Journal of Bioenergetics and Biomembranes, Vol. 24, No. 6, 1992

Function and Biosynthesis of Gas Vesicles in Halophilic *Archaea*

Felicitas Pfeifer[1] and Christoph Englert[1]

Received May 30, 1992; accepted June 15, 1992

The proteinaceous gas vesicles produced by various microorganisms including halophilic *Archaea* are hollow, gas-filled structures with a hydrophobic inner and a hydrophilic outer surface. The structural components of gas vesicles and their biosynthesis are still under investigation; an 8-kDa polypeptide appears to be the major constituent of the gas-vesicle envelope. Genetic analysis of the halobacterial gas-vesicle synthesis revealed an unexpected complexity: about 14 genes organized in three transcription units are involved in gas-vesicle structure, assembly, and gene regulation. Here we describe the comparison of three different genomic regions encoding gas vesicles in *Halobacterium salinarium* (p-vac and c-vac regions) and *Haloferax mediterranei* (mc-vac region) and speculate on the function of the gene products involved in gas-vesicle synthesis.

KEY WORDS: Vac region; *Halobacterium*; *Haloferax*; Vac mutants.

INTRODUCTION

Gas vesicles are hollow, gas-filled protein structures that are synthesized by various bacteria such as cyanobacteria, *Ancylobacter aquaticus*, and by the *Archaea Halobacterium* spp., *Haloferax mediterranei*, and *Methanosarcina barkeri* (Walsby, 1972; Stoeckenius and Kunau, 1968; Van Ert and Staley, 1971; Rodriguez-Valera *et al.*, 1983; Archer and King, 1983). They provide buoyancy to the cell and allow the positioning at a favorable depth for growth in an aquatic environment.

The hollow space of the gas vesicle is filled by passive diffusion of gasses dissolved in the cytoplasma. The hydrophilic outer surface of the wall protects the gas vesicle from surface tension, while the hydrophobic inner surface prevents liquid water from entering (Walsby, 1972). Gas vesicles are seen as 0.2–1.5-μm-long regular-shaped structures under the electron microscope, while in phase-contrast microscopy aggregates of gas vesicles (referred to as gas vacuoles) are seen as bright, light-refractile bodies. Among halophilic *Archaea*, the extremely halophilic species *Halobacterium salinarium* (formerly known as *H. halobium* and *H. salinarium species*) and related isolates (*Halobacterium* spp. GN101, GRB, GRA, YC819-9, and SB3) produce gas vesicles, as well as the moderate halophile *Haloferax mediterranei* (Houwink, 1956; Stoeckenius and Kunau, 1968; Rodriguez-Valera *et al.*, 1983; Surek *et al.*, 1988; Horne *et al.*, 1988). In addition, various *Haloarcula* strains ("square bacteria") isolated from hypersaline pools on the Sinai Peninsula contain gas vesicles (Parkes and Walsby, 1981). Unfortunately none of the latter strains could be grown in the laboratory.

This paper summarizes our results on the analysis of halobacterial gas vesicles at the molecular level. We investigated the structure of gas vesicles in various halobacterial strains under the electron microscope and analyzed their protein composition. The genetic determinants necessary for gas vesicle synthesis in two different halobacterial species, *H. salinarium* and *Hf. mediterranei*, were identified by transformation experiments and characterized. Experimental details have been published elsewhere (Horne *et al.*, 1988;

[1]Max-Planck-Institut für Biochemie, W-8033 Martinsried, Germany.

0145-479X/92/1200-0577$06.50/0

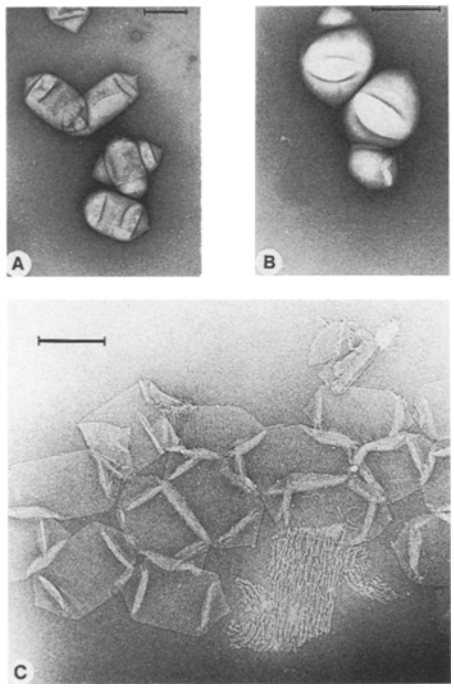

Fig. 1. Electron micrographs of isolated gas vesicles. (A) cylindrical gas vesicles of *Hf. mediterranei*; (B) spindle-shaped gas vesicles of *H. salinarium* PHH1; (C) collapsed gas vesicles of *Hf. mediterranei* indicating the ribbed structure. At the bottom of this picture there is a disintegrated gas vesicle showing the cylindrical part and the cap structures. The bar represents 0.2 μm.

Horne and Pfeifer, 1989; Blaseio and Pfeifer, 1990; Englert *et al.*, 1990; Horne *et al.*, 1991).

