Biochem. J. (1983) **209**, 809–815
Printed in Great Britain

Molecular weight of gas-vesicle protein from the planktonic cyanobacterium *Anabaena flos-aquae* and implications for structure of the vesicle

John E. WALKER* and A. E. WALSBY†
**M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K., and †Department of Botany, University of Bristol, Bristol BS8 1UG, U.K.*

(*Received 4 November 1982/Accepted 3 December 1982*)

The gas vesicle of the planktonic cyanobacterium *Anabaena flos-aquae* is a cylindrical shell made of protein enclosing a gas-filled space. Protein sequence analysis shows that the vesicle is made from a single protein. By gel electrophoresis and amino acid analysis its molecular weight was estimated to be 20600. Taken with previously obtained X-ray data, a simple interpretation of its molecular structure is of the polypeptide snaking in six pairs of antiparallel chains, three in each layer. The molecule would repeat along the ribs of the vesicle at intervals of 3.4nm.

The gas vesicle of planktonic cyanobacteria is a hollow cylindrical structure formed from a shell of protein enclosing a gas-filled space [see Walsby (1972, 1978) for reviews]. Protein is the only component of the structure (Walsby & Buckland, 1969; Jones & Jost, 1970), and electrophoretic analysis has revealed only a single species of protein, although not all of the protein sample entered the gel (Jones & Jost, 1971; Falkenberg *et al.*, 1972; Walsby, 1977). X-ray-diffraction studies have shown that the protein is arranged in a crystalline structure along ribs (Blaurock & Walsby, 1976). However, interpretation of the structure has been hindered by uncertainty of the molecular weight of the gas-vesicle protein (GVP).

Previous estimates of the molecular weight of GVP from various planktonic prokaryotes have included 7500, 14000, 15000 and 22000–55000. In reviewing the evidence, Armstrong & Walsby (1981) concluded that although "one of these values may be the correct one ... selecting it on the basis of the evidence currently available would be no more than a lucky guess". The principal obstacle in obtaining a reliable estimate of molecular weight has been the difficulty in dissolving gas vesicles in SDS and in separating the protein by SDS/polyacrylamide-gel electrophoresis. The electrophoretic systems that had been used to demonstrate a single GVP species employed phenol/acetic acid/urea mixtures in which the mobility of the protein was determined by two factors, molecular weight and charge density, both unknown. Therefore no reliance could be placed on the molecular-weight estimate of 14000 or 15000 suggested by the mobility of GVP relative to various standard marker proteins.

In the present paper we report that sequence analysis in the *N*-terminal region of GVP confirms that the vesicle is made from a single protein species. By comparison with partial sequences of GVP from another cyanobacterium, *Microcystis aeruginosa* (Weathers *et al.*, 1977), it appears that the protein is highly conserved. Also, we have been able to dissolve gas vesicles freshly isolated from *Anabaena flos-aquae* in solutions containing SDS. Aged samples are poorly soluble. From the electrophoretic mobility of the protein in gels we estimate a molecular weight for GVP of about 21000. This estimate is confirmed by quantitative amino acid analysis, which indicates that if the least abundant amino acid, phenylalanine, occurs as only one residue per GVP molecule, then there are 196 residues present, giving a molecular weight of 20600 for the whole GVP molecule.

Abbreviations used: GVP, gas-vesicle protein; SDS, sodium dodecyl sulphate; h.p.l.c., high-pressure liquid chromatography. The one-letter notation used for amino acid sequence is given in *Biochem. J.* (1969) **113**, 1–4.

