

***orfatrAb* - Caracterização molecular do gene que codifica um ativador de transcrição da família LysR em *Azospirillum brasilense*¹**

*Samanta Bolzan de Campos*², *Irene Silveira Schrank*³ e *Luciane Maria Pereira Passaglia*²

Resumo

orfatrAb - Caracterização molecular do gene que codifica um ativador de transcrição da família LysR em *Azospirillum brasilense*

O nitrogênio, elemento-chave na composição de diversas biomoléculas como aminoácidos e ácidos nucléicos, é abundante na atmosfera, onde, sob forma de N₂, não está disponível para a maioria dos seres vivos. Entretanto, existem microrganismos, chamados de diazotróficos, que são capazes de converter o nitrogênio atmosférico em formas assimiláveis para os outros organismos, através do processo de Fixação Biológica do Nitrogênio. Bactérias do gênero *Azospirillum* são microrganismos diazotróficos de vida livre ou endofíticos, que se associam com gramíneas como trigo, arroz, cana-de-açúcar, entre outras. Uma das ferramentas mais utilizadas na identificação de genes envolvidos na fixação de nitrogênio em *Azospirillum* foi a mutagênese sítio-dirigida com o transposon Tn5. Essa metodologia possibilitou o isolamento de um transconjugante de *A. brasilense*, cuja capacidade de fixação de nitrogênio *in vitro* foi bastante elevada em relação à da linhagem de tipo selvagem. No local de inserção do transposon no DNA desse transconjugante foram encontradas duas fases abertas de leitura, *orfatrAb* e *orf281*, tendo sido, essa última, interrompida pelo Tn5. O objetivo deste trabalho foi determinar a seqüência de nucleotídeos de *orfatrAb*, e, a partir desta, a seqüência de aminoácidos, a qual apresentou grande similaridade, principalmente em sua região amino-terminal, com proteínas da família de ativadores de transcrição LysR. *ORFatrAb* possui 297 resíduos, peso molecular de 33.581,6 kDa, ponto isoelétrico de 7,26 e índice de instabilidade de 33,29. Na região reguladora de *orfatrAb* foram identificados elementos de seqüência consensuais para a ligação de proteínas ativadoras de genes envolvidos no metabolismo e na fixação biológica do nitrogênio.

Palavras-chave: fixação biológica do nitrogênio, organismos diazotróficos, genes atr, *Azospirillum brasilense*.

Abstract

orfatrAb - Molecular characterization of one transcription activator coding gene of the LysR family in *Azospirillum brasilense*

Nitrogen, which is a component of several biomolecules such as nucleic acids and amino acids, is abundant in the atmosphere where, in its gaseous form (N₂), it is unavailable for most organisms.

1 Apoio Financeiro: FAPERGS e CNPq. **Autor para correspondência:** Luciane Maria Pereira Passaglia. (E-mail: lpassaglia@terra.com.br). Departamento de Genética/UFRGS – Av. Bento Gonçalves, 9500, C. P. 15053, Porto Alegre, RS, CEP 91501-970, prédio 43323, sala 207.

2 Departamento de Genética

3 Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS.

However, there are microorganisms, called diazotrophs that are able to convert the atmospheric nitrogen into an assimilable form to other organisms through a process called biological nitrogen fixation. Bacteria of the genus *Azospirillum* are diazotrophs, free-living or endophytic that are able to associate with grasses like wheat, rice, sugar cane, etc. One of the most used tools for the identification of genes involved in the nitrogen fixation in *Azospirillum* was the site-direct mutagenesis using the Tn5 transposon. This methodology allowed the isolation of one *A. brasilense* transconjugant with a capacity of biological nitrogen fixation *in vitro* that was higher than that of the wild type strain. In the place of transposon insertion into the DNA of this mutant two open reading frames were identified, *orfatrAb* and *orf281*, which was interrupted by Tn5. The aim of this work was to determinate the *orfatrAb* complete nucleotide sequence and the corresponding amino acid sequence which showed great similarity, especially in the amino-terminus, with proteins of the LysR family of transcriptional activators. ORFATRAb has 297 residues, molecular weight of 33,581.6 kD, isoelectric point of 7.26 and instability index of 33.29. Sequence consensus elements of gene activator proteins involved in the metabolism and in the biological nitrogen fixation were identified in the *orfatrAb* promoter region.

Keywords: biological nitrogen fixation, diazotroph organisms, atr genes, *Azospirillum brasilense*.

Introdução

Fixação Biológica do Nitrogênio (FBN)

O nitrogênio é um elemento-chave na composição de diversas biomoléculas, como proteínas, aminoácidos e ácidos nucléicos. Este nitrogênio é abundante na atmosfera, onde, sob forma de N₂ (com uma ligação tripla entre os dois átomos de N), é muito resistente ao ataque químico, e, portanto, não está disponível para a maior parte dos seres vivos. Porém, existem microrganismos, como bactérias e algas azuis, que são capazes de converter este N₂ em formas que possam ser assimiladas por outros organismos, como a amônia. Estes microrganismos, chamados de diazotróficos, são responsáveis por aproximadamente 60% do nitrogênio fixado e disponível no ambiente. O processo de conversão do nitrogênio atmosférico em compostos orgânicos assimiláveis recebe a denominação de Fixação Biológica do Nitrogênio (FBN) (Stryer 1994).

A alternativa industrial para a conversão do nitrogênio molecular em compostos orgânicos reativos foi estabelecida, em 1910, por Fritz Haber. Embora utilizada na fabricação de fertilizantes, a condição de quebra da ligação só é obtida a uma temperatura de 500°C e uma pressão de 300 atmosferas (Stryer 1994), o que torna o processo bastante caro e praticamente inviável.

O processo de FBN por microrganismos diazotróficos é catalisado pelo complexo enzimático nitrogenase. Um dos produtos da reação catalisada pela nitrogenase é a amônia (NH₃) que a bactéria diazotrófica fornece como fonte assimilável de nitrogênio para os outros organismos ou diretamente para o solo. A relevância deste processo se dá, além da redução dos custos com adubos nitrogenados, pela importância do nitrogênio como um elemento vital, aumentando a produtividade das plantas (Peters *et al.* 1995).