RESULTS AND DISCUSSION

Structure and Shape of Gas Vesicles in Halophilic *Archaea*

Gas vesicles can be easily isolated from halobacteria by cell lysis in distilled water followed by low-speed centrifugation of the lysate at 60 × *g* for several hours (Simon, 1981). Under these conditions the halobacterial cells disintegrate, whereas gas vesicles are stable and can be collected as a pellicle from the surface after centrifugation. Depending on the organisms used for this analysis, different gas-vesicle shapes are seen under the electron microscope (Fig. 1): *H. salinarium* PHH1 wild-type (formerly *H. halobium* NRC817) contains spindle-shaped gas vesicles (Fig. 1B) throughout the growth cycle, whereas gas vesicles isolated from the related species GN101, YC819-9, GRB, and SB3, as well as from *H. salinarium* mutants with certain deletions in plasmid pHH1, are of cylindrical shape. These latter species and mutants synthesize gas vesicles only in the late stage of growth (Horne and Pfeifer, 1989). Similarly, *Hf. mediterrnei* gas vesicles are of cylindrical shape (Fig. 1A); their synthesis occurs in stationary growth, but only in cultures grown at salt concentrations higher than 17%; *Hf. mediterranei* grown at 15% total salt does not synthesize gas vesicles, indicating the salt-modulated expression (Englert *et al.*, 1990).

The gas vesicles shown in Fig. 1A,B are still intact; application of pressure such as centrifugation at more than 30,000 × *g* leads to the collapse of gas vesicles which then reveal the typical rib-shaped structure with a period of 4.6 nm (Fig. 1C). These ribs run perpendicular to the long axis of the gas vesicles. A similar structure can be seen with cyanobacterial gas vesicles (Walsby, 1972; Walsby and Hayes, 1989). So far it is not clear whether these ribs are constituted by a stack of hoops or a helix of low pitch.

Protein Composition of the Gas Vesicle

The lipid-free gas vesicle envelope consists mainly of a single protein or rather polypeptide (GvpA) with a molecular mass of about 8 kDa. The amino acid sequence of the GvpA protein is highly conserved among cyanobacteria and halobacteria (Fig. 2) (Walsby and Hayes, 1989; Englert *et al.*, 1990). The halobacterial GvpA sequences shown in Fig. 2 are deduced from the respective *gvpA* genes (see below). Determination of the N-terminal amino acid sequence of a gas vesicle preparation reveals exclusively the sequence of the GvpA protein encoded by the respective *gvpA* gene expressed (Surek *et al.*, 1988; Horne and Pfeifer, 1989; Walsby and Hayes, 1989; Englert *et al.*, 1990). In addition, the amino acid composition of a *Hf. mediterranei* gas-vesicle preparation is in good agreement with the amino acid composition of the mc-GvpA protein deduced from the mc-*gvpA* sequence (Englert *et al.*, 1990). It is possible, however, that other constituents located on the outer gas vesicle surface is washed off during the purification procedure.

Unfortunately the protein component(s) of the gas vesicle cannot be separated as subunits on SDS-polyacrylamide gels. Gas vesicles strongly resist solubilization in the aqueous solutions necessary for electrophoresis; gas vesicles resuspended in formic acid, lyophilized, and boiled in loading buffer prior to electrophoresis do not indicate separated protein subunits

```
Hmc    MV   QPDSS SLAEVLDRVL DKGVVVDVWA RISLVGIEIL TVEARVVAAS VDTFLHYAEE IAKIEQAELT AGAEA   AP TPEA     78 aa
Hsc    .A   ..... ..........  ..........  ..........  ..........  ..........  ..........  .....PEP..   ..     79 aa
Hsp    .A   ..... G.........  ..........  .V........  ..........  ..........  ..........  .....   ..   ..     76 aa
Cal    ..VEKTNS.. .....I..I.  ...I...A.V  .V......L.  AI...I.I..  .E.Y.K...A VGLTQS.AVP .          71 aa
Ana    ..VEKTNS.. .....I..I.  ...I.I.A.V  .V.....QL.  AI...I.I..  .E.Y.K...A VGLTQS AVP .A        71 aa
Mic    ..VEKTNS.. .....I..I.  ...I.I.A..  .V......L.  AI.....I..  .E.Y.K...A VGLTQX.XXA ?         70 aa
```

Fig. 2. Comparison of GvpA sequences of halobacteria and cyanobacteria. The sequences derive from: Hmc, *Hf. mediterranei*; Hsc, *H. salinarium* c-vac encoded; Hsp, *H. salinarium* p-vac encoded; Cal, *Calothrix*; Ana, *Anabaena flos-aquae*; and Mic, *Microcystis*.

on SDS-polyacrylamide gels; almost the whole material is found as a smear near the top of the separating gel. Using phenol/urea/acetic acid gels, a minor 19–20-kDa protein is detectable (Simon, 1980; Surek *et al.*, 1988) which, however, does not correspond in size to the 8-kDa GvpA protein. So far, the nature of this 20-kDa protein is unclear.