Experimental

Isolation of gas vesicles

Gas vesicles were isolated from large batch cultures of *Anabaena flos-aquae*, the same strain (C.C.A.P. strain 1403/13f) used in previous studies on gas vesicles (e.g. Falkenberg *et al.*, 1972; Walsby, 1982). The filaments were lysed in 0.7M-sucrose solution, and the gas vesicles were separated from other components, by gentle centrifugation, to

the surface of the suspension, followed by repeated washings, further centrifugation and membrane-filtration. The methods used were those described in detail by Buckland & Walsby (1971), with the modifications described by Walsby & Armstrong (1979). Determination of the amount of GVP present in purified samples was made from the volume contraction on compressing the intact gas vesicles, or from measurements of pressure-sensitive absorbance, using the relationship determined by Walsby & Armstrong (1979) that a suspension containing 1mg of gas vesicles·ml^{-1} gives a pressure-sensitive absorbance of 20.84cm^{-1} at a wavelength of 500nm.

Gel electrophoresis

Gel electrophoresis of GVP was carried out on vertical slab gels prepared by the method of Laemmli (1970). The polyacrylamide gels were polymerized in a solution containing 0.13% SDS and the protein samples were applied in solutions containing 2.5% SDS.

The molecular weight of GVP was estimated from its relative mobility compared with that of horse heart myoglobin (mol.wt. 17200), and the following enzymes isolated from *Bacillus stearothermophilus*: superoxide dismutase (mol.wt. 22700) (Brock & Walker, 1980), triosephosphate isomerase (25000) (Artavanis-Tsakonas & Harris, 1980), and D-glyceraldehyde-3-phosphate dehydrogenase (35000) (Harris *et al.*, 1980).

Amino acid analysis

Amino acid compositions were determined with the aid of a Durrum D-500 analyser after hydrolysis *in vacuo* at 105°C with 6M-HCl containing 0.1% phenol (Thompson & Sanger, 1963).

'Succinylation' (3-carboxypropionylation)

The amino groups of GVP (84nmol) were allowed to react with a 50-fold molar excess of succinic anhydride by a procedure described previously (Walker *et al.*, 1980).

Tryptophan cleavage

Unmodified (420nmol) and succinylated (84nmol) GVP were cleaved with dimethyl sulphoxide/HBr by the modification of the method of Savige & Fontana (1977) described by Wachter & Werhahn (1980). In one experiment the protein was dissolved in 600μl of formic acid. To this were added dimethyl sulphoxide (5μl), 35% (w/v) HCl (260μl) and water (40μl). After 30min reaction at room temperature, 49% (w/v) HBr (100μl) and dimethyl sulphoxide (5μl) were added; 30min later the reaction mixture was diluted with water (2ml) and evaporated to dryness in a stream of N_2 at 50°C. The residue was dissolved in 25% (v/v) formic acid (1ml) and applied to a column [120cm × 2cm i.d. (internal diameter)] of Sephadex G-50 (superfine grade) packed in the same solvent (see Fig. 1).

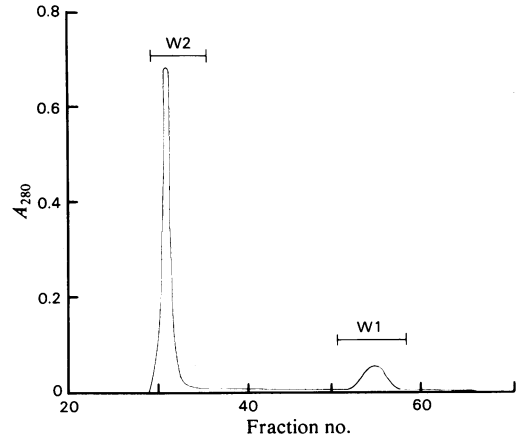

Fig. 1. *Fractionation of a tryptic digest of GVP (500nmol) on Sephadex G-50*

The column (110cm × 2.5cm i.d.) was eluted with 25% formic acid. The pooled fractions containing peptides used in sequence analysis are indicated by bars.

