

PATRÍCIA CASTELLEN

**SEQÜENCIAMENTO E ANÁLISE DE REGIÕES
GENÔMICAS DE *Azospirillum brasilense***

Monografia apresentada como requisito parcial
para a obtenção do título de Bacharel no curso
de Ciências Biológicas, Setor de Ciências
Biológicas, Universidade Federal do Paraná.

Curitiba
2002

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 ao Prof. Emanuel e à Prof.^a. Roseli Wassem que me ensinaram a ver a pesquisa científica como alguém que faz parte dela. Obrigada pelas horas que gastaram me ajudando a ver o caminho, eu nunca vou esquecer. Sem vocês eu não teria conseguido.

Aos colegas do Anexo: Angelita, Karen, Daniela, Ana Paula, Helisson, Leo, Jú Ramos, Humbert, Estefo, Rose Adele, Gustavo, Daniel e Valter. Obrigada pelas discussões sócio-econômico-políticas que tivemos nas horas vagas. Não esquecendo das conversas sobre astronomia e também pelas risadas que aliviaram os dias mais pesados.

A todos os colegas do núcleo: Liloca, Mari, Luciano, Luíza, Adriana Lacerda, Adriana I., Ângela, Fabi, Marcelo, Lauren, Renata, Carol, Andréa, Juliana Schmidt, André e Rafael. Pelas informações, dicas e por termos convivido muito bem.

Às professoras Elaine, Giseli, Cynthia, Leda e Berenice por dividirem seu conhecimento.

A D. Jú e Roseli Prado pelas ajudas e pelos conselhos também.

À Jú Inaba, Ana Claudia e Lys por serem minhas amigas, eu sei que vou levar vocês pelo resto da vida, sempre vão estar comigo. Jú, um dia vão reconhecer o verdadeiro valor da nossa música.

Aos amigos que a vida me deu nesse período: Jana, Carol, Ale, Pry, Mari, James, Jack e Puko. Sei que tive sorte por conhecê-los. Aprendi mais do que esperava com cada um de vocês. Jana, se não fosse você e sua família...

À minha família, que é a melhor do mundo inteiro! 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 que não sei o que dizer.

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.....	VII
Resumo.....	IX
1. Introdução.....	1
1.1 <i>Azospirillum brasilense</i>	1
1.2 A fixação biológica de nitrogênio.....	2
1.3 O sistema <i>ntr</i>	4
1.4 Regulação da fixação de nitrogênio em <i>Azospirillum brasilense</i>	6
2. Objetivos	9
3. Material e Métodos.....	10
3.1 Bactérias e Plasmídeos.....	10
3.2 Condições de cultivo.....	11
3.3 Composição dos meios de cultura.....	11
3.4 Purificação de DNA plasmidial.....	12
3.5 Eletroforese e purificação de fragmentos de DNA.....	13
3.6 Condições de digestão de DNA com enzimas de restrição.....	14
3.7 Nebulização.....	14
3.8 Preparo de DNA com pontas não coesivas (“blunt”).....	14
3.9 Digestão dos vetores.....	15

3.10	Ligação de DNA.....	15
3.11	Mutagênese com Transposons.....	15
3.12	Transformação por eletroporação.....	16
3.12.1	Preparo de células competentes.....	16
3.12.2	Transformação bacteriana.....	16
3.13	Seqüenciamento.....	16
3.13.1	Purificação.....	16
3.13.2	Reação de seqüenciamento.....	17
4.	Resultados e Discussão.....	18
4.1	Estratégias de clonagem.....	18
4.2	Seqüenciamento e análise das seqüências.....	20
4.3	Fator “Trigger”.....	23
4.4	Lon, ClpP e ClpX: proteases dependentes de ATP.....	25
4.5	A proteína HU.....	28
4.6	NADH-desidrogenase.....	29
4.7	Considerações finais... ..	31
5.	Referências Bibliográficas.. ..	52

LISTA DE FIGURAS

	Pág
01 Regulação da Atividade de GS e NtrC em Resposta à Concentração de Nitrogênio.....	05
02 Mapa de Restrição do Plasmídeo pAB441 e seus Subclones.....	19
03 Seqüência da Região à Montante do Operon <i>glnBA</i> Obtida a Partir de pH/P13.....	22
04 Resultado da Análise do Padrão de Organização.....	24
05 Seqüência de Nucleotídeos Obtida e seu Produto de Tradução.....	32
06 Alinhamento da Seqüência de Aminoácidos da Proteína TF <i>A. brasilense</i> com Diversos Organismos.....	40
07 Alinhamento da Seqüência de Aminoácidos da Proteína ClpP de <i>A. brasilense</i> com Diversos Organismos.....	42
08 Alinhamento da Seqüência de Aminoácidos da Proteína ClpX de <i>A. brasilense</i> com Diversos Organismos.....	43
09 Alinhamento da Seqüência de Aminoácidos da Protease Lon de <i>A. brasilense</i> com Diversos Organismos.....	45
10 Alinhamento da Seqüência de Aminoácidos da Proteína HU de <i>A. brasilense</i> com Diversos Organismos.....	48
11 Seqüência de Nucleotídeos Encontrada para <i>nuoB</i> e seu Produto de Tradução	49
12 Alinhamento da Seqüência de Aminoácidos da Subunidade β da NADH-desidrogenase de <i>A. brasilense</i> com Diversos Organismos.....	50
13 Genes Identificados no Plasmídeo pAB441.....	51

LISTA DE TABELAS

Pág.

01 Bactérias e Plasmídeos.....

10

LISTA DE ABREVIATURAS

µg	micrograma
µL	microlitro
Amp	ampicilina
AMP	adenosina monofosfato
ATP	adenosina trifosfato
ATPase	adenosina trifosfatase
CRP	receptor de AMP cíclico
D. O.	densidade óptica
DNA	ácido desoxiribonucleico
EDTA	ácido etilenodiamino-tetra-acético
g	grama
h	hora
kb	quilo pares de bases
Km	canamicina
L	litro
mg	miligrama
min	minuto
miniprep	minipreparação de plasmídeos
mL	mililitro
nm	nanômetro
°C	temperatura em graus Celsius
pb	pares de bases
r	resistência à
RNA	ácido ribonucléico
RNase	ribonuclease
rpm	rotações por minuto
Sm	estreptomicina

Tc	tetraciclina
Tris	tris (hidroximetil)-aminometano
V	volts

RESUMO

As bactérias do gênero *Azospirillum* são encontradas associadas à raízes de gramíneas de interesse comercial como milho, trigo, arroz e sorgo. São aeróbias e diazotróficas, ou seja, capazes de utilizar N₂ atmosférico como fonte única de nitrogênio, reduzindo o N₂ a NH₄⁺ sob condições de baixo oxigênio. Este processo, conhecido como fixação biológica de nitrogênio é altamente regulado por um complexo mecanismo em cascata, o sistema Ntr. Em *A. brasilense* a proteína PII, produto do gene *glnB*, funciona como sensor dos níveis de amônio intracelular controlando a atividade de NifA, proteína responsável pela ativação dos genes envolvidos na síntese do complexo da nitrogenase. A proteína glutamina sintetase, que também faz parte do sistema Ntr, é codificada pelo gene *glnA* e converte amônio a glutamina. O plasmídeo pAB441 foi isolado de um banco genômico de *Azospirillum brasilense*, apresentando um inserto de 20 Kb. Este plasmídeo contém o operon *glnBA* e estudos de complementação genética de mutantes de *A. brasilense*, feitos por Vitorino e colaboradores (2001), sugerem que pode conter outros genes cujo produto participa da regulação da fixação de nitrogênio. O objetivo deste trabalho é identificar estes genes a partir da subclonagem de um fragmento de seqüência desconhecida de aproximadamente 5,0 kb do pAB441 e posterior seqüenciamento. Os resultados permitiram a identificação de cinco genes que codificam para proteínas envolvidas na resposta a choque térmico por frio, são elas: fator “trigger”, ClpP, ClpX, protease Lon e HUβ. O gene que codifica para a subunidade β da NADH-desidrogenase também foi identificado à jusante do gene para HU. Neste trabalho, nenhum gene diretamente envolvido com o metabolismo de nitrogênio foi identificado.

1. INTRODUÇÃO

1.1 *Azospirillum brasilense*

As bactérias do gênero *Azospirillum* (TARRE *et al.*, 1978) fazem parte da subdivisão α das Proteobactérias (YOUNG, 1992). Este gênero é composto por sete espécies: *Azospirillum brasilense*, *Azospirillum lipoferum* (TARRE *et al.*, 1978), *Azospirillum halopraeferans* (REINHOLD *et al.*, 1987), *Azospirillum amazonense* (MAGALHÃES *et al.*, 1983), *Azospirillum irakense* (KHAMMAS *et al.*, 1989), *Azospirillum largimobile* (DEKHIL *et al.*, 1997) e *Azospirillum doebereineriae* (ECKERT *et al.*, 2001). Estes são microrganismos, aeróbios, diazotróficos, ou seja, são capazes de utilizar o dinitrogênio atmosférico como única fonte de nitrogênio e também têm a capacidade de se associar endofiticamente à raízes de plantas de interesse agrícola (DÖBEREINER e DAY, 1976; DÖBEREINER, 1991; PEDROSA, 1988).

O mecanismo de interação entre a planta e *Azospirillum* não é completamente conhecido. As evidências sugerem que a invasão envolve uma etapa inicial reversível na qual a bactéria adere à raiz através do flagelo polar e em seguida secreta polissacarídeos para aderir fortemente à planta (STEENHOUDT e VANDERLEYDEN, 2000).

Vários estudos mostram que a interação *Azospirillum*-planta tem efeito sobre a fisiologia da planta, aumentando sua produtividade. *Azospirillum* é capaz de estimular o aumento do número e do tamanho dos pêlos absorventes das raízes, o aparecimento de raízes laterais e o aumento do volume e da superfície radicular (KAPULNIK *et al.*, 1985; PEDROSA, 1988). Essas mudanças fisiológicas podem estar relacionadas à produção, pelas bactérias, de fitormônios tais como o ácido indolacético, giberelinas e auxinas, que

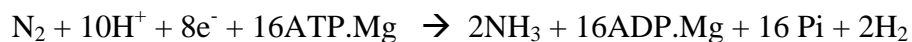
umentam a assimilação de nutrientes pela planta, principalmente em estágios iniciais de desenvolvimento (PEDROSA, 1988; BASHAN e HOLGUIN, 1997).

1.2 A fixação biológica de nitrogênio

As bactérias diazotróficas podem ser classificadas em bactérias de vida livre, associativas e endofíticas. A fixação biológica de nitrogênio é o processo de redução do dinitrogênio atmosférico (N_2) a amônia (NH_3) e as bactérias diazotróficas são os únicos organismos capazes de realizar esta conversão, disponibilizando a amônia para a síntese de moléculas que contenham nitrogênio (BURRIS, 1991). Este processo é de grande importância econômica, pois o nitrogênio é um dos fatores limitantes para o desenvolvimento de plantas. O uso da simbiose entre gramíneas e bactérias fixadoras de nitrogênio possibilita uma substancial economia de fertilizantes nitrogenados na agricultura (PEDROSA, 1987). A fixação de nitrogênio requer condições ambientais favoráveis tais como a ausência de nitrogênio fixado, baixas tensões de oxigênio, presença de molibdênio e temperatura adequada (POSTGATE, 1982). O processo de fixação de nitrogênio é altamente regulado tanto ao nível de atividade enzimática quanto ao nível de expressão gênica, pois envolve um alto gasto energético (FISCHER, 1994).

A reação de redução do dinitrogênio a amônia é catalisada pelo complexo enzimático nitrogenase. Este é formado por duas proteínas codificadas por genes cuja transcrição é regulada pela proteína NifA. São elas: 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. A dinitrogenase redutase é um homodímero γ_2 codificado pelo gene *nifH* (DEAN *et al.*, 1993).

A reação catalisada pela nitrogenase é a seguinte:



A amônia produzida nesta reação é utilizada na síntese de glutamina e glutamato, que servem como doadores de nitrogênio para as reações biossintéticas (MERRICK & EDWARDS, 1995).

A glutamato sintase catalisa a aminação redutiva do α -cetoglutarato, 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-transcricional. Pelo menos seis produtos finais do metabolismo da glutamina, além da alanina e da glicina, são inibidores alostéricos da enzima. Além da regulação alostérica há a inibição por adenililação, que aumenta a sensibilidade da enzima para os inibidores alostéricos (MAGASANIK, 1993). No modelo para bactérias entéricas, baseado em *Klebsiella pneumoniae*, a adenililação e desadenililação da GS são promovidas pela adenilil transferase, que tem sua atividade modulada pela ligação a uma proteína sinalizadora dos níveis de NH_4^+ chamada PII. PII ocorre na forma não modificada (não uridililada), sinalizando altos níveis de NH_4^+ , e na forma uridililada (PII-UMP), sinalizando baixos níveis de NH_4^+ . O resultado final deste mecanismo complexo de regulação é um decréscimo da atividade da glutamina sintetase, quando os níveis de glutamina são altos, e um aumento dessa atividade, quando os níveis de glutamina são baixos e estão disponíveis os substratos α -cetoglutarato e ATP (MERRICK & EDWARDS, 1995).

1.3 O sistema *ntr*

O sistema *ntr* compreende um conjunto de genes cujos produtos estão envolvidos na cascata que regula a expressão dos genes do metabolismo nitrogenado.

Em enterobactérias este sistema é composto pelas seguintes proteínas: NtrB (gene *ntrB*), NtrC (gene *ntrC*), GlnD (gene *glnD*), proteína PII ou GlnB (gene *glnB*), adenililtransferase ou GlnE (gene *glnE*) e glutamina sintetase (gene *glnA*). Essas proteínas interagem em um complexo mecanismo em cascata representado esquematicamente na figura 1 (MERRICK & EDWARDS, 1995).

A proteína PII interage com três destas proteínas, GlnD (Urase/UR), GlnE e NtrB. Em todos os casos as interações dependem dos níveis de carbono e nitrogênio da célula, os quais são sinalizados pela presença de glutamina, α -cetogluturato e ATP (ARCONDÉGUY *et al.*, 2001). A uridililação de PII ocorre em presença de α -cetogluturato e ATP sob baixas concentrações de glutamina, e a desuridililação também requer α -cetogluturato e ATP mas é estimulada pela presença de glutamina (KAMBEROV *et al.*, 1995).

A proteína GlnD, produto do gene *glnD*, atua sobre PII de duas formas dependendo do nível intracelular de nitrogênio fixado (indicado pela concentração de glutamina no interior da célula): em condições limitantes de NH_4^+ a proteína GlnD possui atividade de uridililtransferase e catalisa a adição do grupamento UMP à proteína PII; em condições de excesso de NH_4^+ , GlnD se torna uma enzima removedora de UMP e converte PII-UMP na forma não uridililada (DE MEL *et al.*, 1994).

REGULAÇÃO DA ATIVIDADE DE GS E NtrC EM RESPOSTA À CONCENTRAÇÃO DE AMÔNIO

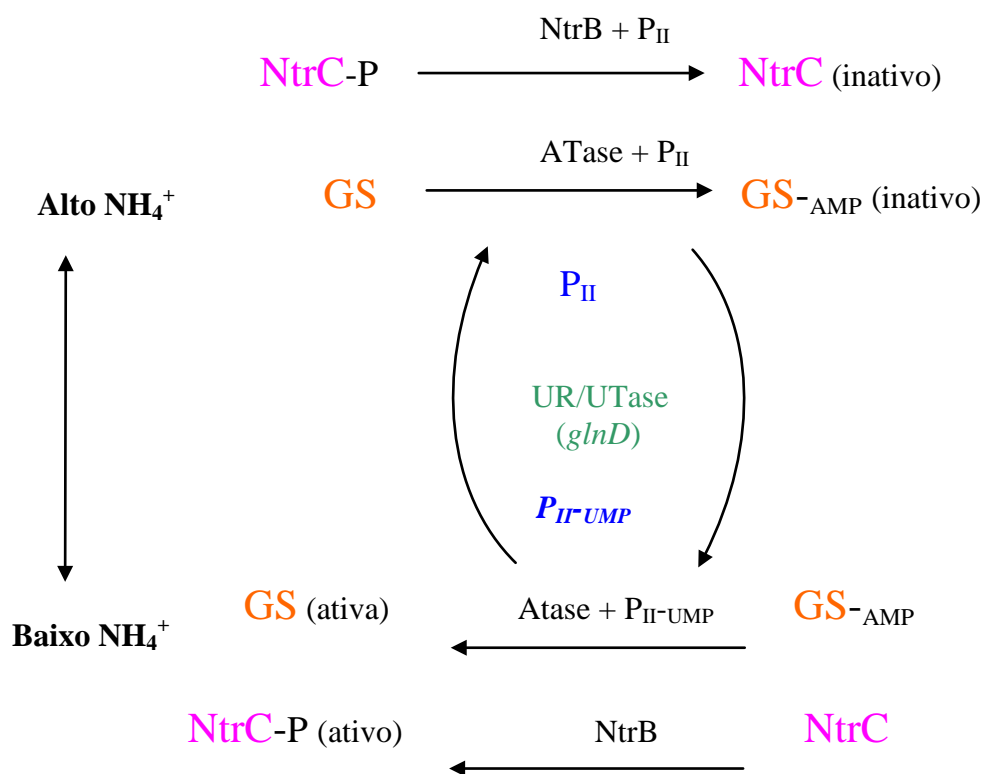


Figura 1: Modelo esquemático da regulação das atividades da glutamina sintetase e sistema Ntr em resposta à pressão de nitrogênio no meio. A Utase (*GlnD*) catalisa a uridililação e deuridililação da GS. A NtrB catalisa a fosforilação e defosforilação da NtrC (MERRICK & EDWARDS, 1995).

