Nucleotide sequence and characterization of the *Rhodobacter sphaeroides* *glnB* and *glnA* genes

V. Zinchenko, Y. Churin, V. Shestopalov and S. Shestakov

Author for correspondence: V. Zinchenko. Tel: +7 095 939 4555. Fax: +7 095 939 5022.
e-mail: libro(a)GeneBec.msu.su subject: for Zinchenko

Department of Genetics,
Biology Division, Moscow
State University, Moscow,
119899, Russia

The *glnA* gene of *Rhodobacter sphaeroides* encoding glutamine synthetase (GS) has been cloned and sequenced. Molecular analysis revealed that there is a *glnB* gene upstream of *glnA*, in a single *glnBA* operon. A putative *glnAp1*-type promoter sequence, a consensus *ntrC* gene product binding site and a consensus upstream activator sequence were detected upstream of the *glnB* gene. The deduced amino acid sequences of the GS and GlnB proteins of *R. sphaeroides* showed strong homology with the same proteins from other Gram-negative bacteria. The sequence of the *glnA* gene isolated from glutamine auxotroph Gln83 was also determined. The *glnA83* mutation was shown to result in premature termination of GS synthesis and formation of a 17 kDa C-truncated GS which could be complemented by a 5'-truncated *glnA* gene which encodes a 30 kDa N-truncated GS. This phenomenon is characteristic for interallelic complementation.

Keywords: *Rhodobacter sphaeroides*, glutamine synthetase, *glnA*, *glnB*

INTRODUCTION

The purple non-sulphur photosynthetic bacterium *Rhodobacter sphaeroides* fixes nitrogen microaerobically or anaerobically when the level of intracellular fixed nitrogen is limited. Glutamine synthetase (GS; EC 6.3.1.2) plays a key role in nitrogen assimilation. In enteric bacteria, the gene *glnA* encoding GS is a member of the *glnA*–*ntrBC* (*glnALG*) operon and is controlled by a complex nitrogen-regulatory system (*ntr*) that coordinately regulates a number of genes involved in nitrogen metabolism (Gussin *et al.*, 1986; Leonardo & Goldberg, 1980; Pahel *et al.*, 1982; Reitzer & Magasanik, 1985). For transcription, activation of *glnA*, the *ntrA* (*glnF*) and *ntrC* (*glnG*) gene products is required.

In *Rhodobacter capsulatus*, closely related to *R. sphaeroides*, the *glnA* gene is a member of the *glnBA* operon and is regulated in a manner fundamentally different from that in the *Enterobacteriaceae* (Kranz *et al.*, 1990). Gene *nifR4*, the *ntrA* equivalent in *R. capsulatus*, is not required for activation of *glnBA*. *R. capsulatus* gene *nifR1*, which was shown to be homologous to *ntrC* and which is required

for transcription of nitrogen fixation genes, was responsible for approximately 50% of the *glnBA* transcription. The *glnB* gene and the 5'-terminal region of the *glnA* gene from *R. capsulatus* have been sequenced and a consensus *ntrA*-binding site was not found (Kranz *et al.*, 1990). It is possible that some genes, such as *glnA*, commonly referred to as *ntr* or nitrogen-regulated genes in the *Enterobacteriaceae*, are only partially nitrogen sensitive in photosynthetic bacteria.

In this paper we report the cloning and sequencing of *R. sphaeroides* *glnB* and *glnA* genes, and their organization as an operon unit. The role of several characteristic sequences that have been identified upstream of the *glnBA* operon in regulation of *glnBA* expression are discussed.

METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids utilized in this study are listed in Table 1.

Media and growth conditions. *Escherichia coli* strains were grown either in complex LB medium or in minimal A medium (Miller, 1972) with appropriate supplements. *R. sphaeroides* strains were grown in Ormerod medium (Ormerod *et al.*, 1961) containing 30 mM sodium DL-malate and 10 mM glutamine as sole nitrogen source. For growth tests, glutamine was replaced by 10 mM (NH₄)₂SO₄ or 10 mM glutamate. For determination of enzyme activity, *R. sphaeroides* was grown under anaerobic conditions and illuminated with a bank of three 60W in-

Abbreviations: GS, glutamine synthetase; UAS, upstream activator sequence.

The GenBank/EMBL/DBJ accession number for the sequence reported in this paper is X71659.

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant properties*	Source or reference
<i>Escherichia coli</i>		
C600	<i>thi-1 thr-1 leuB6</i>	Maniatis <i>et al.</i> (1982)
P3478	<i>thy polA1</i>	Saano & Zinchenko (1987)
POII1734	MC1040 (Mu _{cts}) with Mu dII1734 <i>lac</i> ⁺	Castilho <i>et al.</i> (1984)
M8820 Mu	M8820 with Mu	Castilho <i>et al.</i> (1984)
M8820rif Mu	As M8820 Mu but rifampicin resistant	This work
P678-54	<i>thr ara leuB1 supE lacY minA minB</i>	Adler <i>et al.</i> (1967)
ET8894	(<i>rba-glnA-ntrBC</i>) 1703::Mu _{cts} 62	Hawkes <i>et al.</i> (1985)
<i>Rhodobacter sphaeroides</i>		
2R	Wild-type	Laboratory collection
Gln83	<i>glnA83</i>	Shestakov <i>et al.</i> (1988)
Plasmids		
pTZ18R	Ap, multiple cloning site in <i>lacZ'</i>	Pharmacia LKB
RP1-7	IncP Ap Tc Tra ⁺	Laboratory collection
pAS8-121Δ3	Ap Km Ice1 Tra ⁺	Dubeikovskiy & Kameneva (1984)
pBR325	Ap Tc Cm	Bolivar (1978)
pJV27	Cm Km <i>glnA</i> (<i>R. capsulatus</i>)	Scolnik <i>et al.</i> (1983)
pVZ209	IncQ Su Sm Km Mob ⁺ , broad host-range vector	Zinchenko <i>et al.</i> (1985)
p202	IncQ Su Sm <i>glnB.A</i> (<i>R. sphaeroides</i>) Mob ⁺	This work
p152	IncQ Su Km Mob ⁺ (contains 5' truncated <i>glnA</i> gene of <i>R. sphaeroides</i>)	This work
p229	IncQ Su Km <i>glnB</i> (<i>R. sphaeroides</i>)	This work
p110EP	IncQ Su Mob ⁺ (contains 5' truncated <i>glnA</i> gene of <i>R. sphaeroides</i>)	This work
p110ES	IncQ Su Mob ⁺	This work
p171	Ap Tc (contains 5' truncated <i>glnA</i> gene of <i>R. sphaeroides</i>)	This work
p276	Ap Cm <i>glnB.A</i> (<i>R. sphaeroides</i>)	This work
p300	Ap Cm <i>glnA</i> (<i>R. sphaeroides</i>)	This work