Digestion with trypsin

Before digestion the gas vesicles were denatured with formic acid as follows: formic acid was added dropwise to a suspension of intact gas vesicles in water (18mg of protein/ml). The first addition of acid caused collapse of the gas vesicles, indicated by a large decrease in turbidity and release of gas bubbles. When the concentration of the acid reached 80% (w/v), the remaining slight turbidity (generated by the collapsed vesicles) disappeared, indicating that the protein had dissolved. The solution was dialysed against water (6h) and then centrifuged at 45000**g** for 2h. This pelleted all the protein as a gel. The supernatant was removed and the gel dispersed in 0.5% NH_4HCO_3. Trypsin (4mg·ml^{-1} in 1mM-HCl; 10μl; substrate/enzyme, 50:1, w/w) was added. After 3h digestion at 37°C the gel had disappeared, leaving a clear solution. After 4h the digest was freeze-dried and the peptides were fractionated on a column (120cm × 2cm i.d.) of Sephadex G-50 (superfine grade) (see Fig. 2). Fractions containing ninhydrin-reactive material were pooled. Those pools used in subsequent sequence analysis are indicated in Fig. 2. One of the products of trypsinolysis, peptide T1, was purified from fractions 62–67 by h.p.l.c. on a Waters C18 μ-Bondapak column equilibrated in 0.1% trifluoroacetic acid eluted with a linear gradient of aceto-

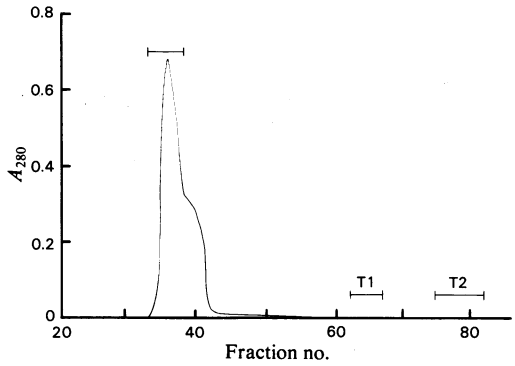

Fig. 2. *Fractionation of tryptophan-cleavage peptides from GVP protein (420 nmol) on Sephadex G-50*
The column (110 cm × 2.5 cm i.d.) was eluted with 25% formic acid. The fractions encompassed by the bars were pooled and used in sequence analysis.

nitrile; peptide T1 eluted at 47% (v/v) acetonitrile. Peptide T2 was taken without further purification from fractions 75–82 (see Fig. 2).

Sequence analysis

Manual sequence analysis of peptides was performed by a scaled-down version of the method of Chang *et al.* (1978), which has been described elsewhere (Runswick & Walker, 1983). Intact protein and other peptides were sequenced with the aid of a Beckman sequencer (Runswick & Walker, 1983). Polybrene (Tarr *et al.*, 1978; Klapper *et al.*, 1978) (2 mg) was precycled for two cycles in the spinning cup before addition of the sample. The degradation program was that described by Brauer *et al.* (1975), 0.1 M-Quadrol being used. Conversion of thiazolinones into phenylthiohydantoins was effected with an autoconvertor as described by Wittman-Liebold *et al.* (1976). Amino acid phenylthiohydantoin derivatives were identified by h.p.l.c. (Brock & Walker, 1980). In addition, many samples were hydrolysed with $HCl/SnCl_2$ (Mendez & Lai, 1975) and subjected to amino acid analysis.

Results

Amino acid analysis

The amino acid composition of the entire GVP molecule is shown in Table 1. The composition is broadly similar to that obtained in previous analyses (Falkenberg *et al.*, 1972). The same 15 amino acids were found in much the same rank order; cysteine, methionine and histidine were absent. The major difference was in isoleucine, now estimated to be 8.2 mol/100 mol of GVP, whereas it was previously estimated to be 10.1 mol/100 mol. None of the estimates of the other amino acids differed by as much as 1 mol/100 mol from the previous values.

Table 1. *Amino acid composition of GVP from Anabaena flos-aquae and its N-terminal tryptophan peptide (W1)*
Abbreviation used: n.d., not determined.