Em presença de baixas concentrações de α -cetogluturato e alta concentração de amônio e glutamina, PII se liga a NtrB induzindo a atividade de fosfatase da proteína. Por outro lado, sob altas concentrações de α -cetogluturato e baixas concentrações de glutamina, PII é ligada a UMP que não interage com NtrB. Na forma livre, NtrB tem atividade de quinase e fosforila NtrC (JIANG e NINFA, 1999).

A proteína NtrC fosforilada ativa a transcrição do operon *glnAntrBC*, *nifLA* e de outros sistemas alternativos de assimilação de fonte nitrogênio (MERRICK & EDWARDS, 1995). Em contrapartida, PII-UMP interage com a adenililtransferase (ATase, produto do gene *glnE*) que catalisa a desadenililação da glutamina sintetase (KUTSU *et al.*, 1971; KAMBEROV *et al.*, 1995 e MERRICK & EDWARDS, 1995). A glutamina sintetase desadenililada passa então a catalisar a produção de glutamina (KEENER & KUSTU, 1988).

Em condições de excesso de nitrogênio a proteína GlnD, desuridilila PII-UMP (de MEL *et al.*, 1994). A proteína PII desuridililada estimula a atividade de fosfatase da proteína NtrB, desfosforilando NtrC-P e inativando-a. Ao mesmo tempo, PII desuridililada age sobre a ATase (GlnE), que passa a catalisar a adenililação da glutamina sintetase. A glutamina sintetase adenililada é inativa (HAWKES *et al.*, 1985 e MERRICK & EDWARDS, 1995).

1.4 Regulação da fixação de nitrogênio em *Azospirillum brasilense*

A expressão dos genes do controle do metabolismo de nitrogênio em *A. brasilense*, como em outros diazotrofos, está sob o controle de fatores ambientais como fonte de nitrogênio e concentração de oxigênio (DOBEREINER e PEDROSA, 1987). A regulação da fixação de nitrogênio em *A. brasilense* apresenta algumas diferenças em relação à

regulação em *K. pneumoniae*. No caso de *A. brasilense*, a expressão de *nifA* não é regulada pelas proteínas NtrB e NtrC (MACHADO *et al.*, 1995; ELMERICH *et al.*, 1997; FADEL-PICHET *et al.*, 1999). Em *A. brasilense*, o gene *nifA* é expresso na presença de alta concentração de NH_4^+ ou de O_2 , mas NifA é inativa em altos níveis de NH_4^+ e O_2 (LIANG, 1991). Além disso, a atividade de NifA não é controlada por NifL como em γ -proteobactérias, mas é auto-inibida por modificações conformacionais de seus motivos estruturais (ARSENE *et al.*, 1996). A proteína PII de *A. brasilense* é sintetizada em altos níveis sob condições de nitrogênio limitante e em baixos níveis sob condições de excesso de nitrogênio (de ZAMAROCZY *et al.*, 1993) e não é essencial para a adenililação da glutamina sintetase (de ZAMAROCZY, 1998).

O gene que codifica para a glutamina sintetase foi isolado de uma biblioteca genômica de *A. brasilense* por complementação de mutantes. (BOZOUKLIAN *et al.*, 1986). Mapeamento por hibridização com sonda contendo o gene *glnA* de *K. pneumoniae* e por mutagênese por Tn5 permitiu a localização do gene *glnA* de *A. brasilense* no inserto de 20kb do plasmídeo pAB441. Este plasmídeo é capaz de restaurar o crescimento independente de glutamina e o fenótipo Nif^+ em mutantes $\text{Gln}^- \text{Nif}^-$ (sem atividade de GS e nitrogenase) e mutantes $\text{Gln}^- \text{Nif}^c$ (constitutivos, ou seja, com atividade de nitrogenase independente do nível de nitrogênio fixado do meio) de *A. brasilense*. O seqüenciamento da região a montante do gene *glnA* indicou a presença de uma ORF (“open reading frame” ou fase aberta de leitura) homóloga ao gene *glnB*, e portanto codifica para a proteína PII (de ZAMAROCZY *et al.*, 1990).

Estirpes mutantes de *A. brasilense* com atividade de nitrogenase constitutiva, foram isoladas por Machado e colaboradores (1991). Anteriormente Vitorino e colaboradores

(2001) mostraram que os mutantes HM26, HM053 e HM210, contendo o plasmídeo pAB441, foram complementados para crescimento dependente de nitrato mas perderam a atividade da nitrogenase tanto na ausência como na presença de amônio. A atividade da nitrogenase no mutante HM14 e na estirpe selvagem não foi alterada em presença do plasmídeo pAB441 (VITORINO *et al.*, 2001). Dessa forma é levantada a hipótese de que a perda da atividade da nitrogenase em alguns mutantes pode ser um efeito conjunto de suas mutações e outros genes presentes no inserto de 20 kb envolvidos no metabolismo de nitrogênio, além dos genes *glnB* e *glnA*.

2. OBJETIVOS

Resultados de experimentos de complementação realizados por Vitorino e colaboradores (2001) sugerem a existência de um ou mais genes envolvidos na fixação de nitrogênio no plasmídeo pAB441, além dos genes que codificam para as proteínas glutamina sintetase e PII. Assim os objetivos deste trabalho são:

- Subclonar um fragmento *HindIII/PstI* de aproximadamente 5,0kb do plasmídeo pAB441;
- Seqüenciar completamente o subclone obtido;
- Analisar as seqüências obtidas e identificar possíveis genes envolvidos na regulação da fixação de nitrogênio em *Azospirillum brasilense*.

3. MATERIAL E MÉTODOS

3.1 Bactérias e Plasmídeos

Estirpe ou plasmídeo	Propriedades	Referência ou Fonte
Estirpe		
<i>E. coli</i> DH10B	Sm ^r F ⁻ [<i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) <i>lacZ</i> ΔM15]	Invitrogen
Plasmídeo		
Vetor pTZ 18R	Cb <i>lacZ</i> f1 IG	MEAD <i>et al.</i> , 1986
pAB441	20kb do genoma de <i>A. brasilense</i> em vetor pVK100	BOZOUKLIAN <i>et al.</i> , 1986
pH/P13	Subclone <i>HindIII/PstI</i> de 5 kb do pAB441	Este trabalho
pH/P13+TnA04	pH/P13 com inserção de transposon	Este trabalho
pH/P13+TnB13	pH/P13 com inserção de transposon	Este trabalho
pH/P13+TnB15	pH/P13 com inserção de transposon	Este trabalho
pH/P13 <i>SstI</i> 3	Subclone <i>SstI</i> pH/P13 com 500 pb	Este trabalho
pABAA02	Subclone de banco aleatório	Este trabalho
pABAC01	Subclone de banco aleatório	Este trabalho
pABAC08	Subclone de banco aleatório	Este trabalho
pABAD07	Subclone de banco aleatório	Este trabalho
pABAH04	Subclone de banco aleatório	Este trabalho
pABCA06	Subclone de banco aleatório	Este trabalho
pABCC02	Subclone de banco aleatório	Este trabalho
pABCC10	Subclone de banco aleatório	Este trabalho
pABCE12	Subclone de banco aleatório	Este trabalho
pABCF05	Subclone de banco aleatório	Este trabalho

3.2 Condições de cultivo

As estirpes de *E. coli* foram cultivadas em meio líquido Luria-Broth (LB) (SAMBROOK *et al.*, 1989) 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 (SAMBROOK *et al.*, 1989) 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.

3.3 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)

Extrato de Levedura	24,0 g/L
Bactotriptona	12,0 g/L
Glicerol	4,0 g/L

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

Antibióticos

Os antibióticos usados foram:

	Concentração final
Ampicilina (Amp)	250 µg/mL
Estreptomicina (Sm)	80 µg/mL
Canamicina (Km)	50 µg/mL
Tetraciclina (Tc)	10 µg/mL

3.4 Purificação de DNA plasmidial

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 com ampicilina e estreptomicina até uma D.O. de aproximadamente 2,0. A cultura foi transferida para tubos plásticos de 1,5 mL, as células coletadas por centrifugação (13.000 x g, 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). 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 10 µg/mL e analisadas em gel de ágar.

3.5 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 (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 com solução de brometo de etídeo (0,5 µg/mL) em transiluminador de luz ultravioleta (312 nm) e o perfil eletroforético foi registrado fotograficamente ou utilizando um sistema de vídeo imagem acoplada a equipamento de impressão a seco (GDS 5000). 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 cortadas sob iluminação luz ultravioleta e o fragmento de DNA foi extraído do gel por tratamento com fenol (pH 7-8) como descrito (BANKIER *et al.*, 1987).

3.6 Condições de digestão de DNA com enzimas de restrição

As condições utilizadas para a digestão de DNA com enzimas de restrição foram aquelas especificadas pelo fabricante ou por SAMBROOK *et al.* (1989).

Usualmente 1 U 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.

3.7 Nebulização

O plasmídeo pAB441 foi nebulizado seguindo-se o protocolo do “TOPO Shotgun Subcloning Kit” (Invitrogen) conforme as especificações do fabricante.

3.8 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) acetilada, 0,17 mmol/L de dNTP's 0,4U de T4 DNA polimerase e 1µg DNA.

3.9 Digestão dos vetores

Os vetores foram linearizados com a enzima de restrição adequada. Após a restrição, a enzima foi desnaturada por calor ou por extração com fenol-clorofórmio-álcool isoamílico (25:24:1). O DNA extraído foi precipitado com etanol absoluto, lavado, seco e dissolvido em água milli-Q estéril.

3.10 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 a 16°C durante a noite. No caso de ligação em terminal coesivo, a mistura vetor-inserto foi aquecida a 65°C por 5 minutos e resfriada em gelo antes da adição do tampão e da enzima (SAMBROOK *et al.*, 1989)

3.11 Mutagênese com transposons

Para seqüenciar o fragmento de 5 kb do plasmídeo pAB441 o transposon EZ::TN TET 1 foi integrado aleatoriamente no plasmídeo pH/P13. 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 0,2 µg de plasmídeo, 0,05 pmol de EZ::TN <TET-1> Transposon, 1U de transposase, 1µL de tampão de reação e água estéril. Os clones obtidos foram purificados, analisados e utilizados como molde em reações de sequenciamento, utilizando oligonucleotídeos iniciadores complementares às extremidades do transposon.

3.12 Transformação por eletroporação

3.12.1 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 10%, 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.

3.12.2 Transformação bacteriana

O método utilizado foi indicado pelo fabricante do eletroporador (Gibco-BRL).

3.13 Seqüenciamento

3.13.1 Purificação

A purificação de DNA plasmidial para posterior seqüenciamento foi realizada através da lise alcalina conforme descrita no item 3.4. Após dissolvido em água, o DNA foi tratado com RNase (concentração final de 1mg/mL) por 1 hora a 37°C. Foi feita, em seguida, uma extração com 70µL de fenol-clorofórmio-álcool isoamílico e uma extração com clorofórmio-álcool isoamílico. O DNA foi precipitado com 42L de acetato de amônio (7,5 M, pH 8,0) e 250 µL de etanol absoluto, lavado com etanol 80%, seco e ressuspensado em 20µL de água.

3.13.2 Reação de Seqüenciamento

O seqüenciamento é baseado no método de 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 - Amershan Biosciences), num volume final de 10µL. A reação foi feita utilizando-se um termociclador e os parâmetros variavam conforme o “primer” utilizado na reação.

Após a reação de seqüenciamento o produto da reação foi purificado com 2µL de acetato de sódio 2M e 50µL de etanol absoluto, lavado com etanol 70%, seco e submetido a eletroforese no Seqüenciador Automático de DNA (Perkin-Elmer Modelo ABI377). As seqüências obtidas foram transferidas eletronicamente ao microcomputador e analisadas com os programas BLASTX (ALTSHUL *et al.*, 1997), Auto Assembler (Perkin-Elmer), ORF Finder (NCBI), BioEdit Sequence Alignment Editor (HALL, 1999) e Vector NTI (InforMax, Inc.).

4. RESULTADOS E DISCUSSÃO

4.1 Estratégias de clonagem

Com o objetivo de identificar genes relacionados ao metabolismo de nitrogênio presentes no plasmídeo pAB441, deu-se início ao seqüenciamento do inserto de 20kb deste plasmídeo cujo tamanho total é de 40kb. Para que este objetivo fosse alcançado, diversas estratégias foram utilizadas. Inicialmente, foram construídas bibliotecas aleatórias (“shotgun” total) do plasmídeo pAB441 clivado com a enzima de restrição *Sau3AI* ou por nebulização. Fragmentos de DNA de 1 a 3 kb obtidos por ambos os métodos foram clonados no vetor de seqüenciamento pTZ18R e transformadas em *E. coli* estirpe DH10B. As colônias de *E. coli* transformantes contendo os plasmídeos foram coletadas e repicadas em placas de 96 poços para facilitar a purificação e permitir o seqüenciamento de um grande número de clones em menor tempo. Os clones de duas placas de 96 amostras cada, tiveram uma de suas extremidades seqüenciadas e analisadas. Os resultados mostraram que mais de 80% dos clones possuíam insertos de DNA de *Escherichia coli*. A clonagem de DNA de *E. coli* ocorreu, provavelmente, em virtude da co-purificação de DNA genômico da *E. coli* hospedeira do pAB441. Os contíguos 39 e 44 apresentados na Figura 2 foram obtidos a partir do alinhamento de seqüências obtidas utilizando a técnica de “shotgun”. Em função das dificuldades encontradas na purificação de DNA cosmidial, optou-se pela mudança de estratégia.

A segunda estratégia, adotada para facilitar a obtenção de DNA para o seqüenciamento, envolveu a subclonagem de fragmentos de restrição do inserto de 20 kb no vetor pTZ18R. Utilizando o mapa de restrição do pAB441 (BOZOUKLIAN *et al.*, 1986) foram realizadas diversas subclonagens. Entre os subclones obtidos foi isolado o

MAPA DE RESTRIÇÃO DO PLASMÍDEO pAB441 E SEUS SUBCLONES

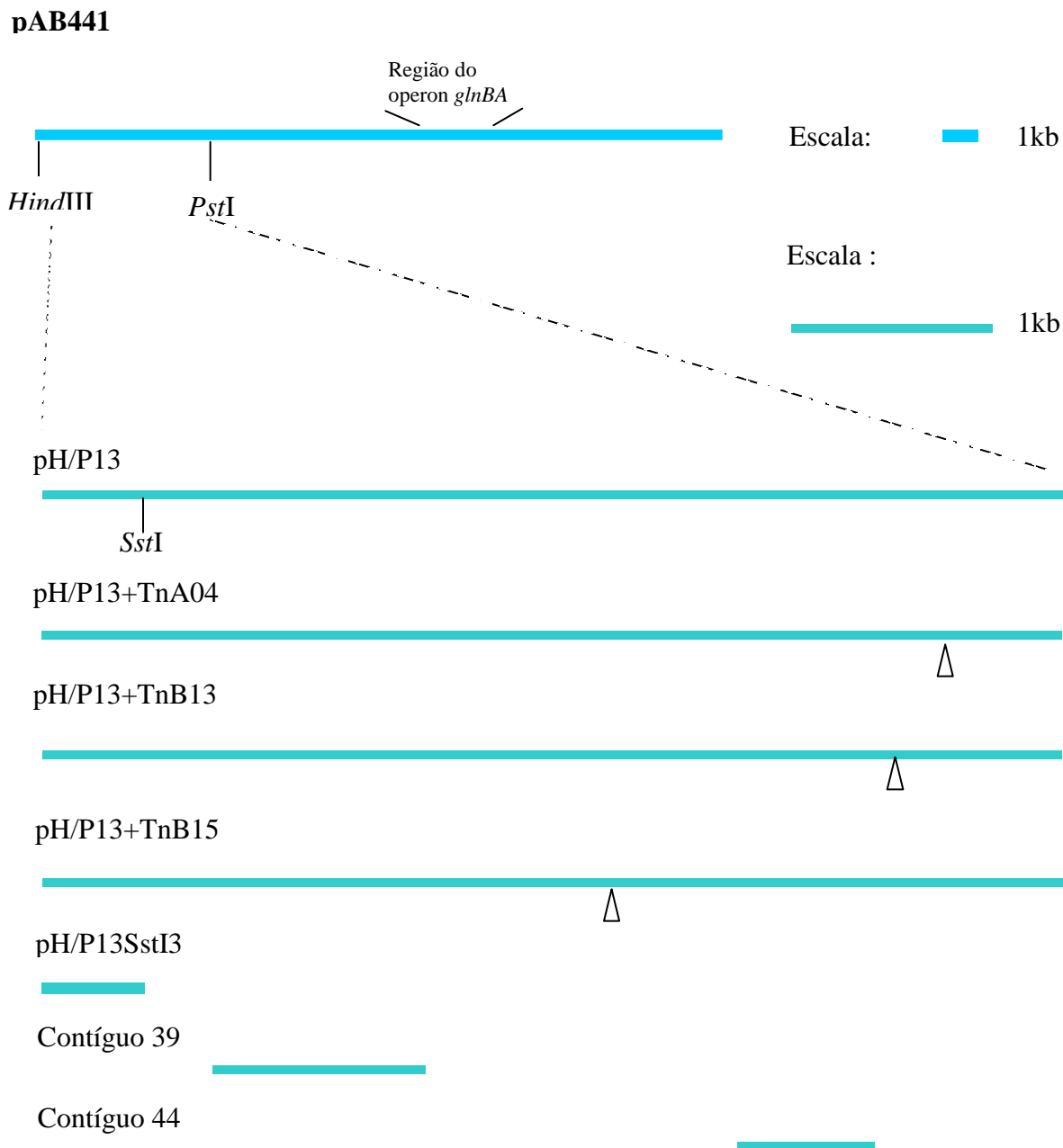


Figura 2: Posições dos insertos de clones obtidos a partir do plasmídeo pAB441. O pH/P13 é um subclone *Hind*III/*Pst*I com aproximadamente 5 kb de inserto; pH/P13+TnA04, pH/P13+TnB13 e pH/P13+TnB15 são clones obtidos a partir da inserção de transposons nas posições indicadas por um triângulo; pH/P13*Sst*I3 é um clone obtido a partir de pH/P13 digerido *Sst*I e religado que possui aproximadamente 500 pares de bases de inserto; os contíguos mostrados são resultado de alinhamento de seqüências de clones do banco aleatório.

pH/P13 que possui o fragmento *HindIII/PstI* de aproximadamente 5kb localizado a montante do operon *glnBA* (Figura 2), numa região pouco conhecida do inserto de 20kb.