*Ice1, immunity to colicin E1; Inc, plasmid incompatibility group; Ap, ampicillin resistance; Cm, chloramphenicol resistance; Km, kanamycin resistance; Sm, streptomycin resistance; Su, sulphonamide resistance; Tc, tetracycline resistance; Tra⁺, encodes transfer functions; Mob⁺, mobilizable plasmid. The species origin of cloned genes is shown in parentheses.

candescent lamps. Antibiotics were used at the following concentrations ($\mu\text{g ml}^{-1}$): ampicillin, 150; chloramphenicol, 25; rifampicin, 50; tetracycline, 20 for *E. coli* and 1 for *R. sphaeroides*. The concentrations of kanamycin and streptomycin for all strains were $50 \mu\text{g ml}^{-1}$. The concentration of sulfanilamide (sulfathiazole-Na) was $200 \mu\text{g ml}^{-1}$.

Cloning of the *glnA* region. Two gene banks of *R. sphaeroides* 2R in the lambda vector L47.1 (Loenen & Brammar, 1980) were constructed. Chromosomal *R. sphaeroides* DNA was completely digested with *Eco*RI or partially with *Sau*3A, then ligated to lambda L47.1 digested with *Eco*RI or *Bam*HI respectively, and packed in lambda particles *in vitro* as described by Hohn (1979). Plasmid pJV27, containing the *R. capsulatus glnA* gene, was used as a hybridization probe against the lambda banks.

Bacterial mating. This was done as described previously (Zinchenko *et al.*, 1984). Plasmid pAS8-121Δ3, which is not able to replicate in *R. sphaeroides* (Dubeikovskiy & Kameneva, 1984), was used as a conjugative mobilizing plasmid.

Mini-Mu mutagenesis. Plasmid p276 was introduced by transformation into *E. coli* strain POII1734, which contains both a Mu_{cts} prophage and the mini-Mu dII1734 (Castilho *et al.*, 1984). After thermal induction, the plasmid p276-derivatives

with mini-Mu insertions were mobilized into Mu-immune *E. coli* strain M8820rif Mu by the conjugative plasmid RP1-7. Transconjugants were selected on medium supplemented with rifampicin and kanamycin. The plasmid p276-derivatives were then transferred into *E. coli* strain C600 by transformation.

Construction of plasmid cointegrates *in vivo*. It has been shown previously that cointegrates between pBR325 and RP1-derivatives are formed by reciprocal recombination and may be selected by using strains in which the plasmids based on ColE1 replicon are not maintained (Zinchenko *et al.*, 1984). Plasmids p171, p300, p276 and p276-derivatives with mini-Mu insertions were mobilized from *E. coli* strain C600 by the conjugative plasmid RP1-7 into *E. coli* strain P3478 containing the *polA1* mutation that prevented replication of pBR325. Agarose gel electrophoresis of plasmid DNA from transconjugants confirmed the presence of cointegrates. Transconjugants in which mobilized plasmids existed in an autonomous state were never obtained.

Routine molecular biology techniques. Plasmid DNA isolation, restriction endonuclease digestion, agarose gel electrophoresis, ligation, nick-translation and Southern blot hybridization were performed using established techniques (Maniatis *et al.*, 1982). DNA sequences were obtained by dideoxy sequencing using Sequenase (USB). The problem of

Fig. 1. Physical map of the cloned *R. sphaeroides* DNA in recombinant phages and plasmids and complementation analysis of the mutant Gln83. The upper line indicates the DNA region in which *glnB* and *glnA* genes are located. The positions of genes shown are based on sequence analysis. Phages and plasmids are described in the text. Fragments of lambda phage L47.1 are indicated by double lines. Cloning vectors are shown in brackets. The capacity of the different DNA fragments to complement strain Gln83 for prototrophic growth, for restoration of GS activity and for restoration of nitrogenase activity to be repressed by ammonium ions is shown to the right of the figure: +, complementation; -, no complementation. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; P, *Pvu*II; S, *Sma*I; Sa, *Sau*3A. Only relevant *Sau*3A sites are shown.

Fig. 2. Location of mini-Mu dII1734 insertion sites in the 4.79 kb *Bam*HI-*Hind*III fragment of p276 and complementation analysis of the mutant Gln83. Numbers refer to the distance in kb from the *Bam*HI site. The arrows show the direction of transcription of the *lacZ* gene in mini-Mu insertions. Other abbreviations are the same as in Fig. 1.

band compressions in the sequencing gel was overcome by making use of deaza-dGTP (Pharmacia LKB). Minicells were used to visualize labelled proteins synthesized by various plasmids as described by Allibert *et al.* (1987).