O complexo enzimático nitrogenase possui duas sub-unidades: a Fe-proteína e a MoFe-proteína. Durante a reação ocorre o transporte de elétrons da Fe-proteína para

uma das sub-unidades da MoFe-proteína. Este processo utiliza uma grande quantidade de energia: pelo menos 16 moléculas de ATP são necessárias para cada molécula de N_2 reduzida à amônia. Devido a isso, existem diferentes níveis de controle na expressão dos genes envolvidos, e, também, mecanismos de controle em nível posterior à transcrição (Fu *et al.* 1989).

Os genes envolvidos no processo de FBN são chamados *nif* (de *nitrogen fixation*) e foram primeiramente caracterizados na bactéria diazotrófica de vida livre *Klebsiella pneumoniae*. A maioria destes genes está envolvida na montagem e estabilização do complexo enzimático da nitrogenase. Existe uma grande similaridade em estrutura, organização e função dos genes *nif* de diversos organismos diazotróficos, com estes identificados em *K. pneumoniae*. Os genes *nif* desta bactéria estão localizados em uma região de DNA de aproximadamente 24.000 pares de bases, próximos aos genes da biossíntese de histidina, e organizados em oito operons (Arnold *et al.* 1988). Os genes estruturais da nitrogenase estão situados em um mesmo operon (*nifHDK*), onde o gene *nifH* codifica a subunidade da Fe-proteína e os genes *nifDK* codificam a MoFe-proteína. Muitos dos outros genes estão envolvidos na síntese do cofator da enzima, o MoFeco, entre eles *nifE*, *nifN*, *nifV*, *nifB*, *nifQ*. Existem, também, genes envolvidos no transporte de elétrons específicos para o sistema de FBN, como os genes *nifJ* e *nifF* (Dixon 1984; Arnold *et al.* 1988).

Entre os genes *nif* de *K. pneumoniae* existem dois cujos produtos apresentam funções reguladoras, que são os genes *nifL* e *nifA*. Nessa bactéria, o produto do gene *nifL* apresenta a função de reprimir a atividade do produto do gene *nifA*, quando as condições para a síntese da nitrogenase não são adequadas. A proteína reguladora NifA, produto do gene *nifA*, reconhece seqüências UAS (seqüências ativadoras a montante, do inglês *upstream activator sequences*) no DNA, aproximadamente 100 nucleotídeos a montante do sítio de ligação da RNA-polimerase, ativando a trans-

crição dos genes *nif*. As seqüências UAS, que possuem a seqüência consensual TGT-N₁₀-ACA, estão presentes em todas regiões promotoras dos genes *nif* já caracterizadas, evidenciando, portanto, que estes genes estão sujeitos à ativação da transcrição através ligação da proteína NifA (Dixon 1984; Arnold *et al.* 1988).

Além deste sistema de regulação com os genes *nifLA*, existem, em *K. pneumoniae*, outros genes cujos produtos apresentam funções reguladoras gerais e que também regulam os genes *nif*, como os genes *ntrB* e *ntrC*. Os produtos dos genes *ntrBC* podem estar envolvidos na regulação do metabolismo de nitratos, regulação posterior à tradução da nitrogenase e podem também ativar a transcrição do gene *nifA* (Arnold *et al.* 1988; Fu *et al.* 1989; Merrick & Edwards 1995). Um outro mecanismo de controle, que não é encontrado em *Klebsiella pneumoniae*, ocorre em diversos diazotróficos através dos produtos dos genes *draT* e *draG*. Os produtos desses genes possuem funções de regulação posterior à tradução da nitrogenase, agindo diretamente sobre a enzima, inativando-a quando o nível de oxigênio celular está muito elevado (Fu *et al.* 1989; Zhang *et al.* 1997).

A região promotora dos genes envolvidos no metabolismo e na fixação biológica do nitrogênio possui uma seqüência de reconhecimento para a RNA-polimerase, que difere daquelas presentes em genes que desempenham outras funções. O fator sigma (σ) da RNA-polimerase de procariotes é o responsável pelo reconhecimento da região promotora e se liga a ela para promover o início da transcrição. A região promotora dos genes envolvidos no processo de fixação do nitrogênio é constituída por uma seqüência consensual rica nos nucleotídeos C e G, localizada nas posições -24/-12, que é reconhecida pelo fator sigma alternativo da RNA-polimerase, σ^{54} (Schrank 2003). A RNA-polimerase em complexo com o fator σ^{54} necessita da presença de proteínas acessórias para a transcrição dos genes, que, no caso dos genes envolvidos na fixação do nitrogênio de bactérias diazotróficas, é a proteína ativadora

NifA, juntamente com a proteína IHF (fator de integração do hospedeiro) (Barrios *et al.* 1999). Com a ligação da proteína ativadora NifA na região UAS e o reconhecimento e ligação da RNA-polimerase- σ^{54} , a proteína IHF promove uma curvatura no DNA, aproximando essas duas proteínas. A interação da proteína NifA com o complexo da RNA-polimerase promove o início da transcrição.

Gênero *Azospirillum*

O gênero *Azospirillum* compreende bactérias diazotróficas, de vida livre ou endofíticas, que pertencem a um grupo de bactérias que promove o crescimento vegetal. Triagens em campo demonstraram que, sob certas condições ambientais, a inoculação de *Azospirillum* tinha efeitos benéficos nas plantações. Porém, apesar de todos os estudos realizados, o modo de ação pelo qual *Azospirillum* aumenta o crescimento das plantas ainda não foi completamente elucidado (Steenhoudt & Vanderleyden 2000; Vande-Broek & Vanderleyden 1995).

O gênero possui sete espécies: *Azospirillum brasilense*, *A. lipoferum*, *A. amazonense*, *A. irakense*, *A. halopraeferens*, *A. largimobile* e *A. doebereineriae*, as quais se associam, preferencialmente, com gramíneas (Martin-Didonet *et al.* 1999; Potrich *et al.* 2001).

As culturas de *Azospirillum* produzem fitormônios, como auxinas, principalmente IAA (Vande-Broek & Vanderleyden 1995; Zimmer *et al.* 1998; Vande-Broek *et al.* 1999), citocinas e giberelinas (Vande-Broek & Vanderleyden 1995). Dos fitormônios produzidos, a auxina é tida como a de maior interesse, pois é ela que promove o crescimento das plantas associadas com a bactéria (Vande-Broek & Vanderleyden 1995).