Para realizar o seqüenciamento do inserto de aproximadamente 5kb do plasmídeo pH/P13 optou-se pela inserção aleatória do transposon EZ::TN <TET-1> (Epicentre) no inserto e deleção de um fragmento de 500 pares de base com a enzima *SstI*. A inserção do transposon em regiões aleatórias do inserto permitiu o seqüenciamento a partir da ponta de inserção, através da utilização de oligonucleotídeos iniciadores específicos para as extremidades do transposon. As regiões a partir da inserção dos transposons que resultaram em ampliação da seqüência estão representadas na Figura 2.

4.2 Seqüenciamento e análise das seqüências

Os clones obtidos foram seqüenciados e suas seqüências foram alinhadas utilizando-se o programa Auto Assembler (Applied Biosystems) e comparadas com as seqüências depositadas no banco de dados GenBank, utilizando-se o programa BLAST (ALTSCHUL *et al.*, 1997).

O alinhamento das seqüências obtidas a partir de subclones do pH/P13 (Figura 3) com as seqüências já depositadas no banco de dados, resultou numa região contígua de aproximadamente 7,6 kb. Alinhando-se a seqüência resultante com outras regiões contíguas obtidas a partir do banco aleatório, obteve-se uma seqüência de aproximadamente 7,9 kb.

A análise das seqüências mostrou que uma região de aproximadamente 4kb, correspondente a uma parte do inserto do plasmídeo pH/P13, já havia sido seqüenciada e depositada no banco de dados GenBank (MORI *et al.*, 1996, INDORATO *et al.*, 1999). Esta é a região codificadora das proteínas fator “trigger”, ClpP, ClpX e Lon-protease. Em

vista disso, deu-se preferência, neste trabalho, à obtenção de seqüências das regiões do pH/P13 que não haviam sido obtidas anteriormente.

SEQÜÊNCIA DA REGIÃO À MONTANTE DO OPERON *glnBA* OBTIDA A PARTIR DE pH/P13

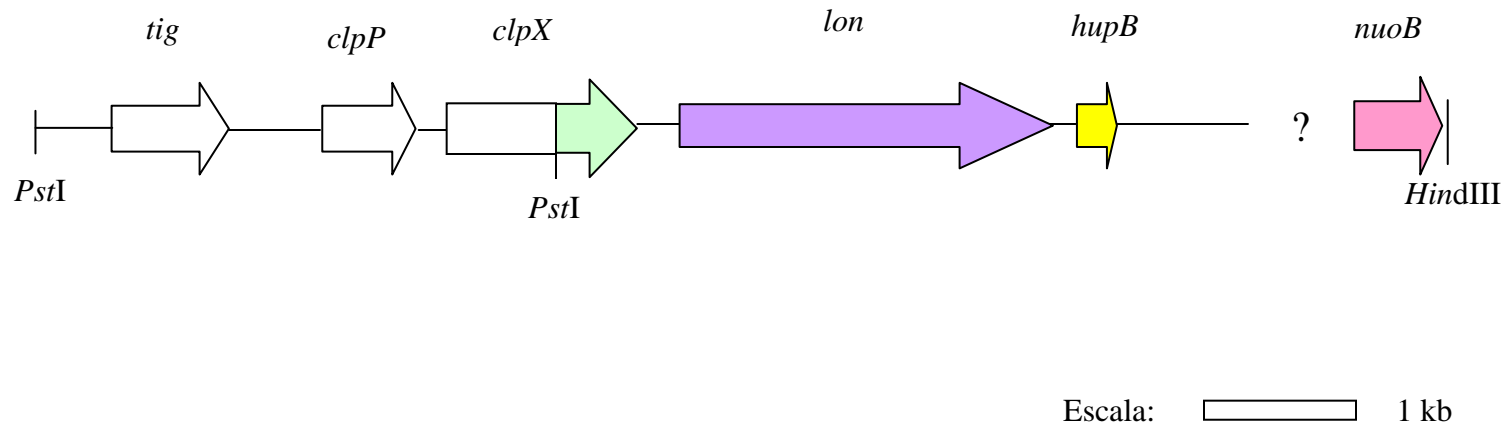


Figura 3: Esquema mostrando a região do pAB441 seqüenciada a partir de pH/P13. As regiões intergênicas estão representadas por linhas e os genes por setas largas. As setas em branco representam seqüências depositadas no banco de dados GeneBank e as setas coloridas representam os genes encontrados no inserto do plasmídeo pH/P13. O gene *lon* e parte de *clpX* fazem parte do inserto e também se encontram depositados no banco de dados.

Os resultados das análises da seqüência final utilizando os programas Orf Finder (HALL, 1999) e BLAST indicam a presença de pelo menos cinco proteínas relacionadas à resposta ao choque térmico por frio (“cold shock”) no plasmídeo pAB441. São elas: Fator “Trigger”, as proteases ClpP e ClpX, protease Lon e a subunidade β da proteína HU (Figura 3). Uma análise posterior, utilizando o programa STRING, revelou que este padrão de organização de genes é altamente conservado entre as bactérias (Figura 4).

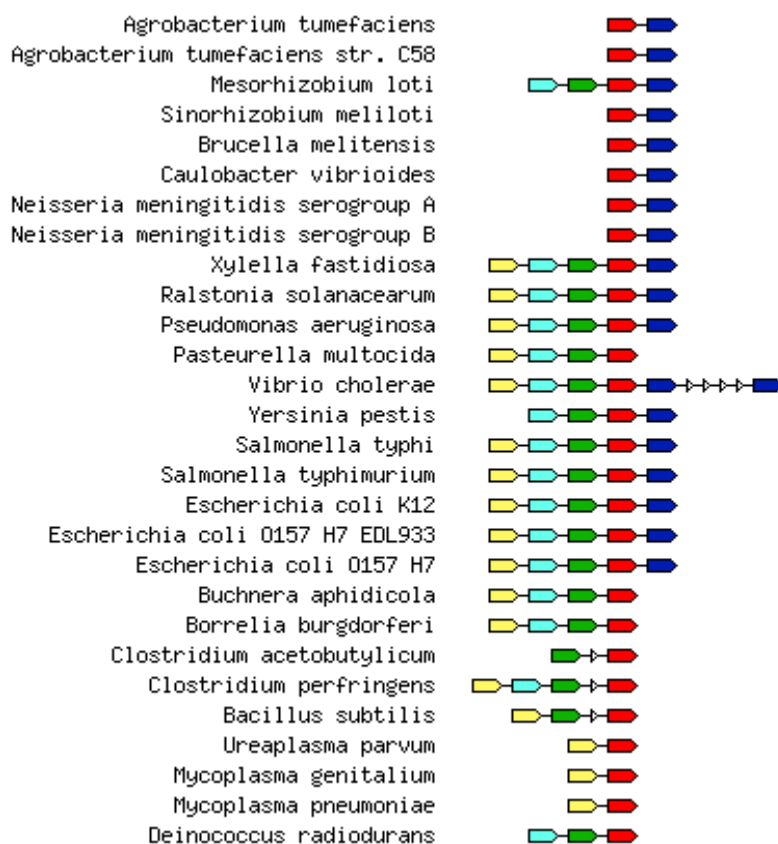
O seqüenciamento do subclone pH/P13 (Figura 3) não foi finalizado. A extremidade *HindIII* deste subclone contém parte do gene *nuoB*, que codifica para a subunidade β da NADH-desidrogenase. Não foi possível determinar com exatidão o tamanho da região entre HU e NADH-desidrogenase (subunidade β). Esta região provavelmente contém o gene que codifica para a subunidade α da NADH-desidrogenase, pois os genes que codificam para subunidades desta proteína encontram-se em operons.

4.3 “Fator Trigger”

O “trigger factor” (TF), codificado pelo gene *tig*, é uma prolil isomerase que acelera a isomerização de resíduos de prolina entre as formas cis e trans durante a aquisição da conformação tridimensional das proteínas recém-sintetizadas. Ele se associa com cadeias polipeptídicas nascentes no ribossomo e catalisa o dobramento num processo conjunto com diversas chaperonas (SCHOLZ *et al.*, 1997).

Diferentemente da maioria das chaperonas, TF 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 nativa sob baixas temperaturas, condição na

ORGANIZAÇÃO GENÔMICA DO GENE *lon* E SEUS VIZINHOS EM DIVERSOS ORGANISMOS



Your Input:

COG0466 - ATP-dependent Lon protease, bacterial type

Predicted Functional Associations:

	Score
COG0544 - FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor)	0.896
COG1219 - ATP-dependent protease Clp, ATPase subunit	0.896
COG0740 - Protease subunit of ATP-dependent Clp proteases	0.875
COG0776 - Bacterial nucleoid DNA-binding protein	0.801

Figura 4: A análise foi feita pelo programa String e mostra que a organização dos genes vizinhos ao *lon* é bem conservada.

qual a solubilidade, agregação, e propriedades conformacionais das proteínas são diferentes quando comparadas às temperaturas normais. Além disso, TF pode ser importante na manutenção de proteínas celulares pré-existentes na sua conformação funcional – promovendo a re-naturação de proteínas desnaturadas 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 provável fator “trigger” de *A. brasilense*, identificado neste trabalho, é uma proteína de 444 aminoácidos localizada a aproximadamente 500 pares de base à montante do operon que codifica para ClpPX (Figura 5). O alinhamento das seqüências de aminoácidos desta proteína em diversos organismos mostra que ela apresenta baixo nível de conservação (Figura 6).

Não foi localizado promotor σ^{70} ou σ^{32} à montante do gene *tig* de *A. brasilense* (Figura 5). Além disso, não foi encontrada outra provável ORF numa região de 400 pares de base à montante do gene *tig*, isto está de acordo com o padrão encontrado em outros organismos (Figura 4 e 5). Uma provável região terminadora independente do fator Rho foi identificada à jusante do gene *tig* (Figura 5).

4.4 Lon, ClpP e ClpX: proteases dependentes de ATP

As proteases dependentes de ATP estão envolvidas na regulação dos níveis ou concentrações de várias proteínas no interior da célula (GOTTESMAN e MAURIZI, 1992). Além disso, muitas proteínas incorretamente dobradas são degradadas por estas proteases (SCHMIDT *et al.*, 1999). A protease Lon de *E. 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 desnaturadas 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), que isolaram e seqüenciaram o gene *lon*. Uma estirpe de *A. brasilense* mutante para o gene *lon*, obtida por mutação com Tn5, apresentou crescimento normal mas foi afetada negativamente 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 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).

A protease Lon de *A. brasilense* é constituída por 810 aminoácidos e a análise comparativa das seqüências de aminoácidos (Figura 9) revela que existem grandes regiões de homologia entre as proteínas de diversos organismos. As regiões com menor similaridade entre as diferentes seqüências estão próximas ao N-terminal e ao C-terminal.

A seqüência da região à montante da protease Lon obtida neste trabalho, permitiu a identificação de um provável promotor dependente de σ^{32} , cuja seqüência consenso é CCCTTGAA_{-N13-15}-CCCGATNT (Figura 5). A seqüência obtida nesta região diverge, em algumas bases, da seqüência depositada no banco de dados GenBank. Entretanto a seqüência depositada no banco de dados também é de baixa qualidade, uma vez que nucleotídeos não identificados estão indicados pela letra N nesta região. Mori e colaboradores (1996), autores da seqüência depositada, não identificaram uma provável região promotora para o gene *lon*, mas, determinaram os sítios de início de transcrição. Um destes sítios encontra-se na região esperada para o início de transcrição a partir do provável promotor proposto a partir da seqüência obtida neste trabalho. A seqüência

apresentada na figura 5 diverge da sequência depositada no banco de dados e nas sequências representadas a seguir, as diferenças na região do provável promotor estão indicadas em letras vermelhas.

```

4166 ttggg gttct ggccc t-gaa agggg g-cta tggct gtgcc -cgtt aga
      ||||| ||||| ||||| | ||| ||||| | |||| ||||| ||||| ||||| ||||| ||| |||
GenBank ttggg gttct ggccc t-gaa -gggg ggcta tggct gtgcc acgtt aga

```

Ensaio de mutagênese utilizando o transposon Tn5 resultaram em estirpes com atividade de NifA aumentada, alguns destes apresentavam mutação no gene *lon*. Estes resultados sugerem que a Lon protease tem efeito sobre a degradação e inativação da proteína NifA em *Rhizobium meliloti* sob condições aeróbicas (HUALA *et al.*, 1991). Este resultado pode ser de interesse numa futura investigação sobre o efeito de eliminação da atividade da nitrogenase causado pela presença do plasmídeo pAB441 nos mutantes Nif^c.

Na região à montante do gene *lon*, foram encontradas duas ORFs similares aos genes *clpX* e *clpP* de *E. coli* (MORI *et al.*, 1996; INDORATO *et al.*, 1999). A análise destas sequências indica que, assim como em *E. coli* e outros organismos, *clpP* e *clpX* estão aparentemente organizados em um operon em *A. brasilense* (Figura 5). Além disso, foi identificado um provável promotor dependente da subunidade ³² da RNA polimerase à montante do gene *clpP* e não foi identificado terminador ou promotor na região intergênica. 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 diversos organismos indica que esta não é uma proteína altamente conservada (Figura 7). Já a proteína ClpX é constituída por 422 aminoácidos e a

comparação com outros organismos mostra que é uma proteína altamente conservada (Figura 8).

4.5 Proteína HU

A proteína HU é uma proteína com função similar às das histonas dos eucariotos, que também contribui para a compactação do genoma de procariotos 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 α HU e β HU, 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 fissão do DNA na iniciação da replicação (KANO *et al.*, 1991; SHARPLES *et al.*, 1994), a cisão e re-ligação 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). Apesar da 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 al.*, 1994).

A transcrição de genes envolvidos na fixação de nitrogênio, na maioria dos casos, é dependente da unidade σ^{54} da RNA polimerase. Proteínas que dobram DNA, como a proteína IHF, ligam-se à região promotora destes genes e freqüentemente estimulam sua transcrição (HOOVER *et al.*, 1990). Ensaios de transcrição a partir do promotor *glnHp2*, que é dependente de σ^{54} , mostram que a proteína HU facilita a formação do complexo aberto para a transcrição do gene *glnH* sob condições de fixação de nitrogênio. Os mesmos

ensaios foram realizados com o promotor do gene *glnAp2* e este efeito não foi observado (CARMONA e MAGASANIK, 1996).

