A transposition-induced mutant of *Nostoc ellipsosporum* implicates an argininebiosynthetic gene in the formation of cyanophycin granules and of functional heterocysts and akinetes

F. Leganés,¹ F. Fernández-Piñas¹ and C. P. Wolk²

Author for correspondence: F. Leganés. Tel: +34 91 397 8176. Fax: +34 91 397 8344. e-mail: francisco.leganes@uam.es

In strain NE1 of Tn5-1058-mutagenized Nostoc ellipsosporum, the transposon was found within a gene whose translation product is similar in amino acid sequence to the arginine-biosynthetic protein N-acetylglutamate semialdehyde dehydrogenase encoded by argC of Bacillus subtilis. The argC reported from Anabaena sp. strain PCC 7120 hybridized to a sequence different from the one interrupted by the transposon in NE1. The newly identified gene from N. ellipsosporum was denoted argL. The argL mutation renders certain processes in strain NE1 conditionally dependent on provision of L-arginine. Heterocysts and apparent akinetes that formed in the absence of added L-arginine failed to fix dinitrogen or to germinate, respectively, and lacked granules of cyanophycin, composed of copolymers of arginine and aspartic acid. However, apparent akinetes that differentiated upon growth of the mutant in the presence of L-arginine plus nitrate formed cyanophycin granules and could regenerate a new culture.

Keywords: akinete, arginine, cyanophycin, differentiation, heterocyst

INTRODUCTION

When certain filamentous cyanobacteria are grown in the presence of fixed nitrogen, all of the cells have similar morphology, and are known as vegetative cells. When the filaments are deprived of nitrogen, vegetative cells at semiregular intervals along the filaments differentiate into nitrogen-fixing heterocysts (Buikema & Haselkorn, 1993; Wolk *et al.*, 1994). Different types of environmental stress (Sutherland *et al.*, 1979; Wolk, 1965) lead to the formation of another kind of specialized cell, the akinete or spore.

Several observations suggest that the alternative differentiation processes may be related. In filaments of some organisms, akinetes arise where heterocyst formation might have been expected (Geitler, 1925). In Anabaena species, the envelopes of heterocysts and akinetes contain equivalent polysaccharides (Cardemil & Wolk,

Abbreviation: DW, dry weight.

1976, 1979, 1981). Soriente *et al.* (1993) reported that akinetes contain glycolipids characteristic of heterocyst envelopes (Winkenbach *et al.*, 1972). The further suggestion has been made (Wolk *et al.*, 1994) that heterocysts may be evolutionary derivatives of akinetes. Evidence has been presented that genes *hetR* and *hepA* are involved in the differentiation of both heterocysts and akinetes (Leganés *et al.*, 1994; Leganés, 1994). To pursue analysis of the differentiation of akinetes, and to test further the presumptive relationship between the differentiation of heterocysts and akinetes, we made use of *Nostoc ellipsosporum*, a genetically manipulable cyanobacterium that can form both these types of cells (Leganés *et al.*, 1994).

Cyanophycin is a nitrogenous reserve material, easily visualized in the form of refractile granules by phasecontrast microscopy, that normally consists of copolymers of arginine and aspartic acid (Simon, 1971; see, however, Merritt *et al.*, 1994). It is often present in the polar regions of heterocysts (Lang *et al.*, 1972) and is frequently abundant in akinetes (Herdman, 1987). Here, we report analysis of transposon mutant NE1 of

¹ Departamento de Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain

² MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA

The GenBank accession number for the sequence reported in this paper is U48355.

N. ellipsosporum, whose akinetes and heterocysts are morphologically abnormal and, as is true also of the vegetative cells of the mutant, lack cyanophycin granules in the absence of exogenous L-arginine. The lack of these granules correlates with impaired functionality of both types of differentiated cells.

METHODS

Strains, plasmids and culture conditions. Axenic cultures of N. ellipsosporum derived from strain B1453-7 of the algal culture collection of the Plant Physiology Institute of the University of Göttingen, and strains of Escherichia coli, were grown in liquid media or on agar-solidified media as previously described (Hu et al., 1981; Wolk et al., 1984). AA/8 is one-eighth strength AA medium. Akinetes were formed in the presence of standard sporulation medium supplemented, when specified, with 25 mM calcium glucuronate obtained from Mann Research Laboratories (Wolk, 1965). The concentrations of antibiotics used in the initial selection for transposon-bearing exconjugants, upon transfer of pRL1058 to N. ellipsosporum (Leganés et al., 1994), were 200 µg neomycin sulfate ml⁻¹ and 10 µg streptomycin sulfate ml⁻¹. Suspension cultures of mutant NE1 were supplemented with 40 µg neomycin sulfate ml^{-1} and 5 µg streptomycin sulfate ml^{-1} . Cultures of strain DR1505 (see below) were supplemented with 2.5 µg streptomycin sulfate ml⁻¹ and 2.5 µg spectinomycin ml⁻¹. Where indicated, cultures were supplemented with 1 or 4 mM L-arginine hydrochloride (Sigma).

Analytical methods. For dry weight (DW) determinations, cells were collected and dried at 70 °C for 48 h. Nitrogenase activity was determined by the acetylene reduction method (Stewart *et al.*, 1967). Cyanophycin granules were isolated by the method of Allen (1988), which is based on that of Simon (1971). Isolated cyanophycin granules were treated with 0.1 M HCl, and the arginine from these granules was measured by the modified Sakaguchi reaction of Messineo (1966).

Recovery of DNA contiguous with the transposon, reconstruction of the mutation, and Southern analysis. DNA contiguous with Tn5-1058 was recovered from the genomic DNA of cyanobacterial mutant NE1 by excision with EcoRV, circularization and transfer to E. coli by electroporation (Wolk et al., 1991), producing plasmid pRL1429. In an effort to reconstruct the original mutation that is present in NE1, pRL1429, cut with PstI and BamHI to eliminate most of the transposon, was ligated to pRL759D (Black et al., 1993) that had been cut with PstI and BamHI, producing plasmid pRL1441D. Plasmid pRL271 (Black et al., 1993) was cut with Nrul and Scal, and the large fragment, containing sacB and determinants for resistance to chloramphenicol and erythromycin, was introduced into the EcoRV site of pRL1441D, producing pRL1501. Cassette C.S4, present in plasmid pRL65 (Fernández-Piñas et al., 1994), confers resistance to streptomycin and spectinomycin. The lux-bearing fragment of pRL1501 bounded by sites for BamHI and BglII was replaced by cassette C.S4 (bracketed by BglII sites) from pRL65, producing pRL1505. This plasmid was introduced into wildtype N. ellipsosporum by conjugation (Elhai & Wolk, 1988). Double recombinant strains were selected as described by Cai & Wolk (1990) on AA medium with nitrate (Allen & Arnon, 1955), streptomycin, spectinomycin, 5% sucrose and 1 mM Larginine. For Southern analysis of chromosomal DNA from these strains, DNA probes were labelled with digoxigenin-11dUTP (Boehringer Mannheim) from random primers (Gibco-BRL).

