Volume 13 Number 20 1985 Nucleic Acids Research

Molecular cloning and nucleotide sequence of a developmentally regulated gene from the cyanobacterium *Calothrix* PCC 7601: a gas vesicle protein gene

Nicole Tandeau de Marsac*, Didier Mazel, Donald A.Bryant+ and Jean Houmard

Unité de Physiologie Microbienne, Département de Biochimie et Génétique Moléculaire, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France, and +Department of Molecular and Cell Biology, S-101 Frear Building, The Pennsylvania State University, University Park, PA 16802, USA

Received 16 August 1985; Accepted 30 September 1985

ABSTRACT

Since the gas vesicle protein (GVP) is highly conserved among the different gas-vacuolate prokaryotes, a 29-mer oligonucleotide corresponding to a portion of the Anabaena flos-aquae GVP gene was synthesized and used to isolate the GVP structural gene from Calothrix PCC 7601 (= Fremyella diplosiphon). Gas vacuole production in this filamentous cyanobacterium is restricted to hormogonia which occur at a specific stage during the developmental cell cycle. The GVP gene (gvpA) was localized on a 709 bp HindIII-HincII fragment. Nucleotide sequence analysis revealed a 213 bp open reading frame whose deduced amino-acid sequence shows a very high homology with that of the Anabaena flos-aquae GVP. Assuming that the first methionine residue is proteolytically processed, the molecular mass of the Calothrix GVP is 7375 daltons. Sequences resembling the Escherichia coli consensus promoter were found upstream from the gvpA gene. The initiator codon of the gvpA gene is preceded by a polypurine sequence assumed to be the ribosome binding site. Southern hybridizations with a probe specific for the gvpA gene indicated that this gene is not plasmid-borne, and that another homologous gene is present in the Calothrix genome.

INTRODUCTION

Gas vacuoles are gas-filled structures providing many different aquatic prokaryotes with buoyancy and allowing their positioning at the most favorable depth for growth (for reviews see 1-4). They are made up of clusters of cylindrical electron-transparent structures termed "gas vesicles". These minute hollow cylinders with conical ends are generally 0.07-0.3 µm in diameter with a length ranging from 0.3-1 µm depending on the strain. The membrane of the vesicles is 2 nm thick and is constituted of a monolayer of protein molecules forming ribs ∿ 4.5 nm wide that wind around in a cylindrical structure. These vesicles are impermeable to water, but permeable to gases. However, when gas vesicles are collapsed by pressure, they cannot be reinflated and are presumably resynthesized by recycling of the protein from disaggregated vesicles (5).

Gas vesicles can be isolated and purified in an intact state, provided

that exposure to pressures is avoided (6). Detailed biochemical analysis of isolated gas vesicles showed that protein, possibly a single polypeptide species, is their sole organic constituent. Consequently, the gas vesicle membrane differs from classical cell unit membranes but resembles the proteinaceous coats of viruses (2). The gas vesicle protein (GVP) is one of the most abundant proteins in some cyanobacterial cells (3.6 % - 10 % of the total protein content) (5,7,8) and can be easily purified. However, its high hydrophobicity has impeded an accurate determination of its molecular mass by gel electrophoresis (2). Based on quantitative amino-acid analyses, molecular masses ranging from $\sim$ 7300 up to $\sim$ 20,600 daltons have been proposed (2,9). Interestingly, comparison of partial amino-acid sequences (9), as well as immunological cross reaction experiments (9,10) have shown that the GVP is a highly conserved protein among prokaryotes, including eubacteria and archaebacteria.

Gas-vacuolate prokaryotes are most commonly planktonic and synthesize gas vacuoles constitutively (1,2). However, there exist nonplanktonic filamentous cyanobacteria in which gas vacuole formation occurs at a specific stage in their developmental cell cycle. Gas vacuole formation in these organisms is restricted to specialized filaments called hormogonia which differentiate at the tapering ends of trichomes (1). In contrast to vegetative trichomes, hormogonia are generally motile and supposedly play a role in the dispersal of a species in its natural habitat. Changes in environmental factors, such as light intensity and/or wavelength, as well as nutrient availability have been shown to induce formation of gas-vacuolate hormogonia (1,11-13 and G. Guglielmi, personal communication). However, the precise regulatory mechanism of this differentiation process remains unknown.

In this report, we present the isolation and the nucleotide sequence of the structural gene encoding the GVP of the filamentous cyanobacterium Calothrix PCC 7601. To our knowledge, this is the first gene encoding GVP which has been cloned and the first molecular marker available to study hormogonia differentiation in cyanobacteria.