	Composition (mol/mol)			
		GVP	W1	
Amino acid	Native GVP	eluted from gel	Observed	From sequence
Aspartic acid	11.2	11.1	3.7	4
Threonine*	10.1	10.2	1.1	1
Serine*	21.0	22.3	3.2	4
Glutamic acid	25.0	25.2	2.1	2
Proline	2.8	3.2	0	0
Glycine	8.6	15.0	1.0	1
Alanine	31.8	34.0	3.2	3
Cysteine	n.d.	n.d.	n.d.	n.d.
Valine	22.8	27.1	2.2	3
Methionine	0	0	0	0
Isoleucine†	15.9	15.1	3.2	4
Leucine†	20.8	20.1	2.1	2
Tyrosine	5.7	6.2	0	0
Phenylalanine	1.0	1.1	0	0
Histidine	0	0	0	0
Lysine	9.0	8.1	1.7	2
Arginine	8.0	8.1	0.8	2
Tryptophan	n.d.	n.d.	n.d.	n.d.

* Extrapolated to zero time.
† After hydrolysis for 72 h.

Falkenberg *et al.* (1972) observed that eight of the nine least-abundant residues occurred in near-integer ratios, and on the basis of this proposed an empirical molecular weight for the protein of 15100, which happened to be close to the estimate suggested by electrophoresis. Our improved estimate of the least-abundant residue, phenylalanine, suggests an empirical molecular weight which is 1.36 times larger. We estimate that, assuming there is only one phenylalanine residue (see below), there is a total of 196 amino acid residues in GVP, with a total molecular weight of 20600.

Electrophoretic analysis

When a milky-white suspension of intact gas vesicles was mixed with SDS to a final concentration of 5% (w/v), the suspension remained turbid, indicating that a substantial proportion of the vesicles present was not solubilized. However, when the suspension was applied to an electrophoresis gel, a single mobile band was obtained. Solubility was increased by freezing the suspension overnight or heating it at 100°C for a few minutes. In each case only the same single band was obtained on electrophoresis. On a 15%-polyacrylamide slab gel run with various standard proteins of known molecular

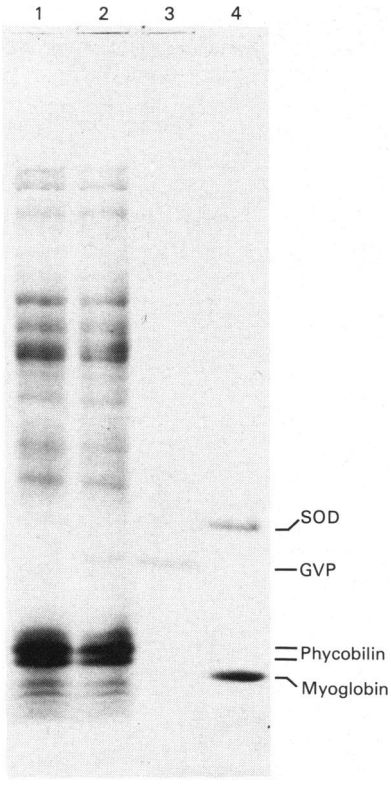

Fig. 3. *Polyacrylamide-gel electrophoretograms of GVP* SDS/polyacrylamide slab gel showing GVP (track 3) running midway between markers (track 4) of mol.wt. 17200 (myoglobin) and 22700 (superoxide dismutase, SOD). Track 2 shows the whole protein extract of *Anabaena flox-aquae* wild-type cells showing the subunits of phycobilin proteins (mol.wt. ~18000) running ahead of GVP; track 1 shows the extract of a GVP$^-$ mutant that does not produce GVP.

weight, the GVP ran slightly slower than myoglobin (mol.wt. 17200) and slightly faster than superoxide dismutase (22700) (see Fig. 3), and plots of mobility against log (molecular weight) indicated the molecular weight of the GVP band to be about 20900, in close agreement with the value obtained by amino acid analysis.