Em *E. coli*, a subunidade β da proteína HU tem sua expressão estimulada por 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 e homodímeros HU β em relação a homodímeros HU α , 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 subunidade β da proteína HU (Figura 5) em *A. brasilense* apresenta um provável sítio de ligação para a subunidade σ^{70} da RNA polimerase (Consenso: TTGACA_{-N13-17}-TATAAT). Diferentemente do que ocorre em *E. coli*, (CLARET e ROUVIÈRE-YANIV, 1996) não foi encontrado nesta região um sítio de ligação para CRP, que é um regulador de transcrição que liga AMPc (BRUCKNER e TITGEMEYER, 2002). Uma provável região terminadora independente do fator Rho também foi identificada. A subunidade β da proteína HU de *A. brasilense* apresenta 90 aminoácidos (Figura 10) e a comparação de sua estrutura primária entre diferentes organismos mostra que esta não é uma proteína altamente conservada entre os organismos.

4.6 NADH-desidrogenase

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

Os genes que codificam para as subunidades do complexo I em *E. coli* 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 funcional das subunidades codificadas (WEIDNER *et al.*, 1993). Em *E. coli*, *nuoC* e *nuoD* encontram-se fundidos e codificam para a subunidade NuoCD (FRIEDRICH, 1998).

A seqüência que apresenta similaridade com a NADH-desidrogenase, obtida neste trabalho, não compreende a subunidade B inteira (Figura 11). Entretanto, a seqüência de aminoácidos deduzida a partir da seqüência de nucleotídeos obtida encontra-se altamente conservada (Figura 12). Em função da grande similaridade entre as proteínas dos diversos organismos e também da posição dos genes nos diferentes genomas, é provável que o gene *nuoA* esteja presente na região não seqüenciada localizada à jusante do gene *hupB* (Figuras 3 e 13).

4.7 Considerações Finais

Além dos genes descritos, presentes no fragmento *HindIII/PstI*, foram obtidas seqüências de outras regiões do plasmídeo pAB441. Além destas seqüências, uma região à montante do operon *glnBA* também havia sido seqüenciada (CHEN *et al.*, 2000). A análise de todos estes dados nos permitiu identificar a maioria dos genes presentes neste plasmídeo, apesar de não terem sido completamente seqüenciados (Figura 13).

Assim como em *E. coli*, os genes *clpP* e *clpX* encontram-se aparentemente organizados em um operon, uma vez que não foi encontrado terminador para *clpP* nem promotor para *clpX* na região intergênica.

Foram encontradas duas seqüências promotoras σ^{32} (consenso: CCCTTGAA-N₁₃₋₁₅-CCCGATNT), uma na região regulatória do operon *clpPX* e outra na região regulatória do gene *lon*, sugerindo que em *A. brasilense*, como ocorre em *E. coli*, estas são proteínas expressas em condições de estresse ao choque térmico por frio (YAMANAKA, 1999).

Diferente de *clpPX* e *lon*, o gene *hupB* apresenta uma seqüência consenso reconhecida pela subunidade σ^{70} da RNA polimerase, o que pode indicar que *hupB* é transcrito constitutivamente. Em *A. brasilense* ainda não existem indícios de que a subunidade β da HU esteja envolvida na resposta ao choque térmico por frio, a não ser pela sua proximidade aos genes *lon*, *clpP* e *clpX*.

Nenhum dos novos genes identificados no pAB441 está diretamente envolvido com o metabolismo do nitrogênio. O gene ou genes responsáveis pelo efeito de eliminação da atividade de nitrogenase observado por Vitorino e colaboradores (2001) nos mutantes HM053, HM26 e HM210 aparentemente não pertencem ao conjunto de genes identificados até o momento no fragmento seqüenciado.

SEQÜÊNCIA DE NUCLEOTÍDEOS OBTIDA E SEUS PRODUTOS DE TRADUÇÃO

1 GGCGG GATGA TGATC AATTA CCGTT TCAAG CCGGG GGAAA TCGAC GCCAA
 51 CCACG AAGCG TACGG GGGGA GGGCA AGGTC AACGC CTCCT CCTCC ATCCG
 101 GTCGC TGGGG AAGGG GGGCT GGATG AGCCG GTTTG CCGCC CGCCG GGGCG
 151 TCGTC GTCGT CGCGT CCGGC GCCCC CCACC AATGG CGTGC GGAGC GCGCG
 201 CCATT TTGCC TGTTT GCGGC GCGAA CCCTG TGCTA TAGAG CGGAC TTCGC
 251 CGGCA GTCCG TCGGG TGTTT CGCAC CCGTC GCGGG CTGTC TATGC CATTG
 301 CGGGC GTGGC GGAAC TGGTA GACGC GCCCG ATTTA GGTTT CCGTG ACTTC
 351 GGTCG TGGGG GTTCG AGTCC CTTCG CCCGC ACCAT GGCAA CAGCT TGGAC
 401 CGCAT GGGAA CACAC GGTTC CCGGC CGCCC GGACG GCGCG CGCCG GGCCG
 451 GACAA CGAGA **TAGGC** TTTTCG TTCCG Gene *tig*

476 **atg**aacatcaccgagaccagcgcgacggcctgaagcgcgagtac
 M N I T E T S A D G L K R E Y
 521 aaggtcgtcatttccgcccaggatatcgaacagaaggtgcagggc
 K V V I S A Q D I E Q K V Q G
 566 cggctcgaagagctgcgcccggacgggtgcagctccccggcttccgt
 R L E E L R R T V Q L P G F R
 611 cggggcaaggtgccggttgccgtgatcaagcagcgcctacggcggc
 P G K V P V A V I K Q R Y G G
 656 agcgtgctggcccaggctctggaggatgcatcgccgacagctcg
 S V L A E A L E D A I A D S S
 701 cgccaggcgcctgaacgagcgcggcctgcatcgccatgcagccg
 R Q A L N E R G L R I A M Q P
 746 aagatcaacgtcgagaagtagcaggacggcggcgacctgtcctac
 K I N V E K Y E D G G D L S Y
 791 accatgggctcgagcttctgcccggacatcgagccgggctgatctc
 T M G V E L L P D I E P G D L
 836 agcggcctggagcttgagaagccggctgccacggctcgaggactcg
 S G L E L E K P V A T V E D S
 881 gccgtcgacgaggcgcctgaccgctctggcctccgcccattcggtg
 A V D E A L T R L A S A H S V
 926 caggccccggtcaccgaggaccgcgcgscgagaagggcgacatc
 Q A P V T E D R A X E K G D I
 971 gccgtgatcgattcgccggttcggctcgactggcgaggccctgccg
 A V I D S P V R S T G E A L P
 1016 ggcattggacggcaaggactaccgctggagctgggcccgaaccag
 G M D G K D Y P L E L G A N Q

1061 ttcgttccgggcttcgaggaccagctgggtcggcgccaaggccggc
 F V P G F E D Q L V G A K A G
 1106 gagcaccgcacggtgaagggtcaccttccccggccgactaccgcac
 E H R T V K V T F P A D Y P H
 1151 gaccgcctgaagggcgcgacaccgtggttcgaggtggacgtcaag
 D R L K G A D T V F E V D V K
 1196 gaactccgcaagaacgtccccggctgaggtcaacgacgaccttgcc
 E L R K N V P A E V N D D L A
 1241 aaggagttcggcatggagagcctggagaagctgcgcgaggccgctc
 K E F G M E S L E K L R E A V
 1286 ggcgaccgcatcaagggcgaatacgccaacgtgtcgcgacctgcg
 G D R I K G E Y A N V S R L R
 1331 gtcaagcgccagctcctcgataagctggccgaggcgcactccttc
 V K R Q L L D K L A E A H S F
 1376 gaagtccccgcccggcatggtggacgtcgagttcgagggcatctgg
 E V P P G M V D V E F E G I W
 1421 cagcgccttcagcaggagctccagaacggcaccgcccgggtgaggac
 Q R L Q Q E L Q N G T A G E D
 1466 gccggcaagccggaagaggagctgaagaccgagtaccgcggcatc
 A G K P E E E L K T E Y R G I
 1511 gccgagcgcgctgctgctgctgctgctgctgctgctgctgctgctg
 A E R R V R L G L L L S E I G
 1556 cgccgcaacgacatccaggtgaccaggacgagatcaaccgcgcc
 R R N D I Q V T Q D E I N R A
 1601 ctgatcgcggaagcccgcgcttccccggccaggagcgtcaggtg
 L I A E A R R F P G Q E R O V
 Sítio de ligação para σ^{32}
 1646 ttcgagttcttcaagcagaaccgggaagccctggaaaacctccgc
 F E F F K Q N R E A L E N L R
 1691 gccccgatcttcgaggacaagggtcgtcgactacatcctggacc
 A P I F E D K V V D Y I L D Q
 1736 gccaaaggtgagcgagaagccgggtgtccgccgaagagctgatgaag
 A K V S E K P V S A E E L M K
 1781 gaccccgacgaggaggccgaggcggcctga 1810
 D P D E E A E A A *

CGTC AGAAA GATTG AGGGG CGGCC TACCG GTCGC CCCTT
 1850 TCCAT TCCTA TATGT AGACT GCTCA GACGA CTAAC CGAGC TCGCG GAGCA
 1900 **GGGAA GACCA** CC Gene *clpP*

1913 atgtacgacttogaaccgaagatgaatgctctggtcccgatggtc
 M Y D F E P K M N A L V P M V
 1958 atcgagcagaccaaccggggcgaacgaggctacgatatttactcg
 I E Q T N R G E R G Y D I Y S
 2003 cgcctgctgaaggaacggatcatcttcttgatcgggtggcgtgaac
 R L L K E R I I F L I G G V N
 2048 gacgccgtcgccagcctgatctgctcgcagctgctgcttccctgaa

D A V A S L I C S Q L L F L E
 2093 tcggaaaacccgagcaaggacatcgcgctctacatcaattcgccg
 S E N P S K D I A L Y I N S P
 2138 ggcggatacgtctcggccggcctcgccatctacgacaccatgcag
 G G Y V S A G L A I Y D T M Q
 2183 tacatccgtccgcaggtgtcgcggtgtgcatggggcagggcggcc
 Y I R P Q V S T V C M G Q A A
 2228 tcgatgggctcgctgctgctcgccgcccgggtgctccgggcaagcgc
 S M G S L L L A A G A P G K R
 2273 ttctcgctgccgaacagccggatcatgatccaccagccgtcggggc
 F S L P N S R I M I H Q P S G
 2318 ggcgcccagggtcaggcctcggacatcgagatccaggctcaggag
 G A Q G Q A S D I E I Q A Q E
 2362 atcctgaagctgcgttcgcgcctcaacgacatctacgtcaagcac
 I L K L R S R L N D I Y V K H
 2408 accggccagtcgctcgacacgatcgaagcgnccatggagcgtgac
 T G Q S L D T I E A X M E R D
 2453 aagttcatgtcgcgagggagggccaaggccttcgggctgatcgac
 K F M S P E E A K A F G L I D
 2498 gaggtgggtggagaagcgcggcgggttcgatcggcgacggcgccgccc
 E V V E K R P G S I G D G A A

2543 tga

*

GGG

2549 GCGGC GCGAA CTAAT GATTG ATTGC GCCGG TTCCC TGTTG TTGAA TGATG
 2599 GGGAA CCGGT CCGAT GGATC GGGTT CGAAG GATGA GCGGA TCGGT TAACC
 2601 GAACC GGTCC GATGG ATCGG GTTCG AAGGA TGAGC GGATG CGTTA ACCTT
 2651 GTCAC GGGCC ATGCC CGGGA CAACC TTTGA AAGGA CCGGC GTCGG GGGAC
 2701 GCCGC AGGCG GTTGA CAGGT TATAC GGATC AGTCC CCCGC GAGAT GACGG
 2751 AGTAC CG

Gene *clpX*

2758 atgagcaagtccagcagcggcgattcgaagaacacgctctactgc
 M S K S S S G D S K N T L Y C
 2803 tccttttgcggaagagccagcacgaggtccgcaagctgattgcc
 S F C G K S Q H E V R K L I A
 2848 ggtccgacggtgttcatctgcgatgaatgcgctcgagctgtgcatg
 G P T V F I C D E C V E L C M
 2893 gacatcattcgcgaggagaacaagacgaccctcgtgaagtcggcgc
 D I I R E E N K T T L V K S R
 2938 gacggcgtgccgaccccgcgacatccatgcggtactggacgat
 D G V P T P R D I H A V L D D
 2983 tacgtgatcggccagcatcacgcgaaacgggtcctttctgttgcg
 Y V I G Q H H A K R V L S V A

3028 gtgcacaaccactacaagcggttggcgcacggcacgaagcacaac
 V H N H Y K R L A H G T K H N
 3073 gacgtcgaactggcgaagtccaacatcctgctggtcggccccacg
 D V E L A K S N I L L V G P T
 3118 ggctgcggcaagacgctgctcgcccagacgctcgcccggatcatc
 G C G K T L L A Q T L A R I I
 3163 gacgttcccttcacgatggccgacgccacgaccctgaccgaggcg
 D V P F T M A D A T T L T E A
 3208 ggctatgtccgcgaggacgctcgagaacatcatcctcaagctgctc
 G Y V R E D V E N I I L K L L
 3253 caggccgcgcgactacaacgacgagcggcgcagcgcggcatcgtc
 Q A A R L Q R R A A Q R G I V
 3298 tacatcgacgaggctcgacaagatcagccgcaagtccgacaacccg
 Y I D E V D K I S R K S D N P
 3343 tcgatcacgcgcgacgtgctcggggcagggcgctgcagcaggccctg
 S I T R D V S G E G V Q Q A L
 3388 ctgaagatcatggagggcaccgtcgcctccgtgcctccgcagggg
 L K I M E G T V A S V P P Q G
PstI
 3433 ggccgcaagcaccgcgacgaggagttcctgcaggtcgcacacgagc
 G R K H P Q Q E F L Q V D T S
 3478 aacatcctgttcatctgcggcggcgccttcgccggggtggacaag
 N I L F I C G G A F A G L D K
 3523 atcatcgcgacgcgaggcaagggcacgtcgatcggcttcggcgct
 I I A Q R G K G T S I G F G A
 3568 gatgtgcgcggccccggacgagcgttcgaccggcgacatcctgcgt
 D V R G P D E R S T G D I L R
 3613 gaggtcgagcccgaggatctgctgaaattcggcctgatccccgag
 E V E P E D L L K F G L I P E
 3658 ttcateggctcgtctgccggtggctgccaccctgtccgacctcgac
 F I G R L P V V A T L S D L D
 3703 gaaacggctctggctcgagatcctcaccaagccgaagaatgccctg
 E T A L V E I L T K P K N A L
 3748 gtcaagcagtaccagcgcctgttcgagatggaggacgtccgtctg
 V K Q Y Q R L F E M E D V R L
 3793 gaattctccgacgacgccctgcgccaccatctcccacaaggccatc
 E F S D D A L R T I S H K A I
 3838 cagcgggaagaccggcgcacgcggcctgcggtcgcacatcggagtcg
 Q R K T G A R G L R S I M E S
 3883 atcctgctcgacccgatgttcgatctgccggcctgtcgggggtg
 I L L D P M F D L P G L S G V
 3928 gaaagcatcctgggtcaacaaggaggtggctcgagggggcgcgccaag
 E S I L V N K E V V E G R A K
 3973 ccgctttacgtccacgccgaacgccgcggggagcagcaggcaccg
 P L Y V H A E R R G E Q Q A P
 4018 ggtgcctga
 G A *

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

4051 GGTAC GAGGA Agggc gcctt tgggc gccct ttgtc gtttc cggct tttgc

4101 agACA TCGAG GGCGC CAGCC GGGCC GCGGG GCGGT CGGAT CGGCA TGTCG

4151 CACCG CCATA CGGAA TTggg gttct ggccc tgaaa ggggg ctatg gctgt

Sítio de ligação para $\sigma 32$

4201 gcccc ttaga AAAAG TGCCG GTTTT CGATC CCTGT GCTCA TCCCC CGAGG

4251 CGGGG GTCTT CCCAG GCCCA TCCCC GGCTG ACGGC GTCCC AATGA TGAAA

4301 GAGGC CCAAT CA

Gene *lon*

4312 atgttcgaatccct
M F E I P

4328 cgtggtgccctctatccggtcccgcgctccgcgacatcgtggt
R G A L Y P V P P L R D I V V

4373 ttccccacatgatcgtgcctcttttcgctcggccgtgagaaatcc
F P H M I V P L F V G R E K S

4418 gtgcgcgccctggaagacgtgatgaaggacgataagcagatcctt
V R A L E D V M K D D K Q I L

4459 ctgctcaccagaagaacgccgcgaggacgatccgacgccggcc
L V T Q K N A A Q D D P T P A

4508 gatatctacagcgtcggcaccggtgggaccggtggtgcagctgctg
D I Y S V G T V G T V L Q L L