As bactérias do gênero *Azospirillum* fixam nitrogênio atmosférico sob condições de baixa tensão de oxigênio e falta de amônia (Vande-Broek & Vanderleyden 1995). O controle do processo ocorre tanto em nível de síntese da nitrogenase (Vande-Broek & Vanderleyden 1995), quanto em nível da sua ativi-

dade (Fu *et al.* 1989; Vande-Broek & Vanderleyden 1995). O íon amônio (NH_4^+) possui ação inibitória rápida, porém reversível, da atividade da nitrogenase em uma grande quantidade de microrganismos diazotróficos, inclusive nas bactérias do gênero *Azospirillum*, que diminuem a atividade da nitrogenase quando inoculadas com NH_4Cl no meio de cultura. O provável modo de ação do íon seria uma modificação covalente em uma das sub-unidades da Fe-proteína (Fu *et al.* 1989).

O gênero *Azospirillum* apresenta uma estrutura genômica complexa, com múltiplos replicons, e estruturas parecidas com cromossomos (*chromosome-like*), no qual o genoma se divide. O operon *nifHDK*, aparentemente, está localizado em um único e grande replicon (Martin-Didonet *et al.* 1999).

Testes com oligonucleotídeos específicos em isolados de bactérias fixadoras de nitrogênio de diversas plantas verificaram que as diversas espécies do gênero *Azospirillum* apresentam associações com gramíneas de importância agrônômica e econômica como trigo, arroz, cana-de-açúcar, palmeiras, aveia, entre outras (Kirchhof *et al.* 1997).

Mutação sítio-dirigida com transposon Tn5

A técnica de mutação sítio-dirigida por transposon foi muito importante na identificação e no estudo do funcionamento de genes, principalmente no início dos anos 80. O transposon Tn5 e seus derivados se mostraram como boas ferramentas de biologia molecular, pois possuem uma alta estabilidade de transposição, baixa frequência de rearranjos e um gene de resistência a um antibiótico, *nptII*, que confere resistência à canamicina, tornando mais fácil a identificação dos mutantes (Araújo 1988; Lewin 2001).

O transposon Tn5 possui 5,7kb de tamanho, com duas seqüências de inserções IS50 nas suas extremidades e o gene *nptII* na região central (Lewin 2001).

Araújo, em 1988, induziu diversas inserções do Tn5 na linhagem selvagem Sp7 de

Azospirillum brasilense com o objetivo de identificar e isolar o operon da nitrogenase, *nifHDK*, dessa bactéria. Os mutantes obtidos foram isolados a partir da resistência ao antibiótico canamicina e da manifestação do fenótipo Nif^- . A maior parte dos mutantes com fenótipo Nif^- apresentou a inserção do transposon em um fragmento de aproximadamente 6,5kb de tamanho, resultado da clivagem do DNA total da bactéria com a enzima de restrição *EcoRI*, e que continha os genes *nifHDK*. Alguns dos mutantes, no entanto, apresentaram fenótipo Nif^+ . Em um desses, denominado Sp7::Tn5-81, a inserção do transposon foi posicionada em um fragmento de aproximadamente 2,4kb, resultado da clivagem do DNA total de *Azospirillum* com as enzimas de restrição *EcoRI* e *HindIII*, adjacente ao fragmento de *EcoRI* de 6,5Kb. Esse resultado ocorreu, provavelmente, devido a um evento de recombinação homóloga. Em outro mutante com o fenótipo Nif^+ analisado, denominado de Sp7::Tn5-33, acredita-se que tenha ocorrido um evento de recombinação não-homóloga, pois, apesar do fragmento gerado ter apresentado o tamanho esperado, de aproximadamente 12,3kb (5,7 kb do transposon e 6,5kb do genoma de *Azospirillum*), a sequência interrompida não correspondeu ao operon *nifHDK* (Araújo 1988; Revers 1996; Blah a 2001).

O seqüenciamento da região correspondente à inserção do Tn5 no genoma do

mutante Sp7::Tn5-33 identificou duas fases abertas de leitura (ORF, do inglês *open reading frame*), as quais foram denominadas de *orf281* e *orfatrAb*, respectivamente, conforme mostrado na Figura 1. Estas duas ORFs apresentam sentidos contrários de transcrição. A primeira delas, *orf281*, apresentou similaridade de seqüência de aminoácidos com uma família de proteínas responsivas a estresse, USP (*universal stress proteins*) (Revers *et al.* 2000; Blaha 2001). A seqüência parcial de aminoácidos, deduzida a partir da seqüência de nucleotídeos, da outra ORF, *orfatrAb*, apresentou grande similaridade de seqüência com membros da família de proteínas ativadoras de transcrição em procariotos LysR (Revers 1996; Blaha 2001).

LysR: ativadores de transcrição em procariotos

A família de ativadores de transcrição de procariotos LysR (*autoregulatory transcriptional regulators* – ATR) possui o motivo hélice-volta-hélice em seu domínio de ligação ao DNA na extremidade amino-terminal, a qual apresenta, devido a isso, uma grande conservação entre os membros da família. A extremidade carboxi-terminal não apresenta grande conservação, podendo ser bem variada entre as diferentes ATRs (Henikoff *et al.* 1988; Schell 1993). Em um estudo comparativo com diversos membros da família LysR, Henikoff

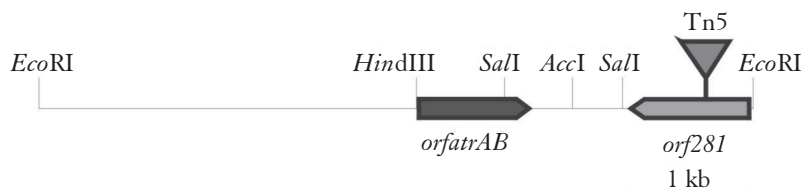


Figura 1 – Mapa físico do fragmento de DNA contendo a ORF281 e ORFatrAb isolado de *Azospirillum brasilense* Sp7, mostrando o local de inserção do transposon Tn5 e os sentidos de transcrição das ORFs, os quais estão indicados pelas setas que acompanham os genes.

et al. (1988) propuseram que uma homologia de 20% na seqüência de aminoácidos com outros membros da família LysR e/ou a presença do motivo hélice-volta-hélice em uma região amino-terminal com alto grau de conservação caracterizariam essa seqüência como correspondente a de uma ATR.

Outra característica desta família é o conteúdo dos aminoácidos lisina e arginina. A maioria dos membros da família LysR apresenta uma quantidade reduzida do aminoácido lisina (Lys) e um aumento na quantidade do aminoácido arginina (Arg), quando comparado com outras proteínas procarióticas (Viale *et al.* 1991).

