L de Tris-HCl pH8,0; 10mmol/L de EDTA pH8,0). Em seguida, foi acrescentado SDS 1% (v/v) e a solução foi incubada a 50°C por 20 minutos. Foram acrescentados então 200µg/mL de pronase E, e a solução foi incubada por aproximadamente 12 horas a 37°C. Após esse período, foram adicionados 200µL de fenol:clorofórmio:álcool isoamílico (25:24;1) e a solução foi homogeneizada por inversão

durante 5 minutos. Após centrifugação por 15 minutos a 13.400 rpm, o sobrenadante foi transferido para tubo novo, precipitado com 1 volume de isopropanol e centrifugado novamente a 13.400 rpm durante 20 minutos. O DNA foi lavado com etanol 80%, seco e ressuspensão em 50µL de água.

3.3.4.3 Ligação de DNA

A mistura vetor-inserto, numa proporção de 1:5 (mol:mol), foi aquecida a 65°C por 5 minutos e resfriada em gelo, em seguida, foi adicionado tampão de ligação e 0,8 U T4 DNA ligase. A incubação foi feita a 18°C durante a noite. (SAMBROOK et al., 1989)

3.4 Transferência de plasmídeo por conjugação

Os plasmídeos a serem transferidos para as estirpes de *A. brasilense* foram mantidos em *E. coli* S17.1 (tra⁺). As estirpes de *A. brasilense* (receptoras) e *E. coli* foram cultivadas nos meios NFbLactatoHPN e LB, respectivamente, a partir de uma cultura fresca, na ausência de antibióticos, até uma D.O₆₀₀ de aproximadamente 0,7. Células das estirpes doadora e receptora foram misturadas na proporção de 1:10 (100µL da estirpe doadora para 1mL da estirpe receptora). A mistura foi centrifugada em tubo plástico de 1,5 mL estéril, as células foram ressuspensas em 200µL de NFb e plaqueadas em meio sólido NFbHPN/LA 1:1. As placas foram incubadas a 30°C por 24 horas. A massa de células formada foi coletada com palito estéril e transferida assepticamente para um tubo plástico de 1,5mL estéril, ressuspensas em 1,0 mL do meio líquido NFbHPN com auxílio da pipeta. Uma alíquota de 100µL da suspensão foi plaqueada em NFbHPN contendo os antibióticos necessários para seleção. As colônias transconjugantes foram transferidas para placas contendo meio NFbHPN com os antibióticos, analisadas quanto a presença dos plasmídeos (KADO & LIU, 1981), e mantidas em meio NFbHPN sólido em temperatura ambiente.

3.5 Dosagem de proteínas

A dosagem de proteínas das culturas de *A. brasilense* foi feita pelo método de LOWRY et al., (1951), após lise alcalina das células utilizando NaOH 0,5 mol/L por 30 minutos à temperatura ambiente. Foi utilizada uma curva padrão de soroalbumina bovina e a absorbância foi determinada a 660nm.

3.6 Atividade da Nitrogenase em Meio Semi-sólido

A atividade da nitrogenase foi determinada em culturas crescidas em 4mL de meio semi-sólido, após 24-48 horas de incubação a 30°C ou após formação de película na superfície do meio. Os frascos foram fechados com rolha de borracha e foi injetado acetileno no volume de 10% da fase gasosa. Em seguida, as culturas foram incubadas novamente a 30°C por 1 hora e o etileno formado foi quantificado por cromatografia gasosa. Para isso, alíquotas de 0,5 mL foram retiradas da fase gasosa e injetadas em cromatógrafo (Cromatógrafo Varian, modelo 3400, equipado com coluna Porapak N e detector de ionização de chama) (DILWORTH, 1966; SCHOLHORN e BURRIS, 1967). Nitrogênio superseco foi o gás de arraste empregado a um fluxo de 20mL/min, temperatura da coluna de 130°C, do detector de 200°C. Como padrão foi utilizado etileno 100 ppm.

A atividade específica da nitrogenase foi expressa em nanomoles de etileno formado por miligrama de proteína por minuto.

3.7 Seqüenciamento do operon *glnBA* dos mutantes HM26, HM053 e HM210

3.7.1 Amplificação de DNA por PCR

A região *glnBA* dos mutantes HM26, HM053 e HM210 foi amplificada a partir de seu DNA cromossomal, purificado conforme o item 3.3.4.2. A amplificação foi realizada utilizando o sistema de reação em cadeia da polimerase (PCR) (KOCHER e WILSON, 1991). Em um tubo de 200µL, foram acrescentados aproximadamente 20ng de DNA molde, tampão de PCR (20mmol/L Tris HCl pH 8,4 e 50mmol/L KCl), 0,2 mmol/L dNTP's, 10pmol de cada oligonucleotídeo, 1,5mmol/L MgCl₂, e 2,5U de *Taq* DNA

polimerase (Life Technologies) em um volume final de reação de 25 μ L. A reação foi feita utilizando termociclador Thermocycler (epENDORF) e os parâmetros estão descritos a seguir: um período de desnaturação a 95°C por 2 min, 30 ciclos com as seguintes temperaturas: 94°C por 30 seg, T_{ann} por 30 seg, 72°C por 3 min e extensão adicional dos produtos de PCR a 72°C durante 5 min, onde T_{ann} é a temperatura de anelamento dos oligonucleotídeos iniciadores (“primers”), que varia conforme a combinação utilizada, como mostra a tabela que segue:

| Combinação | T_{ann} |
|---------------------|-----------|
| p2glnB x promglnB | 63°C |
| pglnA1 x internoGS2 | 65°C |
| promglnB x pglnA2 | 60°C |
| pglnA1 x AbGSfim | 60°C |

A região de anelamento dos “primers”, suas seqüências e temperaturas de anelamento calculadas são mostradas na figura 3 e tabela 2.

REGIÃO *glnBA* DE *A. brasilense* E LOCALIZAÇÃO DOS PRIMERS

-660

CGCCTCCACCCGCCCAATCCCGCATCACCGTCACCCGGCACAATAAGCACCGCGACGGCCGCAGCATAGCCGGGC
 promglnB 5' GCACAATAAGCACCGCGA 3'

ACAGGGGGAATGAAAGGACTGGCCTGATTGGCACAGGAATTGGAAGAAAGGACTGGGGCGATGGATGGGACTCGAA
 CCCACGACCGCTCGGACCACAACCGAGAGCTCTACCAACTGAGCTACCACCGCCACCGGGGACGAAAAGATTTCGTC
 CGCAAAACACCCGACCCGTCGCCCGACCGATGCAACGCCGGCGTGAAGTCTCACGAAACCGACACATCCGCCGTC
 AAGGAGGACTGGGGCGATGGATGGGACTCGAACCACGACCGCTCGGACCACAACCGAGAGCTCTACCAACTGAGC
 -310

TACCACCGCCACCGGGTTCGCTATTTGCCCAAGGGACGGGGCGGGTCAAGCTCTATTTTCCACCTTGTCCATCA
 upstream 3' GAACGTCGTACAA

ATGTGAAAAGCCGGGTGGAAAGCCGGCGCTACCCCATCCGAACGGACAAGGCTGCGCGGATTGCTTGCAGCATGTG
 ATGTTTATAACAA 5' downstream 5' CTTGCAGCATGTT
CACAAATATTGTTCACTTCGCCCTATCCGTTGGCAGTCGCGTGCCGCCACGGTGGCATCGGCCCGCTGCGGCCG
 TACAAATATTGTT 3' p2glnB 3' CCTGGCAGCCGCTCTCGGCA

GCATCCACGCGACCCGGATCGCAACCGCTCCGATTGGCAGCAACGTGCTTTACATCGGACCGTCGGCAGAGCCGT
 SigmaN 5' CCGCTCCGATTGGCAGCAACGTG 3' -60

CCTG 5'
GGACGGTGCCTGCCGACGTTGGCACCTCGTACACGAGAGACCCATGAAGAAGATCGAAGCCATCATTAAGCCGTTG
 AAACTCGACGAAGTGAAGGAAGCCCTTCACGAAGTCGGCATCAAGGGCATCACCGTCACCGAGGCCAAGGGCTTCG
 GCCGTCAGAAGGGGCACACCGAGCTGTACC GCGCGCGGAGTATGTGGTCGACTTCCTGCCGAAGGTGAAGATCGA
 GGTGGTGATGGAGGACTCCCTGGTGGAGCGGGCGATCGAGGCGATCCAGCAGGCCGCCACACCGGCCGCATCGGC
 GACGGCAAGATCTTCGTCACCCCGTGGAGAAGTTGTCCGCATCCGGACCGGCGAGAAGGGAGGCGACGCGATCT
 GATCGCGTCCCGAGAACC CGACCACCGGCCACCATGAACCAGCCCCGCCACAACCGACGCAGAGCCGGCCGACTC
 pglA1 5' -AAGCGTCTAGAGAACC CGACC-3'

CGATCGGGCCGACGCAGACCGTCGCAACCGGAGCGGAACGATCCGTGCGTGGATGCCCATCCGGGCGCGACTTTT
 TGTGCTGTACGTTGTGTGCTTGGGAAAGGCATGACATAACGTCCGCCGCGAGCGCCGTACCCGCCCTTGGTGC
 GS+ 5' GTGAATTCTTGGGAAAGGCATGACATAACG 3' pglA2 3' CAGAAGCTGGAC
 GTTTGACGCAGCGGGACCCAAACCTCTTAAGCGTGGAAAAAGAGATATGTCCGACATCAGCAAGGTCTTCGACCTG
 TAGATCTTTGTG 5'
ATCAAGGAACACGACGTCAAGTACGTGGACCTCCGCTTCACCGACCGCGCGGCAAGCTGCACCACACCGCTCAGC
 ACGTCTCGACCATCGACGAGGACGTGTTTGAAGACGGCATCATGTTTCGACGGCTCCTCGATCGCCGTTGGAAGGC
 GATCAACGAGTCGGACATGATCCTCCAGCTCGACCCGACCACCGCCGTCATGGACCCGTTTTTCGGCCAGCCGACG
 CTGAACATCCTCTGCGACGTGTACGAGCCGTCGACCGCCAGCCCTACGCCCGCTGCCCGCGCGGTATCGCGAAGG
 CCGCCGAGAAGTACATGGCGTCGGCCGGCATCGGCGACACCGCCTACTTCGGCCCGGAAGCCGAGTTCTTCGTCTT
 CGACGACGTCAAGTTCAAGGTCGAGATGAACAAGGTGTCTACGAGTTCGACTCGGAGGAAGCCCGTACACCTCG
 GACAAGGACTATGAGGACGGCAACCTGGGCCACCGCCCGGGCGTCAAGGGCGGCTACTTCCCGGTCGCTCCGGTGC
 ACAGCGGCTCGGACCTGCGCGCCGAGATGCTGAGCGTGCTCGCCGAGATGGGCGTGCCGGTCGAGAAGCACCACCA
 internoGS1 3' CCGGCGAAGCGTCGTACT 5'
 CGAGGTGGCCGCTTCGACGATGAGCTGGGCATCAAGTTCGACACGCTGGTCCGCACCGGCGACAACATGCAGTAC
 TACAAGTACGTGTCACAATGTCGCCACGCTACGGCAAGACCGGACCTTCATGCCGAAGCCCGTCTTCGGCG
 ACAACGGCTCGGGCATGCACATGCACCAGTCGATCTGGAAGGAAGGCCAGCCGCTGTTTCGCCGGAACCAAGTACGC

CGACCTGTCGGAAGTGGCGCTGTACTACATCGGGCGCATCATCAAGCACGCCAAGGCGCTGAACGCCTTCACCAAC
 CCGACGACCAACTCGTACAAGCGTCTGGTCCCGGGCTACGAGGCTCCGGTTCTGCTGGCCTACTCGGCCCGCAACC
 internoGS2 3' CCCGTTTCGCGCAGCTCCA 5'
 GTTCGGCCTCCTGCCGCATCCCCTACGTCGCCTCGCCGAAGGGCAAGCGCTCGAGGTCCGCTTCCCGACCCGTC
 GGCCAACCCGTACCTGGCCTTCGCCGCCCTGCTGATGGCCGGTCTGGACGGCATCCAGAACAAGATCCATCCGGGC
 GAGGCGATGGACAAGAACCTGTACGACCTGCCGGCCGAAGAGCTGGCCAAGGTTCCGACGGTTTGGCGCTCGCTGC
 GCGAGGCCCTGGACAGCCTGAAGGCCGACAGCGCCTTCCTGCAGAAGGGCGACGTGTTACCAAGGACATGATCGA
 GTCCTACATCGACCTGCGCACCGAGGAGCTGCTGGCCTTCGAGACCATGCCGCACCCGATCGAGTACAAGATGTAC
 GS- 3' GCCTAAGGCCACAGGCGGGCTTAAGTG 5'
 TACTCGGTCTGA**TCCATCGGACTTCGGTCCACCGGATTCCGGTGTCCGCCCGCCCTCG**CGGCGCGGGCGGGCACCA
 AbGSFIM 3' GCCGCCCTAGGT
 CCGGAGGGCCCGGACCGCGGAGCAGAGCACAAAAAAGCCACCGGCGACACCGTCCGGTGGGCTTTTTTGTGTCCAG
 GGCTCC 5'
 GACGTGAAGAGATGCC

Figura 3: Seqüência correspondente à região *glnBA* de *A. brasilense*. As letras em verde representam à região codificadora do gene *glnB* e as letras em vermelho correspondem à região codificadora do gene *glnA*. Os locais de anelamento dos oligonucleotídeos iniciadores estão indicados por letras azuis. E as letras sublinhadas representam trocas de base para inserção de sítios de restrição. As letras em negrito e em maior tamanho delimitam as regiões em que foram obtidas seqüências confiáveis (conforme o texto).

TABELA 2: SEQUÊNCIA DOS PRIMERS E SUAS TEMPERATURAS DE ANELAMENTO

| Primer | Seqüência | T_{ann} calculada |
|--------------------------|--------------------------------------|--------------------------------------|
| promglnB | 5' -GCACAATAAGCACCGCGA-3' | 56°C |
| Mutagênico upstream | 5' -AACAAATATTTGTAAACATGCTGCAAG-3' | 68°C |
| Mutagênico downstream | 5' -CTTGCAGCATGTTTACAAATATTGTT-3' | 68°C |
| sigmaN | 5' -CCGCTCCGATTGGCACGCAACGTG-3' | 80°C |
| p2glnB | 5' -GTCCACGGCTCTGCCGACGGTCC-3' | 60°C |
| pglnA1 | 5' -AAGCGTCTAGAGAACCCGACCACC-3' | 66°C |
| pglnA2 | 5' -GTGTTTCTAGATCAGGTCTGAAGAC-3' | 60°C |
| GS+ | 5' -TGAATTCTTGGGAAAGGCATGACATAACG-3' | 61°C |
| GS- | 5' -GTGAATTCGGGCGGACACCGGAATCCG-3' | 48°C |
| internoGS1 | 5' -TCATGCTGCGAAGCGGCC-3' | 60°C |
| internoGS2 | 5' -ACCTCGACGCGCTTGCCC-3' | 62°C |
| AbGSFIM | 5' -CCTCCGGTGGATCCCGCCG-3' | 58°C |

3.7.2 Seqüenciamento

Foi feito conforme item 3.2.5.

3.7.3 Análise das seqüências

As seqüências obtidas foram comparadas utilizando o programa BioEdit (HALL, 1999).

4. RESULTADOS E DISCUSSÃO

Estudos de complementação genética dos mutantes *Nif⁻* de *A. brasilense*, denominados HM26, HM053 e HM210 (MACHADO, 1988), utilizando o plasmídeo pAB441 resultaram na perda da atividade da nitrogenase nos transconjugantes tanto na ausência como na presença de amônio (VITORINO, 2001). Este efeito deletério sobre a atividade da nitrogenase poderia ser decorrente da presença de outros genes regulatórios desconhecidos, incluindo os genes *glnBA*, presentes no plasmídeo pAB441. Com o objetivo de identificar o gene ou conjunto de genes responsáveis por este efeito, uma biblioteca genômica (aproximadamente 900 clones) do plasmídeo pAB441 no vetor pUC18 foi construída e seqüenciada tendo como resultado a seqüência completa do inserto de DNA de *A. brasilense* no pAB441 (Figura 5).

A biblioteca aleatória do plasmídeo pAB441 foi construída utilizando fragmentos de 1 a 3 kb do DNA nebulizado clonados em vetor pUC18. Após a clonagem, os plasmídeos foram transformados em *E. coli*, estirpe TOP10. As colônias transformantes foram coletadas e repicadas em placas de 96 poços. Estas colônias foram crescidas, seu DNA foi purificado e submetido a reações de sequenciamento.

4.1 Seqüenciamento do inserto do plasmídeo pAB441

Aproximadamente 2100 seqüências foram obtidas e analisadas. O alinhamento utilizando os programas Phred/Phrap e Consed resultou em uma região de 21606 pares de base que corresponde ao inserto do plasmídeo pAB441 e a regiões adjacentes do vetor pVK100. A análise mais detalhada desta região revelou que o tamanho do inserto do plasmídeo pAB441 é de 19173 pares de base (Figura 4 – Anexo 1). Utilizando o programa FramePlot (ISHIKAWA e HOTTA, 1999), foi possível identificar 15 prováveis seqüências codificadoras de proteínas (ORF's) nesta região. A comparação das seqüências dos produtos de tradução das ORF's com aquelas depositadas no banco de dados GenBank,

através do programa BLASTP, permitiu identificar as prováveis proteínas codificadas por estes genes. As proteínas identificadas apresentam similaridade superior a 50% com as seqüências de outros organismos já depositadas no GenBank (Tabela 3).

As proteínas codificadas por genes dessa região apresentam diferentes funções: três são proteases dependentes de ATP: protease Lon, ClpP e ClpX; duas estão envolvidas na fosforilação oxidativa: subunidades α - e β - respectivamente da NADH-desidrogenase I; duas no metabolismo de nitrogênio: proteína GlnB (PII) e glutamina sintetase; uma prolil isomerase: o fator “trigger”; uma proteína ligante a DNA: a subunidade β - da proteína HU; uma chaperona: a proteína de choque térmico Hsp70; uma provável monoamina oxidase: MaoC; uma provável histona acetiltransferase: Hpa2; e cinco proteínas hipotéticas: as ORF's com alta homologia a Magn029418 e Magn029417, uma provável quinase de carboidratos, uma proteína similar a proteínas envolvidas na síntese de fenazina, uma ORF conservada (COG5457) e parte de uma ORF que codifica para uma proteína de membrana (COG4325).

Uma parte da seqüência, já obtida por outros grupos, estava depositada no banco de dados GenBank: a região da protease Lon e do fator “trigger” com aproximadamente 4 kb, além das seqüências dos genes *glnB* e *glnA* (MORI et al., 1996; INDORATO et al., 1999; de ZAMAROCZY et al., 1990).

As prováveis funções das proteínas codificadas pelos genes encontrados no inserto do pAB441 são descritas abaixo.

COG4325

Esta ORF apresenta similaridade com uma região conservada que corresponde a uma provável proteína de membrana de função desconhecida. Este domínio é encontrado em 4 espécies do filo Actinobacteria. Apenas parte desta ORF está presente no inserto do plasmídeo pAB441 e codifica para 268 aminoácidos da região C-terminal desta proteína.

Esta seqüência de aminoácidos apresenta 51% de similaridade com a seqüência codificada pela ORF correspondente em *Nostoc punctiforme* (Tabela 3).

COG5457

Esta região apresenta alta similaridade com a região codificadora para uma proteína pequena, conservada e não caracterizada. Esta proteína conservada é encontrada em 5 espécies da subdivisão gama e em 4 espécies da subdivisão alfa das proteobactérias. Em *A. brasilense* esta proteína conservada apresenta 71 resíduos de aminoácidos e 70% de similaridade com a seqüência da proteína de *Pseudomonas fluorescens* (Tabela 3).

Acetiltransferase Hpa2

Hpa2 é membro da superfamília de N-acetiltransferases, uma família de enzimas com diversos substratos, inclusive histonas, outras proteínas e aminoglicosídeos. Estudos *in vitro* mostraram que a Hpa2 de *Saccharomyces cerevisiae* é capaz de acetilar resíduos específicos de lisina em histonas H3 e H4. Hpa2 forma um dímero estável em solução e forma tetrameros quando ligada a acetil-CoA (ANGUS-HILL et al., 1999). Em *Salmonella enterica* foi identificada uma N-acetiltransferase denominada AAC(6')-ly, que forma dímeros com estrutura similar à de Hpa2. Esta enzima catalisa sua auto N-acetilação dependente de acetil-CoA e a acetilação de histonas de eucariotos. Em bactérias gram-negativas patogênicas as N-acetiltransferases conferem resistência a aminoglicosídeos utilizados como antibióticos, sendo responsável pela maioria das estirpes resistentes a estes compostos (VETTING et al., 2004). A provável proteína Hpa2 de *A. brasilense* apresenta 130 resíduos de aminoácidos e 70% de similaridade com a de *Magnetospirillum magnetotacticum* (Tabela 3).

MAPA FÍSICO E GENÉTICO DO INSERTO DO PLASMÍDEO pAB441

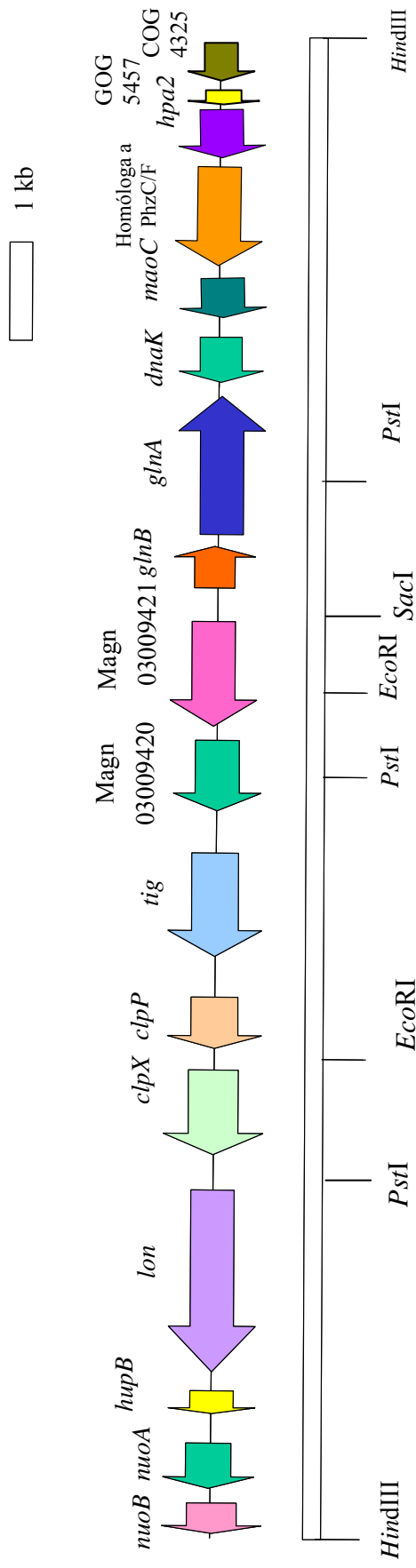


Figura 5. Mapa físico e genético da região *glnBA* de *A. brasilense* clonada no plasmídeo pAB441. Além do operon *glnBA* quatorze ORF's completas e uma incompleta foram localizadas no inserto de 19,6kb.

TABELA 3: SIMILARIDADE DE SEQÜÊNCIA ENTRE AS PROTEÍNAS CODIFICADAS PELAS ORF`s DO PLASMÍDEO pAB441 E SEQÜÊNCIAS DE PROTEÍNAS DO GENBANK

| Proteína | Resíduos | Similaridade | Identidade | Organismos |
|--------------------|-----------------|---------------------|-------------------|---|
| NuoB | 181 | 76% | 67% | <i>Rhodospirillum rubrum</i> |
| NuoA | 118 | 88% | 72% | <i>Rhodospirillum rubrum</i> |
| HU | 90 | 86% | 74% | <i>Rhodospirillum rubrum</i> |
| Lon | 810 | 87% | 77% | <i>Rhodospirillum rubrum</i> |
| ClpX | 422 | 91% | 83% | <i>Magnetospirillum magnetotacticum</i> |
| ClpP | 210 | 82% | 73% | <i>Magnetospirillum magnetotacticum</i> |
| TF | 444 | 69% | 50% | <i>Rhodospirillum rubrum</i> |
| Magn03009420 | 476 | 60% | 50% | <i>Magnetospirillum magnetotacticum</i> |
| Magn03009421 | 433 | 60% | 51% | <i>Magnetospirillum magnetotacticum</i> |
| PII | 117 | 94% | 87% | <i>Rhodospirillum rubrum</i> |
| GS | 413 | 84% | 74% | <i>Rhodospirillum rubrum</i> |
| DnaK | 276 | 66% | 56% | <i>Magnetospirillum magnetotacticum</i> |
| MaoC | 144 | 69% | 50% | <i>Dechloromonas aromática</i> |
| Homóloga PhzC/PhzF | 221 | 64% | 51% | <i>Pseudomonas fluorescens</i> |
| Hpa2 | 130 | 70% | 60% | <i>Magnetospirillum magnetotacticum</i> |
| COG4325 | 268 | 51% | 36% | <i>Nostoc punctiforme</i> |
| COG5457 | 71 | 70% | 59% | <i>Pseudomonas fluorescens</i> |

Epimerase homóloga a PhzC/PhzF

Esta ORF apresentou alta similaridade com as proteínas do grupo de proteínas ortólogas COG0384 e da família de proteínas pfam02567, que corresponde a uma epimerase homóloga às proteínas PhzC e PhzF. Estas proteínas participam de dois dos passos da síntese de fenazinas. Fenazinas são moléculas heterocíclicas que contêm nitrogênio, com atividade antibiótica de amplo espectro contra várias espécies de bactérias e fungos (SMIRNOV e KIPRIANOVA, 1990). A proteína encontrada é similar à outra depositada no banco de dados CDD (MARCHLER-BAUER et al., 2005). Esta está depositada sob os códigos Pfam 05544 e COG0253. Pfam0544 corresponde à família das prolina racemases. Estas enzimas catalisam a interconversão de L- e D- prolina em bactérias. A racemização de aminoácidos é importante para fornecer substratos para a construção de peptideoglicanas que contêm resíduos de D-aminoácidos. COG0253 corresponde à diaminopimelato epimerase. Esta enzima cataliza a epimerização da forma L,L- para a D,L-*meso*-diaminopimelato, precursora imediata da L-lisina em bactérias (CIRILLI et al., 1998). Em *A. brasilense*, esta provável proteína apresenta 221 aminoácidos com similaridade de 64% à de *Pseudomonas fluorescens* (Tabela 3).

Monoamina oxidase MaoC

A síntese de monoamina oxidase em bactérias é controlada por carbono, nitrogênio e compostos aromáticos (OKAMURA et al., 1977).

Monoamina oxidases catalizam a deaminação de monoaminas, como tiramina e catecolaminas a aldeído, peróxido de hidrogênio e amônia. Esta enzima apresenta especificidade por vários substratos, e seu papel fisiológico não está bem definido, embora ela esteja provavelmente envolvida no metabolismo de aminas biogênicas (YAMASHITA et al., 1996).