The failure to separate gas-vesicle protein constituents on detergent-containing gels could be due to the high degree of hydrophobic amino acids present in the GvpA protein (Simon, 1980). Further investigations in our laboratory, however, showed that this was not the case: antibodies raised against gas-vesicle preparations detect the 8-kDa GvpA monomer in total protein preparations of *H. salinarium* and *Hf. mediterranei* separated on SDS-polyacrylamide gels, indicating that GvpA can be identified as a monomer before being inserted into the gas vesicle (C. Englert, unpublished). Once this protein is incorporated, the gas vesicle structure cannot then be disintegrated. Thus, the analysis of the protein subunits of gas vesicles is hampered by the stability and rigidity of this unique structure.

Analysis of the Gene Encoding the Major Gas Vesicle Protein GvpA

The *gvpA* gene has been isolated from various *H. salinarium* strains (i.e., *H. salinarium* PHH1 containing the 150-kbp plasmid pHH1, and *H. salinarium* strain 5 containing three different plasmids including a 65-kbp pHH1 derivative), as well as from *Hf. mediterranei* (Horne *et al.*, 1988; Englert *et al.*, 1990). Surprisingly, each *H. salinarium* strain contains two different but related *gvpA* loci: one (p-*gvpA*) is located on the respective plasmid (pHH1, or the 65-kbp pHH1 derivative); the second gene, c-*gvpA*, is found in the chromosome (Horne *et al.*, 1988; Horne and Pfeifer, 1989). The products encoded by these two *gvpA* genes differ only by two amino acid substitutions (positions 8 and 29) and a small insertion of three amino acids (PEP) near the carboxy terminus of c-GvpA (Fig. 2). The *gvpA* gene present on the 200-kbp plasmid pNRC100 of *H. salinarium* NRC-1 published by Das-Sarma *et al.* (1987) encodes an identical polypeptide (published as differing in the last 15 amino acids, but later corrected).

The product of the chromosomal mc-*gvpA* gene of *Hf. mediterranei* is almost identical to c-GvpA of *H. salinarium*, except for the PEP insertion, but contains a different insertion (TP) near the carboxy terminus and an A to V exchange at positon 2 (Fig. 2) (Englert *et al.*, 1990). Gas-vesicle proteins of cyanobacteria are highly related to each other and are smaller than the halobacterial GvpA proteins (Fig. 2). Their sequences also indicate a high degree of similarity to the halobacterial counterparts, especially in the central region of the sequence. The high degree of similarity was already recognized when the N-terminal amino acid sequences of these proteins were compared, and led to the speculation that cyanobacterial and halobacterial gas vesicles are homologous.

One species of GvpA proteins is found in gas vesicles from different *H. salinarium* strains: determination of the N-terminal sequence of the spindle-shaped gas vesicles of *H. salinarium* PHH1 indicates the p-GvpA sequence, whereas the cylindrical gas vesicles of a *H. salinarium* strain lacking the pHH1 plasmid and of the related *Halobacterium* species GN101, YC819-9, GRB, and SB3 indicate the sequence of the chromosomal c-GvpA product (Surek *et al.*, 1988). Due to their small size, the different GvpA proteins can be distinguished in Tricine–sodium dodecyl sulfate–polyacrylamide gels according to Schägger and von Jagow (1987): Western blots of total proteins from halobacteria probed with polyclonal antibodies raised against gas vesicles of *Hf. mediterranei* indicate the slightly different GvpA sizes (unpublished).

Transcription of the Various *gvpA* Genes during Growth Cycle and Promoter Analysis

Investigation of transcripts produced from both

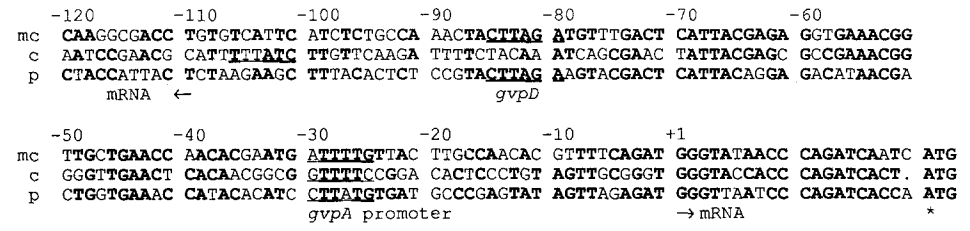

Fig. 3. Comparison of the DNA sequences between the two oppositely arranged *gvpA* and *gvpD* genes. The ATG start codon of *gvpA* is indicated by a star, and the start of the *gvpA* mRNA is labelled +1. The box A promoter motifs of *gvpA* (centered around −28) and *gvpD* (around position −80) pointing in the opposite direction are underlined. Arrows depict the mRNA start sites. Bases identical in all three vac regions are emphasized by bold letters. The regions derive from: mc, *Hf. mediterranei*; c, p, chromosomal and plasmid located vac sequence of *H. salinarium*.