Enzyme assays. Nitrogenase activity was determined by measurement of acetylene reduction as described by Willison & Vignais (1982). GS was determined by the γ -glutamyltransferase assay (Shapiro & Stadtman, 1970). The protein content of

GGATCGCCCCGCTCCACCTCCGAGTGCGGTGAGCGGACGGCAGCTGCCGCCCTCGTCCGCCAGCGCCGCGCATTCGTGACGGCATCCGTGCCCTG 100
 BamHI PvuII
 TCGGGCGCGCACGGTCCGGGCGGGCGGAGCTCGGGGAACACCGCTCCCAACCCCAAGGCGAAGACCTGAACCTGCCAGCCGCGCCGCCAGCAGACGC 200
 GCCACCACGAAGCGTCGCCCCGTGTGTGCGGGGCGGAGAGGACGACCGCCAGCGCGCGGGCCAGTCCGGAACGGCGCGCCGCCACCGCTCTC 300
 CCGCGCGCTCCATCAGCTGCAGTCCGGTCCGCCGTCCAGCGGCTATCGCCGGTCTCAATGGCGCGCATTTGGGCAGACGTGACAAAGTTGTACCACTTC 400
 PvuII -35 -10
 CGGGGCTGTTCGCTTCATATTCGCTCATATTTACGCCGATTCACAAAAATAGGGCGCACCCCGAGATTCCTCCCTCGAGATGGACAGCAGCGACCC 500
 glnB ---->
 V K K I E A I I K P F
 TACCCAATGTTCCTCGGACGAGAAGTGGTCACTGGCTTTCCGACAAGGCAGAGGGGAGTCGAGCCGTGAAGAAGATCGAGGCGATCATCAAGCGTT 600
 K L D E V K E A L Q A A G V Q G L S V T E V K G F G R Q K G H T E
 CAAGCTCGACGAAGTGAAGAGGCGCTGCAGGCGCGGGCTGACGGGACTTTGGTGACCGAGGTCAAGGGCTTCGGACGCCAGAAGGCCACACCGAG 700
 L Y R G A A Y V V D F L P K V K I E V V L A D D M V E A A V E A I V
 CTTTACCGCGGGCGGCATATGTCTGGACTTCCTCCGAAGGTGAAGATCGAGTCTGCTCCGCGGACGATATGGTCGAGGCCGCGGTGCAAGCGATCG 800
 S A S R T D K I G D G K I F I S P V E Q A I R I R T G E T G E D A
 TCTCCGCTCGCGACCCGACAAGATCGGCGAGGCAAGATCTTATTAGCCCGTTCGAGCAGGCGATCCGCATCCGACCGCGGAGACCGGGCAAGACGC 900
 BglII
 V * glnA ---->
 M S K V
 GGTCTGACCCCGGACCCGACGACGACCATGACGGGCGGGACCGGACGACCTGGCCCTCCGACATCGATCCTTCAGCAAGGGAAGACAATGAGCAAGT 1000
 A D A L K L M K D E E V E Y V D I R F T D P R G K L Q H V T L V A
 AGCTGATGCTCTCAAACTGATGAAGACGAGGAGGTCAATATGTGACATCCGCTTCACCGACCCGCGGGCAAGCTCCAGCATGTGACGCTGGTGGCC 1100
 D L V D E D F F E E G F M F D G S S I A G W K S I D Q S D M K L I P
 GATCTCGTCGACGAGGACTTCTTCGAGGAAGGCTTCATGTTGACGGCTCGTCCATCGCCGGCTGGAAGTCGATCGACGAGTCCGACATGAAGCTGATCC 1200
 D A G S V Y I D P F Y A E K T L C V H C N V V E P D T G E A Y S R
 CGGATGCGGGCTCGGTCTACATCGACCGGTTCTATGCCGAGAAGACGCTCTGCGTGCAATGCAACGTCGTCGAGCCCGACACTGGCGAGGCGCTACAGCCG 1300
 D P R G A A V K A E A Y L K A S G I G D V A Y F G P E A E F F I F
 CGACCCGCGCGCGCGGCGGCTCAAGGCGGAAGGCTACCTGAAGGCTTCGGGCATCGGTGACGTGGGCTATTTTCGGGCGGAAGCGAATTCCTCATCTTC 1400
 EcoRI
 ORF3 ---->
 D D V R Y S V T P A K V A Y Q I D A D A G A W N T D S E Y E M G N L
 GACGACGTGCGTATTTCGGTGACCGCGCGGAAGGTGGCTACAGATCGACGCCGACCGCGCGGTGGAACACCGACTCGGAATATGAGATGGGCAACC 1500
 A (glnA83)
 A H R A G H K G G Y F P V N P I D E A Q D L R G E M L S T M K R M
 TCGCCACCGCGCGCGGCACAAAGGGCGGTACTTCCGGTGAACCGATCGACGAGGCGCAGGACCTGCGCGCGGAGATGCTCTCGACCATGAAGCGCAT 1600
 G M K V D K H H E V A T C Q H E L G L I F G G L T E Q A D N I L
 GGGCATGAAGGTGGACAAGCACCAACACGAAGTGGCCACCTGCCAGCAGCAACTCGGGCTGATCTTCGGCGGGCTCACCAGCAGGCGGCAACATCTG 1700
 K Y K Y V I H N V A G M H G K T V T F M P K P M K G D N G S G M H V
 AAATACAGTATGTATCCACAACGTCGCGGCGATGACGGCAAGACCGTGACCTTCATGCCGAAGCCATGAAGGGCGACACGGCTCGGGGATGACG 1800
 N M S I W K E Q A L F A G D K Y A D L S Q E A L W F I G G I L K Q
 TCAACATGTGATTCGAAGGAGCAAGCGCTCTCGCTGGCGACAAATATCCGACCTGAGCCAGGAAGCCCTGTGGTTCATCGCGCGCATCTGAAGCA 1900
 P S V N A L T N P A T N S Y K R L I P G F E A P V L R A Y S A R N
 GCCAAGCGTGAACGCGTACGAACCCGGCCACCAACAGCTACAAGCGGTGATCCCGGCTTCGAGGCTCCGGTGTGCGCGCTATTTCGGCCCGCAAC 2000
 SmaI
 R S G C V R I P W T E S P N A K R V E A R F P D P S A N P Y L A F A
 CGCTCGGGCTCGGTCGCTATTCGTTGGACCGAGTCCGCGAACGCCAAGCGCGTGAAGCCCGTTCCCGATCCGTGCGCGAACCCGTATCTCGCTTCG 2100
 A L L M A G L D G I K N K I D P G P A S D K D L Y D L P P E E L A
 CGGCGCTGCTGATGGCCGGCTCGACGGCATCAAGAACAAGATCGATCCGGGCGGGCATCGGACAAGGACCTCTACGACCTGCCCGCGGAAGAACTGGC 2200
 A I P T V C G S L R E A L E E L E K D H D F L L A G D V F T K D Q
 CGCCATCCGACCGCTGCGGCTCGCTCCGCGAGGCGCTGAAGAGCTGGGAAGGACACGACTTCCTGCTCGCCGGCGACGTGTTACCAAGGACGAG 2300
 L E G Y M A L K W E E V Y A Y E H T P H P V E Y Q M Y Y S C *
 CTCGAGGGCTACATGGCGCTCAAGTGGGAAGAGGTCTACGCTACGAGCACACCGCGACCCGGTGAATATCAGATGTACTACAGCTGCTGATCCAGGC 2400
 PvuII
 GCGAGGGACAGGAAGGGCGGCTCCGGGCGTTCTTCTTGGCGTTACGACGAAGAGGGT 2460