MATERIALS AND METHODS

Materials

Restriction enzymes, calf intestinal alkaline phosphatase, T4 DNA ligase and T4 polynucleotide kinase were purchased from Genofit. Klenow fragment of DNA polymerase I, [α-^{32}P] dATP ($\sim$ 400 Ci/mmol) and [γ-^{32}P] ATP

(∿ 3000 Ci/mmol) were from Amersham. M13 forward sequencing universal primer was obtained from Biolabs. Nucleotides and dideoxynucleotides triphosphates were purchased from PL-Pharmacia. All enzymes were used according to the manufacturer's instructions. Phage λ DNA was a gift from A. Klier (Institut Pasteur). Phages M13 mp8 and mp9 were gifts from G. Lenzen and A. Labigne-Roussel, respectively (Institut Pasteur). All chemicals were reagent grade.

Culture conditions and cyanobacterial DNA purification

Calothrix sp. strain PCC 7601 (= Fremyella diplosiphon, UTEX 481) was grown in BG11 medium (14), at 25° C. Cultures were bubbled continuously with 1 % CO_2/99 % air and illuminated with cool-white fluorescent tubes (50 μ $E.m^{-2}.s^{-1}$). Cells were harvested in early exponential growth phase and chromosomal DNA was isolated as previously described (15). Plasmid DNAs were extracted from cells in late-exponential growth phase, by CsCl gradient centrifugation as described by van den Hondel et al. (16) or evidenced by the in-well lysis technique as described by Rosenberg et al. (17).

Library construction and DNA subclonings

Genomic DNA from Calothrix PCC 7601 was partially digested with Sau 3A and size fractionated on a 10 %- 40 % sucrose gradient (18). Fragments ranging from ∿ 14 to ∿ 20 kilobase pairs (Kb) were dephosphorylated, then ligated into the λ EMBL3 vector (19) previously digested with EcoRI and BamHI. Ligated DNA was packaged in vitro into the heads of bacteriophage λ and the E. coli strain K 802 (20) was infected with this packaged DNA (21). Large-scale phage purification and DNA extraction were performed as described (22).

Recombinant λ DNA fragments were electroeluted from agarose gels and subcloned either into the E. coli plasmid pUC9 or into the M13 phage vectors mp8 and/or mp9 (23) previously digested with the appropriate restriction enzymes. Strains used to propagate recombinant, pUC9 plasmids and M13 phages, were RDP 211 [exconjugant of RDP153 (24); F128 from JM103 (25)] and TG1, a restrictionless derivative of JM101 (25), respectively. Plasmid DNA extraction was performed as described (26).

Synthesis of oligonucleotides

The 29-mer oligonucleotide used to probe the gvpA gene, as well as specific 14-mer and 17-mer oligonucleotides used as sequencing primers were synthesized by an Applied Biosystems Model 380A automated DNA synthesizer as described previously (27).

Hybridization with ^{32}P-labelled probes

Restricted genomic DNA and recombinant DNAs were electrophoresed in 0.7 % agarose gel containing 89 mM Tris-base, 89 mM boric acid, 3.2 mM acid EDTA, pH 8.3. Cyanobacterial plasmid DNAs were electrophoresed as described (16,17). Prior to Southern transfer, gels containing undigested plasmid DNAs were soaked in 250 mM HCl for 15 mn.

The 29-mer oligonucleotide probe was labelled at the 5' end with [γ-^{32}P] ATP using T_4 polynucleotide kinase and the 236 bp HindIII-HincII DNA probe was nick-translated (28). Conditions for Southern blots and *in situ* plaque hybridizations were as previously described (29), except that the temperatures used were 45° C with the 29-mer probe and 60° C with the nick-translated probe.

Nucleotide sequence analysis

Sequencing in pUC9 was carried out by the dideoxy chain termination method (30) according to Heidecker *et al.* (31), whereas the method reported in the Amersham instruction booklet was followed for the sequencing in the M13 vectors. The 29-mer initially used to probe the GVP gene and the 14 and 17-mers, specially synthesized as sequencing on one strand progressed, as well as the universal 15-mer M13 forward primers, were used to complete the sequence on both strands and to overlap restriction sites used in subcloning. Amino-acid sequence and codon frequencies were deduced from the nucleotide sequence by using the computer program of Staden (32) as modified by B. Caudron (Institut Pasteur).