Sequence analysis. Automated sequencing (Applied Biosystems) was performed on restriction fragments that were subcloned from pRL1429 to pUC18 and pUC119 (Yanisch-Perron *et al.*, 1985; Vieira & Messing, 1987). Both strands of DNA were sequenced; database comparisons and alignments of the DNA and of predicted sequences were performed with the default settings of the algorithm developed by Altschul *et al.* (1990) using the network service programs BLASTN and BLASTP of the National Center for Biotechnology Information (NCBI).

RESULTS AND DISCUSSION

Phenotype of mutant NE1

Transposon-generated mutants of N. ellipsosporum affected in the development of heterocysts were identified by their inability to assimilate N₂ (Ernst et al., 1992). When one such mutant, strain NE1, was transferred to medium without combined nitrogen, the doubling time, estimated from the increase in dry weight, of the mutant was significantly greater than that of the wild-type (Table 1); after 48 h, cultures of the mutant started to bleach. Heterocysts of the mutant had partially detached envelopes, lacked cyanophycin granules at their poles (Fig. 1b), and lacked detectable nitrogenase activity (Table 1). Concordantly, cyanophycin was not detected in strain NE1 that had been deprived of fixed nitrogen for 48 h (Table 1). Thus, interruption of argL by the transposon in strain NE1 (see below) and perhaps, as suggested by Gupta & Carr (1981), the resulting absence of cyanophycin granules rendered the structure of the heterocyst abnormal and prevented N2 assimilation by the heterocysts. An alternative interpretation, that newly fixed NH₃ that is not consumed in the synthesis of arginine may accumulate and uncouple ATP synthesis, and that the heterocyst envelope may become distorted as a consequence of ATP deficiency, appears invalid for the following reason: when wild-type N. ellipsosporum and strain NE1 were deprived of fixed nitrogen under an atmosphere of argon, strain NE1 developed heterocysts with the same envelope distortion as observed under N₂replete conditions, whereas heterocysts of the wild-type strain showed no such distortion.

When strain NE1 was grown in sporulation medium with nitrate, virtually every cell appeared to become an akinete eventually. Like akinetes of the wild-type strain, these cells were larger than vegetative cells, were detached from other cells, and had a thicker envelope than did vegetative cells (Fig. 1e). They appeared at the same time as did akinetes in the wild-type strain. However, they were smaller and less rounded than the wild-type akinetes, had a homogeneous appearance (Fig. 1e) due to the lack of cyanophycin granules, and failed to germinate. A small amount of cyanophycin was detectable analytically (Table 1). Thus, interruption of argL by Tn5-1058 (see below), and perhaps the absence of cyanophycin granules, prevented normal functioning of akinetes in N. ellipsosporum. Although stored cyanophycin is mobilized during the process of akinete germination, Sutherland et al. (1985) concluded that cyanophycin is not the primary nitrogen source for de

Table 1. Effect of the nitrogen source on the growth, nitrogenase activity and cyanophycin content of cultures of *N. ellipsosporum* wild-type and mutant NE1

The culture medium used for growth on N_2 was that of Hu *et al.* (1981), and that for growth on NO_3^- was (HEPES-buffered) standard sporulation medium (Wolk, 1965). The doubling time, nitrogenase activity and cyanophycin content were determined at 48 h, except that the cyanophycin content of cultures grown with NO_3^- and NO_3^- + 1 mM Arg (L-arginine) was determined after 20 d culture, when akinetes had differentiated (see Fig. 1). The results shown are the means \pm standard error of the mean of three independent experiments with duplicate samples. ND, Not detected.

<i>Nostoc</i> strain	Nitrogen source	Doubling time (h)	Nitrogenase activity [nmol C ₂ H ₄ (mg DW) ⁻¹ h ⁻¹]	Cyanophycin content [µg L-Arg (mg DW) ⁻¹]
Wild-type	N ₂	41.6 ± 2.3	159 ± 13.5	2.3 ± 0.7
NE1	N ₂	65·8±2·9*	ND*	ND*
Wild type	NO ₃	23.1 ± 2.1	ND	4.1 ± 0.5
NE1	NO_{3}^{-}	23.4 ± 2.1	ND	$0.8 \pm 0.2^{*}$
Wild type	$NO_{3}^{-} + 1 \text{ mM Arg}$	$24 \cdot 4 \pm 2 \cdot 1$	ND	15.7 ± 2.0
NE1	$NO_3^- + 1 \text{ mM Arg}$	$25 \cdot 2 \pm 1 \cdot 7$	ND	4·7 <u>+</u> 1·4*

* Differed statistically (Student's t test, P < 0.05) from the wild-type strain under the same culture conditions.

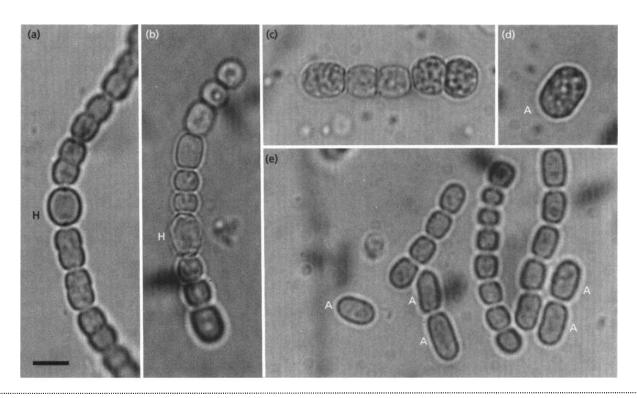


Fig. 1. Heterocysts (H) in wild-type *N. ellipsosporum* (a) and strain NE1 (b) 48 h after transfer to AA/8 medium lacking both nitrate and arginine. The heterocysts of strain NE1 lack polar cyanophycin granules and have an abnormal envelope. Wild-type *N. ellipsosporum* (c, d) and NE1 (e) after 20 d growth in standard sporulation medium (which contains nitrate) supplemented with calcium glucuronate. Apparent akinetes (A) in NE1 (e) are smaller and less rounded than akinetes in the wild-type strain (d) and their interior appears homogeneous due to the lack of cyanophycin granules. Whereas vegetative cells of wild-type *N. ellipsosporum* have visible cyanophycin granules, vegetative cells of NE1 do not (compare c and e). Bar, 5 μm.

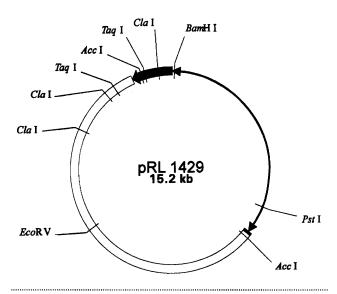


Fig. 2. Physical map of plasmid pRL1429 bearing transposon Tn5-1058 (5.4 kb; thin curved line) together with genomic DNA recovered from the DNA of strain NE1 by restriction with *Eco*RV, ligation, and transfer to *E. coli*. Insertion of the transposon duplicated bp 56–64 in the 1056 bp *argL* gene. Black boxes, *argL*; open box, other *N. ellipsosporum* DNA. Only the positions of restriction sites used for subcloning and sequencing are depicted.

novo synthesis of protein. Therefore, the inability of akinetes of NE1 to germinate may not have been due to a lack of cyanophycin as a nitrogen source for protein synthesis.