whole cells was measured by the Lowry method after lysis of cells with 0.2 M NaOH for 1 min in a boiling water bath.

RESULTS

Cloning and localization of the *glnA* region

Two recombinant phages, E91 and E101, containing different *EcoRI* fragments of 3.4 and 4.4 kb, and one recombinant phage, S215, containing a *Sau3A* fragment of about 12.6 kb, were isolated from *R. sphaeroides* gene banks, as described in Methods. Restriction and hybridization analyses revealed that the *Sau3A* fragment completely overlaps the 4.4 kb *EcoRI* fragment and partially overlaps the 3.4 kb *EcoRI* fragment (Fig. 1).

To localize the *glnA* gene within the cloned fragments, different deletion and insertional derivatives of these fragments were generated and their ability to restore the wild phenotype of *R. sphaeroides* glutamine-dependent mutant Gln83 was tested. This mutant with a lesion in the *glnA* gene has a complex phenotype: absence of GS activity, derepressed nitrogenase synthesis and nitrogenase activity in the presence of ammonium ions (Zinchenko *et al.*, 1991).

For deletion analysis, the 3.4 and 4.4 kb *EcoRI* fragments of recombinant phages E91 and E101 respectively, and the 10.9 kb *HindIII* fragment of the recombinant phage S215 and its subfragments were recloned into the multicopy broad host range vector pVZ209 (Fig. 1); the resultant plasmids were transferred by mobilization into *R. sphaeroides* Gln83 and *E. coli* ET8894. As shown in Fig. 1, the 1.0 kb *EcoRI*–*PvuII* subfragment (p110EP) was the minimal fragment tested that was sufficient for the restoration of the wild phenotype of the mutant Gln83. Neither plasmid restored the prototrophic phenotype of *E. coli* ET8894 (data not shown).

The molecular mass of the *R. sphaeroides* GS subunit, determined by SDS-gel electrophoresis is 50 kDa (Engelhardt & Klemme, 1982). Therefore, the *R. sphaeroides glnA* gene should encompass no less than 1.35 kb. Thus, these data indicate that the complementing 3.4 kb *EcoRI* and 1.0 *EcoRI*–*PvuII* fragments could comprise a truncated *glnA* gene.

For insertional analysis, the phage mini-MuIII1734 was used. The instability of mini-Mu insertional derivatives of recombinant plasmids based on pVZ209 was revealed in preliminary experiments. Therefore, for insertional analysis, the 3.4 kb *EcoRI* fragment from plasmid p152, and the 4.79 kb *BamHI*–*HindIII* and 3.95 kb *BglII*–*HindIII* subfragments from plasmid p202 were recloned into the vector pBR325 resulting in plasmids p171, p276 and p300

(Fig. 1). Although plasmids with ColE1-like replicons are not maintained in *R. sphaeroides*, the inheritance of these plasmids may be obtained by cointegration with mobilizing plasmids (Zinchenko *et al.*, 1984). Cointegrates between plasmids p171, p276 and p300 and the conjugative plasmid RP1-7 were made, as described in Methods, and were introduced into strain Gln83. As shown in Fig. 1, mutant Gln83 was complemented by p276, but not by p171 or p300. In further mini-Mu insertional mutagenesis experiments, only plasmid p276 was used. From 110 mini-Mu insertions in p276, ten were chosen for further study. Others were either located in the vector or were duplicates of the ten selected. Each of these ten mini-Mu inserts was mapped by digestion with *EcoRI*, *HindIII* and *BamHI*. The results of the mapping and complementation analysis (Fig. 2) indicate that the *glnA* gene is located within the 2.2 kb region between the insertions at 0.5 and 2.7 kb.