RESULTS AND DISCUSSION

Partial N-terminal amino-acid sequences of the GVP purified from two taxonomically distinct cyanobacteria, *Anabaena flos-aquae* and *Microcystis aeruginosa* are extremely homologous (33), confirming that GVP is a highly conserved protein. Based upon these sequences, we synthesized an oligonucleotide probe to isolate the corresponding gene from another cyanobacterium, *Calothrix* PCC 7601. Instead of short oligonucleotide sequences corresponding to all possible combinations of codons, we chose to synthesize a unique, 29-bases long, oligonucleotide. For each codon, the base composition was predicted on the basis of codon usage information obtained from the nucleotide sequence analysis of the *Calothrix* PCC 7601 *ps2B-1* (=*psbA*) gene (34) and of the *Anabaena* PCC 7120, *nifH*, *nifD*, *nifK* (35,37), *glnA* (38) and *rbcA* (39) genes. This synthetic oligonucleotide was used to probe a Southern blot of *Calothrix* PCC 7601 total genomic DNA

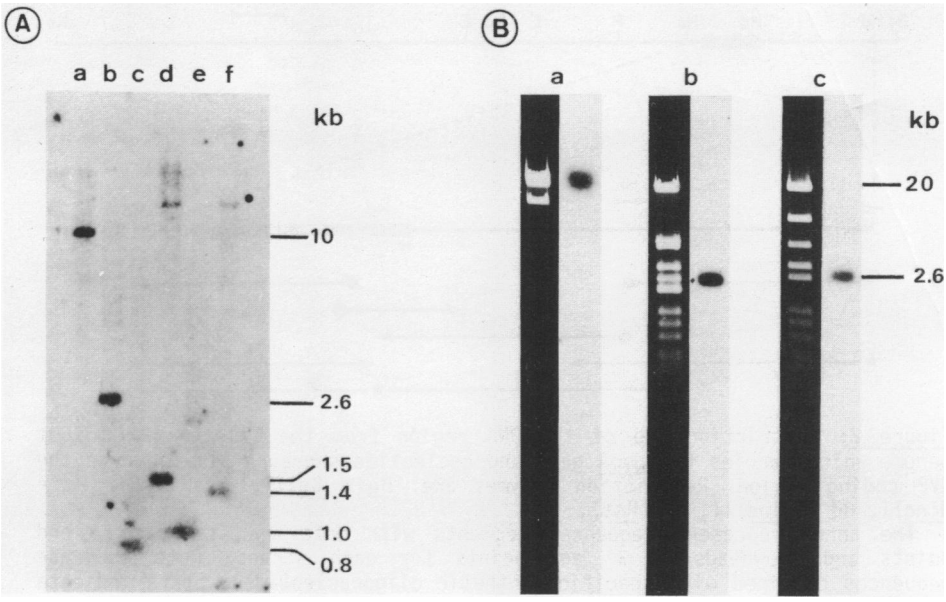

Figure 1: (A) Autoradiogram of a Southern blot of Calothrix PCC 7601 total DNA proved with the 29-mer oligonucleotide. DNA digested with EcoRI (lane a), HindIII (lane b), EcoRI and HindIII (lane c), BglII (lane d), BglII and HindIII (lane e) or BglII and EcoRI (lane f). Hybridizing bands of high molecular masses in lanes d and f are due to partial hydrolysis with BglII. (B) Agarose gel and autoradiogram of a Southern blot of the recombinant GVP7 DNA probed with the 29-mer oligonucleotide, after digestion with SalI (lane a), SalI and HindIII (lane b) or HindIII (lane c). The size of the hybridizing fragments, in kilobases (Kb), are shown on the right.

digests. As shown in Fig. 1A, a single hybridization band was obtained with the different restriction enzymes tested. These results demonstrated the high specificity of the 29-mer probe with regards to hybridization with total genomic DNA from Calothrix PCC 7601.

To clone the sequence homologous to the probe, a genomic library from Calothrix PCC 7601 was constructed in the λ EMBL3 vector and screened. Five λ recombinants which strongly hybridized with the 29-mer probe were selected and one of them, λGVP7, was studied further. This recombinant carried a total insert of ∿ 20 Kb as revealed by SalI digestion (Fig. 1B, lane a). After digestion with SalI and HindIII or HindIII alone (Fig. 1B, lanes b and c), the same 2.6 Kb band exhibited hybridization; this fragment had the size expected from hybridization with total genomic DNA digests (Fig. 1A, lane b). This 2.6 Kb HindIII DNA fragment was subcloned into the E. coli vector pUC9 and a partial nucleotide sequence was obtained by using