Em *Klebsiella aerogenes*, esta enzima é altamente específica para tiramina, β -feniletilamina e dopamina (YAMASHITA et al., 1993). O gene que codifica para monoamina oxidase, neste microrganismo, é o *maoA*. Este gene encontra-se no operon *mao* que inclui o gene *maoC*, o qual tem função desconhecida, mas também é induzido por tiramina e monoaminas relacionadas (SUGINO et al., 1992). Esta enzima, em *A. brasilense* é composta por 144 aminoácidos e apresenta similaridade de 69% com a seqüência correspondente de *Dechloromonas aromatica* (Tabela 3).

DnaK

DnaK é uma proteína induzida por estresse, homóloga às chaperonas do tipo Hsp70. As proteínas de choque térmico do tipo Hsp70, em conjunto com co-chaperonas, são necessárias para o dobramento de proteínas citosólicas, translocação de proteínas através de membranas, união e separação de oligômeros, re-dobramento de proteínas desnaturadas, degradação de proteínas instáveis e controle da atividade de proteínas regulatórias. Estes papéis se baseiam em sua habilidade de se associar a segmentos hidrofóbicos de polipeptídeos desnaturados de maneira ATP-dependente (HESTERKAMP e BUKAU, 1998).

Em *E. coli*, é provável que um sistema composto por DnaK (Hsp70) e pelas co-chaperonas DnaJ e GrpE seja responsável pelo dobramento das proteínas citosólicas durante a tradução (HENDRICK et al., 1993). De acordo com o modelo proposto, DnaJ é a primeira chaperona a interagir com o polipeptídeo nascente, prevenindo o dobramento incorreto e direcionando DnaK para seu substrato. Através de ciclos de ligação e liberação de DnaK, controlados por ATP, a cadeia polipeptídica alcança seu estado nativo ou é transferida para GroEL para o dobramento final (FRYDMAN e HARTL, 1996). A seqüência deduzida de aminoácidos para esta proteína em *A. brasilense* apresenta 66% de similaridade com a de *M. magnetotacticum* com extensão de 276 resíduos de aminoácidos.

Magn03009420

Esta é uma ORF conservada encontrada em vários microrganismos. A ORF de *A. brasilense* apresenta 60% de similaridade com a ORF de *M. magnetotaticum*, mas não se sabe a provável função desempenhada pela proteína codificada (Tabela 3). Em *A. brasilense* esta ORF codifica para uma proteína de 433 aminoácidos e a análise utilizando o programa Blast identificou similaridade com a família pfam01256, que representa uma família de carboidrato quinases. O programa SMART (LETUNIC et al., 2004) indicou apenas a presença de regiões de baixa complexidade.

Magn03009421

Esta ORF é encontrada à montante da ORF Magn030009420 tanto em *M. magnetotaticum* como em *A. brasilense*, a similaridade entre as duas é de 60%, assim, esta também uma ORF conservada (Tabela 3). A exemplo da primeira, também não se sabe a função da proteína de 473 aminoácidos. A análise utilizando o programa SMART resultou na indicação de regiões de baixa complexidade.

“Fator Trigger”

O “trigger factor” (TF) em *E. coli* é uma prolil isomerase que acelera a isomerização dos resíduos de prolina, durante a aquisição da conformação tridimensional das proteínas recém-sintetizadas. As proteínas em seu estado desnaturado apresentam seus resíduos de prolina nas conformações *cis* e *trans*, mas, em sua forma nativa, a conformação destes resíduos é geralmente fixa em um dos estados (DODGE e SCHERAGA, 1996). O TF se associa a cadeias de polipeptídeos nascentes no ribossomo e acredita-se que catalise o seu dobramento num processo conjunto com diversas chaperonas (SCHOLZ et al., 1997).

A proteína TF também é uma proteína de choque térmico por baixas temperaturas e é essencial para a viabilidade celular nesta condição. A indução de TF por frio pode representar uma adaptação que permite a continuação da síntese de proteínas e a aquisição

de sua conformação sob baixas temperaturas, condição na qual a solubilidade, agregação, e propriedades conformacionais das proteínas são diferentes comparadas a temperaturas normais (KANDROR e GOLDBERG, 1997). Além disso, TF pode ser importante na manutenção de proteínas celulares pré-existentes na sua conformação funcional, promovendo a re-dobramento de proteínas prejudicadas pelo frio. Este tipo de manutenção ou reparo pode ser importante sob baixas temperaturas, pois nessas condições a síntese de novas proteínas é lenta (KANDROR e GOLDBERG, 1997).

O fator “trigger” de *A. brasilense*, identificado neste trabalho, é uma proteína de 444 aminoácidos com similaridade de 69% com a de *Rhodospirillum rubrum* (Tabela 3).

As proteases dependentes de ATP Lon, ClpP e ClpX

As proteases dependentes de ATP estão envolvidas na regulação da quantidade de várias proteínas normais no interior da célula (GOTTESMAN e MAURIZI, 1992). Além disso, muitas proteínas incorretamente dobradas são degradadas por estas proteases (NEIDHARDT e VANBOGELEN, 1987). A protease Lon (ou La) de *Escherichia coli*, codificada pelo gene *lon*, é induzida por choque térmico, situação na qual sua transcrição aumenta várias vezes. A preferência da protease Lon por proteínas anormais como substrato e sua indução por choque térmico leva à hipótese de que ela pode ajudar a eliminar proteínas desnaturadas em condições de estresse (PARSELL e LINDQUIST, 1993). O gene que codifica para esta protease em *Azospirillum brasilense* foi caracterizado por Mori e colaboradores (1996) e uma ORF homóloga ao gene *clpX* de *E. coli* foi encontrada à montante deste gene. Uma estirpe de *A. brasilense* mutante para o gene *lon*, obtida por mutação com Tn5, apresenta crescimento normal mas é afetada na resposta a choque térmico e na captação de ferro. Os dados de Mori e colaboradores sugerem que o gene *lon* esteja envolvido na síntese (ou na regulação da síntese) de uma ou mais proteínas de membrana induzidas por deficiência de ferro. Quanto à resposta ao choque térmico, os

mutantes *lon* apresentam um fenótipo inesperado, pois a taxa de sobrevivência dos mutantes é maior do que a da estirpe selvagem após exposição a altas temperaturas (MORI et al., 1996).

O seqüenciamento, neste trabalho, da região à montante do gene para ClpX resultou na identificação do gene para ClpP e sua provável região regulatória. Assim como em *E. coli* e outros organismos, estes dois genes estão aparentemente organizados em um operon em *A. brasilense*. Foi identificado um provável promotor dependente da subunidade σ^{32} da RNA polimerase à montante do gene *clpP* e não foi identificado terminador ou promotor na região intergênica, ao passo que um provável terminador independente de Rho foi encontrado na região à jusante do gene *clpX*.

A proteína ClpP de *A. brasilense* é constituída por 210 aminoácidos e a comparação de sua estrutura primária com a de *M. magnetotaticum* indica que existe 82% de similaridade entre elas. Já a proteína ClpX é constituída por 422 aminoácidos e a comparação de seqüências mostra que a similaridade entre a de *A. brasilense* e a de *M. magnetotaticum* é de 91% (Tabela 3). A protease Lon de *A. brasilense* é constituída por 810 aminoácidos e apresenta 87% de similaridade com a de *Rhodospirillum rubrum*.

Proteína HU

A proteína HU é uma proteína com função similar às das histonas dos eucariotos, que também contribui para a compactação do genoma em estruturas semelhantes a nucleossomos (ROUVIERE-YANIV et al., 1979). A HU de *E. coli* é uma proteína dimérica, pequena e de caráter básico, composta por duas subunidades, HU α e HU β , codificadas pelos genes *hupA* e *hupB*, localizados a 90 e 10 min, respectivamente no cromossomo de *E. coli* (KANO et al., 1988; KANO et al., 1986).

A proteína HU também participa em um grande número de mecanismos celulares como a modulação da expressão de genes específicos (MANNA e GOWRISHANKAR,

1994; WU e DATTA, 1995), a abertura do DNA na iniciação da replicação (KANO et al., 1991; SHARPLES et al., 1994), a cisão e religação do DNA em reações de transposição e inversão (HAYKINSON e JOHNSON, 1993; LAVOIE e CHACONAS, 1993) e recombinação homóloga (DRI et al., 1992; KANO e IMAMOTO, 1990). Além disso, apesar de a proteína HU não reconhecer uma seqüência específica de DNA, ela pode agir em locais muito precisos no DNA cromossomal, ligando-se a estruturas secundárias particulares do DNA (BONNEFOY et de ligação para a al., 1994).

Em *E. coli*, a subunidade β da proteína HU tem sua expressão estimulada durante choque térmico por frio, enquanto a expressão da sua subunidade α diminui consideravelmente. Isto indica que a razão alterada na expressão das subunidades, que resulta num aumento de heterodímeros HU α /HU β ou homodímeros HU β , pode ter um papel importante na adaptação ao choque térmico por frio (GIANGROSSI et al., 2002).

A região promotora do gene que codifica para a proteína HU apresenta um provável sítio subunidade σ^{70} da RNA polimerase. Diferentemente do que ocorre em *E. coli*, (CLARET e ROUVIÈRE-YANIV, 1996) não foi encontrado nesta região um sítio de ligação para CRP. Também foi identificada na seqüência obtida uma provável região terminadora independente do fator Rho. A proteína HU de *A. brasilense* apresenta 90 aminoácidos com 86% de similaridade com *Rhodospirillum rubrum*. A análise utilizando o programa SMART resultou na identificação de uma região conservada que corresponde à toda extensão da proteína. Esta região é identificada como semelhante às de histonas, como já esperado.

NADH-desidrogenase

A NADH desidrogenase corresponde ao complexo I da cadeia respiratória. As desidrogenases coletam elétrons das reações oxidativas do complexo da piruvato desidrogenase, do ciclo do ácido cítrico, da via da β -oxidação e das etapas oxidativas do

catabolismo dos aminoácidos e os canalizam na forma de pares de elétrons para a cadeia respiratória. Elas usam piridina nucleotídeos (NAD ou NADP) ou flavina nucleotídeos (FMN ou FAD) como receptores de elétrons. A maioria das desidrogenases é específica para o NAD^+ como receptor de elétrons, mas algumas requerem NADP^+ . A glutamato desidrogenase, por outro lado, pode reagir tanto com o NAD^+ quanto com o NADP^+ .

As desidrogenases ligadas ao NAD removem dois átomos de hidrogênio dos seus substratos. Um deles é transferido como um íon hidreto ao NAD^+ e o outro aparece como H^+ no meio (MALMSTRÖM, 1989).

Em geral, o complexo I da cadeia respiratória em bactérias é composto por 14 subunidades diferentes (FRIEDRICH et al., 1995; YAGI et al., 1998). Sete delas são proteínas periféricas, incluindo as subunidades que contém os grupamentos redox. As subunidades restantes são proteínas hidrofóbicas e formam, em conjunto, o domínio transmembrana responsável pela translocação do próton.

Os genes que codificam para as subunidades do complexo I são chamados *nuoA* a *nuoN* e encontram-se organizados em um operon. A organização dos loci para os genes *nuo* é conservada, como revela a comparação com outros genomas bacterianos e genomas de cloroplastos de plantas superiores. A ordem dos genes está relacionada ao arranjo topológico das subunidades codificadas (WEIDNER et al., 1993). Em *E. coli*, particularmente, *nuoC* e *nuoD* encontram-se fusionados, codificando para a subunidade NuoCD (FRIEDRICH, 1998). As subunidades codificadas pelos genes *nuoA* e *nuoB* de *A. brasilense* apresentam similaridade de 88% e 76% respectivamente com as seqüências de *Rhodospirillum rubrum* (Tabela 3).

4.2 Análise fisiológica dos mutantes HM26, HM053 e HM210 de *A. brasilense*

Com exceção dos genes *glnBA*, nenhum dos genes identificados no inserto do plasmídeo pAB441 apresenta similaridade com genes envolvidos no metabolismo de

nitrogênio. Para tentar identificar o(s) responsável(is) pelo efeito deletério observado nos mutantes HM26, HM053 e HM210, contendo o plasmídeo pAB441, regiões menores do inserto foram sub-clonadas no vetor pLAFR3.18 resultando nos plasmídeos pL+13 e pL+7.

O plasmídeo pL+13 contém o fragmento *HindIII/PstI* de 4,6 kb que carrega os genes *nuoB*, *nuoA*, *hupB* e *lon* enquanto o plasmídeo pL+7 contém o fragmento *SacI/HindIII* de 7,0 kb com os genes *glnA*, *dnaK*, *maoC*, *phzC*, *hpa2* e COG5457 (Figura 6).

O efeito dos plasmídeos pL+13, pL+7 e pLAFRglnA2 sobre a atividade da nitrogenase na estirpe selvagem FP2 e mutantes HM26, HM053 e HM210 foi determinado e os resultados da atividade da nitrogenase dos transconjugantes contendo estes plasmídeos são apresentados nas Figuras 7, 8 e 9. Os transconjugantes contendo o plasmídeo pL+13 apresentaram atividade de nitrogenase tanto na presença quanto na ausência de amônio (Figura 7). Estes resultados indicaram que esta região do plasmídeo não é o responsável pelo efeito repressor observado na presença do plasmídeo pAB441 (VITORINO, 2001). A diminuição na atividade da nitrogenase observada nos experimentos pode indicar que os fragmentos testados são parcialmente responsáveis pelo efeito de perda total da atividade da nitrogenase em presença do plasmídeo pAB441. Entretanto, um maior número de repetições é necessário para confirmar esta hipótese.

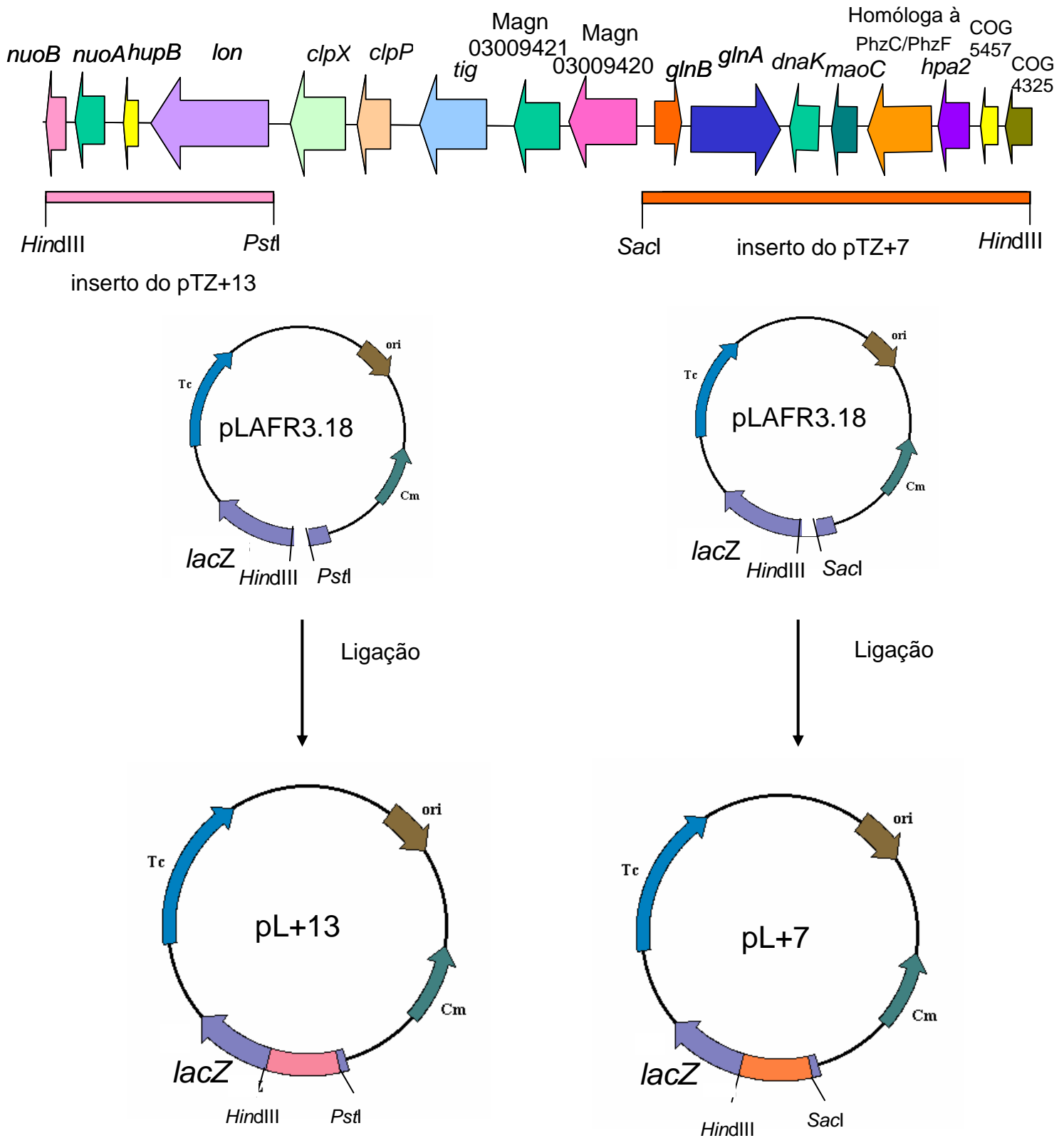


Figura 6: Esquema da clonagem de regiões do inserto de pAB441 em pLAFR3.18

EFEITO DE pL+13 SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES Nif^c DE *A. brasilense*

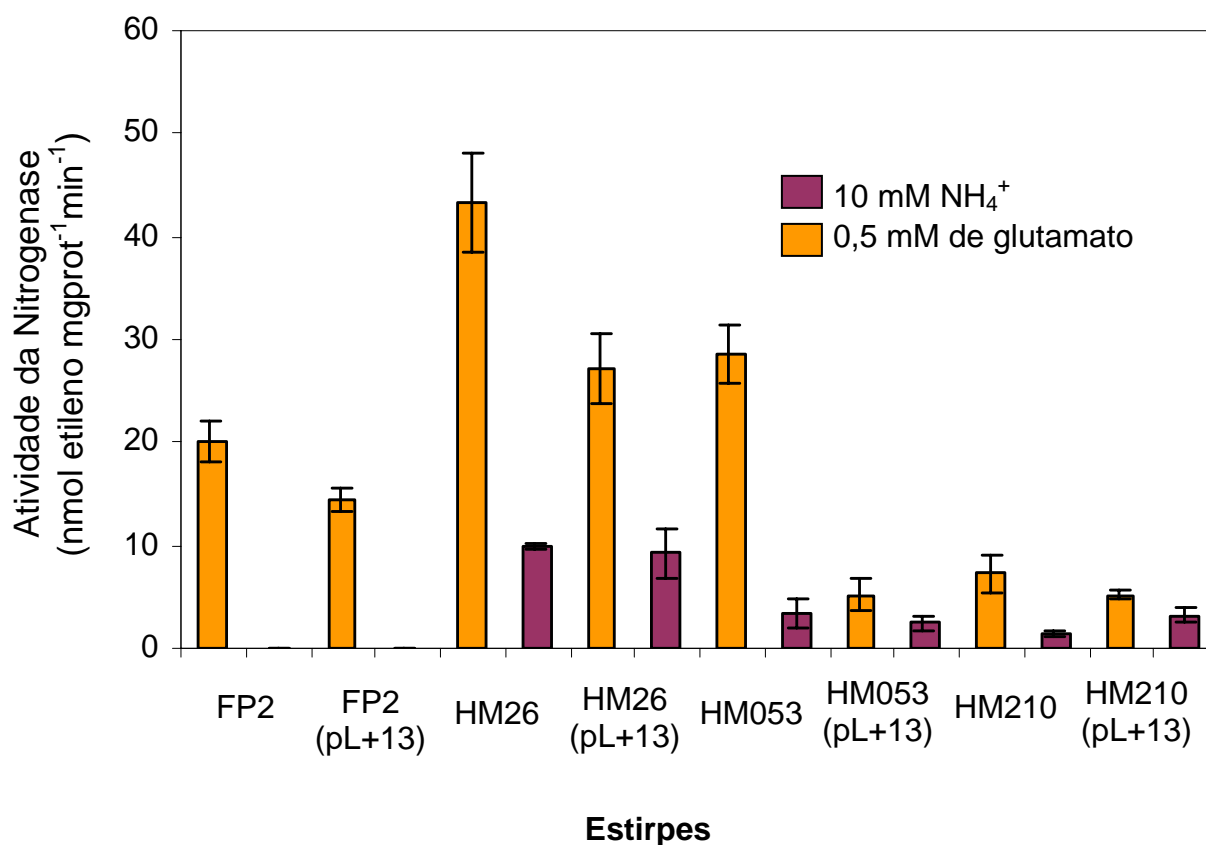


Figura 7: Atividade da nitrogenase da estirpe selvagem FP2, mutantes HM26, HM053, HM210 e transconjugantes FP2(pL+13), HM26(pL+13), HM053(pL+13) e HM210(pL+13) em presença de 10mM de amônio e presença de 0,5mM de glutamato. Média de três experimentos.

EFEITO DE pL+7 SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES Nif^c DE *A. brasilense*

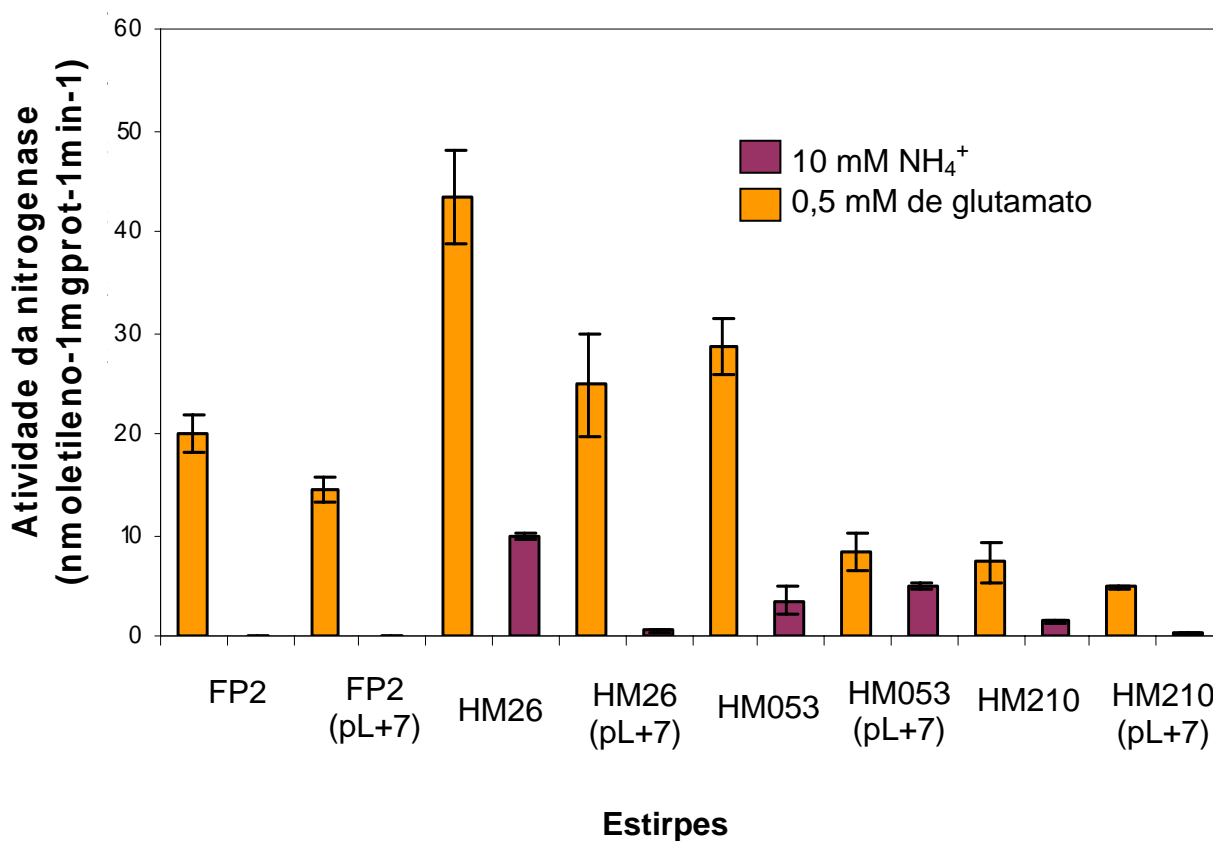


Figura 8: Atividade da nitrogenase da estirpe selvagem FP2, mutantes HM26, HM053, HM210 e transconjugantes FP2(pL+7), HM26(pL+7), HM053(pL+7) e HM210(pL+7) em presença de 10mM de amônio e presença de 0,5mM de glutamato. Média de três experimentos.

EFEITO DE pLAFRglnA2 SOBRE A ATIVIDADE DA NITROGENASE DOS MUTANTES Nif^c DE *A. brasilense*

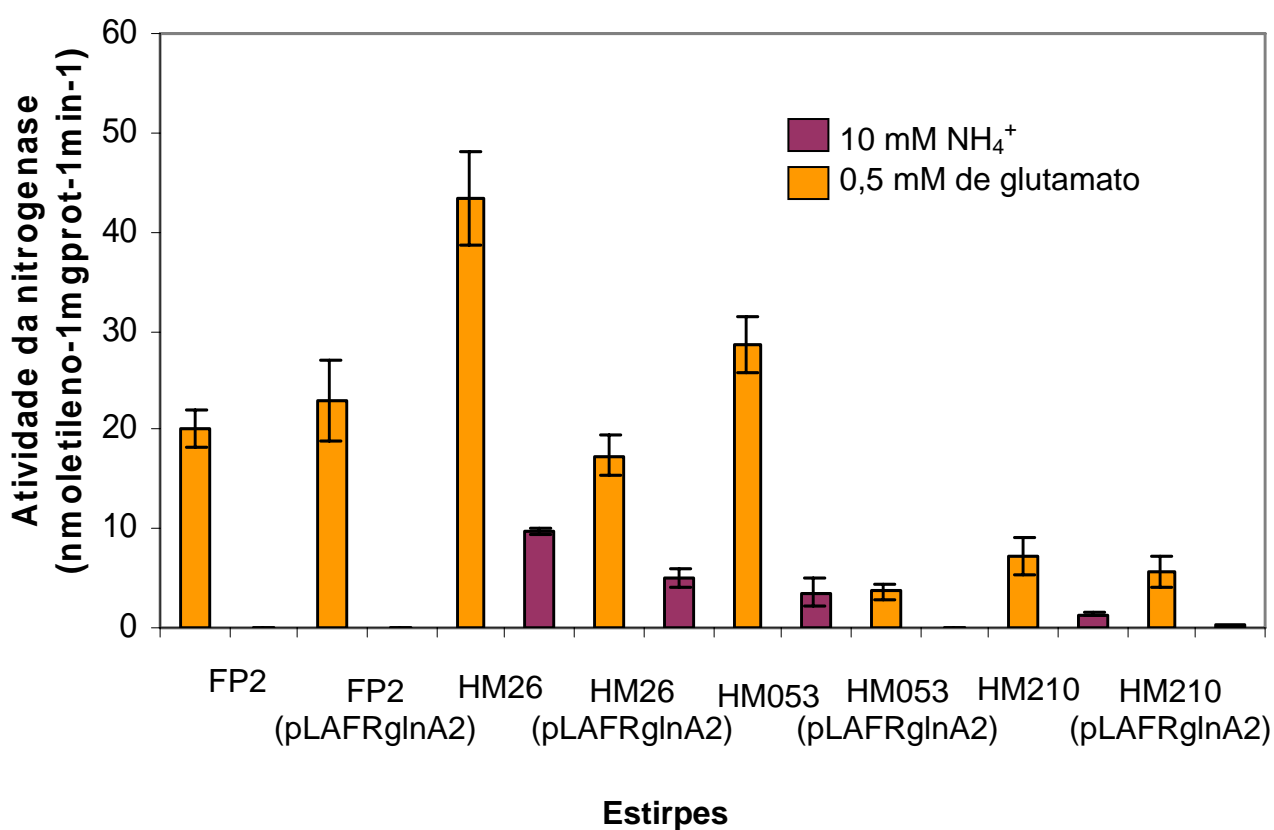


Figura 9: Atividade da nitrogenase da estirpe selvagem FP2, mutantes HM26, HM053, HM210 e transconjugantes FP2(pLAFRglnA2), HM26(pLAFRglnA2), HM053(pLAFRglnA2) e HM210(pLAFRglnA2) em presença de 10mM de amônio e presença de 0,5mM de glutamato. Média de três experimentos.

