


RESEARCH ARTICLE

Open Access



The model cyanobacteria *Anabaena* sp. PCC 7120 possess an intact but partially degenerated gene cluster encoding gas vesicles

Kun Cai¹, Bo-Ying Xu², Yong-Liang Jiang¹, Ying Wang¹, Yuxing Chen¹, Cong-Zhao Zhou^{1†} and Qiong Li^{1*†} 

Abstract

Background: Bacterial gas vesicles, composed of two major gas vesicle proteins and filled with gas, are a unique class of intracellular bubble-like nanostructures. They provide buoyancy for cells, and thus play an essential role in the growth and survival of aquatic and soil microbes. Moreover, the gas vesicle could be applied to multimodal and noninvasive biological imaging as a potential nanoscale contrast agent. To date, cylinder-shaped gas vesicles have been found in several strains of cyanobacteria. However, whether the functional gas vesicles could be produced in the model filamentous cyanobacteria *Anabaena* sp. PCC 7120 remains controversial.

Results: In this study, we found that an intact *gvp* gene cluster indeed exists in the model filamentous cyanobacteria *Anabaena* sp. PCC 7120. Real-time PCR assays showed that the *gvpA* gene is constitutively transcribed *in vivo*, and its expression level is upregulated at low light intensity and/or high growth temperature. Functional expression of this intact *gvp* gene cluster enables the recombinant *Escherichia coli* to gain the capability of floatation in the liquid medium, thanks to the assembly of irregular gas vesicles. Furthermore, crystal structure of GvpF in combination with enzymatic activity assays of GvpN suggested that these two auxiliary proteins of gas vesicle are structurally and enzymatically conserved, respectively.

Conclusions: Our findings show that the laboratory strain of model filamentous cyanobacteria *Anabaena* sp. PCC 7120 possesses an intact but partially degenerated gas vesicle gene cluster, indicating that the natural isolate might be able to produce gas vesicles under some given environmental stimuli for better floatation.

Keywords: Gas vesicle, Cyanobacteria, Natural isolate, Heterologous expression, Crystal structure, ATPase activity

* Correspondence: liqiong@ustc.edu.cn

†Cong-Zhao Zhou and Qiong Li contributed equally to this work.

¹Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230027, Anhui, China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Gas vesicles (GVs), a unique class of intracellular bubble-like nanostructures, are found in many aquatic and soil microbes including halophilic archaea, photosynthetic bacteria, and heterotrophic bacteria [1]. Ambient gases could freely diffuse into and out of GV, whereas water is impermeable, making the GV a gas-filled organelle [2, 3]. GV could regulate the buoyancy of microbial cells, enabling the vertical floatation to an appropriate depth in aqueous environments for a better access of oxygen, light and even nutrients [4]. As an organelle composed of only proteins, GV adopts a spindle-shaped cylinder with conical end caps, usually of 45 ~ 250 nm in width and 100 ~ 2000 nm in length [5]. The unique physical properties allow GV to serve as a potential nanoscale contrast agent for ultrasound and magnetic resonance imaging, which yields multimodal and noninvasive biological imaging with high spatial and temporal resolution [6].

As previously reported, formation of GV is related to a conserved cluster of 8 ~ 14 genes (termed gas vesicle protein gene cluster, or *gvp* gene cluster for short), encoding two major structural proteins and several essential minor components that might putatively function as chaperones, nucleators and regulators [2, 5, 7]. The primary structural protein GvpA and the external scaffold protein GvpC constitute the 2-nm-thick outer amphiphilic shell of the GV [2, 5, 8]. GvpA, a 7.5-kDa highly conserved and hydrophobic protein, assembles into tandem arrays that form 4.6-nm-wide characteristic ribs running nearly perpendicular to the long axis of the GV [9, 10]. Notably, most cyanobacteria possess multiple copies of *gvpA* gene, for example, two in *Calothrix* sp. [11], three in *Microcystis aeruginosa* [12] and five in *Anabaena flos-aquae* [13]. In contrast, GvpC is a less-abundant, not conserved, and highly hydrophilic protein [14]. GvpC usually contains a number of conserved 33-residue repeating motif (33RR), and functions to connect GvpA molecules in the same and/or adjacent ribs to strengthen and stabilize the shell of GV [15]. In vitro experiments demonstrated that removal of GvpC leads to a three-fold decrease of the critical collapse pressure of GV, whereas addition of GvpC helps GV to restore normal strength [16, 17]. In addition, GvpF is reported to be a structural protein localized at the inner surface of GV [18].

To date, a series of cyanobacteria have been found to produce GV, such as *A. flos-aquae*, *Calothrix* sp. PCC 7601, *M. aeruginosa* PCC 7806, *Oscillatoria* sp. 6412, *Pseudanabaena*, *Nostoc* sp. 6705 [12, 19]. Notably, filamentous cyanobacteria *Calothrix* and *Nostoc* can differentiate hormogonia upon environmental stimuli, the process of which is characterized by the formation of GV [2, 20]. Despite the laboratory strain of model

filamentous cyanobacteria *Anabaena* sp. PCC 7120 fails in differentiating hormogonia [19, 21], it remains unknown whether the natural isolate could differentiate hormogonia and produce GV. Here we found that *Anabaena* sp. PCC 7120 possesses an intact *gvp* gene cluster, which shares an organization similar to that of previously identified GV-forming cyanobacteria. The results of real-time PCR showed that *gvpA* is constitutively transcribed in vivo, and its expression level could be augmented at an altered light intensity and growth temperature. The complete *gvp* gene cluster could be heterologously expressed and assembled into irregular GV in *Escherichia coli*. Moreover, structural combined with enzymatic investigations suggested that GvpF and GvpN are structurally and enzymatically conserved, respectively. These findings indicated that the natural isolate of *Anabaena* sp. PCC 7120 is most likely able to produce GV under some given environmental stimuli.

Results

Organization and conservation of the *gvp* genes in *Anabaena* sp. PCC 7120

The entire genomic sequence of the model filamentous nitrogen-fixing cyanobacteria *Anabaena* sp. PCC 7120 was reported in 2001, which consists of a single circular genome of 6,413,771 bp and six plasmids [22]. Eight out of the 5368 putative open reading frames in the genome were annotated as *gvp* genes: *gvpA*, *gvpB*, *gvpC*, *gvpN*, *gvpJ*, *gvpK*, *gvpF* and *gvpG*, without annotations of *gvpV* and *gvpW* compared to some other *gvp* gene clusters. Using BlastP program, we found that the proteins encoded by *alr2246* and *alr2245*, two genes at the downstream of *gvpG*, share a sequence similarity of 62% and 65% to GvpV and GvpW of *M. aeruginosa* PCC 7806, respectively. Thus we assigned *alr2246* and *alr2245* to *gvpV* and *gvpW*, respectively (Fig. 1). It suggested that *Anabaena* sp. PCC 7120 possesses an intact *gvp* gene cluster, which shares a gene organization similar to that in the previously reported GV-forming cyanobacteria, such as *A. flos-aquae* and *M. aeruginosa* PCC 7806. Notably, most of the *gvp* genes in GV-forming *Haloarchaea* and other bacteria are highly conserved [5], despite the gene organizations vary a lot (Fig. 1).