The proportion of GVP that formed a mobile band varied from one sample to another. Usually a proportion of the protein remained in the well at the top of the gel. When the gas vesicles were first dissolved in 80% formic acid (in which they formed a clear solution) the formic acid was removed by dialysis and the precipitated protein resuspended in 5% SDS, the same mobile band was obtained in addition to what appeared to be an oligomeric series of dimer, trimer etc., back to a streak of high-molecular-weight material at the origin of the running gel (see Fig. 3).

In order to determine whether the composition of the mobile band was similar to that of the original gas-vesicle sample, a gel was run loaded with gas vesicles in 5% SDS in a band 8cm wide at the origin. After running the gel, the band which had migrated was located by dunking the slab gel in a solution of 1M-KCl (Nelles & Bamburg, 1976). The band showed as a clear area surrounded by a white turbid background. The band was cut from the gel and the protein was eluted and purified on Bio-Gel P6 by the methods described by Walker *et al.* (1982). Hydrolysis yielded the composition shown in Table 1, which, apart from minor differences, was the same as the composition of the starting material. Therefore it was concluded that the migrating band was representative of the whole sample.

It has been observed that samples of intact gas vesicles stored in aqueous suspension for 1 year or more can no longer be dissolved in SDS. The protein appears to become denatured on aging, although the vesicles, while becoming weaker, retain their essential physical properties of rigidity, permeability to gas and impermeability to liquid water.

N-Terminal sequence

The amino acid sequence of the first 34 residues of GVP (see Fig. 4) was obtained from a 1mg sample of intact gas vesicles (equivalent to 50nmol) in the spinning-cup sequencer. The start of the sequence, AVEK, was identical with that obtained by Weathers *et al.* (1977) for *Microcystis* GVP. The yield of the first residue, alanine, was 41nmol, or 82%, of the starting material, assuming that GVP has a molecular weight of 20600.

Sequences of tryptophan-cleavage products

The products of the tryptophan-cleavage reaction were partially separated on Sephadex G-50 and were eluted in 25% (w/v) formic acid. The A_{280} of the eluate indicated that most of the peptide material was eluted in the exclusion volume of the Sephadex column, with a minor constituent being eluted subsequently (see Fig. 1).

The minor component started with the *N*-terminal sequence of GVP, AVEK . . ., and was evidently the *N*-terminal peptide embracing residues 1–28 (tryptophan).

The major component gave two sequences in the spinning cup: (*a*) residue 29 onwards, as indicated by comparison with the previous sequence from the entire protein (VRVS . . .); and (*b*) the *N*-terminal sequence (AVEK . . .). This second sequence could have been either the incompletely separated peptide (residues 1–28) or the entire GVP left by incomplete reaction with the cleavage reactants. The latter was