In sporulation medium with nitrate, wild-type N. *ellipsosporum* accumulated cyanophycin in the stationary phase of growth (Table 1) with granules visible in vegetative cells and akinetes (Fig. 1c). In contrast,

prolonged incubation of strain NE1 in sporulation medium resulted in death of the culture. Death was prevented if the differentiation of akinetes was avoided by subculturing to fresh nitrate-supplemented liquid medium (normally $AA/8 + NO_3^-$; Hu *et al.*, 1981) every 4–5 d (on solid medium, once per week). Thus, unlike *argC* mutants CS335 and CS336 of *Anabaena* sp. strain PCC 7120 (Floriano *et al.*, 1992), NE1 is not a strict Larginine auxotroph, but rather a conditional auxotroph.

Because the gene mutated in NE1 showed great similarity to a gene involved in the synthesis of L-arginine (see below), growth media were supplemented with this amino acid. When both L-arginine and nitrate were present, strain NE1 formed cyanophycin (Table 1) and small, presumptive akinetes, some of which contained cyanophycin granules. Cultures of strain NE1 in sporulation medium supplemented with L-arginine reinitiated growth when transferred to fresh medium, although normal germination of akinetes was never observed. Addition of L-arginine, as nitrogen source, to a medium lacking combined nitrogen prevented heterocyst differentiation in both the wild-type and strain NE1, and nitrogenase activity was undetectable (data not shown).

Reconstruction of the mutation present in mutant NE1

To determine whether the phenotype of NE1 was the result of insertion of the transposon, rather than the result of a secondary mutation, the transposon insertion was reconstructed (see Black *et al.*, 1993). Transposon Tn5-1058 (5.4 kb) and 9.8 kb of contiguous DNA were recovered from mutant NE1 upon excision with *Eco*RV, circularization, and transfer to *E. coli*, producing plasmid pRL1429 (Fig. 2), from which plasmid pRL1505 was constructed (see Methods). Sequencing outward

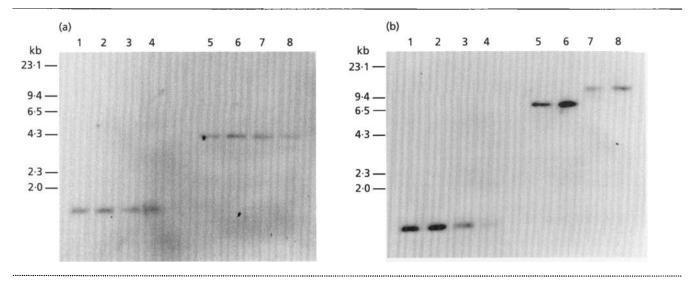


Fig. 3. Southern analysis of DNA from Anabaena sp. strain PCC 7120 (lanes 1, 5), wild-type N. ellipsosporum (lanes 2, 6), strain NE1 (lanes 3, 7), and strain DR1505 (lanes 4, 8) cut with ClaI (lanes 1–4) or EcoRV (lanes 5–8), and probed with (a) a 1.6 kb ClaI fragment that contains argC from Anabaena sp. strain PCC 7120 (Floriano et al., 1992) or (b) a 1.3 kb ClaI fragment that bears part of the gene mutated in strain NE1 (argL; see text). The two probes hybridized to DNA from both sources, but targeted bands of different sizes.