Em resposta a diferentes co-indutores, as ATRs ativam a transcrição de genes ou regulons com sentidos divergentes ao seu e com as mais diversas funções na célula (Schell 1993). Esta família possui mais de cinquenta membros conhecidos com tamanhos similares (em torno de 300 aminoácidos) e, que, aparentemente, evoluíram de um ancestral distante para subfamílias, que são encontradas em diversos gêneros de procariotos, possuindo representantes em diversos microrganismos (Schell 1993). Muitos genes, cujos produtos são ATRs, já foram identificados, como, por exemplo, o gene *rbcR* em *Chromatium vinosum* (Viale *et al.* 1991) e em *Anabaena* sp. (Mori *et al.* 2002), os genes *nac* e *cysB* em *Klebsiella aerogenes* (Bender 1991; Schwacha *et al.* 1993; Pompiello & Bender 1995; Tyrrell *et al.* 1997), o gene *nodD* em bactérias simbióticas como *Rhizobium*, *Bradyrhizobium*, *Azorhizobium*, *Sinorhizobium* e *Mesorhizobium* (Loh & Stacey 2002), o gene *syrM* em *Rhizobium meliloti* (Schell 1993) e o gene *oxyR* em *Escherichia coli* (revisado em Blaha 2001).

A ORF encontrada no mutante Sp7::Tn5-33 (*orfatrAb*) demonstrou, através de suas seqüências de nucleotídeos e de aminoácidos, diversas características que a posicionam como pertencente à família LysR, como a conservação da região amino-terminal, com a presença do motivo de ligação ao DNA, hélice-volta-hélice, e o conteúdo anormal da pro-

porção dos aminoácidos Lys/Arg (Revers 1996; Blaha 2001).

Objetivo

Este trabalho teve como objetivo a obtenção da seqüência completa de nucleotídeos do gene *orfatrAb* e da seqüência de amino-ácidos da proteína (ORFAtrAb) correspondente, para uma análise comparativa mais abrangente com as seqüências protéicas de outros membros da família LysR já caracterizadas.

Material e métodos

Reações de Seqüenciamento

O plasmídeo pKS6,0 (Revers 1996, com. pessoal), o qual possui um fragmento de 6,0 kb de *EcoRI* contendo uma parte do gene *orf281* e a seqüência completa do gene *orfatrAb* de *A. brasilense*, foi utilizado como DNA-molde nas reações de seqüenciamento deste mesmo gene.

As reações de seqüenciamento manual foram realizadas com o kit Thermo Sequenase Radiolabeled Terminator Cycle Sequencing da Amersham Pharmacia Biotech. A estratégia adotada foi a utilização de oligonucleotídeos internos, que foram sintetizados à medida que a seqüência do gene *orfatrAb* foi sendo obtida (conforme demonstrado na Figura 2). O seqüenciamento automático foi realizado no *MegaBACE DNA Analysis System* (Amersham Biosciences) e utilizado para a finalização da seqüência do gene, com o oligonucleotídeo: Atrseq: 5' TTC CTG GTG GAC GGC AGC AA 3'.

Análise das seqüências de nucleotídeos e aminoácidos obtidas

As seqüências obtidas foram analisadas utilizando-se os programas BLASTN (<http://www.ncbi.nlm.nih.gov>), ExpASy (*Expert Protein Analysis System* – <http://us.expasy.org>;

Gasteiger *et al.* 2003), ClustalX versão 1.81 (Thompson *et al.* 1997) e BioEdit versão 5.0.6 (ftp://iubio.bio.indiana.edu/molbio/seqpup). Os parâmetros utilizados nas diferentes análises em todos os programas foram sempre os do modo “*default*” de cada programa.

```

ttc cag gaa gcc gga caa ccg ggg ncn aag gaa gaa tgn tgg gtc 47
gcc ctc ccc caa tgt att cgt cgg cgn cgg gtn caa ctg cgg cct 92
cca tca att cgg caa ggc gtt gcg ccc gtg cca atg atg ctg cga 137
tgc nca tgt cgt cca ggc gcc cgc aga tgt cga tca cct gat cgc 182
ggg gag cgc cgg ctn atg ctg atc ggg cat tcg gaa ccc cct cct 227
tcc aaa ttc cgc atc ccc acg att agc ggc ccg tct ccc ggg ccg 272
gcg caa cgg cgc aag ggt gga agc ccc gcc gca aca agg ggg atg 317
ctc tct ccc gag tcc cct ggt ccc gca ccg gcc ctc gat tgt cct 362
ccg agc aaa tcc agc gcg gcg acg tcg tcc ttc ggt tgc gaa ata 407
ggc aaa aat ccc aca atg cgg gat aag caa atg tgc gct ctt cgc 452
agc aca acg gaa aat tct tac caa gct tgg ctc ttt ttc gtg cgg 497
tga gcc ttt ggt tgg ggg tgc gtt cgc cgc gcg ttc ccg gta taa 542
atg agc atg gcg gct atg cgc cgt ctt cag aat aaa cgg gag gct 587
ccc cgc cag ggc gcg ggc agc aag aca gcc atg atg gac tgg gac 632
                                     M D W D . 3
aag ctt cgg gtc ttc cac gcc gtg gcc gag gct gga agt ttc acg 677
K* L* R* V F* H A V* A E* A G S* F T 18
cac gcc ggc gag acg ctg aac ctc agc cag tcg gcg gtc agc cgc 722
H A* G E* T L* N* L* S* Q* S* A* V* S* R 33
cag atc agc gcg ctg gag gag agc ctc ggc gtg ccg ctg ttt cac 767
Q* I* S A L* E* E* S L* G* P L* F* H 48
cgg cac gcg cgc ggc ctg atc ctc cgg atc ctg acc gag cag gcc 812

```

Continuação...