A atividade de nitrogenase dos transconjugantes contendo o plasmídeo pL+7 também foi determinada. O padrão de atividade da nitrogenase nos transconjugantes foi semelhante ao dos respectivos mutantes tanto na presença quanto na ausência deste plasmídeo (Figura 8), o que indica que pL+7 foi incapaz de reproduzir o efeito observado na presença do plasmídeo pAB41.

Nos transconjugantes contendo o plasmídeo pLAFRglnA2, a atividade constitutiva da nitrogenase também foi mantida (Figura 9), o que indica que esta região do plasmídeo pAB441 também não é a responsável por seu efeito sobre a atividade de nitrogenase dos mutantes.

Uma vez que o efeito repressor produzido pelo plasmídeo pAB441 nos transconjugantes não foi observado com os plasmídeos pL+7, pL+13 e pLAFRglnA2, é possível que o(s) gene(s) responsável esteja(m) localizado(s) na região central do inserto, que não foi subclonada.

4.3 Análise das seqüências de aminoácidos da glutamina sintetase dos mutantes HM26 e HM210

As análises da seqüência e dos experimentos de complementação genética não revelaram um provável gene responsável pelo efeito do plasmídeo pAB441 nas estirpes mutantes Nif^C. Entretanto, GAUTHIER e ELMERICH (1977) isolaram um mutante com fenótipo semelhante ao dos mutantes HM (estirpe 7028). Recentemente, DOMMELEN e colaboradores (2003) mostraram que esta estirpe apresenta mudança de base no gene *glnA*, sugerindo que os mutantes HM possam apresentar mutações nos genes *glnBA* como sugerido originalmente (MACHADO, 1988). VITORINO (2001) seqüenciou a região codificadora dos genes *glnB* e *glnZ* dos mutantes e não encontrou nenhuma mutação, mas não determinou a seqüência da região regulatória e nem do gene *glnA*. Com o objetivo de verificar a hipótese de MACHADO (1988), o operon *glnBA* dos mutantes HM26, HM053 e

HM210 foi amplificado por PCR e parcialmente seqüenciado. Diversas combinações de oligonucleotídeos iniciadores foram utilizadas para permitir a amplificação de fragmentos da região -660 até o final do operon *glnBA* e seu seqüenciamento (ver item 3.7 de Materiais e Métodos, Figura 3 e Figura 4). Seqüências de alta confiabilidade foram obtidas para grande parte do operon com exceção das regiões: promotora de *glnBA* entre os nucleotídeos -310 até -60 e região a jusante do primer interno GS1 (Figura 3). Não foram obtidos produtos de amplificação na região regulatória do operon *glnBA* compreendida entre os nucleotídeos -310 a -60. Portanto, mutações nesta região não puderam ser identificadas. Entretanto, é provável que exista uma mutação nesta região no mutante HM053, pois este apresenta baixos níveis de GS total, sugerindo um defeito na expressão de *glnA* (VITORINO, 2001).

A análise das seqüências obtidas permitiu a identificação de duas mutações no gene *glnA*, uma no mutante HM26 e outra no HM210. Em ambos os casos foi detectada uma transversão G→T no códon GGC (glicina) para TGC (cisteína).

Na glutamina sintetase do mutante HM26 a mutação levou à troca de um resíduo de glicina na posição 53 foi substituído por um resíduo de cisteína, enquanto no mutante HM210 o resíduo de glicina 129 foi trocado por um resíduo de cisteína. Os dois aminoácidos são conservados em praticamente todas as seqüências de glutamina sintetase depositadas no GenBank. Entretanto, o resultado da análise de predição de estrutura secundária da GS dos mutantes HM26 e HM210, utilizando o programa SOPMA (www.npsa-pbil.ibcp.fr), indicou que a mudança dos resíduos de aminoácidos nestas posições não altera a estrutura secundária. A troca de um resíduo de glicina por um resíduo de cisteína implica na mudança de uma cadeia lateral composta por um átomo de hidrogênio não polar, por uma cadeia lateral alifática contendo um átomo de enxofre mais polar.

A análise de estrutura cristalográfica por Raio-X da glutamina sintetase de *Salmonella typhimurium* permitiu a identificação do sítio de ligação para íons amônio. A formação deste sítio envolve os resíduos de ácido glutâmico 327 e 212, tirosina 179 e os resíduos de ácido aspártico 50 e serina 53 (em *A. brasilense* está localizado na posição 54) de uma subunidade adjacente (Figura 10). No modelo proposto, o íon amônio faz pontes de hidrogênio com a serina 53 e o ácido aspártico 50 (LIAW e EISENBERG, 1994). Estes autores sugerem, portanto, que estas regiões são importantes para a atividade da GS, uma vez que o íon amônio é um de seus substratos.

A estirpe mutante de *A. brasilense* 7029 apresenta menos de 1,5% da atividade da glutamina sintetase, tem baixíssima atividade de nitrogenase e é excretora de amônio. A GS desta estirpe apresenta a troca de um resíduo conservado de ácido aspártico por um resíduo de asparagina na posição 52 (van DOMMELEN et al., 2003), ou seja, no sítio de ligação a amônio proposto por LIAW e colaboradores (1995) (Figura 10). No modelo proposto por estes autores, a carga negativa do resíduo de ácido aspártico é crucial para o mecanismo catalítico da GS, pois facilita a desprotonação do íon amônio para gerar espécies ativas de NH_3 , as quais atacam o γ -glutamil fosfato formado na reação biossintética da GS. Este resíduo de ácido aspártico também aumenta a afinidade da GS por amônio (LIAW e EISENBERG, 1994).

A troca de glicina por cisteína na posição 53 da GS do mutante HM26 é vizinha ao resíduo essencial para a ligação com o substrato, sugerindo que a baixa atividade de GS neste mutante se deve a esta mutação, por impedimento estérico ou alteração na flexibilidade do sítio de ligação ao amônio.

ALINHAMENTO DE AMINOÁCIDOS DA PROTEÍNA GS

```

HM26      MSDISKVFDLIKEHVDKYVDLRFDPGRGLHHTAQHVSTIDEDVVFEDGIMFDCSSIAGWK 60
HM210    MSDISKVFDLIKEHVDKYVDLRFDPGRGLHHTAQHVSTIDEDVVFEDGIMFDGSSIAGWK 60
GS_SALTY S-AEHVLTMLNEHEVKFVDLRFDTKQKEQHVTIPAHQVNAEFFEFGKMFDGSSIAGWK 58
          *  .:.*:  :::*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
          *  .:.*:  :::*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

HM26      AINESDMILQLDPTTAVMDPFSAQPTLNILCDVYEPSTGQPYARCPRGIAKAAEKYMASA 120
HM210    AINESDMILQLDPTTAVMDPFSAQPTLNILCDVYEPSTGQPYARCPRGIAKAAEKYMASA 120
GS_SALTY GINESDMVLMFPDASTAVIDPFFADSTLIIRCDILEPGTLQGYDRDPRSIAKRAEDYL RAT 118
          .*****:*  *.:***:*:* *.:** * **:* **.* * * * * **.*** **.*: :
          .*****:*  *.:***:*:* *.:** * **:* **.* * * * * **.*** **.*: :

HM26      GIADTAYFGPEAEFFVFDVDFKFKVEMNKVSYEFDSEEGPYTSDKDYEDGNLGHRRPGVKGG 180
HM210    GIADTAYFCEAEFFVFDVDFKFKVEMNKVSYEFDSEEGPYTSDKDYEDGNLGHRRPGVKGG 180
GS_SALTY GIADTVLFGPEPEFFLFDIRFGASISGSHVAIDIEGAWNSSTKYEGGNKGHRRPGVKGG 178
          *****. * **.***:*:*:*:* * ..:  :*. **.:*...**.* **.***.***.***
          *****. * **.***:*:*:*:* * ..:  :*. **.:*...**.* **.***.***.***

HM26      YFPVAPVDSGSDLRAEMLSVLAEMGVPVEKHHHEVAAS-QHELGIKFDTLVVRTGDNMQYY 239
HM210    YFPVAPVDSGSDLRAEMLSVLAEMGVPVEKHHHEVAAS-QHELGIKFDTLVVRTGDNMQYY 239
GS_SALTY YFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQNEVATRFNTMTKKADEIQIY 238
          *****.***.***:*:*:* * :*:**:* ** ******: :*:**.: **:*:**.:**.*: *
          *****.***.***:*:*:* * :*:**:* ** ******: :*:**.: **:*:**.:**.*: *

HM26      KYVVHNVAHAYGKTATFMPKPVFGDNGSGMHMHQSIWKEGQPLFAGNQYADLSELALYYI 299
HM210    KYVVHNVAHAYGKTATFMPKPVFGDNGSGMHMHQSIWKEGQPLFAGNQYADLSELALYYI 299
GS_SALTY KYVVHNVAHRFGKTATFMPKPMFGDNGSGMHCHMSLAKNGTNLFSGDKYAGLSEQALYYI 298
          ***** * :*****:*:*:*** * * :*:** * **.*:*.**.*** * **.***
          ***** * :*****:*:*:*** * * :*:** * **.*:*.**.*** * **.***

HM26      GGIKHKAKALNAFTNPTTNSYKRLVPGYEAPVLLAYSARNRSASCRIPYVASPKGKRVEV 359
HM210    GGIKHKAKALNAFTNPTTNSYKRLVPGYEAPVLLAYSARNRSASCRIPYVASPKGKRVEV 359
GS_SALTY GGVIKHAKAINALANPTTNSYKRLVPGYEAPVMLAYSARNRSASIRIPVVASPKARRIEV 358
          **:******:*:*:*****:******:******:****** * ** * **.*:*.**.*
          **:******:*:*:*****:******:******:****** * ** * **.*:*.**.*

```

Figura 10: Alinhamento utilizando o programa ClustalW. Os aminoácidos idênticos são indicados por asteriscos, as substituições por aminoácidos de alta similaridade são indicadas por dois pontos e as substituições por aminoácidos de baixa similaridade são indicadas por um ponto. Foram utilizadas para o alinhamento as estirpes mutantes de *A. brasilense* HM26 e HM210 e a espécie *Salmonella typhimurium*. As mutações encontradas neste trabalho estão destacadas por fundo em vermelho. Os resíduos importantes para a síntese de glutamina, de acordo com o modelo proposto por LIAW e colaboradores (2000), estão destacados por fundo em amarelo.

O resíduo de glicina 129 da GS de *A. brasilense* está localizado numa região altamente conservada nas GS de diversas espécies de bactérias e eucariotos. Este domínio conservado apresenta, na posição 131 um resíduo de ácido glutâmico. Estudos realizados com a glutamina sintetase de *Salmonella typhimurium* mostraram que o resíduo de ácido glutâmico equivalente (Glu-129) tem um papel importante no mecanismo de reação da GS, pois é parte do sítio n_2 , que é responsável pela coordenação de íons metálicos, como o Mn^{+2} e o Mg^{+2} . O íon metálico coordenado neste sítio, participa na ligação de ATP e ADP. A carga positiva do íon metálico é estabilizada pelas cadeias laterais dos resíduos de aminoácidos do sítio n_2 e pelo fósforo γ do ATP. A coordenação do íon metálico pelo ATP aumenta sua afinidade pela GS e reduz a densidade eletrônica do átomo de P, facilitando o ataque da carboxila γ do substrato glutamato, permitindo a formação do intermediário γ -glutamil fosfato, passo essencial na síntese da glutamina pela GS (LIAW, VILLAFRANCA e EISENBERG, 1993). O íon amônio então, faz ataque nucleofílico ao carbono-5 da γ -glutamil fosfato, formando glutamina e PO_4^- .

No mutante HM210 a glicina 129 da glutamina sintetase foi substituída por uma cisteína. A presença de um átomo de enxofre próximo ao glutamato (131 na GS de *A. brasilense* e 129 na GS de *S. typhimurium*) do sítio n_2 , pode desestabilizar a coordenação de íons metálicos, levando à uma diminuição da afinidade de GS por ATP. Além disso, caso os íons metálicos não interajam com a GS, o fosfato- γ da molécula de ATP estaria menos susceptível ao ataque nucleofílico do glutamato, dificultando a formação do γ -glutamil fosfato e a síntese de glutamina.

A atividade de GS dos mutantes HM26 e HM210 foram determinadas anteriormente por MACHADO et al. (1988) e VITORINO et al. (2001) e são mostradas na tabela 5. Os níveis de GS total nas duas linhagens é semelhante às da estirpe selvagem FP2, mas nos mutantes ela encontra-se quase completamente adenililada. As atividades biossintéticas

destes mutantes são respectivamente 12 e 23% (Tabela 5). Estes valores são semelhantes aos obtidos para o mutante 7028, cuja atividade de nitrogenase também é constitutiva. Por outro lado, a linhagem 7029 é Nif⁻ e sua atividade de GS biossintética é apenas 1,5% da selvagem (GAUTHIER e ELMERICH, 1977). Estes resultados confirmam um importante papel da GS na regulação da fixação de nitrogênio em *A. brasilense*. Esta função regulatória de GS provavelmente deve-se à sua influência direta nos níveis de glutamina, que é o principal sinalizador intracelular do status de nitrogênio (ADLER et al., 1975).

TABELA 4: ATIVIDADE DE GLUTAMINA SINTETASE NOS MUTANTES Nif^C**DE***Azospirillum brasilense*

| Estirpe | Fonte de Nitrogênio | Atividade biossintética | Atividade transferásica | | |
|---------|-----------------------------------|-------------------------|-------------------------|-------------------|----------------------------|
| | | | -Mg ⁺² | +Mg ⁺² | Fração não adenililada (%) |
| FP2 | Glu 5mM | 0,239 | 4,60 (1,12) | 2,20 (0,89) | 47,8 |
| | NH ₄ ⁺ 20mM | 0,028 | 1,17 (0,15) | 0,19 (0,05) | 16,6 |
| HM26 | Glu 5mM | 0,028 | 3,19 (0,38) | 0,03 (0,01) | 1,1 |
| | NH ₄ ⁺ 20mM | 0,007 | 1,01 (0,08) | 0,03 (0,01) | 3,3 |
| HM210 | Glu 5mM | 0,056 | 5,60 (0,15) | 0,16 (0,01) | 2,9 |
| | NH ₄ ⁺ 20mM | 0,007 | 1,23 (0,08) | 0,05 (0,01) | 4,4 |

FONTES: MACHADO, 1988 e VITORINO, 2001.

5. CONCLUSÃO

Os resultados obtidos até o momento não permitem identificar o gene ou grupo de genes responsável(is) pelo efeito observado em presença do plasmídeo pAB441 sobre a atividade da nitrogenase dos mutantes HM26, HM053 e HM210.

FIGURA 4 - ANEXO 1: SEQUÊNCIA DO INSERTO DO PLASMÍDEO pAB441

COG4325: Predicted membrane protein

| | | |
|--------|---|------|
| | A R S I I Y D N V V Q R V A A D L R D V T A A L L P E | |
| AAGC 1 | TTGCCCCGCTCGATCATCTACGACAATGTGGTGCAGCGGGTTCGCCGCGGATCTGCGCGACGTGACCGCAGCCCTGCTTCCC | 80 |
| TTCG 1 | AACGGGCGAGCTAGTAGATGCTGTTACACCACGTGCGCCAGCGGCGCCTAGACGCGCTGCACTGGCGTCTGGGACGAAGGG | 80 |
| | R G D G P P E I P P D L R G D F A W V D L D Q D G Y | |
| 81 | GAGCGCGGCGACGGTCCCCCGGAGATACCCCCGACCTGCGCGGGGATTTGCGCTGGGTGGACCTCGACCAGGACGGCTA | 160 |
| 81 | CTCGCGCCGCTGCCAGGGGGCTCTATGGGGGGCTGGACGCGCCGCTAAAGCGGACCCACCTGGAGCTGGTCTGCCGAT | 160 |
| | V Q A I D L D G L V A T A Q N A D S V I R L D I R P G | |
| 161 | CGTCCAGGCGATCGACCTGGACGGTCTGGTTCGCCACCGCCAAAACGCCGATTCGGTGATTCGCCTTGATATCCGGCCCCG | 240 |
| 161 | GCAGGTCCGCTAGCTGGACCTGCCAGACCAGCGTGGCGGGTTTTGCGGCTAAGCCACTAAGCGGAACTATAGGCCGGGC | 240 |
| | H Y V I G G G E H V A V F P A G A C T A D L S K A V R | |
| 241 | GCCATTACGTATCGGCGGGCGGAGCATGTGGCCGTGTTTCCGGCCGGCGCCTGCACGGCGGACCTGTGCAAGGCGGGT | 320 |
| 241 | CGGTAATGCAGTAGCCGCCCGCTCGTACACCGGCACAAAGGCCGGCCGCGGACGTGCCGCTGGACAGCTTCCGCCAG | 320 |
| | G A F I V G S E R T P T Q D V E F G I R Q L V E M A | |
| 321 | CGCGGCGCTTTCATCGTTCGGTTCGGAGCGTACCCCCGACCCAGGATGTGGAGTTCGGCATCCGCCAACTGGTGGAGATGGC | 400 |
| 321 | GCGCCGCGAAAGTAGCAGCCAAGCTCGCATGGGGCTGGTCTACACCTCAAGCCGTAGGCGGTTGACCACCTCTACCG | 400 |
| | T R A L S P G I N D V F T A L A V I K N L S A S M A | |
| 401 | CACCCGCGCCCTGTGCGCGGAATCAACGACGTGTTACCGCCTTGGCGGTGATCGACAACCTCTCGGCCTCGATGGCCC | 480 |
| 401 | GTGGGCGGGACAGCGGCCCTTAGTTGCTGCACAAGTGGCGGAACCGCCACTAGCTGTTGGAGAGCCGGAGCTACCGGG | 480 |
| | R I F D R A I E L A V L R D S A G T V R V L R D V T G | |
| 481 | GCATCTTCGACCGCGCGATCGAACTGGCGGTTCTGCGCGACTCCGCCGGAACGGTCCGGGTCTTGCGGGACGTGACCGGA | 560 |
| 481 | CGTAGAAGCTGGCGCGCTAGCTTGACCGCCAAGACGCGCTGAGGCGGCCTTGCCAGGCCCAGAACGCCCTGCACTGGCCT | 560 |
| | Y D G F V G A A F D Q I R Q A G S G N A A V L I R L V | |
| 561 | TACGACGGCTTCGTGCGCGCCCTTCGATCAGATCCGCCAGGCCGGGAGCGGAAACGCCGCGGTGCTCATCCGGCTTGT | 640 |
| 561 | ATGCTGCCAAGCAGCCGCGGCGGAAGCTAGTCTAGGCGTCCGGCCCTCGCCTTTGCGGCGCCACGAGTAGGCCGAACA | 640 |
| | D A I I R L A P R V R L K A Q R E P L R E Q L D M I | |
| 641 | CGACGCCATCATTCGGTGGCCCCGCGCGTCCGCTTGGACGCCCAGCGGGAACCGCTGCGGGAGCAGTTGGACATGATCC | 720 |
| 641 | GCTGCGGTAGTAAGCCGACCGGGGCGCGCAGGCGAACCTGCGGGTTCGCCCTTGGCGACGCCCTCGTCAACCTGTACTAGG | 720 |
| | L A A G E Q N I A I P R D L A I L R A R Q R Q A T E R | |
| 721 | TGGCGGCGGGCGAACAGAACATCGCGATTCGCCGATCTTGCATTCCTTCGAGCCCCTTGGAGGCAGGCGACCGAGCGG | 800 |
| 721 | ACCGCCGCCGCTTGTCTTGTAGCGTAAGGGGCGCTAGAACGGTAAGAAGCTCGGGCAACCTCCGTCCGCTGGCTCGCC | 800 |
| | L E R * | |
| 801 | CTGGAGCGCTGACGCCTGTGCGACCACGCAACCCGGTTCGGCGCGGGGCGCGGGTGTGGCCGCCGGGATAACATCCGAC | 880 |
| 801 | GACCTCGCGACTGCGGACAGCCTGGTGCCTGGGCCAAGCCGCGCCCGCCGCCCCACACCGGCGGCCCTATGTAGGCTG | 880 |
| | promotor σ 70 | |
| 881 | ATCGGAAATGTAATGGTGAACAATAACGACATTGACCAGCGGATTTGTATCGACATAATCGCCATTGTCTCGATTGTTCCG | 960 |
| 881 | TAGCCTTTACATTACCACTGTTATTGCTGTAACCTGGTTCGCGCTAACATAGCTGTATTAGCGGTAACAGAGCTAACAAAGCC | 960 |
| | orf. conservada | |
| | M R T T E A T R S S G S S S F T A | |
| 961 | GGCTGAATGTTTTGACGAAGGAGTTTCATGCGCACTACGGAAGCGACCCGCTCGTCCGGATCGTCTCGTTTCACGGC | 1040 |
| 961 | CCGACTTACAAAACCTGCTTCTCCAAGTAGTACGCGTGATGCCTTCGCTGGGCGAGCAGGCCTAGCAGGAGCAAGTGCCG | 1040 |
| | V F K V I A R A L S V G I R P V W F V M E A G L N G Y | |
| 1041 | CGTCTTCAAGGTCATCGCGCGCGCCCTCTCCGTTCGGCATCCGCCCGTCTGGTTCGTGATGGAGGCCGGCCCTGAACGGCT | 1120 |
| 1041 | GCAGAAAGTTCCAGTAGCGCGCGGGGAGAGGCAGCCGTAGGCGGGGACACCAAGCACTACCTCCGGCCGGACTTGCCGA | 1120 |

E R W R Q R Q A L M R L D D M L L K D I G V S R A D

1121 ATGAACGGTGGCGGCAGCGGCAGGCGCTGATGCGGCTGGACGATCATCTGCTGAAGGACATCGGCGTGTCCCGCGCCGAT 1200
 1121 TACTTGCCACCGCCGTGCGCGTCCGCGACTACGCCGACCTGCTAGTAGACGACTTCTGTAGCCGCACAGGGCGCGGCTA 1200

HPA2

V S R A D V D S E V S K * M S A D A I L Y A V E P

1201 GTGGACAGCGAAGTCAGCAAGCCGTTCTGGCGGGGTTGATGCGATGAGCGCCGATGCGATCCTCTATGCCGTGAGCCCCG 1280
 1201 CACCTGTGCTTCAGTCGTTGCGCAAGACCGCCCAACTACGCTACTCGCGGCTACGCTAGGAGATACGGCAGCTCGGGC 1280

D L T A D A F I D V L H R S G L A E R R P V A D R P R

1281 ACCTGACCGCCGACGCGCTTCATCGACGTGCTTACCCTCCGGGCTGGCCGAACCGCCGTCCTGTTGGCGGACCGCCACGG 1360
 1281 TGGACTGGCGGCTGCGGAAGTAGCTGCACGAAGTGGCGAGGCCGACCGGCTTGGCGCAGGGCACCCTGCGGGGTGCC 1360

V E A M L R N A G L I V T A R N A D G A L V G V A R S

1361 GTGGAGGCGATGCTGCGCAACCGCCGCTGATCGTCACCGCCGCAACCGGACGGAGCCTTGGTGGAGTCGCCCCGGTC 1440
 1361 CACCTCCGCTACGACGCTTGGCGCCGACTAGCAGTGGCGGGCGTTGCGCCTGCCTCGGAACCAGCCTCAGCGGGCCAG 1440

V T D F A F C C Y L S D L A V D R A C Q G R G I G K

1441 CGTCACCGATTTGCGCTTCTGCTGCTATCTGTCCGACCTCGCCGTCGACCGCGCCTGTCAGGGGCGCGGCATCGGCAAGG 1520
 1441 GCAGTGGCTAAAGCGGAAGACGACGATAGACAGGCTGGAGCGGCAGCTGGCGCGGACAGTCCCCGCGCCGTAGCCGTTCC 1520

E L M R R T R A A V G E G V T C L L L S A P K A I T F

1521 AGCTGATGCGCCGACCCGCGCCGCGTGGGGGAGGGCGTGACCTGCCTTCTCCTGTCCGCGCCCAAGGCGATCACCTTC 1600
 1521 TCGACTACGCGGCGTGGGCGGCGCCACCCCTCCGCACTGGACGGAAGAGGACAGGCGGGGTTCCGCTAGTGAAG 1600

COG0384: Predicted epimerase, PhzC/PhzF homolog

Y E A A G M E R H A Q A F L F T E K P * G G T V R L P

1601 TACGAGGCGGCGGGGATGGAGCGGCACGCCAGGCGTTTCTGTTTACCAGAAAACCCTAAGGAGGCACCGTGGCCTTCC 1680
 1601 ATGCTCCGCCGCCCTACCTCGCCGTGCGGGTCCGCAAGGACAAGTGGCTTTTTGGGATTCTCCGTGGCACGCGGAAG 1680

L Y Q V D A F A D A V F A G N P A A V V P L E D W L

1681 CCTGTATCAGGTCGATGCCTTCGCGGACGCGGTGTTTCGCGGGCAACCCGGCGGCGGTGTCGCCGCTGGAGGACTGGCTTC 1760
 1681 GGACATAGTCCAGTACGGAAGCGCCTGCGCCACAAGCGCCGTTGGGCGCGCCAGCACGGCGACCTCCTGACCGAAG 1760

P D A T L Q A I A A E N N L A E T A F L V R R G D G Y

1761 CCGACGCGACTTCCAGGCCATCGCGGCGGAAAACAACCTGGCGGAAACCGCCTTCTCCTGTCGCCGCGGTTGACGGTTAC 1840
 1761 GGCTGCGCTGAGAGGTCCGGTAGCGCCGCTTTTGTGGACCGCCTTGGCGGAAGGAGCAGGCGGCGCCACTGCCAATG 1840