Ne MENFERVPVGIVGASGYGGVQLVRLIMDHPEIELVYLGGESSVGKSFASLYPHLAHAVKLSIEEVDPEVIA 71 Ss MEN.EKARVGIIGASGYGGUQLVRLIMDHPEIELVYLGGESSVGKSFASLYPHLAHAVKLSIEEVDPEVIA 71 Ss MEN.EKARVGIIGASGYGGUQLVRLIMDHPEIELVYLGGESSVGKSFASLYPHLAHAVKLSIEEVDPEVIA 71 An MNKPKIFIDGEAGTTGLQIYSRLMEPDUELYYLAGHSSAGKPYSDLYPHLTHKVNLTIEPIDLEAIA 70 An MNKPKIFIDGEAGTTGLQIYSRLMERDDIELVSIAASKRKDADERAKLL 49 Bs HEIDIMFLAAPPGVSSELTPKLADAGITVIDLSGDLRIKEPAEYEKWYKRTAAPKAVIQEAVYGLAELNQL 136 Ne RRCEVVFLSMPNGLACQIVPTLLEKGCKVLDLSADYRFNULTYTTWYGVERSDRTTADTAIYGJPELYRD 142 Ss SCDAVFLGLPNGLACDMAPALLAKGCKVLDLSADYRFNULTYTTWYGVERSDRTTADTAIYGJPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDASTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL. DESFVIVDANTGVSGEGGEGEAKTYLLLAEDNDFKIYKVN 206 Ne RISEAQLVCCPGSYPTASLLALSPLLAQGFLVPE.TAIIDAKSGTSGGGEAKTYLLLAEDNDFKIYKVN 206 Ne RIBAQLIGCPGCYPTASLLALSPLLAQGFLVPE.TAIIDAKSGTSGGGEAKTYLLLAEDNSLAPFSVV 212 Ss AIRDAQLIGCPGCYPTASLLALSPLLAQGFIVFE.TAIIDAKSGTSGGGEKAKYNLLAEAGSLAPKYNV 206 Ne RIBAQLVCCPGSYPTASLLALSPLLAQGFIVFE.TAIIDAKSGTSGGGGEGGEAKYNNLLAEAGSLGAYGA 211 An • FGIYGQGKGAGAGA	Bs	MKIGIVGATGYGGTELVRILSHHPHABECILYSS8GEGNVYSEGYPHLTGLADQQLKPIDMNTIK	65	
An MNKPKIFIDGEAGTTGLQIYSRLNERDDIELVSIAASKRKDADERAKLL 49 Bs HEIDIMFLAAPPGVSSELTPKLADAGITVIDLSGDLRIKEPAEYEKWYKRTAAPKAVIQEAVYGLAELNQL 136 Ne RRCEVVFLEMPNGLACQIVPTLLEKGCKVLDLSADYRFNNLTYTTWYGVERSDRTTADTAIYGLPELYRD 142 Ss SRCDAVFLGLPNGLACQIVPTLLEKGCKVLDLSADYRFNNLTYTWYGVERSDRTTADTAIYGLPELYRD 142 Ss SRCDAVFLGLPNGLACDMAPALLAKGCKVLDLSADYRFNNLTYTWYGVERSDRTTADTAIYGLPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDASTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL. DESFVIVDANTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGFIVPE.TAIYDAKTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGFIVPE.TAIYDAKTGVSGGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGFIVPE.TAIYDAKTGVSGGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGFIVPE.TAIYDAKTGTSGGGREAKVNMLLAEADNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVPE.TAIYDAKTGTSGGGREAKVNMLLAEADNSLAPSVV 211 NKIANAFVSNPCYPTSLCVRFLAQUFTPHLKQGFIVPE.TAIYDAKTGTSGGGREAKVNMLLAEAGSLGAYGVA 211 NKIANAFVSNPCYPTSLCVRFLAQUFTPHLVRYTAGTIATNY.TRLTCDLTADDLHDLYSEFYQDSYFVRVP 276 NR KIANAQFVSNPCYPTSLCVRFTHILPUVGTPHLVRGILATVATLRDPGLVGDDLTTYTAFYRNSPWVKVCE 233 Ss KNRHTPTIEQIASDLAGTELKVQTTPHLIPWRGILATVATLRDPGLVGDLITIYSAFYRASPFVKILP 242<	Ne	MONFGRVFVGIVGASGYGGVOLVRLIMDHPEIELVYLGGESSVGKSFASLYPHLAHAVKLSIEEVDPEVIA	71	
Bs HEIDIMFLAAPPGVSSELTPKLADAGITVIDLSGDLRIKEPAEYEKWYKRTAAPKAVIQEAVYGLAELNQL 136 Ne RRCEVVFLEMPNGLACQIVPTLLEKGCKVLDLSADYRFNLATYTTWYGVERBORTTADTAIYGLPELYRD 142 SscDavflglpnglacdmapallangckvldlsadyrfnulntytewykkdrodqatnnqavyglpelyrd 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDASTAYRTAQGWVYGFPEMNPGQRE Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGLVVP.TAYDANSGTGGGGREARTYLLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVP.TAYDANSGTGGGGREARTYLLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVP.TAYDANSGTGGGGREARTYLLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVP.TAYDANSGTGGGGREARTYLLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVP.TAYDANSGTGGGGREARTYLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVP.TAYDANSGTGGGGREARTYLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLASPLLKQGFIVP.TATINAVSGYSGGGGREARTYLLAEDNSLAPSVV 212 Sa AIRDAQLIGCPGCYPTASLLASPLLKQGFIVP.TATINAVSGYSGGGGREARTYLLAEDNSLAPSVV 212 Sa AIRDAQLYCCPGYPTASLLASPLLKQGFIVP.TATINAVSGYSGGGGREARTYLLAEDNSLAPSVV 212 Sa FGIYGLQRAGH Sa FGIYGLQRAGH Sa FGIYGLQRAGH Sa KHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATYATLROPGLVGDDLTTYTAFYRNSPWVKVCZ <t< td=""><td>Ss</td><td>MCN. EKARVGIIGASGYGGIQLVRLLLEHPQVELTYLAGHSSACKPYSDLYPHLTHRVNLTIEPIDLEAIA</td><td>70</td><td></td></t<>	Ss	MCN. EKARVGIIGASGYGGIQLVRLLLEHPQVELTYLAGHSSACKPYSDLYPHLTHRVNLTIEPIDLEAIA	70	
Ne RRCEVVFLSMPNGLACQIVPTLLENGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 142 Ss SRCDAVFLGLPNGLACDMAPALLANGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAUTYTEWYKKDRDQOATNNQAVYGLPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISBAQLVGCPGSYPTASLLALSPLLNQGILVPE.TAIVDANSGTSGGGREAKVNNLLABADNSLAPYSVV 212 Ss AIRDAQLIGCPCYPTASLLALSPLLNQGFLYPE.TAIVDANSGTSGGGREAKVNNLLABAEGSLGAYGVA 211 n KANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREALUQKYDSFHEQQKGATSDYP 175 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NKHKHTPEIEQQLANEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTHTEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLTTYTAFYRNSPWVKCE 283 Ss KHRHTPEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLITYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLIPWRGILATVATLRDPGLV	An	M NKPKIFIDGEAGTTGLQIYSRLNERDDIELVSIAASKRKDADERAKLL	49	
Ne RRCEVVFLSMPNGLACQIVPTLLENGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 142 Ss SRCDAVFLGLPNGLACDMAPALLANGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAUTYTEWYKKDRDQOATNNQAVYGLPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISBAQLVGCPGSYPTASLLALSPLLNQGILVPE.TAIVDANSGTSGGGREAKVNNLLABADNSLAPYSVV 212 Ss AIRDAQLIGCPCYPTASLLALSPLLNQGFLYPE.TAIVDANSGTSGGGREAKVNNLLABAEGSLGAYGVA 211 n KANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREALUQKYDSFHEQQKGATSDYP 175 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NKHKHTPEIEQQLANEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTHTEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLTTYTAFYRNSPWVKCE 283 Ss KHRHTPEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLITYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLIPWRGILATVATLRDPGLV				
Ne RRCEVVFLSMPNGLACQIVPTLLENGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 142 Ss SRCDAVFLGLPNGLACDMAPALLANGCKVLDLSADYRFNLTYTTWYGVERSDRTTADTAIYGLPELYDD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAUTYTEWYKKDRDQOATNNQAVYGLPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANTGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISBAQLVGCPGSYPTASLLALSPLLNQGILVPE.TAIVDANSGTSGGGREAKVNNLLABADNSLAPYSVV 212 Ss AIRDAQLIGCPCYPTASLLALSPLLNQGFLYPE.TAIVDANSGTSGGGREAKVNNLLABAEGSLGAYGVA 211 n KANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREALUQKYDSFHEQQKGATSDYP 175 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NKHKHTPEIEQQLANEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 NRHTHTEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLTTYTAFYRNSPWVKCE 283 Ss KHRHTPEIEQASDLAGHETVQTFPHLIPWRGILATVYATLRDPGLVRDDLITYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLIPWRGILATVATLRDPGLV	_		176	
Ss SRCDAVFLGLPNGLACDMAPALLARGCRVLDLSADYRFRNLNTYTEWYKKDRQDQATNNQAVYGLPELYRD 141 An NSVDVAILCLPDDAAREAVSLVNSSQVKILDASTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDAATGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCGGYPTASLLALSPLLNQGLVVF.TAVDARGGTSGGGREARYNLLLAEADNSLAFYSVV 212 Ss AIRDAQLIGCFGCYPTASLLALSPLLNQGIVFF.TAILDARGGTSGGGREARYNLLLAEAGSLGAYGVA 211 An KIANAQFVSNPGCYPTASLLALSPLLNQGIVFF.TAIIDARSGTSGGGREARYNNNLLAEAEGSLGAYGVA 211 An KIANAQFVSNPGCYPTASLLALSPLLNQGIVFF.TAIIDAKSGTSGGGREARYNNLLAEAEGSLGAYGVA 211 An FGIYGLQRGFLACVRPLIAQGILPSSFPITINAVSGYSGGGRSLIQKYDSFHEQQKGATSDYP 175 An • FGIYGLQRAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQD8YFVRVP 276 Ne RHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVRDDLITIYAFYRNSPWVKVCE 283 Ss KHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVRDDLITIYAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLUWTLDNPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNIMNGNMEETGLLTIPIYP 345 Ne SGIYPQTKKEVYGSNFCDIAVTLDERTNRVIVLSATIONSVIDNLIRGGAGQAVQCNFNIMGKMEETGLFNGGFYP	-			
An NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAQGWVYGFPEMNPGQRE 104 Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANTGVBGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVCCPGSYPTASILALBPLINGGLVPF.TAIVDAR <u>SGTBGGGREARTYILLARADNSLAPISVV 212</u> Ss AIRDAQLIGCPGCYPTASILALBPLINGGLVPF.TAIVDAR <u>SGTBGGGREARTYILLARADNSLAPISVV 212</u> Ss AIRDAQLIGCPGCYPTASILALBPLINGGLVPF.TAIVDAR <u>SGTBGGGREARTYILLARADNSLAPISVV 212</u> Ss AIRDAQLIGCPGCYPTASILALBPLINGGLVPF.TAIVDAR <u>SGTBGGGREARTYILLARADNSLAPISVV 212</u> Ss AIRDAQLIGCPGCYPTASILALBPLINGGLVPF.TAIVDARSGTBGGGREARTYILLARADNSLAPISVV 212 Ss AIRDAQLIGCPGCYPTASILALBPLINGGLVPF.TAIVDARSGTBGGGREARTYILLARADNSLAPISVV 212 Ss AIRDAQLIGCPGCYPTASILALBPLINGGILPF.TAIIDARSGTBGGGREARTYILLARADNSLAPISVV 212 Ss FGIYGLQFPTGFLACVRPLIAGGILPSSFPITINAV <u>SGYBGGGRSLICKVDSFHEQQRGATSDYP 175</u> An • FGIYGLQFAGH 186 Bs EHQHTPEIEQICSDLAGHEVTVQFTPHLINGVINGILATVATIRDPGLVGDDLTHIYAFYRNSPWXVCC 283 Ss KHRHTPEIEQICSDLAGHEVTVQFTPHLIPMVRGILATVYATIRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGJFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVS <u>VIDNLMRGAAGQAVQNFNILMIGMDETGLITIPIYP</u> 345 NSVDVARGNLCYIGIEVDPRTGRVIVLS <u>AIDNLVRGAAGQAVQCNNFNILMIGMDETGLIFNGGTYP</u> 352 Ss HGVYQTKNAMGTNLCYIGIEVDPRTGRVIVLS <u>AIDNLVRGAAGQAVQCINILMIGMDETGLIP</u> AGJFP				