R* H A R G L* I L R I L T E Q G	63
gag ctt ctg cac cgc acc gcc cgc gat <u>gtc ttc gcc aag ctg tcg</u>	857
E L L H R T A R D V F A K L S	78
<u>atg</u> acc gaa gcg atg ctg acc gaa agc cgg gag cat ccg aag ggt	902
M T E A M L T E S R E H P K G	93
ccg ctg cgc gtg acc acc acg gtg gcc ttc ggc tcg acc tgg ctg	947
P L R V T T T V A F G S T W L	108
acc ccg cgc gtc aac gag ttc ctg tca atc tac ccg gac atc cag	992
T P R V N E F L S I Y P D I Q	123
ctc acc ctg ctg atc gat gac aat gaa ctg gat ctg gcg atg cgc	1037
L T L L I D D N E L D L A M R	138
gag gcg gac atc gcc atc cgc atg aac acg ccg cgc cag ccc gac	1082
E A D I A I R M N T P R Q P D	153
ctg atc cag cgc cac ctg atg tcg gtg cat ttc cat ctc tac gcc	1127
L I Q R H L M S V H F H L Y A	168
cac cag gag tat ctg aag aag cgc ggc gcc ccg aag atc ccc gcc	1172
H Q E Y L K K R G A P K I P A	183
gac ctg gac ggg cac gac atc gtc gcc tat ccg ccg gac gtc cgc	1217
D L D G H D I V A Y P P D V R	198
gcg ccc atc gcc aac gtc aac tgg atc aac gag gtg ggc gac ccg	1262
A P I A N V N W I N E V G D P	213
gca ccg ggc gtg cgc aag ccg att ctg cag gtc aac agc atc tat	1307
A P G V R K P I L Q V N S I Y	228
gcc atg tac cgc gcc gtg aga gcg gct ggg cat cgc cgc ctg ccc	1352
A M Y R A V R A A G H R R L P	243
gac <u>ttc ctg gtg gac ggc agc aag</u> gag ctg acg cgc gtc ctg ccg	1397
D F L V D G S K E L T R V L P	258
gac gtg gac ggc ctc aag gtc gac gcc tac ttc gtc tat gcc gag	1442
D V D G L K V D A Y F V Y A E	273

Continuação...

gag ttg cgt cat tcc aag cgc atc gcg gtg ttc cgc gat ttc ctc	1487
E L R H S K R I A V F R D F L	288
gtg aag aag gtg gcg gag tcg gcc ttc tga cca agc tca gcc ttt	1532
V K K V A E S A F	297
<u>ggc nca gac atg atg ttn gcc</u> tcg ttt tna tnc acn gcc gtn	1574

Figura 2 – Seqüência completa do gene *orfatrAb*. Em letras maiúsculas está a seqüência de aminoácidos, deduzida a partir da seqüência de nucleotídeos, em código de uma letra. Em negrito estão mostrados o códon de iniciação da tradução (ATG) e o códon de Terminação (TGA); sublinhados e em negrito estão a seqüência do possível promotor σ^{54} da RNA-polimerase (GG-N₁₀-GC), o sítio de ligação ao ribossomo (RBS: AGC AAG A) e a possível seqüência da região de terminação da transcrição (GGC NCA GAC ATG ATG TTN GCC), respectivamente. Sublinhados e em itálico estão demonstrados os oligonucleotídeos utilizados para o seqüenciamento dessa região de DNA. Os asteriscos (*) indicam os resíduos que apresentam similaridade com a porção N-terminal das proteínas ATRs da família LysR, onde se localiza o motivo hélice-volta-hélice (Schell 1993).

Resultados e discussão

Seqüência do gene *orfatrAb*

A Figura 2 mostra a seqüência completa do gene *orfatrAb*, que, da região promotora até o códon de terminação da proteína correspondente, apresenta em torno de 1.500 pares de base. A seqüência de aminoácidos, deduzida a partir da seqüência de nucleotídeos corresponde a uma proteína de 297 aminoácidos, a qual foi denominada ORFatrAb.

Através da análise computacional da seqüência de aminoácidos de ORFatrAb foram deduzidos o peso molecular desta proteína, 33.581,6 kDa, o ponto isoelétrico (pI) de 7,26 e o índice de instabilidade (II) de 33,29, o que classifica essa proteína como sendo estável (de acordo com o programa ExpASy).

Na região promotora (5') do gene *orfatrAb* foi encontrada uma região de ligação para o fator sigma alternativo da RNA-polimerase σ^{54} , o códon de iniciação da tradução (ATG) da ORFatrAb e o sítio de ligação ao ribossomo (RBS) (Figura 2). O fato de apenas uma região com a seqüência de reconhecimento para o fator σ^{54} ter sido identificada nessa seqüên-

cia sugere que este gene seja transcrito sob o controle do promotor reconhecido por σ^{54} (Barrios *et al.* 1999). Os genes que estão sob o controle de σ^{54} , em *Azospirillum brasilense*, geralmente estão envolvidos no processo de fixação biológica do nitrogênio, embora outros grupos de genes já tenham sido descritos como sendo regulados por este mesmo promotor, como alguns dos genes envolvidos na síntese do flagelo, por exemplo, o gene *lafI* (revisado em Revers 1996). Além das regiões anteriormente descritas também foram encontradas, na região 5' do gene *orfatrAb*, seqüências para a ligação de uma outra proteína reguladora, NtrC, que faz parte do sistema global de regulação do nitrogênio, cuja seqüência de reconhecimento consensual é: GCAC-N₅-TGGTGCA (Merrick & Edwards 1995; Steenhoudt & Vanderleyden 2000; Pedrosa *et al.* 2001). Nesse sistema, a proteína NtrC é fosforilada por um outro componente do sistema (NtrB) em um de seus domínios, ligando-se ao DNA e ativando a transcrição de diversos genes, como, por exemplo, reguladores da fixação do nitrogênio, *nifLA*, em *Klebsiella pneumoniae*, e o gene de regulação do nitrogênio, *nac*, em *K. aerogenes* (Merrick & Edwards 1995).

Na seqüência de aminoácidos da ORFatrAb podem ser encontradas algumas

das características comuns aos membros da família de ativadores de transcrição LysR, como a proporção desigual dos aminoácidos lisina (K) e arginina (R) (Viale *et al.* 1991), que possivelmente é resultante do alto conteúdo de G+C da seqüência nucleotídica dos genes. Uma vez que os códons de lisina são compostos apenas de A+T e os de arginina de G+C, isso gera a desproporção na seqüência de aminoácidos (Henikoff *et al.* 1988; Schell 1993).

Outra característica comum aos membros da família LysR é a capacidade de regular a expressão de genes com sentidos de transcrição contrários ao seu (Henikoff *et al.* 1988; Schell 1993). Como pode ser visto na Figura 1, esta característica também é encontrada na ORFatrAb, onde o possível alvo da regulação dessa proteína é o gene *orf281* (Revers 1996; Revers *et al.* 2000; Blaha 2001). A análise da região promotora do gene *orf281* demonstrou a presença de sítios para ligação de diversas proteínas reguladoras, incluindo um possível sítio para a ligação da proteína Nac, membro da família LysR (Bender 1991; Schwacha & Bender 1993; Pomposiello & Bender 1995), assim como a seqüência consensual para a ligação de outras proteínas da família LysR (Schell 1993).