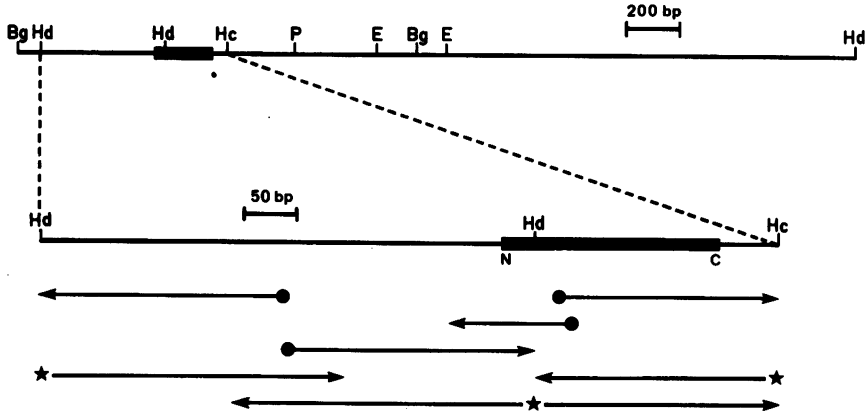

Figure 2: Restriction map of the DNA region from the Calothrix PCC 7601 genome which carries the gvpA gene and nucleotide sequence strategy for the GVP coding region. Restriction enzymes are: Bg = BglII, E = EcoRI, Hc = HincII, Hd = HindIII, P = PstI.
The arrow represent sequenced segments with dots and stars at 5' end points and anowheads at 3' end points for each strand. Dots indicate sequences obtained with specific synthetic oligonucleotides. Stars indicate sequences obtained using commercial M13 forward sequencing 15-mer primer. The thick lines shows the gvpA coding region, with NH_2 and COOH termini indicated (N and C, respectively); bp = base pairs.

the 29-mer oligonucleotide as a primer. Comparison of the deduced amino-acid sequence with that determined for the Anabaena flos-aquae GVP (33) confirmed that the 2.6 Kb HindIII DNA fragment effectively carried the gene encoding the GVP. However, the nucleotide sequence obtained from the other strand using a specific 17-mer oligonucleotide, revealed that the 5' end of the GVP gene was missing in this fragment. Further analysis showed that the entire GVP gene was carried by a 1.5 Kb BglII fragment.

To complete the sequence on both strands, smaller DNA fragments were subcloned into M13 phage vectors. The sequencing strategy is shown in Fig. 2. The sequence of 709 nucleotides extending from the HindIII site, located on the left part of the restriction map, to the unique HincII site just following the 3' end of the GVP coding region is shown in Fig. 3. In this sequence, a 213 bp open reading frame was found. As shown below, 25 out of the 29 nucleotides of the sequence used as a probe match the corresponding sequence found in the gene encoding GVP in Calothrix PCC 7601 (gvpA gene):

29-mer 5' ATC TTG GAC AAA GGT ATC GTT ATC GAC GC-3'
Ile Leu Asp Lys Gly Ile Val Ile Asp Ala

gvpA gene 5' ATC TTA GAC AAA GGT ATC GTT GTA GAT GC-3'
Ile Leu Asp Lys Gly Ile Val Val Asp Ala

The amino acid sequence deduced from the nucleotide sequence of the gvpA gene shows a very high homology with those determined for Anabaena flos-aquae GVP (N-terminal residues 1 to 64 and C-terminal tryptic peptide) (9,33). Only few differences are observed: i) a replacement of the amino-acids Ile (ATC/T/A) by Val (GTA) at residue 25 (nucleotides 76-78) and Gln (CA A/G) by Glu (GAA) at residue 37 (nucleotides 112-114), which might require only one base change in the nucleotide sequence of the Anabaena flos-aquae GVP gene; ii) in the Calothrix GVP amino-acid sequence, there are two successive Ala residues corresponding to nucleotides at position 199 to 205, whereas in Anabaena flos-aquae GVP there is only one residue at this position, the second one being located at the end of the amino-acid sequence. Although in the gvpA gene from Calothrix PCC 7601, the initiator codon is an AUG, no methionine residue was found in any of the N-terminal amino-acid sequences determined for the various cyanobacterial and bacterial GVPs (9). This suggests that, in Calothrix PCC 7601 GVP, the methionine is most likely proteolytically processed as it has been reported for some proteins in other prokaryotic or eukaryotic organisms. More interestingly, a stop codon UAA was located at nucleotides 214-216. This implies that the protein encoded by the gvpA gene in Calothrix PCC 7601 contains no more than 70 amino-acid residues with a total molecular mass of 7375 daltons, if one excludes the first methionine residue. This result is in agreement with the assumption of Walker et al. (9) who observed that the sequence of the C-terminal tryptic peptide of Anabaena flos-aquae GVP overlapped the N-terminal sequence by eight residues and concluded that this protein could possibly contain only 70 residues with an overall molecular mass of 7380 daltons. A similar molecular mass was also predicted from X-ray diffraction studies (40). Our results, together with these previous observations, provide an answer to the question pending for years about the molecular mass of the GVP purified from different prokaryotes. Indeed, numerous attempts to determine its molecular mass by gel electrophoresis analyses led to a very confusing situation. Konopka et al. (41) estimated a molecular mass of 50,000 daltons for the GVP purified from the heterotrophic bacterium Microcyclus aquaticus. For two archaebacteria, Halobacterium halobium and Halobacterium salinarium, the molecular mass varied from 12,100 to 18,200 according to the authors (42,43). These latter values were similar to those determined for GVPs from the cyanobacteria Microcystis aeruginosa [14,300 (44); 15,700 (45); 14,600 (45)] , Anabaena-