Multiple-sequence alignment showed that *gvpB* is nearly identical to *gvpA* in the *gvp* gene cluster of *Anabaena* sp. PCC 7120, suggesting that *gvpB* is an isoform of *gvpA*. Accordingly, *gvpA* and *gvpB* should be re-annotated to *gvpA1* and *gvpA2*, respectively (Fig. 1). Moreover, GvpA of *Anabaena* sp. PCC 7120 shares a sequence similarity to those of other cyanobacterial strains up to 90% (Fig. 2a), indicating that the primary structural protein GvpA exhibits a rarely high conservativity in cyanobacteria. Further sequence analysis revealed that the external scaffold protein GvpC of *Anabaena* sp. PCC 7120 contains only three conserved 33RRs (Fig. 2b),

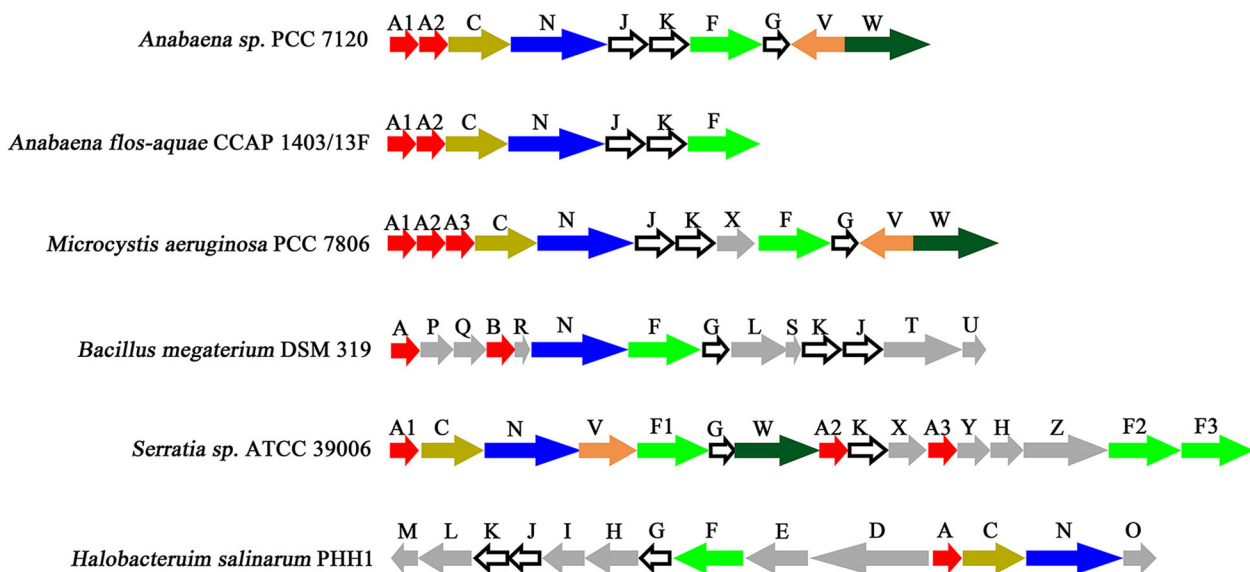


Fig. 1 Organizations of the *gyp* gene cluster from different species of bacteria. Each alphabet above the arrow represents a *gyp* gene. Transcription direction of each gene is indicated by the arrow. The *gyp* genes absent in *Anabaena* sp. PCC 7120 are shown as grey arrows