Comparação da seqüência de aminoácidos da ORFatrAb com outras proteínas da família LysR

A comparação da seqüência de aminoácidos deduzida do gene *orfatrAb* com as seqüências de aminoácidos de proteínas pertencentes à família LysR de outros microrganismos revelou uma grande similaridade entre as seqüências, principalmente na região amino-terminal, onde, nos membros desta família, a conservação é maior (Henikoff *et al.* 1988; Viale *et al.* 1991; Schell 1993), conforme está demonstrado na Figura 3. É nesta região, altamente conservada, que se encontra o motivo de ligação ao DNA, hélice-volta-hélice (Henikoff *et al.* 1988; Viale *et al.* 1991; Schell 1993).

Na região carboxi-terminal, aproximadamente entre os resíduos 236-246, que é muito variável entre os membros dessa família, existe apenas uma região que demonstra alguma similaridade de seqüência, onde provavelmente ocorre a ligação da molécula indutora (co-indutor). Foi demonstrado que mutações nessa região levam à perda de função em diversas ATRs (Schell 1993).

As diferentes ATRs, nos mais diversos microrganismos regulam diferentes tipos de genes (Henikoff *et al.* 1988; Schell 1993). Essas proteínas podem regular genes envolvidos na biossíntese de aminoácidos, como no caso de *metR*, *ilvY*, *cysB*, em *Escherichia coli* (Schell 1993) e *cysB* em *Klebsiella aerogenes* (Tyrrell *et al.* 1997), na nodulação de bactérias simbióticas, como *nodD* em *Rhizobium Sinorhizobium*, *Mesorhizobium* e *Bradyrhizobium* (Loh & Stacey 2002), ou, ainda, podem regular genes responsivos ao estresse oxidativo, como *oxyR* em *Escherichia coli* (Schell 1993) e *oxyS* em *Mycobacterium tuberculosis* (Domenech *et al.* 2001).

Com base na estrutura de raios-X da proteína CysB de *K. aerogenes*, determinada por Tyrrell *et al.* (1997), a qual foi cristalizada a partir de seus resíduos 88 a 324, com uma resolução de 1,8 Å, foi possível prever a estrutura tridimensional da proteína ORFatrAb, onde foi identificado o domínio de ligação ao DNA, hélice-volta-hélice, característico de proteínas reguladoras da transcrição (Figura 4).

Apesar da estrutura descrita acima, nenhum trabalho, até o momento, descreve uma caracterização mais detalhada da seqüência reguladora de um gene correspondente a uma ATR, apresentando os sítios de proteínas ativadoras e outras seqüências de consenso. O que se encontra na literatura são os tipos de promotores aos quais as ATRs respondem, como, por exemplo, os estudos realizados com o gene *nac* de *Klebsiella aerogenes*, onde se verificou que este gene responde também ao fator σ^{54} (Bender 1991; Schwacha & Bender 1993; Pomposiello & Bender 1995).