4549 aagctgcccagcgaacggtgaaggtgctcgtggagggcgccaa
K L P D G T V K V L V E G G Q

4594 cgcgctccatcaccaagttcgccgagaacgaggatttcttccag
R A S I T K F A E N E D F F Q

4639 gccacgccgacctcgtcgaggagaaggtcggggaaagccaggaa
A H A D L V E E K V G E S Q E

4684 cttgagggcgtgggacgcgcggtcgtctcgcaattcgagcagtac
L E A L G R A V V S Q F E Q Y

4729 atcaagctgaacaagaagatcccgcggaggtcctggtctcgatc
I K L N K K I P P E V L V S I

4774 aaccagatcgaagagccgggaaagctggcggacaccgctcgcctcc
N Q I E E P G K L A D T V A S

4819 cacctcgcgctgaagattccggaaaagcagcagcttctggaatgc
H L A L K I P E K Q Q L L E C

4864 gccacggtttcggagcgggttgagcgggtctacgccttcatggaa
A T V S E R L E R V Y A F M E

4909 ggcgaaatcgggtgtcctccaagtggaaaagcgcacccgcaaccgc
G E I G V L Q V E K R I R N R

4954 gtcaagcggcagatggagaagaccagcgcgagtagtactacctcaac
V K R Q M E K T Q R E Y Y L N

4999 gaacagctcaaggcgatccagaaggaactcggcgagaccgaggac
E Q L K A I Q K E L G E T E D

5044 ggccgtgacgagtcggccgagctggaagagaagatcaacaagacc
G R D E S A E L E E K I N K T
5089 cgcttctccaaggaagcccgcgacaaggccctggccgagctgaag
R F S K E A R D K A L A E L K
5134 aagctgcgctccatgagcccgatgtcggccgaggcgacggtggtg
K L R S M S P M S A E A T V V
5179 cgcaactatctggactggatgctgtccattccgtggaagaagcgc
R N Y L D W M L S I P W K K R
5224 accaaggtgaagaaggacctgaagctcgcgcagaagatcctcgac
T K V K K D L K L A Q K I L D
5269 gccgaccactacggcctggagaagggtcaaggaacgcatcctcgag
A D H Y G L E K V K E R I L E
5314 tatcttcgggtccagaaccgcatgaacaagggtcaaggggcccgatc
Y L R V Q N R M N K V K G P I
5359 cwgwgccctcgtcgggcccggcgcttggaagacctcgctcggc
X X L V G P P G V G K T S L G
5404 aagtcgatcgccaagtccaccggccgcaatttcgtgcgcatgtcg
K S I A K S T G R N F V R M S
5449 ctcggcggcgtgcgggacgaggctgagggtccgcggccaccgacgc
L G G V R D E A E V R G H R R
5494 acctacatcggctcgatgcccggcaagggtcatccagggcatgaag
T Y I G S M P G K V I Q G M K
5539 aaggcgaagtcgtccaaccgctgttcctgctggatgagatcgac
K A K S S N P L F L L D E I D
5584 aagctcggcgcgactggcgcggcgaccctcgtcggccctgctt
K L G A D W R G D P S S A L L
5629 gaggttctcgatcccgagcagaacggcactttcaacgaccattat
E V L D P E Q N G T F N D H Y
5674 ctggagggtcgattacgacctgtcggacgtgatgttcgtctgcacg
L E V D Y D L S D V M F V C T
5719 gccaacacgatgcgcatgccgcagccgctgctggaccgcatggag
A N T M R M P Q P L L D R M E
5764 atcatccgcgtcgcggctataccgaggatgaaaagggtcgagatt
I I R V A G Y T E D E K V E I
5809 tccaagcgcacctgatcgaaaagcagggtggaagccaacggcttg
S K R H L I E K Q V E A N G L
5854 aagaagggcgagttcgccatttccgacgacgcgctgcgcgacctg
K K G E F A I S D D A L R D L
5899 atccgctattacacgcgggaagccggcgtccgcagcctggagcgc
I R Y Y T R E A G V R S L E R
5944 gagatcgccaacctctgccgcaaagccgtgaaggagatcctgatg
E I A N L C R K A V K E I L M
5989 aagggctccgccggcgccaaggtttcggtcacgcgccggaacctg
K G S A G A K V S V T R R N L
6034 gacaagtacgctggcgtccgccgcttccacttcggcgaggcggag
D K Y A G V R R F H F G E A E
6079 ctggaggatctggtgggcgtcaccaccgggctggcctggacggaa

L E D L V G V T T G L A W T E
 6124 gtcggaggcgagctgctgtcgatcgaggcggtcagcctgcccggc
 V G G E L L S I E A V S L P G
 6169 aaggggcgggtgaccaccaccggcaagctgggacgacgtcatgaag
 K G R V T T T G K L G D V M K
 6214 gagtcgggtccaggcggcggagagctacgtcaagtcgcgggccacc
 E S V Q A A E S Y V K S R A T
 6259 gccttcggcatcaagccgacgctcttcgagaagcgggacatccac
 A F G I K P T L F E K R D I H
 6304 gtccacgtgcccgaaggcggcactccgaaggacggcccctcggcg
 V H V P E G A T P K D G P S A
 6349 ggcgtggcgatgatcacctcgatcgtttcggctcctgaccggtatc
 G V A M I T S I V S V L T G I
 6394 gcggtccgcaaggacgtggcgatgaccggtgagatcacctgcg
 A V R K D V A M T G E I T L R
 6439 ggccgggtgcttccgatcggcgggtctgaaggagaagctgctggct
 G R V L P I G G L K E K L L A
 6484 gctctgcgcgggcggcctcaagcatgtgctgatcccgaaggacaac
 A L R G G L K H V L I P K D N
 6529 gagaaggatctcgccgagatcccggacaacgtgaagcgcgggctg
 E K D L A E I P D N V K R G L
 6574 gagatcatcccagtcagcaccgtcgacgacgtcctgaagcacgcc
 E I I P V S T V D D V L K H A
 6619 ctggttcgggaagtcgagcccatcgagtggaaggagcccggaagcg
 L V R E V E P I E W K E P E A
 6664 gtcgagccggcggctcgccaagccgcagactgacggcggaggcgag
 V E P A V A K P Q T D G G G E
 6709 gtgcttcgtcactga 6723
 V L R H *

6724 TcCAG CTtGG CGTTT AaAGC ACCCt TAAAG CqcTA CCATG cTTCG CATCG
 Sítio de ligação para σ^{70}
 6774 CCCGC CCGTC TGCAA TTGAC GGgCG GGCGA TGTCG CGTTT ATCGT TTGAA
 6824 TGCGT GATGT CATTG AGTGC CTTTT CTTC CTTTT CTTTG CCGGA CTGAG
 6874 AAGGG **GGGAA** GTAA

6888 **GTG** AAC AAG AAC GAT CTG GTG GCG CAC GTC GCC GAC GCC GTG GGT
 1 **V** N K N D L V A H V A D A V G
 6933 TTG TCC AAA ACC GAC GCG ACC AAG GCG GTC GAC GCA ATT TTC GAC
 16 L S K T D A T K A V D A I F D
 6978 GGC ATC GCC GAC TCG CTC AAG AAC GGG GAG GAG GTC CGT CTG GTC
 31 G I A D S L K N G E E V R L V


```

7023 GGA TTC GGC ACA TTC GCG GTT TCC GAA CGC GCG GCC AGC GAG GGA
46  G  F  G  T  F  A  V  S  E  R  A  A  S  E  G

7068 CGC AAC CCG CGC ACC GGC GAA AAG ATC GCC ATA CCC GCG TCG AAA
61  R  N  P  R  T  G  E  K  I  A  I  P  A  S  K

7113 CAG CCG AAG TTC AAG CCG GGC AAG ACC TTG AAG GAC GGG TTG AAG
76  Q  P  K  F  K  P  G  K  T  L  K  D  G  L  K

7158 TGA
91  *

          TGA CGTCT TGAAC GGATA GGGCA TTCCG GGCAT AGTGC
7199 GCGCC GGTCA GCACG GGCTG GCCGG CTTTG CGTTG TCTGG AACGT CAATC
7249 CGGTC AACGA CGCGG CGCGG GCGAT TAGCT CAGCG GTAGA GCATC TCGTT
7299 TACAC CGAGA GGGTC GGGGG TTCGA AACCC TCATC GCCCA CCATC GCGCC
7349 CTGGC CATTG CGGCC GGCAC TTTTC AAATT CGTGA TTGAC AGGCC ATGCG
7399 TCGGG CGGCT AGAAA CCGCC CATCG GCGAC GGGGC GTGGC CCCGG AAACT
7449 GCGCT GCGAT CCTAT GACAC GGACG CCTGT TTTTG CGCAT TTCAC ATCGC
7499 CGTGT CGCGC GGGTG TAGCT CAGTT GGTTA GAGCG CCGGC CTGTC ACGCC
7549 GGAGG CCGCG GGTTG AAGTC CCGTC ACTCG CGCCA TTCTC TCCCT CACCA
7599 TTCTC CTTTT TGCAA TGAAA ATGAA GCCTT CATGG CTTTA CCGGC AGGCA
7649 CGCGA TCCCC ATATC CGATT AGCTC GACCC GCATA TCCGA TAAGC CGGAT
7699 AGAAG AAGGG TCATT TAAAA AGTTT GGTA TACAA GGACA TCCGT CATTG
7749 CGAAA GTGGA TTCnT TGGCC GAAAG CGATT CAAGG CTTTC CTGAT CAGCT
7799 CTGTC ATATT TCCGG CCCGC TGATG GGTA TCAAG GATTC AGCAA AAACC
7849 GnGCG TC

```

Figura 5: O produto da tradução derivado das seqüências de nucleotídeos (exceto *hupB*) foi obtida pelo programa ORF finder. A seqüência de aminoácidos do gene *hupB* foi obtida através do programa BioEdit. Os prováveis sítios de ligação para ribossomo estão indicados em negrito. Os códons de iniciação são indicados em azul e os códons de terminação em vermelho. As seqüências consenso para ligação de promotores σ^{32} e σ^{70} encontram-se sublinhadas e indicadas. Prováveis terminadores independentes de Rho estão indicados por setas invertidas e estão representados em verde. As bases em letras minúsculas nas regiões intergênicas são de baixa qualidade. O sítio *PstI* não mapeado no trabalho de Bozouklian e colaboradores (1986) encontra-se indicado em vermelho. O sítio de início de transcrição do gene *lon*, encontrado por Mori e colaboradores (1996) está indicado por uma seta vertical.

ALINHAMENTO DA SEQÜÊNCIA DE AMINOÁCDOS DA PROTEINA TF DE A.

*brasile*se COM DIVERSOS ORGANISMOS

```

Azbras -----MNITETSADGLKREYKVVISAQDIEQKVQGRLEELRRTVQLPGF 44
Rorubr MAGFRLSTRKVVREQPMQVTETVNESLKREYKIVI PAADIAERSARRLEELKGQMLPGF 60
Brmeli -----MTRSEGLNMQVTETLNEGLKREIKVVVPAGDLEAKLAERLETARGRARINGF 52
Meloti -----MQVTETLNSGLKREIKITVPAGDMEAKLMARLSDARNKVRINGF 44
Simeli -----MQVIETLAQGLKRELKVVIPADEM QARMNERLVEVKDRVRINGF 44
Rhcaps -----MQVTQTLNEGLKRGYTTITLTGAELDAKVTEKLI EVQPEVEIKGF 44
Rasola -----MSTVIENLGLDRKVT LAVPKAEVEQEQSRLARLSKTVKMSGF 44
                * .          * . * . . : : . : *          . : **

Azbras  RPGKVPVAVIKQRYGGSVLAEALEDIAIDSSRQALNERGLRIAMQPKIN-VEKYEDGG-- 101
Rorubr  RPGKVPMSMLRQRFGKSVLGEVLEKAVQESIREVMTSHELK PATQPDIDLISEVEEGK-- 118
Brmeli  RPGKVPPTAHLRKMVGKSFMAEIVNEILNDSSRSILAERNEKSATQPEVIMSEDEKEAEKV 112
Meloti  RPGKVPVQHLRKVYGKSFMAEIVNEILNDSTRSIIITGRGEKAAMQPEVIMTEDEKEAEKI 104
Simeli  RPGKVPVAHLKKVYGKSIMADLVNEIVREKPT EILTSRGEKSATQPEIAMTEDEAEADKI 104
Rhcaps  RKGKVPMAMLRKNFNGDRVLGDV LNESVDGAIKDLLDQSGDRPALQPKVEMENKGDWKPG- 103
Rasola  RPGKVPMKMVEKQYGOQVEMEVRFDKAARQFFDITNEQGVKVAGQPRFELKSEGVADD-- 102
                * * * * * : : : * . : . . : * * * . .

Azbras  -----DLSYTMGVELL PDI EPGLDLSGLELEKPVATVEDSAVDEALTRLASAHSVQAPVTE 156
Rorubr  -----DVEFTLAL EVLPEI GETDF SALALEREVAEVA AEKIEEALET LRQQSKTHEPVT D 173
Brmeli  LDGKADFV FSLN YEVLP AIEVKDFSKIAV TREVV DISDEEVDEQVKRIASSTRTFETKKG 172
Meloti  LAGGTD FEFRLN YEVIPA IEIKDFSDIKVTRQVFDV PDAEIDDQVKRVAESARSYEPKTG 164
Simeli  LNAQADFEFTVAYEII PPIELKDASGIKVTR EVDVSEDEVNEQIQRIAESARTYESKKG 164
Rhcaps  ----TDAIFTVS YEALPPIPAFDRATV TLERLVVKADEASVTEALENLA KSAQAFEDRKK 159
Rasola  -----QFAFDATFEVYPEVKIGDLAGAELTRTKTEITDAEIDKTIDILRKQRVHYHPRGE 157
                : : * * : * : : : : : . : : .

Azbras  D-----RAAEKGDI AVIDSPVRSTGEALPGMDGKDYPLEL GANQFVPGFEDQLVGAK 208
Rorubr  G-----RAAAGGDLVVIDF IGKLDGEAFEGGSAEAYDLELGSNSFIPGFEDQLIGAT 225
Brmeli  -----KAENEDRVTIDYLGKLDGEPFEGGADNDAQLV LGSQFIPGFEEQLIGLK 222
Meloti  -----KAAEGDRVSIDYV GKIDGEAFAGGAGTDQLLV LGSKEFIPGFEDQLIGTK 214
Simeli  -----KAANGDRVTVDYLGKVDGEAFDGGKDEDAELV LGSNRFIPGFEEQLVGVK 214
Rhcaps  G-----TKAKDGDQVVIDFEGFLGDEPFEGGKGEEYPLV LGSNSFIPGFEDQLVGAK 211
Rasola  AGDHGDGGEAVAQNGDRVTVD FVGTIDGVEFAGGKAEGFGFV LGEGRMLPEFEQATLGLK 217
                * * . : *          . : *          : **          : * * : * .

Azbras  AGEHRTVKVTFPADYPHDR LK GADTVFEVDVKELRKNVPAEVNDDLAK EFGMES--LEKL 266
Rorubr  AGEARTVTVSFPEDYPA AHLGKETVFDVTVKEVKQAVVPELDDDLAKAFGKES--AEAL 283
Brmeli  AGDEKVITVTFPAEYGA AHLGKEATFDIKVKEVAKPNELVLDDETA KKLGI ES--LERL 280
Meloti  AGDEKQVTVTFPENYQA AHLGKEATFDVTVKEVSKPGELEINDETA KNLGLES--LERL 272
Simeli  AGDEKTITVTFPADYPA ANLAGKEATFDITVKDVAAAAP IEINDELATK LGL ES--VDKL 272
Rhcaps  AGEDVEVKVTFPAEYGA AHLGKEATFKCHVHAVKAPKPAEIDDELAKKFGAAD--LEAL 269
Rasola  QGESKVFPLAFPADYHG KD VAGKTA EFTVTLKQVEWAHLPEINGAF AQLSLGIADGSLDKM 277
                . : * * : *          : *          : *          : :          : *          : *          : *          : :