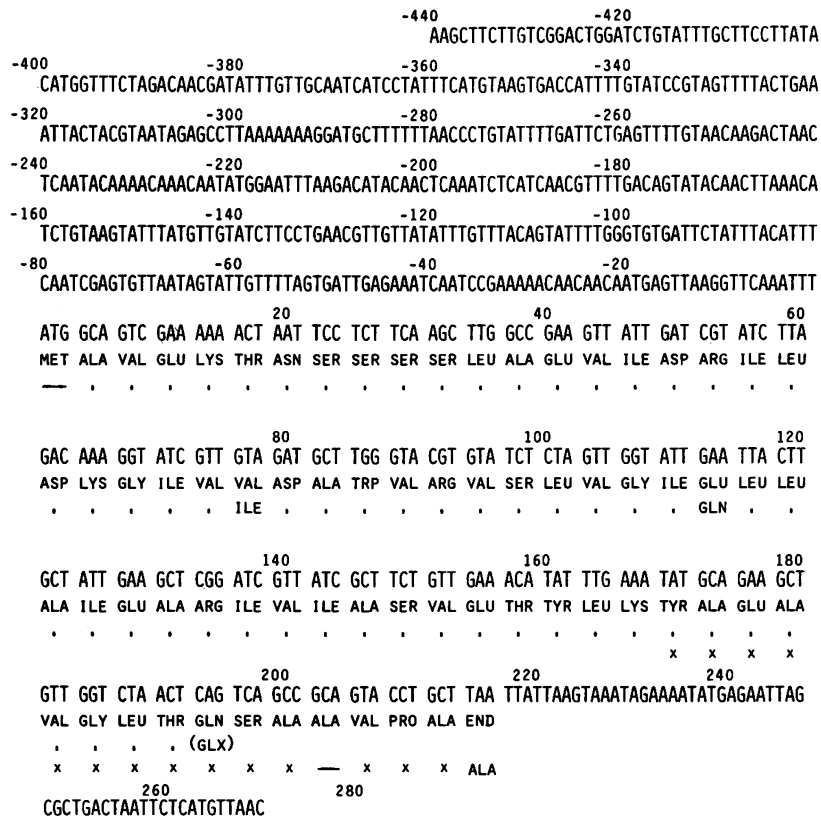

Figure 3: Nucleotide sequence of the GVP coding region from the genome of Calothrix PCC 7601. The deduced amino-acid sequence (second line) for the Calothrix GVP is juxtaposed under the nucleotide sequence (first line), as well as the N-terminal sequence (dots) and the C-terminal tryptic peptide sequence (crosses) of the Anabaena flos-aquae GVP. Dots and crosses indicate identity with the Calothrix sequence at that position. Dashes denote gaps introduced to maximize homology. Parentheses indicate uncertainty for the corresponding amino-acid residue.

flos aquae [15,100 (46); 14.700 (42); 20,600 (33)] and Nostoc muscorum PCC 6719 [20,000 (12)]. The discrepancy in the apparent molecular mass is probably due to the low solubility of this hydrophobic protein which tends to aggregate during electrophoresis, leading to multimers of the GVP subunit.

An examination of the 5' nucleotide sequence upstream to the gvpA gene revealed several sequences analogous either to the -35 or to the -10

promoter sequences of the "nif gene-type" or the "E. coli-type" as found in cyanobacterial genes examined so far (47-50). However, the sequences which better fulfill the requirements for a promoter with regards to both sequences and spacing (51) are of the "E. coli-type": 5' TTGTAT 3' (nucleotides -136 to -141) and 5' TATATT 3' (nucleotides -117 to -122) which are separated by 15 bp. We are nonetheless aware of the fact that only the determination of the start site for transcription will allow the precise localization of the gvpA gene promoter sequences.

Nine nucleotides upstream from the AUG initiator codon, the sequence AAGG is found. This sequence is complementary to a portion of the 3' end of the only reported cyanobacterial 16S rRNA sequence 5' UCACCUCCUUU 3' (52,53) and might correspond to the ribosome binding site.