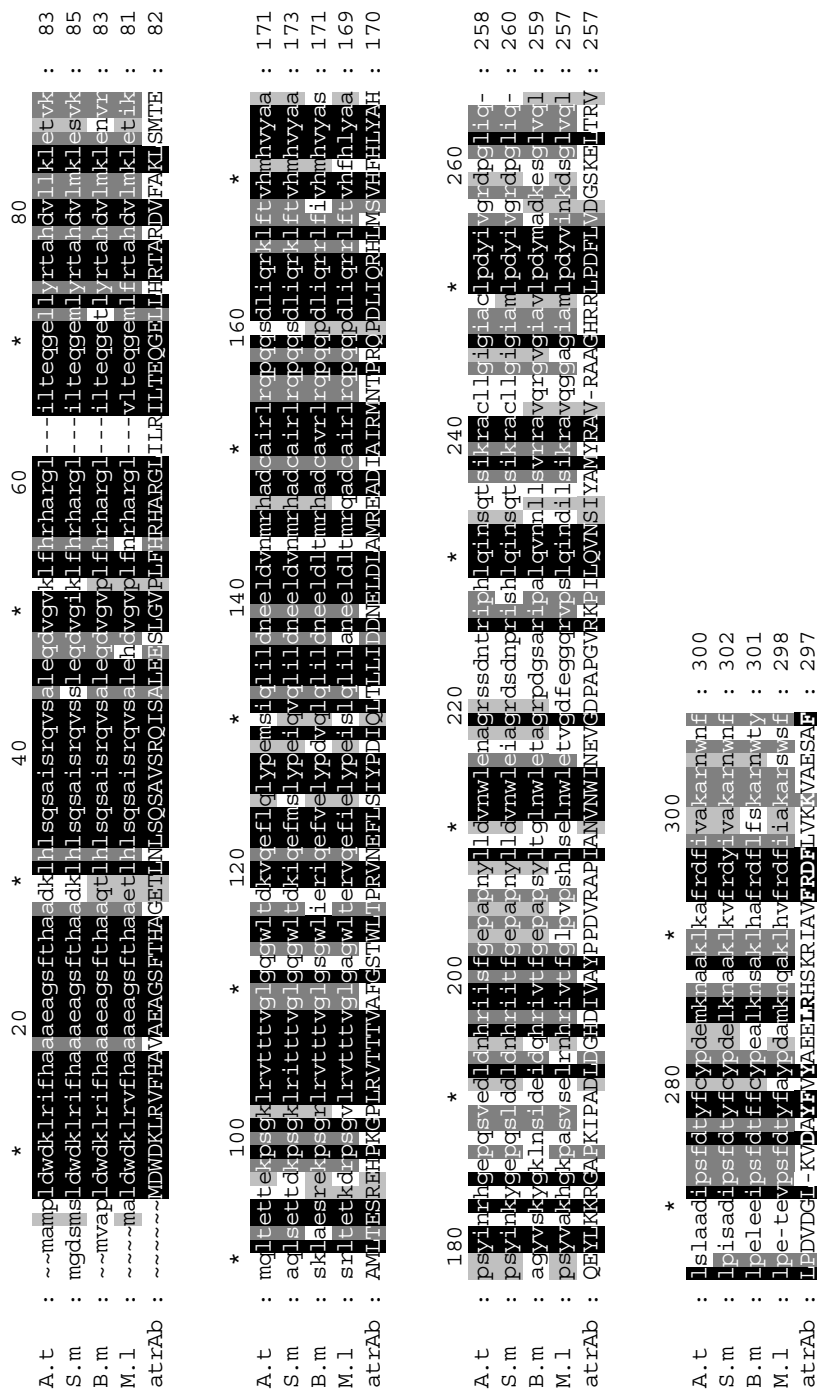


Figura 3 – Comparação da seqüência de aminoácidos da ORFatrAb com seqüências de aminoácidos de outros membros da família LysR de outros microrganismos, disponíveis no GenBank. A. t: *Agrobacterium tumefaciens* (nº de acesso: NP_355147); S. m: *Sinorhizobium meliloti* (nº de acesso NP_385672); B. m: *Brucella melitensis* (nº de acesso NP_539430); M. l: *Mesorhizobium loti* (nº de acesso NP_103870). Os blocos de aminoácidos em preto significam 100% de similaridade; blocos cinza apresentam 80% de similaridade.

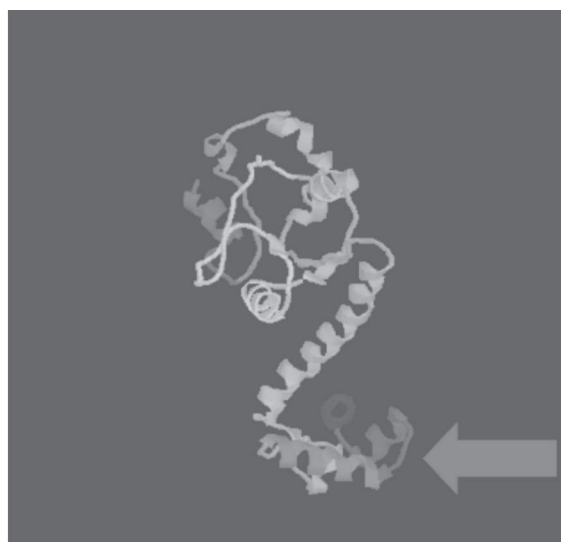


Figura 4 – Predição da estrutura tridimensional da proteína ORFatrAb utilizando o programa Swiss-Model (de acordo com o programa ExPASy). Como molde foi utilizada a proteína CysB de *Klebsiella aerogenes* já cristalografada (P45600). A seta indica a região do domínio de ligação ao DNA, hélice-volta-hélice.

Conclusões

A seqüência nucleotídica completa do gene *orfatrAb* foi obtida. Essa seqüência corresponde a aproximadamente 1.500 nucleotídeos e apresenta na região promotora um único sítio de reconhecimento para o fator σ^{54} da RNA-polimerase. Além desse sítio foram também encontradas seqüências consensuais para a ligação de proteínas ativadoras de genes envolvidos no metabolismo e fixação biológica do nitrogênio (NtrC e NifA).

A seqüência de aminoácidos deduzida da seqüência de nucleotídeos apresentou 297 resíduos. A análise desta seqüência demonstrou que a proteína correspondente possui um peso molecular de 33.581,6KDa, ponto isoelétrico de 7,26 e índice de instabilidade de 33,29, o que classifica essa proteína como sendo estável.

Na seqüência de aminoácidos foi encontrado o motivo de ligação ao DNA, hélice-volta-hélice, característico de proteínas que regulam a transcrição de determinados genes. Este motivo localiza-se na porção amino-ter-

minal da proteína, justamente onde ocorre a maior homologia com outros membros da família de ativadores de transcrição LysR. A estrutura tridimensional predita apresenta a localização desse domínio em ORFatrAb.

A proteína ORFatrAb possivelmente ativa a transcrição de uma outra ORE, ORF281, que no mutante Sp7::Tn5-33, encontra-se interrompida pelo transposon Tn5. A ORF281 já foi caracterizada como uma proteína responsiva a estresse, da família USP (do inglês *universal stress protein*).

Referências

- ARAÚJO, E. F. (1988). Caracterização de segmentos de DNA adjacentes aos genes *nifHDK* de *Azospirillum brasilense* Sp7, por mutagenese sítio-direcionada com transposon Tn5. *Tese de Doutorado*. Programa de Pós-Graduação em Genética e Biologia Molecular, UFRGS.
- ARNOLD, W.; RUMP, A.; KLIPP, W.; PRIEFER, O. & PÜHLER, A. (1988). Nucleotide sequence of a 24,206 base-pair DNA fragment carrying