```

```

Azbras REAVGRPHQGRIRQRVALRVKRQLLDKLAEAHSFEVPPGMVDVEFEGIWQRLQQELQNGT 326
Rorubr REAVKADLQGELDEVSKTKLKRKLLDALADGHDFPVPQTLVDAEFEGIWAQIEKAKTDGQ 343
Brmeli RQVVREQIESQYGQITRQKVQRQLLDALDGDYQFETPQKLVDAEFNNIWAQINFDLQQAG 340
Meloti REIVRGQIENQFGSMTRQKVQRQLLDQLDAAYSFEAPSCLKLIDAEFNNIWAQVTRDLEAAG 332
Simeli KEIVRGQIESQFGSITRQKVQRQLLDQLDELYQFDTPERLVDAAEFENIWRQINTDLQQAG 332
Rhcaps KGQVASRLEAEYKASRAILKRALLDILDAQVKFDLPPSLVEAEAGQIAHQWHEENPDH 329
Rasola RADIRENLEREVKRRTHALLKDQVMDALLKVAELDVPKSLIEQDQERLVEARRDLEARG 337
: : : . : * : * * . : * : : : : :

Azbras AG-EDAGKPEEELKTEYRGIARRVRLGLLLSEIGRRNDIQVTQDEINRALIAEARRFPG 385
Rorubr LDEEDAAKSDEDLRADYRKIAARRVRLGLLLADVGQRAQVTVQEDLNKALMRELRRFPG 403
Brmeli RTFEDEETTEEAAREEYRKLAAARRVRLGLVLAEIGEKAGVEVTEEELQRAVYDQVRRYPG 400
Meloti RTFADEETTEEEARAAYRLAARRVRLGLVLAEIGEKAGVTVSDEELQRGLFEQVRRFPA 392
Simeli KTFADEDTTEEEARAAYRKLAAARRVRLGLVLAEIGEKAGVQVSDDQMQRSLFEQLRQFPG 392
Rhcaps HG---HDHGAVEPTTEHKTLAARRVRLGLLLAELGRNEKIEVTDAMTQAVLAAARQYPG 386
Rasola MP---NVKNMPIPAEMFAQQAERRVKLGLVLAEVVKVNALEAKPEQIKAEIEEFAKSYED 394
. *****:***:***: . : . : : : :

Azbras -QERQVFEFFKQNREALNLRAPIFEDKVVDYILDQAKVSEKPVSAEELMKDPDEEEAEA 444
Rorubr -QEAAVINYYRNNQQAMDNLRAPVFEDKVCALALATVTDKPVSVEDLMKDPDEDATPT 462
Brmeli -QEKEIYDFLRRTPDVANLRAPIFEEKVVDHLLANINVTDKKVSKEELTAEDEDAAS-- 457
Meloti NQQQEAFAFEFYRSNPEALNLRAPMFEEKVVDHLLGQISVTDVKVSKEELMADDEDAETTT 452
Simeli -QEKEILEYFRNTPGAAASLRAPLFEEKVVDHLLTEVSVTDKKVSKEELTADDEGDEKPA 451
Rhcaps -HEREFFEFVKGNAQMQQIRAPLYEEKVIDFIVAGAAVTEKEVSKEELQKAI EALDEL- 444
Rasola --PKEVIRWYYGDQQRLEAMEAYVLENNVVNFVCDKAKVADKAMSFEELTATQGAQG--- 449
: : . : * : ***: . : * : : : * * : *

Azbras -----
Rorubr A----- 463
Brmeli EAKPAKAAAAKKKAAPKKK--AEEGKSEEA----- 485
Meloti KAKPAKAAAAKKAEAKANEDAEAEPPKKAAPKKAADAE 492
Simeli KKTASKKAAAACAAGEGEEAAAPKKAAPKKAADSAE 491
Rhcaps -----
Rasola -----

```

Figura 6: 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Rhodospirillum rubrum* (Rorubr), *Brucella melitensis* (Brmeli), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), *Rhodobacter capsulatus* (Rhcaps) e *Ralstonia solacacearum* (Rasola).

ALINHAMENTO DA SEQÜÊNCIA DE AMINOÁCIDOS DA PROTEÍNA

ClpP DE *A. brasilense* COM DIVERSOS ORGANISMOS

```

Azbras -----MYD-FEPKMNALVPMVIEQTNRGERGYDIYSRLLKERIIFLIGGVND 46
Meloti -----MKNPVETYMN-LVPMVVEQTNRGERAYDIFSRLKERRIFITGPVED 46
Brmeli -----MRDPIETVMN-LVPMVVEQTNRGERAYDIFSRLKERRIFVNGPVED 46
Rorubr -----MIDRDPIDVFNNTLVPMVVEQTNRGERSYDIYSRLLKERIIFLTGQVHD 49
Rasola MIRNELIDQLARTQASALETQGLGLVPMVVEQSGRGERAYDIYSRLLKERVIFMVGEVND 60
Simeli -----MRNDDDQE-----EKKTELPLGKETEANLFKRSIFIYGTITQ 38
          .   :                               *: . . . : : .*:*. * **: * : :

Azbras AVASLICSQLLFLESENPSKDIALYINSPGGYVSAGLAIYDTMQYIRPQVSTVCMGQAAS 106
Meloti GMATLVCAQLLFLEAENPKKEINLYINSPGGVVTSGMAIYDTMQFIKPAVSTLCIGQAAS 106
Brmeli GMSMLVCAQLLFLEAENPKKEINMYINSPGGVVTSGMAIYDTMQFIRPPVSTLCMGQAAS 106
Rorubr GVASLICAQLLFLESENPSKDISFYINSPGGVVTSGMAIYDTMQYIRSPVSTVCIGQAAS 109
Rasola QTANLVVAQLLFLESENPKDISLYINSPGGSVSAGLAMAYDTMQFVKPDVSTLCMGMAAS 120
Simeli ELAQKVCSQLVALAAAS-DDDIRLFVNSPGGHVESGDSIHDMIKFVKPKVWTIGTGWVAS 97
          :   : **: * : . . :* ::***** * :* :*: : : : . * * : * .**

Azbras MGSLLLAAGAPGKRFLPNSRIMIHPQSGGAQGGASDIEIQAEIILKLRSLNDIYVKHT 166
Meloti MGSLLLTAGHKDMRFATPNARIMVHQPSGGFQGGASDIERHAMDIVKLRRLNEVYVKHT 166
Brmeli MGSLLLTAGATGHRYALLNARIMVHQPSGGFQGGASDIERHAQDIKMKRRLNEIYVKHT 166
Rorubr MGSLLLCAGEAGKRYATPNARIMIHPQSGGFQGGAAADIEIQAREILALRERLNRIYVKHT 169
Rasola MGAFLLAAGAKGKRYALPNSRIMIHPQPLGGARGQASDIEIQAREIILYLRERLNTILSEVT 180
Simeli AGALIYVAAPKEQRLCLPNTRFLLHQPSGGTRGMASDIEIQAREI IKMNERLNRIIFSEAT 157
          *::: * . * . *:::*** ** :* *::*: * :*: :. *** : : *

Azbras GQSLDTIEAXMERDKFMSPEEAKAFGLIDEVVEKRPGSIGDGAA--- 210
Meloti GKSYYEIERTLDRDHFMATADEAKDFGLIDKVISSREPAESAVA---- 209
Brmeli GRDYDTIERTLDRDHFMATAQEAEFGLIDKVVEARDVSADESK---- 209
Rorubr GQPLETIERAMDRDNMYTAEESRAFGLTDSVIERRALSDDGDTKSSS 216
Rasola GQPVDKIARDTDRDNFMMSGDQAKEYGLIDKVLARRGA----- 217
Simeli GQPVDKIAKDTDRDYWLGAEAEAKAYGLVSRIVTSIAEI----- 195
          *: : * **: : : : :* . : :

```

Figura 7: 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Rhodospirillum rubrum* (Rorubr), *Brucella melitensis* (Brmeli), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), e *Ralstonia solanacearum* (Rasola).

ALINHAMENTO DA SEQÜÊNCIA DE AMINOÁCIDOS DA PROTEÍNA

ClpX DE *A. brasilense* COM DIVERSOS ORGANISMOS

```

Azbras MSKS---SSGDSKNTLYCSFCGKSQHEVRKLIAGPTVVICDECVELCMDIIREENKTTLV 57
Meloti MSKV-GNNSGDSKNTLYCSFCGKSQHEVRKLIAGPTVVICDECVELCMDIIREENKTSMV 59
Simeli MSKVSNGSNGGDSKNTLYCSFCGKSQHEVRKLIAGPTVVICDECVELCMDIIREENKTSMV 60
Rorubr MSKS---SSGDSKNTLYCSFCGKSQHEVRKLIAGPTVVICDECVELCMDIIREEHKTSLV 57
Mamagn MTKS---TSGESKNTLYCSFCGKSQHEVRKLIAGPTVVICDECVELCMDIIREEHKTHLV 57
*:* .:*:*****:***:*

Azbras KSRDGVPTPRDIHAVLDDYVIGQHHAKRVLSVAVHNHYKRLAHGTKHN-DVELAKSNILL 116
Meloti KSRDGVPTPQEILKVLDDYVIGQPYAKRVLSVAVHNHYKRLAHAGKNN-DVELAKSNILL 118
Simeli KSRDGVPTPQEILKVLDEYVIGQQQAKRILSVAVHNHYKRLAHAAKSS-DVELAKSNIML 119
Rorubr KSRDGVPSPKDILKVLDDYVIGQQHAKKVLVAVHNHYKRLSAGKNS-DVELAKSNILL 116
Mamagn RSRDGVPTPKDQCQLDDYVIGQGHAKKVLVAVHNHYKRLQHGGKNNNEVELAKSNILL 117
:***:***:~::~* ~***:***** ~*~:***** ~. * ~:*****:~

Azbras VGPTGCGKTLLAQTLARIIDVPFTMADATTLTEAGYVREDVENIILKLLQAARLQRRAAQ 176
Meloti IGPTGCGKTLLAQTLARIIDVPFTMADATTLTEAGYVGEDVENIILKLLQSADYNVERAQ 178
Simeli VGPTGCGKTYLAQTLARIIDVPFTMADATTLTEAGYVGEDVENIILKLLQAADYNVERAQ 179
Rorubr IGPTGCGKTLLAQTLARILDVPFTMADATTLTEAGYVGEDVENIILKLLQASDYNVERAQ 176
Mamagn IGPTGCGKTLLAQTLARILDVPFTMADATTLTEAGYVGEDVENIILKLLQAAEYNVERAQ 177
:***** ~*****:***** ~***** ~*~:*****:~ . * ~:*****:~

Azbras RGIVYIDEVDKISRKSDNPSITRDVSGEGVQQALLKIMEGTVASVPPQGGRKHPQQEFLQ 236
Meloti RGIVYIDEIDKISRKSDNPSITRDVSGEGVQQALLKIMEGTVASVPPQGGRKHPQQEFLQ 238
Simeli RGIVYIDEVDKISRKSDNPSITRDVSGEGVQQALLKIMEGTVASVPPQGGRKHPQQEFLQ 239
Rorubr RGIVYIDEIDKISRKSDNPSITRDVSGEGVQQALLKIMEGTVASVPPQGGRKHPQQEFLQ 236
Mamagn RGIVYIDEVDKISRKSDNPSITRDVSGEGVQQALLKIMEGTVASVPPQGGRKHPQQEFLQ 237
*****:*****:*****:*****:*****:*****:*****:*****:*****

Azbras VDTSNILFICGGAFAGLDKIIAQRGKGTSIGFGADVRGPDERSTGDILREVEPEDLLKFG 296
Meloti VDTANILFICGGAFAGLDKIIISDRGRKTSIGFGATVASPEDRRTGDI FRQVEPEDLLKFG 298
Simeli VDTTNILFICGGAFAGLDKIIISARGEKTSIGFGATVRAPEDRRVGEVLRLEPEDLVKFG 299
Rorubr VDTTNILFICGGAFAGLDRIIISARGKSSIGFGADVRDPEARRTGEVLRRAVEPEDLLKYG 296
Mamagn VDTTNILFICGGAFSGLDKIIISRRGRGTSIGFGADVRGPDERRTGEILREVEPEDLLKFG 297
***:*****:***:~*~. *~. ~***** * *~: * ~*~:~* ~*****:~*~

Azbras LIPEFIGRLPVVATLSDLDETALVEILTKPKNALVKYQRLFEMEDVRLEFSDDALRTIS 356
Meloti LIPEFVGRPLVPLATLEDLDEPALIQILTEPKNALVKYQRLFEMENVDLTFHENALSAIA 358
Simeli LIPEFIGRLPVVATLLEDLDEDALIQILSEPKNALVKYQRLFEMEDVELNFHEDALREIA 359
Rorubr LIPEFVGRPLVPLATLDDLEDALVDILTTPKNALVKYQRLFEMESTQLSFKEDALRCIA 356
Mamagn LIPEFVGRPLVPLATLEDLDVDALIDILSKPKNALVKYQRLFEMEDTRLSFSDDALKAIA 357
*****:*****:***.*** ~*~:~*~: ***** ~*~. * * ~:~*~ *~

Azbras HKAIQRKTGARGRLRSIMESILLDPMFDLPGLSGVESILVNKEVVEGRAKPLYVHAERGER 416
Meloti KRAIERKTGARGRLRSIMEAILLDTMFELPALEGVREVVISEEVVTGSARPLYIYSEQKEK 418
Simeli RRAIVRKTGARGRLRSIMEKILLDTMFELPTLEGVREVVISDEVVKGTARPLYIYSERSEE 419
Rorubr LKAIARKTGARGRLRSIMEGILLDTMFDLPGMEGVVEVVVNKEVVEGRAKPLFIYSDQRTS 416
Mamagn EKAIARKTGARGRLRSIMETILLDTMFDLPGDAVDEVVINKEVVEGRAKPLYIYSERR-E 416
:~*~ ***** ~*~.*** ~*~:~*~ ~*~.~*~ ~*~:~*~ ~*~:~*~ ~*~:~*~ ~*~:~*~ ~*~:~*~

```

```
Azbras QQAPGA 422
Meloti KGNVSA 424
Simeli KTNVSA 425
Rorubr ADGASA 422
Mamagn DVGSSA 422
      .*
```

Figura 8: 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Rhodospirillum rubrum* (Rorubr), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), *Magnetospirillum magnetotaticum* (Mamagn).

ALINHAMENTO DA SEQUÊNCIA DE AMINOÁCIDOS DA PROTEASE Lon DE

A. *brasiliense* COM DIVERSOS ORGANISMOS

Azbras -----MFEIPRG-----ALYPVPLRDI VVFPHMIVPLFVGREKSVR 37
Rhrubr -----MNAIPT-----QVFPVPLRDI VVFPHMIVPLFVGREKSVR 36
Brmeli MGRKARERTSVMTGIEQKTPVGGSETGGADGLYAVLPLRDI VVFPHMIVPLFVGREKSIR 60
Meloti -----MKGWTMAKISKAP-----SDGVFAVLPLRDI VVFPHMIVPLFVGREKSIK 45
Simeli -----MTNKTSPATESAT-----YPVLPPLRDI VVFPHMIVPLFVGREKSIR 41
Rasola -----MSGTQLLP AEQIR-----LPLLPLRDVVVFPHMVIPLFVGRPKSIK 41
: *****:*****:***** **::

Azbras ALEDVMKDDKQ ILLVTQKNAAQDDPTPADIYSVGT VGTVLQLLKL PDGTVKVLVEGGQRA 97
Rhrubr ALEDVMREDKQ ILLVAQKNATQDDPGDDIYTVGT VSTVLQLLRL PDGTVKVLVEGSHRA 96
Brmeli ALEEVMGVDKQ ILLLATQKNAADDDPAPDAIYE IGTIANVLQLLKL PDGTVKVLVEGTARA 120
Meloti ALEEVMGQEKQ ILLLATQMNAADDDPEPDAIFD IGTLANVLQLLKL PDGTVKVLVEGASRA 105
Simeli ALEEVMGTDKQ IMLVTQINATDDDDPEPSAIYKVGT IANVLQLLKL PDGTVKVLVEGRSRA 101
Rasola ALEAAMEAGKS IMLVAQKTAAKDEPTDKDLYEVGC IANILQMLKL PDGTVKVLVEGTQRA 101
*** . * * . * : * : * . * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

Azbras SITKFAENEDFFQAHADLVEEKVGSQELEALGRAVVSQFEQYIKLNKKIPPEVLVSINQ 157
Rhrubr RIGAYTAREDDFFAEAEATILADDEGDHQEIEALGRSVINQFEQYIKLNKKIPPEVLVSINQ 156
Brmeli KISKFTDREDYHEAYAAALQEPEDAVEIEALARSVVSDFENYVKNKKISPEVVG TASQ 180
Meloti KIVSFTDRPDFHEARATALVEPEEEEEVEVEALARSVVTDFENYVKNKKISPEVVG AASQ 165
Simeli EIERYTPRDDFYEAMAHALPEPEDPVEIEALSRSVVSDFEFESYVKNKKISPEVVG VASQ 161
Rasola NILSVTDDESHFHCEAMP IGPEPTESAETEALRRRAIVSQFDQYVKNKKIPPEILTSLSG 161
* : * : : * *** * : : : * : : * : : * : : * : : * : : *

Azbras IEEPGKLADTVASHLALKIPEKQQLLECATVSERLERLVYAFMEGEIGVLQVEKRIRNRVK 217
Rhrubr IEDAAKLADTVASHLVLKIADKQELLEIETIAERLERVVSFMESEIGVLQVEKKIRNRVK 216
Brmeli IDDYSKLADTVASHLAIKIPEKQEMLSVLSVRRERLEKALSFMEAEISVLQVEKRIRSRVK 240
Meloti IDDYSKLADTVASHLAIKIPEKQEMLATLSVKERLEKAMGFMEAEISVLQVEKRIRSRVK 225
Simeli IEDYSKLADTVASHLSIKIVEKQEMLETTSVKMRLEKALGFMEGEISVLQVEKRIRSRVK 221
Rasola IDEPGRLADTIAAHLPIKLEQKQKILEMFNVTERLESLLSQLEGEIDILQVEKRIRGRVK 221
* : : * : : * : : * : : * : : * : : * : : * : : * : : * : : * : : *