Bs QIQQAKLIANPGCFPTAVLLGLAPLAQKKLL.DESFVIVDANIGVSGAGRKASMGTHFSELNDNFKIYKVN 206 Ne RISEAQLVGCPGSYPTASLLALSPLLKQGLIVPE.TAIVDANIGTSGGGREANTYLLLAEADNSLAPYSVV 212 SATRDAQLICCPCCYPTASLLALSPLLKQGLIVPE.TAIVDANIGTSGGGREANTYLLLAEADNSLAPYSVV 212 SATRDAQLICCPCCYPTASLLALSPLLKQGLIVPE.TAIVDANIGTSGGGREANTYLLLAEADNSLAPYSVV 212 An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREANTYLLLAEADNSLAPYSVV 212 An * FGIYGLQFAGH Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RIRHTPEIEQICGDLACHEVTVQFTPHLIPMVRGILATVYATIRDPGLVRDDLTTIYTAFYRNSPWKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATIRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGJFEQGMLVQIPPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KQQYPQTKEVYGSNFCDIAVTLDERTNKVTIVSVIDNIAKGAAGQAVQNFNIANGNNEETGLTITPIYP 345 Ne SGIYPQTKKANGTNLCYIGIEVOPRTGRVIVLSAIDNIAKGAAGQAVQNFNIANGNDETGLIPMGFYP 352 Ss HGVYPQTKKANGTNLCYIGIEVOPRTGRVIVLSAIDNIKGAAGQAVQCINIAMGMEESLGLPQLAFYP 351				
Ne RISEAQLVGCPGSYPTASLLALSPLLKQGLIVFF.TAIVDARSGTSGGGREAKTYLLLALADNSLAPYSVV 212 Ss AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVFF.TAIIDARSGTSGGGREAKVNNLLARAEGSLGAYGVA 211 An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREAKVNNLLARAEGSLGAYGVA 211 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVCDDLTTIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGHEVTVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLIRGAAGQAQAVQNFNILMNGNNEETGLTITPIYP 345 N SGIYPQTKWAAGSNLCYIGVPRTGRVIVLSAIDNLIKGQAGQAVQCNINLANGMEETGLFUGPKMGTYP 352 Ss HGVYPQTKMAMGTNLCYIGIEVDPRTGRVIVLSAIDNLIKGQAGQAVQCUNILMGMEESLGLPQLAFYP 351	An	NSVDVAILCLPDDAAREAVSLVNSSQVKILDABTAYRTAQGWVYGFPEMNPGQRE	104	
Ne RISEAQLVGCPGSYPTASLLALSPLLKQGLIVFF.TAIVDARSGTSGGGREAKTYLLLALADNSLAPYSVV 212 Ss AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVFF.TAIIDARSGTSGGGREAKVNNLLARAEGSLGAYGVA 211 An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREAKVNNLLARAEGSLGAYGVA 211 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVCDDLTTIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGHEVTVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLIRGAAGQAQAVQNFNILMNGNNEETGLTITPIYP 345 N SGIYPQTKWAAGSNLCYIGVPRTGRVIVLSAIDNLIKGQAGQAVQCNINLANGMEETGLFUGPKMGTYP 352 Ss HGVYPQTKMAMGTNLCYIGIEVDPRTGRVIVLSAIDNLIKGQAGQAVQCUNILMGMEESLGLPQLAFYP 351				
Ne RISEAQLVGCPGSYPTASLLALSPLLKQGLIVFF.TAIVDARSGTSGGGREAKTYLLLALADNSLAPYSVV 212 Ss AIRDAQLIGCPGCYPTASLLALSPLLKQGFIVFF.TAIIDARSGTSGGGREAKVNNLLARAEGSLGAYGVA 211 An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGREAKVNNLLARAEGSLGAYGVA 211 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVCDDLTTIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGHEVTVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHRHTPEIEQICADLAGTELRVQFTPHLIPWVRGILATVYATLRDPGLVRDDLITIYAFYRNSFWVKVCE 283 Ss KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLIRGAAGQAQAVQNFNILMNGNNEETGLTITPIYP 345 N SGIYPQTKWAAGSNLCYIGVPRTGRVIVLSAIDNLIKGQAGQAVQCNINLANGMEETGLFUGPKMGTYP 352 Ss HGVYPQTKMAMGTNLCYIGIEVDPRTGRVIVLSAIDNLIKGQAGQAVQCUNILMGMEESLGLPQLAFYP 351	Bs	0100AKLIANPGCFPTAVLLGLAPLAOKKLL, DESFVIVDARTGVSGAGRKASMGTHFSELNDNFKIYKVN	206	
Ss AIRDAQLICCPCCYPTASLLALSPLLEQGFIVFE.TAIIDARSCTSCGGREAKUNNLLALAEGSLGAYGVA 211 An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGKSLIQKYDSFHEQQKGATSDYP 175 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLICDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RIRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATIROPGLVGDDLTTIYAFYRNSPWKVCE 283 Ss KHRHTPEIEQICSDLAGHEVTVQFTPHLIPMVRGILATVYATIROPGLVGDDLTTIYAFYRNSPWKVCE 283 Ss KHRHTPEIEQICSDLAGHEVTVQFTPHLIPMVRGILATVYATIROPGLVGDDLTTIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGJFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMRGAAGQAVQNFNIANGWNEETGLTITPIYP 345 Ne SGIYPQTKKAVGSNLCYIGVPPTGRVICKSVIDNLIRGQAGQAQQNOFNIANGWEETGLFUGPFWGFYP 352 Ss HGVYPQTKWAMGTNLCYIGIEVDPRTGRVIVLSAIDHLYKGQAGQAVQCINILMGWEESLGLPQLAFYP 351	Ne			
An KIANAQFVSNPGCYPTGFLACVRPLIAQGILPSSFPITINAVSGYSGGGKSLIQKYDSFHEQQKGATSDYP 175 An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RIRHTPEIEQICGDLAGHETVQFTPHLVPIVRGILATVYATIRDPGLVRDDLTTIYTAFYRNSPWXVCE 283 Ss KHRHTPEIEQIASDLAGHETVQFTPHLIPMVRGILATVYATIRDPGLVRDDLITIYTAFYRNSPWXVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATIRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPMVTLDNPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNIANGWNEETGLTITPIYP 345 Ne SGIYPQTKKAVGTNLCYIGIEVDPRTGRVICMSVIDNLIRGQAGQAJQCINIMGCMEETGLPMGFYP 352 Ss HGVYPQTKKAMGTNLCYIGIEVDPRTGRVIVLSAIDNLIKGQAGQAVQCINILMGMEESLGLPQLAFYP 351	Ss	AIRDAOLIGCPGCYPTASLLALSPLLKOGFIVPE, TAIIDAKSGTSGGGREAKVNMLLAEAEGSLGAYGVA	211	
An • FGIYGLQFAGH 186 Bs EHQHTPEIEQALNEWQPGLGPITFSAHLFPMTRGIMATMY.TRLTCDLTADDLHDLYSEFYQDSYFVRVRP 276 Ne RHRHTPEIEQICSDLAGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQICSDLAGHEVTVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIABDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIABDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Sn KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLUMTLDNPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQQAVQNFNIMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAGSNLCYIGVEVPRTGRVICMSVIDNLIRGQAGQAVQCINIMGGNEETGLIPMGFYP 352 Ss HGVYPQTKWAMGTNLCYIGIEVDPRTGRVIVLSAIDNLVRGQAGAVQCINIMGGNEESLGLPQLAFYP 351	An			
Ne RHRHTPEIEQICSDLÄGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNLMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAAGSNLCYIGVEVDPRTGRVIGMSVIDNLIRGQAGQAJQCLNLMMGMEETGLPKMGFYP 352 Ss HGVYPQTKWAAGTNLCYIGIEVDPRTGRVIVLSAIDNLVKGQAGQAVQCLNLMMGMEESLGLPQLAFYP 351	An	• FGIYGLQFAGH	186	
Ne RHRHTPEIEQICSDLÄGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNLMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAAGSNLCYIGVEVDPRTGRVIGMSVIDNLIRGQAGQAJQCLNLMMGMEETGLPKMGFYP 352 Ss HGVYPQTKWAAGTNLCYIGIEVDPRTGRVIVLSAIDNLVKGQAGQAVQCLNLMMGMEESLGLPQLAFYP 351				
Ne RHRHTPEIEQICSDLÄGHEVTVQFTPHLVPIVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVGDDLTTIYTAFYRNSPWVKVCE 283 Ss KHRHTPEIEQIASDLAGTELRVQFTPHLIPMVRGILATVYATLRDPGLVRDDLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNLMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAAGSNLCYIGVEVDPRTGRVIGMSVIDNLIRGQAGQAJQCLNLMMGMEETGLPKMGFYP 352 Ss HGVYPQTKWAAGTNLCYIGIEVDPRTGRVIVLSAIDNLVKGQAGQAVQCLNLMMGMEESLGLPQLAFYP 351	n -		276	
Ss KHRHTPEIEQIASDLAGTELRVOPTPHLIPMVRGILATVYATLROPGUVRODLITIYSAFYRASPFVKILP 282 An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNLMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAAGSNLCYIGVEVVPRTGRVICMSVIDNLIRGQAGQAVQCLNLMMGWEETGLTITPIYP 352 Ss HGVYPQTKWAAGSNLCYIGIEVDPRTGRVIVLSAIDNLVKGQAGQAVQCLNLMMGWEESLGLPQLAFYP 351				
An KHVKEMHQHSGLLFIPAVGDFEQGMLVQIPLPLWTLDNPPSGEEIHQAIAQYYQGEKFVQVAPFKDPSLLR 257 Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMRGAAGQAVQNFNLMNGWNEETGLTITPIYP 345 Ne SGIYPQTKWAAGSNLCYIGVEVDPRTGRVICMSVIDNLIRGQAGQAYQCINIMGWEETGLTUTPIYP 352 Ss HGVYPQTKWAWGTNLCYIGIEVDPRTGRVIVLS <u>AIDNL</u> VKGQAGQAVQCINIMGWEESLGLPQLAFYP 351				
Bs KGQYPQTKEVYGSNFCDIAVTLDERTNRVIIVS <u>VIDNLWRGAAGQAVQNFNLMNGWNEETGL</u> IITPIYP 345 Ne sgiypqtrkaagsnlcyigytvoprtgrvigas <u>vidnlirggaggaiqcinimggdetigi</u> pragyp 352 Ss HgypqtrkaagsnlcyigievDprtgrvivls <u>aidnl</u> vrg <u>Qaggavqcinimggdetigi</u> pqafyp 351				
Ne sciypotrvaacsnicyiguevoprtorvicasyidniirooacoaiociniiacopeticipragfyp 352 Ss hgvypotrvaaginicyigievoprtorvivis <u>aidni</u> vkooacoavociniaasm <u>eesicip</u> olafyp 351	An	NEVREMEQUOGULE TEXAGLE EQUIDA QIFLE DE INDEFOGEDINQAINQIIQUERE AQVAETROFOLDA	23,	
Ne sciypotrvaacsnicyiguevoprtorvicasyidniirooacoaiociniiacopeticipragfyp 352 Ss hgvypotrvaaginicyigievoprtorvivis <u>aidni</u> vkooacoavociniaasm <u>eesicip</u> olafyp 351				
Ss HGVYPOTKMAWGTNLCYIGIEVDPRTGRVIVLSAIDNLVKGDAGOAVOCLNLMMGWEESLGLPOLAFYP 351	Bs	KGQYPQTKEVYGSNFCDIAVTLDERTNRVTIVSVIDNLMKGAAGQAVQNFNLMNGWNEETGLTITPIYP	345	
	Ne	SGIYPQTRWAAGSNLCYIGVEVDPRTGRVIGMSVIDNLIKGOAGOAIOCINIMMGNDETLGLPRMGFYP	352	
	Ss	HGVYPQTKWAWGTNLCYIGIEVDPRTGRVIVLSAIDNLVKGQAGQAVQCLNLMMGWEESLGLPQLAFYP	351	
An DGTFLDATAVNGTNIVQVFVFGNDNTKEALLVA <u>RLDNL</u> GK <u>GASGAAVONLNIM</u> LGL <u>PEELGL</u> 321	An	DCTFLDATAVNCTNIVQVFVFGNDNTKEALLVARLDNLGKGASGAAVONLNIMLGLPEELGL	321	