E I R W F T P T V E V D L C G H A T L A S A F V I A T

1841 GAGATCCGCTGGTTACCCCGACGGTTCGAGGTGGACCTGTGCGGTACGCCACGCTGGCCTCGGCCTTCGTATCGCGAC 1920
 1841 CTCTAGGCGACCAAGTGGGGCTGCCAGTCCACCTGGACACGCCAGTGGCGGTGCGACCGGAGCCGGAAGCAGTAGCGCTG 1920

M L E W G C P R I D F A T R E A G T L S V T R D G D

1921 GATGCTGGAGTGGGGCTGCCCGCGCATCGACTTCGCCACCCGGGAAGCCGGAACCCTCAGCGTCACCCGCGACGGCGACC 2000
 1921 CTACGACCTCACCCCGACGGGCGGCTAGCTGAAGCGGTGGGCCCTTCGGCCTTGGGAGTCGAGTGGGCGCTGCCGCTGG 2000

L Y T L D F P S R P P E P L A G P P A G L L D A L G G

2001 TTTACACGCTGGATTTCCCCAGCCGTCGCCGGAACCGCTGGCCGGGCGCCGCGCCGCGCCTGCTCGACGCGCTGGGCGGA 2080
 2001 AAATGTGCGACCTAAAGGGTTCGGCAGGCGGCTTGGCGACCGGCCGCGGCGGCGGACGAGCTGCGCGACCCGCT 2080

P P P T A L L K A R D L V V V Y D D E A T V R G L T P

2081 CCGCCGCCGACGGGCGTGTCAAGGCGCGGACCTCGTGGTCTGCTATGACGACGAGGCGACGGTGGCGGGCTTGACCCC 2160
 2081 GGCGGCGGCTGCCGCGACGAGTTCGCGCGCTGGAGCACCAGCAGATACTGCTGCTCCGCTGCCACGCGCCGAACCTGGGG 2160

D M S A L A R L E D F Y A V I V T A P G S G G V D F

2161 CGACATGAGCGCCCTGGCCCGGCTGGAGGATTTCTACGCGGTATCGTGACGGCGCCGGGACGCGGTGGCGTGGACTTCG 2240
 2161 GCTGTACTCGCGGGACCGGGCCGACCTCCTAAAGATGCGCCAGTAGCACTGCCGCGGCCCCGTGCCACCGCACCTGAAGC 2240

V S R F F A P A Q G I P E D P V T G S A H C T L I P Y

2241 TCTCCCGCTTCTTCGCGCCGCGCCAGGGGATTTCCGAAGACCCGGTGACCGGCTCCGCCATTGCACGCTGATCCCCTAT 2320
 2241 AGAGGGCGAAGAAGCGCGGCGGGTCCCTAAGGGCTTCTGGGCCACTGGCCGAGGCGGGTAACGTGCGACTAGGGGATA 2320

| | | |
|------|---|------|
| | W A R R L G K E R L K A R Q V S A R G G S L S C E L V | |
| 2321 | TGGGCGGGCGGCTGGGCAAGGAGCGGCTGAAGGCCCGGCAGGTCTCGGCCCGGGCGGTTCCCTGTCTGCGAGCTGGT | 2400 |
| 2321 | ACCCGCGCCGCCGACCCGTTCTCGCCGACTTCCGGGCCGTCCAGAGCCGGGCGCCGCCAAGGGACAGGACGCTCGACCA | 2400 |
| | G D R V K I A G K A V L Y M D G A V Y I * | |
| 2401 | GGGCGACCGTGTGAAGATCGCCGGGAAAGCGGTTCTCTATATGGATGGCGCCGTCTACATTTAGCCCCGTCCATACGAGG | 2480 |
| 2401 | CCCGCTGGCACACTTCTAGCGGCCCTTTCGCCAAGAGATATACCTACCGCGGCAGATGTAATCGGGGCAGGTATGTCTCC | 2480 |
| | COG2030: Acyl dehydratase MaoC | |
| | M A G R Y F E D F R V G D V I D T E | |
| 2481 | ACCAATGGAACGGGAGGGCAGCGGGCATGGCGGGGCGCTATTTTCGAGGATTTCCGCGTGGTGACGTGATCGACACCGAG | 2560 |
| 2481 | TGGTTACCTTGCCCTCCCGTGCCTCCCGTACCGCCCCGCATAAAGCTCCTAAAGGCGCAGCCACTGCACTAGCTGTGGCTC | 2560 |
| | G A T V T D S Q I M D F A Q R F D P Q P F H M D A V A | |
| 2561 | GGGCGACGGTACCGACAGCCAGATCATGGATTTTCGCCAGCGCTTCGACCCGCAACCCTTCCACATGGACGCGGTGGC | 2640 |
| 2561 | CCCCGTGCCAGTGGCTGTGGTCTAGTACCTAAAGCGGGTTCGCGAAGCTGGGCGTTGGGAAGGTGTACCTGCGCCACCG | 2640 |
| | A A E G P F G G L I A S G F H T L S L T F R L F R D | |
| 2641 | CGCCGCGGAGGGGCGGTTTCGGCGGGCTGATCGCCAGCGGCTTCCACACCCTGTCCCTGACCTTCCGGCTGTTCCGCGACA | 2720 |
| 2641 | GCGGCGCTCCCCGGCAAGCCGCCGACTAGCGGTTCGCCAAGGTGTGGGACAGGGACTGGAAGGCCGACAAGGCGCTGT | 2720 |
| | T G A I T G T S L G G S G G D E L R W L R P V R P G D | |
| 2721 | CCGGGGCCATACCGGCACCAGCCTGGGCGGTTCCGGCGGCGACGAGTTGCGCTGGCTGCGCCCGGTGCGGCCCGGGC | 2800 |
| 2721 | GGCCCCGTAGTGGCCGTGGTTCGACCCGCCAAGCCCGCCGCTGCTCAACGCGACCGACGCGGGCCACGCGGGCCGCTG | 2800 |
| | T L R V R V E V V E A I P S R R G D R G T V R L A Y T | |
| 2801 | ACGCTGCGCGTCCGGGTCGAGGTGGTGGAGGCCATCCCGTCCCGCCGCGGCGACCGCGGCACGGTCCGGCTGGCCTACAC | 2880 |
| 2801 | TGCGACGCGCAGGCCAGCTCCACCACCTCCGGTAGGGCAGGGCGGCGCCGCTGGCGCCGTGCCAGGCCGACCGGATGTG | 2880 |
| | T L N Q H G E A V M T I T L N H I V A T R S I P A E | |
| 2881 | CACGCTGAACCAGCACGGCGAGGCGGTGATGACCATCACCTTGAACCACATCGTGGCGACCCGAGCATTCCGGCGGAGT | 2960 |
| 2881 | GTGCGACTTGGTGTGCCGCTCCGCCACTACTGGTAGTGGAACTTGGTGTAGCACCCTGGGCGTCGTAAGGCCGCTCA | 2960 |
| | * | |
| 2961 | AGCAAACACCGACGGCCTGTTCCGTGCGCGCTTTGATCCGCGAGGCCCGGCCCTGTATAGCCTCGCCCCCTTCCATCACG | 3040 |
| 2961 | TCGTTTGTGGCTGCCGGACAAGGCAGCGCGCAGAACTAGGCGTCCGGGCCGGGGACATATCGGAGCGGGGAAGGTAGTGC | 3040 |
| | probable heat-shock protein Hsp 70 | |
| | M T A C G L D F G T S N T T L G V R M A A G A | |
| 3041 | GCGGGGTTCCCATGACGGCTTGCAGCCTCGATTTTCGGCACGTCCAACACGACGCTCGGCGTGCAGGATGGCGGGCGGCGC | 3120 |
| 3041 | CGCCCCAAGGGTACTGCCGAACGCCGAGCTAAAGCCGTGCAGGTTGTGCTGCGAGCCGCACGCTACCGCCGCCCGCG | 3120 |
| | D T G G G A Y G L L A L D G T R T T L P S A V F F D | |
| 3121 | CGACACGGGAGGGGGGCTTACGGCCTGCTTGCCCTCGACGGCACACGGACGACCCTGCCAGCGCCGTCTTCTTCGATT | 3200 |
| 3121 | GCTGTGCCCTCCCCCGAATGCCGGACGAACGGGAGCTGCCGTGTGCTGCTGGGACGGGTGCGGGCAGAAGAAGCTAA | 3200 |
| | F E A D A V T V G Q A A I D A Y A S G V E G R L M R S | |
| 3201 | TCGAAGCCGACGCGGTGACGGTTCGGGACGGCGCCATCGACGCCTATGCGTCCGGCGTTCGAGGGGCGTTTGTGCGTTCC | 3280 |
| 3201 | AGCTTCGGTTCGCGCACTGCCAGCCGTCGCCCGGTAGCTGCGGATACGCAGCCCGCAGCTCCCCGAAACTACGCAAGG | 3280 |
| | I K S M L G T D L I D A D T A L R K G R I T F R A V I | |
| 3281 | ATCAAGTCGATGCTGGGCACCGACCTGATCGACGCCGACACCGCGCTGCGCAAGGGACGCATCACCTTCCGCGCGGTCAT | 3360 |
| 3281 | TAGTTCAGCTACGACCCGTGGCTGGACTAGCTGCGGCTGTGGCGCGACGCGTTCCTGCGTAGTGGAAAGGCGCGCCAGTA | 3360 |
| | A T F L A A V K E R A E T A L G G E L T D V V L G R | |
| 3361 | CGCCACCTTCTGGCGGCGTCAAGGAACGGGCGGAGACCGCGCTCGGCGGAGAGCTGACCGACGTGGTGTCTGGGCGCC | 3440 |
| 3361 | GCGGTGAAAGACCGCCAGTTCCTTGCCCGCTTGGCGCGAGCCGCTCTCGACTGGCTGCACCACGAGCCCGCGG | 3440 |
| | P V H F V D G D P A G D A A A E E V L G E I A R S V G | |
| 3441 | CGGTGCATTTTCGTCGACGGCGACCCGGCGGGCGACGCCGCGGCGGAGGAAGTGTGGGCGAGATCGCCCGCTCGGTTCGGA | 3520 |
| 3441 | GCCACGTAAAGCAGCTGCCGCTGGGCGCCCGCTGCGGCGCCGCTCCTTACGACCCGCTCTAGCGGGCGAGCCAGCCT | 3520 |

F R H I S F Q Y E P I A A A L D Y E Q Q V E R E E L A
 3521 TTCCGCCACATCTCCTTCCAATATGAGCCGATCGCCGCCGCGCTCGACTACGAGCAGCAGGTGGAGCGCGAGGAGCTGGC 3600
 3521 AAGCGGTGTAGAGGAAGGTTATACTCGGCTAGCGGGCGCGAGCTGATGCTCGTCCACCTCGCGCTCCTCGACCG 3600

L I A D I G G G T S D F S I V R I G P E R R G R A D
 3601 GCTGATCGCCGATATCGGGCGGCGGTACGTCCGACTTCTCCATCGTCCGCATCGGGCCGAGCGGGCGGGGCGGGCCGACC 3680
 3601 CGACTAGCGGCTATAGCCCGCCCATGCAGCCTGAAGAGGTAGCAGGCGTAGCCCGGCTCGCCGCGCCCGCCGCGGCTGG 3680

R S A D I L A N D G I R V G G T D F D R D L S L A T A
 3681 RCAGCGCCGACATCCTGGCCAATGACGGCATCCGGGTGCGGGCACCGATTTTCGACCGTGACCTGTGCTGGCGACGGCC 3760
 3681 CGTCCGCGGCTGTAGGACCGGTTACTGCCGTAGGCCAGCCGCGTGGCTAAAGCTGGCACTGGACAGCGACCGCTGCCGG 3760

M P L L G H G S A M K R P G L L A P K H L F F D L A T
 3761 ATGCCGCTGTGGGCCACGGCAGCGCCATGAAGCGCCCGGTCTGTGGCCCCAAGCATCTGTTCTTCGATCTGGCGAC 3840
 3761 TACGGCGACGACCCGGTGCCGTGCGGTTACTTCCGCGGGCCAGACGACCGGGGGTTCGTAGACAAGAAGCTAGACCGCTG 3840

W S K I N F L Y T A K A M A D L E R L K R D S A T P
 3841 CTGGTCAAGATCAACTTCTTACACCGCAAGGCGATGGCCGACCTTGAGCGCCTGAAGCGGATTCCGCGACTCCGG 3920
 3841 GACCAGCTTCTAGTTGAAGGAGATGTGGCGGTTCCGCTACCGGCTGGAACCTCGCGGACTTCCGCGTAAGGCGCTGAGGCC 3920

D R I E R L I G V V E H R L G H A L A M A V E R T K I
 3921 ACCGCATTGAGCGGTGATCGGCGTGGTTCGAGCACCAGGCTGGGCCATGCCCTGGCCATGGCGGTGGAGCGCACCAAGATC 4000
 3921 TGGCGTAACTCGCCGACTAGCCGCACCAGCTCGTGGCCGACCCGGTACGGGACCGGTACCGCCACCTCGCGTGGTTCTAG 4000

A L S D G P E A G L L L D W G G G M T R P V S V G E
 4001 GCCCTGTCCGACGGGCCGAGGCCGGCTGCTGCTGGATTGGGGCGGTGGCGGCATGACCGCCCCGGTATCGGTGGGCGA 4080
 4001 CGGGACAGGCTGCCCGCCTCCGCGCGACGACCTAACCCTCGCCACCGCGTACTGCGCGGGCCATAGCCACCCGCT 4080

L N E A T A I L A G R I G G R V R A C L A E A G T L
 4081 GTTGAACGAGGCGACGGCGATCCTGGCCGGCCGCATCGGCGGGCGGGTGCAGCGCCTGTCTGGCCGAGGCCGGCACGCTGC 4160
 4081 CAACTTGTCCGCTGCCGCTAGGACCGCCGGCGTAGCCGCCCGCCACGCGCGGACAGACCGGCTCCGCGCGTGCAGC 4160

P A E I D A V F L T G G S T L L P Q V R A A I L A E L
 4161 CGGCGGAGATCGACCGGTGTTTCTGACCGCGGCTCCACGCTGCTGCCGAGGTGCGGGCGGCCATCCTGGCCGAGCTG 4240
 4161 GCCGCTCTAGCTGCGCCACAAGGACTGGCCGCCGAGGTGCGACGACGGCGTCCACGCCCCGCGGTAGGACCGGCTCGAC 4240

P G A R V V E G N I F G S V G T G L T I E A A R R F G
 4241 CCCGGCGCCCGTGTGGTGGAGGGCAACATCTTCGGCTCGGTCCGGCACCAGGCTGACCATCGAGGCAGCGCGGCGCTTCGG 4320
 4241 GGGCCGCGGGCACACCACCTCCGTTGTAGAAGCCGAGCCAGCCGTTGGCCGACTGGTAGCTCCGTCGCGCCGCGAAGCC 4320

G T A * terminador
 4321 CGGCACCGCTGAGGCATCTCTTACGTCTTGGACACAAAAAGCCCACCGACGGTGTGCGCCGGTGGGCTTTTTTTGTGC 4400
 4321 GCCGTGGCGGACTCCGTAGAGAAGTGCAGGACCTGTGTTTTTTCGGGTGGCTGCCACAGCGGCCACCCGAAAAAACACG 4400

4401 TCTGCTCCGCGGTCCGGCCCTCCGGTGGTGCCCGCCGCGCCGAGGGCGGGGCGGACACCGGAATCCGGTGGACCGAAG 4480
 4401 AGACGAGGCGCCAGGCCGGAGGCCACCACGGGCGGCGGGCGCTCCCGCCCGCCTGTGGCCTTAGGCCACCTGGCTTC 4480

4481 TCCGATGGATCAGACCGAGTAGTACATCTTGTACTCGATCGGGTGCGGCATGGTCTCGAAGGCCAGCAGCTCCTCGGTGC 4560
 4481 AGGCTACCTAGTCTGGCTCATCATGTAGAACATGAGCTAGCCCACGCCGTACCAGAGCTTCCGGTTCGTCGAGGAGCCACG 4560
*** V S Y Y M K Y E I P H P M T E F A L L E E T R**

4561 GCAGGTCGATGTAGGACTCGATCATGTCCTTGGTGAACACGTGCCCTTCTGCAGGAAGGCGCTGTGCGGCTTCAGGCTG 4640
 4561 CGTCCAGTACATCCTGAGCTAGTACAGGAACACTTGTGCAGCGGGAAGACGTCTTCCGCGACAGCCGGAAGTCCGAC 4640
L D I Y S E I M D K T F V D G K Q L F A S D A K L S

4641 TCCAGGGCCTCGCGCAGCGAGCCGCAAACCGTCCGAACCTTGGCCAGCTCTTCGGCCGGCAGGTTCGTACAGGTTCTTGTG 4720
 4641 AGGTCCCGGAGCGGTCGCTCGGCGTTCGGCAGCCTTGAACCGGTTCGAGAAGCCGGCCGTCACGATGTCCAAGAACAG 4720
D L A E R L S G C V T P V K A L E E A P L D Y L N K D

4721 CATCGCCTCGCCCGATGGATCTTGTCTGGATGCCGTCCAGACCGGCCATCAGCAGGGCGGCGAAGGCCAGGTACGGGT 4800
 4721 GTAGCGGAGCGGCTACCTAGAACAAGACCTACGGCAGGTCTGGCCGGTAGTCGTCCCGCCGCTTCCGGTCCATGCCCA 4800

M A E G P H I K N Q I G D L G A M L L A A F A L Y P N
 4801 TGGCCGACGGGTCCGGGAAGCGGACCTCGACGCGCTTGCCTTCGGCGAGGCGACGTACGGGATGCGGCAGGAGGCCGAA 4880
 4801 ACCGGCTGCCAGGCCCTTCGCTGGAGCTGCGCGAACGGGAAGCCGCTCCGCTGCATGCCCTACGCCGTCTCCGGCTT
 A S P D P F R V E V R K G K P S A V Y P I R C S A S
 4881 CGGTTGCGGGCCGAGTAGGCCAGCAGAACCGGAGCCTCGTAGCCCGGGACCAGACGCTTGTACGAGTTGGTCGTCGGGTT 4960
 4881 GCCAACGCCCGGCTCATCCGGTCGTCTTGGCCTCGGAGCATCGGGCCCTGGTCTGCGAACATGCTCAACCAGCAGCCCAA 4960
 R N R A S Y A L L V P A E Y G P V L R K Y S N T T P N
 4961 GGTGAAGGCGTTCAGCGCCTTGGCGTGCTTGATGATGCCGCCGATGTAGTACAGCGCCAGTTCCGACAGGTCGGCGTACT 5040
 4961 CCACTTCCGCAAGTTCGCGGAACCGCACGAACACTACTACGGCGGCTACATCATGTCGCGGTCAAGGCTGTCCAGCCGCATGA 5040
 T F A N L A K A H K I I G G I Y Y L A L E S L D A Y Q
 5041 GGTTCGCCGGCGAACAGCGGCTGGCCTTCCTTCCAGATCGACTGGTGCATGTGCATGCCCGAGCCGTTGTGCGCCGAAGACG 5120
 5041 CCAACGGCCGCTTGTGCGCCGACCGGAAGGAAGGTCTAGCTGACCACGTACACGTACGGGCTCGGCAACAGCGGCTTCTGC 512
 N G A F L P Q G E K W I S Q H M H M G S G N D G F V
 5121 GGCTTCGGCATGAAGTTCGCGGTCTTGCCGTAGGCGTGGGCGACATTGTGCACGACGTACTTGTAGTACTGCATGTTGTC 5200
 5121 CCGAAGCCGTACTTCCAGCGCCAGAACGGCATCCGCACCCGCTGTAACACGTGCTGCATGAACATCATGACGTACAACAG 5200
 P K P M F T A T K G Y A H A V N H V V Y K Y Y Q M N D
 5201 GCCGGTTCGGGACCAGCGTGTGCAACTTGATGCCAGCTCATGCTGCGAAGCGGCCACCTCGTGGTGGTGCTTCTCGACCG 5280
 5201 CGGCCACGCTGGTTCGCACAGCTTGAACACTACGGGTCGAGTACGACGCTTCGCCGGTGGAGCACCACCACGAAGAGCTGGC 5280
 G T R V L T D F K I G L E H Q S A A V E H H H K E V P
 5281 GCACGCCCATCTCGGCGAGCAGCTCAGCATCTCGGCGCGCAGGTCCGAGCCGCTGTGACCGGAGCGACCGGGAAGTAG 5360
 5281 CGTGGGGTAGAGCCGCTCGTGGGAGTCGTAGAGCCGCGCTCCAGGCTCGGCGACAGCTGGCCTCGCTGGCCCTTCATC 5360
 V G M E A L V S L M E A R L D S G S D V P A V P F Y
 5361 CCGCCCTTGACGCCCCGGGCGGTGGCCAGGTTGCCGTCCTCATAGTCCTTGTCCGAGGTGTACGGGCCTTCCCTCCGAGTC 5440
 5361 GGCGGAACTGCGGGCCCGCCACCGGTTCCAACGGCAGGAGTATCAGGAACAGGCTCCACATGCCCGGAAGGAGGCTCAG 5440
 G G K V G P R H G L N G D E Y D K D S T Y P G E E S D
 5441 GAACTCGTAGGACACCTTGTTCATCTCGACCTTGAACCTGACGTCGTCGAAGACGAAGAACTCGGCTTCCGGGCCGAAGT 5520
 5441 CTTGAGCATCCTGTGGAACAAGTAGAGCTGGAACCTGAACTGCAGCAGCTTCTGCTTCTTGGAGCCGAAGGCCCGGCTTCA 5520
 F E Y S V K N M E V K F K V D D F V F F E A E P G F Y
 5521 AGGCGGTGTCGCCGATGCCGGCCGACGCCATGTACTTCTCGGCGGCCTTCGCGATACCGCGGGCAGCGGGCGTAGGGC 5600
 5521 TCCGCCACAGCGGTACGGCCGGCTGCGGTACATGAAGAGCCCGCGGAAGCGCTATGGCGCGCCCGTTCGCCCCATCCCG 5600
 A T D G I G A S A M Y K E A A K A I G R P C R A Y P
 5601 TGGCCGGTTCGACGGCTCGTACACGTCGAGAGGATGTTTACGCTCGGCTGGGCCGAGAACGGGTCCATGACGGCGGTGGT 5680
 5601 ACCGGCCAGCTGCCGAGCATGTGCAGCGTCTCCTACAAGTCGAGCCGACCCGGCTCTTGGCCAGGTAAGTCCGCCACCA 5680
 Q G T S P E Y V D C L I N L T P Q A S F P D M V A T T
 5681 CGGGTCGAGCTGGAGGATCATGTCCGACTCGTTGATCGCCTTCCAACCGGCGATCGAGGAGCCGTCGAACATGATGCCGT 5760
 5681 GCCCAGCTCGACCTCCTAGTACAGGCTGAGCAACTAGCGGAAGGTTGGCCGCTAGCTCCTCGGCAGCTTGTACTACGGCA 5760
 P D L Q L I M D S E N I A K W G A I S S G D F M I G D
 5761 CTTCAAACACGTCCTCGTCGATGGTCGAGACGTGCTGAGCGGTGTGGTGCAGCTTGCCGCGGGTTCGGTGAAGCGGAGG 5840
 5761 GAAGTTTGTGCAGGAGCAGCTACCAGCTCTGCACGACTCGCCACACCACGTCGAACGGCGCGCCAGCCACTTCGCTCC 5840
 E F V D E D I T S V H Q A T H H L K G R P D T F R L
 5841 TCCACGTAAGTTCGTCGATGGTTCCTTGATCAGGTCGAAGACCTTGCTGATGTCGGACATATCTTTTTTCCACGCTTAAG 5920
 5841 AGGTGCATGAACTGCAGCACAAGGAACCTAGTCCAGCTTCTGGAACGACTACAGCCTGTATAGAGAAAAAGGTGCGAATTC 5920
 D V Y K V D H E K I L D F V K S I D S M
 GLUTAMINE SYNTHETASE (GLUTAMATE--AMMONIA LIGASE)
 5921 AGGTTTGGGTCCCGCTGCGTCAAACCGACCAAGGGGCGGTTGACGGCGCTCGCGGGCGGACGTTATGTCATGCCTTTCCCA 6000
 5921 TCAAACCCAGGGCGACGAGTTTGGCTGGTTCCCCGCCACTGCCGCGAGCGCCGCTGCAATACAGTACGGAAAGGTT 6000

6001 AGCACACAACGTGACAGCACAAAAAGTCGCGCCCGGATGGGCATCCACGCACGGATCGTTCCGCTCGCGGTTGCGACGGT 6080
6001 TCGTGTGTTGCACTGTCTGTTTTCAGCGCGGGCCTACCCGTAGGTGCGTGCCTAGCAAGGCGAGCGCCAACGCTGCCA 6080

6081 CTGCGTCGGCCCCGATCGGAGTCGGCCGGCTCTGCGTCCGGTGTGGGCGGGGCTGGTTCATGGTGGCCGGTGGTCCGGTTC 6160
6081 GACGCAGCCGGGCTAGCCTCAGCCGGCCGAGACGCAGCCAACACCCGCCCGACCAAGTACCACCGGCCACCAGCCCAAG 6160

6161 TCGGGACGCGATCAGATCGCGTCGCCTCCCTTCTCGCCGGTCCGGATGCGGACAACCTTCTCCACGGGGGTGACGAAGAT 6240
6161 AGCCCTGCGCTAGTCTAGCGCAGCGGAGGGAAGAGCGGCCAGGCCTACGCCTGTTGAAGAAGGTGCCCCACTGCTTCTA 6240
* I A D G G K E G T R I R V V E E V P T V F I

6241 CTTGCCGTCGCCGATGCGGGCCGGTGTGGGCGGCCTGCTGGATCGCCTCGATCGCCCGCTCCACCAGGGAGTCTCCATCA 6320
6241 GAACGGCAGCGGCTACGCCGGCCACACCCGCCGGACGACCTAGCGGAGCTAGCGGGCGAGGTGGTCCCTCAGGAGGTAGT 6320
K G D G I R G T H A A Q Q I A E I A R E V L S D E M V

6321 CCACCTCGATCTTACCTTCGGCAGGAAGTCGACCACATACTCCGCGCCCGGTACAGCTCGGTGTGCCCTTCTGACGG 6400
6321 GGTGGAGCTAGAAGTGAAGCCGTCCTCAGCTGGTGTATGAGGCGCGGCCATGTCGAGCCACACGGGGAAGACTGCC 6400
V E I K V K P L F D V V Y E A G R Y L E T H G K Q R

6401 CCGAAGCCCTTGGCCTCGGTGACGGTGATGCCCTTGATGCCGACTTCGTGAAGGGCTTCCCTCACTTCGTCGAGTTTGAA 6480
6401 GGCTTCGGGAACCGGAGCCACTGCCACTACGGGAACCTACGGCTGAAGCACTTCCCGAAGGAAGTGAAGCAGCTCAAATT 6480
G F G K A E T V T I G K I G V E H L A E K V E D L K F