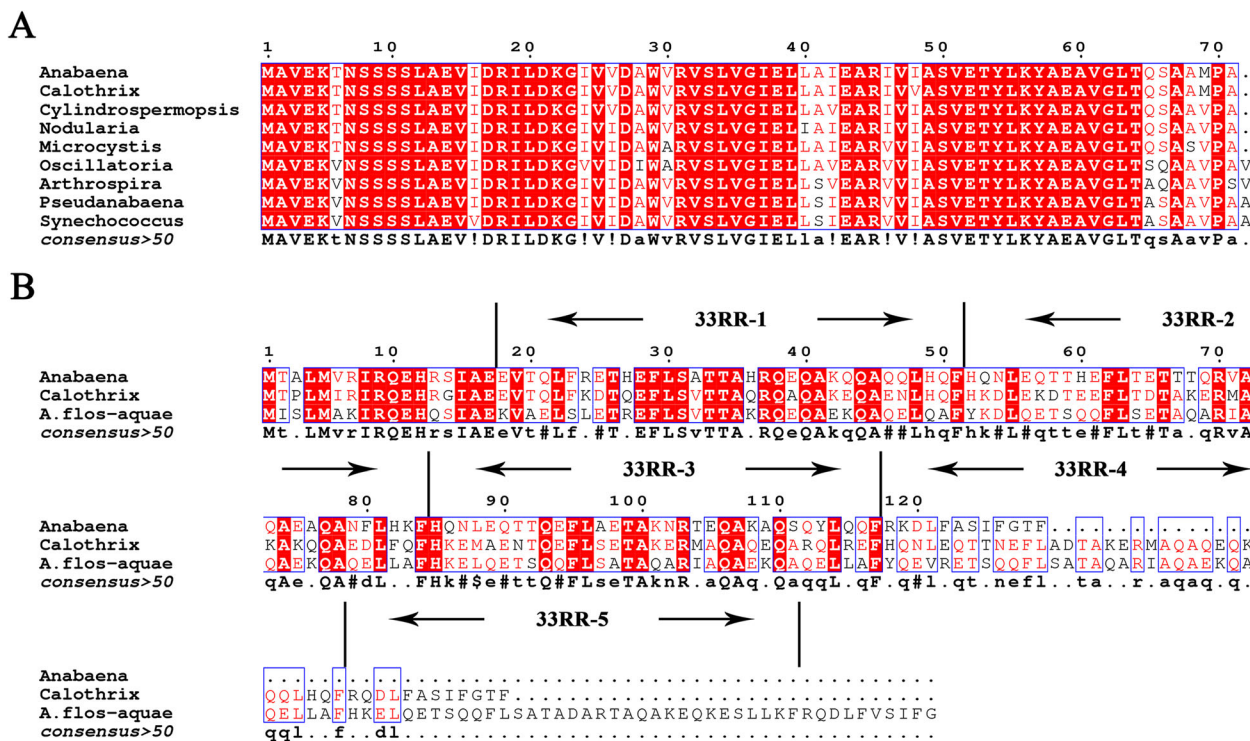


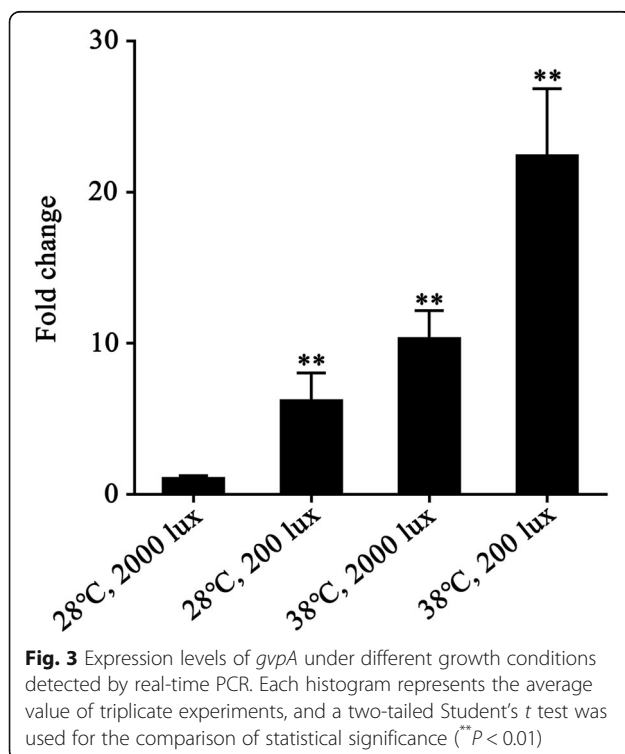
Fig. 2 Consistency of GvpA and GvpC. **a** Multiple-sequence alignment of GvpA from different cyanobacterial strains. The alignment was performed with the program Multalin. All sequences were downloaded from the NCBI database (www.ncbi.nlm.nih.gov) with the following accession numbers: *Anabaena* sp. PCC 7120, WP_010996411; *Calothrix* sp. PCC 7103, WP_011316976; *Nodularia spumigena* CCY9414, AHJ27872; *Microcystis aeruginosa* PCC 7806, WP_084989880; *Cylindrospermopsis raciborskii*, WP_057178839; *Oscillatoria* sp. PCC 10802, WP_017721733; *Arthrospira platensis*, WP_006616598; *Pseudanabaena* sp. SR411, WP_009626980; *Synechococcus* sp. PCC 7502, WP_015167036. **b** Multiple-sequence alignment shows the 33RRs of GvpC. All sequences were downloaded from the NCBI database with the following accession numbers: *Anabaena* sp. PCC 7120, WP_010996410; *A. flos-aquae*, AAA58710; *Calothrix* sp. PCC 7103, EKf01074

which probably result in GVs of smaller diameter. In fact, a previous report revealed that *A. flos-aquae* GVs with a GvpC of five 33RRs have a larger diameter compared to those of *Calothrix* sp. PCC 7601 with a GvpC of four 33RRs [23].

The *gvpA* gene is upregulated at low light intensity and high temperature

Considering that the transcription of *gvp* genes is the prerequisite of GV formation, we investigated whether the *gvpA* gene, encoding the major structural component of GV, could be transcribed in vivo. The total RNA was extracted from *Anabaena* sp. PCC 7120 cells grown at a light intensity of 2000 lux at 28 °C, and applied to real-time PCR assays. The results showed that an expected fragment of *gvpA* could be detected (Fig. 3), suggesting that the *gvpA* gene is constitutively transcribed in *Anabaena* sp. PCC 7120 under normal laboratory growth condition.

Afterwards, we shifted *Anabaena* sp. PCC 7120 cells to various external stimuli and detected the expression profiles of *gvpA* gene. The expression level of *gvpA* in *Anabaena* sp. PCC 7120 upon a single stimulus of low light intensity at 200 lux or high temperature at 38 °C were elevated to 6 and 11 folds, respectively (Fig. 3), compared to the constitutive expression level. Moreover, when *Anabaena* sp. PCC 7120 cells were grown under the condition of double stimuli of both low light intensity and high temperature, the expression level of *gvpA*



was upregulated approximately 23 folds (Fig. 3). Considering that GvpA is the primary structural component of GV, we speculated that prototype GVs could be produced in *Anabaena* sp. PCC 7120 at some given conditions. However, we failed in observing the floatation of *Anabaena* sp. PCC 7120 cells in response to the above double stimuli. It implied that mature and functional GVs do not exist in the laboratory strain of *Anabaena* sp. PCC 7120, in consistence with its incapability of differentiating hormogonia.

Expression of the *gvp* gene cluster of *Anabaena* sp. PCC 7120 in *E. coli*

The *gvp* genes of *Anabaena* sp. PCC 7120 were constructed in the expression vectors and then transformed to *E. coli* cells. Interestingly, we observed that the *E. coli* cells transformed with the recombinant *gvp* plasmids exhibit a buoyancy phenotype, whereas the cells carrying the control vectors sink to the bottom of the tube (Fig. 4a). Then, the turbidity measurements showed that the upper fraction of the culture medium of the experimental group has an absorbance of 0.57 at 600 nm, compared to 0.02 of the control. The results suggested that GVs might be produced in the recombinant *E. coli* cells that harbor the *gvp* gene cluster of *Anabaena* sp. PCC 7120.

Afterwards, the recombinant GVs were purified from *E. coli* cells and applied to transmission electron microscopy, following a previously reported protocol [18]. The image displayed a large number of round and oval bubble-like structures (Fig. 4b), similar to those irregular GVs via expressing *Bacillus megaterium* *gvp* gene cluster in *E. coli* [24]. In contrast, similar bubble-like structures were absent from the *E. coli* cells that expressing the control vectors (Fig. 4b). Western blot assays further showed that the purified GVs possess enriched His-tagged GvpA proteins (Fig. 4c). Altogether, transformation of the *gvp* gene cluster of *Anabaena* sp. PCC 7120 enabled *E. coli* to display a buoyancy phenotype, because of the assembly of GVs, indicating that *Anabaena* sp. PCC 7120 possesses an intact *gvp* gene cluster capable of heterogeneously producing irregular but functional GVs.

The crystal structure of GvpF

To further investigate the putative GVs of *Anabaena* sp. PCC 7120 from the structural point of view, all Gvp proteins, except for GvpA and GvpC, were successfully overexpressed, purified and applied to crystal screening; however, only the crystal structure of GvpF was eventually solved at 2.55 Å resolution in the space group C222₁ (Table 1). *Anabaena* sp. PCC 7120 GvpF is composed of two structurally separated domains, both of which display a fold in $\alpha + \beta$ class, followed by a C-terminal tail