gvpA genes present in *H. salinarium* PHH1 wild-type (p-*gvpA* and c-*gvpA*) indicates a high level of p-*gvpA* in the logarithmic growth phase (Horne and Pfeifer, 1989). The gene is transcribed producing a 340-nucleotide mRNA with a maximal level in the mid-logarithmic growth phase. In the stationary phase, the level of p-*gvpA* mRNA drops significantly. No c-*gvpA* mRNA can be detected, indicating that the chromosomal gene is not expressed in wild-type. However, in strains lacking the p-*gvpA* gene (i.e., pHH1 deletion variants), c-*gvpA* transcripts of 340 nt length are observed, with a maximal level in the early stationary growth phase. The time point of c-*gvpA* expression in the related *Halobacterium* species GN101, GRB, YC819-9, and SB3 (which all contain plasmid species unrelated to pHH1 and thus no p-*gvpA* gene) is similar to that of *H. salinarium* pHH1 deletion variants, and corresponds in time point and degree to the phenotypic occurrence of gas vacuoles in these strains (Horne and Pfeifer, 1989).

In *Hf. mediterranei* the mc-*gvpA* gene is expressed at a 7-fold higher level in cells grown in media containing more than 17% total salt compared to cells grown in 15% salt (Englert *et al.*, 1990). The maximal mc-*gvpA* mRNA level is seen in the early stationary growth phase, 40 h prior to the maximal gas vacuolation in the late stationary growth.

Differences in the *gvpA* promoters and respective upstream regions might account for these differences in *gvpA* mRNA synthesis during the growth cycle (Fig. 3). A high degree of sequence similarity is found in the mRNA leader. Upstream of the mRNA start site (position +1 in Fig. 3), however, the DNA sequence similarity drops significantly, and the promoter sequence (centered around position −28) is different in all three regions. The promoter of the p-*gvpA* gene (expressed in the log phase) resembles the consensus sequence of the archaeal promoter (Reiter *et al.*, 1990) to a greater degree than the promoters of either the c-*gvpA* or mc-*gvpA* genes (both are expressed in stationary growth). In line with the observed similarity between the p-*gvpA* and the archaeal consensus promoter, the p-*gvpA* promoter was shown to be recognized *in vitro* by the DNA-dependent RNA polymerase of the thermophilic archaeon *Sulfolobus shibatae* (Hüdepohl *et al.*, 1991). In this *in vitro* transcription experiment the mRNA start site was found to be identical to that identified *in vivo*. In contrast to this, neither the promoter of the c-*gvpA* nor of the mc-*gvpA* gene were recognized by the *Sulfolobus in vitro* transcription system. Both promoters appear to be only active in stationary growth and show a stretch of four thymidines in the center (Fig. 3).

Vac Mutant Analysis of *H. salinarium* PHH1

H. salinarium PHH1 exhibits an extremely high mutation frequency with respect to the gas-vesicle phenotype (Vac). Vac mutants occur with a frequency of 1% and are easily visible as transparent colonies on agar plates (Pfeifer *et al.*, 1981). In most cases, these Vac mutants are due to the integration of an insertion element near the p-*gvpA* gene (Pfeifer *et al.*, 1981, 1988; Horne *et al.*, 1991). Similar results were obtained with *H. salinarium* NRC-1 (Jones *et al.*, 1989). A minor amount (about 5%) of the Vac mutants are due to the deletion of a region encompassing the p-*gvpA* gene in plasmid pHH1 (Pfeifer *et al.*, 1988). The extremely high frequency of insertions and deletions in plasmid pHH1 is due to the activity of a variety of transposable elements (Pfeifer *et al.*, 1984; Pfeifer and Blaseio, 1989, 1990; Horne *et al.*, 1991). Ten different insertion elements of *H. salinarium* are known so far; these halobacterial ISH elements have

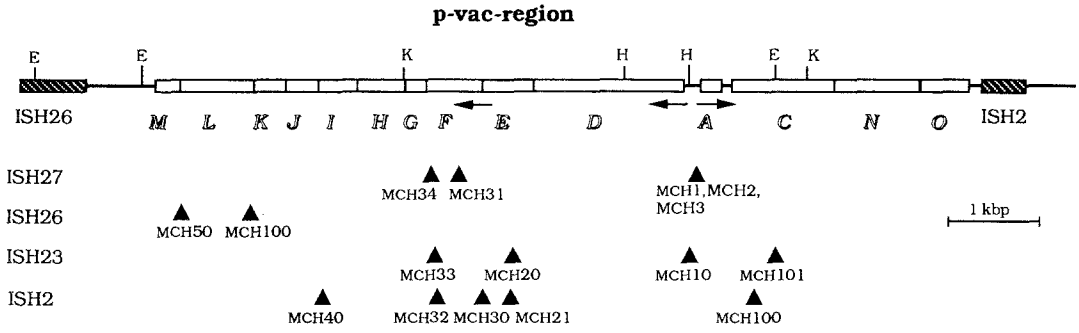

Fig. 4. Integration sites of insertion elements in various Vac mutants of *H. salinarium* PHH1. The *gvpA* gene (labelled A) and the various ORFs (labelled C through M) representing the p-vac region in plasmid pHH1 are indicated by white boxes. The ISH26 element located near p-*gvpM* and an ISH2 element adjacent to p-*gvpO* are indicated. Arrows depict the direction of transcripts. Triangles show the integration sites of ISH2, ISH23, ISH26, and ISH27 in the various Vac mutants (designated MCH). Restriction sites marked are: *Eco* RI (E), *Hin* dIII (H), and *Kpn* I (K).

similar structural features similar to those in *E. coli* IS elements: an element-characteristic inverted repeat is present at the termini, and during replicative transposition they duplicate target DNA upon integration (Pfeifer *et al.*, 1984; Pfeifer, 1985; Pfeifer and Blaseio, 1989; Pfeifer and Ghahraman, 1991).