Fig. 4. Comparison of the predicted sequences of (Ne) the protein encoded by argL from N. ellipsosporum and those of the N-acetylglutamate semialdehyde dehydrogenases from (Bs) B. subtilis (Smith et al., 1990), (An) Anabaena sp. strain PCC 7120 (Floriano et al., 1992), and, by presumption, (Ss) Synechocystis sp. strain PCC 6803 (Kaneko et al., 1996). The bold-face characters indicate amino acid residues that are identical in the predicted protein from N. ellipsosporum. The underlined regions are presumptively involved in the binding of nucleotides: the upper two are similar to the 'P-loop' in ATP- and GTP-binding proteins (Floriano et al., 1992; Saraste et al., 1990); the lower three are collectively similar to the $\beta\alpha\beta$ fold involved in binding of NAD, NADP and FAD by some oxidoreductases (Floriano et al., 1992). The asterisk indicates a cysteine residue conserved in Anabaena sp. strain PCC 7120, B. subtilis and Synechocystis sp. strain PCC 6803, but replaced by a serine in N. ellipsosporum.

from the transposon in pRL1429 showed no vestige of pRL1058 outside of the transposon, indicating that strain NE1 was derived from a true transposition of Tn5-1058 (data not shown). Colonies derived from presumptive double recombination of pRL1505 with the wild-type strain were shown by Southern hybridization to have EcoRV fragments that differed from those of the wild-type strain but matched that of the original mutant strain NE1 (Fig. 3b). The phenotype of these colonies also matched that of strain NE1. One such strain was designated DR1505.