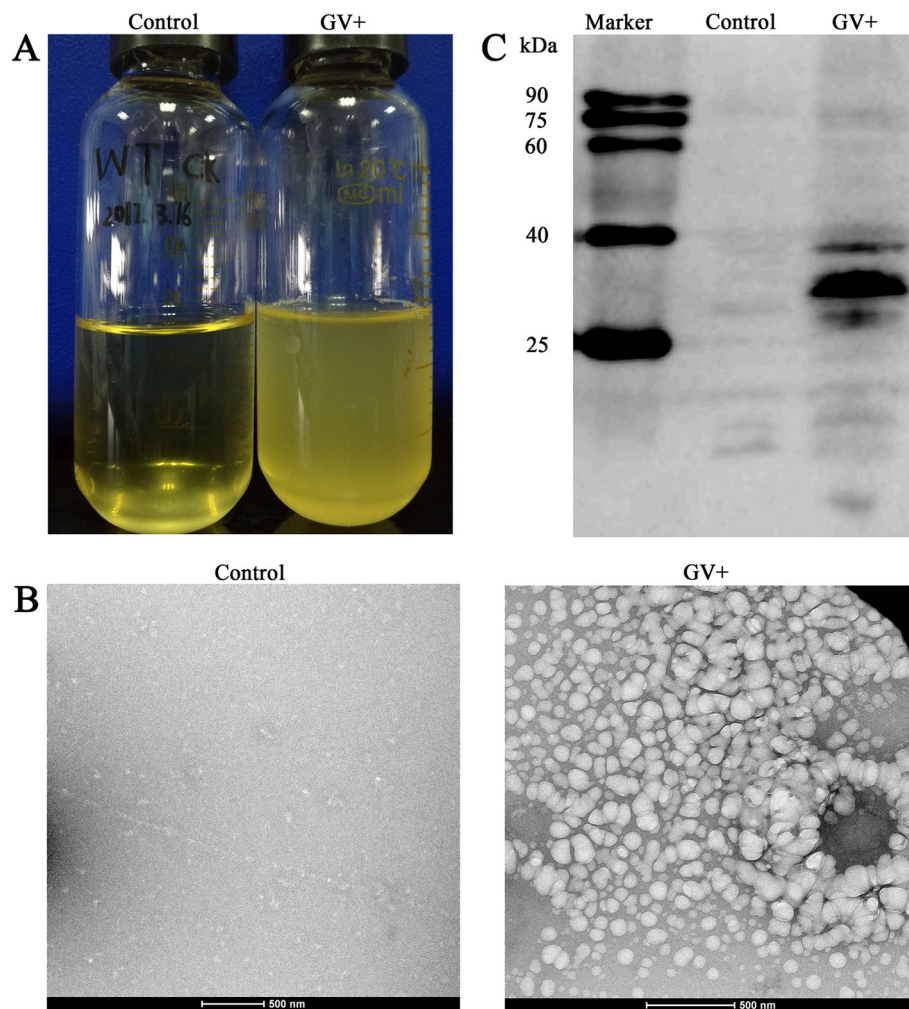


Fig. 4 Expression of the *gvp* gene cluster of *Anabaena* sp. PCC 7120 in *E. coli*. **a** Photographs of *E. coli* cells transformed with recombinant *gvp* plasmids and control vectors, respectively. **b** Negative-staining electron microscopy images of the putative gas vesicles purified from *E. coli* cells expressing the *gvp* genes (right) and the control vectors (left). **c** Western blot of the purified gas vesicles. The probe is anti-His antibody. The prestained protein standards are displayed in the lane marked Marker and their molecular masses are indicated in kDa. The lanes marked control and GV+ indicate the control and the experimental group, respectively

inserted into the middle area of the two domains (Fig. 5a). Further structural analysis showed that the N-domain of GvpF displays an architecture in which a six-stranded β -sheet ($\beta 1$ - $\beta 6$) is sandwiched by two α -helices ($\alpha 1$ - $\alpha 2$) and the helix $\eta 1$, whereas the C-domain adopts a modified ferredoxin fold owing to an extension region consisting of three consecutive helices ($\alpha 4$, $\alpha 5$ and the N-terminal segment of $\alpha 6$) (Fig. 5a). Moreover, the additional C-terminal tail provides an interface for the N-domain and C-domain to pack against each other, resulting in the structural stability and correct folding of GvpF (Fig. 5a).

DALI search [25] revealed that *Anabaena* sp. PCC 7120 GvpF shares a high structural homology to the previously reported GvpF of *M. aeruginosa* PCC 7806 (PDB code: 4QSG, Z score 27.5, sequence identity 67%), with a

root-mean-square deviation of 1.8 Å over 238 C α atoms. Superposition of the two structures showed a very similar structure, except that GvpF of *Anabaena* sp. PCC 7120 possesses a shorter helix $\alpha 5$ in the C-domain (Fig. 5b). Besides, structure-based multiple-sequence alignment revealed that GvpF proteins are highly conserved among diverse species of cyanobacteria (Fig. 5c). It indicated that *Anabaena* sp. PCC 7120 GvpF might also function as a structural protein involved in forming GVs as that of *M. aeruginosa* PCC 7806 [18].

GvpN is an active ATPase

Sequence analysis against the Pfam database [26] showed that *Anabaena* sp. PCC 7120 GvpN contains an ATPases Associated with various cellular Activities (AAA) domain at the N-terminus, which was previously

Table 1 Crystal parameters, data collection, and structure refinement

	GvpF
Data collection	
Space group	C222 ₁
Unit cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	96.40, 147.09, 106.58
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90.00, 90.00, 90.00
Resolution range (Å)	50.00–2.55 (2.62–2.55) ^a
Unique reflections	24, 900 (2, 440)
Completeness (%)	99.9 (100)
<I/σ(I)>	16.3 (2.9)
<i>R</i> _{merge} ^b (%)	10.6 (69.1)
Average redundancy	10.9 (10.9)
Structure refinement	
Resolution range (Å)	44.46–2.55
<i>R</i> _{factor} ^c / <i>R</i> _{free} ^d (%)	19.57/25.61
Number of protein atoms	3, 962
Number of water atoms	28
RMSD ^e bond lengths (Å)	0.015
RMSD bond angles (°)	1.709
Mean B factors (Å ²)	53.189
Ramachandran plot (residues, %) ^f	
Most favored	95.80
Allowed	4.00
Outliers	0.20
Protein Data Bank entry	6L5D

^a Highest resolution shell is shown in parenthesis. $R_{sym} = \sum_h \sum_i |I_{h,i} - \bar{I}_h| / \sum_h \sum_i I_{h,i}$, where \bar{I}_h is the mean intensity of the i observations of symmetry related reflections of h . $R = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, where $F_{obs} = F_p$ and F_{calc} is the calculated protein structure factor from the atomic model. RMSD in bond lengths and angles are the deviations from ideal values