Azbras RQMEKTQREYYLNEQLKAIQKELGETEDGRDESAEELEEKINKTRFSKEARDKALAE LKKL 277
Rhrubr RQMEKTQREYYLNEQLKAIQKELGEGEDGRDELQLEEDRIAETKLTKEAKEKALGELKKL 276
Brmeli RQMEKTQREYYLNEQMKAIQKELGDS EDGRDEVAEIEERITKTKLSKEAREKALAE LKKL 300
Meloti RQMEKTQREYYLNEQMKAIQKELGEGEDGRDEAAEIEARIKKTKLSKEAREKAEAE LKKL 285
Simeli RQMEKTQREYYLNEQMKAIQKELGDS EDGRDEMAELEERISKTKLSKEAREKADAELKKL 281
Rasola RQMEKSQREYYLNEQVKAIQKELGEGEEGAD-LEELD KKIKAARMPKEAKKKADSEFKKL 280
*****:*****:*****: * : * * * : : * : : * : : * : : * : : * : : *

Azbras RMSPM SAEATVVRNYLDWMLSIPWKKR TKVKKDLKLAQKILDADHYGLEKVKERILEYL 337
Rhrubr RNMPMSAEATVVRNYLDWMLAIPWQKR TRVKKDLKIAHNVLNTDHYGLDKVKDRILEYL 336
Brmeli RMSPM SAEATVVRNYLDWLLSIPWGK KSKVKQDLNFAQEVLD AEHFGLGKVKERIVEYL 360
Meloti RTMSPMSAE STVVRNYLDWLLSIPWGK KSKVKQDLAYAQNVLD TDHFGLDKVKERIVEYL 345
Simeli RQMSPMSAEATVVRNYLDWLLGLPWG KSKVIKTDLNHA EKVLDTDFHGLDKVKERIVEYL 341
Rasola KLMPMSAEATVVRNYIDTLVGLPWRK KSKVNNDLSNAEQVLDQDHYGLEKVKERILEYL 340
: *****:*****: * : : * : : * : : * : : * : : * : : * : : * : : *

Azbras RVQNRMNKVKGP IQSLVGP PGVGTSLGKSI AKSTGRNFV RMSLGGVRDEAEVRGHRRTY 397
 Rhruabr AVQLRTQKIKGP ILCLVGP PGVGTSLGKSI ARATGRNFV RVS LGGVRDESE IRGHRRTY 396
 Brmeli AVQARSTKIKGP ILCLVGP PGVGTSLARSIAKATGREYV RMSLGGVRDEAE IRGHRRTY 420
 Meloti AVQSRQKKLKGP ILCLVGP PGVGTSLGKSI AKATGREFIRMALGGVRDEAE IRGHRRTY 405
 Simeli AVQARSSKIKGP ILCLVGP PGVGTSLAKSIAKATGREYIRMALGGVRDEAE IRGHRRTY 401
 Rasola AVQQRVDK LKAPILCLVGP PGVGTSLGQSVARATNRKFV RMALGGVRDEAE IRGHRRTY 400
 ** * * : : . * . * * . * * * * * * * * : : * : * : * : * : * * * * * : * : * * * * *

Azbras IGSM PGKVIQGMK KAKSSNPLFLLDEIDKLGADWRGDPSSALLEVLDP EQNGTFNDHYLE 457
 Rhruabr IGSM PGKIIQGM RKAKSSNPFLLDEIDKLGNDWRGDPSSALLEVLDP EQNNTTFNDHYLE 456
 Brmeli IGSM PGKVIQSMK KAKKSNPLFLLDEIDKMGQDFRGDPSSAMLEVLDPKQ NATFMDHYLE 480
 Meloti IGSM PGKVIQSMK KAKKSNPLFLLDEIDKMGQDFRGDPSSALLEVLDP EQNSTFMDHYLE 465
 Simeli IGSM PGKVQSMK KAKKSNPLFLLDEIDKMGQDFRGDPSSALLEVLDP EQNSTFMDHYLE 461
 Rasola IGSM PGKILQSLTKVGVRNPLFLLDEIDKMGADFRGDPSSALLEVLDP EQNHTFQDHYIE 460
 * * * * * : : * . * * : : * * * * * : * * : * * * * * * * * * * * * * * * * *

Azbras VDYDLS DVMFVCTANTMRMPQPLLD RMEIIRVAGYTEDEKVEI SKRHLIEKQVEANGLKK 517
 Rhruabr VDYDLS DVMFMTTANSLRMPQPLLD RMEIIRLSGYTEDEKIEI AKRHLLIDKQVKANGLRR 516
 Brmeli VEYDLS NVMFVTTANTMNIIPG PLLDRMEIIRIAGYTEDEKLEI AKRHLLPKAIKDHALQP 540
 Meloti VEYDLS SVMFVTTANTLNIIPAPLM DRMEIIRIAGYTEDEKIEI AKRHLLMPKVIIRDHALQP 525
 Simeli VEYDLS NVMFITANTLNIIPPLMDRMEVIRIAGYTEDEKREI AKRHLLPKAIRDHALQP 521
 Rasola VDFDLS DVMFVATSNSLNIIPAPLLDRMEVIRLSGYTEDEKVNI AQRYLLPKQTKNNGLKG 520
 * : : * * . * * * : * : * : . : * * : * * * * * : * : * : * : * * . : . * :

Azbras GEFAISDDALRDLIRYYTREAGVRSLEREIANLCRKAVKEILMKGSAGAKVSVTRRLDK 577
 Rhruabr GEWSISDDALRDLIRYYTREAGVRSLERELANLTRKATKEILLRKGAKVGVKVTTRANLEK 576
 Brmeli KEFSVTE DALRN VIRHYTREAGVRSLERELMLARKAVTEILKTK -- KKSVKITDKNLS 598
 Meloti KEFSVTE DAIIRAI IQTYTREAGVRSLERELMKLARKAVTEILKTK -- KKTVSITANLAD 583
 Simeli NEFSVTE DGA LMAVIQNYTREAGVRNFERELMKLARKAVTEILKTK -- TKKVEVTAENIHD 579
 Rasola GEIEVAEGALRDIIRYYTREAGVRSLEREVSKICRQVVKQLLLKKEAGASVKVDSNDLNDK 580
 * : : . * * : * * : * * * * * . : * * * * : : * * . . : * * . : * * . :

Azbras YAGVRRFHFGEAELEDLVGVT TGLAWTEVGGELLSIEAVSLPGKGRVTTT GKLGDVMKES 637
 Rhruabr YSGVRKYRYGAEEREDLVGVT TGLAWTEVGGELLSIEAVMMPGKGQVTLTGQLGDVMKES 636
 Brmeli YLGVEKFRFGQIDGEDQVGVT TGLAWTEVGGELLTIEGVMMPGKGRMTVTGNLRDVMKES 658
 Meloti YLGVQRYRFGQVEADDQVGVT TGLAWTEVGGELLTVEGVMMPGKGRMTVTGNLRDVMKES 643
 Simeli YLGVPRFRHGEAERDDQVGVT TGLAWTEVGGELLTIEGVMMPGKGRMTVTGNLRDVMKES 639
 Rasola FLGVRRFDFGLAGKEDQVGQVT TGLAWTEVGGDLLTIEAALMPGKGNITRTGSLGDVMKES 640
 : * * : : . * : * * . * * * * * * * * : * : * . : * * * . * * * * * * *

Azbras VQA AESYVKS RATAFGIKPTLFEK RDIHVHVPEGATPKDGPSAGVAMITSIVSVLTGIAV 697
 Rhruabr VQA ARSFVRSRAV TYGIKPTQFEK GDIHVHVPEGATPKDGPSAGVAMCTSIVSAMTGI AV 696
 Brmeli ISAAASYVRSRAIDFGIEPPLF DDKRDIHVHVPEGATPKDGPSAGIAMVTAIVSVLTGIPV 718
 Meloti ISAAASYVRSRALDFGIEPPLF DDKRDIHVHLPPEGATPKDGPSAGAAAMATAIVSVLTGIPV 703
 Simeli ISAAASYVRSRAIDFGIEPPLF DDKRDIHVHVPEGATPKDGPSAGVAMATAIVSVMTGIPI 699
 Rasola VEAARSVVRSR SARLGIKDEMFEK RDIHVHVPEGATPKDGPSAGIAMTTALVSVLTGIPV 700
 : . * * * * : * * * : * * : * :

Azbras R KDVAMTGEITLRGRVLP IGG LKEKLLAALRGG LKHVLI PKDNEKDLAEIPDNV KRGLEI 757
 Rhruabr R KDIAMTGEITLRGRVLP IGG LKEKLLAALRGG IKT VLI PKDNEKDLAEIPDNV KRG LTI 756
 Brmeli R KDIAMTGEVTLRGRVLP IGG LKEKLLATLRGG IKKVLI PEENAKDLAEIPDNV KNNLEI 778
 Meloti R ADVAMTGEITLRGRILP IGG LKEKLLAALRGG IKKVLI PEDNAKDLAEIPDNV KNGMEI 763
 Simeli S KDVAMTGEITLRGRVLP IGG LKEKLLAALRGG IKKVLI PEENAKDLADIPDNV KNSLEI 759
 Rasola R ADVAMTGEITLRGEVLP IGG LKEKLLA AHRGG IKL VLI PEENVKDLADIPDNV KNSIEI 760
 * : * * * * : * * * . : * * * * * * * * : * * * : * * * * * * * * . : * * :


```

Azbras  IPVSTVDDVVKHALVREVEPIEWKEPEAVEPAVAKPQTDGG---GEVLRH  804
Rhrubr  IPVGHADDEVLELALVRKLTPIEWISPDEAEEEARLKSKAAEAPAEGDVVTH  806
Brmeli  VPVSRVGEVLKHALVRQPEPIEWTEQENPTAVP--PVEDEA---GASLAH  823
Meloti  IPVSRVGEVLRHALVRMPEPIEWVEPVNPPATT--DSADDA---GKSLAH  808
Simeli  IPVSRMGEVIAHALRLRLPEPIEWDPASQPAALPSVDSQDEA---GTSIAH  806
Rasola  MPVRWIDKVLLELALVRKPEPLPEEEPKPAPEAGKTEGGDAH----EMVHH  806
      : **      . . * :      * * : *      * :      : *

```

Figura 9: 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Rhodospirillum rubrum* (Rorubr), *Brucella melitensis* (Brmeli), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), *Ralstonia solanacearum* (Rasola).

ALINHAMENTO DA SEQÜÊNCIA DE AMINOÁCIDOS DA SUBUNIDADE β DA PROTEÍNA HU DE *A. brasilense* COM DIVERSOS ORGANISMOS

```

Azbras  VNKNLVAHVADAVGLSKTDATKAVDAIFDGIADSLKNGEEVRLVGFVGFVAVSERAASEG 60
Rhcaps  MNKNLIIAAVADSAGLSKADAAKAVDGVFDAITGALKDDQEVRLVGFVGFVAVTKRAATEG 60
Simeli  MNKNELVAAVADKAGLSKADASSAVDAVFETIQGELKNGGDIRLVGFVGFVAVSRREASKG 60
Meloti  MNKNELVSAVADAASISKGDAQSAVDVAVFSVITGELKKGGDVRLVGFVGFVAVTVSKRAASTG 60
Psaeru  MNKSELIDAI AASADIPKDVAGRALDAVIESVTGALKAGDSVVLVGFVGFVAVKEPAARTG 60
Rasola  MNKTDLIDHVAAQTDMSKAAAGRAIDALIGGIKDALCDGGSVTLVGFVGFVAVLQVSARAG 60
      :*:.:*:  :*  .:.:*  *  *:*.:.:  :  .  *  .  .:  *****.*  *  .  *  *

Azbras  RNPRTGEKIAIPASKQPKFKPGKTLKDGLK- 90
Rhcaps  RNPRTGETIAIPASKQPKFKPGKGLKDAVNA 91
Simeli  RNPSTGAEVDIPARNVPKFTAGKGLKDAVN- 90
Meloti  RNPQTGAEVKIPARTVPKFSAGKGLKDAVN- 90
Psaeru  RNPQTGKPIKIA-AKIPGFKAGKALKDAVN- 89
Rasola  RNPRTGATIKIEGAKVAKFKPGKALKDALN- 90
      *** **  :  *  .  .  *..**  ***.:

```

Figura 10: Alinhamento utilizando-se 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), *Rhodobacter capsulatus* (Rhocaps), *Pseudomonas aeruginosa* (Psaeru) e *Ralstonia solacacearum* (Rasola).

SEQÜÊNCIA DE NUCLEOTÍDEOS ENCONTRADA PARA *nuoB* E SEU PRODUTO DE TRADUÇÃO

1 TTACA ATTTA ATACG ACTCA CTATA GGGAA TTCGA **GCTCT** GGA

Gene *nuoB*

44 atg **ggagtagaggctgccaagctggcgccgattccgcccggaccg**
M G V E A A K L A P I P P G P

89 gaacaggacgcctatctccgcgcggttaccgaggagatccaggaa
E Q D A Y L R A V T E E I Q E

134 aagggttcatcacggcgaagtacgaggacgtgctcgcctgggccc
K G F I T A K Y E D V L A W A

179 cgcaccgggtccctgtggccgatgaccttcggtctggcctgctgc
R T G S L W P M T F G L A C C

224 gcggtggagatgatccacgcctacatgagccggtacgacctggac
A V E M I H A Y M S R Y D L D

269 cgtttcggcgcgcattccgcgccccagcccgcgccagtcgactgc
R F G V I P R P S P R Q S D C

314 atgatcgtggccggcacgctgaccaacaagatggccccgcgctg
M I V A G T L T N K M A P A L

359 cgcaaggctctatgaccagatgccggaaccgcgctgggtgatctcg
R K V Y D Q M P E P R W V I S

404 atgggctcctgcgccaacggcggcggtactatcactattcctac
M G S C A N G G G Y Y H Y S Y

449 tcggtggtgcgcggtgcgaccggatcgttccggtc 484
S V V R G C D R I V P V

Figura 11: O produto de tradução do gene *nuoB* derivado da seqüência de nucleotídeos foi obtida pelo programa ORF Finder. O provável sítio de ligação de ribossomo está indicado em negrito. O códon de iniciação encontrado é indicado em letras azuis. A seqüência deste gene não está completa pois o gene não está inteiro no plasmídeo pH/P13.

**ALINHAMENTO DA SEQUÊNCIA DE AMINOÁCIDOS DA SUBUNIDADEβ DA
PROTEÍNA NADH-DESIDROGENASE DE *A. brasilense* COM DIVERSOS
ORGANISMOS**

```

Azbras -----MGVEAAKLAPIPPGPEQDAYLRAVTEEIQEKGFIITAKY 38
Brmeli MEEGGAGMGLTGTNTTLVAPQPKGILDPRTG--KPVGSDDAFFNDLNGELSDKGFIVTSA 58
Simeli MELASG-----TTLVAPQPKGILDPATG--KPIGSNDAFFGEINNELADKGFVLTST 50
Meloti MEKRSAGMGLNDSSGTLVAPKPKGIIDPNTG--RPVGEDDPFFLEINNELADKGFVLTST 58
Rhcaps -----MGVMTGSNT--AAVDMD-VAAAALNRDLQDKGFLLTSTA 35
Rasola -----MAIEGVLN-EGFVTTTA 16
                                     :   :   :** :   .

Azbras EDVLAWARTGSLWPMTFGLACCAVEMIHAYMSRYDLDRFGVIPRPSRQSDCMIVAGTLT 98
Brmeli DALITWARTGSLMWMTFGLACCAVEMMHISMPRYDAERFGIAPRASPRQSDVMIVAGTLT 118
Simeli DELINWARTGSLMWMTFGLACCAVEMMQMSMPRYDAERFGFAPRASPRQSDVMIVAGTLT 110
Meloti EALITWARSGLMFMTFGLACCAVEMIHTSMRYDSERFVGFAPRASPRQSDIMIVAGTLT 118
Rhcaps EDIINWARNGSLHWMTFGLACCAVEMMHTAMPRYDVERYGFAPRASPRQSDVMIVAGTLT 95
Rasola DKLINWTRTGLWPMTFGLACCAVEMMHAGASRYDLDRFGVVFRPSRQSDVMIVAGTLC 76
      :  :  : ** . ***      ***** :  :  : . *** : ** . * . ***** *****


Azbras NKMAPALRKVYDQMPEPRWVISMGSCANGGGYYHYSYSVVRGCDRIVPV----- 151
Brmeli NKMAPALRKVYDQMPEPRYVISMGSCANGGGYYHYSYSVVRGCDRVVPVDIYVPGCPPTA 178
Simeli NKMAPALRKVYDQMPEPRYVISMGSCANGGGYYHYSYSVVRGCDRVVPVDIYVPGCPPTA 170
Meloti NKMAPALRKVYDQMPEPRYVISMGSCANGGGYYHYSYSVVRGCDRIVPVDIYVPGCPPSA 178
Rhcaps NKMAPALRKVYDQMPEPRYVISMGSCANGGGYYHYSYSVVRGCDRIVPVDIYVPGCPPSA 155
Rasola NKMAPALRKVYDQMAEPRWVISMGSCANGGGYYHYSYSVVRGCDRIVPVDVYVPGCPPTA 136
      ***** . *** : ***** : *****

Azbras ----- 155
Brmeli EALLYGILLLLQKKIRRTGTIER-- 200
Simeli EALLYGVLLLLQKKIRRTGTIER-- 192
Meloti EALLYGILLLLQKKIRRTGTIER-- 200
Rhcaps EALMYGILQLQKIRRTGTLVR-- 177
Rasola EALIYGIIQLQAKIRRTNTIARKG 160

```

Figura 12: 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. As espécies utilizadas para o alinhamento foram: *Azospirillum brasilense* (Azbras), *Brucella melitensis* (Brmeli), *Mesorhizobium loti* (Meloti), *Sinorhizobium meliloti* (Simeli), *Rhodobacter capsulatus* (Rhocaps) e *Ralstonia solanacearum* (Rasola).