The nucleotide sequences upstream and downstream from the gvpA gene are significantly richer in A and T residues than the coding region (69.6 %, 74.1 % and 59.7 % respectively, mean DNA base composition of the Calothrix PCC 7601 genome being 42.7 mol % GC (54)). As expected in such AT rich regions, several palindromic structures can be found. Whether the secondary structures found upstream from the gvpA gene play a role in the regulation of its transcription is unknown. Downstream from it, the largest palindromic structure (nucleotides 237-264) has a 11 bases perfect inverted repeat sequence and a loop of 6 bases. However, this sequence does not have the common features of the transcription terminator sequences described by Rosenberg and Court (51), since it is neither GC rich nor followed by a series of T residues. As shown in Halobacteria (55-57), the production of gas vacuoles in cyanobacteria could be determined or controlled by plasmid genes (8). To verify the location of the gvpA gene in the Calothrix PCC 7601 genome, plasmids were extracted and Southern blot analysis of both plasmid and chromosomal DNAs were performed using a 236 bp HindIII-HincII DNA fragment as a probe. This DNA fragment contains 183 bp within the gvpA gene coding region and 54 bp of downstream flanking sequence. No hybridization was obtained with the plasmids (data not shown) suggesting that the gvpA gene is not harbored by one of the endogeneous plasmids found in this strain (the plasmids species mentioned in ref. 11 and few others of higher molecular weight, from 90 to $\sim$ 135 Kb). Interestingly however, the 236 bp probe hybridized with two HindIII chromosomal DNA fragments (2.6 Kb and 2.3 Kb) (data not shown). The 2.6 Kb fragment corresponds to the cloned fragment in which we found the gvpA gene. The presence of another hybridization band strongly suggests that a second gene, highly homologous

```
                          1         10        20        30        40
(a) Calothrix PCC 7601    MAVEKTNSSSSLAEVIDRILDKGIVVDAWVRVSLVGIELLAIEARIV
(b) Anabaena flos-aquae   -.....................I...........Q.........  ...
(c) Aphanizomenon flos-aquae -.....................I...........P/E....P....  ...
(d) Oscillatoria agardhii PCC 7821 -....V.....................I...........P..S.P....  ...
(e) Dactylococcopsis salina PCC 8305 -..........G..V.......V...L.........E...L...V.  ...
(f) Halobacterium halobium -.--QPD-..G....L..V....V...V.A.....P/EI.TV...V.  ...
(g) Halobacterium salinarium -.--QPD-..GV..ML..V....V...V.A ...
```

Figure 4: Comparison of the N-terminal amino-acid sequences of GVPs from Calothrix PCC 7601 (a) and from other cyanobacterial (b-e) and halobacterial (f,g) strains (b-g, data taken from ref. 9). A = alanine; D = aspartic acid; E = glutamic acid; G = glycine; I = isoleucine; K = lysine; L = leucine; M = methionine; N = asparagine; P = proline ; Q = glutamine; R = arginine; S = serine; T = threonine; V = valine; W = tryptophan. Dots indicate identity of the residue at that position. Dashes denote gaps introduced to maximize homology. Amino acids corresponding to the portion of the sequence chosen to synthesize the 29-mer oligonucleotide is underlined.

to the gvpA gene, exists. Based on chemical analysis of purified GVPs from different organisms, it has been assumed that the gas vesicles could be constituted of only one protein species (2). However, the isolation of mutants able to produce gas vesicles with only the conical end caps, but not the cylindrical part (41), as well as of mutants partially defective in gas vesicle formation (8,43,58) suggested that there could exist several

Table 1. Codon frequencies in Calothrix PCC 7601 GVP coding sequence

Phe	UUU	0	Ser	UCU	3	Tyr	UAU	2	Cys	UGU	0
Phe	UUC	0	Ser	UCC	1	Tyr	UAC	0	Cys	UGC	0
Leu	UUA	2	Ser	UCA	2	Term	UAA	1	Term	UGA	0
Leu	UUG	2	Ser	UCG	0	Term	UAG	0	Trp	UGG	1
Leu	CUU	1	Pro	CCU	1	His	CAU	0	Arg	CGU	2
Leu	CUC	0	Pro	CCC	0	His	CAC	0	Arg	CGC	0
Leu	CUA	2	Pro	CCA	0	Gln	CAA	0	Arg	CGA	0
Leu	CUG	0	Pro	CCG	0	Gln	CAG	1	Arg	CGG	1
Ile	AUU	3	Thr	ACU	2	Asn	AAU	1	Ser	AGU	0
Ile	AUC	4	Thr	ACC	0	Asn	AAC	0	Ser	AGC	1
Ile	AUA	0	Thr	ACA	1	Lys	AAA	3	Arg	AGA	0
Met	AUG	1	Thr	ACG	0	Lys	AAG	0	Arg	AGG	0
Val	GUU	6	Ala	GCU	6	Asp	GAU	2	Gly	GGU	3
Val	GUC	1	Ala	GCC	2	Asp	GAC	1	Gly	GGC	0
Val	GUA	4	Ala	GCA	3	Glu	GAA	6	Gly	GGA	0
Val	GUG	0	Ala	GCG	0	Glu	GAG	0	Gly	GGG	0

genes encoding gas vesicle proteins. This hypothesis is presently under study.