In Vac mutants, four different types of insertion elements (ISH2, ISH23, ISH26, ISH27) are found in a 6-kbp region encompassing the p-*gvpA* gene (Fig. 4) (Horne *et al.*, 1991). The mutation always leads to a Vac^- phenotype, indicating that more than just the p-*gvpA* gene is necessary for gas-vesicle synthesis. Most of these insertion mutants stay Vac^- throughout the growth cycle, indicating that the chromosomal c-*gvpA* gene region does not complement these mutants.

Determination of the Size of the Genomic Region Required for Vac Synthesis by Transformation Experiments

Transformation experiments confirmed the hypothesis that the respective *gvpA* gene is not sufficient for gas-vesicle synthesis. Various fragments derived from the genomic region surrounding the mc-*gvpA* and p-*gvpA* gene were cloned into the halobacterial vector plasmids pWL102 or pUBP2 and analyzed for the ability to transform the Vac^- species *Haloferax volcanii* to Vac^+ (Blaseio and Pfeifer, 1990; Horne *et al.*, 1991). Whereas a 12.7-kbp *Pst* I fragment (PP) of *Hf. mediterranei* containing the mc-*gvpA* gene revealed Vac^+ transformants, two subfragments (HP and PS, see Fig. 5) revealed Vac^- transformants. A more detailed deletion analysis of the 12.7-kbp mc-vac fragment finally defined the minimal size of the mc-vac region: the 9.5-kbp fragment BN contained all genes required for gas-vesicle synthesis in *Hf. mediterranei* (Englert *et al.*, 1992).

Similar results are found with the plasmid-located p-vac region of *H. salinarium*: the two subfragments KK and EE did not reveal Vac^+ transformants, while the larger fragment EB contained all the information required for gas-vescile synthesis (see Fig. 5) (Horne *et al.*, 1991; Englert *et al.*, 1992). The Vac^+ phenotype of the latter *Hf. volcanii* transformants also demonstrates that the p-vac region is sufficient and does not depend on products of the chromosomal c-vac-region.

Comparative DNA Sequence Analysis of the Three vac Regions

The DNA sequence determination of the three regions surrounding the respective *gvpA* genes and subsequent comparison of these sequences reveals similarity throughout a region of approximately 9 kbp (Horne *et al.*, 1991; Englert *et al.*, 1992). Comparison of the DNA sequences further upstream or downstream indicates no significant similarity. The homologous regions are most likely comprised of the genes necessary for gas-vesicle synthesis, and are designated mc-vac region (*Hf. mediterranei*), p-vac region (*H. salinarium* plasmid pHH1), and c-vac region (*H. salinarium* chromosome) (Horne *et al.*, 1991; Englert *et al.*, 1992). The size of the vac regions correlates well with the results obtained by transformation experiments and with the distribution of insertion elements in the p-vac region of Vac mutants. In plasmid pHH1

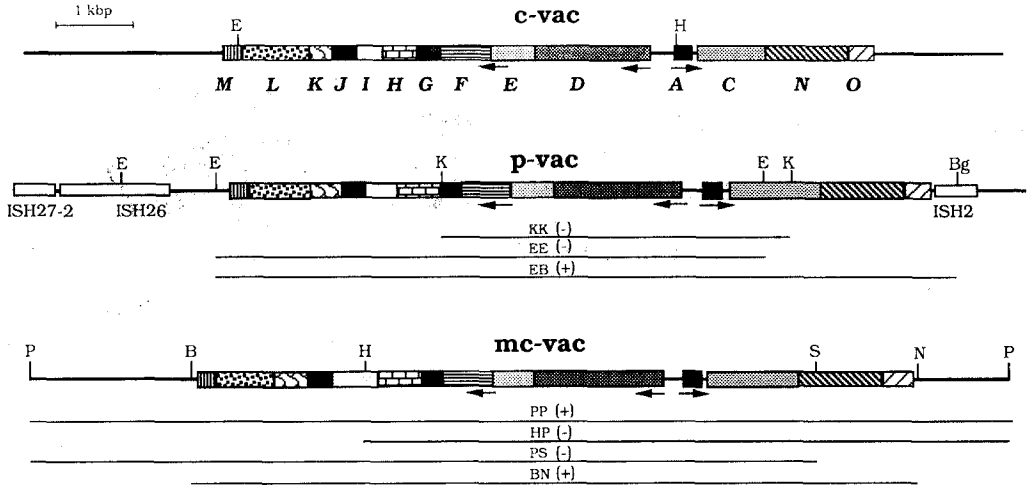

Fig. 5. Comparison of the three halobacterial vac regions and location of the various subfragments used in transformation experiments. The *gvp* ORFs representing each vac region are indicated by patterned boxes and labelled C through O. The *gvpA* gene is indicated by a black box. Arrows indicate the direction of transcripts. Restriction sites indicated are: *Bgl* I (B), *Bgl* II (Bg), *Eco* RI (E), *Hin* dIII (H), *Kpn* I (K), *Nhe* I (N), *Pst* I (P), and *Sph* I (S). The lines underneath the p-vac and mc-vac region represent fragments used in transformation experiments. Fragments marked with (+) revealed Vac^+ transformants, whereas (−) indicates fragments that revealed Vac^- *Hf. volcanii* transformants.