6481 CGGCTTAATGATGGCTTCGATCTTCTTCATGGGTCTCTCGTGTACGAGGTGCCAACGTCGGCAGGCACCGTCCACGGCTC 6560
6481 GCCGAATTACTACCGAAGCTAGAAGAAGTACCCAGAGAGCACATGCTCCACGGTTGCAGCCGTCCGTGGCAGGTGCCGAG 6560
P K I I A E I K K M
Nitrogen regulatory protein P-II

6561 TGCCGACGGTCCGATGTAAAGCACGTTGCGTGCCAATCGGAGCGGTTGCGATCCGGCTGCGTGGGATGCCGGGCGCAGC 6640
6561 ACGGCTGC CAGGCTACATTTTCGTGCAACGCACGGTTAGCCTCGCCAACGCTAGGCCGACGCACCCTACGGCCCCGGCTCG 6640
-60

6641 GGGCCGATGCCACCGTGGGCGGCACGCGACTGCCAACGGATAGGGCGAAGTGAACAATATTTGTGCACATGCTGCAAGCA 6720
6641 CCCGGCTACGGTGGCACCCGCGGTGCGCTGACGGTTGCCTATCCCGCTTCACTTGTATAAACACGTGTACGACGTTTCGT 6720

6721 ATCCGCGCAGCCTTGTCCGTTCCGATGGGGTAGCGCCGGCTTCCACCCGGCTTTTTCACATTGATGGACAAGGTGGAAAA 6800
6721 TAGGCGCGTCGGAACAGGCAAGCCTACCCCATCGCGGCCGAAAGGTGGGCCGAAAAGTGTAACTACCTGTTCCACCTTTT 6800

6801 TAGAGCTTGACCCGCCCCGTCCTTGGGGCAAATAGGCGACCCCGGTGGCGGTGGTAGCTCAGTTGGTAGAGCTCTCGGT 6880
6801 ATCTCGAACTGGGCGGGCAGGGAACCCGTTTATCCGCTGGGGCCACCGCCACCATCGAGTCAACCATCTCGAGAGCCA 6880
-310

6881 TGTGGTCCGAGCGGTCGTGGGTTTCGAGTCCCATCCATCGCCCCAGTCCCTTTCGAGCGGATGTGTTCGGTTTCGTGAGC 6960
6881 ACACCAGGCTCGCCAGCACCAAGCTCAGGGTAGGTAGCGGGGTCAGGAGGAAGTCCCGCTACACAGCCAAAGCACTCG 6960

6961 AGTTCACGCCGGCGTTGCATCGGTTCGGGGCGACGGGTCGGGTGTTTGCAGCAATCTTTTCGTCGCCCGGTGGCGGTGGT 7040
6961 TCAAGTGCGGCCGCAACGTAGCCAGCCCCGCTGCCAGCCACAAACGCCTGCTTAGAAAAGCAGGGGCCACCGCCACCA 7040

7041 AGCTCAGTTGGTAGAGCTCTCGGTTGTGGTCCGAGCGGTCGTGGGTTTCGAGTCCCATCCATCGCCCCAGTCCCTTCTTCC 7120
7041 TCGAGTCAACCATCTCGAGAGCCAACACCAGGCTCGCCAGCACCCAAGCTCAGGGTAGGTAGCGGGGTCAGGAAAGAAGG 7120
-660

7121 AATTCCTGTGCCAATCAGGCCAGTCCCTTTCATTCCCCCGTGTCGCCGGCTATGCTGCGGCCGTCGCGGTGCTTATTGTGCC 7200
7121 TTAAGGACACGGTTAGTCCGGTCAGGAAAGTAAGGGGGACACGGGCCGATACGACGCCGGCAGCGCCACGAATAACACGG 7200
Magn03009421: Predicted sugar kinase
M D E L L S V A E M Y R A D

7201 GGGTGCAGGTTGATGCGGGATTGCGGGCGGGTGGAGGCGATGGACGAAGTCTTTCGTCGGCGGAGATGTACCGTGCAGC 7280
7201 CCCACTGCCACTACGCCCTAACGCCGCCACCTCCGCTACCTGCTTACGAAAGGCACCGCCTCTACATGGCACGGCTGC 7280

A L T M A G G V P G P V L M E A A G A A V V R A V C E

7281 CGCTGACCATGGCGGGCGGCTGCCGGGCGTGTGCTGATGGAAGCGCGGGGGCCGCGGTGGTGCAGCGGGTCTGCGAG 7360

7281 GCGACTGGTACCGCCCGCCGACGGCCCCGGACACGACTACCTTCGCCGCCCGCCCGGCGCCACCACGCGCGCCAGACGCTC 7360
R W A P H P T V V L C G P G N N G G D G F V I A R L L
7361 CGCTGGGGCGCCGACCCGACCGTCTGTCCTGTGCGGTCCGGGCAACAATGGCGGAGATGGCTTCGTATCGCGCGCCTTCT 7440
7361 GCGACCCGCGGGCTGGGCTGGCAGCAGGACACGCCAGGCCGTTGTTACCGCCTCTACCGAAGCAGTAGCGCGCGGAAGA 7440
L E A G W P V R L A L L G S R S A L R G D A A V A A
7441 TCTCGAAGCGGGCTGGCCGGTGC GGCTGGCCCTGCTGGGGTCCCGCTCCGCCCTGCGCGGCGACGCGGGCGGTGGCGGCCG 7520
7441 AGAGCTTCGCCCAGCCGGCCACGCCAGCCGGGACACCCAGGGCGAGGCGGGACGCGCCGCTGCGCCGCCACCGCCGGC 7520
G H W T G P V E A A D P W L L E G N P L V I D A L F G
7521 GTCACTGGACCGGGCCGGTGGAGGCCGCCACCCCTGGCTGCTGGAGGGCAACCCCTGGTCATCGACGCGCTGTTCCGGG 7600
7521 CAGTGACCTGGCCCGCCACCTCCGGCGGCTGGGGACCGACGACCTCCCGTTGGGGGACCAGTAGCTGCGCGACAAGCC 7600
A G L A R P L D G M A K E V V E A M A G R T V V A V D
7601 GCCGGGCTGGCCGTCGCTCGACGGCATGGCCAAGGAGTCTGTCGAGGCGATGGCGGGGCGCACCGTCTGTCGCCGTGGA 7680
7601 CGCCCCAGCCGGGCGAGCTGCCGTACCGGTTCTCCAGCAGCTCCGCTACCGCCCCGCGTGGCAGCAGCGGCACCT 7680
V P S G L H G D S G Q V M G V A P Q A A L T V T F F
7681 CGTGCCAGTGGTCTGCACGGCGACAGCGGTCAAGTGATGGGGTGGCCCCGAGGCGGGCGCTGACCGTCACCTTCTTCC 7760
7681 GCACGGGTACCAGACGTGCCGCTGTGCCAGTTCCTACCCCCACCGGGGCGTCCGCCGCGACTGGCAGTGAAGAAGG 7760
R R K P G H L L L P G R A L C G E V V V A D I G I P D
7761 GCCGCAAGCCCGCCATCTGCTGCTGCCGGGCCGCGCCCTGTGCCGGCAGGTTGTGGTCGCCGACATCGGCATTCCCGAC 7840
7761 CGCGTTCGGGCCGGTAGACGACGACGGCCCCGGCGGGACACGCCGCTCCAACACCAGCGGCTGTAGCCGTAAGGGCTG 7840
T V L D A I A P Q T L V N G P E L W R Q C F P W P R L
7841 ACCGTGCTGGACGCCATCGCGCCGAAACCCTGGTGAACGGCCCGGAGTTGTGGCGGCAGTGCTTTCCCTGGCCACGGCT 7920
7841 TGGCACGACCTGCGGTAGCGCGGCGTTTGGGACCACCTGCCGGGCCTCAACACCGCCGTCACGAAAGGGACCGGTGCCGA 7920
D A H K Y A R G H A V V L G G A R M T G A A R L S A
7921 CGACGCTCACAAATACGCGCGGGGGCATGCGGTGGTGTCTCGGCGGGGCGCGGATGACCGGTGCGGCGCGCTTGTGCGGCG 8000
7921 GCTGCGAGTGTTTATGCGCGCCCCCGTACGCCACCACGAGCCGCCCGCGCCTACTGGCCACGCCGCGCAACAGCCGCG 8000
R G A L R A G A G L V T V A C P P D S V P I Y A A G S
8001 GCGGTGCCCTGCGCGCGGGGGCGGGGCTGGTCACGGTGGCCTGTCCGCCGACTCGGTGCCGATCTACGCCCGCGGCTCG 8080
8001 CGCCACGGGACGCGCGCCCCCGCCCCGACCAGTGCCACCGGACAGGCGGGCTGAGCCACGGCTAGATGCGGCGGGCCGAGC 808
P S L I V A D V A D G A A F E A L L A D P R K N A A L
8081 CCCAGCCTGATCGTGGCTGACGTGGCGGACGGCGCCGCTTCGAGGCGCTGCTGGCGGACCCACGCAAGAACGCGGCGCT 8160
8081 GGGTCGGACTAGCACCAGCTGCACCGCTGCCGCGGCGGAAGCTCCGCGACGACCGCCTGGGTGCGTTCTTGCGCCGCGA 8160
L G P G A G V G A D T R S H V V A A L G A G K A C V
8161 GCTGGGGCCGGGCGCCGGGCTCGGTGCGGACACGCGGAGCCATGTGGTCGCGGCTCTTGGCGCCGAAAGGCGTGCCTGC 8240
8161 CGACCCCGCCCCGCGGCCGAGCCACGCCTGTGCGCCTCGGTACACCAGCGCCGAGAACC GCGGCCTTTCGCGACGCACG 8240
L D A D A L T S F A E S P S A L L D R L N G R C L L T
8241 TCGACGCCGATGCGCTGACCAGCTTCGCCGAATCGCCATCCGCTCTGCTCGACCGGCTGAACGGGCGCTGCCTGCTGACC 8320
8241 AGCTGCGGCTACGCGACTGGTCTGAAGCGGCTTAGCGGTAGGCGAGACGAGCTGGCCGACTTGCCCCGCGACGGACGACTGG 8320
P H E G E F A R I F P D I A V Q D V G K L A R A R A A
8321 CCGCACGAGGGCGAGTTCGCGCGGATCTTCCCGACATCGCGGTCCAGGACGTTGGCAAACCTGGCGCGCGCCCCGCGCCGC 8400
8321 GGCGTCTCCCGCTCAAGCGCGCCTAGAAGGGCCTGTAGCGCCAGGTCCTGCAACCGTTTGACCGCGCGGGGCGCGGCG 8400
A A R C G A V V L L K G A D T V V A A P D G R A V V
8401 GGCGGCGCGCTGCGGCGCGGTCTGTCCTTCTGAAGGGCGCCGACACGGTGGTTCGCGGCGCCGGACGGACGGGCGGCTCGTCA 8480
8401 CCGCCGCGGACGCGCGCCAGCAGGAAGACTTCCCGCGGCTGTGCCACCAGCGCCGCGGCCTGCCTGCCCGGCAGCAGT 8480
N A N A P A D L A T A G S G D V L A G M A L G L I A Q
8481 ACGCCAACGCTCCGGCGGATCTGGCGACCCCGGGTCCGGGAGACGTGCTGGCGGGCATGGCGCTCGGGCTGATCGCTCAA 8560
8481 TGCGGTTGCGAGGCCGCTAGACCCTGGCGGCCAGCCCTCTGCACGACCCCGTACC GCGAGCCCGACTAGCGAGTT 8560

8561 G M N T F D A A C A A V W L H G A A A T A F G P G L I
 8561 GGAATGAACACGTTTCGACGCCGCTGCGCCGCGGTCTGGCTGCATGGCGCGGGCAACCGGTTTCGGGCGGGTCTGAT 8640
 8561 CCTTACTTGTGCAAGCTGCGGCGGACGCGGCCAGACCGACGTACCGCGCCGCGTGGCGCAAGCCCGGCCAGACTA 8640

A E D L P D A L P G V L K R L K A G A R *
 8641 CGCCGAGGATTTGCCGGATGCGCTGCCGGGGTCTGAAGCGGCTCAAGCGGGGGCGCGCTGACCCTGTTTCGCCAAGG 8720
 8641 GCGGCTCTAAACGGCCTACGCGACGGCCCCACGACTTCGCCGAGTTCGCCCGCGCGACTGGGACAAGCGGGTTCC 8720

8721 CATAACATACGCTGGACAGGGAGGCTTACTCTTAAGGCCGGCTCAAGCTATGGCTTGGCGAGAGCGCCCCGACGCCA 8800
 8721 GTATGTATGCGGACCTGTCCCTCCGAATGAGAATTCGGGCCGAGTTCGATACCGAAGCGCTCTCGGGGGCGTGGCT 8800

Magn03009420

M G
 8801 GACTCTGGGTATGAGCGAGACGCCAGAAATCAAATCCGCCGAGCCCGTTCGGTCCCAGGCCCAACCTCGGGGATGGGAG 8880
 8801 CTGAGACCATACTCGCTCTGCGGTCTTTAGTTTAGGCGGCTCGGGCAAGGCAGGGTCCGGGGTTGGAGCCCTACCCTC 8880

P P G L L D R T L D T L R A G W R E I A A S A R G V
 8881 GCCGCCGGGCTGCTCGACCGCACCTGGACACGCTGCGCGCGGGTGGCGCGAGATCGCCGCTTCGCCCGCGGGGTG 8960
 8881 CGGGCGCCCGGACGAGCTGGCGTGGGACCTGTGCGACGCGCGCCGACCGCGCTCTAGCGGCGAAGGCGGGCGCCCCAC 8960

V G A A P R A N L P K E D A E R L K E Q I E A C L E G
 8961 GTCGGCGCCGCGCCGCGCGCAATCTGCCAAGGAGGACGCGGAGCGGCTGAAGGAGCAGATCGAAGCCTGTCTGGAGGG 9040
 8961 CAGCCGCGGCGCGCGCGGTTAGACGGTTCTCTGCGCTCGCCGACTTCTCTGTCTAGCTTCGGACAGACCTCCC 9040

R G G E V S A R A R A A Q L G R T Y L A L D G Q G R
 9041 CCGCGCGGCGAGGTGTCGGCGCGGGCACGGGCAGCGCAGCTCGGCCGACCTATCTGGCGCTGGACGGGCAGGGGCGGC 9120
 9041 GGCGCCGCGCTCCACAGCCGCGCCGTGCCGTGCGCTCGAGCCGGCGTGGATAGACCGCGACCTGCCGTCCCCGCCG 9120

R N F L M M L A R D F D V D R A A V D Q A M A T F Q T
 9121 GCAATTTCTGATGATGCTGGCGCGGACTTCGACGTGGACCGCGCCGCTCGACCAGGCGATGGCGACCTTCCAGACG 9200
 9121 CGTTAAAGGACTACTACGACCGCGCGTGAAGCTGCACCTGGCGCGGCGGACGCTGGTCCGCTACCGCTGGAAGGTCTGC 9200

A G D P V S R G H A E R A L R R A L E P P R L R L L T
 9201 GCGGGCGATCCGGTGTGCGCGGGCCACGCCGAGCGCGCCCTGCGCCGCGCCTTGGAAACCGCCGCGGCTGCGGCTGCTCAC 9280
 9201 CGCCGCTTAGGCCACAGCGCGCGGTGCGGCTGCGCGGGACGCGGCGGGAACCTTGGCGGCGCCGACGCCGACGAGTG 9280

Q F N G L P E G V K F L V D L R A E L M E M A R G E
 9281 TCAGTTCAACGGCCTGCCGGAAGGGTGAATTCCTGGTGGACCTGCGCGCCGAGCTGATGGAGATGGCACGCGGCGAAC 9360
 9281 AGTCAAGTTGCCGACGGCCTTCCCACCTTAAGGACCACCTGGACGCGCGGCTCGACTACCTCTACCCTGCGCCGCTTG 9360

P L L Q A L E D D L K S L L R N W F D V G F L E M R R
 9361 CGCTGCTCCAGGCGCTGGAGGACGATCTCAAGAGCCTGCTGCGCAACTGGTTCGACGTGGCTTCTGGAGATGCGCCGG 9440
 9361 GCGACGAGTCCGCGACCTCCTGCTAGAGTTCGCGACGCGGTTGACCAAGCTGCAGCCGAAGGACCTCTACGCGGCC 9440

I T W D S P A S L L E K L I K Y E A V H E I S G W R D
 9441 ATCACCTGGGACAGCCCGGCTCGCTCCTGGAAAAGCTCATCAAATACGAGGCGGTGCACGAGATCAGCGGTTGGCGGGA 9520
 9441 TAGTGGACCTGTGCGGGCCGAGCGAGGACCTTTTCGAGTAGTTTATGCTCCGCCACGTGCTCTAGTCGCCAACCGCCCT 9520

L K D R L D S D R R C F A F F H P R M P N E P L I F
 9521 CCTGAAGGACCGCCTCGACTCCGACCGCGCTGCTTCGCCTTCTTCCACCCGCGCATGCCGAACGAGCCGCTCATCTTCG 9600
 9521 GGACTTCTGGCGGAGCTGAGGCTGGCGGCGACGAAGCGGAAGAAGGTGGGCGGCTACGGCTTGTCTGGCGAGTAGAAGC 9600

V E V A L V Q G M S D N V Q A L L D P S A P V C D P K
 9601 TCGAGGTGCTCTGGTCCAGGGCATGTGCGACAACGTCCAGGCGCTGCTCGACCCCTCCGCCCGGTTGCGACCCGAAG 9680
 9601 AGCTCCAGCGAGACCAGTCCCGTACAGCCTGTTGCAGGTCCGCGACGAGCTGGGGAGGCGGGGCCAGACGCTGGGCTTC 9680

S A D S A I F Y S I S N A Q V G L A G I S F G N F L I
 9681 AGCGCCGACAGCGCCATCTTCTATTCATCTCCAACGCCAGGTGGGCTGGCGGGAATCAGCTTCGGCAACTTCTCAT 9760
 9681 TCGGGCTGTGCGGTTAGAAGATAAGGTAGAGGTTGCGGGTCCAGCCGACCGCCCTTAGTCGAAGCCGTTGAAGGAGTA 9760

K R V V D G L T G E F P N I K C F A T L S P I P G F

9761 CAAGCGGGTGGTGGACGGGCTGACCGGCGAATTCGCCAACATCAAATGCTTCGCCACCCTGTGCGCGATTCCCAGGCTTCC 9840
9761 GTTCGCCACCACCTGCCCCACTGGCCGCTTAAGGGGTTGTAGTTTACGAAGCGGTGGGACAGCGGCTAAGGGCCGAAGG 9840

R R W L D R E L A D R G D S L L S P P E A E R I A E A

9841 GCCGCTGGCTGGACCGGGAGCTGGCGGACC GCGGCGATTTCGCTGCTCAGCCCCCGAGGCGGAGCGCATCGCCGAGGCC 9920
9841 CGGCGACCGACCTGGCCCTCGACCGCTGGCGCCGCTAAGCGACGAGTCGGGCGGGCTCCGCCTCGCGTAGCGGCTCCGG 9920

M T A R H D G G E S E T E G D A E E R P V L A G A L A

9921 ATGACCGCCCCGGCACGATGGTGGCGAATCGGAGACGGAAGGCGACGCGGAGGAACGTCCGGTGCTGGCCGGCGCCTTGGC 10000
9921 TACTGGCGGGCGGTGCTACCACCGCTTAGCCTCTGCCTTCCGCTGCGCCTCCTTGAGGCCACGACCGGCCGCGGAACCG 10000

L P D W H T D E E L Q K Q L R G P L M R L C A H Y L

10001 TCTGCCCGACTGGCACACC GATGAGGAGCTTCAGAAGCAGCTTCGCGGCCCGTTGATGAGGCTGTGCGCGCATTACTTGA 10080
10001 AGACGGGTGACCGTGTGGCTACTCCTCGAAGTCTTCGTCGAAGCGCCGGGCAACTACTCCGACACGCGCGTAATGAACT 10080

T T A R A S D R R S E R P G A A P A R A L D P V A H F

10081 CGACCGCGCGCGCTCGGACCGCCGAGCGAACGCGCGGGCGCGCGCCGGCGCGAGCGCTCGACCCGGTCCGCGCACTTC 10160
10081 GCTGGCGCGCGCAGCTGGCGGCGTCGCTTGCCGGCCCGCGCCGCGCGCTCGCGAGCTGGGCCAGCGCGTGAAG 10160

H L S N G A R M E R L N W L G D R S A K G V R Q S C G

10161 CATCTCAGCAACGGGCGCCGCATGGAGCGGCTGAACTGGCTGGGTGACCGTTTCGGCCAAGGGCGTGCGCCAATCCTGCGG 10240
10161 GTAGAGTCGTTGCCGCGGGCGTACCTCGCCGACTTGACCGACCCACTGGCAAGCCGGTTCGCCGACGCGGTTAGGACGCC 10240

M M I N Y R Y K P G E I D A N H E A Y R G E G K V N

10241 GATGATGATCAACTACCGCTACAAGCCGGGGAAATCGACGCCAACACGAAGCGTACCGGGGGGAGGGCAAGGTCAACG 10320
10241 CTACTACTAGTTGATGGCGATGTTTCGGCCCCCTTTAGCTGCGGTTGGTGCTTCGCATGGCCCCCTCCCGTTCCAGTTGC 10320

A S S S I R S L G K G G G *

10321 CCTCCTCCTCCATCCGGTTCGCTGGGGAAAGGGGGCGGATGAGCCGGTTTTCGCCCGCCGCGGGGCGTTCGTCGTCGTCGCT 10400
10321 GGAGGAGGAGGTAGGCCAGCGACCCCTTCCCCCGCCTACTCGGCCAAACGGCGGGCGGCCCGCAGCAGCAGCAGCGCA 10400

10401 CCGGCGCCCCACCAATGGCGTGGCGAGCGGCGCCATTTTGCCTGTTTTCGGCGCGAACCCCTGTGCTATAGAGCGGAC 10480
10401 GGCCGCGGGGGTGGTTACCGCACGCTCGCCGCGGTAACGCGGACAAACGCCGCGCTTGGGACACGATATCTCGCCTG 10480

10481 TTCGCCGGCAGTCCGTTCGGGTGTTCCGCACCCGTCGCGGGCTGTCTATGCCATTGCGGGCGTGGCGGAACTGGTAGACGC 10560
10481 AAGCGGCCGTCAGGCAGCCACAAGGCGTGGGCAGCGCCGACAGATACGGTAACGCCCGCACCGCCTTGACCATCTGCG 10560

10561 GCCGGATTTAGGTTCCGGTGACTTCGGTTCGTGGGGGTTTCGAGTCCCTTCGCCCGCACCATGGCAACAGCTTGGACCGCAT 10640
10561 CGGCCTAAATCCAAGGCCACTGAAGCCAGCACCCCAAGCTCAGGGAAGCGGGCGTGGTACCCTTGTGCAACCTGGCGTA 10640

Trigger factor (TF)

M N I

10641 GGGAACACACGGTTTCGGGCGCCCGGACGGGCGCCGCGGGCCGACAAACGAGATAGGCTTTTCGTTCCGATGAACATCA 10720
10641 CCCTTGTTGTCGCAAAGCCCGGCGGGCTGCCGCGGGCGCCGCGCTGTTGCTCTATCCGAAAGCAAGGCTACTTGTAGT 10720

T E T S A D G L K R E Y K V V I S A Q D I E Q K V Q G

10721 CCGAGACCAGCGCCGACGGCTGAAGCGCGAGTACAAGGTGCTCATTTCCGCCAGGATATCGAACAGAAGGTGCAGGGC 10800
10721 GGCTCTGGTTCGCGGCTGCCGACTTCGCGCTCATGTTCCAGCAGTAAAGGCGGGTCTATAGCTTGTCTTCCACGTCCCC 10800

R L E E L R R T V Q L P G F R P G K V P V A V I K Q R

10801 CGGCTCGAAGAGCTGCGCCGACGGTGCAGCTCCCCGGCTTCCGTCCGGGCAAGGTGCCGGTTGCCGTGATCAAGCAGCG 10880
10801 GCGGAGCTTCTCGACGCGGCTGCCACGTCGAGGGGCCGAAGGCAGGCCGTTCCACGGCAACGGCACTAGTTTCGTCGC 10880

Y G G S V L A E A L E D A I A D S S R Q A L N E R G

10881 CTACGGCGGCAGCGTGTGGCCGAGGCTCTGGAGGATGCGATCGCCGACAGCTCGCGCCAGGCGCTGAACGAGCGCGGCC 10960
10881 GATGCCCGGTCGACGACCGGCTCCGAGACCTCCTACGCTAGCGGCTGTGAGCGCGGTCCGCGACTTGTCTGCGCCGG 10960