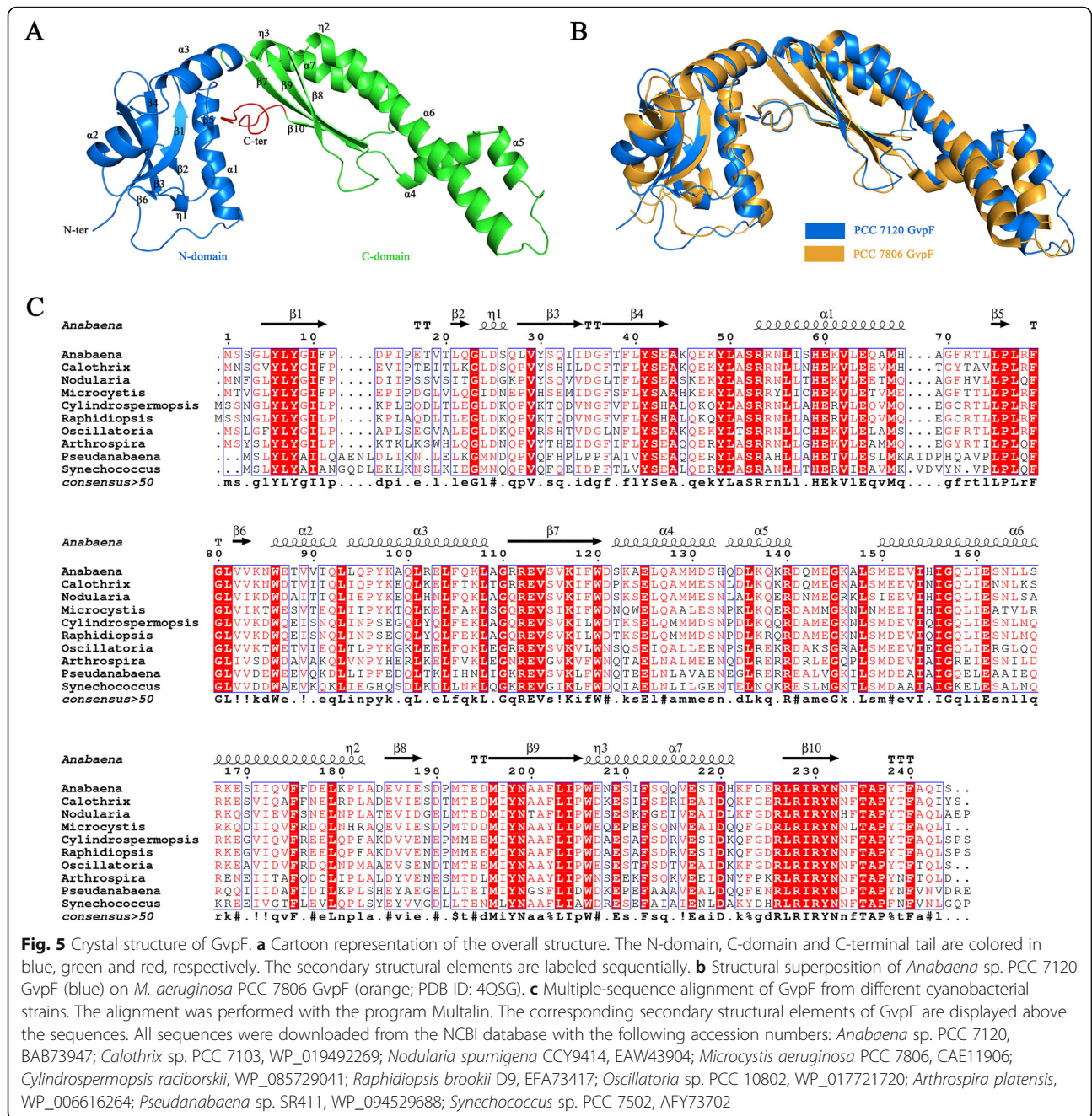
classified in the AAA+ protein superfamily of the ring-shaped P-loop NTPases [27]. Therefore, the recombinant GvpN of *Anabaena* sp. PCC 7120 was overexpressed in *E. coli* and purified (Fig. 6a), which was applied to the ATPase activity assays. Upon the addition of recombinant GvpN, the substrate ATP was gradually hydrolyzed over time (Fig. 6b). Upon the increase of GvpN added to the reaction, the ATP was hydrolyzed at a higher rate (Fig. 6b), suggesting that GvpN indeed possesses the ATPase activity. Using the Hanes-Woolf plot method (Fig. 6c), we determined the Michaelis-Menten parameters of GvpN towards ATP at a K_m of $3.9 \pm 1.5 \mu\text{M}$, k_{cat} of $35 \pm 2 \text{ s}^{-1}$ and k_{cat}/K_m of $8.97 \text{ s}^{-1} \mu\text{M}^{-1}$. Moreover, multiple-sequence alignment showed that the AAA domain of GvpN is highly conserved among different cyanobacterial species (Fig. 6d), indicating that the ATPase activity is a common feature of GvpN. In fact, a previous report showed that deletion of GvpN in

Serratia sp. ATCC 39006 led to small bicone-shaped GVs and lack of cell buoyancy [28]. All together, it suggested that *Anabaena* sp. PCC 7120 possesses an enzymatically active GvpN, which might be necessary for the formation of mature GVs.

Discussion

GVs play an essential role in the survival of prokaryotic species, and thus should be assembled or disassembled properly at the right time in response to diverse external stimuli. Actually, the formation of GVs or the expression of genes encoding Gvp proteins are affected by various environmental stimuli, such as temperature, light intensity, oxygen supply, pH and salinity, cell density and carbon source [5, 29–32]. For the cyanobacteria *A. flos-aquae*, *Calothrix* sp. PCC 7601 and *Microcystis* sp. BC 84/1, low light intensity could induce more GVs in order to enable the cells to move towards the surface of the aqueous habitat [5]. Mlouka and colleagues found that lack of nutrition, especially CO₂ and light irradiance, leads to an augmented production of GVs in *M. aeruginosa* [12]. Moreover, for enterobacteria *Serratia* sp. ATCC39006, the formation of GVs depends on cell density under the control of quorum-sensing signals, and is responsive to oxygen shortage, resulting in facilitating the buoyancy of cells [28, 33]. In addition, for haloarchaea, two regulatory proteins GvpD and GvpE were shown to be involved in regulating the expression of *gvp* genes at both transcriptional and translational level [34]. Altogether, the formation of GVs is necessary for some bacteria in response to various environmental conditions.

The model filamentous and heterocyst-forming cyanobacteria *Anabaena* sp. PCC 7120 were isolated from the Lake Michigan in the late 1960s [22]. Theoretically, the environmental stimuli mentioned above should probably induce the formation of GV in *Anabaena* sp. PCC 7120, which possesses the *gvp* gene cluster. However, no one has observed GVs in the past 50 years of studying the physiology of this model organism. In this study, we found that *Anabaena* sp. PCC 7120 has an intact *gvp* gene cluster similar to those GV-forming cyanobacterial strains (Fig. 1). In fact, *A. flos-aquae* and *M. aeruginosa* PCC 7806 were proved to be able to produce cylinder-shaped GVs, the formation of which was regulated by light intensity [2, 12]. It is most likely that *Anabaena* sp. PCC 7120 is also capable of forming GVs under some given environmental stimuli. Indeed, the primary structural gene *gvpA* of *Anabaena* sp. PCC 7120 is constitutively transcribed, and could be upregulated at low light intensity and high temperature (Fig. 3). Heterologous expression of the intact *Anabaena* sp. PCC 7120 *gvp* gene cluster enabled *E. coli* cells to gain the capacity of floatation, thanks to the formation of irregular GVs (Fig. 4).



However, we failed in either setting up a reproducible procedure to enable the floatation of *Anabaena* sp. PCC 7120, or seeing GVs inside the cells, or purifying GVs from this laboratory strain, after extensive trials of various stimuli and combinations. It indicated that the *gvp* gene cluster of *Anabaena* sp. PCC 7120 was partially degenerated in the 50-year laboratory culture.

Notably, formation of GVs is a key feature accompanied with the differentiation of hormogonia, which has already been proved in filamentous cyanobacteria *Calothrix* and *Nostoc* [2, 20]. Recently, Gonzalez and colleagues reported

that hormogonia differentiation is regulated by a hierarchical sigma factor cascade in the filamentous cyanobacteria *Nostoc punctiforme*, which retain the developmental complexity of natural isolates [21]. In detail, the sigma factor sigJ activates the expression of both *sigC* and *sigF* genes, as well as other hormogonium-specific genes; meanwhile, sigJ controls the transcription of *gvpA* gene via binding at the -10 region, which is a consensus sigJ-dependent promoter (designated as J-Box, GGGaAtaT) [21]. However, we found that the highly conserved GGG stretch of J-Box was mutated to AGC at the upstream promoter region of