Sequence of the *glnA* region

The fact that the 3.4 and 4.4 kb *EcoRI* fragments of the *R. sphaeroides* genome identified by hybridization with the *R. capsulatus glnA* gene are linked suggested that their joint *EcoRI* sites lie within the *glnA* gene. A DNA fragment of 2460 bp overlapping this *EcoRI* site was sequenced on both strands (Fig. 3). Two open reading frames (ORFs) were found in this sequence. ORF1, encoding a 112 amino acid polypeptide of 12067 Da, was assigned to *glnB* according to its similarity with other eubacterial GlnB amino acid sequences. The homology of the *R. sphaeroides glnB* gene product ranged from 64.3% identity with *E. coli* GlnB (Son & Rhee, 1987) to 84.8% with *R. capsulatus* GlnB (Kranz *et al.*, 1990).

Following the termination codon of ORF*glnB*, there is no inverted repeat structure which could play a role in transcription termination.

ORF2, distal to *glnB*, showed significant homology with prokaryote GS amino acid sequences. The homology of the *R. sphaeroides glnA* gene product ranged from 52.7% for the GS of *Anabaena* sp. strain 7120 (Tumer *et al.*, 1983) to 90.0% for the N-terminal region of *R. capsulatus* GS (Kranz *et al.*, 1990). ORF*glnA* encodes a 467 amino acid polypeptide of 51961 Da.

A possible *rho*-independent termination signal was found close to the 3' end of ORF*glnA* (Fig. 3). Both *glnB* and *glnA* ORFs are preceded by a putative ribosome-binding site AAAGGG.

Codon usage frequencies for ORF*glnB* and ORF*glnA* reflect a G + C content, characteristic for *R. sphaeroides*

Fig. 3. The nucleotide and deduced amino acid sequences of *glnB* and *glnA* genes from *R. sphaeroides*. Relevant restriction endonuclease recognition sites are indicated. The putative *glnAp1* type promoter sequence and the putative *ntnC* gene product binding site are indicated by double lines above and beneath sequences, respectively. The putative upstream activator sequence (UAS) is indicated by asterisks. The possible *rho*-independent terminator is indicated by overlined sequences. The putative ribosome binding sites preceding start codons are underlined. Arrows above the start codons show the direction of transcription. The *glnA83* mutation (G to A) is indicated at bp 1469.

(67%; Marmur & Doty, 1962), possessing strong bias toward the use of codons in which G and C predominate. In the *R. sphaeroides glnA* and *glnB* genes, G and C are used almost everywhere in the third position, except tyrosine, in which there is no obvious preference between TAT and TAC. Both *glnB* and *glnA* exhibit a G + C content of 63.9% and 63.1% respectively. The G + C content for the intergenic regions varies from 57.3% (position 419–568) to 65.9% (position 905–989).

The absence of complementing activity of p300, based on pBR325 and containing the distal part of *glnB* and the total sequence of *glnA* (Figs. 1 and 3), indicates the lack of suitable promoters both in the vector plasmid and in the *glnB*–*glnA* intergenic region.

Sequence analysis of *glnA* confirmed the *Eco*RI site inside the *glnA* gene (Fig. 3) and that the plasmids p152 and p110EP, which are able to complement the mutant Gln83, contain a truncated *glnA* gene. This truncated gene contains only the 3'-terminal region of *glnA* without its 397 bp 5'-terminal region. The proximal portion of the 3'-terminal region (downstream of *Eco*RI) contains several ATG codons (at positions 1491, 1578, 1590, 1599 and 1605) encoding methionine. All of these ATG codons are preceded by sequences similar to ribosome-binding sites (Fig. 3).

It is possible that one of these ATG codons may be used as a transcriptional start codon in the presence of a suitable promoter. Such a promoter in the plasmids p152 and p110EP, based on the RSF1010 replicon, may be the promoter of the *sul*II gene. Plasmid p171, based on pBR325 and containing a 3.4 kb *Eco*RI fragment similar to p152, lacks a suitable promoter.

One of the putative ORFs, corresponding to the longest possible translation, begins with ATG at position 1491 and ends with TGA at position 2391, and was designated ORF3 (Fig. 3). The molecular mass of the deduced amino acid sequence of ORF3 is 33426 Da.

The 4.4 kb *Eco*RI fragment of the plasmid p152 contains ORF3 and downstream nucleotide sequences, while the 1.0 kb *Eco*RI–*Pvu*II fragment of plasmid p110EP contains a truncated ORF3 without the last 4 bp and stop codon since the *Pvu*II cleavage site lies at a distance of 4 bp upstream of the end of ORF3. However, in the resultant plasmid p110EP, restoration of the *Pvu*II site was accompanied by elimination of the last TGC codon and formation of a new TGA codon, according to the nucleotide sequence of plasmid RSF1010 around the *Pvu*II site (Scholz *et al.*, 1989).

Identification of the plasmid p110EP-encoded proteins

The synthesis of plasmid-encoded proteins was studied in a mini-cell system. One novel polypeptide with apparent molecular mass of 30 kDa was synthesized from plasmid p110EP (as compared with the vector plasmid pVZ209), (Fig. 4). The data obtained indicate that this protein is

Fig. 4. Expression of the 5'-truncated *glnA* gene in *E. coli* minicells. Proteins were labelled with [^{35} S]methionine. Crude extracts were then analysed by SDS-PAGE, and the gel was autoradiographed. The minicell-producing strain P678-54 contained the following plasmids: lane 1, p110EP; lane 2, pVZ209. Molecular mass is indicated in kDa.

encoded by ORF3, since this would encode a protein with a predicted molecular mass of 33426 Da. This result, taken in conjunction with complementation and sequence analyses (see above), suggests that the 30 kDa protein is the C-terminal portion of GS.

Gel electrophoresis under non-denaturing conditions of total protein from *R. sphaeroides* wild-type and Gln83 cells with plasmids p110EP or p152, followed by activity staining revealed that the molecular masses (about 600 kDa) of the GS from these strains are indistinguishable (data not shown).