- the entire nitrogen fixation cluster of *Klebsiella pneumoniae*. *Journal of Molecular Biology*, 203 (3): 715-738.
- BARRIOS, H.; VALDERRAMA, B.; MORRET, E. (1999). Compilation and analysis of σ^{54} -dependent promoter sequences. *Nucleic Acids Research*, 27 (22): 4305-4313.
- BENDER, R. A. (1991). The role of the NAC protein in the nitrogen regulation of *Klebsiella aerogenes*. *Molecular Microbiology*, 5 (11): 2575-2580.
- BLAHA, C. G. (2001). Estudos microbiológicos, fisiológicos e moleculares de um mutante de *Azospirillum brasilense* superfixador de nitrogênio: papel dos genes *orf281* e *orfatrAb*. Tese de Doutorado. Programa de Pós-Graduação em Genética e Biologia Molecular, UFRGS.
- DIXON, R. A. (1984). The genetic complexity of nitrogen fixation. *Journal of Genetics Microbiology*, 130: 2745-2755.
- DOMENECH, P.; HONORÉ, N.; HEYM, B. & COLE, S. T. (2001). Role of OxyS of *Mycobacterium tuberculosis* in oxidative stress: Overexpression confers increased sensitivity to organic hydroperoxides. *Microbes and Infection*, 3: 713-721.
- FU, H.; HARTMANN, A.; LOWERY, R. G.; FITZMAURICE, W. P.; ROBERTS, G. P. & BURRIS, R. H. (1989). Post-translational regulatory system for nitrogenase activity in *Azospirillum spp.* *Journal of Bacteriology*, 171: 4679-4685.
- GASTEIGER, E.; GATTIKER, A.; HOOGLAND, C.; IVANYI, I.; APPEL, R.D. & BAIROCH, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31: 3784-3788.
- HENIKOFF, S.; HAUGHN, G. W.; CALVO, J. M. & WALLACE, J. C. (1988). A large family of bacterial activator proteins. *Proceedings of the National Academy of Sciences*, 85: 6602-6606.
- KIRCHHOF, G.; SCHLOTTER, M.; ABMUS, B. & HARTMANN, A. (1997). Molecular microbial ecology approaches applied to diazotrophs associated with non-legumes. *Soil Biology Biochemistry*, 29 (5/6): 853-862.
- LEWIN, B. (2001). *Genes VII*. Ed. Artes Médicas, Porto Alegre, 955 p.
- LOH, J. & STACEY, G. (2002). Nodulation gene regulation in *Bradyrhizobium japonicum*: a unique integration of global regulatory circuits. *Applied and Environmental Microbiology*, 69 (1): 10-17.
- MARTIN-DIDONET, C. C.; CHUBATSU, L. S.; SOUZA, E. M.; KLEINA, M.; REGO, F. G.; RIGO, L. U.; YATES, M. G. & PEDROSA, F. O. (1999). Genome structure of the genus *Azospirillum*. *Journal of Bacteriology*, 182: 4113-4116.
- MERRICK, M. J. & EDWARDS, R. A. (1995). Nitrogen control in Bacteria. *Microbiological Reviews*, 59(4): 604-622.
- MORI, S.; CASTORENO, A. & LAMMERS, P. J. (2002). Transcript levels of *rbcR1* and *ntcA*, and *rbcL/S* genes in cyanobacterium *Anabaena* sp. PCC 7120 are downregulated in response to cold and osmotic stress. *FEMS Microbiology Letters*, 213: 167-173.
- PEDROSA, F. O.; BENELLI, E. M.; YATES, M. G.; WASSEM, R.; MONTEIRO, R. A.; KLASSEN, G.; STEFFENS, M. B. R.; SOUZA, E. M.; CHUBATSU, L. S. & RIGO, L. U. (2001). Recent developments in the structural organization and regulation of nitrogen fixation genes in *Herbaspirillum seropedicae*. *Journal of Biotechnology*, 91: 189-195.
- PETERS, J. W.; FISCHER, K. & DEAN, D. R. (1995). Nitrogenase structure and function: A biochemical-genetic perspective. *Annual Reviews of Microbiology*, 49: 335-366.
- POMPSIELLO, P. J. & BENDER, R. A. (1995). Activation of the *Escherichia coli lacZ* promoter by the *Klebsiella aerogenes* Nitrogen Control Assimilation (NAC), a LysR family transcription factor. *Journal of Bacteriology*, 177 (16): 4820-4824.

- POTRICH, D. P.; PASSAGLIA, L. M. P. & SCHRANK, I. S. (2001). Partial characterization of *nif* genes from bacterium *Azospirillum amazonense*. *Brazilian Journal of Medical and Biology Research*, 34: 1105-1113.
- REVERS, L. F. (1996). Análise molecular de mutantes de *Azospirillum brasilense* Sp7 com alta atividade de nitrogenase obtidos por mutagênese com transposon Tn5. *Dissertação de Mestrado*. Programa de Pós-Graduação em Genética e Biologia Molecular, UFRGS.
- REVERS, L. F.; PASSAGLIA, L. M. P.; MARCHAL, K.; FRAZZON, J.; BLAHA, C. G. & VANDERLEYDEN, J.; SCHRANK, I. S. (2000). Characterization of an *Azospirillum brasilense* Tn5 mutant with enhanced N₂ fixation: the effect of ORF281 on *nifH* expression. *FEMS Microbiology Letters*, 183: 23-29.
- SCHELL, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annual Reviews of Microbiology*, 47: 597-626.
- SCHRANK, A. (2003). Síntese de RNA - Transcrição. In: ZAHA, A., FERREIRA, H. B. & PASSAGLIA, L. M. P. (Eds). *Biologia Molecular básica*. Ed. Mercado Aberto, Porto Alegre, 421 p.
- SCHWACHA, A. & BENDER, R. A. (1993). The *nac* (Nitrogen Assimilation Control) gene from *Klebsiella aerogenes*. *Journal of Bacteriology*, 175 (7): 2107-2115.
- STEENHOUDT, O. & VANDERLEYDEN, J. (2000). *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiology Reviews*, 24: 487-506.
- STRYER, L. (1994). *Bioquímica*. Ed. Guanabara Koogan S. A., Rio de Janeiro, 1000 p.
- THOMPSON, J. D.; GIBSON, T. J.; PLEWNIAK, F.; JEANMOUGIN, F. & HIGGING, D. G. (1997). The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality tools. *Nucleic Acids Research*, 25: 4876-4882.
- TYRRELL, R.; VERSCHUEREN, K. H.; DODSON, E. J.; MURSHUDOV, G. N.; ADDY, C. & WILKINSON, A. J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure*, 15: 1017-1032.
- VANDE BROEK, A. & VANDERLEYDEN, J. (1995). Genetics of the *Azospirillum*-plant root association. *Critical Reviews in Plant Sciences*, 14 (5): 445-466.
- VANDE-BROEK, A.; LAMBRECHT, M.; EGGERMONT, K. & VANDERLEYDEN, J. (1999). Auxins upregulate expression of the Indole-3-Pyruvate Decarboxylase gene in *Azospirillum brasilense*. *Journal of Bacteriology*, 181 (4): 1338-1342.
- VIALE, A. M.; KOBAYASHI, H.; AKASAWA, T. & HENIKOFF, S. (1991). *rbcR*, a gene coding for a member of the LysR family of transcriptional regulators, is located upstream of the expressed set of Ribulose 1,5-Biphosphate Carboxylase/Oxygenase genes in the photosynthetic bacterium *Chromatium vinosum*. *Journal of Bacteriology*, 173 (16): 5224-5229.
- ZHANG, Y.; BURRIS, R. H.; LUDDEN, P. W. & ROBERTS, G. P. (1997). Regulation of nitrogen fixation in *Azospirillum brasilense*. *FEMS Microbiology Letters*, 152: 195-204.
- ZIMMER, W.; WESCHE, M. & TIMMERMANS, L. (1998). Identification and isolation of the Indole-3-Pyruvate Decarboxylase gene from *Azospirillum brasilense* Sp7: sequencing and functional analysis of the gene locus. *Current Microbiology*, 36(6): 327-331.

Recebido em 25/06/2004

Aceito em 27/07/2004