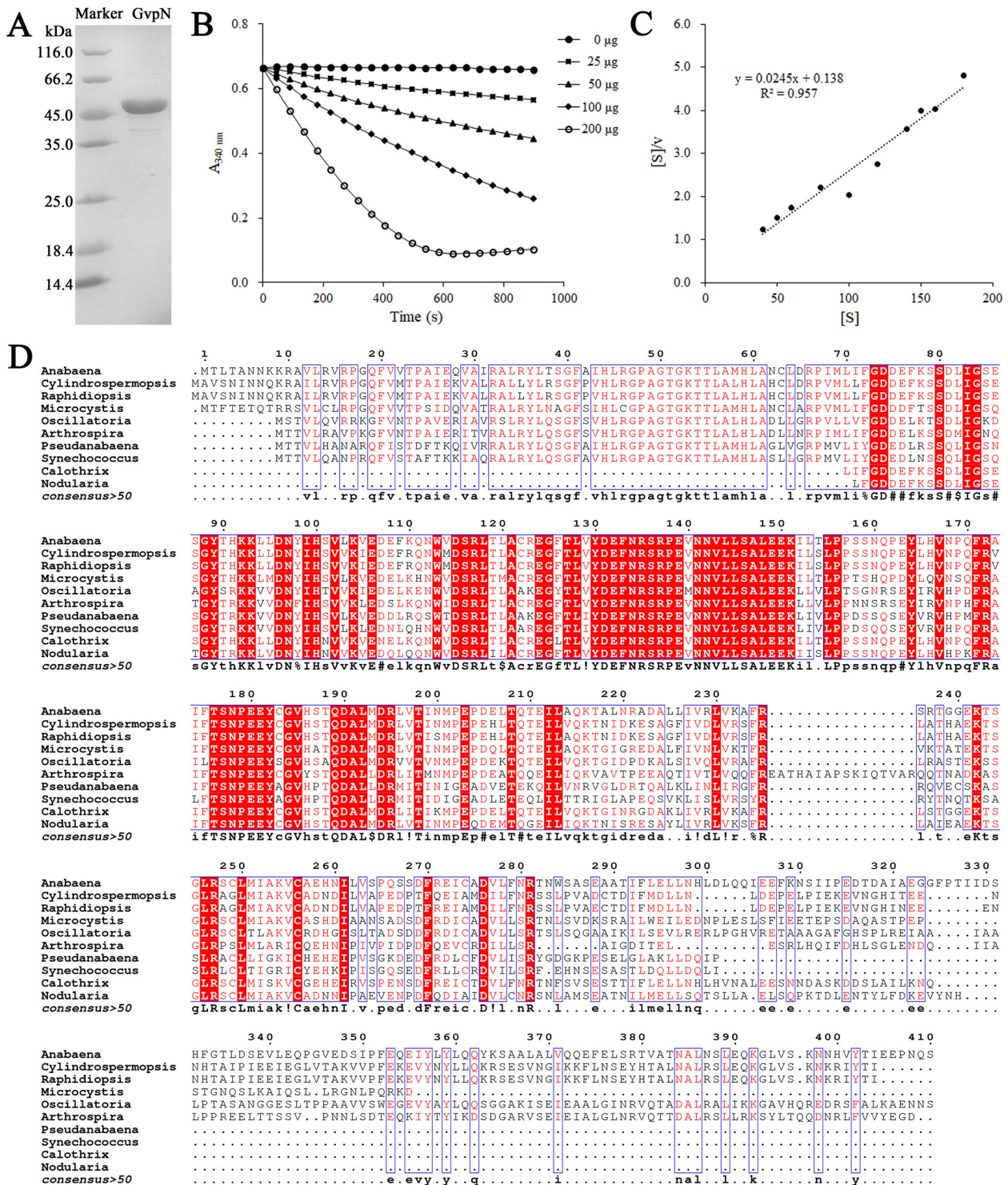


Fig. 6 GvpN is an active ATPase. **a** SDS-PAGE of the purified GvpN protein. **b** The enzymatic profiles of GvpN. The final amounts of recombinant GvpN in the 200-µL system are 0, 25, 50, 100 and 200 µg, respectively. The decrease in absorbance at 340 nm was monitored using a DU800 spectrophotometer. **c** The Hanes-Woolf plot of GvpN. **d** Multiple-sequence alignment of GvpN from different species of cyanobacteria. The alignment was performed with the program Multalin. All sequences were downloaded from the NCBI database with the following accession numbers: *Anabaena* sp. PCC 7120, WP_010996409; *Calothrix* sp. PCC 7103, WP_019492266; *Nodularia spumigena* CCY9414, AHJ27875; *Microcystis aeruginosa* PCC 7806, WP_002747926; *Cylandrospermopsis raciborskii*, WP_061547066; *Raphidiopsis brookii* D9, EFA73420; *Oscillatoria* sp. PCC 10802, WP_017715028; *Arthrospira platensis*, WP_006616595; *Pseudanabaena* sp. SR411, WP_094529416; *Synechococcus* sp. PCC 7502, WP_015167038

gvpA in the laboratory strain of *Anabaena* sp. PCC 7120, which might result in the altered binding affinity towards sigJ, and eventually the failure of producing functional GVs. It indicated that the natural isolate of *Anabaena* sp. PCC 7120, without the mutations at the regulatory region of *gvp* gene cluster, might be capable of differentiating hormogonia and producing GVs for better floatation in response to some given environmental stimuli. More investigations including comparative genomics analyses might help us to clearly elucidate which mutations in the present laboratory strain of *Anabaena* PCC 7120 lead to the loss of function.

Conclusions

In this study we demonstrated that the laboratory model filamentous cyanobacterium *Anabaena* sp. PCC 7120 indeed possesses an intact but partially degenerated gene cluster encoding gas vesicles, which gives us the hint that its natural isolate was most likely able to produce GVs under some given environmental stimuli. Owing to the fast growth and non-toxicity of the model strain sp. PCC 7120, investigations that enable the large production of GVs in this strain will benefit the potential application of GVs in biological imaging.

Methods

RNA extraction and real-time PCR

The *Anabaena* sp. PCC 7120 cells were grown at 28 °C under a light intensity of 2000 lux (supplied from top) with a 12/12 photoperiod in BG-11 medium to an OD_{730nm} of 0.8, and then induced with 200 lux light intensity, 38 °C and both for 24 h, respectively. The stressors were selected according to a previous report summarizing the environmental conditions that could induce the formation of gas vesicles [5]. The cells were harvested by centrifugation and washed twice with the PBS buffer. The total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The residual genomic DNA was removed by RNase-free DNase (Takara, Shiga, Japan) at 37 °C for 2 h. PCR assays were conducted to confirm the absence of genomic DNA contamination. The RNA quality was checked by agarose gel electrophoresis. The cDNA synthesis was carried out by reverse transcription using the PrimeScript™ RT reagent Kit (Takara, Shiga, Japan).

For real-time PCR, amplification was performed using the FastStart universal SYBR Green Master (Roche, Basel, Switzerland) with the StepOne™ Real-Time PCR System (Applied Biosystems, Carlsbad, USA). The primers for *rnpB* are 5'-GCGATTATCTATCTGGGACG and 5'-CAACTCTTGTTAAGGGTGC, whereas those for *gvpA* are 5'-TGGCAGAAGTTATTGACC and 5'-GAGAAACACGTACCCAAG. Notably, the *rnpB*

gene encoding RNaseP subunit B was used as the internal reference gene according to previous real-time PCR experiments concerning cyanobacteria [35]. The PCR conditions were as follows: 1 cycle at 95 °C for 10 min, 40 cycles at 95 °C for 15 s, 60 °C for 60 s, and 72 °C for 20 s; then the melting curve stage was performed rising from 60 °C to 95 °C by every 0.3 °C. The transcription ratios of *gvpA* to *rnpB* were calculated using the relative quantification analysis module of 2^{-ΔΔCt} method based on Ct values [36]. All real-time PCR experiments were performed in triplicate.