Analysis of the gene mutated in strain NE1

The transposon in strain NE1 was found to interrupt an open reading frame of 1056 bp encoding a protein of 352 amino acids (Fig. 4) that shows 40% identity (58%) similarity) to N-acetyl-y-glutamyl-phosphate reductase (N-acetylglutamate semialdehyde dehydrogenase), the product of the arginine biosynthetic gene, argC, from Bacillus subtilis (Smith et al., 1990); 72% identity (84% similarity) to the hypothetical product of what is presumed to be argC of Synechocystis sp. strain PCC 6803 (Kaneko et al., 1996); but only 18 % identity (26 % similarity) to the predicted product of argC from Anabaena sp. strain PCC 7120 (Floriano et al., 1992). The amino acid sequences deduced from argC of N. ellipsosporum (Ne), Anabaena sp. strain PCC 7120 (An) and B. subtilis (Bs), and of presumptive argC of Synechocystis sp. strain PCC 6803 (Ss) are compared in Fig. 4. Southern analysis of DNA from Anabaena sp. PCC 7120, N. ellipsosporum wild-type, strain NE1 and strain DR1505 using the argC genes from Anabaena sp. PCC 7120 (Floriano et al., 1992) and N. ellipsosporum (this paper) as probes showed different patterns of hybridization (Fig. 3): one copy of each gene is present in each of the cyanobacteria. We denote the newly identified gene from N. ellipsosporum, argL.

Highly conserved regions (underlined in Fig. 4) of these proteins are similar to domains involved in the binding of nucleotides (Floriano et al., 1992; Saraste et al., 1990). Two domains that resemble a 'P-loop' (ATP-/GTP-binding site motif) are present in ArgL of N. ellipsosporum (amino acids 39-46 and 184-192). A presumptive 'P-loop' in the C-terminal part of the protein is common to all four predicted proteins depicted in Fig. 4, but a second presumptive 'P-loop' in the N-terminal part of the protein is present only in ArgL and in the protein from Synechocystis sp. strain PCC 6803. A βαβ fold that is involved in the binding of the adenine nucleotide moiety of NAD, NADP and FAD by some oxidoreductases is present in all four proteins. The cysteine that has been proposed to be involved in the reaction mechanism of N-acetylglutamate semialdehyde dehydrogenase (Fig. 4, asterisk; Parsot et al., 1988) is substituted by serine S155 in N. ellipsosporum but may be functionally replaced by the nearby cysteine C152.

As in the case of numerous genes from heterocystforming cyanobacteria, the open reading frame that is interrupted by the transposon in NE1 is positioned close to a series of repeated sequences, here 3' from the gene, that may possibly serve as binding sites for regulatory proteins (Holland & Wolk, 1990). A 7 nt repeat (located 65 bp after the termination triplet), CCCCAGT, occurring in five perfect and two imperfect sets, differs from that found 3' from *argC* in *Anabaena* sp. strain PCC 7120, CTAATGA (Floriano *et al.*, 1992), but closely resembles the sequence CCCCAAT found in *Anabaena* sp. strain PCC 7120 downstream from *hepA* (Holland & Wolk, 1990; gene renamed by Ernst *et al.*, 1992). Also, 10 bp after the termination triplet of *argL* and prior to the CCCCAGT repeats are three consecutive repetitions of the sequence GGGTATT.

Like insertions within *hetR* and *hepA* (Leganés *et al.*, 1994; Leganés, 1994), a transposon insertion in *argL* blocks normal differentiation and function of heterocysts and akinetes. Our results are consistent with the idea that *N. ellipsosporum* has two different genes, *argC* and *argL*, whose products are involved in the biosynthesis of arginine. We suggest that the product of *argC* is involved primarily in the synthesis of arginine for protein biosynthesis whereas the product of *argL* may be involved specifically in formation of cyanophycin.

ACKNOWLEDGEMENTS

F. Leganés was a recipient of an FPU postdoctoral fellowship (Subprograma General en el Extranjero, Modalidad Normal). This work was supported by NSF grant IBN-9118152, DOE grant DE-FG02-90ER20021 and DGICTYT no. PB 93-0274.

REFERENCES

Allen, M. B. & Arnon, D. I. (1955). Studies on nitrogen-fixing bluegreen algae. I. Growth and nitrogen fixation by *Anabaena cylindrica* Lemm. *Plant Physiol* 30, 366–372.

Allen, M. M. (1988). Inclusions: cyanophycin. Methods Enzymol 167, 207–213.

Altschul, F. F., Warren, G., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. J Mol Biol 215, 403–410.

Black, T. A., Cai, Y. & Wolk, C. P. (1993). Spatial expression and autoregulation of *betR*, a gene involved in the control of heterocyst development in *Anabaena*. Mol Microbiol 9, 77–84.