Comparison of the deduced N-terminal amino-acid sequence of the GVP from Calothrix PCC 7601 with those obtained from four other cyanobacterial strains and two halobacterial strains shows an extremely high homology. The low degree of divergence of the GVPs could be explained, at least in part, by the numerous physical constraints for this type of structure. As pointed out by Walker et al. (9), GVPs from cyanobacteria and halobacteria form two distinct groups, but nevertheless their degree of homology strongly suggests that they are phylogenetically related. Most of the differences consist in conservative replacements among aliphatic residues. The wider polymorphism in the gas vesicles of Halobacteria (2) could be the reflect of such variations.

GVP is a small protein which consists of 51 % of hydrophobic amino-acids. According to Walsby (2) and Armstrong et al. (12), it represents one of the most abundant protein in many cyanobacterial cells. For highly expressed genes, in E. coli, one of the criteria for codon usage optimization is the preferential utilization of pyrimidines at the third position for certain amino-acids. This has been defined by a parameter called the P2 index (59). When calculated for the gvpA gene, the value obtained (0.67) is close to those determined by de Lorimier et al. (29) for the highly expressed genes encoding the phycobiliprotein subunits (α-PC: 0.79; β-PC: 0.70) in the cyanobacterium Agmenellum quadruplicatum PR6. Generally, in the gvpA gene, the codon frequencies for the most abundant amino-acids are rather similar to those encountered in the few other genes already sequenced from the same (34) or from other cyanobacterial strains (29,35-39,47,48,60-65).

With the gene encoding the GVP now isolated and sequenced, one can address a number of questions regarding the gas vacuole formation process. The gvpA gene will provide an hybridization probe to determine the number of gvp genes and their molecular organization, not only in cyanobacterial strains, but also in other prokaryotes such as archaebacteria. It will also become possible to study the regulation of the gvp gene(s) in different physiological conditions and this should help to determine precisely the factor(s) which trigger(s) gas vacuole production. Lastly, this gene provides the first molecular probe to study the differentiation of vegetative trichomes into hormogonia during the developmental cycle of a number of filamentous cyanobacteria.

ACKNOWLEDGMENTS

We are grateful to A.-M. Castets for her skilled assistance, particularly in purification and hybridization of cyanobacterial plasmid DNA, to V. Stirewalt and J. Dubbs for their technical help during the initiation of this work, and to G. Guglielmi for helping in the preparation of the photographs. We would like to thank Drs. G. Cohen-Bazire, I. Saint-Girons and P. Marlière for valuable comments on this manuscript and M. Kaczorek for helpful discussions. This work was supported by an A.T.P.-C.N.R.S. grant (Microbiologie 955353, 1984).

*To whom correspondence should be addressed

REFERENCES

1. Walsby, A.E. (1981) in The Prokaryotes, Starr, M., Stolp, H., Trüper, H., Balow, A., Schlegel, H.G. Eds., pp. 224-235, Springer-Verlag, Berlin Heidelberg New York.
2. Walsby, A.E. (1978) Symposia of the Society for General Microbiology 28, 327-358.
3. Walsby, A.E. (1975) Ann. Rev. Plant Physiol. 26, 427-439.
4. Walsby, A.E. (1972) Bacteriol. Rev. 36, 1-32.
5. Hayes, P.K. and Walsby, A.E. (1984) J. Gen. Microbiol. 130, 1591-1596.
6. Walsby, A.E. (1974) in Methods in Enzymology, Fleischer, S. and Packer, L. Eds., pp. 678-686, Academic Press, New York.
7. Jones, D.D. and Jost, M. (1970) Arch. Microbiol. 70, 43-64.
8. Walsby, A.E. (1977) Arch. Microbiol.114, 167-170.
9. Walker, J.E., Hayes, P.K., Walsby, A.E. (1984) J. Gen. Microbiol., 130 2709-2715.
10. Konopka, A.E., Lara, J.C. and Staley, J.T. (1977) Arch. Microbiol. 112, 133-140.
11. Tandeau de Marsac, N. (1983) Bull. Inst. Pasteur 81, 201-254.
12. Armstrong, R.E., Hayes, P.K. and Walsby, A.E. (1983) J. Gen. Microbiol. 128, 263-270.
13. Rippka, R. and Herdman, M. (1985) Ann. Inst. Pasteur/Microbiol. 136A, 33-39.
14. Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M. and Stanier, R.Y. (1979) J. Gen. Microbiol. 111, 1-61.
15. Kallas, T., Rebière, M.C., Rippka, R. and Tandeau de Marsac, N. (1983) J. Bacteriol. 155, 427-431.
16. van den Hondel, C.A.M.J.J., Keegstra, W., Borrias, W.E. and van Arkel, G.A. (1979) Plasmid 2, 323-333.
17. Rosenberg, C., Casse-Delbart, F., Dusha, I., David, M. and Boucher C. (1982) J. Bacteriol. 150, 402-406.
18. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, U.S.A.
19. Frischauf, A., Lehrach, H., Poustka, A. and Murray, N. (1983) J. Mol. Biol. 170, 827-842.
20. Wood, W. (1966) J. Mol. Biol. 16, 118-133.
21. Hohn, B. and Murray, K. (1977) Proc. Natl. Acad. Sci. U.S.A 74, 3259-3263.
22. Davis, R.W., Botstein, D. and Roth, J.R. (1980) in Advanced Bacterial Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, U.S.A.
23. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.