Analysis of the *glnBA* leader region

Several characteristic sequences have been identified upstream of the *glnB* gene (Fig. 3). The sequence GCACN₅TAGGGC, position 444 to 458, possesses 80% homology with the repressor binding site GCACN₅TGGTGC for the *ntrC* gene product (Dixon, 1984). Expression of the *glnA* gene in the *Enterobacteriaceae* is regulated at tandem promoters (Dixon, 1984; Reitzer & Magasanik, 1985). In *E. coli*, the upstream *glnAp1* promoter has a sequence similar to the –35 and –10 RNA polymerase-binding consensus sequence, and is repressed by the *ntrC* gene product and requires CAP (Reitzer & Magasanik, 1985). The repressor binding site for the *ntrC* gene product overlaps the –35 RNA polymerase contact site and the transcriptional start site of

glnAp1. The downstream promoter *glnAp2* resembles the NtrA-dependent promoters with conserved residues at -24 and -12 with respect to the transcriptional start site. In *R. sphaeroides*, the *glnB* leader region consensus sequence TTGCAC-N₂₃-TTCCCC at position 442 to 476 possesses a high degree of homology with the *E. coli glnAp1* promoter sequence (-35)TTGCAC-N₁₈-(-10)TTCCAT. The proposed -35 consensus sequence of *R. sphaeroides* is overlapped, like *glnAp1* in *E. coli*, with the *ntrC* gene product binding site.

The *R. sphaeroides* sequence TGTTG-N₆-CCGCA, localized at position 226 to 241, possesses 80 % homology to the consensus upstream activator sequence (UAS) TGTYR-N₆-YRACA of *Klebsiella pneumoniae* (Cannon *et al.*, 1991; Gussin *et al.*, 1986). The motif TCTGCCAGA proposed earlier may be involved in binding the *R. capsulatus ntrC* gene product (Preker *et al.*, 1992); however, no sequence resembling this 9 bp motif was found upstream of *glnB* and *glnA* in *R. sphaeroides*.

Analysis of the previously sequenced *R. capsulatus glnB* leader region revealed the absence of an *ntrA* consensus promoter (Kranz *et al.*, 1990). Our analysis of this region revealed characteristic sequences similar to *E. coli* -35 of *glnAp1*, *ntrC*-binding site and UAS (data not shown).

Cloning and sequencing the *glnA83* mutant allele

Fragments of approximately 3–4 kb were isolated after separation of *EcoRI*-digested chromosomal DNA from *R. sphaeroides* strain Gln83 on a 0.7 % agarose gel and ligated into *EcoRI*-digested pTZ18R. Recombinant plasmids containing fragments of *glnA83* were isolated by hybridization with the 3.4 and 4.4 kb *EcoRI* fragments of p152 and p229, respectively, as *glnA* probes.

Sequence analysis revealed that the mutation resulted in the substitution of G for A at position 1469 and in the formation of a stop codon instead of Trp at position 159 (Fig. 3). The deduced molecular mass of the truncated GS in *glnA83* was 17527 Da.

DISCUSSION

In *Escherichia coli* and *Klebsiella pneumoniae*, the *glnA* gene is a member of the *glnAntrBC* operon (Leonardo & Goldberg, 1980; Pahel *et al.*, 1982), whereas *glnB* is not linked with these genes and is located in a different region of the chromosome (Magasanik, 1982). In a number of micro-organisms – *Rhodobacter capsulatus*, *Rhizobium leguminosarum*, *Bradyrhizobium japonicum* and *Azospirillum brasilense* – the *glnA* gene is distal to *glnB* (Kranz *et al.*, 1990; Colonna-Romano *et al.*, 1987; Carlson *et al.*, 1985; de Zamaroczy *et al.*, 1990). In addition, *R. capsulatus glnB* and *glnA* genes belong to one operon, *glnBA* (Kranz *et al.*, 1990).

The absence of expression of the *glnA* gene in plasmid p300, containing a DNA fragment with a deletion of the *glnB* 5'-terminus (Figs. 1 and 3) indicates that the *R. sphaeroides glnA* and *glnB* genes are organized in a *glnBA* operon. The absence of a recognizable terminator struc-

ture between *glnA* and *glnB* supports this conclusion. Therefore, unlike the *Enterobacteriaceae*, purple bacteria and probably some other micro-organisms in which *glnB* and *glnA* genes are tightly linked may show coordinated synthesis of *glnB* and *glnA* gene products as one of the regulatory mechanisms of nitrogen metabolism.

Analysis of the upstream regions of the *R. sphaeroides* and *R. capsulatus glnB* coding regions revealed similar characteristic sequences which may take part in *glnBA* operon regulation. In the *Enterobacteriaceae*, the *glnA* gene is transcribed from tandem promoters, in which the upstream promoter *glnAp1* is negatively regulated and the downstream promoter *glnAp2* is positively regulated by the products of the *ntr* system (Dixon, 1984; Reitzer & Magasanik, 1985). In *R. sphaeroides* and *R. capsulatus*, the *glnBA* operon leader regions contain only one promoter element, similar to *glnAp1*. As in *glnAp1*, the -35 consensus of this promoter is overlapped by the binding site of the *ntrC* gene product. The absence of the NtrA-dependent promoter with conserved residues at -24 and -12 in the upstream region of *glnBA* operon is in good agreement with the data obtained on *R. capsulatus*, showing the *glnBA* operon to be NtrA (NifR4)-independent (Kranz *et al.*, 1990). The absence of the canonical *ntr*-regulated promoter is a good explanation for the lack of expression of *R. capsulatus* (Scolnik *et al.*, 1983) and *R. sphaeroides glnA* genes in *E. coli* cells. In purple bacteria, synthesis of GS is repressed when the cells are grown in media containing excess ammonia, and is derepressed in ammonia-limited conditions (Caballero *et al.*, 1985; Johansson & Gest, 1976; Sakhno *et al.*, 1981). Transcription of the *glnBA* operon is activated by the *ntrC* (*nifR1*) gene product (Kranz *et al.*, 1990; V. Zinchenko, unpublished). Based on these data, the existence of a second, positively regulated promoter may be proposed.