L R I A M Q P K I N V E K Y E D G G D L S Y T M G V E

10961 TGCGCATGCCATGCAGCCGAAGATCAACGTCGAGAAGTACGAGGACGGCGGCGACCTGTCCTACACCATGGGCGTGCAG 11040

| | | |
|-------|---|-------|
| 10961 | ACGCGTAGCGGTACGTCGGCTTCTAGTTGCAGCTCTTCATGCTCCTGCCGCCGCTGGACAGGATGTGGTACCCGCAGCTC | 11040 |
| | L L P D I E P G D L S G L E L E K P V A T V E D S A V | |
| 11041 | CTTCTGCCGGACATCGAGCCGGGCGATCTCAGCGGCCCTGGAGCTTGAGAAGCCGGTCGCCACGGTCGAGGACTCGGCCGT | 11120 |
| 11041 | GAAGACGGCCTGTAGCTCGGCCCGCTAGAGTCGCCGGACCTCGAACTCTTCGGCCAGCGGTGCCAGCTCCTGAGCCGGCA | 11120 |
| | D E A L T R L A S A H S V Q A P V T E D R A A E K G | |
| 11121 | CGACGAGGCGCTGACCCGTCTGGCCTCCGCCATTCGGTGCAGGCCCGGTACCCGAGGACCGCGCGGCCGAGAAGGGCG | 11200 |
| 11121 | GCTGCTCCGCGACTGGGCAGACCCGGAGCGGGTAAGCCACGTCCGGGGCCAGTGGCTCCTGGCGCGCCGGCTCTTCCCGC | 11200 |
| | D I A V I D F A G S V D G E A L P G M D G K D Y P L E | |
| 11201 | ACATCGCCGTGATCGACTTCGCCGGTTCGGTTCGATGGCGAGGCCCTGCCGGCATGGACGGCAAGGACTACCCGCTGGAG | 11280 |
| 11201 | TGTAGCGGCACTAGCTGAAGCGGCCAAGCCAGCTACCGCTCCGGGACGGCCGTACCTGCCGTTCTGATGGGCGACCTC | 11280 |
| | L G A N Q F V P G F E D Q L V G A K A G E H R T V K V | |
| 11281 | CTGGGCGCAACCAGTTCGTTCCGGGCTTCGAGGACCAGCTGGTCCGGCGCAAGGCCGGCGAGCACCGCACGGTGAAGGT | 11360 |
| 11281 | GACCCGCGGTTGGTCAAGCAAGGCCCGAAGCTCCTGGTGCACCAGCCGCGGTTCCGGCCGCTCGTGGCGTGCCACTTCCA | 11360 |
| | T F P A D Y P H D R L K G A D T V F E V D V K E L R | |
| 11361 | CACCTTCCCGCCGACTACCCGCACGACCGCCTGAAGGGCGCGGACACCGTGTTCGAGGTGGACGTCAAGGAACTCCGCA | 11440 |
| 11361 | GTGGAAGGGCCGGCTGATGGGCGTGTGGCGGACTTCCCGCGCCTGTGGCACAAGCTCCACCTGCAGTTCCTTGAGGCGT | 11440 |
| | K N V P A E V N D D L A K E F G M E S L E K L R E A V | |
| 11441 | AGAACGTCCCGGCTGAGGTCAACGACGACCTTGCCAAGGAGTTCGGCATGGAGAGCCTGGAGAAGCTGCGCGAGGCCGTC | 11520 |
| 11441 | TCTTGCAGGGCCGACTCCAGTTGCTGCTGGAACGGTTCCTCAAGCCGTACCTCTCGGACCTCTTCGACGCGCTCCGGCAG | 11520 |
| | G D R I K G E Y A N V S R L R V K R Q L L D K L A E A | |
| 11521 | GGCGACCGCATCAAGGGCGAATACGCCAACGTGTGCGCCTGCGCGTCAAGCGCCAGCTCCTCGATAAGCTGGCCGAGGC | 11600 |
| 11521 | CCGCTGGCGTAGTTCCCGCTTATGCGGTTGCACAGCGCGGACGCGCAGTTCGCGGTCGAGGAGCTATTTCGACCGGCTCCG | 11600 |
| | H S F E V P P G M V D V E F E G I W Q R L Q Q E L Q | |
| 11601 | GCACTCCTTCGAAGTCCCGCCGGGCATGGTGGACGTGAGTTCGAGGGCATCTGGCAGCGCCTTCAGCAGGAGCTCCAGA | 11680 |
| 11601 | CGTGAGGAAGCTTCAGGGCGGCCCGTACCACCTGCAGCTCAAGCTCCCGTAGACCGTCGCGGAAGTCGTCTCGAGGTCT | 11680 |
| | N G T A G E D A G K P E E E L K T E Y R G I A E R R V | |
| 11681 | ACGGCACCCCGGTGAGGACGCCGGCAAGCCGGAAGAGGAGCTGAAGACCGAGTACC GCGGCATCGCCGAGCGCCGCGTG | 11760 |
| 11681 | TGCCGTGGCGGCCACTCCTGCGGCCGTTCCGCCCTTCTCCTCGACTTCTGGCTCATGGCGCCGTAGCGGCTCGCGGCGCAC | 11760 |
| | R L G L L L S E I G R R N D I Q V T Q D E I N R A L I | |
| 11761 | CGCCTGGGCTGCTGCTGTCGGAGATCGGCCGCCGCAACGACATCCAGGTGACCCAGGACGAGATCAACCGCGCCCTGAT | 11840 |
| 11761 | GCGGACCCGGACGACGACAGCCTTAGCCGGCGGCGTGTGCTGTAGGTCCACTGGGTCTGCTCTAGTTGGCGCGGGACTA | 11840 |
| | A E A R R F P G Q E R Q V F E F F K Q N R E A L E N | |
| 11841 | CGCGAAGCCCGCTTCCCGGCCAGGAGCGTCAGGTGTTTCGAGTTCCTCAAGCAGAACC GGGAAAGCGCTGGAAAACC | 11920 |
| 11841 | GCGCCTTCGGGCGGCAAGGGGCCGTCCTCGCAGTCCACAAGCTCAAGAAGTTCGTCTTGGCCCTTCGCGACCTTTTGG | 11920 |
| | sítio para σ^{32} | |
| | L R A P I F E D K V V D Y I L D Q A K V S E K P V S A | |
| 11921 | TCCGCGCCCCGATCTTCGAGGACAAGGTGTCGACTACATCCTGGACCAAGCCAAGGTGAGCGAGAAGCCGGTGTCCGCC | 12000 |
| 11921 | AGGCGCGGGCTAGAAGCTCCTGTTCCAGCAGCTGATGTAGGACCTGGTTCGGTTCCTACTCGCTCTTCGGCCACAGGCGG | 12000 |
| | E E L M K D P D E E A E A A * | |
| 12001 | GAAGAGCTGATGAAGACCCCGACGAGGAGGCCGAGGCCGCTGACGTGAGAAAGGTTGAGGGGCGGCCTACCGGTGCGCC | 12080 |
| 12001 | CTTCTGACTACTTCTGGGGCTGCTCCTCCGGCTCCGCCGACTGCAGTCTTTCCAACCTCCCGCCGGATGGCCAGCGG | 12080 |
| | ClpP | |
| | M Y D F E | |
| 12081 | CCTTTCATTCCTATATGTAGACTGCTCAGACGACTAACCGAGCTGCGGGAGCAGGGAAGACCACCATGTACGACTTCGA | 12160 |
| 12081 | GGAAAGGTAAGGATATACATCTGACGAGTCTGCTGATTGGCTCGACGCCCTCGTCCCTTCTGGTGGTACATGCTGAAGCT | 12160 |

| | | |
|-------|---|-------|
| | P K M N A L V P M V I E Q T N R G E R A Y D I Y S R | |
| 12161 | ACCGAAGATGAATGCTCTGGTCCCGATGGTCATCGAGCAGACCAACCGGGGCGAACGAGCGTACGATATTTACTCGCGCC | 12240 |
| 12161 | TGGCTTCTACTTACGAGACCAGGGCTACCAGTAGCTCGTCTGGTTGGCCCCGCTTGCTCGCATGCTATAAATGAGCGCGG | 12240 |
| | L L K E R I I F L I G G V N D A V A S L I C S Q L L F | |
| 12241 | TGCTGAAGGAACGGATCATCTTCTTGATCGGTGGCGTGAACGACGCCGTGCGCCAGCCTGATCTGCTCGCAGCTGCTGTTC | 12320 |
| 12241 | ACGACTTCCTTGCCTAGTAGAAGAAGTAGCCACCGCACTTGCTGCGGCAGCGGTGCGACTAGACGAGCGTGCACGACAAG | 12320 |
| | L E S E N P S K D I A L Y I N S P G G Y V S A G L A I | |
| 12321 | CTGGAATCGGAAAACCCGAGCAAGGACATCGCGCTCTACATCAATTCGCCGGGCGGATACGTCTCGGCCGGCCTCGCCAT | 12400 |
| 12321 | GACCTTAGCCTTTTGGGCTCGTTCCTGTAGCGCGAGATGTAGTTAAGCGGCCCGCCTATGCAGAGCCGGCCGGAGCGGTA | 12400 |
| | Y D T M Q Y I R P Q V S T V C M G Q A A S M G S L L | |
| 12401 | CTACGACACCATGCAGTACATCCGTCCGCAGGTGTGACGGTGTGCATGGGGCAGGCGGCCCTCGATGGGCTCGCTGCTGC | 12480 |
| 12401 | GATGCTGTGGTACGTCATGTAGGCAGGCGTCCACAGCTGCCACACGTACCCCGTCCGCCGGAGCTACCCGAGCGACGACG | 12480 |
| | L A A G A P G K R F S L P N S R I M I H Q P S G G A Q | |
| 12481 | TCGCCCGCGGTGCTCCGGGCAAGCGCTTCTCGCTGCCGAACAGCCGGATCATGATCCACCAGCCGTGCGGCGGGCGCCAG | 12560 |
| 12481 | AGCGGCGGCCACGAGGCCGTTGCGGAAGAGCGACGGCTTGTCGGCCTAGTACTAGGTGGTTCGGCAGCCCGCCGGGTC | 12560 |
| | G Q A S D I E I Q A Q E I L K L R S R L N D I Y V K H | |
| 12561 | GGTCAGGCCTCGGACATCGAGATCCAGGCTCAGGAGATCCTGAAGCTGCGTTCGCGCCTCAACGACATCTACGTCAAGCA | 12640 |
| 12561 | CCAGTCCGGAGCCTGTAGCTCTAGGTCCGAGTCTCTAGGACTTCGACGCAAGCGCGGAGTTGCTGTAGATGCAGTTTCGT | 12640 |
| | T G Q S L D T I E A A M E R D K F M S P E E A K A F | |
| 12641 | CACCGGCCAGTCGCTCGACACGATCGAAGCGGCCATGGAGCGTGACAAGTTCATGTGCGCCGGAGGAGGCCAAGGCCTTCG | 12720 |
| 12641 | GTGGCCGTCAGCGAGCTGTGCTAGCTTCGCCGGTACCTCGCACTGTTCAAGTACAGCGGCCTCCTCCGGTTCGGGAAGC | 12720 |
| | G L I D E V V E K R P G S I G D G A A * | |
| 12721 | GGCTGATCGACGAGGTGGTGGAGAAGCGCCCCGGTTCGATCGGCGACGGCGCCGCCTGAGGGGGCGCGCAACTAATGAT | 12800 |
| 12721 | CCGACTAGCTGCTCCACCACCTCTTCGCGGGGCCAAGCTAGCCGCTGCCGCGGCGGACTCCCCGCGCGCTTGATTACTA | 12800 |
| 12801 | TGATTGCGCCGGTTCCTGTTGTTGAATGATGGGGAACCGGTCCGATGGATCGGGTTCGAAGGATGAGCGGATGCGTTAA | 12880 |
| 12801 | ACTAACCGGCCAAGGACAACAACCTACTACCCCTTGGCCAGGCTACCTAGCCCAAGCTTCTACTCGCTACGCAATT | 12880 |
| 12881 | CCTTGTCACGGCCATGCCCGGACAACCTTTGAAAGGACCGGCGTGGGGGACGCCGAGGCGGTTGACAGGTTATACG | 12960 |
| 12881 | GGAACAGTGCCCGGTACGGGCCCTGTTGGAACCTTTTCCTGGCCGACGCCCTGCGGCGTCCGCCAACTGTCCAATATGC | 12960 |
| | ClpX | |
| | M S K S S S G D S K N T L Y C S | |
| 12961 | GATCAGTCCCCCGGAGATGACGGAGTACCGATGAGCAAGTCCAGCAGCGGCGATTTCGAAGAACACGCTCTACTGCTCCT | 13040 |
| 12961 | CTAGTCAGGGGGCGCTCTACTGCCCATGGCTACTCGTTCAGGTGCTCGCCGCTAAGCTTCTTGTGCGAGATGACGAGGA | 13040 |
| | F C G K S Q H E V R K L I A G P T V F I C D E C V E L | |
| 13041 | TTTGCGGCAAGAGCCAGCACGAGTCCGCAAGCTGATTGCCGGTCCGACGGTGTTCATCTGCGATGAATGCGTTCGAGCTG | 13120 |
| 13041 | AAACGCCGTTCTCGGTGCTGCTCCAGGCGTTCGACTAACGGCCAGGCTGCCACAAGTAGACGCTACTTACGCAGCTCGAC | 13120 |
| | C M D I I R E E N K T T L V K S R D G V P T P R D I H | |
| 13121 | TGCATGGACATCATTCGCGAGGAGAACAAGACGACCCTCGTGAAGTCCCAGCAGCGGCTGCCGACCCCGCGGACATCCA | 13200 |
| 13121 | ACGTACCTGTAGTAAGCGCTCCTCTTGTTCGCTGGGAGCACTTCAGGGCGCTGCCGACGGCTGGGGCGCGCTGTAGGT | 13200 |
| | A V L D D Y V I G Q H H A K R V L S V A V H N H Y K | |
| 13201 | TGCGGTACTGGACGATTACGTGATCGGCCAGCATCACGCGAAACGGGTCTTTCTGTTGCGGTGCACAACCACTACAAGC | 13280 |
| 13201 | ACGCCATGACCTGCTAATGCACTAGCCGGTCTAGTGCCTTTGCCAGGAAAGACAACGCCACGTGTTGGTGATGTTTCG | 13280 |
| | R L A H G T K H N D V E L A K S N I L L V G P T G C G | |
| 13281 | GGTTGGCGCACGGCACGAAGCACAACGACGTCGAACTGGCGAAGTCCAACATCCTGCTGGTTCGGCCCCACGGGCTGCGGC | 13360 |
| 13281 | CCAACCGCGTCCGTGCTTCGTGTTGCTGCAGCTTGACCGCTTCAGGTTGTAGGACGACCAGCCGGGGTGCCTGACGCGC | 13360 |

K T L L A Q T L A R I I D V P F T M A D A T T L T E A

13361 AAGACGCTGCTCGCCAGACGCTCGCCCGGATCATCGACGTTCCCTTACGATGGCCGACGCCACGACCCTGACCGAGGC 13440
13361 TTCTGCGACGAGCGGGTCTGCGAGCGGGCCTAGTAGCTGCAAGGGAAGTGCTACCGGCTGCGGTGCTGGGACTGGCTCCG 13440

G Y V G E D V E N I I L K L L Q A A D Y N V E R A Q

13441 GGGCTATGTCGGCGAGGACGTCGAGAACATCATCCTCAAGCTGCTCCAGGCCGCCACTACAACGTCGAGCGGGCGCAGC 13520
13441 CCCGATACAGCCGCTCCTGCAGCTCTTGTAGTAGGAGTTCGACGAGGTCCGGCGGCTGATGTTGCAGCTCGCCCGCTGC 13520

R G I V Y I D E V D K I S R K S D N P S I T R D V S G

13521 GCGGCATCGTCTACATCGACGAGGTCGACAAGATCAGCCGCAAGTCCGACAACCCGTCGATCACGCGGACGTCGCGGC 13600
13521 CGCCGTAGCAGATGTAGCTGCTCCAGCTGTTCTAGTCGGCGTTCAGGCTGTTGGGCAGCTAGTGCAGCTGCACAGCCCG 13600

E G V Q Q A L L K I M E G T V A S V P P Q G G R K H P

13601 GAGGGCGTGCAGCAGGCCCTGCTGAAGATCATGGAGGGCACCGTCGCCTCCGTGCCTCCGCAGGGTGGCCGCAAGCACCC 13680
13601 CTCCCGCACGTCGTCGGGACGACTTCTAGTACCTCCCGTGGCAGCGGAGGCACGGAGGCGTCCCACCGGCGTTCGTGGG 13680

Q Q E F L Q V D T S N I L F I C G G A F A G L D K I

13681 GCAGCAGGAGTTCCTGCAGGTCGACACGAGCAACATCCTGTTTCATCTGCGGGCGGCGCCTTCGCCGGGCTGGACAAGATCA 13760
13681 CGTCGTCTCAAGGACGTCAGCTGTGCTCGTGTAGGACAAGTAGACGCCGCCGGAAGCGGCCCGACCTGTTCTAGT 13760

I A Q R G K G T S I G F G A D V R G P D E R S T G D I

13761 TCGCGCAGCGCGGCAAGGGCACGTCGATCGGCTTCGGCGCTGATGTGCGCGGCCCGGACGAGCGTTCGACCGGCGACATC 13840
13761 AGCGCGTCGCGCCGTTCCCGTGCAGCTAGCCGAAGCCGCGACTACACGCGCCGGGCGTCTCGCAAGCTGGCCGCTGTAG 13840

L R E V E P E D L L K F G L I P E F I G R L P V V A T

13841 CTGCGTGAGGTCGAGCCCGAGGATCTGCTGAAATTCGGCCTGATCCCCGAGTTCATCGGTGCTCTGCCGGTGGTCGCCAC 13920
13841 GACGCACTCCAGCTCGGGCTCCTAGACGACTTTAAGCCGGACTAGGGGCTCAAGTAGCCAGCAGACGGCCACCAGCGGTG 13920

L S D L D E T A L V E I L T K P K N A L V K Q Y Q R

13921 CCTGTCCGACCTCGACGAAACGGCTCTGGTTCGAGATCCTCACCAAGCCGAAGAATGCCCTGGTCAAGCAGTACCAGCGCC 14000
13921 GGACAGGCTGGAGCTGCTTTGCCGAGACCAGCTCTAGGAGTGGTTCGGCTTCTTACGGGACCAGTTCGTTCATGGTTCGCGG 14000

L F E M E D V R L E F S D D A L R T I S H K A I Q R K

14001 TGTTTCGAGATGGAGGACGTCGCTGGAATTCCTCCGACGACGCCCTGCGCACCATCTCCACAAGGCCATCCAGCGGAAG 14080
14001 ACAAGCTTACTCCTGCAGGCAGACCTTAAGAGGCTGCTGCGGGACGCGTGGTAGAGGGTGTTCGGGTAGGTCGCCTTC 14080

T G A R G L R S I M E S I L L D P M F D L P G L S G V

14081 ACCGGCGCACGCGGCCTGCGGTGATCATGGAGTGCATCCTGCTCGACCCGATGTTTCGATCTGCCCGGCCTGTCCGGGGT 14160
14081 TGGCCGCTGCGCCGGACGCCAGCTAGTACCTCAGCTAGGACGAGCTGGGCTACAAGCTAGACGGGCCGGACAGCCCCCA 14160

E S I L V N K E V V E G R A K P L Y V H A E R R G E

14161 GGAAAGCATCCTGGTCAACAAGGAGGTGGTTCGAGGGGCGGCCAAGCCGCTTTACGTCCACGCCGAACGCCGCGGGGAGC 14240
14161 CCTTTCGTAGGACCAGTGTTCCTCCACCAGCTCCCCGCGGGTTCGGCGAAATGCAGGTGCGGCTTGCGGCGCCCCTCG 14240

Q Q A P G A *

14241 AGCAGGCACCGGGTGCCTGAGTTCGTCACGGCCGATCAGCATCGGGTACGAGGAAGGGCGCCTTTGGGCGCCCTTTGTCTG 14320
14241 TCGTCCGTGGCCACGGACTCAAGCAGTGCCGGCTAGTCTGATGCCATGCTCCTTCCCGCGAAACCCGCGGAAACAGC 14320

T T T C C G G T T T T G C A G A C A T C G A G G G C C C A G C C G G G C G G G G C G G T C G G A T C G G C A T G T C G C A C C G C C A T A C G G A A T

14321 TTTCCGGCTTTTGCAGACATCGAGGGGCCAGCCGGGCGCGGGGCGGTCGGATCGGCATGTCGCACCGCCATACGGAAT 14400
14321 AAAGGCCGAAAACGCTCTGTAGCTCCCGGGTTCGGCCCGGCGCCCGCCAGCCTAGCCGTACAGCGTGGCGGTATGCCTTA 14400

T G G G T T C T G G C C T T G A A G G G G G C T A T G G C T G T G C C A C G T T A G A A A A G T G C C G G T T T T C G A T C C C T G T G C T C A T C C C

14401 TGGGTTCTGGCCCTTGAAGGGGGGCTATGGCTGTGCCACGTTAGAAAAAGTGCCGGTTTTTCGATCCCTGTGCTCATCCC 14480
14401 ACCCAAGACCGGGAACCTCCCCCGATACCGACACGGTGAATCTTTTTTACGGCCAAAAGCTAGGGACACGAGTAGGG 14480

Lon

M K E A Q S M F E I P

14481 CCGAGGCGGGGCTTCCCAGGCCATCCCGGGCTGACGGCGTCCCAATGAAAGAGGCCCAATCAATGTTTCGAAATCCCT 14560
14481 GGCTCCGCCCCAGAAGGGTCCGGGTAGGGCCCGACTGCCGAGGGTTACTTTCTCCGGTTAGTTACAAGCTTTAGGGA 14560

terminador

→ ←

R G A L Y P V L P L R D I V V F P H M I V P L F V G R

14561 CGTGGTGCCTCTATCCGGTCCCTGCCGCTCCGCGACATCGTGGTTTTCCCCACATGATCGTGCCTCTTTTCGTGCGGCCG 14640
 14561 GCACCACGGGAGATAGGCCAGGACGGCGAGGGCGTGTAGCACAAAAGGGGGTGTACTAGCACGGAGAAAAGCAGCCGGC 14640

E K S V R A L E D V M K D D K Q I L L V T Q K N A A

14641 TGAGAAATCCGTGCGCGCCCTGGAAGACGTGATGAAGGACGATAAGCAGATCCTTCTCGTCACCCAGAAGAACGCCGCGC 14720
 14641 ACTCTTTAGGCACGCGCGGGACCTTCTGCACTACTTCCCTGCTATTCGTCTAGGAAGAGCAGTGGGTCTTCTTGCGGCGCG 14720

Q D D P T P A D I Y S V G T V G T V L Q L L K L P D G

14721 AGGACGATCCGACGCCGCGGATATCTACAGCGTCGGCACCGTTGGGACCGTGTTCAGCTGCTGAAGCTGCCCGACGGA 14800
 14721 TCCTGCTAGGCTGCGGCCGGCTATAGATGTCGACGCGTGGCAACCCTGGCACAACGTCGACGACTTCGACGGGCTGCCT 14800

T V K V L V E G G Q R A S I T K F A E N E D F F Q A H

14801 ACGGTGAAGGTGCTCGTGGAGGGCGGCCAACGCGCGTCCATCACCAAGTTCGCCGAGAACGAGGATTTCTTCCAGGCCCA 14880
 14801 TGCCACTTCCACGAGCACCTCCCGCCGGTTGCGCGCAGGTAGTGGTTC AAGCGCTCTTGCTCCTAAAGAAGGTCCGGGT 14880

A D L V E E K V G E S Q E L E A L G R A V V S Q F E

14881 CGCCGACCTCGTCGAGGAGAAGTTCGGGAAAGCCAGGAACCTTGAGGCGCTGGGACGCGCGGTCTCTCGCAATTCGAGC 14960
 14881 GCGGCTGGAGCAGTCTCTTCCAGCCCTTTTCGGTCTTGAACCTCCGCGACCCTGCGCGCCAGCAGAGCGTTAAGCTCG 14960

Q Y I K L N K K I P P E V L V S I N Q I E E P G K L A

14961 AGTACATCAAGCTGAACAAGAAGATCCCGCCGGAGGTCTGGTCTCGATCAACCAGATCGAAGAGCCGGGAAAGCTGGCG 15040
 14961 TCATGTAGTTGACTTGTCTTCTAGGGCGGCCCTCCAGGACCAGAGCTAGTTGGTCTAGCTTCTCGGCCCTTTTCGACCGC 15040

D T V A S H L A L K I P E K Q Q L L E C A T V S E R L

15041 GACACCGTCGCCTCCACCTCGCGCTGAAGATTCGGAAAAGCAGCAGCTTCTGGAATGCGCCACGGTTTCGGAGCGGTT 15120
 15041 CTGTGGCAGCGGAGGTGGAGCGGACTTCTAAGGCCTTTTCGTGCTCGAAGACCTTACGCGGTGCCAAAGCCTCGCCAA 15120

E R V Y A F M E G E I G V L Q V E K R I R N R V K R

15121 GGAGCGGGTCTACGCCCTCATGGAAGGCGAAATCGGTGTCTCCCAAGTGGAAAAGCGCATCCGCAACCGCGTCAAGCGGC 15200
 15121 CCTCGCCAGATGCGGAAGTACCTTCCGCTTTAGCCACAGGAGTTACCTTTTCGCGTAGGCGTTGGCGCAGTTTCGCCG 15200

Q M E K T Q R E Y Y L N E Q L K A I Q K E L G E T E D

15201 AGATGGAGAAGACCCAGCGGAGTACTACCTCAACGAACAGCTCAAGGCGATCCAGAAGGAACCTCGGCGAGACCGAGGAC 15280
 15201 TCTACCTTCTTGGTTCGCGCTCATGATGGAGTTGCTTGTGAGTTCCGCTAGGTCTTCTTGGAGCCGCTCTGGCTCCTG 15280

G R D E S A E L E E K I N K T R F S K E A R D K A L A

15281 GGCCGTGACGAGTCCGCGGAGCTGGAAGAGAAGATCAACAAGACCCGCTTCTCCAAGGAAGCCCGCGACAAGGCCCTGGC 15360
 15281 CCGGCACTGCTCAGCCGGCTCGACCTTCTTCTTAGTTGTTCTGGGCGAAGAGGTTTCTTCCGGGCGCTGTTCCGGGACCG 15360

E L K K L R S M S P M S A E A T V V R N Y L D W M L

15361 CGAGCTGAAGAAGTTCGCTCCATGAGCCCGATGTCGGCCGAGGCGACGGTGGTGCGCAACTATCTGGACTGGATGCTGT 15440
 15361 GCTCGACTTCTTCGACGCGAGGTA TCGGGCTACAGCCGGCTCCGCTGCCACCACGCGTTGATAGACCTGACCTACGACA 15440

S I P W K K R T K V K K D L K L A Q K I L D A D H Y G

15441 CCATTCCGTGGAAGAAGCGCACCAAGGTGAAGAAGGACCTGAAGCTCGCGCAGAAGATCCTCGACGCCGACCACTACGGC 15520
 15441 GGTAAGGCACCTTCTTCGCGTGGTTCCACTTCTTCCCTGGACTTCGAGCGCGTCTTCTAGGAGCTGCGGCTGGTGTATGCC 15520

L E K V K E R I L E Y L A V Q N R M N K V K G P I L C

15521 CTGGAGAAGGTCAAGGAACGCATCCTCGAGTATCTTGGCGTCCAGAACC GCATGAACAAGGTCAAGGGGCCGATCCTGTG 15600
 15521 GACCTCTTCCAGTTCTTGGCTAGGAGCTCATAGAACGCCAGGTCTTGGCGTACTTGTTCAGTTCCCGGCTAGGACAC 15600

L V G P P G V G K T S L G K S I A K S T G R N F V R

15601 CCTCGTCCGGGCCCGCCGGCGTGGCAAGACCTCGTTCGGCAAGTTCGATCGCCAAGTCCACCGGCCGCAATTTTCGTGCGCA 15680
 15601 GGAGCAGCCCGGCGGGCCGCAACCCTTCTGGAGCGAGCCGTTT CAGCTAGCGGTTT CAGGTGGCCGGCGTTAAAGCACGCGT 15680

M S L G G V R D E A E V R G H R R T Y I G S M P G K V

15681 TGTCGCTCGGCGCGGTGCGGGACGAGGCTGAGGTCCGCGGCCACCGACGCACCTACATCGGCTCGATGCCCGGCAAGGTC 15760
 15681 ACAGCGAGCCCGCCACGCCCTGCTCCGACTCCAGGCGCCGGTGGCTGCGTGGATGTAGCCGAGCTACGGGCCGTTCCAG 15760

I Q G M K K A K S S N P L F L L D E I D K L G A D W R

15761 ATCCAGGGCATGAAGAAGGCGAAGTCGTCCAACCCGCTGTTCCCTGCTGGATGAGATCGACAAGCTCGGCGCCGACTGGCG 15840
15761 TAGGTCCCCTACTTCTTCCGCTTCAGCAGGTTGGGCGACAAGGACGACCTACTCTAGCTGTTTCGAGCCGCGGCTGACCCG 15840