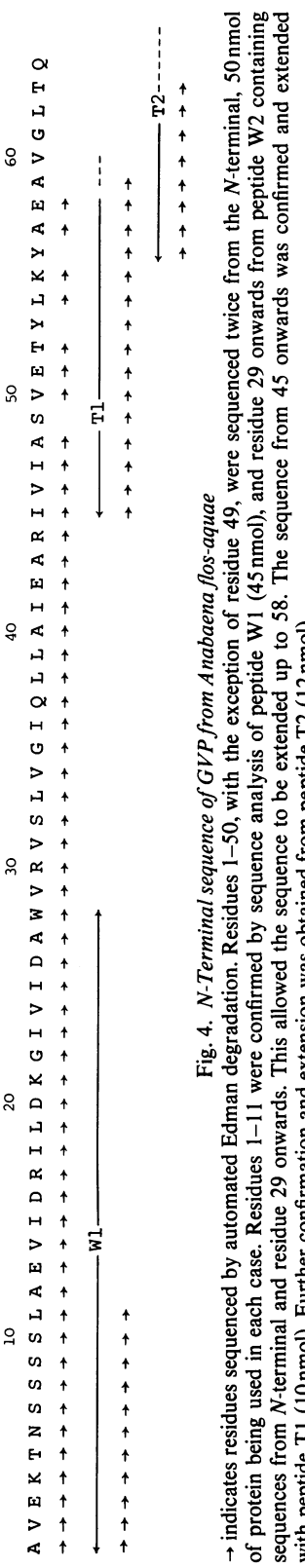

Fig. 4. *N-Terminal sequence of GVP from Anabaena flos-aquae*

→ indicates residues sequenced by automated Edman degradation. Residues 1–50, with the exception of residue 49, were sequenced twice from the *N*-terminal, 50nmol of protein being used in each case. Residues 1–11 were confirmed by sequence analysis of peptide W1 (45nmol), and residue 29 onwards from peptide W2 containing sequences from *N*-terminal and residue 29 onwards. This allowed the sequence to be extended up to 58. The sequence from 45 onwards was confirmed and extended with peptide T1 (10nmol). Further confirmation and extension was obtained from peptide T2 (12nmol).

more likely in view of the elution position of this fraction from the Sephadex G-50 column, and was confirmed by the slight overlap at residues 29 and 30.

Because the *N*-terminal sequence was known, it was possible to deduce the second sequence by subtraction. The spinning cup gave sequence information for 30 cycles, i.e. 1–30 on (*a*) and 29–58 on (*b*), with some uncertainties, for example at cycle 12, where the residue is alanine in both sequences, and at residues 49, 53, 56 and 57. These ambiguities were subsequently resolved by repeating the above experiment with succinylated GVP; the *N*-terminal sequence of uncleaved succinylated GVP was then blocked and the single sequence of residue-29 onwards was obtained in isolation. Further checks were provided by sequences of tryptic peptides obtained from this region.

Sequences of tryptic cleavage products

Incubation of formic acid-denatured protein with trypsin gave incomplete cleavage. From the digest a number of 'partial' products with clean *N*-terminal sequences were isolated. Sequences were identified corresponding to cleavages after residues 4, 17 and 30 (results not shown). Peptides T1 and T2 isolated from the same digest allowed us to extend the sequence derived from the *N*-terminus and tryptophan-cleavage reaction products beyond residue-60. The major peak (Fig. 2) of the initial fractionation of the digest contained only sequences corresponding to the *N*-terminus of the protein and from residue-5, suggesting that most of the protein was little affected by trypsin. Indeed the key to the determination of the rest of the sequence of the protein may be to find conditions leading to more extensive hydrolysis by proteinases.

Discussion

Previous X-ray studies of the crystalline structure of the gas-vesicle wall have shown that GVP molecules are arranged in a regular fashion along 4.6nm-wide ribs that are oriented perpendicular to the long axis of the gas-vesicle cylinder. Crystallography indicated that the walls were 1.9nm thick, formed from two layers of β-sheet with the polypeptide chains in the β-sheet inclined at an angle of 35° to the long axis of the vesicle. A pair of (probably antiparallel) chains in each layer generated a unit cell repeating at intervals of 1.15nm along the rib, with an estimated volume of about 10nm^3 and molecular weight of about 8000 (Blaurock & Walsby, 1976). A slightly smaller unit cell is suggested by measurements indicating that the mean thickness of the wall is only 1.54nm (Walsby & Armstrong, 1979).

Our new estimate of the molecular weight of GVP,

(a) A V E K T N S S S S L A E V I D R I L D K G I V I D A W V R V S L V G I Q L L A I E A R I V I A S V E T Y L K Y A E A V G L T Q
(b) A V E K I L D K/G I V I D A A - R Y A E A V G L T Z

10 20 30 40 50 60

Fig. 5. *Alignment of N-terminal sequence of GVP of Anabaena flos-aquae (a) with peptides isolated from digests of GVP protein from Microcystis aeruginosa (b)*
'–' indicates a deletion in sequence (*b*) and '/' that there is no overlap at this point.