The functional role of the sequences possessing high degree of homology with the UAS (*nifA*-motif) of *K. pneumoniae* and localized in the *glnBA* operon upstream region remains unclear. Consistent with a DNA-looping model, the *nifA* gene product binds at the UAS and contacts the sigma⁵⁴-RNA-polymerase complex at the -24 to -12 region by looping-out of the intervening DNA (Buck *et al.*, 1987). However, no sequence similar to the -24 to -12 consensus sequence for *ntrA*-dependent promoters was identified upstream of the *R. sphaeroides glnB* gene.

The *glnA83* mutation was shown to result in premature termination of GS synthesis and formation of a 17 kDa C-truncated GS which could be complemented by a 5'-truncated *glnA* gene encoding a 30 kDa N-truncated GS. These data indicate that the N- and C-terminal fragments of *R. sphaeroides* GS, which possess a molecular mass essentially equal to the mass of the monomer, can reassemble to form the dodecamer with restoration of the active sites. This phenomenon is characteristic for interallelic complementation.

In connection with the results obtained, it is possible to note that limited proteolysis of *E. coli* GS results in the inactivation of the enzyme, while the nicked protein

remains dodecameric (Dautry-Varsat *et al.*, 1979; Lei *et al.*, 1979). The proteolysis-sensitive site is known to be the central loop region of GS (Almassy *et al.*, 1986), the same region in which the *glnA83* mutation is located in *R. sphaeroides* GS.

ACKNOWLEDGEMENTS

The authors wish to thank Dr R. Haselkorn for plasmid pJV27, Dr R. Dixon for *E. coli* strain ET8894 and Dr W. Martin for correction of the manuscript. This work was supported by grants from the Russian Fundamental Research Foundation and the Russian State Program, 'Frontiers in Genetics'.

REFERENCES

- Adler, H. I., Fisher, W. D., Cohen, A. & Hardigree, A. A. (1967). Miniature *Escherichia coli* cells deficient in DNA. *Proc Natl Acad Sci USA* **57**, 321–326.
- Allibert, P., Willison, J. C. & Vignais, P. M. (1987). Complementation of nitrogen-regulatory (*ntr*-like) mutations in *Rhodobacter capsulatus* by an *Escherichia coli* gene: cloning and sequencing of the gene and characterization of the gene product. *J Bacteriol* **169**, 260–271.
- Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N.-H. & Eisenberg, D. (1986). Novel subunit-subunit interaction in the structure of glutamine synthetase. *Nature* **323**, 304–309.
- Bolivar, F. (1978). Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique *EcoRI* sites for selection of *EcoRI* generated recombinant molecules. *Gene* **4**, 121–136.
- Buck, M., Cannon, W. & Woodcock, J. (1987). Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve DNA loop formation. *Mol Microbiol* **1**, 243–249.
- Caballero, F. J., Cejudo, F. J., Florencio, F. J., Cárdenas, J. & Castillo, F. (1985). Molecular and regulatory properties of glutamine synthetase from the phototrophic bacterium *Rhodospseudomonas capsulata* E1F1. *J Bacteriol* **162**, 804–809.
- Cannon, W., Charlton, W. & Buck, M. (1991). Organization and function of binding sites for the transcriptional activator *NifA* in *Klebsiella pneumoniae* *nifE* and *nifU* promoters. *J Mol Biol* **220**, 915–931.
- Carlson, T. A., Guierinot, M. L. & Chelm, B. K. (1985). Characterization of the gene encoding glutamine synthetase I (*glnA*) from *Bradyrhizobium japonicum*. *J Bacteriol* **162**, 698–703.
- Castilho, B. A., Olfson, P. & Casadaban, M. (1984). Plasmid insertion mutagenesis and *lac* gene fusion with mini-Mu bacteriophage transposons. *J Bacteriol* **158**, 488–495.
- Colonna-Romano, S., Riccio, A., Guida, M., Defez, R., Lamberti, A., Iaccarino, M., Arnold, W., Priefer, U. & Pühler, A. (1987). Tight linkage of *glnA* and a putative regulatory gene in *Rhizobium leguminosarum*. *Nucleic Acids Res* **15**, 1951–1964.
- Dautry-Varsat, A., Cohen, G. N. & Stadtman, E. R. (1979). Some properties of *Escherichia coli* glutamine synthetase after limited proteolysis by subtilisin. *J Biol Chem* **254**, 3124–3128.
- Dixon, R. (1984). Tandem promoters determine regulation of the *Klebsiella pneumoniae* glutamine synthetase (*glnA*) gene. *Nucleic Acids Res* **12**, 7811–7830.
- Dubeikovsky, A. N. & Kameneva, S. V. (1984). Inheritance of hybrid plasmid pAS8-121 in cells of the purple nitrogen fixing phototrophic bacterium *Rhodospseudomonas sphaeroides*. *Genetika* **20**, 1783–1791 (in Russian).
- Engelhardt, H. & Klemme, J.-H. (1982). Purification and structural properties of adenylated and deadenylated glutamine synthetase from *Rhodospseudomonas sphaeroides*. *Arch Microbiol* **133**, 202–205.
- Gussin, G. N., Ronson, C. W. & Ausubel, F. M. (1986). Regulation of nitrogen fixation genes. *Annu Rev Genet* **20**, 567–591.
- Hawkes, T., Merrick, M. & Dixon, R. (1985). Interaction of purified *NtrC* protein with nitrogen regulated promoters from *Klebsiella pneumoniae*. *Mol & Gen Genet* **201**, 492–498.
- Hohn, B. (1979). *In vitro* packaging of λ and cosmid DNA. *Methods Enzymol* **68**, 299–309.
- Johansson, B. C. & Gest, H. (1976). Inorganic nitrogen assimilation by the photosynthetic bacterium *Rhodospseudomonas capsulata*. *J Bacteriol* **128**, 683–688.
- Kranz, R. G., Page, V. M. & Caldicott, I. M. (1990). Inactivation, sequence and *lacZ*-fusion analysis of a regulatory locus required for repression of nitrogen fixation genes in *Rhodobacter capsulatus*. *J Bacteriol* **172**, 53–62.
- Lei, M., Aebi, U., Heidner, E. G. & Eisenberg, D. (1979). Limited proteolysis of glutamine synthetase is inhibited by glutamate and feedback inhibitors. *J Biol Chem* **254**, 3129–3124.
- Leonardo, J. M. & Goldberg, R. B. (1980). Regulation of nitrogen metabolism in glutamine auxotrophs of *Klebsiella pneumoniae*. *J Bacteriol* **142**, 99–110.
- Loenen, W. A. & Brammar, W. Y. (1980). A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. *Gene* **10**, 249–259.
- Magasanik, B. (1982). Genetic control of nitrogen assimilation in bacteria. *Annu Rev Genet* **16**, 135–168.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Marmur, J. & Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol* **5**, 109–118.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Ormerod, J. G., Ormerod, K. S. & Gest, H. (1961). Light-dependent utilization of organic compounds and photoproduction of molecular hydrogen by photosynthetic bacteria; relationships with nitrogen metabolism. *Arch Biochem Biophys* **94**, 449–463.
- Pahel, G., Rothstein, D. M. & Magasanik, B. (1982). Complex *glnA-glnL-glnG* operon of *Escherichia coli*. *J Bacteriol* **150**, 202–213.
- Preker, P., Hubner, P., Schmehl, M., Klipp, W. & Bickle, T. A. (1992). Mapping and characterization of the promoter elements in the regulatory *nif* genes *rpoN*, *nifA1* and *nifA2* in *Rhodobacter capsulatus*. *Mol Microbiol* **6**, 1035–1047.
- Reitzer, L. J. & Magasanik, B. (1985). Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. *Proc Natl Acad Sci USA* **82**, 1979–1983.
- Saano, A. K. & Zinchenko, V. V. (1987). A new IncQ plasmid R89S: properties and genetic organization. *Plasmid* **17**, 191–201.
- Sakhno, O. N., Ivanovskii, R. N. & Kondrat'eva, E. N. (1981). The glutamine synthetase-glutamate synthase system in *Rhodospseudomonas sphaeroides*. *Microbiologiya* **50**, 607–612 (in Russian).
- Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M. & Scherzinger, E. (1989). Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene* **75**, 271–288.
- Scolnik, P. A., Virosco, J. & Haselkorn, R. (1983). The wild-type gene for glutamine synthetase restores ammonia control of