Buoyancy tests

The *gvpABC*, *gvpNJKFG* and *gvpVW* genes of *Anabaena* sp. PCC 7120 were amplified and cloned into pET-Duet, pET-28a and pCDFDuet-1 vectors (with different antibiotic markers), respectively. Notably, a His-tag was fused to the N-terminus of GvpA for detecting the expression of the gene cluster. Next, the three *gvp* recombinant plasmids were co-overexpressed in *E. coli* BL21 (DE3) strain. Cells were grown in liquid LB broth, induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C, and resuspended in 35-mm-diameter test tubes. Then, the tubes were undisturbed at room temperature for about 24 h, at which time the cell buoyancy was determined by the turbidity of the upper fraction of the culture medium. The cells transformed with the empty vectors without *gvp* gene cluster were used as the control.

GV isolation, electron microscopy and western blot

According to a previously described protocol [18], GVs were purified from *E. coli* cells co-overexpressing the three recombinant plasmids that cover the complete *gvp* gene cluster. Carbon-coated copper grids (300-mesh) were immersed in the purified GVs for 1 min and excess liquid was removed with filter paper. GVs were negatively stained with 2% (w/v) uranyl acetate and then examined with a Tecnai G² transmission electron microscopy (FEL, USA) running at 120 kV voltage. Images were taken using a CCD camera attached to the microscopy. The purified GVs were mixed with an equal volume of 2 × sample-loading buffer (100 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 2% β-mercaptoethanol, 0.2% bromophenol blue), boiled for 10 min, and then applied to western blot using anti-His polyclonal antibodies.

Cloning, expression and purification of GvpF and GvpN

The coding region of *gvpF* was amplified from the genomic DNA of *Anabaena* sp. PCC 7120, and cloned into a modified pET-29a vector with a C-terminal 6 × His-tag. The *E. coli* BL21 (DE3) strain was used for the overexpression of recombinant protein. The transformed cells were grown at 37 °C until OD_{600 nm} reached 0.8 and then induced with 0.2 mM IPTG for another 20 h at

16 °C. Cells were harvested by centrifugation (6000×g, 4 °C, 10 min) and resuspended in the lysis buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl). After 10 min of sonication on ice and 30 min of centrifugation at 12,000×g, the supernatant was loaded onto a Ni-NTA column (GE Healthcare, Chicago, USA) equilibrated with the binding buffer, the same as the lysis buffer. The target protein was eluted with 300 mM imidazole and further applied to a Superdex 200 column (GE Healthcare, Chicago, USA) pre-equilibrated with the binding buffer. Fractions containing the target protein were collected and concentrated to 10 mg/mL for crystallization.

GvpN was expressed and purified in the same manner as GvpF. Samples for ATPase activity assays were collected at the highest peak fractions without concentration and stored at – 80 °C with 50% glycerol.

Crystallization, data collection and structure determination

Crystals of GvpF were grown at 16 °C using the hanging drop vapor diffusion method, with a drop of 1 µL protein solution mixed with an equal volume of the reservoir solution. Crystals were obtained against the reservoir solution of 20% (w/v) polyethylene glycol 4000, 0.2 M NaCl, and 0.1 M Tris-HCl, pH 8.0. Then, they were pooled and flash cooled with liquid nitrogen after transferring to cryoprotectant (reservoir solution supplemented with 30% sucrose). The X-ray diffraction data were collected at 100 K using beamline BL17U with an EIGER X 16 M detector at the Shanghai Synchrotron Radiation Facility.

The diffraction data were indexed, integrated and scaled with HKL-2000 to the highest resolution of 2.55 Å. The structure of *M. aeruginosa* PCC 7806 GvpF (PDB code: 4QSG) was used as the search model to determine the structure of *Anabaena* sp. PCC 7120 GvpF by molecular replacement using the Molrep program [37] in the CCP4i program suite [38]. Further refinement was performed by programs REFMAC5 [39] and COOT [40]. The final model was evaluated with MolProbity [41]. Crystallographic parameters and data-collection statistics are listed in Table 1. All structure figures were prepared with PyMOL (<https://pymol.org>).

ATPase activity assays of GvpN

The ATPase activity of GvpN was measured using an ATP/NADH coupled assay [42], in which the decrease of NADH is proportional to the rate of steady-state ATP hydrolysis. The reaction mixture contains 50 mM Tris-HCl, pH 8.0, 20 mM KCl, 5 mM MgCl₂, 2.5 mM ATP, 1 mM phosphoenolpyruvate, 0.1 mM NADH, 12 U/mL pyruvate kinase (Sigma, Saint Louis, USA) and 12 U/mL lactate dehydrogenase (Sigma, Saint Louis, USA). The reaction was initiated by the addition of recombinant GvpN, final amounts of which in a 200-µL system are 0,

25, 50, 100 and 200 µg, respectively. Using a DU800 spectrophotometer (Beckman Coulter, Fullerton, USA), the decrease in absorbance at 340 nm was monitored at 25 °C at 45 s intervals for 15 min. Michaelis-Menten parameters of GvpN were calculated from the data at the concentration of NADH varying from 40 to 200 µM and in the presence of 50 µg GvpN using the Hanes-Woolf plot method.

Abbreviations

GV: Gas vesicle; *gvp* gene cluster: Gas vesicle protein gene cluster; 33RR: 33-Residue repeating motif; IPTG: isopropyl β-D-1-thiogalactopyranoside; AAA protein: ATPases Associated with various cellular Activities protein.

Acknowledgements

We appreciate the assistance of the staff at the Shanghai Synchrotron Radiation Facility (SSRF) and the Core Facility Center for Life Sciences at University of Science and Technology of China.

Authors' contributions

CZZ and QL conceived and designed the experiments; KC and YW performed the experiments; KC and YLJ solved and refined the structure; KC and BYX performed the ATPase activity assays; KC, YC, CZZ and QL analyzed the data; CZZ and QL wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work is supported by the National Natural Science Foundation of China (Grant No. 31500602), the Ministry of Science and Technology of China (Grant No. 2016YFA0400900), National Synchrotron Radiation Laboratory (Grant No. UN2018LHJJ) and Chongqing Research Program of Basic Research and Frontier Technology (Grant No. cstc2015jcyjBX0142). The funding bodies had no role in study design, data collection, analysis and interpretation, and in writing of the manuscript.

Availability of data and materials

Structural coordinate of *Anabaena* sp. PCC 7120 GvpF has been deposited in the Protein Data Bank (<https://www.rcsb.org/>) under the accession number of 6L5D. The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230027, Anhui, China. ²College of Life Sciences, Chongqing Normal University, Chongqing 401331, China.