Buikema, W. J. & Haselkorn, R. (1993). Molecular genetics of cyanobacterial development. Annu Rev Plant Physiol Plant Mol Biol 44, 33–52.

Cai, Y. & Wolk, C. P. (1990). Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *J Bacteriol* 172, 3138–3145.

Cardemil, L. & Wolk, C. P. (1976). The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Methylation analysis and structure of the backbones. J Biol Chem 251, 2967–2975.

Cardemil, L. & Wolk, C. P. (1979). The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Structure of the basic repeating unit. J Biol Chem 254, 736–741.

Cardemil, L. & Wolk, C. P. (1981). Polysaccharides from the envelopes of heterocysts and spores of the blue-green algae *Anabaena variabilis* and *Cylindrospermum licheniforme*. J Phycol 17, 234–240.

Elhai, J. & Wolk, C. P. (1988). Conjugal transfer of DNA to cyanobacteria. Methods Enzymol 167, 747–754.

Ernst, A., Black, T., Cai, Y., Panoff, J.-M., Tiwari, D. N. & Wolk, C. P. (1992). Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. J Bacteriol 174, 6025–6032.

Fernández-Piñas, F., Leganés, F. & Wolk, C. P. (1994). A third genetic locus required for the formation of heterocysts in *Anabaena* sp. strain PCC 7120. J Bacteriol 176, 5277–5283.

Floriano, B., Herrero, A. & Flores, E. (1992). Isolation of arginine auxotrophs, cloning by mutant complementation, and sequence analysis of the *arg*C gene from the cyanobacterium Anabaena species strain PCC 7120. Mol Microbiol 6, 2085–2094.

Geitler, L. (1925). Synoptische Darstellung der Cyanophyceen in morphologischer und systematischer Hinsicht. *Beih Bot Centralbl* 41, 163–294.

Gupta, M. & Carr, N. G. (1981). Enzyme activities related to cyanophycin metabolism in heterocysts and vegetative cells of *Anabaena* spp. J Gen Microbiol 125, 17–23.

Herdman, M. (1987). Akinetes: structure and function. In *The Cyanobacteria*, pp. 227–250. Edited by P. Fay & C. Van Baalen. Amsterdam: Elsevier.

Holland, D. & Wolk, C. P. (1990). Identification and characterization of *betA*, a gene that acts early in the process of morphological differentiation of heterocysts. *J Bacteriol* 172, 3131–3137.

Hu, N. T., Thiel, T., Giddings, T. H. & Wolk, C. P. (1981). New *Anabaena* and *Nostoc* cyanophages from sewage settling ponds. *Virology* 114, 236–246.

Kaneko, T., Sato, S., Kotani, H. & 21 other authors (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3, 109–136, and 3(Suppl.), 185–209.

Lang, N. J., Simon, R. D. & Wolk, C. P. (1972). Correspondence of cyanophycin granules with structured granules in *Anabaena cylindrica*. Arch Mikrobiol 83, 313–320.

Leganés, F. (1994). Genetic evidence that *hepA* gene is involved in the normal deposition of the envelope of both heterocysts and akinetes in *Anabaena variabilis* ATCC 29413. FEMS Microbiol Lett 123, 63–68.

Leganés, F., Fernandez-Piñas, F. & Wolk, C. P. (1994). Two mutations that block heterocyst differentiation have different effects on akinete differentiation in Nostoc ellipsosporum. Mol Microbiol 12, 679–684.

Merritt, M. V., Sid, S. S., Mesh, L. & Allen, M. M. (1994). Variations in the amino acid composition of cyanophycin in the cyanobacterium *Synechocystis* sp. PCC 6308 as a function of growth conditions. *Arch Microbiol* 162, 158–166.

Messineo, L. T. (1966). Modification of the Sakaguchi reaction: spectrophotometric determination of arginine in proteins without previous hydrolysis. *Arch Biochem Biophys* **117**, *534–540*.

Parsot, C., Boyen, A., Cohen, G. N. & Glansdorff, N. (1988). Nucleotide sequence of *Escherichia coli argB* and *argC* genes: comparison of *N*-acetylglutamate kinase and *N*-acetylglutamategamma-semialdehyde dehydrogenase with homologous and analogous enzymes. *Gene* 68, 275–283.

Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990). The P-loop – a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* 15, 430–434.

Simon, R. D. (1971). Cyanophycin granules from the blue-green alga Anabaena cylindrica: a reserve material consisting of copolymers of aspartic acid and arginine. Proc Natl Acad Sci USA 68, 265–267.

Smith, M. C., Mountain, A. & Baumberg, S. (1990). Nucleotide sequence of the *Bacillus subtilis argC* gene encoding N-acetyl-glutamate-gamma-semialdehyde dehydrogenase. *Nucleic Acids Res* 18, 4595.

Soriente, A., Gambacorta, A., Trinone, A., Sili, C., Vincenzini, M. & Sodano, G. (1993). Heterocyst glycolipids of the cyanobacterium Cyanospira rippkae. Phytochemistry 33, 393–396.

Stewart, W. D. P., Fitzgerald, G. P. & Burris, R. H. (1967). Acety-

lene reduction by nitrogen-fixing blue-green algae. Arch Mikrobiol 62, 336-348.

Sutherland, J. M., Herdman, M. & Stewart, W. D. P. (1979). Akinetes of the cyanobacterium Nostoc PCC 7524: macromolecular composition, structure and control of differentiation. J Gen Microbiol 115, 273–287.

Sutherland, J. M., Reaston, J., Stewart, W. D. P. & Herdman, M. (1985). Akinetes of the cyanobacterium Nostoc PCC 7524: macromolecular and biochemical changes during synchronous germination. J Gen Microbiol 131, 2855–2863.

Vieira, J. & Messing, J. (1987). Production of single-stranded DNA. Methods Enzymol 153, 3-11.

Winkenbach, F., Wolk, C. P. & Jost, M. (1972). Lipids of membranes and of the cell envelope in heterocysts of a blue-green alga. *Planta* 107, 69–80.

Wolk, C. P. (1965). Control of sporulation in a blue-green alga. Dev Biol 12, 15-35.

Wolk, C. P., Vonshak, A., Kehoe, P. & Elhai, J. (1984). Construction of shuttle vectors capable of conjugative transfer from *Escherichia* coli to nitrogen-fixing filamentous cyanobacteria. *Proc Natl Acad* Sci USA 81, 1561–1565.

Wolk, C. P., Cai, Y. & Panoff, J. M. (1991). Use of a transposon with luciferase as a reporter to identify environmentally responsive genes in a cyanobacterium. *Proc Natl Acad Sci USA* 88, 5355–5359.

Wolk, C. P., Ernst, A. & Elhai, J. (1994). Heterocyst metabolism and development. In *Molecular Biology of the Cyanobacteria*, pp. 769–823. Edited by D. Bryant. Dordrecht: Kluwer.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119.

Received 24 December 1997; revised 16 February 1998; accepted 5 March 1998.