nitrogen fixation to Gln⁻ (*glnA*) mutants of *Rhodopseudomonas capsulata*. *J Bacteriol* **155**, 180–185.

Shapiro, B. M. & Stadtman, E. R. (1970). Glutamine synthetase (*Escherichia coli*). *Methods Enzymol* **17A**, 910–922.

Shestakov, S., Zinchenko, V., Babykin, M., Kopteva, A., Kameneva, S., Frolova, V., Shestopalov, V. & Bondarenko, O. (1988). Genetic studies on the regulation of nitrogen fixation in *Rhodobacter sphaeroides*. In *Nitrogen Fixation: 100 Years After. Proceedings of the 7th International Congress on Nitrogen Fixation*, pp. 163–169. Edited by H. Bothe, F. J. de Bruijn & W. E. Newton. Stuttgart & New York: Fischer.

Son, H. S. & Rhee, S. G. (1987). Cascade control of *Escherichia coli* glutamine synthetase. Purification and properties of P_{II} protein and nucleotide sequence of its structural gene. *J Biol Chem* **262**, 8690–8695.

Tumer, N. E., Robinson, S. J. & Haselkorn, R. (1983). Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. *Nature* **306**, 337–342.

Willison, J. C. & Vignais, P. M. (1982). The use of metronidazole to isolate Nif⁻ mutants of *Rhodopseudomonas capsulata* and the

identification of a mutant with altered regulatory properties of nitrogenase. *J Gen Microbiol* **128**, 3001–3010.

de Zamaroczy, M., Delorme, F. & Elmerich, C. (1990). Characterization of three different nitrogen-regulated promoter regions for the expression of *glnB* and *glnA* in *Azospirillum brasilense*. *Mol & Gen Genet* **224**, 421–430.

Zinchenko, V. V., Babykin, M. M. & Shestakov, S. V. (1984). Mobilization of non-conjugative plasmids into *Rhodopseudomonas sphaeroides*. *J Gen Microbiol* **130**, 1587–1590.

Zinchenko, V., Saano, A., Babykin, M., Kopteva, A. & Shestakov, S. (1985). Gene cloning vectors for *Rhodopseudomonas sphaeroides*. In *Abstracts of Vth International Symposium on Photosynthetic Prokaryotes*, p. 371 Grindelwald, Switzerland.

Zinchenko, V. V., Kopteva, A. V., Belavina, N. V., Mitronova, T. N., Frolova, V. D. & Shestakov, S. V. (1991). The study of *Rhodobacter sphaeroides* mutants of different type with derepressed nitrogenase. *Genetika* **27**, 991–999 (in Russian).

Received 5 November 1993; revised 1 February 1994; accepted 8 February 1994.