20600, exceeds the size of the X-ray subunits by approx. 3-fold. Therefore a simple interpretation of the molecular structure is of the polypeptide snaking into six pairs of antiparallel chains, three in each layer. A molecule of this size would repeat at intervals of 3.4 nm (3×1.15 nm) along the rib. There is no evidence of a reflexion indicating this periodicity on the rib axis in the diffraction patterns of Blaurock & Walsby (1976), and this seems to suggest that the neighbouring pairs of β-chains in adjacent GVP molecules must be aligned with one another in the same way as the neighbouring pairs within the same molecule, giving a smooth profile.

Each run of the β-sheet chain across the rib must have about the same length (4.6 nm/cos 35°, or 5.6 nm) equivalent to 15 amino acid residues, or 30 residues for an antiparallel pair. It is possible that neighbouring pairs of chains in the same layer might have similar sequences, and if the primary polypeptide were pleated contiguously in one layer before folding to the next, the sequence would repeat at an interval of 30 residues. Such a repeat can be seen in the putative *C*-terminal sequence of the *Microcystis* GVP reported by Weathers *et al.* (1977), though the evidence for it is not strong, since the middle 15 residues (in the *Microcystis* GVP fragment TlPlb) have not been sequenced. The repeating sequence claimed by Weathers *et al.* (1977) is the octapeptide AEAVGLTZ starting 14 and 44 residues from the *C*-terminus of the molecule. A third repeat of this sequence was also reported, though its position has not been determined. Hitherto, no evidence of this repeat has been found in the *Anabaena* sequence.

We have also found this same octapeptide in *Anabaena* GVP, though a short overlap indicates that it starts at residue 57 from the *N*-terminus, which is much further than 30 residues from its next occurrence near the *C*-terminus (95 residues away if GVP is 196 residues long).

As shown in Fig. 5, there is a high degree of homology between the *N*-terminal sequence of *Anabaena* GVP and the peptide fragments of *Microcystis* GVP sequenced by Weathers *et al.* (1977). This homology permits the alignment of the *N*-terminal sequence of *Anabaena* GVP and four of the seven unique peptide fragments of *Microcystis* GVP sequenced by Weathers *et al.* (1977). In these four the homology is 100%. A fifth peptide, the *Microcystis* GVP fragment NPT, shows homology with the *Anabaena* sequence at residues 31–40, but as the identities of 8 out of 15 of the NPT residues are uncertain, it is not possible to be sure of this alignment. Of the two remaining unique *Microcystis* fragments, one is the *C*-terminus sequence and the other is a solitary tryptic lysine residue. The overall degree of homology so far encountered in the GVP of *Anabaena* and *Microcystis* is very high. The two organisms are not closely related, the former being a

filamentous cyanobacterium from the order Nostocales and the latter a colonial unicellular cyanobacterium from the Chroococcales. GVP must therefore be a highly conserved protein. This may reflect the large number of constraints on a single protein providing all the physical properties required of the gas vesicle. Not only must it self-assemble to create the gas-filled structure, but it must provide rigidity, the required resistance to collapse under pressure, a hydrophobic inner surface and a hydrophilic outer one (see Walsby, 1978). It has been proposed that the rigidity and strength is provided by the way in which the two layers of β-sheet interlock (Blaurock & Walsby, 1976), a product of the amino acid sequence. The hydrophilic outer surface must be provided by runs of residues with hydrophilic side chains, like the sequence EKTNSSSS near the *N*-terminus (see Fig. 5).

From the analyses carried out on *Anabaena* and *Microcystis* GVP, sequence information has been recovered only from the first 64 and last 45 residues; there is a central core of over 80 residues from which no peptides have been obtained. However, the overall amino acid composition of the remainder is not dissimilar from that which has been sequenced.

Inevitably, with difficult proteins, one looks at the possibility of sequencing them indirectly from the coding DNA base sequence which in some cases can be isolated by using synthetic oligonucleotide probes. The sequence in *Anabaena* GVP with the lowest codon redundancy is that starting at residue 25; 24 different 16-base oligonucleotides are required to provide all the possible codons for the sequence IDAW and the first two bases of the possible codons for V (valine). Since there are no other tryptophan residues and no methionine in the molecule, the minimum number possible in the unsequenced part of the GVP molecule will be 16 (from a run of four two-codon amino acids).