the p-vac region is surrounded by two insertion elements, ISH2 and ISH26 (Fig. 5).

In each vac region, 13 open reading frames (ORFs) are present in addition to the *gvpA* gene, ten of which (*gvpD/E/F/G/H/I/J/K/L/M*) are found upstream of *gvpA* and in the opposite orientation (Fig. 5). Downstream of the *gvpA* gene we found three ORFs (*gvpC/N/O*) in each case. ORFs *gvpD* through *gvpM* are consecutive, with stop codons frequently overlapping the start codon of the following ORF (Horne *et al.*, 1991, Englert *et al.*, 1992). Part of the p-vac region present on plasmid pNRC100 is published by Jones *et al.* (1991); *gvpN* and *gvpO* were, however, not identified by these authors.

The mc-vac region is the largest (9.4 kbp), while c-vac (8.4 kbp) is the smallest vac region. The mc-vac region encodes, in most cases, the largest Gvp proteins, whereas the products of the c-vac region are often the smallest ones (Englert *et al.*, 1992). Interestingly, the overall similarity between the mc-vac and p-vac region is much higher than between the p-vac and c-vac region presented in the same halobacterial cell (Table I). The highest conservation is observed with the *gvpA* gene (93–97% similarity), while the proteins encoded by *gvpD,F,G,J,K,M* and *gvpN* are more than 80% similar. The lowest conservation (still more than 50%) is found with products of *gvpC*, *gvpE*, and *gvpI* (Table I).

Transcript Analysis of the Different vac Regions

The transcription of these putative genes has been analyzed in *H. salinarium* PHH1 wild-type and p-vac deletion mutants as well as in *Hf. mediterranei.* The overlapping stop and start codons of the ORFs imply an expression as transcription units, with two of them oriented in the opposite direction. In all three regions, however, *gvpD-gvpM* are transcribed as two units: one transcript encompasses *gvpD* and possibly *gvpE*, and a second one *gvpF* through *gvpM*. Interestingly, the latter transcript occurs prior to the *gvpD/E* and *gvpA* mRNA, suggesting that the translated products are involved in early steps of gas-vesicle synthesis (Horne *et al.*, 1991; Englert *et al.*, 1992). In the mc-vac region, *gvpC/N/O* are cotranscribed with the *gvpA* gene, and their transcript occurs at the same time point in growth as *gvpD* mRNA. In contrast to our results with the p-vac region on pHH1, Jones *et al.* (1991) suggest

Table I. Percent Amino Acid Similarity of the Various Gvp Proteins

Gvp	M	L	K	J	I	H	G	F	E	D	A	C	N	O
p-vac/c-vac	88	65	82	83	50	66	85	82	70	81	97	67	82	75
c-vac/mc-vac	85	65	82	88	58	64	82	82	62	81	93	61	80	75
mc-vac/p-vac	89	75	88	81	76	74	83	86	71	83	95	73	84	82

that *gvpD-gvpM* on plasmid pNRC100 are transcribed as one unit.

A detailed mapping of the *gvpD* mRNA start site in the three vac regions reveals that the distance between the oppositely oriented *gvpA* and *gvpD* promoters is 50 bp in the case of the p-vac and mc-vac regions, and 70 bp in case of the c-vac region (Jones *et al.*, 1989; Horne *et al.*, 1991; Englert *et al.*, 1992). This intervening sequence is highly conserved (Fig. 3).

Characteristics of the Various GVP Proteins and Speculations on their Possible Function

The results obtained by comparative DNA sequence analysis, mutant analysis, and transformation experiments demonstrate the unexpected complexity of the genetic determinants involved in gas-vesicle synthesis. Questions arise about the possible role of the various products in Vac synthesis. These Gvp proteins could be minor structural components of the gas vesicles, "chaperone"-like proteins involved in protein folding during gas-vesicle assembly, or enzymes catalyzing energy-dependent steps. Another possible function could be DNA binding and *gvp*-gene regulation

The protein product of the *gvpC* gene located downstream of *gvpA* indicates 6–7 internal amino acid sequence repeats. A similar characteristic is found for a 20-kDa protein GvpC that is located on the outer surface of the cyanobacterial gas vesicle and strengthens its structure (Walsby and Hayes, 1988; Hayes *et al.*, 1992). The cyanobacterial gene encoding this GvpC protein is located downstream of two *gvpA* genes (*gvpA1/A2*) in *Calothrix*, and all three genes are transcribed as one unit (Damerval *et al.*, 1987). As yet, a similar location and function of the halobacterial GvpC protein could not be shown; however, the possession of this structural feature points to a similar function. Expression of the halobacterial *gvpC* gene in *E. coli* for protein purification and antibody formation is in progress and should help to investigate the location of this protein by immuno-electron microscopy.