24. Seifert, M.S. and Porter, R.D. (1984) Mol. Gen. Genet. 193, 269-274.
25. Messing, J. (1983) in Methods in Enzymology, Wu, R., Grossman, L. and Moldave, K. Eds., pp. 20-78, Academic Press, New York.
26. Holmes, D.S. and Quigley, M. (1981) Anal. Biochem. 114, 193-197.
27. Bryant, D.A., de Lorimier, R., Lambert, D.H., Dubbs, J.M., Stirewalt, V.L., Stevens Jr., S.E., Porter, R.D., Tam, J. and Jay, E. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 3242-3246.
28. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
29. de Lorimier, R., Bryant, D.A., Porter, R.D., Liu, W.-Y., Jay, E. and Stevens Jr., S.E. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 7946-7950.
30. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
31. Heidecker, G., Messing, J. and Gronenborn, B. (1980) Gene 10, 69-73.
32. Staden, R. (1978) Nucl. Acids Res. 5, 1013-1015.
33. Walker, J.E. and Walsby, A.E. (1983) Biochem. J. 209, 809-815.
34. Mulligan, B., Schultes, N., Chen, L. and Bogorad, L. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 2693-2697.
35. Mevarech, M., Rice, D. and Haselkorn, R. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 6476-6480.
36. Rice, D., Mazur, B.J. and Haselkorn, R. (1982) J. Biol. Chem. 257, 13157-13163.
37. Mazur, B.J. and Chui, C.F. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6782-6786.
38. Fisher, R., Tuli, R. and Haselkorn, R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 3393-3397.
39. Curtis, S.E. and Haselkorn, R. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 1835-1839.
40. Blaurock, A.E. and Walsby, A.E. (1976) J. Mol. Biol. 105, 183-199.
41. Konopka, A.E., Staley, J.T. and Lara, J.C. (1975) J. Bacteriol. 122, 1301-1309.
42. Simon, R.D. (1980) Electrophoresis 1, 172-176.
43. Simon, R.D. (1981) J. Gen. Microbiol. 125, 103-111.
44. Jones, D.D. and Jost, M. (1971) Planta 100, 277-287.
45. Weathers, P.J., Jost, M. and Lamport, D.T.A. (1977) Arch. Biochem. Biophys. 178,226-244.
46. Falkenberg, P., Buckland, B. and Walsby, A.E. (1972) Arch. Microbiol. 85, 304-309.
47. Curtis, S.E. and Haselkorn, R. (1984) Plant Mol. Biol. 3, 249-258.
48. Pilot, T.J. and Fox, J.L. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 6983-6987.
49. Shinozaki, K. and Sugiura, M. (1985) Mol. Gen. Genet. 200, 27-32.
50. Haselkorn, R., Curtis, S.E., Golden, J.W., Nierzwicki-Bauer, S.A., Robinson, S.J. and Tumer, N.E. (1985) in Molecular Form and Function of the Plant Genome, van Vlotten Doting, L., Groot, G.S.T., Hall, T.C., Eds. Plenum Press.
51. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. 13, 319-353.
52. Borbely, G. and Simoncsits, A. (1981) Biochem. biophys. Res. Commun. 101, 846-852.
53. Tomioka, N. and Sugiura, M. (1983) Mol. Gen. Genet. 191, 46-50.
54. Lachance, M.-A. (1981) Intern. J. System. Bacteriol. 31, 139-147.
55. Simon, R.D. (1978) Nature 273, 314-317.
56. Pfeifer, F., Weidinger, G. and Goebel, W. (1981) J. Bacteriol. 145, 375-381.
57. Weidinger, G., Klotz, G. and Goebel, W. (1979) Plasmid 2, 377-386.
58. Walsby, A.E. (1976) Arch. Microbiol. 109, 135-142.
59. Gouy, M. and Gautier, C. (1982) Nucl. Acids Res. 10, 7055-7074.

60. Lammers, P.J. and Haselkorn, R. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 4723-4727.
61. Reichelt, B.Y. and Delaney, S.F. (1983) DNA 2, 121-129.
62. Shinozaki, K. and Sugiura, M. (1983) Nucl. Acids Res. 11, 6957-6964.
63. Shinozaki, K., Yamada, C., Takahata, N. and Sugiura, M. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 4050-4054.
64. Nierzwicki-Bauer, S.A. and Curtis, S.E. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 5961-5965.
65. Tumer, N.E., Robinson, S.J. and Haselkorn, R. (1983) Nature 306, 337-342.