We thank Mr. F. D. Northrop, Mr. M. J. Runswick and Mr. A. F. Carne for assistance with sequencing and electrophoresis, Mr. P. K. Hayes for running the electrophoretogram shown in Fig. 3, and Dr. R. Henderson for useful discussions on the structure of gas-vesicle protein. This work was supported in part by a grant from the Science and Engineering Research Council to A. E. W.

References

Armstrong, R. E. & Walsby, A. E. (1981) in *Organisation of Prokaryotic Cell Membranes* (Ghosh, B. K., ed), vol. 2, pp. 95–129, CRC Press, Cleveland

Artavanis-Tsakonas, S. & Harris, J. I. (1980) *Eur. J. Biochem.* **108**, 599–611

Blaurock, A. E. & Walsby, A. E. (1976) *J. Mol. Biol.* **105**, 183–199

Brauer, A. W., Margolies, M. H. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035

Brock, C. J. & Walker, J. E. (1980) *Biochemistry* **19**, 2873–2882

Buckland, B. A. & Walsby, A. E. (1971) *Arch. Mikrobiol.* **79**, 327–337

Chang, J. Y., Brauer, D. & Wittmann-Liebold, B. (1978) *FEBS Lett.* **93**, 205–214

Falkenberg, P., Buckland, B. & Walsby, A. E. (1972) *Arch. Mikrobiol.* **85**, 304–309

Harris, J. I., Hocking, J. D., Runswirk, M. J., Suzuki, K. & Walker, J. E. (1980) *Eur. J. Biochem.* **108**, 599–611

Jones, D. D. & Jost, M. (1970) *Arch. Mikrobiol.* **70**, 43–64

Jones, D. D. & Jost, M. (1971) *Planta* **100**, 277–287

Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85**, 126–131

Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685

Mendez, E. & Lai, C. Y. (1975) *Anal. Biochem.* **68**, 47–53

Nelles, L. P. & Bamburg, J. R. (1976) *Anal. Biochem.* **73**, 522–531

Runswick, M. J. & Walker, J. E. (1983) *J. Biol. Chem.* in the press

Savige, W. E. & Fontana, A. (1977) *Methods Enzymol.* **47**, 442–453

Tarr, G. E., Beecher, J. F., Bell, M. & McKean, D. J. (1978) *Anal. Biochem.* **84**, 622–627

Thompson, E. O. P. & Sanger, F. (1963) *Biochim. Biophys. Acta* **71**, 488–491

Wachter, E. & Werhahn, R. (1980) in *Methods in Peptide and Protein Sequence Analysis* (Birr, C., ed.), 21–33, Elsevier/North-Holland Biomedical Press, Amsterdam

Walker, J. E., Carne, A. F., Runswick, M. J., Bridgen, J. & Harris, J. I. (1980) *Eur. J. Biochem.* **108**, 549–565

Walker, J. E., Auffret, A. D., Carne, A., Gurnett, A., Hanisch, P., Hill, D. & Saraste, M. (1982) *Eur. J. Biochem.* **123**, 253–260

Walsby, A. E. (1972) *Bacteriol. Rev.* **36**, 1–32

Walsby, A. E. (1977) *Arch. Microbiol.* **114**, 167–170

Walsby, A. E. (1978) *Symp. Soc. Gen. Microbiol.* **28**, 327–358

Walsby, A. E. (1982) *Proc. R. Soc. London Ser. B* **216**, 355–368

Walsby, A. E. & Armstrong, R. E. (1979) *J. Mol. Biol.* **129**, 279–285

Walsby, A. E. & Buckland, B. (1969) *Nature (London)* **224**, 716–717

Weathers, P. J., Jost, M. & Lamport, D. T. A. (1977) *Arch. Biochem. Biophys.* **178**, 226–244

Wittman-Liebold, B., Graffunder, H. & Cohls, H. (1978) *Anal. Biochem.* **75**, 621–633