The proteins GvpD and GvpN, encoded by the two late transcription units, both indicate a conserved sequence motif (P-loop; Saraste *et al.*, 1990) that is similar to the nucleotide-binding site in ATP/GTP binding proteins (Horne *et al.*, 1991; Englert *et al.*, 1992). The presence of this sequence motif suggests an involvement of these proteins in an energy-consuming step or in the phosphorylation of other Gvp protein(s). Interestingly, deletion of a large part of the mc-*gvpD* gene results in an overexpression of gas vesicles in *Hf. volcanii* transformants, indicating a regulatory role of GvpD in gas-vesicle synthesis (unpublished results).

So far, similar *gvp* genes are not described for cyanobacteria, except for the *gvpA1/A2/C* operon (Damerval *et al.*, 1987). No functional analyses by transformation experiments are described for cyanobacteria; however, according to our analyses, we expect a similar number of genes involved in gas-vesicle synthesis in these bacteria.

Early Steps in Gas-Vesicle Synthesis

The "early" transcription unit encompassing *gvpF-gvpM* encodes two conserved proteins, GvpJ and GvpM, that exhibit a similarity of more than 60% to GvpA, the major structural gas-vesicle protein (Fig. 6) (Horne *et al.*, 1991; Englert *et al.*, 1992). These similarities to GvpA imply that GvpJ and GvpM are structural components of the gas vesicle and, since they are synthesized early, that they are possibly involved in the formation of the "pre" gas-vesicle structure. Gas-vesicle synthesis in the cyanobacterium *Nostoc* starts with the formation of a small biconical structure (Waaland and Branton, 1969). This bicone is successively enlarged to a cylinder, most likely by the addition of more and more GvpA protein. It is possible that GvpJ and GvpM serve as starter molecules in the formation of these bicones, or are even constituents. Another speculative possibility is that both (or one of them) are later substituted by the GvpA protein. The various stages in gas-vesicle assembly are still unclear; the biconial structure could also constitute a "shape forming unit" in which a GvpA protein layer is introduced to form the cylindrical part of the gas vesicle. These hypotheses can be tested by analyzing the putative "pre"-gas vesicle structures formed by products of the early *gvp*-transcription unit in the respective *Hf. volcanii* transformants. Furthermore, expression of the GvpJ and GvpM in *E. coli* for subsequent antisera formation is in progress to investigate the location of these proteins in more detail by immuno-electron microscopy.

The peptides GvpG and GvpK indicate highly hydrophobic N-terminal parts, whereas the rest of the protein is hydrophilic (Englert *et al.*, 1992). These proteins could also serve as minor structural components in gas vesicles, or possibly as chaperones for

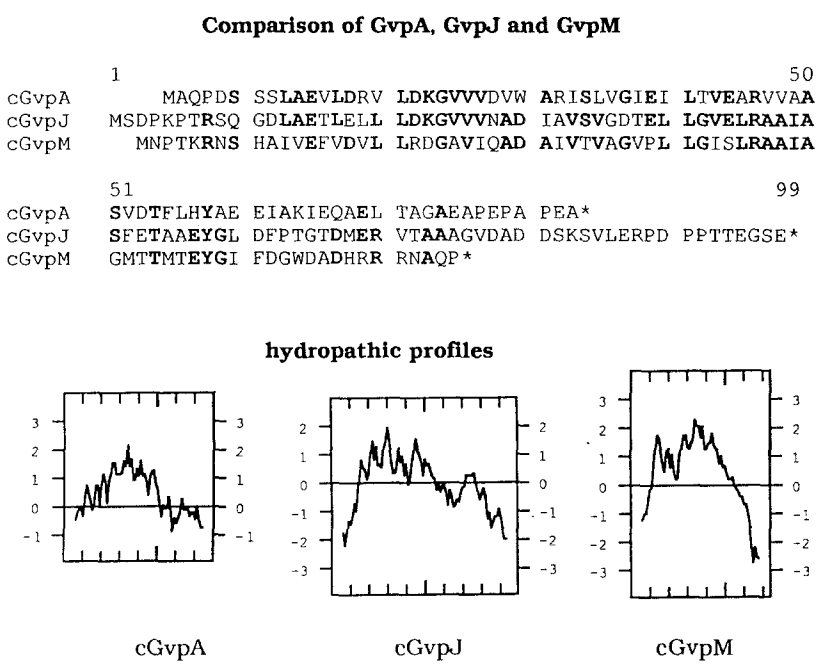

Fig. 6. Alignment of the GvpM, GvpJ, and GvpA protein of the c-vac region, and hydropathic profiles according to Kyte and Doolittle (1982).

GvpJ, GvpM, and/or GvpA, aiding the solubilization of these hydrophobic monomers in the cytoplasm until they are incorporated into the gas vesicle. The product of ORF *gvpI* is rather basic, typical for a DNA-binding protein.