G D P S S A L L E V L D P E Q N G T F N D H Y L E V

15841 CGGCGACCCGTCGTCCGGCCCTGCTTGAGGTTCTCGATCCCGAGCAGAACGGCACTTTCAACGACCATTATCTGGAGGTCG 15920
15841 GCCGCTGGGCGAGCAGCCGGGACGAACTCCAAGAGCTAGGGCTCGTCTTGCCGTGAAAGTTGCTGGTAATAGACCTCCAGC 15920

D Y D L S D V M F V C T A N T M R M P Q P L L D R M E

15921 ATTACGACCTGTCGGACGTGATGTTTCGTCTGCACGGCCAACACGATGCGCATGCCGACGCCGCTGCTGGACCGCATGGAG 16000
15921 TAATGCTGGACAGCCTGCACTACAAGCAGACGTGCCGGTTGTGCTACGCGTACGGCGTCGGCGACGACCTGGCGTACCTC 16000

I I R V A G Y T E D E K V E I S K R H L I E K Q V E A

16001 ATCATCCGCGTCGCCGGCTATACCGAGGATGAAAAGGTCGAGATTTCCAAGCGCCACCTGATCGAAAAGCAGGTGGAAGC 16080
16001 TAGTAGGCGCAGCGGCCGATATGGCTCCTACTTTTCCAGCTCTAAAGGTTTCGCGGTGGACTAGCTTTTTCGTCCACCTTCG 16080

N G L K K G E F A I S D D A L R D L I R Y Y T R E A

16081 CAACGGCTTGAAGAAGGGCGAGTTCGCCATTTCCGACGACGCGCTGCGCGACCTGATCCGCTATTACACGCGGGAAGCCG 16160
16081 GTTGCCGAACCTTCTCCCGCTCAAGCGTAAAGGCTGCTGCGCGACGCGCTGGACTAGGCGATAATGTGCGCCCTTCGGC 16160

G V R S L E R E I A N L C R K A V K E I L M K G S A G

16161 GCGTCCGCGAGCCTGGAGCGCGAGATCGCCAACCTCTGCCGCAAAGCCGCTGAAGGAGATCCTGATGAAGGGCTCCGCCGGC 16240
16161 CGCAGGCGTCGGACCTCGCGCTCTAGCGGTTGGAGACGGCGTTCGGCACTTCTCTAGGACTACTTCCCGAGGCGGGCCG 16240

A K V S V T R R N L D K Y A G V R R F H F G E A E L E

16241 GCCAAGGTTTCGGTCACGCGCCGGAACCTGGACAAGTACGCTGGCGTCCGCCGCTTCCACTTCGGCGAGGCGGAGCTGGA 16320
16241 CGGTTCCAAAGCCAGTGCAGCGGCTTGGACCTGTTTCATGCGACCGCAGGCGGCGAAGGTGAAGCCGCTCCGCCTCGACCT 16320

D L V G V T T G L A W T E V G G E L L S I E A V S L

16321 GGATCTGGTGGGCGTCACCACCGGGCTGGCCCTGGACGGAAGTCGGAGGCGAGCTGCTGTGATCGAGGCGGTCAGCCTGC 16400
16321 CCTAGACCACCCGAGTGGTGGCCCCGACCGGACCTGCCTTCAGCCTCCGCTCGACGACAGCTAGCTCCGCCAGTCGGACG 16400

P G K G R V T T T G K L G D V M K E S V Q A A E S Y V

16401 CCGGCAAGGGGCGGGTGACCACCACCGGAAGCTGGGCGACGTCATGAAGGAGTCGGTCCAGGCGGCGGAGAGCTACGTC 16480
16401 GGCCGTTCCCGCCACTGGTGGTGGCCGTTTCGACCCGCTGCAGTACTTCTCAGCCAGGTCGCCCGCTCTCGATGCAG 16480

K S R A T A F G I K P T L F E K R D I H V H V P E G A

16481 AAGTCGCGGGCCACCGCCTTCGGCATCAAGCCGACGCTCTTCGAGAAGCGGGACATCCACGTCCACGTGCCCGAAGGCGC 16560
16481 TTCAGCGCCCCGGTGGCGGAAGCCGTAGTTCGGCTGCGAGAAGCTCTTCGCCCTGTAGGTGCAGGTGCACGGGCTTCCGCG 16560

T P K D G P S A G V A M I T S I V S V L T G I A V R

16561 CACTCCGAAGGACGGCCCCCTCGGCGGGCGTGGCGATGATCACCTCGATCGTTTCGGTCTGACCGGTATCGCGGTCCGCA 16640
16561 GTGAGGCTTCTTCCGCGGGAGCCGCCCGCACCCTACTAGTGGAGCTAGCAAAGCCAGGACTGGCCATAGCGCCAGGCGT 16640

K D V A M T G E I T L R G R V L P I G G L K E K L L A

16641 AGGACGTGGCGATGACCGGTGAGATCACCCCTGCGCGGCCGGGTGCTTCCGATCGGCGGTCTGAAGGAGAAGCTGCTGGCT 16720
16641 TCCTGCACCGCTACTGGCCACTCTAGTGGGACGCGCCGGCCACGAAGGCTAGCCGCCAGACTTCTCTTCGACGACCGA 16720

A L R G G L K H V L I P K D N E K D L A E I P D N V K

16721 GCTCTGCGCGGCGGCTCAAGCATGTGCTGATCCCGAAGGACAACGAGAAGGATCTCGCCGAGATCCCGGACAACGTGAA 16800
16721 CGAGACGCGCCCGGAGTTCGTACACGACTAGGGCTTCTGTTGCTTCTCCTAGAGCGGCTCTAGGGCCTGTTGCACTT 16800

R G L E I I P V S T V D D V L K H A L V R E V E P I

16801 GCGCGGGCTGGAGATCATCCAGTCAGCACCGTCGACGACGCTCCTGAAGCACGCCCTGGTTCGGGAAGTCGAGCCCATCG 16880
16801 CGCGCCCGACCTCTAGTAGGTCAGTCTGTCGAGCTGCTGCAGGACTTCTGTCGGGACCAAGCCCTTCAGCTCGGGTAGC 16880

E W K E P E A V E P A V A K P Q T D G G G E V L R H *

16881 AGTGAAGGAGCCGGAAGCGGTCGAGCCGGCGGTCGCCAAGCCGACGACTGACGGCGGAGGTCGTTTCGTCCTACTGA 16960
16881 TCACCTTCTCGGCTTCGCCAGCTCGGCCCGCAGCGGTTTCGGCGTCTGACTGCCGCTCCGCTGCACGAAGCAGTGACT 16960

16961 TCCAGCTTGGCGTTTAAAGCACCTTAAAGCGCTACCATGCTTCGCATCGCCCCCGGCTCTGCAATTGACGGGCGGGCGA 17040

terminator

| | | |
|-------|--|-------|
| 16961 | AGGTCGAACCGCAAATTTTCGTGGGAATTTTCGCGATGGTACGAAGCGTAGCGGGCGGGCAGACGTTAACTGCCCGCCCGCT | 17040 |
| 17041 | TGTCGCGTTTTATCGTTTTGAATGCGTGATGTCATTGAGTGCCTTTTCTTCCATTTTCTTTGCCGGACTGAGAAGGGGGAA | 17120 |
| 17041 | ACAGCGCAAATAGCAAACCTTACGCACTACAGTAACTCACGGAAAAGAAGTAAAAGAAACGGCCTGACTCTTCCCCCCTT | 17120 |
| | HU | |
| | V N K N D L V A H V A D A V G L S K T D A T K A V | |
| 17121 | CTAAGTGAACAAGAACGATCTGGTGGCGCACGTCGCCGACGCCGTGGGTTTGTCCAAAACCGACGCGACCAAGGCGGTTCG | 17200 |
| 17121 | CATTCACTTGTCTTGTCTAGACCACCGCGTGCAGCGGCTGCGGCACCCAAACAGGTTTTGGCTGCGCTGGTTCCGCCAGC | 17200 |
| | D A I F D G I A D S L K N G E E V R L V G F G T F A V | |
| 17201 | ACGCAATTTTCGACGGCATCGCCGACTCGCTCAAGAACGGGGAGGAGGTCCGTCTGGTCGGATTTCGGCACATTCGCGGTT | 17280 |
| 17201 | TGCGTTAAAAGCTGCCGTAGCGGCTGAGCGAGTTCTTGCCCCCTCTCCAGGCAGACCAGCCTAAGCCGTGTAAGCGCCAA | 17280 |
| | S E R A A S E G R N P R T G E K I A I P A S K Q P K F | |
| 17281 | TCCGAACGCGCGGCCAGCGAGGGACGCAACCCGCGCACCCGGCGAAAAGATCGCCATACCCGCGTCAAACAGCCGAAGTT | 17360 |
| 17281 | AGGCTTGCGCGCCGCTGCTCCCTGCGTTGGGCGCGTGGCCGCTTTTCTAGCGGTATGGGCGCAGCTTTGTCTGGCTTCAA | 17360 |
| | K P G K T L K D G L K * | |
| 17361 | CAAGCCGGGCAAGACCTTGAAGGACGGGTTGAAGTGATGACGCTTGAACGGATAGGGCATTCCGGGCATAGTGCGCGCC | 17440 |
| 17361 | GTTCGGCCCGTTCTGGAACCTTCTGCCAACTTCACTACTGCAGAACTTGCCTATCCCGTAAGGCCCGTATCACGCGCGG | 17440 |
| 17441 | GGTCAGCACGGGCTGGCCGGCTTTGCGTTGTCTGGAACGTCAATCCGGTCAACGACGCGGCGGGCGATTAGCTCAGCG | 17520 |
| 17441 | CCAGTCGTGCCCCACCGGCCAAACGCAACAGACCTTGCAGTTAGGCCAGTTGCTGCGCCGCGCCCGCTAATCGAGTCGC | 17520 |
| 17521 | GTAGAGCATCTCGTTTACACCGAGAGGGTTCGGGGTTTCGAAACCCATCGCCACCATCGCGCCCTGGCCATTCCGGCC | 17600 |
| 17521 | CATCTCGTAGAGCAAATGTGGCTCTCCAGCCCCAAGCTTTGGGAGTAGCGGGTGGTAGCGCGGGACCGGTAAGGCCGG | 17600 |
| 17601 | GGCACTTTTCAAATTCGTGATTGACAGGCCATGCGTTCGGGCGGCTAGAAACCGCCATCGGCGACGGGGCGTGGCCCCGG | 17680 |
| 17601 | CCGTGAAAAGTTAAGCACTAACTGTCCGGTACGCAGCCCGCCGATCTTTGGCGGGTAGCCGCTGCCCCGCACCGGGGCC | 17680 |
| 17681 | AAACTGCGCTGCGATCCTATGACACGGACGCCGTGTTTTTGCGCATTTACATCGCCGTGTGCGCGGGGTGTAGCTCAGTT | 17760 |
| 17681 | TTTGACGCGACGCTAGGATACTGTGCCGTGCGGACAAAACGCGTAAAGTGTAGCGGCACAGCGCGCCACATCGAGTCAA | 17760 |
| 17761 | GGTTAGAGCGCCGGCCTGTACGCGGAGGCCGCGGGTTCAAGTCCCCTCACTCGCGCCATTCTCTCCCTCACCATTCTC | 17840 |
| 17761 | CCAATCTCGCGGCCGACAGTGCAGCCCTCCGGCGCCCAAGTTCAAGGCGAGTGCAGCGGGTAAGAGAGGGAGTGGTAAGAG | 17840 |
| 17841 | CTTTTTGCAATGAAAATGAAGCCTTCATGGCTTTACCGGCAGGCACGCGATCCCCATATCCGATTAGCTCGACCCGCATA | 17920 |
| 17841 | GAAAAACGTTACTTTTACTTCGGAAGTACCGBAAATGGCCGTCCGTGCGCTAGGGGTATAGGCTAATCGAGCTGGGCGTAT | 17920 |
| 17921 | TCCGATAAGCCGGATAGAGAGAAGGGTCATTTAAAAGTTTGGTAATACCAAGGACATCCGCTCATTTCCGCAGAGTGGGA | 18000 |
| 17921 | AGGCTATTCGGCCTATCTCTTCCAGTAAATTTTTCAAACCATATAGGTTCCCTGTAGGCGAGTAAAGGCGTCTCACCT | 18000 |
| | ligação para σ^{70} | |
| 18001 | TCGTTGGCCGAAAGCGATTCAAAGCGCTTTCCTTGATCACGCTCTCGTCATATATTCGGGCCGCTGTGATGGGCTAATC | 18080 |
| 18001 | AAGCAACCGGCTTTTCGCTAAGTTTCGCGAAAGGAAC TAGTGCGAGAGCAGTATATAAGGCCGGGCAGACTACCCGATTAG | 18080 |
| 18081 | CAAGGCATTCAGCCACGACACCGCGCGTTCGGCCGATTACCGGCCGTACCGGCATCGTGGCGTGGGCTTATGGGAGGAC | 18160 |
| 18081 | GTTCCGTAAGGTCGGTGTGTGGCGCGCAGCCGGCTAATGGCCGGCATGGCCGCTAGCACCGCACCCGAATACCCTCCTG | 18160 |
| 18161 | CGCGGTGACCAACCGCTGATCATGAGTACCTCCGATCCTGGTGTTCCTGCTGATCGGTATCGCGCTGGCCGTGGTGA | 18240 |
| 18161 | GCGCCACTGGTTGGGCGACTAGTAACTCATGGAAGGCTAGGACCACAAGGACGACTAGCCATAGCGCGACCGGCACCACT | 18240 |
| | NADHsub3 | |
| | M V G A S Y I I S P K N P D S E K L S P Y E C G F E P | |
| 18241 | TGGTTGGCGCGTCTTACATCATCAGCCGAAGAATCCGGATTTCGGAGAAGCTCTCCCCCTACGAGTGCGGCTTCGAGCCG | 18320 |
| 18241 | ACCAACCGCGCAGAATGTAGTAGTCGGGCTTCTTAGGCCTAAGCCTCTTCGAGAGGGGGATGCTCACGCCGAAGCTCGGC | 18320 |
| | F E D A R T K F D V R F Y L V S I L F I I F D L E V A | |
| 18321 | TTCGAGGATGCCCGCACGAAGTTTCGACGTGCGGTTTTACCTGGTCTCCATCCTGTTTCATCATCTTCGACCTCGAGGTCCG | 18400 |
| 18321 | AAGCTCCTACGGGCGTGTTCGAAGCTGCACGCCAAAATGGACCAGAGGTAGGACAAGTAGTAGAAGCTGGAGCTCCAGCG | 18400 |

F L F P W A V A L G D I G L F G F W S M I V F L G I
 18401 CTTCTGTCCCGTGGGCGTGGCGCTCGGGGACATCGGCCTCTTCGGCTTCTGGTTCGATGATCGTGTTCCTTGGGATTC 18480
 18401 GAAGGACAAGGGCACCCGGCACCCGAGCCCTGTAGCCGGAGAAGCCGAAGACCAGCTACTAGCACAAGGAACCCTAAG 18480

NADH nq06
 L T I G F I Y E W K K G A L E W E *
 M G V E A A K L A P I P
 18481 TGACCATCGGCTTCATCTATGAGTGGAAAGAAAGGAGCTCTGGAATGGGAGTAGAGGCTGCCAAGCTGGCGCCGATTCCGC 18560
 18481 ACTGGTAGCCGAAGTAGATACTCACCTTCTTTCCCTCGAGACCTTACCCTCATCTCCGACGGTTCGACCGCGGCTAAGGCG 18560

P G P E Q D A Y L R A V T E E I Q E K G F I T A K Y E
 18561 CCGGACCGGAACAGGACGCCTATCTCCGCGCGGTTACCGAGGAGATCCAGGAAAAGGGCTTCATCACGGCGAAGTACGAG 18640
 18561 GGCTGGCCTTGTCTGCGGATAGAGGCGGCCAATGGCTCCTCTAGTCTTTTCCCGAAGTAGTGCCGCTTCATGCTC 18640

D V L A W A R T G S L W P M T F G L A C C A V E M I H
 18641 GACGTGCTCGCCTGGGCCCGCACCGGTTCCCTGTGGCCGATGACCTTCGGTCTGGCCTGCTGCGCGGTGGAGATGATCCA 18720
 18641 CTGCACGAGCGGACCCGGGCGTGGCCAAGGACACCGGCTACTGGAAGCCAGACCGGACGACGCGCCACCTCTACTAGGT 18720

A Y M S R Y D L D R F G V I P R P S P R Q S D C M I
 18721 CGCCTACATGAGCCGGTACGACCTGGACCGTTTCGGCGTTCATTCCGCGCCCCAGCCCGCGCCAGTCCGACTGCATGATCG 18800
 18721 GCGGATGTACTCGGCCATGCTGGACCTGGCAAAGCCGAGTAAGGCGCGGGTTCGGGCGCGGTTCAGGCTGACGTACTAGC 18800

V A G T L T N K M A P A L R K V Y D Q M P E P R W V I
 18801 TGGCCGGCACGCTGACCAACAAGATGGCCCCGCGTGC GCAAGGTCTATGACCAGATGCCGGAACCGCGCTGGGTGATC 18880
 18801 ACCGGCCGTGCGACTGGTTGTTCTACCGGGGGCGGACGCGTTCAGATACTGGTCTACGGCCTTGGCGCGACCCACTAG 18880

S M G S C A N G G G Y Y H Y S Y S V V R G C D R I V P
 18881 TCGATGGGCTCCTGCGCCAACGGCGCGGCTACTATCACTATTCCCTACTCGGTGGTGC GCGGCTGCGACCGGATCGTTCC 18960
 18881 AGCTACCCGAGGACGCGGTTGCCGCCGCGATGATAGTGATAAGGATGAGCCACCACGCGCCGACGCTGGCCTAGCAAGG 18960

V D I Y V P G C P P T A E A L V Y G I L Q L Q K K I
 18961 GGTCGATATTTACGTGCCGGGCTGTCTCCGACCGCGGAAGCGCTGGTCTACGGCATTCTGCAACTCCAGAAGAAGATCC 19040
 18961 CCAGCTATAAATGCACGGCCCACAGGAGGCTGGCGCCTTCGCGACCAGATGCCGTAAGACGTTGAGGTCTTCTTCTAGG 19040

R R G N R I A R *
 19041 GGCGCGCAACCGCATCGCTCGCTGAATAAGAAAAAGGCAGGTCTGGCCATGTCCGAACAGGCGTTGAAGGAACTTGGG 19120
 19041 CCGCGCCGTTGGCGTAGCGAGCGACTTATTCTTTTTTCCGTCCAGGACCGGTACAGGCTTGTCCGCAACTTCTTGAACCC 19120

19121 GACCATATCGCCGCGACGCTCGGCGATGACGTCCTGAAGGTCGAGCTGAAGCTT 19173
 19121 CTGGTATAGCGGCGTGCAGCCGCTACTGCAGGACTTCCAGCTCGACTTCGAA 19173

Fig. 4 Seqüência de nucleotídeos dos genes presentes no inserto do plasmídeo pAB441. As seqüências polipeptídicas deduzidas para cada gene são mostradas acima da região codificadora no caso de proteínas codificadas na direção da seqüência apresentada e as seqüências de aminoácidos das proteínas GlnB (*glnB*) e GS (*glnA*) estão mostradas abaixo da seqüência de nucleotídeos. Códon de parada estão indicados por asteriscos. Os prováveis promotores encontrados estão indicados por letras sublinhadas e os terminadores indicados por flechas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ADLER, S. P.; PURICH, D.; STADTMAN, E. R. Cascade control of *Escherichia coli* glutamine synthetase. Properties of the PII regulatory protein and the uridylyltransferase-uridylyl-removing enzyme **J. Bacteriol.** v. 250 (16), p. 6264-6272, 1975.
- ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D.J. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." **Nucleic Acids Res.** 25:3389-3402, 1997.
- ANGUS-HILL, M. L.; DUTNALL, R. N.; TAFROV, S. T.; STERNGLANZ, R.; RAMAKRISHNAN, V. Crystal structure of the histone acetyltransferase Hpa2: A tetrameric member of the Gcn5-related N-acetyltransferase superfamily. **J Mol Biol.** v. 294, p. 1311-25, 1999.
- ARCONDÉGUY, T.; JACK, R.; MERRICK, M. PII Signal transduction proteins: pivotal players in microbial nitrogen control. **Microbiol Mol. Biol. Rev.** v. 65, n.1, p. 80-105, 2001
- ARSENE, F.; KAMINSKI, P. A.; ELMERICH, C. Modulation of NifA activity by PII in *Azospirillum brasilense*: evidence for a regulatory role of the NifA N-terminal domain. **J Bacteriol.**, v. 178, p.4830-4838, 1996.
- BALDANI, V. L. K.; BALDANI, J. I.; OLIVARES, F. L. DÖBEREINER, J. Identification and ecology of *Herbaspirillum seropedicae* and closely related *Pseudomonas rubrisubalbicans*. **Symbiosis** v. 13, p. 65-73, 1992.
- BANKIER, A.T.; WESTON, K.M.; BARREL, B.G. Random cloning and sequencing by the M13 dideoxynucleotide chain termination method. **Methods Enzymol.**, **155**, 51-93, 1987.

- BASHAN, Y.; HOLGUIN, G. *Azospirillum* – plant relationships: environmental and physiological advances (1990-1996). **Can. J. Microbiol.** V. 43, p. 103-121, 1997.
- BENDER, R. A.; MACALUSO, A.; MAGASANIK, B. Glutamate dehydrogenase: genetic mapping and isolation of regulatory mutants of *Klebsiella aerogenes*. **J. Bacteriol.**, 127:141-8, 1976.
- BONNEFOY, E., TAKAHASHI, M. and ROUVIÈRE YANIV, J. DNA binding parameters of the HU protein of *Escherichia coli* to cruciform DNA. **J. Mol. Biol.** 242: 116-129, 1994.
- BOZOUKLIAN, H. and ELMERICH, C. Nucleotide sequence of the *Azospirillum brasilense* Sp7 glutamine synthetase structural gene. **Biochimie** 68, 1181-1187, 1986.
- BURRIS, R. H. Nitrogenases. **J. Biol Chem.** V.266, p. 9339-9342, 1991
- CLARET, L.; ROUVIÈRE-YANIV, J. Regulation of HU beta by CRP and FIS in *Escherichia coli*. **J Mol Biol.** v. 25, n. 263(2) p.126-39, 1996.
- CIRILLI, M.; ZHENG, R.; SCAPIN, G.; BLANCHARD, J. S. Structural Symmetry: The three-dimensional structure of *Haemophilus influenzae* diaminopimelate epimerase. **Biochemistry**, v. 37, p. 16452-16458, 1998.
- DE MEL, V. S. J.; KAMBEROV, E. S.; MARTIN, P. D.; ZHANG, J.; NINFA, A. J.; EDWARDS, B. F. P. (1994) Preliminary X-ray diffraction analysis of crystals of the PII protein from *Escherichia coli*. **J. Mol. Biol.**, v.243, p. 796-798, 1994.
- DEAN, D. R.; BOLIN, J. T.; ZHENG, L. Nitrogenase metalloclusters: structures, organization and synthesis. **J. Bacteriol**, Washington, v. 175, p. 6737-6744, 1993

- DEKHIL, S.B.; CAHILL, M.; STACKEBRANDT, E.; SLY, L. I. Transfer of *Comglomeromonas largomobilis* subsp. *largomobilis* to the genus *Azospirillum* as *Azospirillum largomobile* comb. nov., elevation of *Comglomeromonas largomobilis* subsp. *Parooensis* to the new type species of *Comglomeromonas*, *Comglomeromonas parooensis* sp. nov. *System. Appl. Microbiol.* v. 20, p. 72-77, 1997.
- DILWORTH, M. J. Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum*. *Bioch. Biophys. Acta*, v. 127, p. 285-294, 1966.
- DIXON, R. The oxygen-responsive NifL-NifA complex: a novel two-component regulatory system controlling nitrogenase synthesis in gamma-proteobacteria. *Arch. Microbiol.* v. 169, p. 371-380, 1998
- DÖBEREINER, J. The genera *Azospirillum* and *Herbaspirillum*. In BALLOWS, A., TRUPPER, H.G., DWORKING, M., HARDER, W. **The Prokaryotes**. 2.ed. vol III., p.2236-3353, Springer-Verlag, 1991.
- DÖBEREINER, J.; DAY, J. M. Associative symbioses in tropical grasses characterization of microorganisms and dinitrogen-fixing sites. In NEWTON, W. E. & NYMAN, C. J. eds. **Proc. of the 1st. Intl. Symp. Nitrogen Fixation**. v.2. Pullman, Washington Univ. Press. p.518-38, 1996.
- DÖBEREINER, J.; PEDROSA, F. O. Nitrogen-fixing bacteria in nonleguminous crop plants. **Science Tech Publishers**, Springer-Verlag, 1987.
- DODGE, R. W.; SCHERAGA, H. A. Folding and unfolding kinetics of the praline-to-alanine mutants of bovine pancreatic ribonuclease A. **Biochemistry**, v.35, p. 1548-1559, 1996.

DONG, Z.; CANNY, M. J.; McCULLY, M. E.; ROBOREDO, M. R.; CABADILLA, C. F.; ORTEGA, E.; RODÉS, R. A nitrogen-fixing endophyte of sugarcane stems. A new role for the apoplast. **Plant Physiol.** v. 105, 1139-1147, 1994.

DOWER, W. J.; MILLER, J. F.; RAGDALE, C.W. High efficiency transformation of *E. coli* by high voltage electroporation. **Nucl. Acids Res.** V. 16, p. 6127-6145, 1988.

DRI, A. M.; MOREAU, P. L.; ROUVIERE-YANIV, J. Role of the histone-like proteins OsmZ and HU in homologous recombination. **Gene.** v. 120(1) p.11-6, 1992.

EADY, R. R. Enzimology in free-living diazotrophs. In: BROUGHTON, W. J.; PUHLER, S. (Ed) **Nitrogen Fixation**, v. 4 p. 1-49, 1986

ECKERT, B.; WEBER, O. B.; KIRCHOF, G.; HALBRITTER, A.; STOFFELS, M.; HARTMANN, A. *Azospirillum doebereinae* sp. nov., a nitrogen-fixing bacterium associated with the C4-grass *Miscanthus*. **Int. J. Syst. Evol. Microbiol.** v. 51, p. 17-26, 2001.

EISENBERG, D.; GILL, H. S.; PFLUEGL, G. M. U.; ROTSTEIN, S. H. Structure-function relationships of glutamine synthetases. **Bioch. et Bioph. Acta** v. 1477, p. 122-145, 2000.

Embrapa: www.embrapa.br

ETCHEGARAY, J. P.; JONES, P. G. and INOUE, M. Differential thermoregulation of two highly homologous cold-shock genes, *cspA* and *cspB*, of *Escherichia coli*. **Genes Cells** 1: 171-178, 1996.

EYDMANN, T.; SODERBACK, E.; JONES, T.; HILL, S.; AUSTIN, S.; DIXON, R. Transcriptional activation of the nitrogenase promoter in vitro: adenosine nucleotides

are required for inhibition of NifA activity by NifL. **J. Bacteriol.** v.177, p.1186-1195, 1995.