Received: 18 December 2019 Accepted: 27 April 2020

Published online: 06 May 2020

References

1. Walsby AE. Structure and function of gas vacuoles. *Bacteriol Rev.* 1972;36(1): 1–32.
2. Walsby AE. Gas Vesicles. *Microbiol Rev.* 1994;58(1):94–144.
3. Shapiro MG, Goodwill PW, Neogy A, Yin M, Foster FS, Schaffer DV, et al. Biogenic gas nanostructures as ultrasonic molecular reporters. *Nat Nanotechnol.* 2014;9(4):311–6.
4. Walsby AE. Buoyancy-providing role of gas vacuoles in an aerobic bacterium. *Arch Microbiol.* 1976;109(1–2):135–42.
5. Pfeifer F. Distribution, formation and regulation of gas vesicles. *Nat Rev Microbiol.* 2012;10(10):705–15.

6. Lakshmanan A, Lu GJ, Farhadi A, Nety SP, Kunth M, Lee-Gosselin A, et al. Preparation of biogenic gas vesicle nanostructures for use as contrast agents for ultrasound and MRI. *Nat Protoc.* 2017;12(10):2050–80.
7. Kinsman R, Hayes PK. Genes encoding proteins homologous to halobacterial Gvps N, J, K, F & L are located downstream of *gvpC* in the cyanobacterium *Anabaena flos-aquae*. *DNA Seq.* 1997;7(2):97–106.
8. Englert C, Pfeifer F. Analysis of gas vesicle gene expression in *Haloferax mediterranei* reveals that GvpA and GvpC are both gas vesicle structural proteins. *J Biol Chem.* 1993;268(13):9329–36.
9. Beard SJ, Hayes PK, Pfeifer F, Walsby AE. The sequence of the major gas vesicle protein, GvpA, influences the width and strength of halobacterial gas vesicles. *FEMS Microbiol Lett.* 2002;213(2):149–57.
10. Walsby AE, Hayes PK. Gas vesicle proteins. *Biochem J.* 1989;264(2):313–22.
11. Damerval T, Houmard J, Guglielmi G, Csiszar K, Tandeau de Marsac N. A developmentally regulated *gvpABC* operon is involved in the formation of gas vesicles in the cyanobacterium *Calothrix* 7601. *Gene.* 1987;54(1):83–92.
12. Mlouka A, Comte K, Castets AM, Bouchier C, Tandeau de Marsac N. The gas vesicle gene cluster from *Microcystis aeruginosa* and DNA rearrangements that lead to loss of cell buoyancy. *J Bacteriol.* 2004;186(8):2355–65.
13. Hayes PK, Powell RS. The *gvpA/C* cluster of *Anabaena flos-aquae* has multiple copies of a gene encoding GvpA. *Arch Microbiol.* 1995;164(1):50–7.
14. Walsby AE, Hayes PK. The minor cyanobacterial gas vesicle protein, Gvpc, is attached to the outer surface of the gas vesicle. *J Gen Microbiol.* 1988;134:2647–57.
15. Hayes PK, Buchholz B, Walsby AE. Gas vesicles are strengthened by the outer-surface protein. *GvpC Arch Microbiol.* 1992;157(3):229–34.
16. Kinsman R, Walsby AE, Hayes PK. GvpCs with reduced numbers of repeating sequence elements bind to and strengthen cyanobacterial gas vesicles. *Mol Microbiol.* 1995;17(1):147–54.
17. Dunton PG, Mawby WJ, Shaw VA, Walsby AE. Analysis of tryptic digests indicates regions of GvpC that bind to gas vesicles of *Anabaena flos-aquae*. *Microbiology.* 2006;152(Pt 6):1661–9.
18. Xu BY, Dai YN, Zhou K, Liu YT, Sun Q, Ren YM, et al. Structure of the gas vesicle protein GvpF from the cyanobacterium *Microcystis aeruginosa*. *Acta Crystallogr D Biol Crystallogr.* 2014;70(Pt 11):3013–22.
19. Damerval T, Castets AM, Guglielmi G, Houmard J, Demarsac NT. Occurrence and distribution of gas vesicle genes among cyanobacteria. *J Bacteriol.* 1989;171(3):1445–52.
20. Damerval T, Guglielmi G, Houmard J, De Marsac NT. Hormogonium differentiation in the cyanobacterium *Calothrix*: a photoregulated developmental process. *Plant Cell.* 1991;3(2):191–201.
21. Gonzalez A, Riley KW, Harwood TV, Zuniga EG, Risser DD. A tripartite, hierarchical sigma factor cascade promotes hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. *Mosphere.* 2019;4(3):e00231-19.
22. Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, et al. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res.* 2001;8(5):205–13 27–53.
23. Dunton PG, Walsby AE. The diameter and critical collapse pressure of gas vesicles in *Microcystis* are correlated with GvpCs of different length. *FEMS Microbiol Lett.* 2005;247(1):37–43.
24. Li N, Cannon MC. Gas vesicle genes identified in *Bacillus megaterium* and functional expression in *Escherichia coli*. *J Bacteriol.* 1998;180(9):2450–8.
25. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010;38:W545–9.
26. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* 2019;47(D1):D427–D32.
27. Erzberger JP, Berger JM. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu Rev Biophys Biomol Struct.* 2006;35:93–114.
28. Tashiro Y, Monson RE, Ramsay JP, Salmond GPC. Molecular genetic and physical analysis of gas vesicles in buoyant enterobacteria. *Environ Microbiol.* 2016;18(4):1264–76.
29. Gao H, Zhu T, Xu M, Wang S, Xu XD, Kong RQ. pH-dependent gas vesicle formation in *Microcystis*. *FEBS Lett.* 2016;590(18):3195–201.
30. Englert C, Horne M, Pfeifer F. Expression of the major gas vesicle protein gene in the halophilic archaeobacterium *Haloferax mediterranei* is modulated by salt. *Mol Gen Genet.* 1990;222(2–3):225–32.
31. Hechler T, Frech M, Pfeifer F. Glucose inhibits the formation of gas vesicles in *Haloferax volcanii* transformants. *Environ Microbiol.* 2008;10(1):20–30.
32. DasSarma P, Zamora RC, Muller JA, DasSarma S. Genome-wide responses of the model archaeon halobacterium sp strain NRC-1 to oxygen limitation. *J Bacteriol.* 2012;194(20):5530–7.
33. Ramsay JP, Williamson NR, Spring DR, Salmond GPC. A quorum-sensing molecule acts as a morphogen controlling gas vesicle organelle biogenesis and adaptive flotation in an enterobacterium. *P Natl Acad Sci USA.* 2011;108(36):14932–7.
34. Pfeifer F. Haloarchaea and the formation of gas vesicles. *Life (Basel).* 2015;5(1):385–402.
35. Pinto F, Pacheco CC, Ferreira D, Moradas-Ferreira P, Tamagnini P. Selection of suitable reference genes for RT-qPCR analyses in cyanobacteria. *Plos One.* 2012;7(4):e34983.
36. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻($\Delta\Delta C_T$) method. *Methods.* 2001;25(4):402–8.
37. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr.* 1997;30:1022–5.
38. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011;67:235–42.
39. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 2011;67:355–67.
40. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004;60:2126–32.
41. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010;66:12–21.
42. Scharschmidt BF, Keeffe EB, Blankenship NM, Ockner RK. Validation of a recording spectrophotometric method for measurement of membrane-associated mg- and NaK-ATPase activity. *J Lab Clin Med.* 1979;93(5):790–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