GENES IDENTIFICADOS NO PLASMÍDEO pAB441

Escala:  1 kb

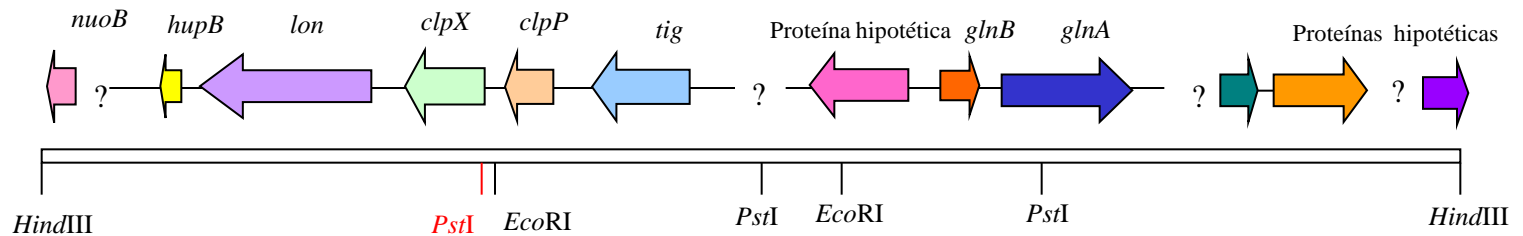


Figura 13: Genes identificados no plasmídeo pAB441. O sítio *PstI* em vermelho não está mapeado no trabalho de Bozouklian e colaboradores (1986). Os pontos de interrogação representam regiões não sequenciadas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D.J. "Gapped BLAST e PSI-BLAST: a new generation of protein database search programs". **Nucleic Acids Res.** v. 25 p. 3389-3402, 1997.
- 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.**, Washington, v. 178, p.4830-4838, 1996.
- BANKIER, A.T.; WESTON, K.M.; BARREL, B.G. Random cloning and sequencing by the M13 dideoxynucleotide chain termination method. **Methods Enzymol.**, v. 155, p. 51-93, 1987.
- BASHAN, Y.; HOLGUIN, G. *Azospirillum* – plant relationships: environmental e physiological advances (1990-1996). **Can. J. Microbiol.** v. 43, p. 103-121, 1997
- BONNEFOY, E., TAKAHASHI, M. e ROUVIÈRE YANIV, J. DNA binding parameters of the HU protein of *Escherichia coli* to cruciform DNA. **J. Mol. Biol.** v. 242, p. 116-129, 1994.
- BOZOUKLIAN, H. ; ELMERICH, C. Nucleotide sequence of the *Azospirillum brasilense* Sp7 glutamine synthetase structural gene. **Biochimie.** v.68, p.1181-1187, 1986.
- BRUCKNER R.; TITGEMEYER, F., Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. **FEMS Microbiol Lett.** v. 209, p. 141-148, 2002.
- BURRIS, R. H. Nitrogenases. **J. Biol Chem.** v.266, p. 9339-9342, 1991.
- CARMONA, M.; MAGASANIK, B., Activation of Transcription at σ^{54} - dependent Promoters on Linear Templates Requires Intrinsic or Induced Bending of the DNA. **J. Mol. Biol.** v. 261, p. 348-356, 1996.

- CHEN,S.; YANG,H.; LI,J. Nucleotide sequence of *glnB* and *glnA* cluster from *Azospirillum brasilense*. Número de acesso no GenBank: AF323964
- DE MEL, V. S. J.; KAMBEROV, E. S.; MARTIN, P. D.; ZHANG, J.; NINFA, A. J.; EDWARDS, B. F. P. 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.; STACKEBRET, E.; SLY, L. I. Transfer of *Comglomeromonas largomobilis* subsp. *largomobilis* to the genus *Azospirillum* as *Azospirillum largomobile* comb. nov., elevation of *Comglomeromonas largomobilis* subsp. *Paroensis* to the new type species of *Comglomeromonas*, *Comglomeromonas parroensis* sp. nov. **System. Appl. Microbiol.** v. 20, p. 72-77, 1997.
- DÖBEREINER, J. The genera *Azospirillum* e *Herbaspirillum*. In BALLOWS, A., TRUPPER, H.G., DWORKING, M., HARDER, W. **The Prokaryotes**. Springer-Verlag, 2^a ed. v. III, p.2236-3353, 1991.
- DÖBEREINER, J.; DAY, J. M. Associative symbioses in tropical grasses and characterization of microorganisms and dinitrogen-fixing sites. In NEWTON, W. E. & NYMAN, C. J. eds. **Proc. of the 1st. Intl. Symp. Nitrogen Fixation**. Pullman, Washington Univ. Press. v.2, p.518-38, 1996.
- DÖBEREINER, J.; PEDROSA, F. O. Nitrogen-fixing bacteria in nonleguminous crop plants. **Science Tech Publishers**, Springer-Verlag, 1987.
- 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.; ROUVIÈRE-YANIV. Role of the histone-like proteins OsmZ and HU in homologous recombination. **Gene**, v. 120, p. 11-16.

- EADY, R. R. Enzymology 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 doebereineriae* sp. nov., a nitrogen-fixing bacterium associated with the C₄-grass *Miscanthus*. **Int. J. Syst. Evol. Microbiol.** v. 51, p. 17-26, 2001.
- ELMERICH, C.; ZAMAROCZY, M. de; ARSENE, F.; PEREG, L.; PAQUELIN, A.; KAMINSKI, A. Regulation of *nif* gene expression and nitrogen metabolism in *Azospirillum*. **Soil Biol. Biochem.** v.29, p. 874-852, 1997.
- ETCHEGARAY, J. P.; JONES, P. G. e INOUYE, M. Differential thermoregulation of two highly homologous cold-shock genes, *cspA* e *cspB*, of *Escherichia coli*. **Genes Cells** v. 1, p. 171-178, 1996.
- FADEL-PICHET, C. M. T.; SOUZA, E. M.; RIGO, L. U.; FUNAYAMA, S.; YATES, M. G.; PEDROSA, F. O. Regulation of *Azospirillum brasilense nifA* gene expression by ammonium and oxygen. **FEMS Microbiol. Lett.** v.179, p. 281-288, 1999.
- FINEL, M. Organization and evolution of structural elements within complex I. **Biochim Biophys Acta**, v. 1364(2), p. 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**, v. 367(2) p. 107-11, 1995.
- FRIEDRICH, T. The NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*. **Biochim Biophys Acta**, v. 1364(2) p. 134-46, 1998.
- FRYDMAN, J. Folding Of Newly Translated Proteins In Vivo: The Role of Molecular Chaperones. **Annu. Rev. Biochem.** v. 70 p. 603-647, 2001
- GIANGROSSI M.; GIULIODORI A. M.; GUALERZI, C. O.; PON C. L. Selective expression of the β -subunit of nucleoid-associated protein HU during cold shock in *Escherichia coli*. **Mol. Microbiol.**, v. 44 (1), p. 205-216, 2002.

- GOTTESMAN, S.; MAURIZI M. R.; Regulation by proteolysis: energy-dependent proteases e their targets. **Microbiol Rev.** v. 56(4), p. 592-621, 1992.
- HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor e analysis program for Windows 95/98/NT. **Nucl. Acids. Symp. Ser.** v. 41, p. 95-98, 1999.
- HANAHAN, D. Studies on transformation of *Escherichia coli* with plasmids. **J. Mol. Biol.** v.166, p. 557-580, 1983.
- HAYKINSON, M. J.; JOHNSON, E R. C. DNA looping and the helical repeat *in vitro* e *in vivo*: effect of HU protein e enhancer location on Hin invertasome assembly. **EMBO J.** v. 12, p.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.
- HOOVER, T. R.; SANTERO, E.; PORTER, S.; KUTSU, S. The Integration Host Factor stimulates interaction of RNA polimerase with NifA, the transcriptional activator of nitrogen fixation operons. **Cell**, v. 63, p. 11-22, 1990.
- HUALA, E.; MOON, A. L.; AUSUBEL, F. M. Aerobic Inactivation of *Rhizobium meliloti* NifA in *Escherichia coli* Is Mediated by *lon* and Two Newly Identified Genes, *snoB* and *snoC*.. **J Bacteriol** v. 173 (1), p. 382-390, 1991.
- INDORATO, C.; GIANNELLI, L.; BAZZICALUPO, M. *Azospirillum brasilense* *tig-clp-clpx-lon*, 1999. Número de acesso no GenBank: AF150957
- INOYUYE, M. Experimental manipulation of gene expression. **Academic Press, Inc.** p. 28-30, 1983.
- JIANG, P.; NINFA, A. J., Regulation of autophosphorylation of *Escherichia coli* nitrogen regulator II by the PII signal transduction protein. **J Bacteriol.** v. 181(6),p. 1906-11, 1999.
- 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 e ATP. **J Biol Chem**, Baltimore, v. 270, p. 17797-17807, 1995.
- KEROR, O.; GOLDBERG A. L. Trigger factor is induced upon cold shock e enhances viability of *Escherichia coli* at low temperatures **Proc. Natl. Acad. Sci. USA** v. 94, p. 4978-4981, 1997
- KANO, Y.; YOSHINO. S.; WADA. M.; YOKOYANA, K.; NOBUHARA, M.; IMAMOTO, F. Molecular cloning e nucleotide sequence of HU-1 gene of *Escherichia coli*. **Mol. Gen. Genet** v. 201, p. 360-362, 1985.
- KANO Y.; WADA, M.; IMAMOTO F. Genetic characterization of the gene *hupA* encoding the HU-2 protein of *Escherichia coli*. **Gene**. v30;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**. v. 89(1), p. 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**. v. 103(1), p.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.
- KUTSU, S.; HIRSCHMAN, J; BURTON, M.; JELESKO, J.; MEEKS, J. C. Covalent modification of bacterial glutamine synthetase: physiological significance. **Mol. Gen. Genet.** v. 197, p. 309-317. 1984.
- LAVOIE, B. D.; CHACONAS, G. Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. **Genes Dev.** v. 7, p. 2510-9, 1993.

- 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 and in the presence of air and ammonia. **Mol. Microbiol.** v.5 , p. 2735-2744, 1991
- MACHADO, H. B.; FUNAYAMA, S.; RIGO, L. U.; PEDROSA, F.O. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. **Can. J. Microbiol.** v.29, p. 549-553, 1991.
- MACHADO, G. 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-btrC* operon e are negatively regulated. **Can. J. Microbiol.** v. 41, p. 674-684, 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.; GOWRISHANKAR, J. Evidence for involvement of protein HU e RpoS in transcription of the osmoreponsive *proU* operon in *Escherichia coli*. **J. Bacteriol.** v. 176, p. 5378-5384, 1994.
- MEAD, D. A.; SZCZENA-SKORUPA, E.; KEMPER, B. Single strand DNA “blue”T7 promoter plasmids a versatile tandem promoter system for cloning e 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* e *nifA*. **EMBO J.** v. 2, p. 39-44, 1983.

- MERRICK, M. J.; EDWARDS, R. A. Nitrogen control in bacteria. **Microbiol. Rev.**, v. 59, p. 604-622, 1995
- MORI E.; FULCHIERI, M.; INDORATO, C.; FANI, R.; BAZZICALUPO, M. Cloning, nucleotide sequencing, e expression of the *Azospirillum brasilense lon* gene: involvement in iron uptake. **J Bacteriol.** v. 178, p. 3440-6, 1996.
- REINHOLD, B.; HUREK, T.; FENDRIK, I.; POT, B.; GILLIS, M.; KERSTERS, K.; THIELEMANS, S.; de LEY J. *Azospirillum halopraeferans* sp. nov., a nitrogen-fixing organism associated with roots of kallar grass (*Leptochloa fusca* (L.) Kunth) **Int. J. Syst. Bacteriol.** v. 37, p. 43-51, 1987.
- ROUVIERE-YANIV, J.; YANIV, M.; GERMONK, J. E. *Escherichia coli* DNA-binding protein HU forms nucleosome-like structure with circular double-strand DNA. **Cell** v. 17, p. 265-274, 1979.
- PARSELL, D. A.; LINDQUIST, S. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. **Annu. Rev. Genet.** v. 27, p. 437-496, 1993.
- PEDROSA, F. O. Fixação biológica de nitrogênio: fértil idéia. **Ciência Hoje**, v. 6, p. 12-13, 1987.
- PEDROSA, F. O. Physiology, biochemistry and genetics of *Azospirillum* and other root-associated nitrogen-fixing bacteria. **Critical Reviews in Plant Sciences.** Boca Raton, v. 6, p. 345-383, 1988.
- POSTGATE, J. R. The fundamentals of nitrogen fixation. Cambridge Univ. Press., p. 252, 1982
- 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.

- SCHMIDT, M.; LUPAS, A. N.; FINLEY, D. Structure and mechanism of ATP-dependent proteases. **Curr. Opin. Chem. Biol.** v. 3, p. 584-91, 1999.
- SCHOLZ, C. G.; STOLLER, T.; ZARNT, G.; FISCHER, E. F.; SCHMID X. Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. **EMBO J.** v. 16 p. 54-58, 1997.
- SHARPLES, G. J.; CHAN, S. N.; MAHDI, A. A.; WHITHY, M. C.; LOYD R. G. Processing of intermediates in recombination e DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. **EMBO J.** v. 13, p. 6133-42, 1994.
- STEENHOUDT, O.; VANDERLEYDEN, J. *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. **FEMS Microbiol Rev.** v. 24, p. 487-506, 2000.
- TARRE, J.J.; KRIEG, N. R.; DOBEREINER, J. A taxonomic study of the *Spirillum lipoferum* group with description of a new genus, *Azospirillum* gen. nov. e two species, *Azospirillum lipoferum* (Beijerinck) comb. nov. e *Azospirillum brasilense* sp. nov. **Can. J. Microbiol.** v. 24, p. 967-980, 1978.
- WEIDNER, U.; GEIER, S.; PTOCK, A.; FRIEDRICH, T.; LEIF, H.; WEISS, H. The gene locus of the proton-translocating NADH-ubiquinone oxidoreductase in *Escherichia coli*. Organization of the 14 genes e relationship between the derived proteins e subunits of mitochondrial complex I. **J Mol Biol**, v. 233, p. 109-22, 1993.
- WU, Y.; DATTA, P. Influence of DNA topology on expression of the *tdc* operon in *Escherichia coli* K-12. **Mol Gen Genet.** v. 247, p. 764-7,1995.
- YAGI, T.; YANO, T.; di BERNARDO. S.; MATSUNO-YAGI, A. Procaryotic complex I (NADH-1), an overview. **Biochim Biophys Acta.** v. 1364, p. 125-33, 1998.
- YAMANAKA, K. Cold-shock response in *Escherichia coli*. **J Mol Microbiol Biotechnol** v. 1, p. 193-202, 1999.
- YOUNG, J. P.W. Phylogenetic classification of nitrogen-fixing organisms. In: **Biological Nitrogen Fixation**. STACEY, G.; BURRIS, R. M.; EVANS, H. J. eds. Chapman & Hall, New York. p. 42-86, 1992.

- 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.** v. 201, p. 199-204, 2001.
- de ZAMAROCZY, M.; DELORME, F.; ELMERICH C. Characterization of three different nitrogen-regulated promoter regions for the expression of *glnB* e *glnA* in *Azospirillum brasilense*. **Mol. Gen. Genet.** v. 224, p. 421-430, 1990.
- de ZAMAROCZY, M.; PAQUELIN, A.; ELMERICH, C. Functional organization of the *glnB-glnA* cluster of *Azospirillum brasilense*. **J. Bacteriol.** 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.** 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. **Molecular Microb.** v. 29, p. 449-463, 1998