EWING, B.; HILLIER, L.; WENDI M. C.; GREEN, P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. **Genome Res.**, v.8, p. 175-185,1998.

EWING, B. & GREEN, P. Basecalling of automated sequencer using phred. II Error probabilities. **Genome Res.**, v.8, p.186-194, 1998.

FADEL-PICHET, C. M. T., SOUZA, E. M., RIGO, L. U., FUNAYAMA, S., YATES, M. G., and PEDROSA, F. O. Regulation of *Azospirillum brasilense nifA* gene expression by ammonium and oxygen. **FEMS Microbiol. Lett.** 179, 281-288, 1999.

FIGURSKI, D. H.; HELINSKI, D. R. Replication of na origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. **Proc. Natl. Acad. Sci. USA.** v. 76, p. 1648-1652, 1979.

FINEL, M. Organization and evolution of structural elements within complex I. **Biochim Biophys Acta**, May 6, 1364(2): 112-21, 1998.

FISCHER, H-M. Genetic Regulation of Nitrogen Fixation in Rhizobia. **Microbiol. Rev.**, v. 58, p. 352-386, 1994.

FRIEDRICH, T., STEINMÜLLER K.; WEISS H. The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. **FEBS Lett**, June 26; 367(2): 107-11, 1995.

FRIEDRICH, T. The NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*. **Biochim Biophys Acta**, May 6; 1364(2): 134-46, 1998.

- FRYDMAN, J.; HARTL, F. U. Principles of chaperone-assisted protein folding: differences between in vitro and in vivo mechanisms. *Science*, v. 7;272(5267), p.1497-502, 1996.
- FRYDMAN, J. Folding Of Newly Translated Proteins In Vivo: The Role of Molecular Chaperones. **Annu. Rev. Biochem.** 70: 603-647, 2001.
- GAUTHIER, D.; ELMERICH, C. Relationship between glutamine synthetase and nitrogenase in *Spirillum lipoferum*. **FEMS Microbiol. Lett.** v. 2; p. 101-104, 1977.
- GIANGROSSI M.; GIULIODORI A. M.; GUALERZI, C. O. and PON C. L. 2002; Selective expression of the β -subunit of nucleoid-associated protein HU during cold shock in *Escherichia coli*. **Mol. Microbiol.**, 44 (1), 205-216, 2002.
- GILLIS, M.; DERSTERS, K.; HOSTE, B.; JANSSENS, E.; KROPPESTEDT, R. M.; SEPHAN, M. P., TEIXEIRA, K. R. S., DÖBEREINER, J. *Acetobacter diazotrophicus* sp. nov., a nitrogen-fixing acetic acid bacterium associated with sugarcane. **Int. J. Syst. Bacteriol.** v. 39 p. 361-364, 1989.
- GILLIS, M.; VAN, T. V.; BARDIN, R.; GOOR, M.; HEBBAR, P.; WILLEMS, A.; SEGERS, P.; DERSTERS, K. Polyphasic taxonomy in genus *Burkholderia* leading to an emended description of the genus and proposition of *Burkholderia vietnamiensis* sp. Nov. for N₂-fixing isolates from rice in Vietnam. **Int. J. Syst. Bacteriol.** v. 45, p. 274-279, (1995)
- GORDON, D.; ABAJIAN, C.; GREENP. Consed: A graphical tool for sequence finishing. **Genome Res.**, v8, p. 195-202,1998.
- GOTTESMAN, S.; MAURIZI M. R.; Regulation by proteolysis: energy-dependent proteases and their targets. **Microbiol Rev.** Dec;56(4):592-621, 1992.

- HAAKER, H.; KLUGKIST, J. The bioenergetics of electron transport to nitrogenase. **FEMS Microbiol. Rev.** Amsterdam, v. 46, p. 57-71, 1987.
- HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucl. Acids. Symp. Ser.** 41:95-98, 1999.
- HANAHAN, D. Studies on transformation of *Escherichia coli* with plasmids. **J. Mol. Biol.** V.166, p. 557-580, 1983.
- HARTMANN, A.; FU, H.; BURRIS, R. H. Regulation of nitrogenase activity by ammonium chloride in *Azospirillum* spp. **J. Bacteriol.** v. 165, n.3, p 864-870, 1986.
- HAYKINSON, M. J., AND R. C. JOHNSON. DNA looping and the helical repeat *in vitro* and *in vivo*: effect of HU protein and enhancer location on Hin invertasome assembly. **EMBO J.** 12:2503-251, 1993.
- HAWKES, T.; MERRICK, M.; DIXON, R. Interaction of purified NtrC protein with nitrogen regulated promoters from *Klebsiella pneumoniae*. **Mol. Gen. Genet.**, v. 201, p. 492-498, 1985.
- HENDRICK, J. P.; LANGER, T.; DAVIS, T. A.; HARTL, F. U.; WIEDMANN, N.; Control of folding and membrane translocation by binding of the chaperone DnaJ to nascent polypeptides. **Proc Natl Acad Sci U S A.** v.90 (21), p. 10216-20, 1993.
- HESTERKAMP, T.; BUKAU, B.; Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E.coli*. **EMBO J.**, v. 17 (16), p. 4818-28, 1998.
- HILL, S.; AUSTIN, S., JONES, T.; DIXON, R. *Azotobacter vinelandii* NIFL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch. **Proc Natl Acad Sci U S A.** v. 93, p. 2143-8, 1996.

- HUERGO, L. F.; **Caracterização da Região Promotora do Operon gleba de *Azospirillum brasilense***. Dissertação (Mestrado em Ciências - Bioquímica). Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2002.
- HUERGO, L. F.; SOUZA, E. M.; STEFFENS, M. B. R.; YATES, G.; PEDROSA, F. O.; CHUBATSU, L. S.; Regulation of *glnB* gene promoter expression in *Azospirillum brasilense* by the NtrC protein. **FEMS Microbiol. Lett.** 223, p. 33-40, 2003.
- HUNGRIA, M.; CAMPO, R.J.; MENDES, I.C. Fixação biológica do nitrogênio na cultura da soja. **Embrapa Soja Circular Técnica**, 35 p. 48, 2001
- INDORATO, C.; GIANNELLI, L.; BAZZICALUPO, M. *Azospirillum brasilense* *tig-clp-clpx-lon*, Código de acesso do GenBank: AF150957
- INOYUYE, M. Experimental manipulation of gene expression, p 28-30. **Academic Press**, Inc., Orlando, Fla. 1983.
- ISHIDA, M. L.; ASSUMPCÃO, M. C.; MACHADO, H. B.; BENELLI, E. M.; SOUZA, E. M.; PEDROSA, F. O. Identification and characterization of the two-component NtrY/NtrX regulatory system in *Azospirillum brasilense*. **Braz. J. Med. Biol. Res.** v. 35(6), p. 651-661, 2002.
- 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 Microbiol. Lett.** v. 15, p. 251-253.
- JIANG, P., NINFA, A. J., Regulation of autophosphorylation of *Escherichia coli* nitrogen regulator II by the PII signal transduction protein. **J Bacteriol.** 1999 Mar;181(6):1906-11, 1999.
- KADO, C. I.; LIU, S. T. Rapid procedure for detection and isolation of large and small. **J. Bacteriol.** v. 145, n. 3, p. 1365-1373, 1981.

KAMMAS, K. M.; AGERON, E.; GRIMONT, P. A. D.; KAISER, P. *Azospirillum irakense* sp. nov., a nitrogen-fixing bacterium associated with rice roots and rhizosphere soil. **Res. Microbiol.** v. 140, p. 679-694, 1989.

KAMBEROV, E. S.; ATKINSON, M.R.; NINFA, A. J. The *Escherichia coli* PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP. **J Biol Chem**, Baltimore, v. 270, p. 17797-17807, 1995.

KANDROR, O. and GOLDBERG A. L. Trigger factor is induced upon cold shock and enhances viability of *Escherichia coli* at low temperatures **Proc. Natl. Acad. Sci. USA** Vol. 94, pp. 4978-4981, May 1997

KANO, Y., YOSHINO. S., WADA. M., YOKOYANA, K., NOBUHARA, M. and IMAMOTO, F. Molecular cloning and nucleotide sequence of HU-1 gene of *Escherichia coli*. **Mol. Gen. Genet** 201: 360-362, 1985.

KANO Y., WADA, M., IMAMOTO F. 1988 Genetic characterization of the gene *hupA* encoding the HU-2 protein of *Escherichia coli*. **Gene**. Sep 30;69(2):331-5, 1988.

KANO, Y., IMAMOTO, F., Requirement of integration host factor (IHF) for growth of *Escherichia coli* deficient in HU protein. **Gene**. Apr 30;89(1):133-7, 1990.

KANO, Y., OGAWA, T.; OGURA, T.; HIRAGA, S.; OKAZAKI, T.; IMAMOTO, Y.
Participation of the histone-like protein HU and of IHF in minichromosomal maintenance in *Escherichia coli*. **Gene**. Jul 15;103(1):25-30, 1991.

KAPULNIK Y.; OKON, Y.; HENIS. Y. Changes in root morphology of wheat caused by *Azospirillum brasilense* inoculation. **Can J. Microbiol.** V. 31, p. 881-887, 1985

KEENER, J.; KUSTU, S. Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the

conserved amino-terminal domain of NTRC. **Proc Natl Acad Sci USA**, v. 85, p. 4976-4980, 1988.

KLASSEN, G.; SOUZA, E. M.; YATES, M. G.; RIGO, L. U.; INABA, J. PEDROSA, F. O. Control of nitrogenase reactivation by the GlnZ protein in *Azospirillum brasilense*. **J. Bacteriol.** v. 183, n. 22, p. 6710-6713, 2001.

KLASSEN, G. et al. -submetido 2005.

KOCHER, T. D.; WILSON, A. C. DNA amplification by the polymerase chain reaction. In BROWN, T. A., **Essential Molecular Biology: A Practical Approach**, IRL Press, Oxford University Press, Oxford, p. 185-208, 1991.

KUTSU, S., HIRSCHMAN, J. BURTON, JELESKO, J. and MEEKS, J. C., Covalent modification of bacterial glutamine synthetase: physiological significance. **Mol. Gen. Genet.** 197: 309-317. 1984.

LAVOIE, B. D. and CHACONAS, G., Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. **Genes Dev.** Dec; 7(12B): 2510-9, 1993.

LETUNIC, I.; COPLEY, R. R.; SCHMIDT, S.; CICCARELLI, F. D.; DOERKS, T.; SCHULTZ, J.; PONTING, C. P.; BORK, P. SMART 4.0: towards genomic data integration. **Nucleic Acids Res.** v. 1;32(Database issue):D142-4. 2004

LIANG Y. Y.; KAMINSKI, P. A.; ELMERICH, C. Identification of a *nifA*- like regulatory gene of *Azospirillum brasilense* Sp7 expressed under conditions of nitrogen fixation in the presence of air and ammonia. **Mol. Microbiol.**, Oxford, v.5 , p. 2735-2744, 1991

- LIANG, Y. Y.; de ZAMAROCZY, M.; ARSENE, F.; PAQUELIN, A.; ELMERICH, C. Regulation of nitrogen in *Azospirillum brasilense* Sp7: involvement of *nifA*, *glnA* and *glnB* gene products. **FEMS Microbiol. Lett.**, v. 79, p. 113-19, 1992.
- LIANG, Y. Y.; ARSENE, F.; ELMERICH, C. Characterization of the *ntrBC* genes of *Azospirillum brasilense* Sp7: their involvement in the regulation of nitrogenase synthesis and activity. **Mol. Gen. Genet.** v. 240, p. 188-196, 1993.
- LIAW, S. H.; PAN, C.; EISENBERG, D. Feedback inhibition of fully unadenylylated glutamine synthetase from *Salmonella typhimurium* by glycine, alanine and serine. **Proc. Natl. Acad. Sci. USA**, v. 90, p. 4996-5000, 1993.
- LIAW, S. H.; EISENBERG, D. Structural Model for the Reaction Mechanism of Glutamine Synthetase, Based on Five Crystal Structures of Enzyme-Substrate Complexes. **Biochemistry**, v. 33, p. 675-681, 1994.
- LIAW, S. H.; KUO, I.; EISENBERG, D. Discovery of the ammonium substrate site on glutamine synthetase, a third cation binding site. **Protein Sci.** v. 4, p. 2358-2365, 1995.
- LIAW, S. H.; VILLAFRANCA, J. J.; EISENBERG, D. A model for oxidative modification of glutamine synthetase, based on crystal structures of mutant H269N and the oxidized enzyme. **Biochem.** v.32(31), p.7999-8003, 1993.
- LOWRY, O. W.; ROSENBROUGH, N. J.; FARR, A. L. L.; RANDALL, R. J. Protein measurement with the folin phenol reagent. **J. Biol. Chem.** v. 193, p. 267-275, 1951.
- MACHADO, H. B. **Isolamento e caracterização de mutantes de *Azospirillum brasilense* constitutivos para fixação de nitrogênio.** Dissertação de mestrado (Ciências – Bioquímica) Setor de Ciências Biológicas, UFPR, Curitiba, 1988.

- MACHADO, H. B.; FUNAYAMA, S.; RIGO, L. U. and PEDROSA, F.O. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. **Can. J. Microbiol.** 29, 549-553, 1991.
- MACHADO, H. B.; YATES, M. G.; FUNAYAMA, S.; RIGO, L. U. STEFFENS, M. B. R.; SOUZA, E. M.; PEDROSA, F. O. The *ntrBC* genes of *Azospirillum brasilense* are part of a *nifR3*-like-*ntrB-ntrC* operon and are negatively regulated. **Can. J. Microbiol.** v. 41, p. 674-684, 1995.
- MACHADO, H. B. **Organização estrutural do operon ORF1-ntrB-ntrC em *Azospirillum brasilense*.** Tese (Doutorado Ciências- Bioquímica) Setor de Ciências Biológicas, Universidade Federal do Paraná. Curitiba, 1995.
- MAGALHÃES, F. M. M.; BALDANI, J. I.; SOUTO, S. M.; KUYKENDALL, J. R.; DOBEREINER, J. A new acid-tolerant *Azospirillum* species. **An. Acad. Bras. Ciên.** v. 55, p. 471-480, 1983.
- MAGASANIK, B. The regulation of nitrogen utilization in enteric bacteria. **J. Cell. Biochem.**, v. 51, p.34-40, 1993
- MALMSTRÖM, B. G. The mechanism of proton translocation in respiration and photosynthesis. **FEBS Lett.** 250, 9-21, 1989.
- MANNA, D.; J. GOWRISHANKAR. Evidence for involvement of protein HU and RpoS in transcription of the osmoreponsive *proU* operon in *Escherichia coli*. **J. Bacteriol.** 176:5378-5384, 1994.
- MARCHLER-BAUER A, ANDERSON JB, CHERUKURI PF, DEWEESE-SCOTT C, GEER LY, GWADZ M, HE S, HURWITZ DI, JACKSON JD, KE Z, LANCZYCKI CJ, LIEBERT CA, LIU C, LU F, MARCHLER GH, MULLOKANDOV M, SHOEMAKER B.A., SIMONYAN V, SONG JS, THIESSEN PA, YAMASHITA

RA, YIN JJ, ZHANG D, BRYANT SH. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res.* v. 33 Database Issue:D192-6. 2005

MEAD, D. A.; SZCZENA-SKORUPA, E.; KEMPER, B. Single strand DNA “blue”T7 promoter plasmids a versatile tandem promoter system for cloning and protein engineering. **Prot. Eng.**, v.1, p. 67-74, 1986.

MERRICK, M. J. Nitrogen control of the *nif* regulon in *Klebsiella pneumoniae*: involvement of the *ntrA* gene and analogies between *ntrC* and *nifA*. **EMBO J.**, v. 2, p. 39-44, 1983.

MERRICK, M. J. Organization and regulation of nitrogen fixation genes. In: Palacios, R.; Mara, J.; Newton, W. E. (eds) **New Horizons in Nitrogen Fixation**. Nijhoff Publishers, p. 1-12, 1992.

MERRICK, M. J.; EDWARDS, R. A. Nitrogen control in bacteria. **Microbiol. Rev.**, v. 59, p. 604-622, 1995

MILLER, J. H. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and Related Bacteria. Cold Spring Harbour Laboratory Press, New York, 1992.

Ministério da Agricultura, Pecuária e Abastecimento: www.agricultura.gov.br

MORI E.; FULCHIERI, M.; INDORATO, C.; FANI, R.; BAZZICALUPO, M. Cloning, nucleotide sequencing, and expression of the *Azospirillum brasilense lon* gene: involvement in iron uptake. **J Bacteriol.** Jun;178(12):3440-6, 1996.

NEIDHARDT, F. C., and VANBOGELEN, R. A. Heat shock response, p. 1334-1345. In: F. C. Neihardt, J.L. Ingraham. K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.) *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. **American Society for Microbiology**, Washington, D. C., 1987.

- OKAMURA, H.; MUROOKA, Y.; HARADA, Y. Tyramine oxidase and regulation of arylsulfatase synthesis in *Klebsiella aerogenes*. **J. Bacteriol.** v. 129 p. 59-65, 1977.
- PARSELL, D. A.; LINDQUIST, S. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. **Annu. Rev. Genet.** 27: 437-496, 1993.
- PAWLOSWSKI, K.; KLOSSE, U.; DE BRUIJN, F. J. Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component system: NtrY/NtrX, involved in nitrogen fixation and metabolism. **Mol. Gen. Genet.** v. 231, p. 124-138, 1991.
- PEDROSA, F. O. Fixação biológica de nitrogênio: fértil idéia. **Ciência Hoje**, v. 6, p. 12-13, 1987.
- PEDROSA, F. O.; YATES, M. G. Regulation of nitrogen fixation (*nif*) genes of *Azospirillum brasilense* by *nifA* and *ntrC* (*glnG*) type genes. **FEMS Microbiol. Lett.** v. 55, p. 95-101, 1984.
- PEDROSA, F. O. Physiology, biochemistry and genetics of *Azospirillum* and other root-associated nitrogen-fixing bacteria CRC **Critical Reviews in Plant Sciences**. Boca Raton, v. 6, p. 345-383, 1988
- POSTGATE, J. R. The fundamentals of nitrogen fixation, Cambridge, **Cambridge Univ. Press.**, p. 252, 1982
- REINHOLD, B.; HUREK, T.; FENDRIK, I.; POT, B.; GILLIS, M.; KERSTERS, K.; THIELEMANS, S. and de LEY J. *Azospirillum halopraeferes* sp. nov., a Nitrogen-Fixing Organism Associated with Roots of Kallar Grass (*Leptochloa fusca* (L.) Kunth) **Int. J. Syst. Bacteriol.** Jan. 37(1):43-51, 1987.
- REINHOLD-HUREK, B.; HUREK, T.; GILLIS, M.; HOSTE, B.; VANCANNEYT, M.; KERSTERS, K.; DE LEY, J. *Azoarcus* gen. nov., nitrogen fixing proteobacteria

associated with root of Kallar grass [*Leptochloa fusca* (L) Knuth], and description of two species, *Azoarcus indigenus* sp. nov. **Int. J. Syst. Bacteriol.** v. 43, p. 574-584, 1993.

RENNIE, R. J. de FREITAS, J. R.; RUSCHEL, P.; VOSE, P. B. Isolation and identification of N₂-fixing bacteria associated with sugar cane (*Saccharum sp.*) **Can. J. Microbiol.** v. 28, p. 462-467, 1982.

ROUVIERE-YANIV, J., YANIV, M., and GERMONK, J. E., *Escherichia coli* DNA-binding protein HU forms nucleosome-like structure with circular double-stranded DNA. **Cell** 17: 265-274, 1979.

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.

SANGER, F.; NICKLEN, S.; COULSON, A.R. Sequencing with chain-terminating inhibitors. **Proc Natl Acad Sci USA**, v. 74, p. 5463-5467, 1977.

SCHOLLHORN, R.; BURRIS, R. H. Acetylene as a competitive inhibitor of N₂ fixation. **Proc. Nat. Acad. Sci. USA**, v. 58, p. 213-216, 1967.

SCHOLZ, C., G. STOLLER, T. ZARNT, G. FISCHER, AND F. X. SCHMID. Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. **EMBO J.** 16:54-58, 1997.

SHARPLES, G. J.; CHAN, S. N.; MAHDI, A. A.; WHITHY, M. C.; LOYD R. G.; Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. **EMBO J.** Dec 15;13(24):6133-42, 1994.

- SIMON, R.; PRIEFER, U.; PUHLER, 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.
- SMIRNOV, V. V.; KIPRIANOVA, E. A. Bacteria of Pseudomonas genus, Naukova Dumka, Kiev, Ukraine, p. 100-111, 1990.
- STEENHOUDT, O. and VANDERLEYDEN, J. *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. **FEMS Microbiol Rev.** Oct;24(4):487-506, 2000.
- SUGINO, H.; SASAKI, M.; AZAKAMI, H.; YAMASHITA, M.; MUROOKA, Y. A monoamine-regulated *Klebsiella aerogenes* operon containing the monoamine oxidase structural gene (*maoA*) and the *maoC* gene. **J. Bacteriol.** v. 174 (8), p. 2485-2492, 1992.
- TARRAND, J.J.; KRIEG, N. R.; DOBEREINER, J. Ataxonomic study of the Spirillum lipoferum group with description of a new genus, Azospirillum gen. nov. and two species, Azospirillum lipoferum (Beijerinck) comb. nov. and Azospirillum brasilense sp. nov. **Can. J. Microbiol.** v. 24, p. 967-980, 1978.
- WEIDNER, U., GEIER, S., PTOCK, A., FRIEDRICH, T., LEIF, H. AND WEISS, H., The gene locus of the proton-translocating NADH: ubiquinone oxidoreductase in Escherichia coli. Organization of the 14 genes and relationship between the derived proteins and subunits of mitochondrial complex I. **J Mol Biol**, September 5, 233(1): 109-22, 1993.
- WESTBY, C. A.; ENDERLIN, C. S.; STEINBERG, N. A.; JOSEPH, C. M.; MEEKS, J. C. Assimilation of $^{13}\text{NH}_4^+$ by *Azospirillum brasilense* grown under nitrogen limitation and excess. **J. Bacteriol.** v. 169, n. 9, p. 4211-4214, 1987.
- WU, Y.; DATTA, P.; Influence of DNA topology on expression of the *tdc* operon in Escherichia coli K-12. **Mol Gen Genet.** Jun 25;247(6):764-7,1995.

- YAGI, T.; YANO, T.; Di BERNARDO, S.; MATSUNO-YAGI, A.; Procaryotic complex I (NDH-1), an overview. **Biochim Biophys Acta**. May 6;1364(2):125-33, 1998.
- YAMASHITA, M.; SAKAUE, N.; IWATA, H.; SUGINO, H.; MUROOKA, Y. Purification and characterization of monoamine oxidase from *Klebsiella aerogenes*. **J. Ferment. Bioeng.** v. 76, p. 289-295, 1993.
- YAMASHITA, M.; AZAKAMI, H.; YOKORO, N.; ROH, J.; SUZUKI, H.; KUMAGAI, H.; MUROOKA, Y. *maoB*, a gene that encodes a positive regulator of the monoamine oxidase gene (*maoA*) in *Escherichia coli*. **J. Bacteriol.** v. 178, p. 2941-2947, 1996.
- YANISCH-PERRON, D., VIEIRA, J.; MESSING, J.; Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, **Gene** 33, 103-119, 1985.
- VANDE BROEK, A., DOBBELAERE, S., VANDERLEYDEN, J., VANDOMELLEN, A., Azospirillum-plant root interactions: signaling and metabolic interactions. In: Prokaryotic Nitrogen Fixation: A Model System for Analysis of a Biological Process. **Horizon Scientific Press**, pg761-777, 2000.
- VAN DOMMELLEN, A.; KEIJERS, V.; WOLLEBRANTS, A.; VANDERLEYDEN, J. Phenotypic Changes Resulting from Distinct Point Mutations in the *Azospirillum brasilense glnA* Gene, Encoding Glutamine Synthetase. **Appl. Environ. Microbiol.** v. 69, n.9, p.5699-5701.
- VETTING, M. W.; MAGNET, S.; NIEVES, E.; RODERICK, S. L.; BLANCHARD, J. S. A bacterial acetyltransferase capable of regioselective N-acetylation of Antibiotics and Histones. **Chemistry & Biology** v. 11, p. 565-573, 2004.
- VITORINO, J. **Análise genética e funcional de genes *nif*, *gln*, e *ntr* em mutantes de *Azospirillum brasilense* constitutivos para fixação de nitrogênio.** Tese de

Doutorado em Ciências-Bioquímica, Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2001.

VITORINO, J. C.; STEFFENS B. R.; MACHADO, H. B.; YATES, G.; SOUZA, E. M.; PEDROSA, F. O. Potential roles for the *glnB* and *ntrXY* genes in *Azospirillum brasilense*. **FEMS Microbiol. Lett.** 201, 199-204, 2001.

VON MERING, C.; HUYNEN, M.; JAEGGI, D.; SCHMIDT, S.; BORK, P.; SNEL, B. STRING: a database of predicted functional associations between proteins. **Nucleic Acids Res.** v. 31(1), p. 258-61, 2003.

de ZAMAROCZY, M.; F. DELORME, e C. ELMERICH. Characterization of three different nitrogen-regulated promoter regions for the expression of *glnB* and *glnA* in *Azospirillum brasilense*. **Mol. Gen. Genet.** 224:421-430, 1990.

de ZAMAROCZY, M.; PAQUELIN, A.; ELMERICH, C. Functional Organization of the *glnB-glnA* Cluster of *Azospirillum brasilense*. **J. Bacteriol.**, Washington, v. 175, p. 2507-2515, 1993.

de ZAMAROCZY, M.; PAQUELIN, A.; PELTRE, G.; FORCHHAMMER, K.; ELMERICH, C. Coexistence of two structurally similar but functionally different PII protein in *Azospirillum brasilense*. **J. Bacteriol.**, Washington, v. 178, p. 4143-4149, 1996

de ZAMAROCZY, M. Structural homologues PII and PZ of *Azospirillum brasilense* provide intracellular signaling for selective regulation of various nitrogen-dependent functions. **Mol. Microb.** v. 29, p. 449-463, 1998

ZHANG, Y. BURRIS, R. H.; LUDDEN, P. W. ROBERTS, G. P. Posttranslational regulation of nitrogenase activity in *Azospirillum brasilense ntrBC* mutants: ammonium and anaerobic switch-off occurs through independent signal transduction pathways. **J. Bacteriol.** v. 176 n. 18, p. 5780-5787, 1994.

ZHANG, Y.; BURRIS, R. H.; LUDDEN, P. W.; ROBERTS, G. P. Regulation of nitrogen fixation in *Azospirillum brasilense*. **FEMS Microbiol. Rev.** v. 152, p195-204, 1997.

ZUMFT, W. Z.; MORTENSON, L. E. The nitrogen-fixing complex of bacteria. **Reviews on Bioenergetics**, Amsterdam v. 416, 1975.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)