CONCLUSIONS

For a long time, gas vesicles were thought to consist of a single protein, since only the N-terminal sequence of GvpA could be determined from a gas-vesicle preparation. Although GvpA appears to be the major constituent, the analysis of the genetic determinants involved in Vac synthesis has revealed the necessity of many other genes and proteins. Gas vesicles appear to be synthesized along a complex pathway, and further analysis should dissect the various steps in the near future.

ACKNOWLEDGMENTS

U. Blaseio, M. Horne, K. Krüger, and S. Offner are thanked for the contributions to the work reviewed in this paper, W. Zillig for support and discussions, W. Baumeister for help in electron microscopy, and M. Holmes for critical reading of the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Pf165/3-1 and Pf165/4-1).

REFERENCES

Archer, D. B., and King, N. R. (1983). *FEBS Microbiol. Lett.* **16**, 217–223.

Blaseio, U., and Pfeifer, F. (1990). *Proc. Natl. Acad. Sci. USA* **87**, 6772–6776.

Damerval, T., Houmard, J., Guglielmi, G., Csiszar, K., and Tandeau de Marsac, N. (1987). *Gene* **54**, 83–92.

DasSarma, S., Damerval, T., Jones, J. G., and Tandeau de Marsac, N. (1987). *Mol. Microbiol.* **1**, 365–370.

Englert, C., Horne, M., and Pfeifer, F. (1990). *Mol. Gen. Genet.* **222**, 225–232.

Englert, C., Offner, S., Krüger, K., and Pfeifer (1992). *J. Mol. Biol.*, **227**, in press.

Hayes, P. K., Buchholz, B., and Walsby, A. E. (1992). *Arch. Microbiol.* **157**, 229–234.

Horne, M., and Pfeifer, F. (1989). *Mol. Gen. Genet.* **218**, 437–444.

Horne, M., Englert, C., and Pfeifer, F. (1988). *Mol. Gen. Genet.* **213**, 459–464.

Horne, M., Englert, C., Wimmer, C., and Pfeifer, F. (1991). *Mol. Microbiol.* **5**, 1159–1174.

Houwink, A. L. (1956). *J. Gen. Microbiol.* **15**, 146–150.

Hüdepohl, U., Gropp, F., Horne, M., and Zillig, W. (1991). *FEBS Lett.* **285**, 257–259.

Jones, J. G., Hackett, N., Halladay, J., Scothorn, D., Yang, C., Ng, W. and DasSarma, S. (1989). *Nucleic Acids Res.* **17**, 7785–7793.

Jones, J. G., Young, D. C., and DasSarma, S. (1991). *Gene* **102**, 117–122.

Kyte, J., and Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.

Parkes, K., and Walsby, A. E. (1981). *J. Gen. Microbiol.* **126**,

503–506.
Pfeifer, F. (1985). *Syst. Appl. Microbiol.* **7**, 36–40.
Pfeifer, F., and Blaseio, U. (1989). *J. Bacteriol.* **171**, 5135–5140.
Pfeifer, F., and Blaseio, U. (1990). *Nucleic Acids Res.* **18**, 6921–6925.
Pfeifer, F., and Ghahraman, P. (1991). *Nucleic Acids Res.* **19**, 5788.
Pfeifer, F., Weidinger, G., and Goebel, W. (1981). *J. Bacteriol.* **145**, 375–382.
Pfeifer, F., Friedman, J., Boyer, H. W., and Betlach, M. (1984). *Nucleic Acids Res.* **12**, 2480–2497.
Pfeifer, F., Blaseio, U., and Ghahraman, P. (1988). *J. Bacteriol.* **170**, 3718–3742.
Reiter, W.-D., Hüdepohl, U., and Zillig, W. (1990). *Proc. Natl. Acad. Sci. USA* **87**, 9509–9513.
Rodriguez-Valera, F., Juez, G., and Kushner, D. J. (1983). *Syst. Appl. Microbiol.* **4**, 369–381.
Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). *Trends Biochem. Sci.* **15**, 430–434.
Schägger, H., and von Jagow, G. (1987). *Anal. Biochem.* **166**, 368–379.
Simon, R. (1980). *Electrophoresis 1980*, Vol. I, DeGruyter, Berlin, pp. 172–176.
Simon, R. (1981). *J. Gen. Microbiol.* **125**, 103–111.
Stoeckenius, W., and Kunau, W. H. (1968). *J. Cell Biol.* **38**, 337–357.
Surek, B., Pillay, B., Rdest, U., Bayreuther, K., and Goebel, W. (1988). *J. Bacteriol.* **170**, 1746–1751.
Van Ert, M., and Staley, J. T. (1971). *J. Bacteriol.* **108**, 236–240.
Waaland, J. R., and Branton, D. (1969). *Science* **163**, 1339–1341.
Walsby, A. E. (1972). *Bacteriol Rev.* **36**, 1–32.
Walsby, A. E., and Hayes, P. K. (1988). *Mol. Microbiol.* **2**, 545–552.
Walsby, A. E., and Hayes, P. K. (1989). *Biochem. J.* **264**, 313–322.